

## 人間の声と顔の表情による AVJ システム

滝澤勇介 飯田栄治  
金沢学院大学美術文化学部情報デザイン学科

一般に VJ(Visual Jockey)とは映像を使ったパフォーマンスであり、コンピュータを利用する VJ システムは数多く発表されている。近年、AVJ(Audio Visual Jockey)と呼ばれる映像と音声を同時に扱うことの出来るシステムが発表され、AVJ パフォーマンスが行われるようになって来た。本研究では既存のシステムとは異なり、人間の声と顔の表情を素材として即座にビデオカメラでコンピュータに取り込む AVJ システムを製作した。本論文ではシステムのコンセプト、構成、操作方法、問題点などを報告する。

### AVJ System using Voices and Facial Expressions

Yusuke Takizawa, Eiji Iida  
Faculty of Fine Arts and Informatics, Kanazawa Gakuin University

Generally VJ (Visual Jockey) machine is a system to give people visual effects by pictures and videos in various live events, i.e. rock concert etc. Many VJ systems using a computer have been invented until now. In recent years, AVJ (Audio Visual Jockey) system, the system that could control pictures and sounds, was invented. We propose the system that was different from existing AVJ systems. The most remarkable feature of our system is to have functions recording and playing back voices and facial expressions. We report the concepts, the constitution, the operation methods and same problems in this system.

#### 1. はじめに

一般に VJ (Visual Jockey)とは映像を使ったパフォーマンスのことを言う[1]。素材となる映像を製作し、複数の映像を切り替えて演奏を行う。以前は映像の再生に複数台のビデオデッキを利用していたが[2]、現在はコンピュータに保存した映像をソフトウェアで切り替えてパフォーマンスを行っている[3]。パフォーマンスは主に既存の音楽や DJ(Disc Jockey)の演奏などを視覚的に盛り上げる為に行われている[4]。近年は映像と音声を同時に扱うことの出来る AVJ(Audio Visual Jockey)と呼ばれるシステムが発表され[5]、AVJ パフォーマンスが行われるようになって来た。

既存の VJ システムは映像を扱う事に特化しており音声を扱うことは出来ない。またビデオカメラなどからの映像入力をそのまま出力することは出来るが、その映像を保存し素材として即座に再利用することは出来ない。事前に複数の映像を用意しなければならない為、長時間の演奏の場合には同じ映像を利用する事が多くなり、パフォーマンスが単調になってしまうという欠点がある。

AVJ は登場してそれほど時間が経っておらず、従来の VJ と DJ を複合しただけのシステムでは操作方法が非常に複雑になるなど、システムとして完成されたものはまだ存在しない。

本研究では、人間の声と顔の表情を素材として即座にビデオカメラでコンピュータに取り込む AVJ システムを製作した。本システムは映像と音声を同時に扱うことができ、ビデオカメラからの入力を素材として保存することが出来る為、事前に複数の映像を用意する必要がない。また本システムは現在、不特定多数の人にパフォーマンスしてもらう事を前提に製作している為、AVJ システムとしての機能の充実よりも手軽に操作出来る事に重点を置いている。

以降、システムのコンセプト、構成、操作方法、問題点などについて報告する。

## 2. コンセプト

本システムは、以下のようなコンセプトにて構築を行った。

- ・コントローラのサイズは、VJとしての操作性を考慮し、大人が両手の指を広げて操作しても無理のない無敵の大きさとする。
- ・ビデオキャプチャー、音声キャプチャー機能を搭載する。
- ・ビデオ再生、音声再生のタイムラグをVJ操作に支障の無い範囲とする。

## 3. システム説明

本システムは図1のとおり、映像の切り替えなどを行うコントローラ、素材となる映像を入力するビデオカメラ、ソフトウェアをインストールしたコンピュータ、映像を出力するプロジェクタ、音声を出力するスピーカーによって構成されている。システムの外観は図2のとおりである。

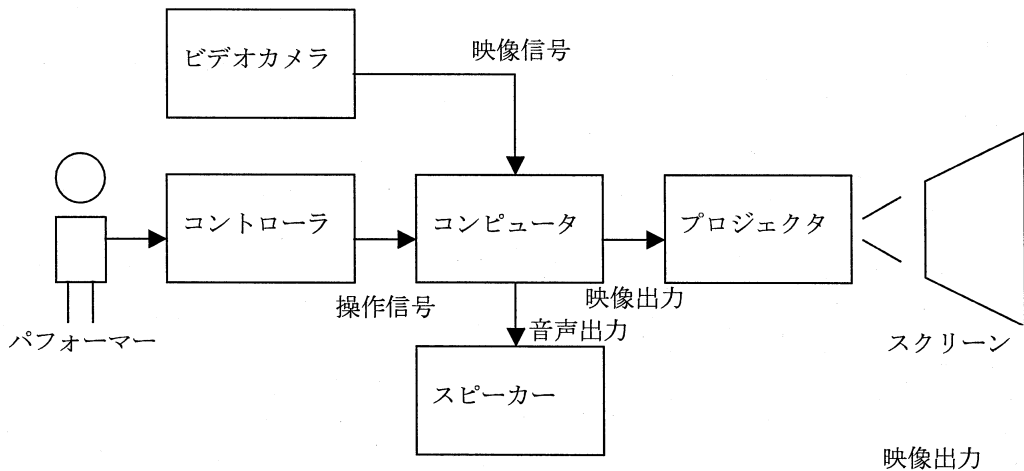


図1 AVJシステム構成

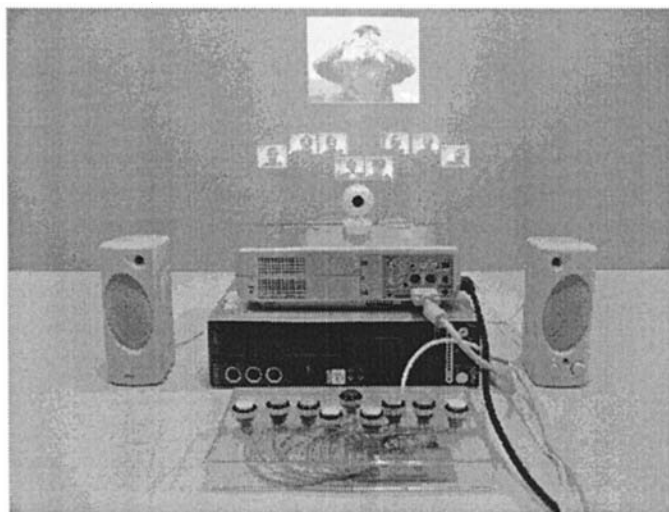


図2 AVJシステム外観

コントローラには図3のとおり PS/2 接続の 104 キーボードの基盤を利用している。基盤には 26 個の端子があり、この端子を 8 行 16 列に組み合わせて各キーを識別する。

ビデオカメラには図4のとおりコンピュータからの制御が容易な、一般的にウェブカメラと呼ばれる USB 接続の物を利用した。このビデオカメラにはマイクロフォンも内蔵されており、映像と同時に音声も取り込むことが出来る。

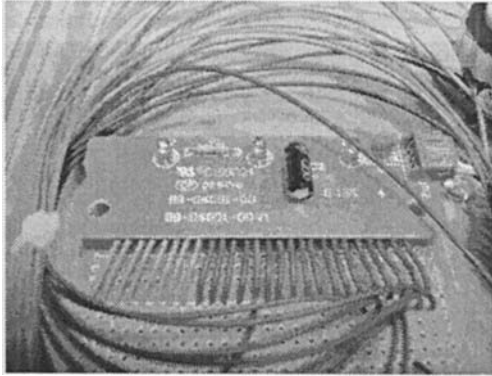


図3 コントローラ基盤

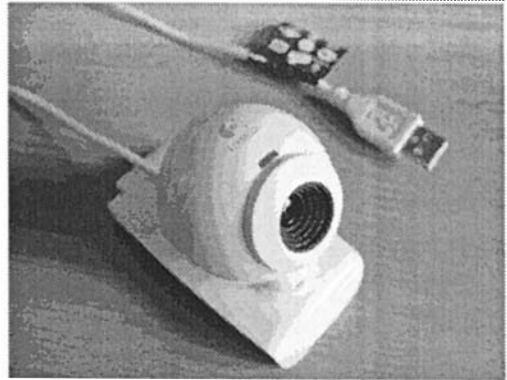


図4 ウェブカメラ

#### 4. 操作方法

コントローラのスイッチの配置は図5、ソフトウェアのインターフェイスは図6のとおりである。

- 1.コントローラの録画スイッチを押し、ソフトウェアの GUI に表示されるプレビューを確認しながら、素材となる映像と音声を録画する。
- 2.録画した映像はコンピュータに保存され、順にコントローラのスイッチに割り当てられる。
- 3.ユーザーはコントローラの再生スイッチを押し、自由に映像を切り替えて AVJ パフォーマンスを行う。

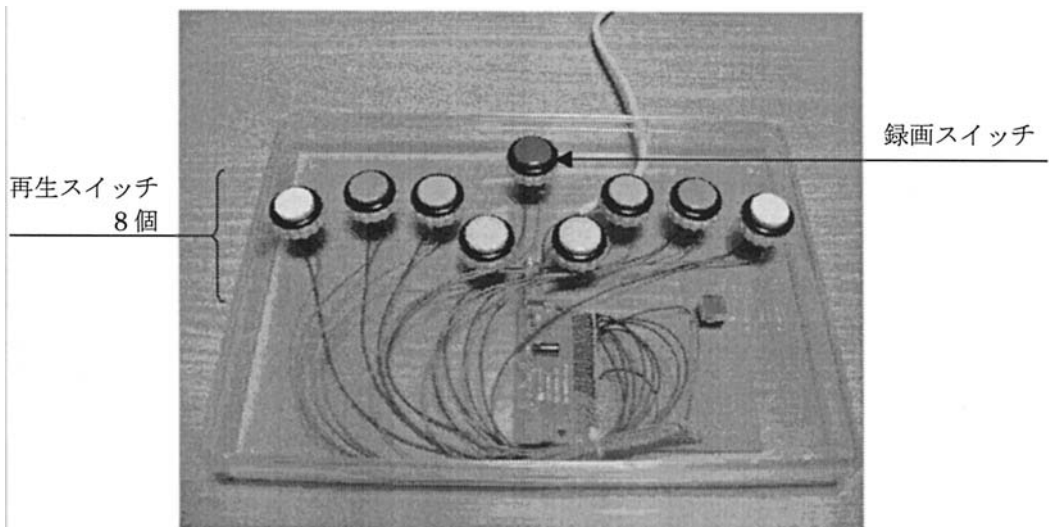


図5 コントローラ



図6 ソフトウェア GUI

## 5. ソフトウェア詳細

本システムのソフトウェアは映像の再生と録画に Microsoft Platform SDK の DirectShow[6]、コントローラからの入力処理に Microsoft Direct X の DirectInput[7]を利用している。

DirectShow は Windows 環境でマルチメディアの再生・キャプチャーを実現するアーキテクチャーである。DirectShow はフィルタと呼ばれるソフトウェアコンポーネントによって構成されており、フィルタの集合を「フィルタグラフ」と言う。フィルタグラフは Platform SDK に付属する「GraphEdit」と言うソフトウェアを使い、ブロック図として表示し簡易的に動作確認することが出来る。図 7 は実際にソフトウェアでウェブカメラから映像と音声を取得しプレビュー表示と録画するフィルタグラフをリモート接続によって表示させたものである。

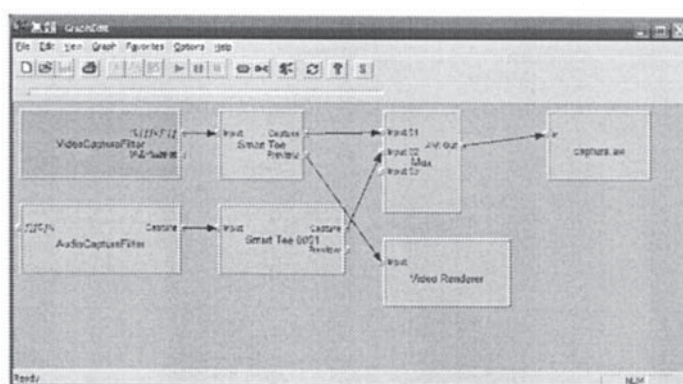


図7 GraphEdit

一般的な Windows アプリケーションはメッセージと呼ばれるループ処理によって動作する為、ゲームなどの高速で正確な入力処理には不向きである。そこで本ソフトウェアではハードウェアドライバと直接通信する DirectInput を利用して、映像の切り替えの際に生じるタイムラグを極力抑えている。

本システムは図 8 のとおり、ビデオカメラからの映像信号をビデオカメラ用フィルタグラフに取り込み、コントローラから入力信号によって、プレビュー画面に表示し素材として保存する。保存された素材を再生用フィルタグラフに取り込み、コントローラからの入力信号によって、メイン画面に表示する。

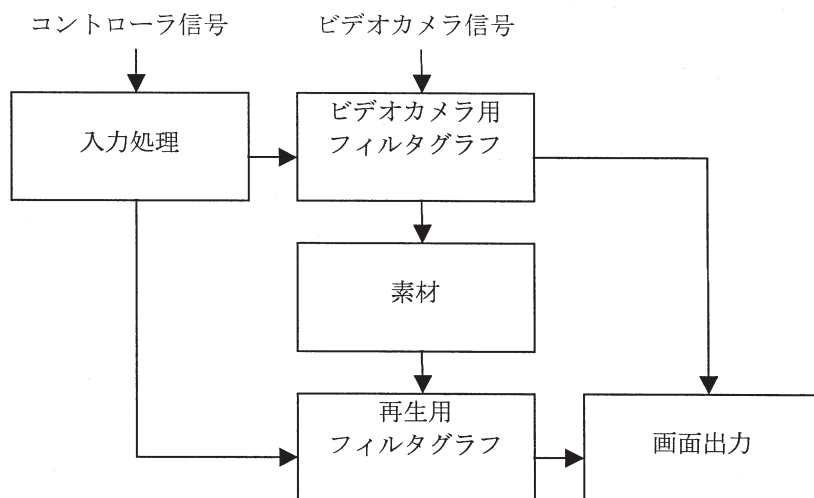


図 7 ソフトウェアブロック図

## 6. 考察

まず、本システムが AVJ システムとして成立しているかという問題であるが、AVJ とは音と映像を使ったパフォーマンスであり、現在の VJ 及び AVJ システムには様々な合成やエフェクトなど、表現の幅を広げる為に様々な機能がある。本システムは簡単な操作を目的としており複雑な機能などは無いが、複数の映像を切り替えて表示する事が出来、最低限のパフォーマンスをする事が出来るので AVJ システムと呼べるであろう。

また、コントローラには録画ボタンを除いて 8 個のスイッチが並んで配置されており、映像素材の音声に音階を付ける事によって、鍵盤楽器としても利用することが可能である。

実際に何人かの VJ 未経験者に演奏してもらったが、はじめの内は探りながらスイッチを押し、再生される映像を見ていたが、数分も経てば映像よりも音に集中し目を閉じてスイッチングによるリズムを刻む様になった。これは DJ の演奏による音を主導として、VJ が映像を演奏するという通常の演奏スタイルと同じである。

操作性について本システムでは当初、図 8 のとおり一般的な VJ と同じくスイッチを押している間だけ映像が再生され、映像が終点に達すると自動的に始点に戻り連続して再生されるというループ再生方式を採用していたが、初心者にもその操作に違和感があると指摘を受け、スイッチを一度押すと映像が最後まで再生される方式に変更し再度演奏してもらった。この方式はピアノなどの鍵盤楽器と同じ感覚で演奏でき、短い映像を連続して切り替えるには都合が良い。しかし、長い映像する再生する場合にはリズムを取り辛く返って演奏が困難になってしまう。また慣れて来ると一つの映像が最後まで再生される前にスイッチングを行ってしまう為、タイミングよく映像の切り替えが出来ないといった不具合が発生する。今回は録画する映像の長さを固定する事でリズムを取り易くし、鍵盤楽器の方式を採用して初心者にも違和感無く演奏出来るようにした。図 8 に再生タイミングイメージを示す。

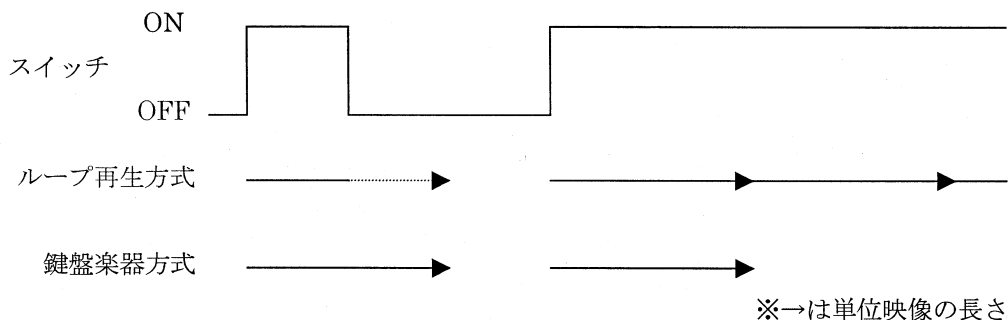


図 8 再生方式

## 7. おわりに

### 謝辞

本システムの操作性試験に協力して頂きました金沢学院大学美術文化学部情報デザイン学科の皆さんに感謝致します。

## 8. 参考文献

- [1] “motordrive VJ”, <http://www5f.biglobe.ne.jp/~motordrive/>
- [2] kom@koms.com, “Visual Remix”, エーアイ出版
- [3] 猪蔵, “VJ:TOKIO”, AGOSTO
- [4] “GLOBAL AUDIO VISUAL JAPAN”, <http://www.audiovisual.jp/>
- [5] “Pioneer DVJ”, <http://pioneer.jp/cdj/dvj/products/>
- [6] “MSDN Japan Microsoft DirectX 9.0c DirectShow”,  
[http://msdn.microsoft.com/library/ja/default.asp?url=/library/ja/DirectX9\\_c/directx/htm/directshow.asp](http://msdn.microsoft.com/library/ja/default.asp?url=/library/ja/DirectX9_c/directx/htm/directshow.asp)
- [7] “MSDN Japan Microsoft DirectX 9.0c DirectInput”,  
[http://msdn.microsoft.com/library/ja/default.asp?url=/library/ja/DirectX9\\_c/directx/inputs/directinputroadmap.asp](http://msdn.microsoft.com/library/ja/default.asp?url=/library/ja/DirectX9_c/directx/inputs/directinputroadmap.asp)