

パノラマ映像生成機能を持つテレビ会議用ビデオカメラの開発 Video Camera for TV Conferences with a Panorama Image Generation Function

長瀬 幸規[†]
Yukinori Nagase

川村 尚生[‡]
Takao Kawamura

菅原 一孔[‡]
Kazunori Sugahara

1. はじめに

本研究では、パノラマ映像生成機能を持つテレビ会議用ビデオカメラを提案する。

1台のビデオカメラでテレビ会議会場を撮影する場合、会場全体を広い範囲に渡って撮影すると、話者の表情を読み取ることが難しい。また逆に、表情が分かるように撮影すると、会場の狭い範囲しか撮影できない。この問題を解決するには、多数のビデオカメラを用いて、会場全体と話者をそれぞれ別々に撮影すれば良いが、多数の話者がいる場合、人数分のビデオカメラを用意するか、もしくは、話者に応じてビデオカメラの向きを変えることが必要となる。しかし、多数のビデオカメラを用いると、コストが増加するうえ接続に手間がかかる。また、話者に応じてビデオカメラの向きを変えると、そのための人員が必要となる。

そこで、会場のパノラマ映像と話者映像を1つの映像に合成することで、この問題の解決を試みたところ、良好な結果が得られたので報告する。

2. パノラマカメラの装置構成

本研究で開発したパノラマカメラは、以下の各装置から成り立っている。図1に各装置と信号の流れを示す。

- 3つの NTSC ビデオカメラを用いた映像撮影装置 (映像入力部)
- NTSC 信号デコーダ (映像入力部)
- FPGA ボード (映像処理部・制御部)

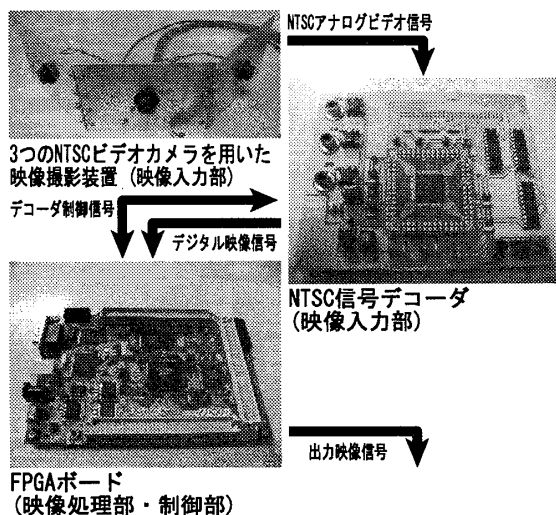


図1: 各装置と信号の流れ

[†]鳥取大学 大学院 工学研究科 知能情報工学専攻

[‡]鳥取大学 工学部 知能情報工学科

パノラマカメラの機能は、20万ゲート相当のFPGA(Field Programmable Gate Array)と呼ばれる書き換え可能な大規模集積回路と、それに独立のバスで接続された2つの36[bit]幅で容量18[Mbit]のSS-RAM(Synchronous SRAM)で実現されており、特別な外付け装置なしで利用可能である。

2.1 3つの NTSC ビデオカメラを用いた映像撮影装置

図2に示す、3つの NTSC ビデオカメラを用いた映像撮影装置とは、3つの超小型 NTSC ビデオカメラを用い、そのカメラ同士の角度が変更可能な装置である。この装置を用いることにより、容易にパノラマ映像中の3つの撮影映像の境界線の変更が可能になる。

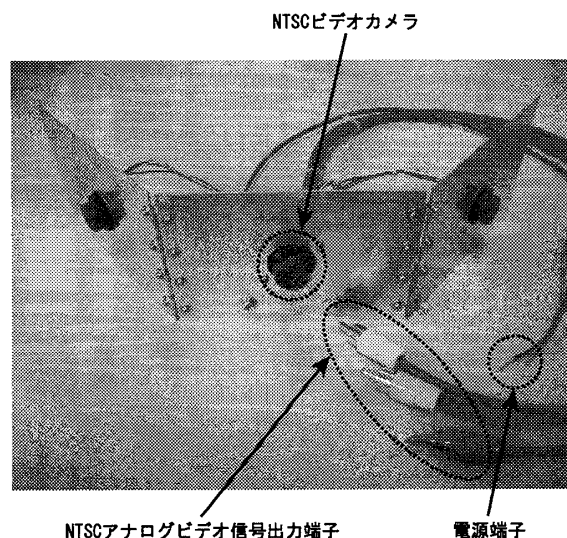


図2: 3つの NTSC ビデオカメラを用いた映像撮影装置

2.2 NTSC 信号デコーダ

3つの NTSC ビデオカメラを用いた映像撮影装置から出力される映像信号は、NTSC アナログビデオ信号であるため、そのままでは映像処理を施すのが困難である。そこで、図3に示す、NTSC アナログビデオ信号をRGB デジタル映像信号に変換する NTSC 信号デコーダを製作した。この NTSC 信号デコーダは、3チャンネルの NTSC アナログビデオ信号入力を持ち、そのチャンネルを切り替えることにより、3つの入力信号をそれぞれ個別の RGB 各 8[bit] のデジタル映像信号に変換する。

製作した NTSC 信号デコーダには、NTSC 信号デコード LSI として、沖電気工業株式会社の「MSM7664B」[1]を使用した。「MSM7664B」は、NTSC または PAL 方式のアナログビデオ信号を RGB デジタル映像信号にデコードする LSI で、コンポジット信号に関しては5チャンネルの入力を持っており、外部からの信号でチャンネル切り替えが可能である。

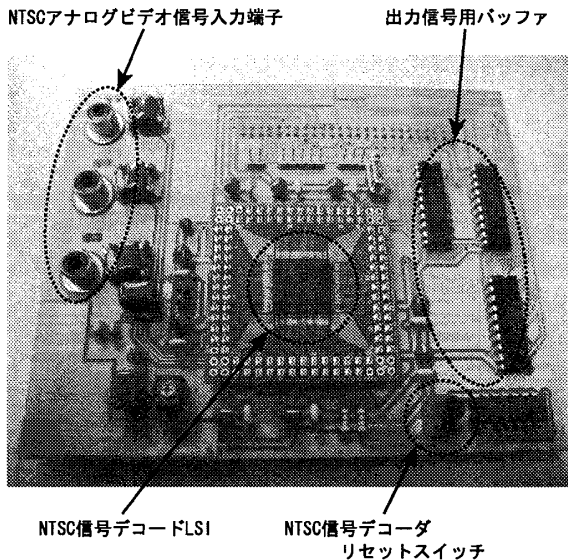


図 3: NTSC 信号デコーダ

2.3 FPGA ボード

パノラマ映像生成機能を持つテレビ会議用ビデオカメラの機能は、FPGA ボードに搭載された FPGA の内部回路やメモリなどで構成される。使用した FPGA ボードを図 4 に、また、その主な仕様を表 1 に示す。

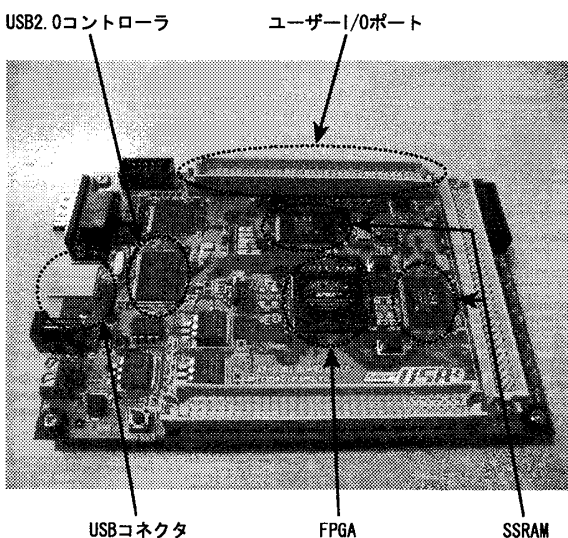


図 4: FPGA ボード

表 1: FPGA ボードの主な仕様

開発元	有限会社 プライムシステムズ
型番	AX-USB2/200C8
FPGA	ALTERA 社 EP20K200CF484C8
メモリ	GSI 社 GS8160Z36T-166
ユーザー I/O	最大 140 本
水晶発振器	48MHz

この FPGA ボードは、PC と USB2.0 インタフェースを用いて接続でき、あらかじめ用意されている PC 用の

ソフトウェアや FPGA 用の内部回路を用いることによって、容易に PC 側から、FPGA のコンフィギュレーションやメモリアクセスが可能である。

FPGA の内部回路の構築は、設計した VHDL(Verilog high speed integrated circuit Hardware Description Language) [2] [3] によるプログラムを基に、ALTERA 社のソフトウェア「QuartusII」Version.2.2 を用いて論理合成および配置配線を行った。

3. パノラマカメラの機能

パノラマカメラで行われている各処理の流れを以下に、図 5 に各処理と信号の流れを示す。

1. 3つの NTSC ビデオカメラで撮影された映像は、それぞれ個別の NTSC アナログビデオ信号として出力される。
2. カメラから出力された 3つの NTSC アナログビデオ信号は、NTSC 信号デコーダによってそれぞれ個別の RGB デジタル映像信号に変換される。
3. デジタル化された 3つの映像信号は、FPGA の内部回路である「FPGA ボード搭載メモリへのアクセス」回路によって、2つのメモリのうち撮影映像用メモリのそれぞれ個別の領域に書き込まれる。
4. メモリに書き込まれた 3つの撮影映像を基に、FPGA の内部回路である「映像サイズの縮小」回路と「出力映像の合成」回路によって、新たに出力映像が合成され、その映像は出力映像用メモリに書き込まれる。
5. 出力映像用メモリに書き込まれた出力映像は、出力映像信号として出力される。

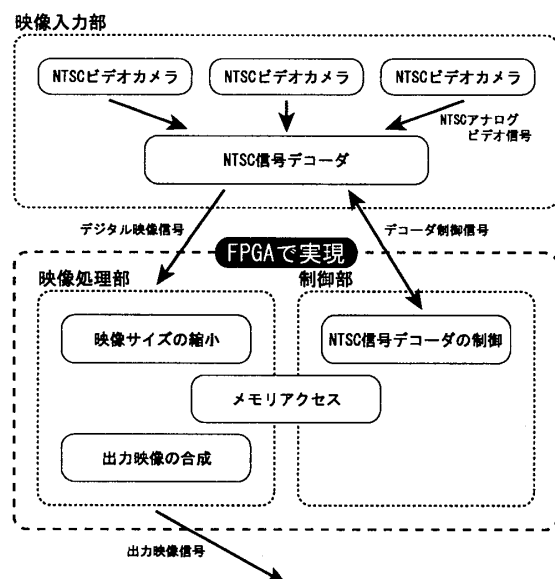


図 5: 各処理と信号の流れ

次に、FPGA で実現した機能について述べる。

3.1 メモリアクセス

FPGA の各内部回路から、FPGA ボードに搭載されている 2つの SSRAM にアクセスするためには、メモリに接続されている各信号線の信号をある決められたタイミングで変化させる必要がある。この各信号線のタイミングを制御し、FPGA の各内部回路からメモリへのアクセスを容易に行うことを可能にさせるのが、FPGA ボード搭載メモリへのアクセスを行う内部回路である。

3.2 映像サイズの縮小

パノラマ映像を生成する際、メモリに保存したままの撮影映像サイズ (320×240[pixel]) でパノラマ化すると、パノラマ映像の映像サイズが想定している出力映像の映像サイズ (640×480[pixel]) よりも大きくなる。そこで、映像サイズを $3/4(240 \times 180[\text{pixel}])$ に縮小するという対策を行った。この処理を行うのが、映像サイズの縮小を行う内部回路である。

320×240[pixel] の映像を $3/4$ サイズの 240×180[pixel] に縮小する処理を行うにあたって、図 6 に示す、縮小前の映像の画素を縦横方向にそれぞれ 4 画素に 1 画素間引くという手法を用いた。

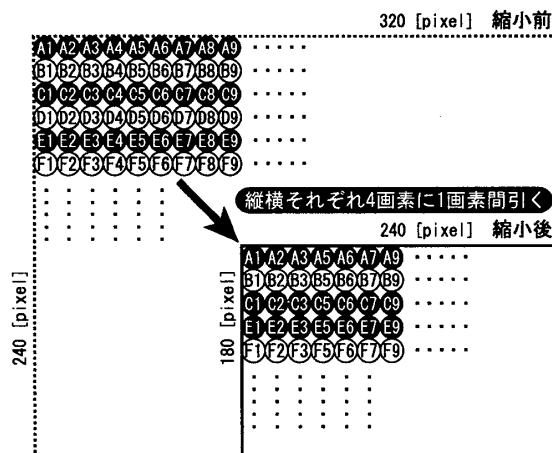


図 6: 映像サイズ縮小の様子

3.3 出力映像の合成

パノラマカメラでは、ある決められた距離と角度で配置された 3つの超小型ビデオカメラで会場を撮影し、それらの映像を $3/4$ サイズに縮小しパノラマ状につなぎ合わせたパノラマ映像と、それらの映像の中から選択された 1つの話者映像とを、合成し出力する機能を持つ。出力映像を合成する様子を図 7 に示す。

また、以下に、パノラマ映像と話者映像について説明する。

- パノラマ映像

パノラマ映像とは、ある決められた距離と角度で配置された複数のビデオカメラで撮影された映像を縮小し、一部分を重ねて合成したもので、1 台のビデオカメラで撮影した映像よりも視野角が広く、より広範囲に渡って会場の状況を撮影可能である。

- 話者映像

パノラマ映像は会場の状況を把握するのには便利だ

が、話者の細かい表情を読み取るためには縮小されたパノラマ映像では難しい。そこで、撮影された縮小される前の映像の中から話者が映っている映像を選び出し、それを話者映像とした。本研究ではこの選択を、FPGA ボード上に搭載されたプッシュスイッチによって行った。

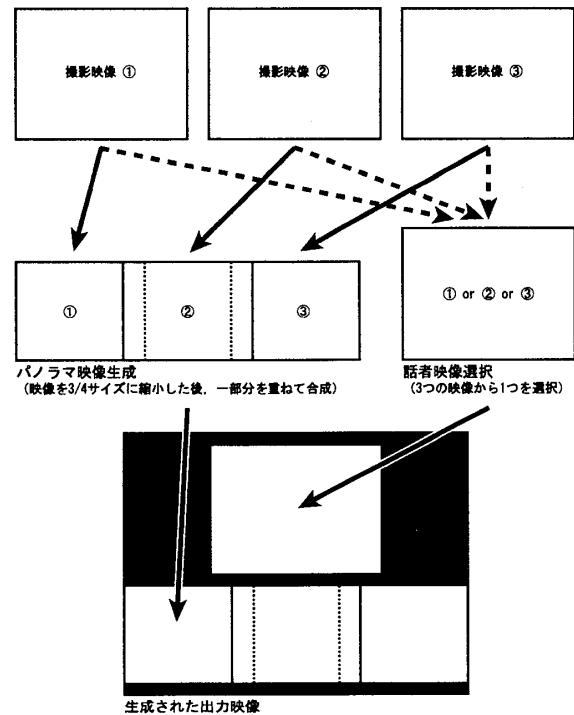


図 7: 出力映像合成の様子

これらの 2つの映像を、図 8 で示す位置に配置し、出力映像として合成するのが、出力映像の合成を行う内部回路である。

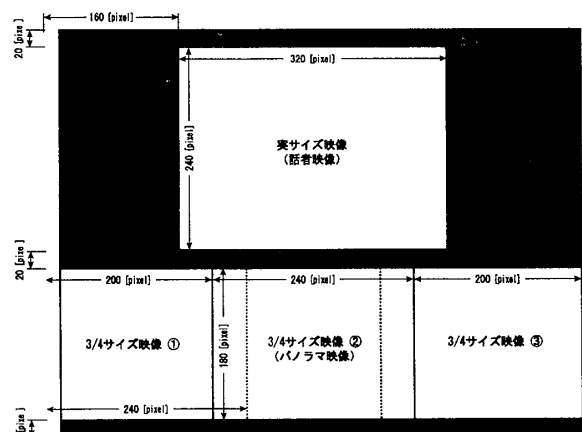


図 8: 出力映像の合成

出力映像の合成は、図 9 に示すように、撮影映像用メモリから出力映像用メモリに映像データをコピーする際に行われる。

まず、映像サイズの縮小を行いながら出力映像用メモ

りの決められたアドレスに、3つの映像データがコピーされ、パノラマ映像が生成される。続いて、選択された話者映像の映像データが出力映像用メモリの決められたアドレスにそのままコピーされる。

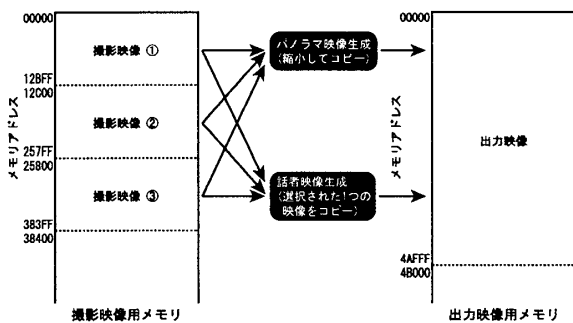


図9: 出力映像合成時のメモリ内部の様子

なお、メモリ上で映像データは、画面の左下から右上へと向かって、下位ビットから BLUE, GREEN, RED の順に各色 8[bit] で保存されており、上位 12[bit] は未使用である。例として、サイズが 8×6[pixel] の映像データのメモリ上での様子を、図10に示す。

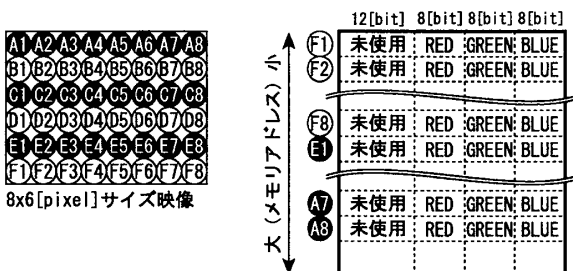


図10: サイズが 8×6[pixel] の映像データのメモリ上での様子

3.4 NTSC 信号デコーダの制御

NTSC 信号デコーダは、3チャンネルの NTSC アナログビデオ信号入力を持っているが、このチャンネル切り替えを行うには、外部から NTSC 信号デコーダに対して決められた信号を送る必要がある。また、NTSC 信号デコーダの電源投入時には、リセット信号を送らなければならない。NTSC 信号デコーダの制御を行う内部回路は、NTSC デコーダに接続された各信号線を使用して、リセットの実行や、適切なタイミングで NTSC 信号デコーダからの映像信号を取り込み、それをメモリに保存することを行う。

なお、NTSC 信号デコーダから出力される映像の有効な垂直ラインは、奇数および偶数フィールドを合わせて計 505 本、また、有効な水平画素は 640[pixel] である。そのため、垂直ラインについては奇数フィールドのみ、また、水平画素については 1[pixel] おきに信号を取得することにより、320×240[pixel] の撮影映像を得た。

4. 動作実験

本研究の動作実験を行うために、PC 上で動作するソフトウェアを作成した。このソフトウェアは、FPGA が

出力映像の合成を完了する度に出力映像用メモリから出力映像データを読み込み、PC の画面上に表示するものである。また、開発した処理回路を FPGA 内部へ書き込むコンフィグレーション作業も行う。

実際に動作実験中の画面の様子を、図11に示す。画面下部の映像は3つの撮影映像を縮小し生成されたパノラマ映像である。また、画面上部の映像は3つの撮影映像の中から FPGA ボード上のプッシュスイッチによって1つ選択された話者映像である。

映像表示速度の測定を行ったところ、現在は 5[フレーム/秒] 程度であり、実際に人物が動いてから画面に反映されるまでに若干の遅延があった。



図11: 動作実験中の画面

5. おわりに

本研究では、3つの NTSC ビデオカメラを用いた映像撮影装置、NTSC 信号デコーダ、FPGA の内部回路を開発し、パノラマカメラの基本的な機能を実現したが、現在のところ、話者映像の選択はプッシュスイッチによる手動操作のため、実際に人員の削減という状況にはまだ至っていない。

そのため今後は話者映像の選択を、現在の手動操作ではなく自動化することを目指す。その具体的な方法としては、複数のマイクロフォンを設置し取得した音声を FPGA の内部で処理することで、話者の方向を推定することを考えている。また、現在の映像表示速度は 5[フレーム/秒] 程度であるため、その高速化のためにメモリアクセス方法を改良し、メモリへのアクセス速度を向上させたい。そして、パノラマ映像を生成する際、3つの NTSC ビデオカメラ映像の境界線を自動的に検出する内部回路を FPGA 上に構築することを考えている。

参考文献

- [1] 沖電気工業株式会社：MSM7664B データシート，沖電気工業株式会社 (2001)
- [2] 長谷川 裕恭：VHDL によるハードウェア設計入門，CQ 出版社 (2002)
- [3] 森岡 澄夫：HDL による高性能デジタル回路設計，CQ 出版社 (2002)