

テレビ会議場のパノラマ映像生成機能を持つビデオカメラの開発 Video Camera for TV Conference with a Panoramic Image Generation Function

長瀬 幸規 †

川村 尚生 ††

菅原 一孔 ††

Yukinori NAGASE †

Takao KAWAMURA ††

Kazunori SUGAHARA ††

† 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

1 はじめに

本研究では、テレビ会議場のパノラマ映像生成機能を持つ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には、そのための人員が必要となる。

そこで本稿では、会議場のパノラマ映像と話者映像を1つの映像に合成する機能をビデオカメラに組み込んだパノラマカメラを開発することでこの問題の解決を試みた。

2 パノラマカメラの装置構成

図1のように、開発したパノラマカメラは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ボード、NTSC信号エンコーダから成り立っている。

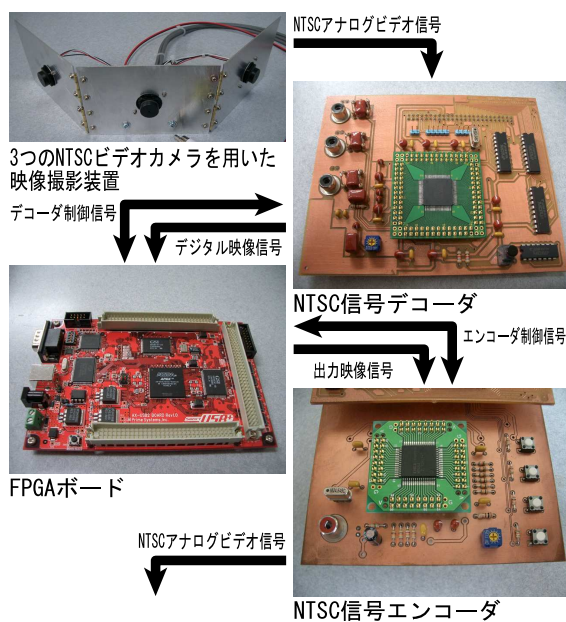


図1: 各装置と信号の流れ

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。

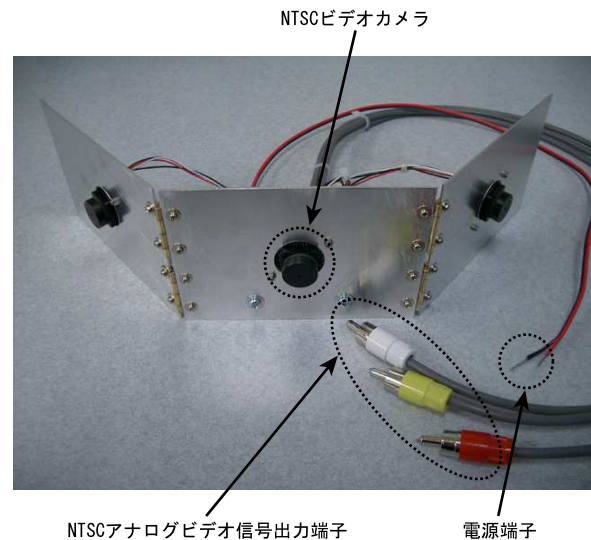


図2: 3つのNTSCビデオカメラを用いた映像撮影装置

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信号デコーダ回路を製作した。この回路で使用したNTSC信号デコーダLSI [1]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RGB各8[bit]のデジタル映像信号に変換する。

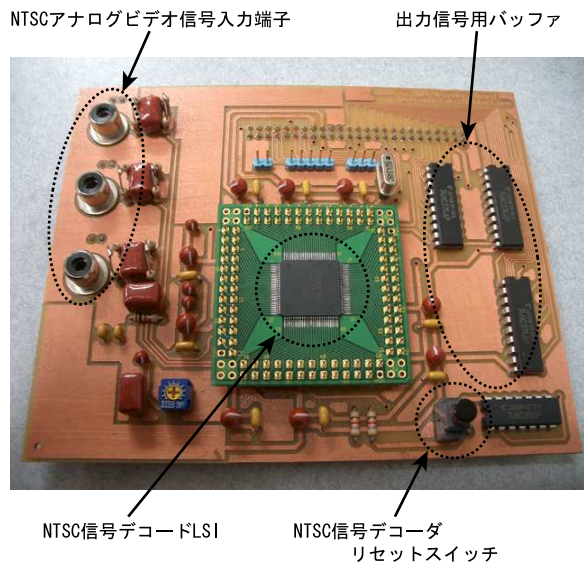


図 3: NTSC 信号デコーダ

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 に示す。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 のコンフィグレーションやメモリアクセスが可能である。

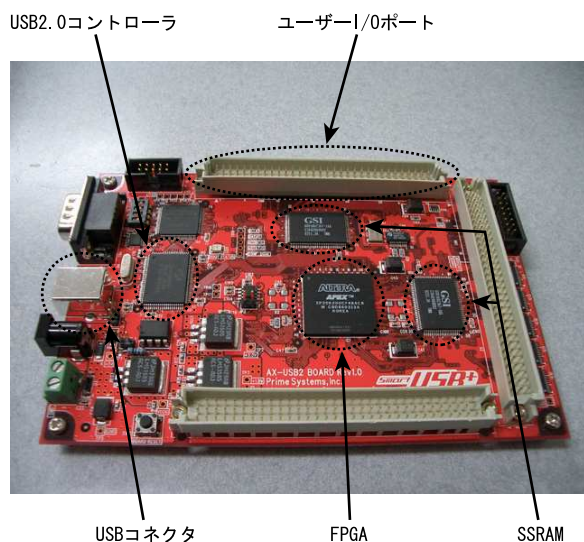


図 4: FPGA ボード

FPGA の内部回路の構築は、設計した VHDL (Very high speed integrated circuit Hardware Description Language) [3] [4] 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。

2.4 NTSC 信号エンコーダ

FPGA 内部で合成された出力映像は RGB デジタル映像信号であるために、その映像をテレビモニタなどに表示させるためには、NTSC アナログビデオ信号に変換する必要がある。そこで図 5 に示す、NTSC 信号エンコーダ回路を製作した。この回路で使用した NTSC 信号エンコーダ LSI [2] は、RGB 各 8[bit] のデジタル映像信号を NTSC アナログビデオ信号に変換する。

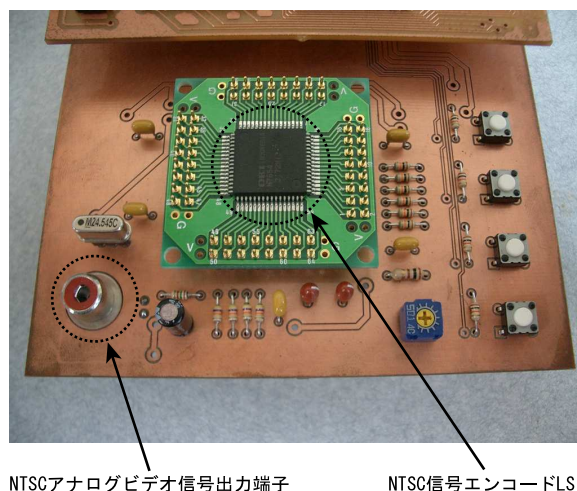


図 5: NTSC 信号エンコーダ

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 6 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、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各信号は、NTSC 信号デコーダによって 640×480 [pixel] の RGB デジタル映像に変換される。
3. デジタル化された 3つの映像は 320×240 [pixel] に縮小され、FPGA ボード上の撮影映像用メモリに独立して書き込まれる。
4. メモリに書き込まれた 3つの撮影映像を基に、FPGA の内部回路を用いて映像サイズの縮小および出力映像の合成が行われ、その映像は 640×480 [pixel] のサイズで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、NTSC 信号エンコーダによって再び NTSC アナログビデオ信号に変換され、出力される。

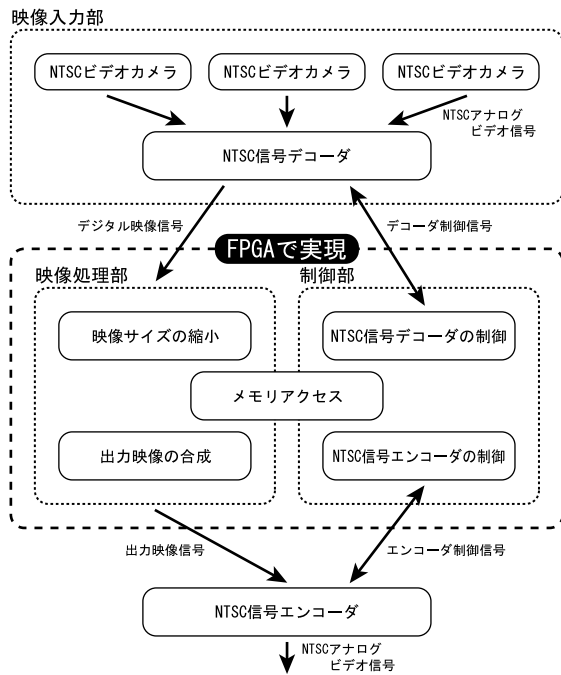


図 6: 各処理と信号の流れ

次に、FPGA で実現した主要な機能について述べる。

3.1 パノラマ映像の合成

合成したパノラマ映像が NTSC アナログビデオ信号の規格に合うよう、3 つの撮影映像を 240×180 [pixel] に縮小し、それを合成することとした。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。

取り込んだ 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 7 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。

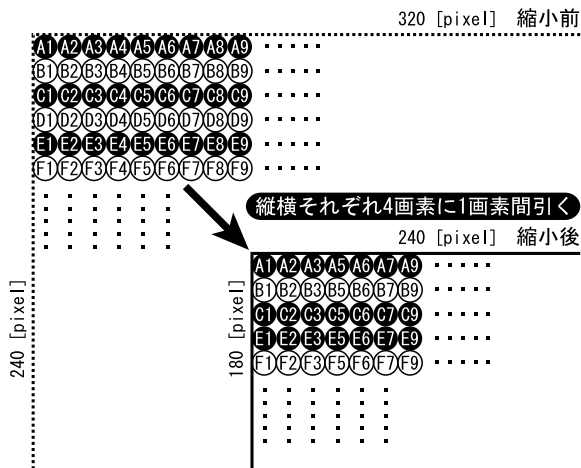


図 7: 映像サイズ縮小の様子

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繋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 8 に示す。

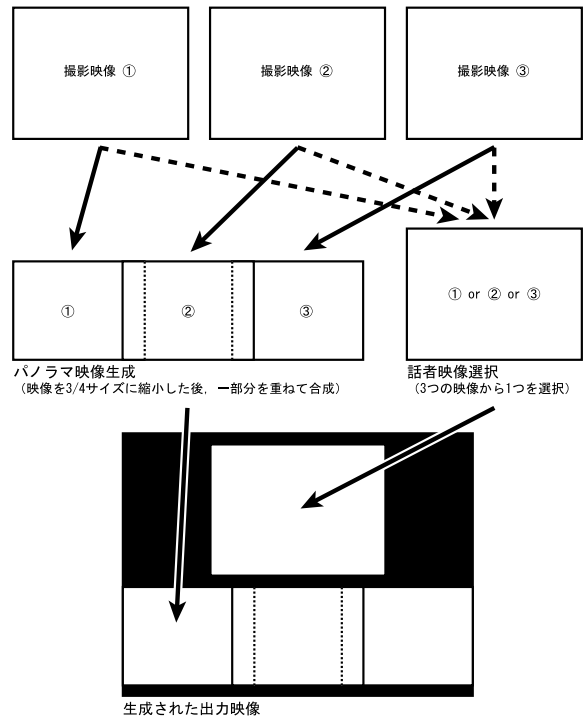


図 8: 出力映像合成の様子

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。

これらの 2 つの映像を、図 9 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。

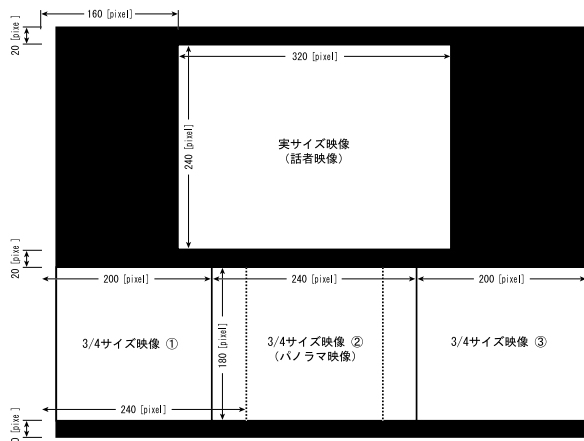


図 9: 出力映像の合成

出力映像の合成は、図 10 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リ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。

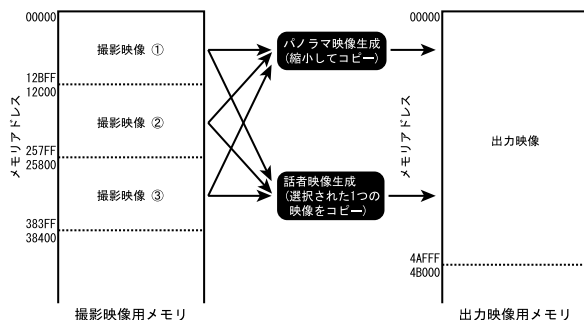


図 10: 出力映像合成時のメモリ内部の様子

3.2 NTSC 信号デコーダ、エンコーダの制御

NTSC 信号デコーダ、エンコーダの出力するタイミング信号に合わせて、デジタル映像信号の取り込みまたは出力する処理を行うのが、NTSC 信号デコーダ、エンコーダの制御を行う内部回路である。

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

動作実験は、装置から出力された NTSC アナログビデオ信号をテレビモニタに入力することで行った。

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、NTSC 信号エンコーダ、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、実用段階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。

6 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] 沖電気工業株式会社：MSM7654 データシート，沖電気工業株式会社 (2003)
- [3] 長谷川 裕恭：VHDL によるハードウェア設計入門，CQ 出版社 (2002)
- [4] 森岡 澄夫：HDL による高性能デジタル回路設計，CQ 出版社 (2002)