

パノラマ映像生成機能のハードウェア実現について

長瀬 幸規[†] 川村 尚生^{††} 菅原 一孔^{††}

[†] 鳥取大学 大学院 工学研究科 〒680-8552 鳥取県鳥取市湖山町南 4-101

^{††} 鳥取大学 工学部 〒680-8552 鳥取県鳥取市湖山町南 4-101

E-mail: †{nagase,kawamura,sugahara}@ike.tottori-u.ac.jp

あらまし 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。そこで、会議場を複数のカメラで撮影した映像を滑らかに結合したパノラマ映像と話者が存在する映像を 1 つの映像に合成する機能を持つビデオカメラを FPGA を用いて試作したので報告する。

キーワード テレビ会議, FPGA

Hardware Realization of Panoramic Image Generation Function

Yukinori NAGASE[†], Takao KAWAMURA^{††}, and Kazunori SUGAHARA^{††}

[†] Graduate School of Engineering, Tottori University

4-101 Koyama-Minami, Tottori-City, Tottori, 680-8552 Japan

^{††} Faculty of Engineering, Tottori University

4-101 Koyama-Minami, Tottori-City, Tottori, 680-8552 Japan

E-mail: †{nagase,kawamura,sugahara}@ike.tottori-u.ac.jp

Abstract In this paper, a video camera for TV conferences are proposed. Proposed camera has panoramic image generation functions to transmit both atmospheres of conference places and speaker's expressions. The camera is implemented as hardware circuits by using FPGA to construct a stand alone system. These panoramic image generation functions are confirmed by the experiments.

Key words TV conference, FPGA

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、会場の雰囲気を知ろうとして会場全体を広範囲に渡って撮影すると、映像中の話者が小さくなるため話者の表情を読み取ることが難しくなる。また逆に、話者の表情を読み取ろうとして話者が大きくなるように撮影すると、会場の狭い範囲しか撮影できなくなるため会場の雰囲気を知ることが難しくなる。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、会議場を複数のカメラで撮影した映像を滑らかに結合したパノラマ映像と話者が存在する映像を 1 つの映像に合成する機能を持つビデオカメラを FPGA を用いて試作したので報告する。

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 [1] [2]、NTSC 信号エンコーダから構成されている。

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 SSRAM(Synchronous SRAM) [3] 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

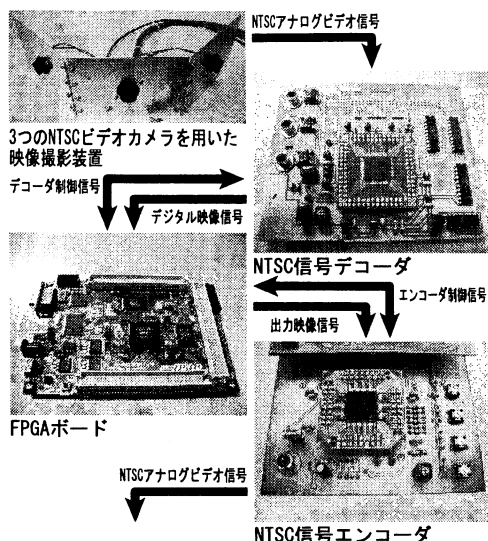


図1 パノラマカメラの装置構成

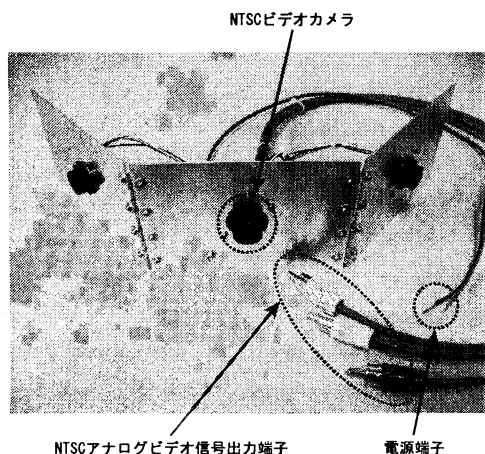


図2 3つのNTSCビデオカメラを用いた映像撮影装置

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。

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[6][7][8]。この回路で使用了NTSC信号デコーダLSI[4]は、4チャンネルのNTSCアナログビデオ信号入力を持ち、そのうちの3個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。

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は、

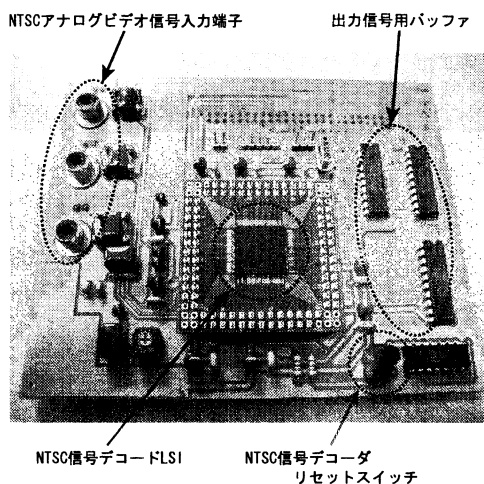


図3 NTSC 信号デコーダ

FPGA ボードに搭載されたFPGAの内部回路やメモリなどで実現される。使用したFPGAボードを図4に示す。このFPGAボードは、パーソナルコンピュータ（以下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。

FPGAの内部回路の構築は、VHDL(Verilog high speed integrated circuit Hardware Description Language)[9][10]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。

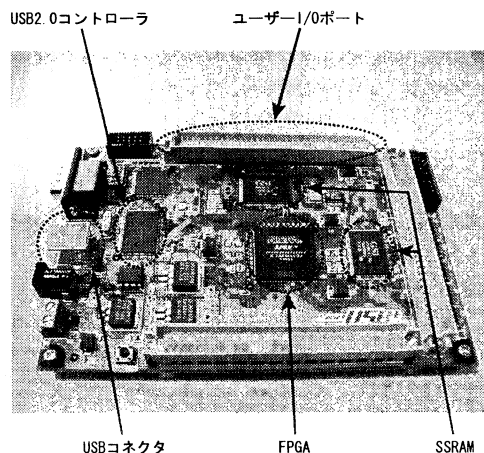
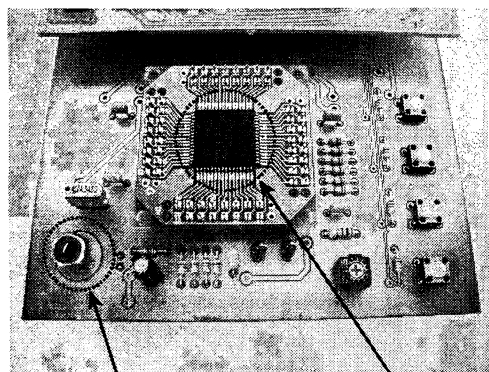


図4 FPGA ボード

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、その映像をテレビモニタなどに表示させるには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[5] 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。



NTSCアナログビデオ信号出力端子 NTSC信号エンコードLSI

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、境界線の自動検出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。

3.1 パノラマ映像と話者映像

パノラマカメラは、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。

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便利であるが、話

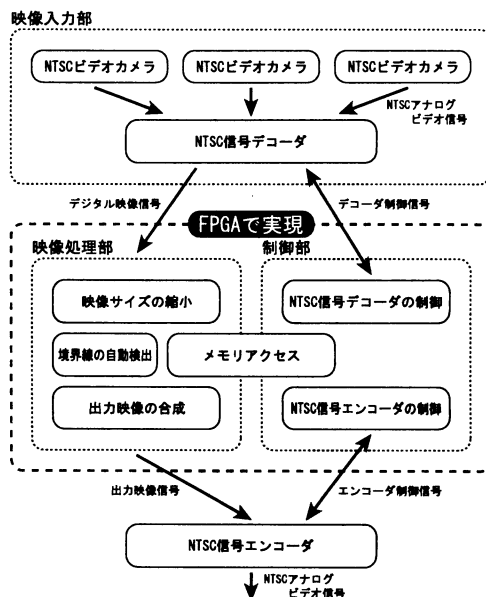


図 6 各処理と信号の流れ

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ブッシュスイッチによって行った。

3.2 映像サイズの縮小

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。

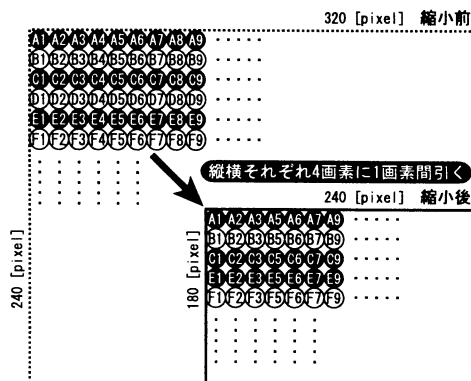


図 7 映像サイズ縮小の様子

3.3 3つのNTSCビデオカメラ映像の境界線の自動検出

パノラマ映像を合成する際、境界線を滑らかに結合するために、3つのNTSCビデオカメラ映像の境界線を自動検出する処理を行った。この処理を行うのが、3つのNTSCビデオカメラ映像の境界線の自動検出を行う内部回路である。

例として、中央と左の映像の境界線を検出する手順を以下に説明する。また、その様子を図8に示す。

(1) 中央映像の左端にある縦方向計180画素の組を基準値として、左映像の右上端の画素から順に中央映像の対応する画素とのRGB値のR値の差を求めて、その絶対値の合計を計算する。

(2) 左映像の右端にある縦方向計180画素の組から順番に、左方向に向かって50組分(1)の計算を行う。

(3) 50組分の計算結果のうち、絶対値の合計が最小である位置が最適な境界線であると決定する。

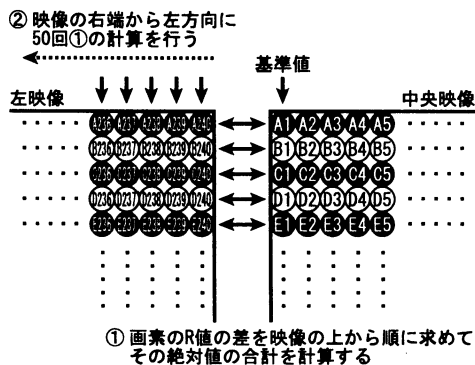


図8 境界線の自動検出の様子

3.4 出力映像の合成

出力映像を合成する様子を図9に示す。

パノラマ映像と話者映像の2つの映像を、図10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。

出力映像の合成は、図11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ことで行われる。

まず、3つのNTSCビデオカメラ映像の境界線の自動検出結果を元に、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。

3.5 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

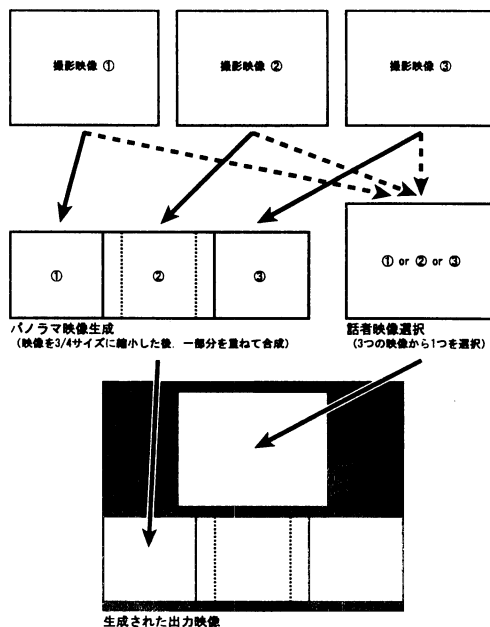


図9 出力映像合成の様子

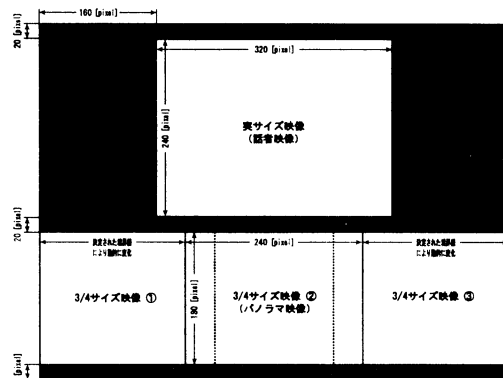


図10 出力映像の合成

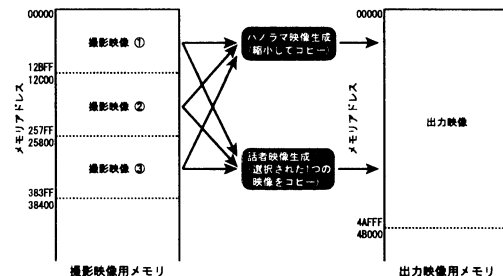


図11 出力映像合成時のメモリ内部の様子

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 12 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 12 動作実験中の画面

また、3 つの NTSC ビデオカメラ映像の境界線の自動検出についても実験を行った。境界線の自動検出を行う前の画面の様子を図 13 に、また、境界線の自動検出を行った後の画面の様子を図 14 にそれぞれ示す。

これらの図より、境界線の自動検出の効果が確認できる。



図 13 境界線の自動検出を行う前の画面



図 14 境界線の自動検出を行った後の画面

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。また、3 つの NTSC ビデオカメラ映像の境界線の自動検出についても、ほぼ最適な結果が得られた。しかし、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で 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。

文 献

- [1] 有限会社 プライムシステムズ：Smart-USB Plus AX-USB2 ボード 取扱説明書，有限会社 プライムシステムズ (2002)
- [2] 有限会社 プライムシステムズ：Smart-USB Plus GPIF 版 インターフェース仕様書，有限会社 プライムシステムズ (2002)
- [3] Giga Semiconductor, Inc. : GS8160Z36T-166 データシート，Giga Semiconductor, Inc. (2002)
- [4] 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)
- [5] 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)
- [6] CQ 出版社：トランジスタ技術 2003 年 6 月号，CQ 出版社 (2003)
- [7] Altium Limited：Introducing Protel DXP，Altium Limited (2003)
- [8] アルティウムジャパン株式会社：Introducing Protel DXP チュートリアル，アルティウムジャパン株式会社 (2003)
- [9] 長谷川 裕恭：VHDL によるハードウェア設計入門，CQ 出版社 (2002)
- [10] 森岡 遼夫：HDL による高性能デジタル回路設計，CQ 出版社 (2002)