

テレビ会議用ビデオカメラのパノラマ映像生成に関する研究

上杉 徹[†] 米本 良[†] 川村 尚生[‡] 菅原 一孔[‡]

[†]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

本研究室で開発しているテレビ会議用のビデオカメラは、会場全体の映像を表示するために、3つのNTSCビデオカメラを用いてパノラマ映像を生成している。本稿では、パノラマ映像を合成する手法として、Alistair J. Fitchらが提案したFast Robust Correlation手法[1]を取り上げ、そのハードウェア実現について述べる。

2. システムの構成

処理の流れは以下の通りである。まず、3つのNTSCビデオカメラから送られてきたNTSCアナログ映像信号をNTSCビデオデコーダLSIによってRGBデジタル映像信号に変換し、FPGAボード上のメモリに蓄積する。そして、FPGA上の各回路において、入力された映像データを基にパノラマ映像生成を行い、デジタルビデオエンコーダでNTSCアナログ映像信号に変換しモニタに表示する。

3. パノラマ映像生成

パノラマ映像は、3つのNTSCビデオカメラで撮影した映像の一部分を重ねて生成している。またパノラマ映像生成の際に、3つの映像の境界を上手く繋ぎ合わせるために、左と中央、中央と右の映像の各々の境界領域で結合位置を自動で検出し、最適な位置に映像を補正する自動補正処理を行っている。

処理方法は以下の通りである。まず、画像 f に対して x 軸方向に m 、 y 軸方向に n ずらした画像 g があると仮定する。このとき、 x, m を $x = (x, y), m = (m, n)$ と定義し、2つの画像 f, g の全体としての画像の差を次のように表す。

$$R(x) = \sum_x h(f(x) - g(x - m)) \alpha_f(x) \alpha_g(x - m) \quad (1)$$

ここで、式(1)に示している $\alpha_f(x)$ と $\alpha_g(x - m)$ はアルファマスクである。また、式(1)の $h(r)$ の値が常に正の値を取るように、 $h(r) \approx \sum_{p=1}^P b_p (1 - \cos(a_p \pi r))$ という偶関数で近似することとする。この式を式(1)に

代入して整理すると式(2)が得られる。

$$R(m) = \sum_x [(\alpha_f(x) \star \alpha_g(x)) \sum_{p=1}^P b_p - \sum_{p=1}^P b_p ((\alpha_f(x) e^{j a_p \pi f(x)}) \star (\alpha_g(x) e^{j a_p \pi g(x)}))] \quad (2)$$

式(2)の \star はコンボリューションを示している。ここで、式(2)を、FFTを用いて求め、そうして得られた $R(m)$ について最小値を求め、最小値となる (m, n) により、パノラマ映像の自動補正処理を行う。

また、FFTのハードウェア化の際、1つのバタフライ演算につき、実数の乗算4回と実数の加減算が6回必要であり、1回の乗算に5クロック、1回の加減算に8クロック必要なため、その計算には68クロック必要である。128×256の画像2枚に対し、式(2)の計算を行うには、おおよそ 1.5×10^8 クロック必要で、48MHzのクロックでは約2.9秒の処理時間が必要となる。

4. 実験結果

図1にNTSCビデオカメラより撮影した3つの入力映像から生成されたパノラマ映像を示す。なお、図1に示す2本の点線は結合位置である。



図1: 生成されたパノラマ映像

5. おわりに

本研究では、パノラマ映像生成の自動補正処理を行うためにFitchらのFast Robust Correlation手法[1]のハードウェア実現を試みた。

参考文献

- [1] Alistair J. Fitch, Alexander Kadyrov, William J. Christmas and Josef Kittler, "Fast Robust Correlation," IEEE (2005).