

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

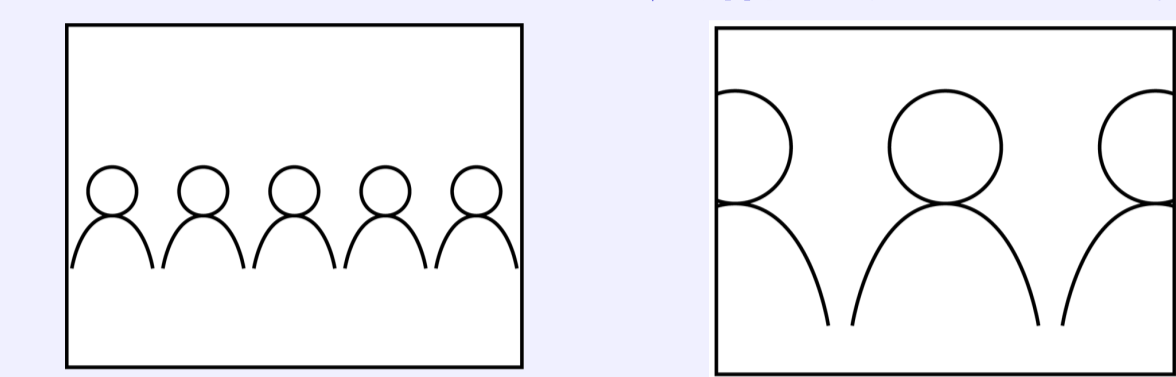
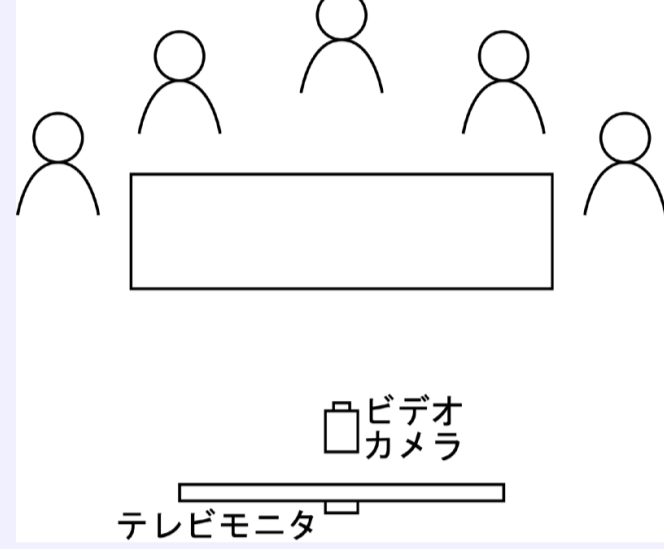
鳥取大学大学院工学研究科 上杉 徹, 米本 良
鳥取大学工学部 川村 尚生, 菅原 一孔

1. はじめに

通常、小規模な会議会場の間でのテレビ会議では、
固定された1台のビデオカメラで会場を撮影

【問題点】

会場全体を撮影 表情が分かるように撮影



話者の表情を読み取ることが難しい!!
会場の狭い範囲しか撮影できない!!

【解決策】

会場全体を複数台のビデオカメラで撮影

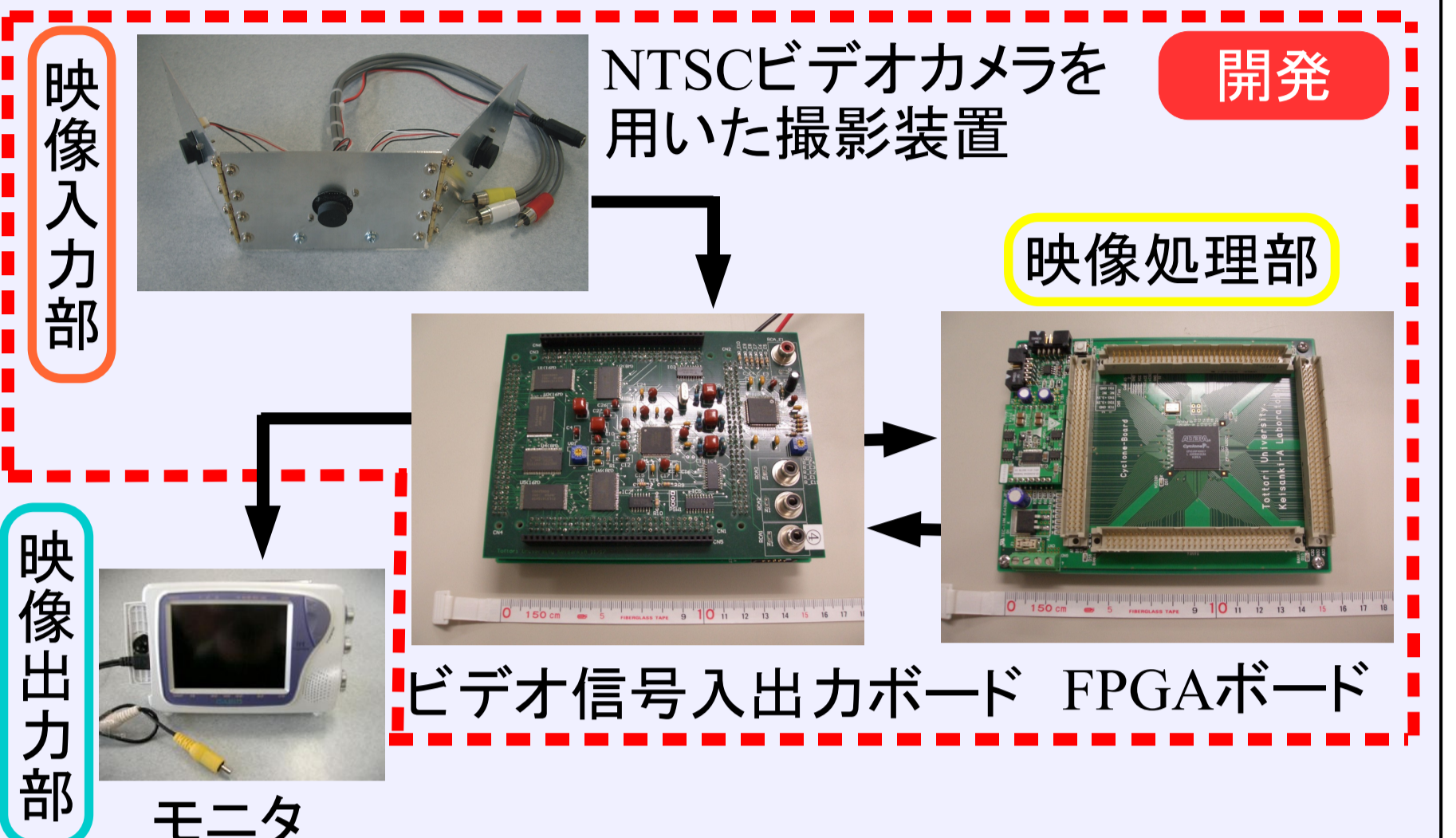
テレビ会議用ビデオカメラ



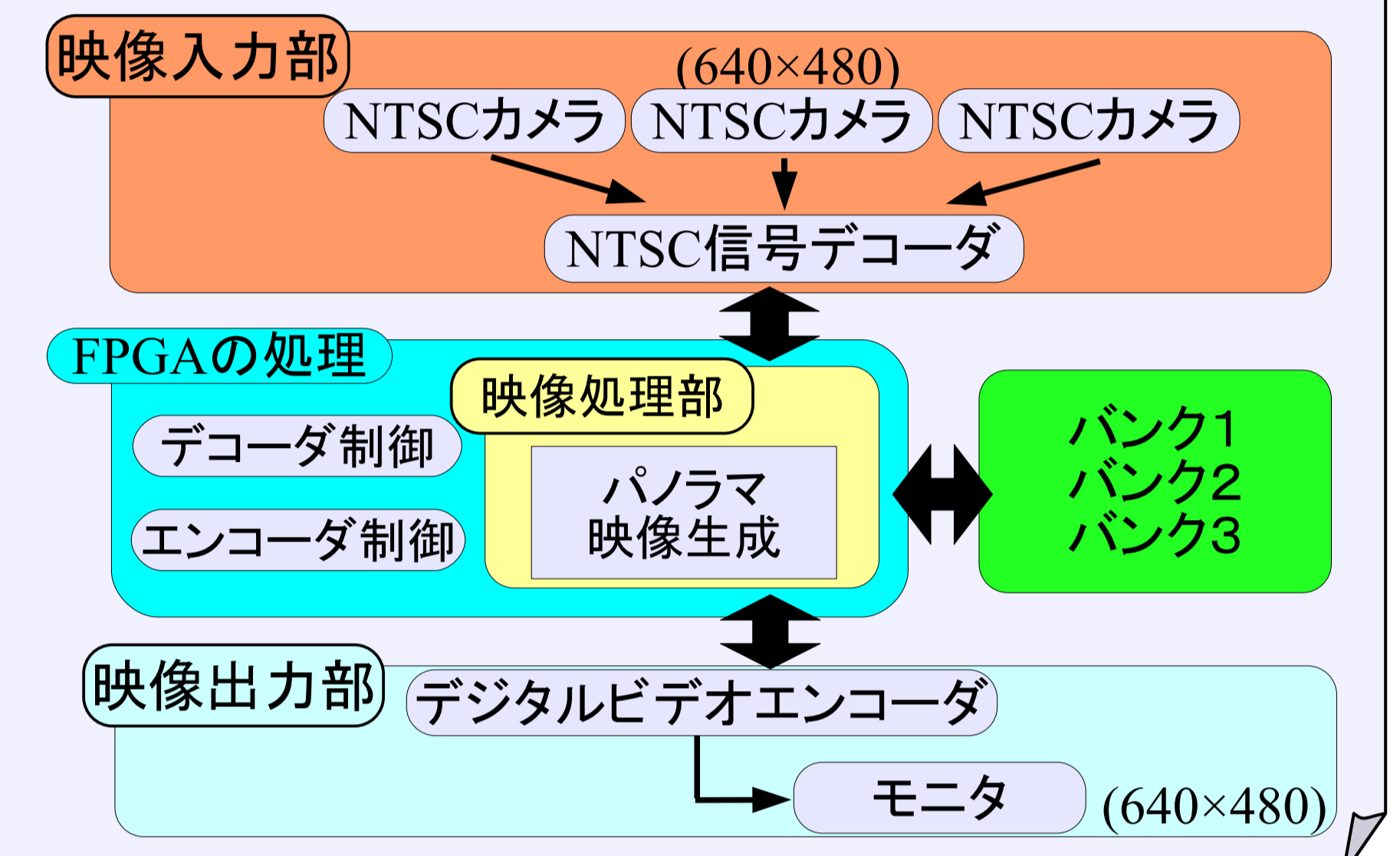
パノラマ映像を生成する際に、映像の
繋ぎ目を上手く合成する必要がある
その手法として、

Fast Robust Correlation 手法のハードウェア実現

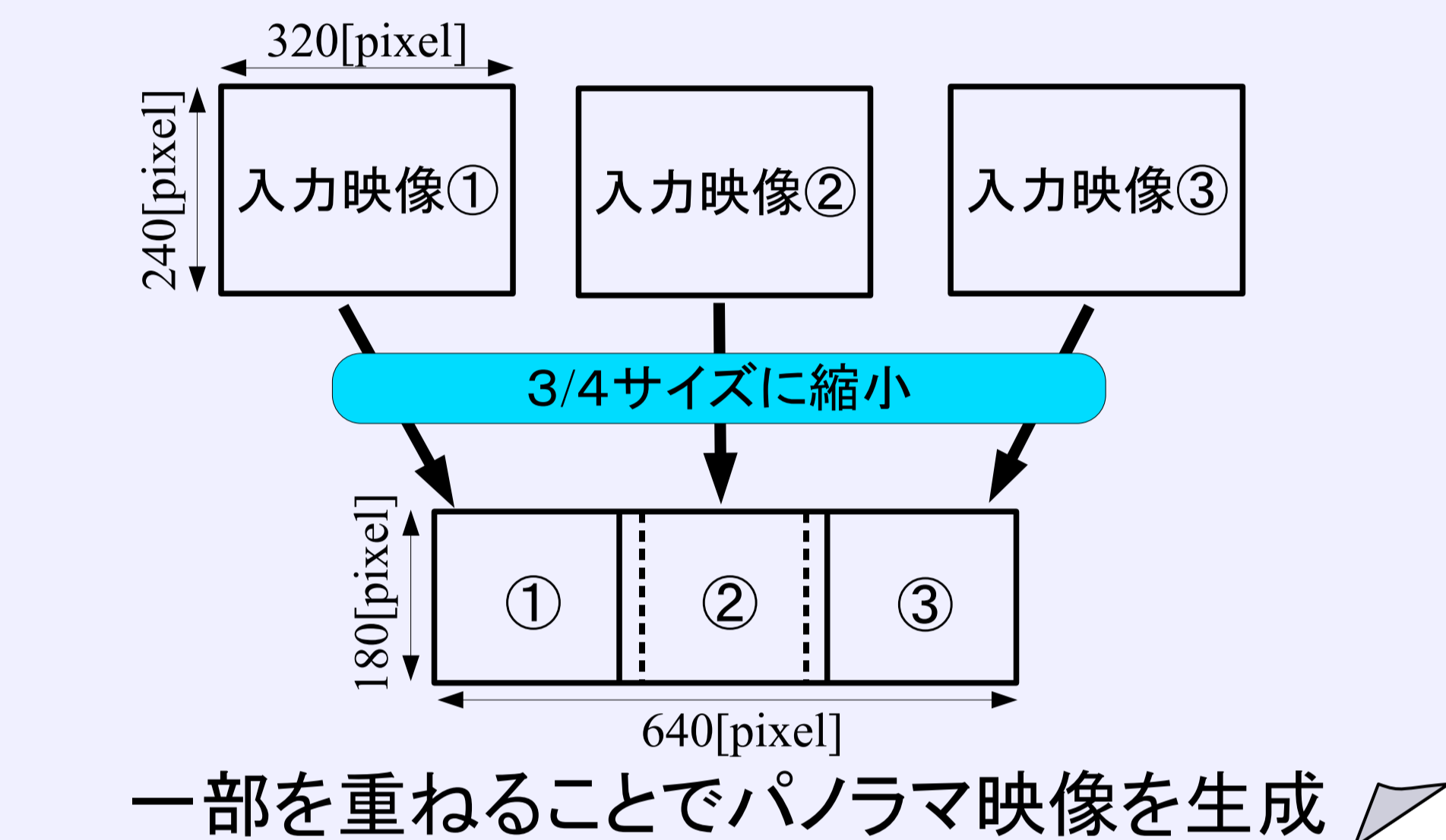
2. システム構成



【処理の流れ】

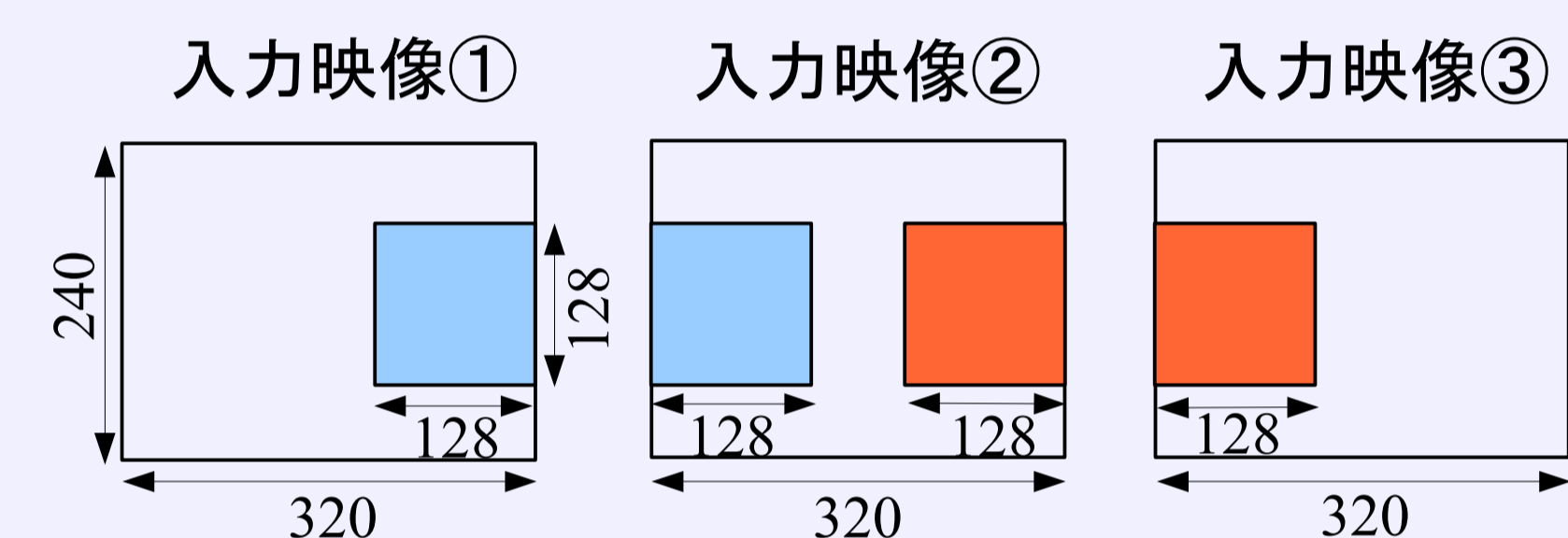


3. パノラマ映像生成

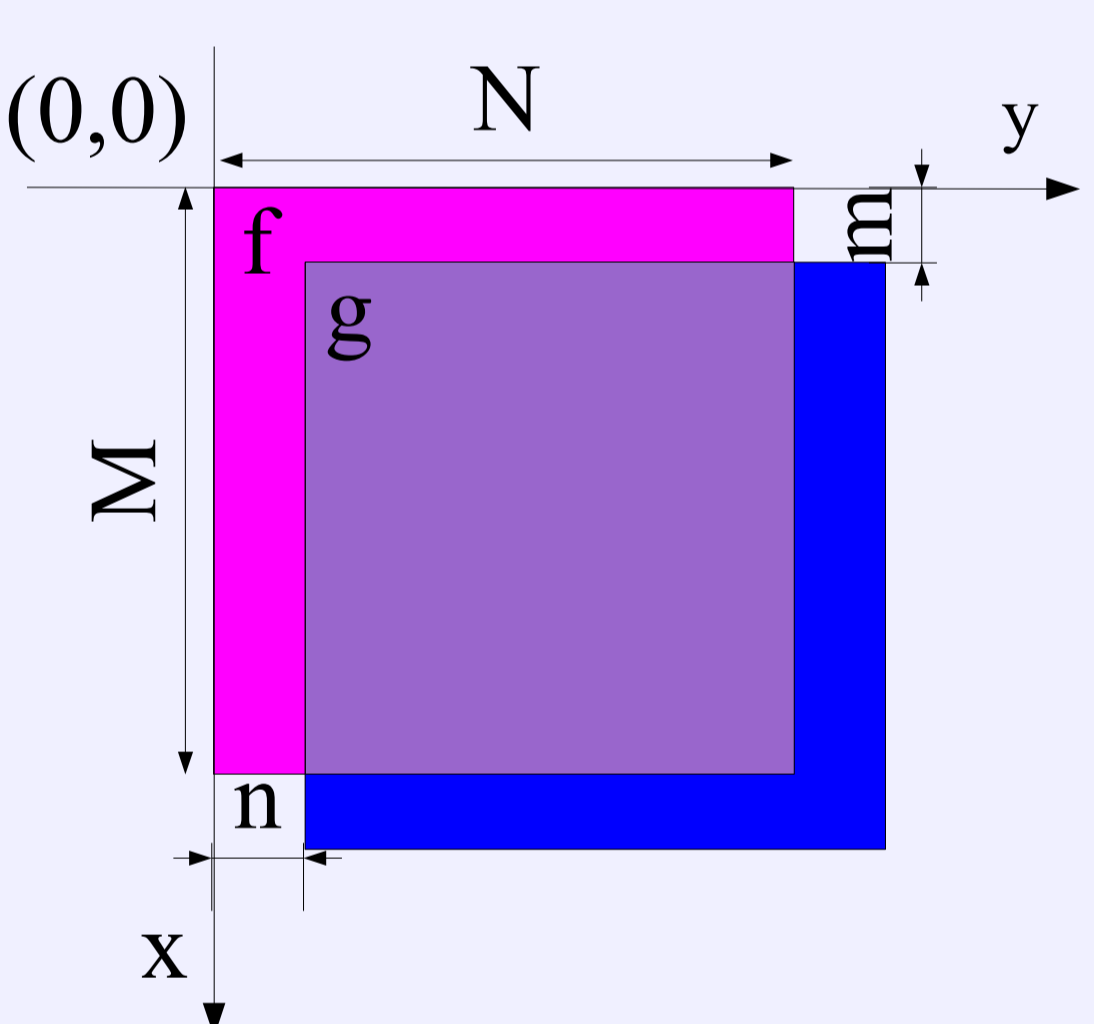


4. 自動補正処理

入力映像①と②, ②と③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



5. Fast Robust Correlation手法



f, gの画素差を **アルファマスク**
 $R = f(x, y) \cdot \alpha_f(x, y) - g(x-m, y-n) \cdot \alpha_g(x-m, y-n)$

f, gの全体としての画像の差は
 $R(m, n) = \sum_{x, y} h(f(x, y) - g(x-m, y-n)) \cdot \alpha_f(x, y) \alpha_g(x-m, y-n)$
偶関数で近似

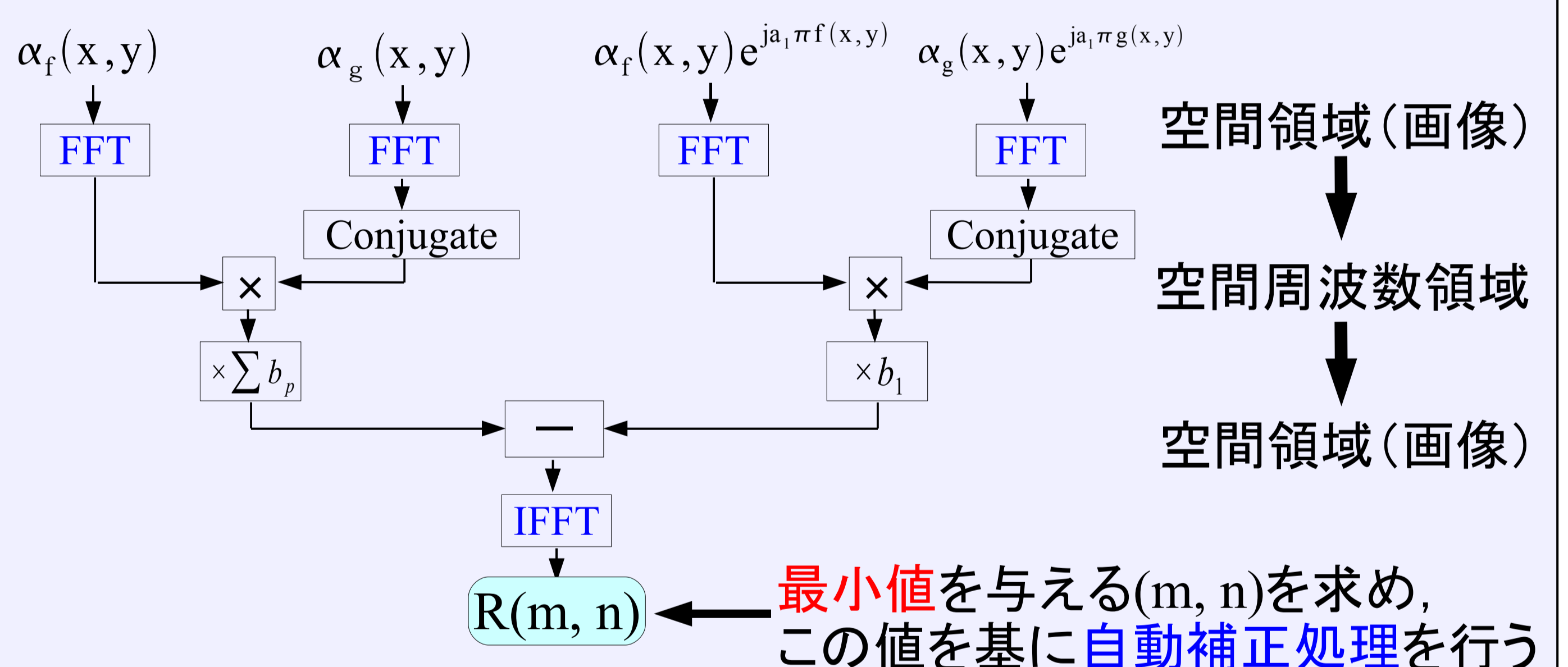
$$h(r) \approx \sum_{p=1}^P b_p (1 - \cos(a_p \pi r))$$

パラメータ

h(r)をR(m, n)に代入して整理すると

$$R(m, n) = \sum_{x, y} [\alpha_f(x, y) \alpha_g(x, y) \sum_{p=1}^P b_p - \sum_{p=1}^P b_p ((\alpha_f(x, y) e^{ja_p \pi f(x, y)}) (\alpha_g(x, y) e^{ja_p \pi g(x, y)}))]$$

【計算の流れ】

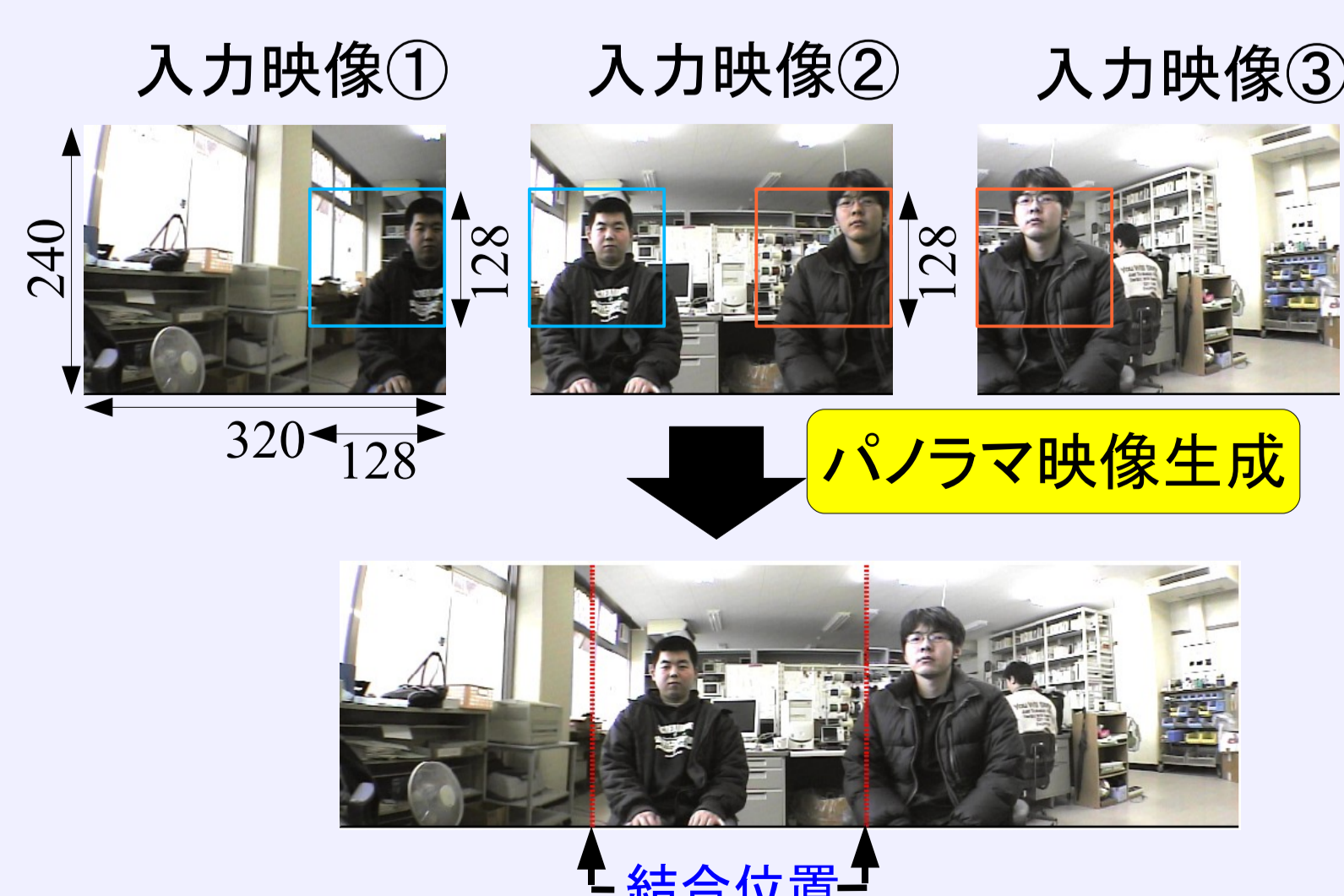


なお、FFTをハードウェア化するには**バタフライ演算**をハードウェア化する必要がある

バタフライ演算を繰り返し行うことで**FFTのハードウェア化**を実現

6. 実験結果

【動作確認】



【回路規模】

Logic Element 数 :20060

映像入力制御部	219
パノラマ映像生成部	12409
映像出力制御部	329
メモリ等	252
合計	13209

使用可能なLogic Element 数の約65%で実現

【処理速度】

- FPGA
 - 動作周波数 :48[MHz]
 - 自動補正処理(片側)
 - クロック数 :約 4.9×10^7
 - 処理時間 :約1.01[秒]
 - パノラマ映像生成
 - クロック数 :約 1.1×10^8
 - 処理時間 :約2.3[秒]

7. FPGAとは

【FPGA】

回路の書き換えが可能なLSIであり、高速動作や並列処理を得意としている。このLSI上で各種映像処理を行っている。

【FPGAの主な仕様】

開発元	ALTERA
型番	EP1C20F400C7
Logic Element 数	20060
最大I/Oポート数	301
パッケージ	400-Pin FineLine BGA
サイズ	21[mm]角