

J-17 ドキュメントを対象としたビデオモザイクの検討

A Study of Video Mosaicing for Documents

杉岡 賢十
Ken Sugioka

山地可城十
Yoshiki Yamaji

石川和弘
Kazuhiro Ishikawa

1. まえがき

昨今では、カメラの付いた携帯電話やPDAが急速に普及してきている。このようなモバイルカメラを利用し、文字が主体として構成されるドキュメントを撮影する場合には、解像度の低さや視野の狭さなど画像品質の点で問題がある。すなわち、ドキュメント全体を撮影フレームへ収まるように撮影すると文字の判読が困難となり、逆に文字を判読可能な状態まで接写するとドキュメント全体が撮影フレームに収まらない。

そこで、カメラを接写した状態で移動させながら撮影した動画からフレーム単位に切り出し、位置決め、貼り合わせを繰り返しながら1枚の広域な静止画を生成するビデオモザイクに着目した。モバイルカメラを用いて、ドキュメントを対象としたビデオモザイクを行うには以下のような課題がある。

- ① 撮影者がカメラを手を持った状態で移動させるためカメラの移動量を画像から推測する必要がある。
- ② 文字のように特徴が細かく、しかも類似した特徴が多数存在する場合の対処が必要である。
- ③ 前記②の影響も有り、位置決めにかかる処理量が大きくなるため高速化を図る必要がある。

本稿では、上記①を満たす処理構成において、②、③を満足する位置決め手法を提案する。さらに、本手法を組み込んで試作したシステムでの評価結果を示す。

2. モザイク処理の流れ

図1に本モザイク処理の構成を示す。

動画を構成するフレーム中の連続する2フレーム間の移動量を検出し、その移動量に基づいて順次モザイク画像を生成するようにしている。なお、これから処理を行うものを現フレーム、前回の処理で既に行ったものを前フレームとする。

特徴領域抽出では、前フレームから位置決めに適した特徴的な領域を複数決定する。

1段目マッチングでは、決定された各々の領域についてブロックマッチング法により現フレーム上での対応位置を検出することで移動ベクトルを求める。

ベクトル同期検出では、各移動ベクトルの信頼性を評価し、信頼性の低い領域については特徴領域抽出まで戻り繰り返し処理を行う。

変換画像作成では、求めた移動ベクトルに基づいて、平面射影変換により前フレームからの移動分を反映した現フレームの変換画像を作成する。

2段目マッチングでは、前フレームまでにモザイク処理で作成された画像(旧モザイク画像)と現フレームの

変換画像とを前記と同様の手段によりマッチングを行い、位置決め微調整を行う。

最後に、画像貼り合わせにおいて、旧モザイク画像と現フレームとを貼り合わせ、新モザイク画像を生成する。

以上の処理を、動画を構成するフレーム全てについて実行し、広域な1枚の静止画を生成する。

以降では、本処理の特徴である特徴領域抽出とベクトル同期検出に関する位置決め手法を詳細に説明する。

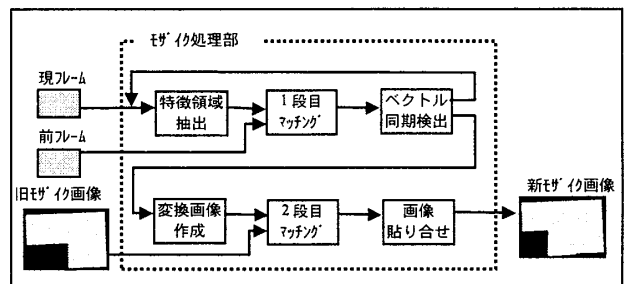


図1 モザイク処理の構成

3. 位置決め手法

3.1 特徴領域抽出

画像の位置決め処理は、処理量が大いいため、粗い画像を用いて大雑把な位置決めを行い、徐々に精度を上げていくようなピラミッド型の階層的な位置決めが行われることが多かった。しかし、単純な文字列だけで構成されたドキュメントでは、文字の特徴が細かいために画像を粗くするとその特徴が失われてしまう可能性が高い。そのため、前記階層的な位置決めでは対処できないという問題がある。

そこで、マッチング対象領域を複数の小領域に限定した位置決めを行うことにした。この小領域は、文字エッジなどの輝度変化の大きい特徴的な画素を多く含み、かつ、次段の変換画像作成における平面射影変換の誤差を少なくするために、フレーム上のなるべく外側に近いところを求めるのが望ましい。

そこで、図2に示すように、まず、前フレーム上で現フレームと重なると推定される領域を、前フレームを位置決めしたときの移動ベクトルから推定する。次に、前フレームの重なり推定領域内を $m \times n$ ドットのメッシュ状に区切り、各領域に対し縦、横、右斜め、左斜めの4方向について輝度差を算出し、その合計を面積で正規化した値を輝度ポイント L_p とした。また、各領域に対し重なり推定領域の4隅および4辺の中心の計8点(図3の●)からの最小距離を算出し距離ポイント D_p とした。次に、

(1) 式より各領域の特徴ポイント F_p を求める。この F_p は、上記8点それぞれに近づくにつれて特徴的な領域かを表している。各点ごとに F_p の高い順に候補をそれぞれ選出する。なお、 N は L_p と D_p のバランスをとる重みである。

$$F_p = L_p - N \times D_p \quad (1)$$

次に、8領域の候補についてそれぞれブロックマッチング法により前フレームからの移動ベクトルを求める。

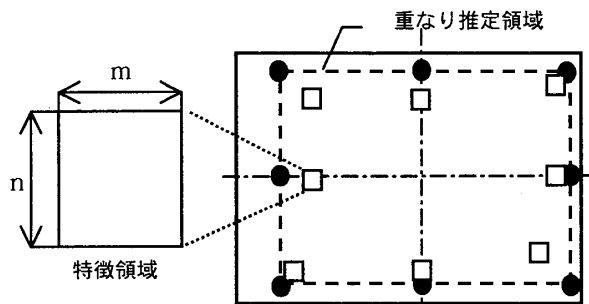


図2 特徴領域の検出

3.2 ベクトル同期検出

人がカメラを移動させながら撮影する場合、カメラの回転、傾き、上下動が原因で、各フレーム間において微小なロテートやズームなどの変動が発生する。このような変動が存在しても被写体がドキュメント等の平面状であれば基本的に平面射影変換により対応できる¹⁾。すなわち、得られた8つの移動先は平面に近似できる。

しかし、文字は類似した特徴を持つものが多く、ブロックマッチングの探索範囲にこのような特徴が複数存在すると、正しい位置決めができない可能性が極めて高い。本手法では8点という少数の領域の位置決めしか行っていないため、その1つでも位置決めを失敗すると、現フレームの貼り付け位置が大きくずれ、それ以降のフレームにも多大な影響を与える。

そこで、得られた8つの移動ベクトルを用いて(2)式に示す平面射影変換行列のa~hの8つのパラメータを最小2乗法により算出し、図3に示すように、i領域の座標(x_i, y_i)の変換先に対応する座標(x'_i, y'_i)を、行列を用いて求め、マッチングによって得られた移動ベクトルの移動先座標(X_i, Y_i)との距離を求める。この距離が大きいときは、i領域と他の領域との移動ベクトルの同期が取れていないことになる。そこで(3)式より、この距離が閾値T_Lを超える場合、i領域の位置決めは信頼性が低いと判断し、i領域付近の2番目の候補(F_pが2番目に大きい領域)から再度位置決めと信頼性の評価を行う。

このような処理を繰り返し行い、8点すべてが(3)式を満たさなかったとき、変換画像作成の処理へ移行する。

$$\begin{pmatrix} x' \\ y' \\ m \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (2)$$

$$\sqrt{(X_i - x'_i)^2 + (Y_i - y'_i)^2} > T_L \quad (3)$$

4. 実験・評価

前述した位置決め手法を用いた実験を行った。

撮影者はモバイルカメラを手を持って操作し、机上に設置されたドキュメントをなぞるようにカメラを移動させながら動画撮影した。また、撮影と同時に映像をPCに伝送し、リアルタイムにモザイク処理を行った。

実験におけるカメラスペックは、フレーム画像が172×144ドットの24ビットフルカラー、フレームレートは15fpsである。

実験のパラメータとして特徴領域サイズをm=n=10、(1)式の重みをN=5、(3)式の閾値をT_L=1と定めた。これは、経験的に求めた値である。

なお、今回は歪曲収差による画像の歪みを抑えるために、事前に、使用するカメラの歪み量を検出し、前処理としてフレームの補正を行っている。

実験結果を図4に示す。この例では、開始フレームから撮影を始め、折り返し点で撮影位置を下げつつ横方向に1往復したときのものである。同図に示すように、文字のみで構成された画像において良好なモザイク画像が得られた。また、モザイク処理に要する処理時間は、前処理も含めて、800MHzのPCで1フレームあたり約50ミリ秒であり、フレームレート15fpsで取得したすべてのフレームをリアルタイムに処理することができた。

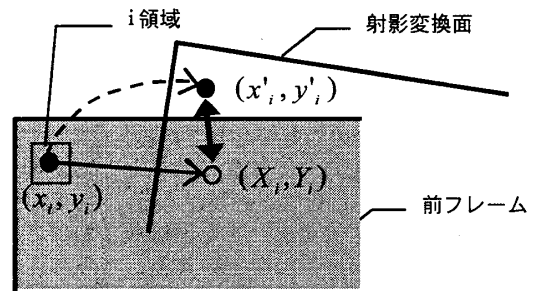


図3 信頼性の評価

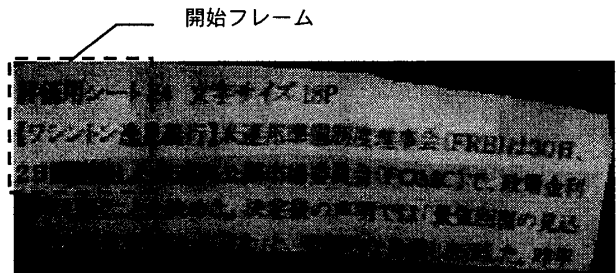


図4 モザイク処理結果

5. むすび

本稿では、ドキュメントのような文字を主体に構成される被写体からのビデオモザイクングにおいて、複数の少数特徴領域の抽出による高速化と、文字画像に対応するため、位置決め結果の同期を取ることにより信頼性の高い位置決め結果を得る手法を提案し、実験・評価によりその有効性を示した。

参考文献

1) R.Szeliski, "Video mosaics for virtual environments" IEEE CG&A, Vol16, No.2, pp.22-30, MARCH 1996.