

SD-7-4 多重解像での残差最小化に基づくパノラマ画像生成と視野角推定

Stitched Panorama using Multiresolution Error Minimization and FoV Estimation

間瀬 健二 西羅 光

Kenji Mase and Hikaru Nishira

ATR 知能映像通信研究所

ATR Media Integration & Communications Research Laboratories

1 はじめに

画像をグラフィックオブジェクトに貼り付けるテクスチャマッピング技術は、オブジェクトを詳細にモデリングしなくても、容易にリアリティを得られるという利点がある。この利点は実写を利用した仮想環境 [1] を構成する際にも活かすことができ、仮想世界のウォークスルーや環境マッピングにも用いることができる。例えば、固定点でカメラを回転させて撮影した画像系列をつないで生成したパノラマ画像を順路の各ポイントに用意するだけで、臨場感のある効果的な仮想環境を提供できる [2]。

このような仮想環境を構成して臨場感を伝えるときに撮影時のカメラパラメータ、特に視野角 (Field of View, FoV) の透過性を保持することは、画像の解像度などのリアリティ以上に、カメラマンが受けた臨場感を時間と場所を越えて伝達する際に重要と思われる。コンピュータビジョンにおけるカメラパラメータの自動抽出は古くからの課題であり受動的ナビゲーションのパラダイムの中で多く検討されてきたが、近年、画像接続という目的に適した image-based アプローチが提案されている [3, 4]。

本稿では画像接続を前提としたカメラパラメータの抽出法を用いて、接続されたパノラマ画像の視野角の推定を行う方法を提案する。多変数関数の残差最小化による透視変換不変量の推定法として、平行移動量の初期値を与える局所相関法と、残差の2次項を利用した収束の早い Levenberg-Marquardt 法 [4] を組み合わせた多重解像的手法でパノラマ画像を生成する手法を提案する。また、実験によりその効果を示す。

2 複数画像のパノラマ化

パノラマ画像の再構成には、位置合わせをする変換パラメータの推定、推定パラメータによる画像の投影面への変換、および変換画像の接続合成という3つのステップがある。本稿では、変換パラメータ推定と画像接続合成の処理について述べる。

2.1 変換パラメータの推定

$\mathbf{x} = (x, y)$ および $\mathbf{x}' = (x', y')$ を、それぞれ変換 M の前後の画素座標を表すものとする。画像変換式は1式となる。

$$\mathbf{x}' = M\mathbf{x} \quad (1)$$

変換パラメータ推定には残差二乗和 (Sum of Squared Differences, SSD) を用いる。

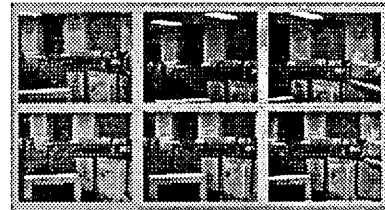


図 1: Input example pictures for mosaic.

2.1.1 平行移動パラメータの推定

最初に平行移動パラメータを局所相関の最小化によって推定する。この場合 SSD は次式の E を最小化する (u, v) のパラメータを与える。

$$E = \sum_i [I'(x_i + u, y_i + v) - I(x_i, y_i)]^2 \quad (2)$$

本手法では (u, v) の推定値を後段の初期値とするので、画像解像度に対して疎な計算点において推定し、頻度分布を求めその最頻値を与える値を次のステップで用いる。なお、本稿では対象世界は静止していることを前提としているが、前景移動物体および背景からなる世界を対象とするパノラマムービー [5] を生成する際には頻度分布の複数モードを利用することもできる。

2.1.2 Levenberg-Marquardt 法による変換パラメータの推定

\mathbf{x} から \mathbf{x}' への変換 M は一般に未知数 8 の非線形方程式で表されるが、ピンホールカメラの理想的回転を仮定すると未知数は 4 となり次式でその関係を表すことができる。

$$x' = \frac{m_0x + m_1y + m_2}{m_6x + m_7y + 1} = f \frac{x + \alpha \sin \theta y + f \alpha \cos \theta}{-\alpha \cos \theta x + \gamma y + f} \quad (3)$$

$$y' = \frac{m_3x + m_4y + m_5}{m_6x + m_7y + 1} = f \frac{-\alpha \sin \theta x + y - f \gamma}{-\alpha \cos \theta x + \gamma y + f} \quad (4)$$

ここで、 θ は光軸周りの傾き、 α は pan、 γ は tilt の角度、 f は焦点距離である。 α および γ を用いてカメラ回転を計算し、パノラマ画像の視野角が決定できる [5]。

Levenberg-Marquardt 法 (以下 LM 法と呼ぶ) はこのような多変数関数の最小化の手法の一つで、次のように最急降下法を利用している。すなわち、一般的多変数関数 $y = f(\mathbf{x})$ が $\mathbf{x} = \mathbf{x}_{min}$ (未知) で最小値をとるものとする。このとき $f(\mathbf{x})$ を $\mathbf{x} = \mathbf{x}_{min}$ のまわりでテーラー展開し 2 次の項までをとると、次のようになる。

$$f(\mathbf{x}) \simeq c - \mathbf{b} \cdot \mathbf{x} + \frac{1}{2} \mathbf{x}^T \cdot \mathbf{A} \cdot \mathbf{x} \quad (5)$$



図 2: An example of mosaic room.

$$b = -\nabla f|_{x_{min}} \quad (6)$$

$$A_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j} \Big|_{x_{min}} \quad (7)$$

最小値をとる x では $\nabla f = 0$ だから

$$x_{min} - x_i = A^{-1}[-\nabla f(x_i)] \quad (8)$$

となるので、ヘッセ行列 A と f の勾配 $[-\nabla f(x_i)]$ より x_{min} を求めることができる。

2.2 画像の接続

画像変換パラメータが得られたら、1つの画像を基準(平面)にして他の画像を変換して接続する。たとえ画像変換が正しく再現されたとしても撮影時の条件により各対応画素の輝度がずれることがある。また、画像変換は重量領域においてつじつまを合わせているので、重ならなかった周辺領域の歪みは反映しきれない。したがって適当な境界を定めて画像をモザイク化して張り合わせたのでは境界において不自然な不連続が生じる。

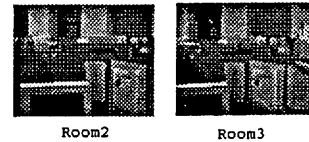
そこで、変換した画像間で輝度線形補間をして合成するわけであるが、周辺領域の歪んだ部分は上記の理由でできるだけ用いたくない。本文では各画像の中心を母点とするドロネー網を作り、個々のドロネー多角形の輝度補間はそれを形成する頂点(母点)の画像のみを用いて補間する方法を提案する。こうすることで各画像の中心付近を重みづけした合成が可能となる。

3 実験結果

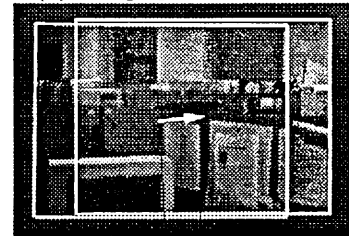
実験には手で支えたカメラをすこしずつ回転させながら撮った静止画像6枚(640x480, 白黒、図1)を用いた。6枚の画像を本手法を用いて貼り合わせたところ図2のような画像が得られた。端の部分は歪みを合成時に吸収できず、まだずれがみられる部分もあるが、全体的には滑らかに合成されている。また、図3は2枚の画像のパノラマ接続について全体の視野角推定をした実験結果である。

4 おわりに

変換パラメータが精度良く求まっている場合には、合成を施す重畳した中心部分を滑らかにつなぐことができた。た



(a) Image pair for mosaic



(b) Mosaic and computed FoV

図 3: Experimental result of computing FoV.

だし合成を施さない画像の端の部分のずれまでは修正することはできない。このずれはLM法の計算誤差と画像変換モデルと実際のレンズ系の透視変換のずれによるものである。画像の重畳部分の割合が大きくなればこの誤差は小さくでき、連続撮影したビデオ画像などからパノラマ画像を作成する際に有利に働くと思われる。

大型プロジェクタ等の普及で、臨場感のあふれる映像鑑賞などが可能となりつつあるが、視野角に関する透過性は無視されることが多い。より臨場感をあげるために画像のフレームにこだわらない新しい映像提示手法と映像表現手法の検討が今後必要である。

なお、著者のうち西羅(現所属: 東京大学工学部計数工学科)は学外実習生としてATR滞在中に本研究に携わった。

参考文献

- [1] K. Mase, R. Kadobayashi and R. Nakatsu: "Meta-museum: A supportive augmented reality environment for knowledge sharing", Intn'l Conf on Virtual Systems and Multimedia'96(1996).
- [2] S. E. Chen: "Quicktime VR - an image-based approach to virtual environment navigation", Proc. SIGGRAPH'95(Los Angeles, CA), pp. 29-38(1995).
- [3] A. Akutsu and Y. Tonomura: "Video tomography: An efficient method for camerawork extraction and motion analysis", in Proc. ACM Multimedia'94, pp. 349-356(1994).
- [4] R. Szeliski: "Video mosaics for virtual environments", IEEE Computer Graphics and Applications, 16, 2, pp. 22-30(1996).
- [5] K. Mase and H. Nishira: "Computing the field-of-view of a stitched panorama to create fov sensitive virtual environment", Proc. ICPR'96(1996).