

非拘束画像マッチングを用いた自動画像補間に基づく 自由視点画像生成法

山崎俊太郎[†] 池内 克史^{††} 品川 嘉久^{†††}

[†] 独立行政法人産業技術総合研究所, 東京都江東区青海 2-41-6

^{††} 東京大学, 東京都目黒区駒場 4-6-1

^{†††} 2021 Beckman Institute University of Illinois 405 N. Mathews Urbana, IL. 61801, USA

E-mail: †shun-yamazaki@aist.go.jp, ††ki@cvl.iis.u-tokyo.ac.jp, †††sinagawa@uiuc.edu

あらまし 本論文では, 自動画像モーフィングを用いた, 2次元画像間の対応点計算手法を提案する. 提案法では, 対象となる物体や風景, 及びカメラパラメータなどの情報が未知の場合にも, 視覚的に自然な画像間の対応を計算し, 仮想的な中間視点画像を合成することが可能である. まず, 入力画像に関する知識なしに画像間の自然な対応を得るために, 人間の初期視覚で用いられているものと同様の線形フィルタを用いて, 画像の特徴を抽出する. 次に, 画像特徴の類似度, 及び対応写像の滑らかさから構成される目的関数を最小化して対応点を計算する. その際, 最適化の計算が局所解に陥りやすいため, 画像の対応写像の滑らかさに関する幾何的な拘束を提案することにより手法の頑健性を向上させる. 実験では, 複数の対象物に対して仮想的な中間視点画像を生成し, 手法の有効性を実証する.
キーワード 非拘束画像マッチング, 自動画像補間, 自由視点画像, 回転偏微分ガウスフィルタ

Unconstrained view-interpolation based on automatic image-morphing

Shuntaro YAMAZAKI[†], Katsushi IKEUCHI^{††}, and Yoshihisa SHINAGAWA^{†††}

[†] The National Institute of Advanced Industrial Science and Technology, 2-41-6, Aomi, Koto-ku, Tokyo, Japan

^{††} The University of Tokyo, 4-6-1 Komaba Meguro-ku, Tokyo, Japan

^{†††} 2021 Beckman Institute University of Illinois 405 N. Mathews Urbana, IL. 61801, USA

E-mail: †shun-yamazaki@aist.go.jp, ††ki@cvl.iis.u-tokyo.ac.jp, †††sinagawa@uiuc.edu

Abstract This paper presents a method of view-interpolation based on automatic image-morphing. In our method, a dense and smooth mapping between two images is automatically determined without a priori knowledge of either the camera pose or objects. In order to find plausible correspondences, images are first convolved with a set of linear filters similar to those that are used in early vision. The vector fields consisting of filter responses are then matched with each other by minimizing the cost function which expresses the similarity of transformed images and mapping smoothness in a multiresolutional hierarchy. Since a point in one image can correspond to its counterpart far away, globally optimized mapping may result in strong local distortion. Such distortions are prevented by guaranteeing the convexity of mapping in nonlinear optimization. In this paper, certain experimental results are discussed for various data sets, including images of a rotating object, objects with occlusion and a translucent object, to demonstrate the usefulness of the proposed method.

Key words Unconstrained image matching, Automatic morphing, View-image synthesis, Rotated Gaussian derivative filters

1. はじめに

複数の画像間の対応点の計算 (対応点問題) は, コンピュータビジョンの主要な研究課題のひとつである. 画像間の対応

点が見られることによって, 形状の復元, 動きの検出, カメラの姿勢の推定など, 広い分野における様々な応用が考えられる. 既存の多くの研究では, 対象物の特徴 (色分布の線形性など), 画像間の関係 (カメラの外部パラメータが既知, 対象物の

変化が微小など), 計測カメラの特性 (カメラの内部パラメータが既知, 反応曲線が同一など) が既知であるという拘束を用いて対応点を計算している。従って, 性能の低いカメラで, キャリブレーションせずに計測して得られた画像から対応点を計算する事は困難である。本論文では, 2 枚の画像間の補間画像を生成することを目的とした, 非拘束の対応点計算法を提案する。

入力画像中の物体が未知の場合, 画像間の対応点計算を行うためには次の 2 つが問題となる。第一に, 画像間の対応を決定するための特徴量を定義する必要がある。特に, 画像の対応点間での輝度が同一であるという保証がない場合, “自然な” 対応を定義する必要がある。第二に, 対応点計算の際に, 局所解を避けるための工夫が必要である。本論文では, 自然な対応を得るために人間の初期視覚で用いられているものと同様のフィルタを用いて画像の特徴を抽出する。また局所解を避けるために, 対応写像の幾何的な拘束を考慮した非線形最適化手法を提案する。

2. 関連研究

対応点計算の手法は, 様々な対象物, 計測環境の組み合わせに対して数多く提案されてきた。この章では, 画像間の密な対応を計算する手法を簡単に紹介する。

a) Optical flow

画像間の視差が小さく, また対応点間で画像輝度が変化しない時, Optical flow を計算することによってカメラまたは対象物の移動を検出できる [1]。Optical flow の計算の際に用いられる, 基本となる拘束式は次の式で与えられる。

$$I_x u + I_y v + I_t = 0 \quad (1)$$

ここで画像輝度 I は, 2 次元画像座標 x, y , 及びフレーム間の時刻 t でパラメータ化されており, 各画素の動きベクトル $u, v = (x_t, y_t)$ を推定することが目的となる。式 (1) は画像間の視差が小さく, また輝度変化が線形近似できる場合にのみ成立する式であり, u, v が大きい場合には推定が困難である。また, 画像輝度 I が変化しないという拘束により光源の変化を伴う場合や, 画像がノイズを含む場合に推定が不正確となる。

b) 画像に関する知識を利用する方法

対象物や計測カメラに関する情報があらかじめ得られている場合, 対応点の計算を低い次元で行うことが可能である。例えば, ステレオ視による 3 次元復元の手法では, カメラの姿勢からエピポーラ拘束を導くことにより各対応点の探索範囲を 1 次元にすることが可能である [2]。また, 顔画像の対応付け [3] や, 医療画像の位置あわせ [6] 手法では, 各画像が類似の形状を持っており, 輝度分布に一定の規則性を仮定できることから, テンプレートの当てはめによって問題を簡単化することが可能である。本手法ではこのような事前知識なしに対応を計算する。

c) 非拘束画像マッチング法

画像の輝度に関する拘束のみを使って対応点を計算する手法として, 非拘束画像マッチング法が品川らによって提案されている [8]。この手法では Critical Point Filter (CPF) と呼ば

れる非線形フィルタを用いて画像特徴を抽出し, 多重解像度計算によって目的関数を最小化することで対応点を計算する。この方法は画像間の輝度分布が似ている場合には有効であるが, 実際の画像では必ずしもその仮定は満たされないため, 対応の精度が低い。また, 入力画像の視差が大きい場合には, 画像の一部分において対応点計算が局所解に陥り, それによって周辺に対応が大きく歪むという問題がある。また, CPF の特徴として画像輝度のピークに非常に強く反応するため, ノイズを含む画像では結果が大きく歪む。提案法は, 輝度の変化やノイズに対して頑健な特徴抽出法を提案し, 局所解に陥りにくい計算手法を提案することによって, 視点画像補間に適した非拘束画像マッチング法を提案する。

3. 対応点問題の定式化

\mathbb{D} を画像座標空間, \mathbb{V} を画像輝度空間, $I: \mathbb{D} \rightarrow \mathbb{V}$ を画像の輝度分布関数とする。このとき, 画像間の対応点問題は, 画像 I_{src} から画像 I_{dst} への写像 $m: \mathbb{D} \rightarrow \mathbb{D}$ のうち, $I_{src}(m(\cdot))$ と $I_{dst}(\cdot)$ との差 ε_D^2 が最小となるようなものを見つけること, として定式化できる。

多くの場合, 画像は離散的な画素の 2 次元集合として与えられるが, 画素を補間して再サンプルすることにより, \mathbb{D} をさらに高解像度で定義できる。また, 写像 m は全単射とする。即ち, m^{-1} は \mathbb{D} の各点で一意に定義できるものとする。

以降では, 議論を単純化するため, 2 次元のグレースケール画像の対応点計算の手法を説明する。提案法は, より高次の, 例えばカラー 2 次元画像, またはグレースケール 3 次元画像 (ボリウム) などへの適用も, 自明な拡張によって容易に実現可能である。カラー画像への拡張は 6. 節で述べる。

4. 目的関数の構成

4.1 特徴抽出

本研究では, 多くのコンピュータビジョンの手法の仮定と異なり, 画像中の輝度は必ずしも画像間で一定ではない状況を想定している。そのため, 画像輝度を直接, 特徴量とすることはできない。その上でさらに, 対応点は人間の視覚的に違和感のないものである事が望ましい。

このような考察から, 人間の初期視覚で用いられる特徴抽出フィルタと類似した, 回転偏微分ガウスフィルタによって特徴抽出を行う。このフィルタのカーネルは次の式で定義される。

$$K_{\sigma, n, \theta}(x, y) = G_{\sigma}^{(n)}(x \cos \theta - y \sin \theta) \times G_{\sigma}^{(0)}(x \cos \theta + y \sin \theta) (2)$$

ここで $G_{\sigma}^{(n)}$ は分散 σ^2 の 1 次元ガウス関数の n 階微分である。

Malik らはこのフィルタを組み合わせるフィルタバンクを構成して画像から特徴量を抽出する手法を提案している [7]。ただし複数のフィルタが線形従属性を持つ場合, フィルタの反応に冗長性が生じて計算効率が下がる。そこで, 多くの異なる θ, σ, n の組み合わせに対するカーネルに対し主成分分析を適用し, カーネル空間の次元を決定して無駄のないフィルタバンクを構成する。実験では, n 階の偏微分フィルタは $n+1$ 次元空間を構成するという結果 [5] に基づき, 図 1 に示すよ

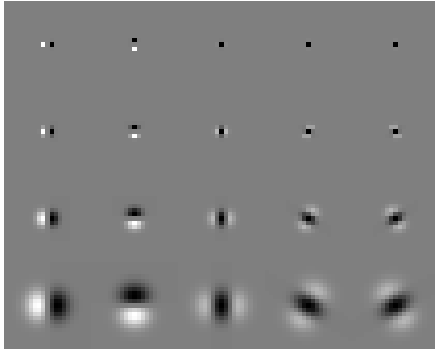


図1 特徴抽出に用いたフィルタバンク.

うな4種類の異なる分散を持つ、1階および2階フィルタから構成されるフィルタバンクを用いた.

画像の対応計算の際には、まず I_{src}, I_{dst} に対し、 N_f 種類 (実験では $N_f = 20$) の異なるカーネル $K_{\sigma, n, \theta}$ から構成されるフィルタバンク $F = \{F_i\} (i = 1, \dots, N_f)$ を適用する. このフィルタの反応は N_f 次元のベクトル画像であり、これを特徴ベクトル画像と呼ぶ. 以降の処理では、特徴ベクトル画像のみを用いて計算を行う.

4.2 目的関数の構成

I_{src} から I_{dst} への最適写像 m は、 $I_{src}(m(\cdot))$ と $I_{dst}(\cdot)$ の類似度を m の滑らかさを保ちながら最大化することにより計算する. そのために、特徴ベクトル画像間の類似度と m の微分から構成される目的関数を構成する.

まず、類似度に関する基準として、次の ε_D^2 を用いる.

$$\varepsilon_D^2 = \int_{p \in \mathbb{D}} \sum_{i=1}^{N_f} \rho_D \left((F_i * I_{src})(p) - (F_i * I_{dst})(m(p)) \right) \quad (3)$$

ここでオペレータ $*$ は畳み込み積分であり、 ρ_D は M 推定量 [4] である. 入力画像の視差が大きい時、5. 節で述べる反復計算における初期対応点の精度は低いことが予想される. M 推定量を用いることによって、精度の低い初期解から始めたときに局所解に陥る可能性を減らすことができる.

実験では、M 推定量として次の式で表される Lorentzian を用いた.

$$\rho(x) = \log \left(1 + \frac{1}{2} \left(\frac{x}{\sigma} \right)^2 \right) \quad (4)$$

ただし ρ_D は、 $F_i * I$ の各要素に対する $\sigma = \sigma_D$ の Lorentzian の和である.

次に画像の滑らかさの基準として、次の ε_S^2 を用いる.

$$\varepsilon_S^2 = \int_{p \in \mathbb{D}} \sum_{d=1}^{dim \mathbb{D}} \rho_S \left(\nabla m_d(p) \right) \quad (5)$$

ここで ρ_S は、 ∇m_d の各要素に対する $\sigma = \sigma_S$ の Lorentzian の和である. 入力画像に遮蔽がある場合、対応写像には不連続性が生じる. M 推定量を用いることにより写像の不連続性を扱うことが可能となる.

式 (3) と式 (5) を重み付け係数 α^2 を用いて線形結合することにより、最終的に最小化する目的関数 ε^2 は次のようになる.

$$\varepsilon^2 = \varepsilon_D^2 + \alpha^2 \varepsilon_S^2 \quad (6)$$

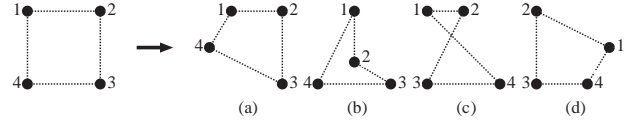


図2 隣接4点からなる四辺形の対応候補.

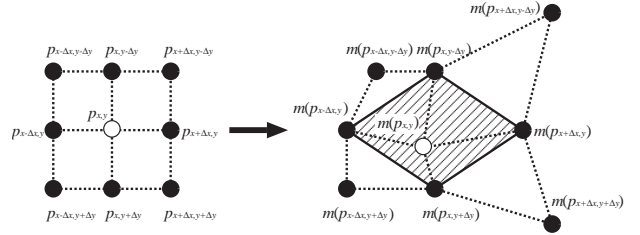


図3 凸写像拘束を満たすための対応点の探索範囲.

本手法で利用している特徴量は N_f 次元の情報を持つため、多くの場合 ε_D^2 だけを用いて対応点の候補を絞る事が可能である. ε_S^2 は特徴ベクトルから写像を決定できない場合、即ち画像輝度分布が一定の領域における写像を確定するために用いられる. 従って、一般に α^2 は非常に小さい値でよい. また、抽出された特徴ベクトルは対応点を決定する唯一の情報であるため、 σ_S は大きめに取る必要がある. σ_D は実験的に決定する必要があり、実装では画像サイズの5%となるように設定した.

5. 目的関数の最小化

入力画像間の差が大きい場合には、初期推定の精度が低くなるために、最適な α^2 を選択することは非常に困難である. この章では、写像が凸性を保つという拘束を考慮することによって、計算する課程で写像の大きな歪みを防ぎ、比較的小さい α^2 でも滑らかな拘束が得る方法を提案する.

I_{src} の隣接する4格子点なる正方形が、写像 m によって I_{dst} に向って変形するとき、 m の全単射性から、変形後の4格子点の位置関係は図2の通りである. 四辺形内部における全単射性を満たすためには、(c) と (d) を候補から除外する必要がある. 現実には、これらの写像は、例えば画像が遮蔽領域を持つ場合に存在しうるが、四辺形が非常に小さくつぶれる場合、大きく広がる場合を考慮することによって、(c) と (d) の場合を無視しても画像間の対応は矛盾なく決定できる. 次に、(b) の変形が存在すると、図中の頂点2を含む隣接四辺形における可能な写像が制限され、全体として写像が局所解に陥りやすい. そこで (b) も目的関数の最小化の際に候補から除外する. 以上より、推定の各段階において (a) の場合の写像のみを考えることにより、目的関数とは別に、最小化の過程において写像の滑らかさを保証することができる. この拘束を凸写像拘束と呼ぶ.

全体として凸写像拘束を満たす写像 m を決定するためには、局所的な写像の推定の際に凸写像拘束が満たされるようにすれば十分である. これは次のようにして実現できる. 図3のように、 I_{src} 中の点 $p_{x,y}$ が、 m によって I_{dst} に対応付けられる場合、 p の対応点の候補を $m(p_{x+\Delta x, y}), m(p_{x, y+\Delta y}), m(p_{x-\Delta x, y}), m(p_{x, y-\Delta y})$

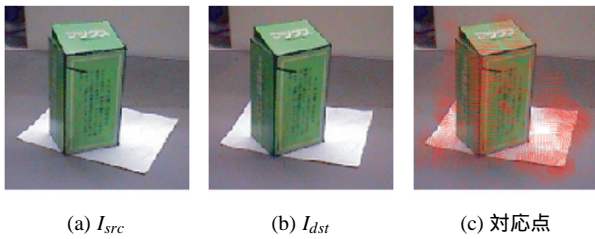


図4 非拘束画像マッチングの例

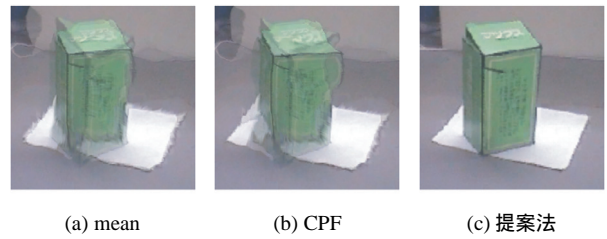


図5 画像輝度の変化に対する頑健性の比較

からなる四辺形の内側に限定すればよい。

6. 実装

凸写像拘束によって導入された写像の隣接対応点の幾何的な依存関係のため、画像中の各点の写像を同時に推定することは困難である。そこで、我々は品川らの手法 [8] を拡張し、反復法による逐次最適化を用いる。まず I_{src} 中の一点 p に対し、 I_{dst} 中で $m(p)$ として選択可能な全ての点に関して目的関数を計算し、最も関数値が小さい点を $m(p)$ とする。この操作を I_{dst} 中の全ての点に関して繰り返すことにより、画像全体に対する写像 m を推定する。また、 p の選択順序による写像の偏りを防ぐために、 p は乱数を用いて選択する。

さらに、局所解を防ぐために、目的関数 ε^2 の最小化を多重解像度計算で行う。対応点計算をする前に、特徴ベクトル画像をダウンサンプルして解像度を下げ、ピラミッド構造を構築しておく。写像の推定時には 2×2 サイズの画像から計算を開始し、各解像度で目的関数が最小化されたときの写像を、一つ上の解像度における推定の初期値とする。これを必要な解像度になるまで繰り返す。実装では、各解像度における推定には反復法を用い、最初の局所最適解をその解像度における最適解とした。

また画像がカラー情報を持つ場合には、4.1 節で述べた特徴抽出フィルタを各カラーチャンネルに対して適用して、特徴ベクトル画像を生成した。即ち、RGB 画像の場合、R,G,B の各チャンネルに対し N_f 次元のフィルタバンクを適用し、全体として $3N_f$ 次元の特徴ベクトル画像を生成した。

現在の実装では、 ε^2 の積分計算は、 \mathbb{D} を離散的にサンプリングして和を取ることで行っている。サンプリング幅は画像の画素サイズと一致する必要はなく、解像度を上げると精度が向上するが、計算時間が増す傾向にある。実験では、画素幅を 1.0 としたとき、サンプル幅を 0.1 から 0.5 程度に設定した。

7. 実験結果

7.1 非拘束画像マッチング

図4は、回転している物体の画像間における推定写像の結果である。画像解像度は 128×128 である。(a) と (b) が入力画像で、(c) が提案法で推定された写像である。

図5では画像の輝度に変化がある場合の推定の頑健性を比較した。安価なカメラで自動ホワイトバランスで計測を行う場合には、このような揺らぎを考慮する必要がある。実験では図4(a)と、輝度を10%明るくした(b)の間の対応点を推定した。平均値フィルタを使った推定法、品川らの手法 [8] による

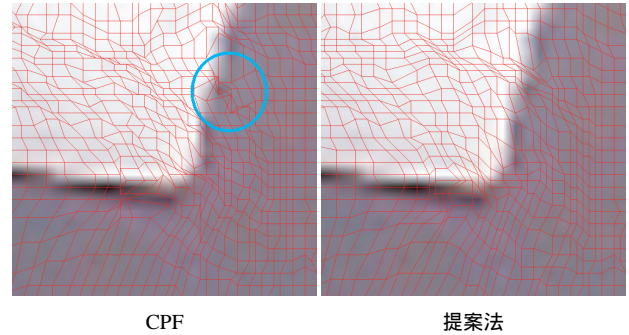


図6 推定写像の滑らかさの比較

推定法、提案法による推定法で対応点を計算し、trilinear 補間によって中間視点画像を生成して結果を比較した。輝度の不変性を仮定する既存手法では、わずかな輝度変化で写像が大きく歪むのに対し、提案法では中間視点画像で写実性が失われていないことが分かる。

図6では、写像の滑らかさに関して比較を行った。図4で用いた画像間の推定写像において、品川らの手法 [8] で推定した写像は部分的に裏返っているのに対し、提案法では滑らかさが維持されていることが分かる。

7.2 視点画像補間

推定された対応写像を基に、入力画像を trilinear 補間して中間画像を生成した結果を示す。

図7は2枚の画像を用いた視点画像補間の結果である。(a) と (b) が入力画像、(c) が画像モーフィングによって生成された中間視点画像である。入力画像の内部、外部カメラパラメータは未知である。対応点は 256×256 解像度で推定した。(d) は、(c) と同一の視点で実際に計測された画像であり、(c) の合成には用いていない。合成画像はほとんどの部分で、実際の計測画像と似ているが、遮蔽が存在する部分では対応する点が存在しないために、補間画像中に本来存在しない色が現れている。この問題を改善するためには、推定した対応写像中で、隣接対応点が近い場所では I_{src} を、離れている場合には I_{dst} を重み付けして補間画像を生成するなどの工夫が考えられる。本論文では対応点計算の手法の提案を主眼としており、補間手法に関しては紙面の都合上これ以上踏み込まないものとする。

図8に、透明物体に対する視点画像補間の例を示す。この物体は、透明で、テクスチャがほとんど存在せず、強い鏡面反射成分を持ち、また蛍光灯下で自動ホワイトバランスを使って安価なカメラで撮影した。このような画像に対して、特徴点抽出に基づくカメラ姿勢の推定を行うことは困難である。また、

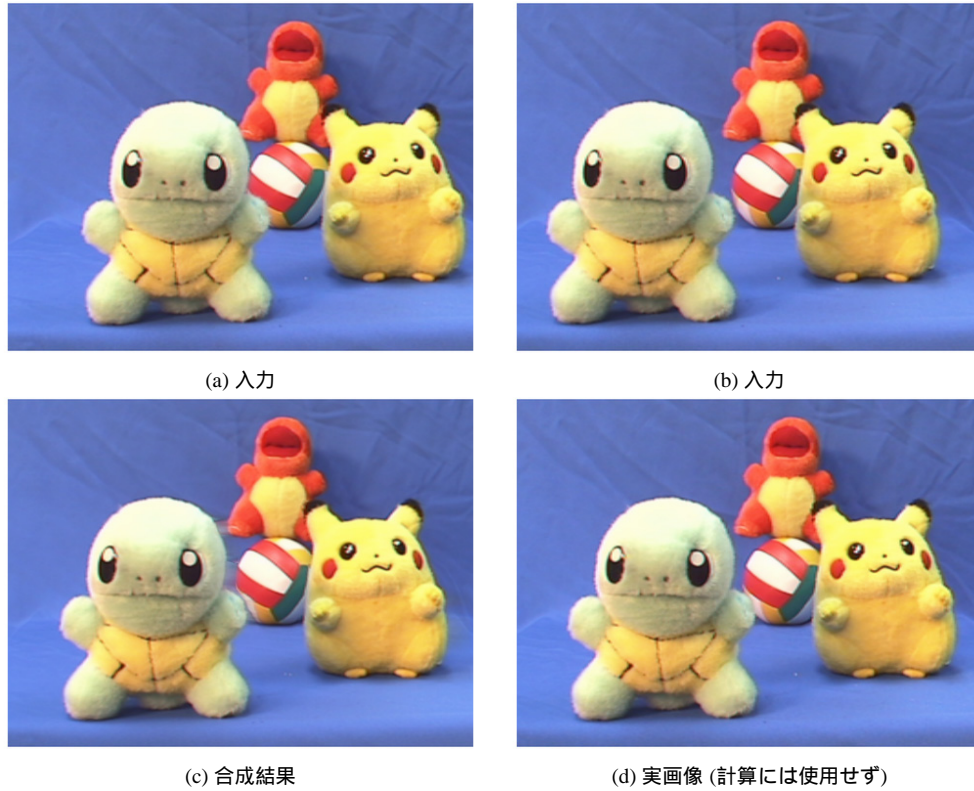


図7 視点画像補間の結果



図8 透明物体に対する視点画像補間の結果

カメラパラメータが未知であるため、ステレオ視による立体復元は現実的ではない。提案法では、画像の対応に基づく視点画像補間を行うため、このような画像に対しても写実性のある中間視点画像を生成することが可能である。ここで、推定される対応点は幾何的に正しい対応となっている保証はないが、合成される中間視点画像は人間の視覚にとって正しく見えることが期待できる。

7.3 性能評価

本論文で提案した手法を Pentium3 1GHz, 主記憶 768MB の PC で実装し、性能を比較した。表 1 は計算速度と対応点の輝度の自乗差である。非拘束画像マッチング法の一つである品川らの手法 [8] と比較した場合、特徴量の次元が高いこと、また推定に反復法を利用していることから計算時間は増加している。一方、得られた写像はより画像輝度差の小さい対応を得ることができていることから、より大域的な最適解に近い推

表 1 性能評価。計算時間の単位は秒、誤差は画像輝度差の二乗和 ($\sum_p |I_{src}(p) - I_{dst}(m(p))|^2$) である。

data (size)	[8]		proposed method	
	time	error	time	error
boxes (128 ²)	2.1	6.0×10^6	8.5	5.0×10^6
textures(128 ²)	2.0	3.8×10^8	10.7	1.6×10^8
faces (256 ²)	8.2	1.3×10^8	51.3	1.1×10^8

定ができていることが推定される。

8. まとめ

本論文では、完全に自動で、画像に関する知識無しに、画像間の密な写像を推定する非拘束画像マッチング手法を提案した。提案法では回転偏微分フィルタを用いた新しい特徴量を提案し、滑らかな写像を推定するための凸写像拘束を提案し

た. 既存手法と比べ, 推定写像の精度が向上していることを示し, 様々な物体に対して視点画像補間を行い手法の有効性を実証した.

文 献

- [1] Steven S. Beauchemin and John L. Barron. The computation of optical flow. *ACM Computing Survey*, 27(3):433–467, 1995.
- [2] Myron Z. Brown, Darius Burschka, and Gregory D. Hager. Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):993–1008, 2003.
- [3] Timothy F. Cootes, Gareth J. Edwards, and Christopher J. Taylor. Active appearance models. In *Proc. ECCV '98*, pages 484–498, 1998.
- [4] Frank R. Hampel, Elvezio M. Ronchetti, Peter J. Rousseeuw, and Werner A. Stahel. *Robust Statistics: The Approach Based on Influence Functions*. John Wiley, 1986.
- [5] David Jones and Jitendra Malik. A computational framework for determining stereo correspondence from a set of linear spatial filters. In *Proc. ECCV '92*, pages 395–410, 1992.
- [6] J. B. Antoine Maintz and Max A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.
- [7] Jitendra Malik, Serge Belongie, Jianbo Shi, and Thomas Leung. Textons, contours and regions: Cue integration in image segmentation. In *Proc. International Conference on Computer Vision '99*, pages 918–925, 1999.
- [8] Yoshihisa Shinagawa and Tosiyasu L. Kunii. Unconstrained automatic image matching using multiresolutional critical-point filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(9):994–1010, 1998.