

3カメラ間の射影幾何を用いた View Morphing

木村誠^{†,††} 斎藤英雄^{†††}

本発表では、3カメラによる撮影画像中の対応点に基づいた画像合成の一手法について述べる。筆者らは従来より、多カメラによる撮影画像とカメラ間の射影幾何を用いて、画像中の対応点に基づいた任意視点映像生成の研究を行ってきた。それらの研究では対応点探索を主眼としており、画像合成処理の原理は、画像中の対応点座標を重み付け平均する、対応点の内挿処理であった。そのような画像合成法は幾何学的正確さに欠け、入力画像によっては合成画像中に不自然な歪みなどが含まれる可能性があることが過去に明らかにされている。そこで、本発表では、2つの入力画像からの画像合成手法“View Morphing”を3画像へと拡張し、カメラ間の射影幾何と画像中の対応点から歪みの無い画像合成を行う。

View Morphing based on Projective Geometry Between Three Cameras

MAKOTO KIMURA^{†,††} and HIDEO SAITO^{†††}

In this paper, we propose a method to synthesize images of novel viewpoints based on projective geometry between three cameras and matching points in their images. We have proposed several methods for image synthesis from multiple input images based on projective geometry between cameras. Because our previous work mainly focused on searching matching points in images, image synthesis was practiced by so called interpolation, which is performed by weighted mean of matching points. It is known that image synthesis by interpolation may produce unnatural distortion of objects in the synthesized images. Therefore, we present an extension of “view morphing”, that is anti-distorted image synthesis using matching points and projective geometry between two cameras.

1. はじめに

コンピュータ・ビジョンの分野においてシーン内の物体の3次元形状を画像から再構成する問題は一般的であり、その中でも複数カメラからの入力画像を用いた手法はステレオ・ビジョンと呼ばれている。ステレオ・ビジョンによる3次元形状の再構成には、カメラ・キャリブレーションと呼ばれるカメラ・パラメータの同定が必須であるが、カメラ・キャリブレーションは複雑な作業であることが知られている。一方、複数の画像を入力として、実際にはカメラの存在しない場所からシーンを撮影したかのような画像を合成するシステムも従来より研究^{1),7),9),11)}されており、それらの手法はImage-Based-Rendering(IBR)と呼ばれている。IBRを実現するシステムの多くは何らかの

形で3次元形状の再構成を行い、再構成形状に基づいて画像合成を行うという方法を用いている。しかしながら、このように画像合成だけを目的としている場合には、画像中の対応点からもっともらしい画像を合成する^{5),7),11)}ことが可能である。つまり、IBRにおいて3次元形状再構成は必須ではなく、したがってカメラ・キャリブレーションも必須ではない。

筆者らは射影幾何^{6),10)}の概念に着目し、弱キャリブレーションと呼ばれる簡易なキャリブレーションからカメラ間のエピソード幾何のみを同定し、エピソード拘束と画像の輝度情報から対応点の探索を行う手法^{2),11)}を研究してきた。それらの手法では、画像中の対応点探索問題を非ユークリッドな3次元空間内における3次元形状再構成問題に置き換えることにより、オクルージョンの検出も行うことができた。したがって、それらの手法では画像間の対応点群が得られるのみならず、入力画像中でオクルージョンが生じている部分について、“どの点がどの点に隠されているか”という隠蔽関係をも得ることができ、そのような隠蔽関係を合成画像に反映することによって仮想視点の移動に伴う物体の隠蔽の変化が再現されていた。それらの画像合成処理はモーフィングで実現されており、具体的には、対応点の各画像中の2次元座標位置と画

[†] 独立行政法人 産業技術総合研究所 デジタルヒューマン研究ラボ
Digital Human Lab., National Institute of Advanced
Industrial Science and Technology

^{††} CREST, 科学技術振興事業団
CREST, Japan Science and Technology Corporation

^{†††} 慶應義塾大学 情報工学科
Department of Information and Computer Science,
Keio University

素の色情報とを重み付け平均する対応点の内挿処理にオクルージョンの再現を導入したものであった。

2 カメラを用いた画像合成において、前述のようなモーフィングによる合成画像中に不自然な歪みが発生する可能性が Seitz によって指摘⁴⁾ されている。その歪みの原因は、内挿による画像合成法がカメラ・モデルなどを全く考慮しておらず、2 次元の座標値を直接内挿している点にある。Seitz は、この問題を解決するために“view morphing”^{4),8)} と呼ばれる手法を提案している。view morphing では、2 カメラ間の射影幾何に基づいて入力画像の変形を行うことにより仮想的な平行カメラによる撮影画像を生成し、その変形画像間でモーフィングを適用した後に、モーフィング画像を仮想カメラ座標系へと再変形することにより合成画像を得る。このように、view morphing は内挿の前後に射影幾何を考慮した処理を適用することにより不自然な歪みを生じない画像合成を行っている。

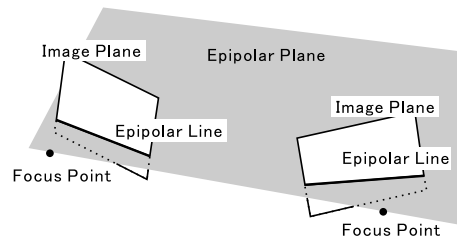
しかしながら、Seitz の提案した view morphing では、カメラの数が 2 つであるということに依存した処理を用いているため、仮想視点の位置は 2 カメラの中間の直線上に限定されてしまう。一方、筆者らの従来の研究²⁾ における画像生成では、仮想視点は 3 カメラで囲まれた三角形の内部に任意に設定可能であったが、幾何学的に正確な画像合成を行っているわけではなかった。そこで本研究では、Seitz 手法を拡張することにより、3 カメラから view morphing を行う手法を提案する。また、Seitz 手法では内挿後の再変形に手動処理が必要とされているが、本研究では Seitz とは異なるアプローチによって自動化を試みた。このように本研究は撮影画像中の対応点に基づいた画像合成を 3 カメラ間の射影幾何を考慮して行ったものである。

2. 射影幾何

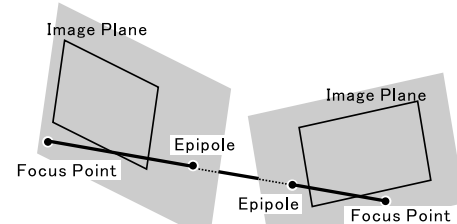
2.1 エピポーラ幾何と Fundamental Matrix

2 つのカメラで共通のシーンを撮影しているような場合を想定すると、カメラの焦点、画像平面、シーンの関係は図 1-(a) のようになる。この場合、2 カメラの焦点を結ぶ直線を想定すると、その直線を通るような任意の平面を定義することができる。このような平面はエピポーラ平面と呼ばれる。エピポーラ平面は 360° の自由度を持つが、エピポーラ平面と画像平面が平行にならない限りは、任意のエピポーラ平面とそれぞれの画像平面は交線を持つ。そのような交線はエピポーラ線と呼ばれる。1 つのエピポーラ平面は、それぞれの画像中に対応するエピポーラ線を決定する。また、図 1-(b) に示すように、別カメラの焦点が画像平面上に投影された点はエピポールと呼ばれる。理論上、全てのエピポーラ線がエピポールを通ることはエピポーラ線の定義より明らかである。

Fundamental Matrix (F 行列)^{6),10)} は、このような 2 カメラ間のエピポーラ幾何を記述する行列であり、あらかじめ F 行列が求まっていれば、2 画像間で任意のエピポーラ線を求めることが可能である。理論上、



(a) 2 カメラとエピポーラ平面およびエピポーラ線



(b) エピポール

図 1 エピポーラ幾何

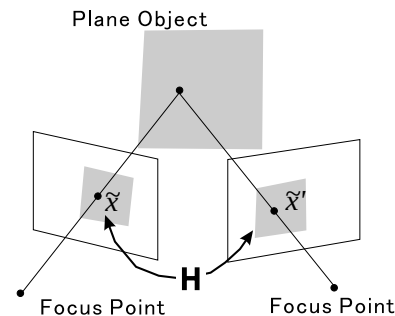


図 2 Homography

F 行列から画像平面中のエピポールの座標を求めることも可能であるが、カメラの内部および外部パラメータを個別に求めるようなことはできない。

2.2 Homography

図 2 のように、3 次元空間中のある同一の平面を 2 つのカメラで撮影した場合を想定し、その平面上の任意の点がそれぞれのカメラで撮影された画像座標を $\tilde{x}$, $\tilde{x}'$ とすると、 3×3 の行列 H とスカラー s を用いて、

$$s\tilde{x} = H\tilde{x}' \quad (1)$$

という関係が成立する。この行列 H は homography¹⁰⁾ と呼ばれ、平面同士の射影変換を記述する行列である。

homography による射影では、直線は直線へと変換されるが平行性や直角性は保存されないことが知られている。理論上、homography を求めるには両画像中の平面上で 4 点以上の対応が必要である。

3. Seitz 手法

複数画像中で正確な対応点が求まっている場合、それらの対応点の 2D 座標および画素値を重み付平均することにより、何らかの画像を合成することが可能である。このような画像合成法は 3D モデルに基づいた画像合成に比べて幾何学的な正確さに欠けるものの、入力カメラが適切に配置されている場合には、人間の目には違和感のない画像を合成することが可能である。特に、平行カメラを入力とした場合には、このような画像合成は正確な画像合成である。Seitz は、このような画像合成法を “image morphing” と呼び、その問題点を明らかにしている。第 5 章の実験で示すように、カメラの配置が image morphing に適切でない場合、合成画像中の物体の形状が歪んでしまう場合がある。これは、入力カメラが平行ではないことが大きく影響している。そこで、この問題を解決するために Seitz は “view morphing” を提案している。view morphing の処理は、以下に示すように、カメラ間の射影幾何を考慮した画像変形をモーフィングの前後に行うものである。

3.1 Pre-Warp

まず、2 つの入力画像のエピポーラ線を平行化する。この処理は画像平面全体を 2 カメラの焦点に平行な仮想平面へ射影することに等しい。言い替えると、仮想的な平行カメラを 2 つ想定し、実カメラの画像平面全体を仮想平行カメラの画像平面へ射影することに等しい。この処理は平面から平面への射影であるので、homography によって実現される。

3.2 仮想平面内でのモーフィング

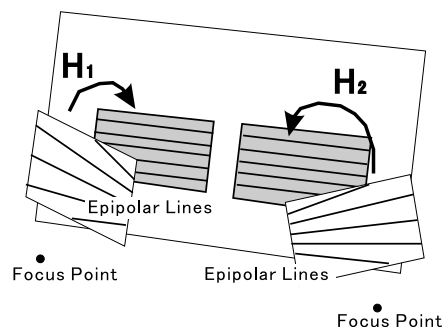
Pre-Warp された画像は、仮想的な平行カメラにより撮影された画像として扱うことができる。そこで、それらの画像間で対応点の内挿を行っても、光学的な歪みを生じることはない。そこで、Pre-Warp された画像内の全対応点について、2 次元座標および画素値を線形内挿することによって、内挿画像を生成する。

3.3 Post-Warp

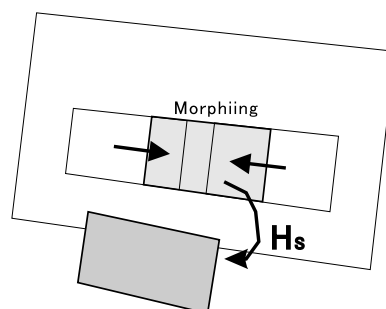
前節で合成された内挿画像を homography H_s を用いて仮想カメラの画像平面へ再射影する。この再射影によって、歪みの無い合成画像が得られる。Seitz 手法ではこの H_s を手動処理によって得ているが、その詳細については後に第 4.3 節で述べる。

4. 本 手 法

本研究では 3 画像を入力としており、各画像中の対応点および各カメラ同士の F 行列は予め求まっているものとする。本研究で 3 画像を用いている理由は、仮想視点位置の自由度にある。2 画像間で内挿処理に基づく画像合成を行う場合、仮想カメラの位置は 2 つの入力カメラの間に限定されてしまうが、3 画像を用いることによって、仮想カメラは 3 カメラで囲まれる三角形内に 2 次元の自由度を持つことができる。



(a) Pre-Warp



(b) 内挿画像生成と Post-Warp

図 3 Seitz 手法: 2 カメラの焦点に平行な平面へと画像全体を射影することにより、エピポーラ線は平行化される。そのようにして合成された仮想的な平行カメラの画像間で対応点の内挿処理(いわゆるモーフィング)を行い、適切な homography(図中の H_s) で仮想視点の画像平面へと再射影することによって、歪みの無い合成画像が得られる

前章で述べた Seitz 手法では、Pre-Warp 処理において 2 画像間のエピポーラ線の平行化によって、仮想平行カメラの画像を生成した。3 カメラの場合においても、仮想平行カメラへの射影を行うが、後述のように、その実際の処理は単純なエピポーラ線の平行化では実現できない。また、Seitz 手法では Post-Warp 処理に用いる homography は手動で解決していた。本研究では、Post-Warp についても自動化を試みた。

4.1 Pre-Warp

Seitz 手法では、2 画像を入力としていたため、平行カメラの画像平面への射影は、エピポーラ線の平行化処理によって実現されていた。3 画像に対して Pre-Warp を適用した場合においても、3 組の画像ペアそれぞれについては、エピポーラ線が平行化されることは明らかである。しかしながら、3 画像の Pre-Warp においては、別組のエピポーラ線がどのような角度で交わるかということと、また、ある組のエピポーラ線同士の間隔と、別組のエピポーラ線同士の間隔がどのような比率になるかについて配慮する必要がある。

まず、別組のエピポーラ線同士の成す角度について

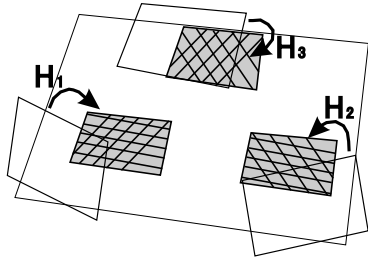


図 4 本手法における Pre-Warp は、3 カメラの焦点に平行な平面へ画像全体を射影することにより実現される

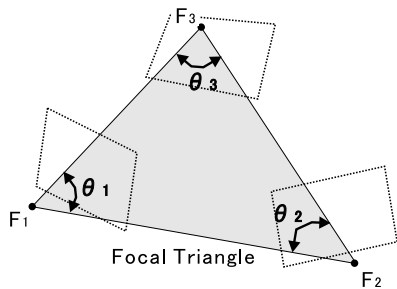


図 5 3 カメラの焦点によって定義される *Focal Triangle*. 図 4 の平行化されたエビポーラ線は、*Focal Triangle* の各辺に平行となる。

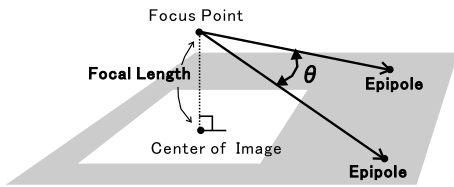


図 6 *Focal Triangle* の形状はカメラ内部パラメータとエビポールによって求めることができる。

考える。実際のカメラ配置およびエビポーラ幾何の原理から、図 4 および図 5 に示すように、各エビポーラ線はカメラ焦点を結ぶ直線に平行になる。したがって、異なるカメラ組のエビポーラ線の成す角度は実際のカメラの配置に依存し、各カメラ間のエビポーラ線により構成される三角形は、3 次元空間内で 3 つのカメラ焦点が成す三角形と相似になる。以降、このカメラ焦点が成す三角形を“*Focal Triangle*”と呼ぶ。

本研究では各カメラ間の F 行列は既知と想定しており、第 2.1 章で述べたようにエビポールは F 行列から求めることが可能である。もしも各カメラの内部パラメータが既知であれば、各カメラについて画像平面を基準とした焦点の位置がわかるので、各カメラ間のエビポールの座標と焦点の位置から、*Focal Triangle* の各頂点の角度を求めることが可能である。F 行列から 2 カメラの内部パラメータを求める方法は Kruppa 方程式^{3),10)}として知られている。Kruppa 方程式は、

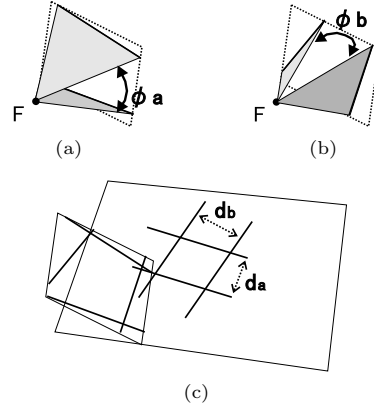


図 7 カメラ焦点とエビポーラ線からエビポーラ平面同士の角度 ϕ_a および ϕ_b が求まるので、それらエビポーラ線が平行化された後のエビポーラ線同士の間隔 d_a および d_b が決定される

“2 カメラの内部パラメータは同一であり、カメラ光軸は画像平面と画像の中心で直交している”という仮定を用いて、2 カメラの焦点距離を求めるものである。本研究では 3 つのカメラについて Kruppa 方程式と同様に“3 カメラの内部パラメータは同一であり、カメラ光軸は画像平面と画像の中心で直交している”という仮定を用い、3 カメラの焦点距離を非常に単純な方法で求める。

図 6 に示すように、*Focal Triangle* の各頂点の角度 $\theta_1, \theta_2, \theta_3$ は、画像平面上のエビポールの座標とカメラの焦点距離から求めることが可能である。また、*Focal Triangle* は三角形であるので、3 頂点の角度の合計は π にならなければならない。本論文では 3 カメラの焦点距離が共通の値であると仮定しているので、焦点距離を f とすると、上記の条件を満たすような f は一意に決定できる。求まった焦点距離から $\theta_1, \theta_2, \theta_3$ が求まるため、*Focal Triangle* の形状が求まり、Pre-Warp 後の画像中で各カメラ組のエビポーラ線が成す角度も決定される。

さらに、このようにして求まった焦点距離を用いて、同一の画像組間で 2 つのエビポーラ線が与えられた時、それらエビポーラ平面の成す角度を求めることができる。例えば、図 7-(a), (b) のように、Pre-Warp の基準となるエビポーラ線を画像内に任意に与えると、それらのエビポーラ線を与えるためのエビポーラ平面同士の成す角度 ϕ_a, ϕ_b が決定される。また、Pre-Warp による射影先の平面は *Focal Triangle* と平行であり、先の図 6 でも明らかなように、カメラ座標系における *Focal Triangle* の法線は焦点距離とエビポールから求まるので、図 7-(c) に示す Pre-Warp 後のエビポーラ線同士の距離 d_a と d_b の比は、 ϕ_a と ϕ_b から求めることが可能である。

このようにして求まった条件に基づき、homography を用いて画像全体を仮想平行カメラの画像平面へと射影することにより、3 画像への Pre-Warp の適用が可能になる。

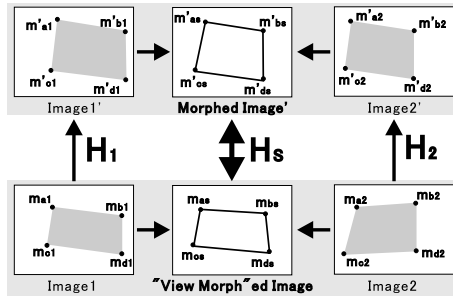


図 8 Seitz 手法における homography 決定方法

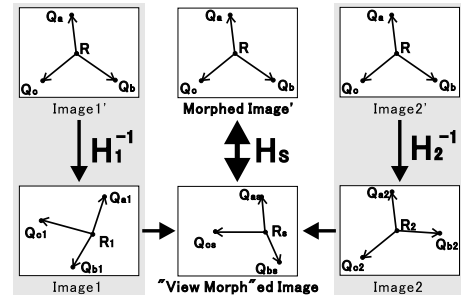


図 9 本手法における homography 決定方法

4.2 仮想平面内での内挿画像生成

前節の処理により、3つの変形画像が得られた。本研究では入力画像間の対応点は予め求まっているものとしている。したがって、変形画像同士の対応点もすでに求まっている。それらの対応点について、3画像中の座標値と画素値それぞれを重み付け平均することにより、内挿画像を合成する。

4.3 Post-Warp

前節で得られた内挿画像は、光学的には、あるカメラ焦点と画像平面による撮影画像を表している。しかし、その焦点と画像平面の位置関係は現実には存在し得ないようなカメラを示している。そこで、実在し得るようなカメラ・モデルによる撮影画像へと再変形する必要がある。この処理について Seitz 手法は手動で解決を行っているが、本研究では自動解決を試みる。

まず Seitz 手法の概要を図 3 に示す。Seitz 手法では、手動で入力画像中の対象物体内に平面領域を探し、その平面領域中の 4 点の対応関係 m_{p1} と m_{p2} ($p = a, b, c, d$) の対応関係を手動で与える。それら 4 対応点について、Pre-Warp および仮想平面内での内挿処理後の m'_{ps} の座標を調べる。次に、入力画像中の 4 対応点座標に線形内挿を適用し、仮想カメラにおける 4 点 m_{ps} の位置を決定する。そして、 m_{ps} と m'_{ps} ($p = a, b, c, d$) を H_s による射影前後の点とみなして Post-Warp に必要な homography を決定する。このように、Seitz 手法では、“実画像間での 4 点のモーフィング”と“仮想平面内での 4 点のモーフィング”から homography を内挿している。

本研究では、Seitz 手法とは異なるアプローチによって、Post-Warp の homography を決定する。第 4.1 節の Pre-Warp は、3つのカメラについて、実在の画像平面から別の平面への射影を homography で実現したものである。そして、その共通の平面と実在の各画像平面の関係は、個別の homography によって記述されている。我々は、それらの 3つの homography から適切な homography を内挿することが可能なのではないかと考えた。第 2.2 節で述べたように、homography は 4 つの対応点から求めることができる。本研究では、Pre-Warp で用いた各 homography について、その homography のによる射影前後の 4 対応点を求めておき、それらの対応関係を内挿することによ

り、Post-Warp に必要な homography を求める。

具体的には、まず Pre-Warp 後の画像中心に点 R を設定し、その周辺に 3 点 Q_p ($p=a,b,c$) を設定する。そして、それらの点が Pre-Warp で射影される前の原画像中の座標 R_x, Q_{px} を 3 入力画像 ($x=1,2,3$) に求める。次に各入力画像中の“ R_x の座標”、“ R_x から 3 点 Q_{px} へのベクトルの大きさ”、“ R_x から 3 点 Q_{px} へのベクトルの向き”を求め、それらを個別に内挿した後に、内挿パラメータを用いて仮想カメラ中の R_s, Q_{ps} を求める。このようにして、平行化内挿画像中の点 R, Q_p と仮想カメラ中の点 R_s, Q_{ps} の対応関係を合成し、homography を求める。

このような処理を行っても、実在のカメラで撮影し得る画像を合成される理論的な保証は無い。しかしながら、次章における実験で示すように、我々の経験上はもっともらしい画像合成を行うことができている。

5. 実験

以下に、実画像を用いた実験を示す。これらの実験は、書籍を撮影した 640×480 画素の実画像 3 枚と、それらを撮影した 3 カメラ間の F 行列を入力とし、“書籍表面が平面である”という条件に基づいて求めた対応点を用いて、対応点座標の単純内挿による合成を行ったものと、本手法を適用して合成を行ったものである。どちらの合成画像においても、仮想視点、3 カメラについて、カメラ間を結び直線上を移動している。仮想視点の位置は 3 カメラで構成される三角形の内部に指定可能であるが、仮想視点の移動位置の分かり易さを考慮し、このような視点移動を行った。

図 11 に示す単純内挿画像では、例えば (a) の左画像と右画像の中間で書籍の大きさが不自然に小さく合成されており、また、(c) の上画像と左画像の中間で書籍の形状が不自然に歪んで合成されている。一方、図 12 に示した本手法による合成画像では、(a) における視点移動中にも書籍の大きさに不自然な変化は無く、(c) における視点移動中にも書籍の形状に歪みは感じられない。このように、本手法における処理では直線性が維持されるため、書籍表面のテクスチャや輪郭などの直線性を保つことにより、見た目の上での不自然さが大幅に軽減されていることが明らかである。

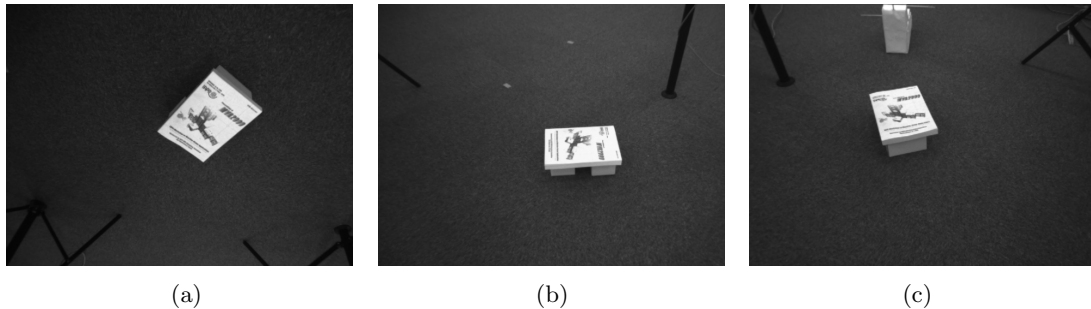


図 10 入力画像: (a) 上カメラ; (b) 右カメラ; (c) 左カメラによる撮影画像

6. 結 論

複数画像間のモーフィングによる画像合成は正確ではなく、入力画像によっては合成画像中に不自然な歪みなどが含まれる。そのような歪みは、合成画像中の直線性を失わせ、あるいは画像内の物体の大きさに不自然な変化を与えるので、合成画像の品質に直接影響する問題であると言える。本研究は、2 画像間で歪みの無いモーフィングを実現するための手法 view morphing を拡張し、3 画像を入力として view morphing を行ったものである。その際、Seitz による view morphing の手動処理部を自動で解くための一手法も提案した。

このような画像合成においても、幾何学的に 3D レンダリングと同等の正確さを得ることは困難である。しかし、実験結果に示したように、通常のモーフィングで合成画像中に発生し得る不自然な歪みは、view morphing によって大幅に改善することができている。この理由として、通常のモーフィングにおける歪みの多くは、合成画像中における直線性の喪失が原因であることが挙げられる。view morphing では、homography を用いることにより、仮想的な平行カメラによる撮影画像を生成している。理論上、homography による射影は直線性を失われない。また、平行カメラ同士でのモーフィングは、理論的に正確な画像合成であると言える。したがって、view morphing における画像合成は直線性を失わない画像合成であり、それ故に不自然な歪みを大幅に軽減することができたと考えられる。

本研究における入力画像を 3 枚であり、仮想視点は入力カメラ 3 つで囲まれた三角形の内部に設定可能である。2 画像間のモーフィングでは仮想視点をカメラ間直線上の 1 次元に限定されるが、本研究における視点位置は 2 次元の自由度を持つと言える。今後の課題として、より自由度の高い 3 次元的な仮想視点の位置設定を考えている。

参 考 文 献

- 1) Buehler, C., Matusik, W., Gortler, S., and McMillan, L., "Creating and Rendering Image-based Visual Hulls", MIT Laboratory for Computer Science Technical Report, MIT-LCS-TR-780, June 1999.
- 2) Kimura, M., and Saito, H., "3D Reconstruction Based on Epipolar Geometry", IEICE Transactions on Information and Systems, vol.E84-D, No.12, pp.1690-1697, Dec. 2001.
- 3) Lourakis, M., and Deriche, R., "Camera Self-Calibration Using the Kruppa Equations and the SVD of the Fundamental Matrix: The Case of Varying Intrinsic Parameters", Research Report 3911, INRIA Sophia-Antipolis, France, 2000.
- 4) Seitz, S.M. and Dyer, C.R., "View morphing", Proceedings of SIGGRAPH '96, pp. 21-30, 1996.
- 5) Wolberg, G., "Image Morphing: A Survey", Visual Computer, vol.14, pp.360-372, 1998.
- 6) Zhang, Z., "Determining the epipolar geometry and its uncertainty: a review", Research Report 2927, INRIA Sophia-Antipolis, France, July 1996.
- 7) 稲本奈穂, 斎藤英雄, "サッカーシーンにおける多視点映像からの中間映像生成", 電子情報通信学会技術研究報告, PRMU2000-151, pp.87-94, 2001.
- 8) 技術編 CG 標準テキストブック編集委員会, "技術編 CG 標準テキストブック", 財団法人画像情報教育振興協会, 1999.
- 9) 北原格, 石川寛享, 渡辺真生, 大田友一, "大規模空間の多視点映像を用いた運動視差の再現可能な自由視点映像の生成・提示方式", 画像電子学会論文誌, Vol.31, No.4, 2002.
- 10) 徐剛, 辻三朗, "3 次元ビジョン", ISBN 4-320-08522-1, 共立出版, 1998.
- 11) 矢口悟志, 木村誠, 斎藤英雄, 金出武雄, "未校正多視点カメラシステムを用いた任意視点画像生成", 情報処理学会コンピュータビジョンとイメージメディア研究会論文誌, Vol.42, No.SIG06-003, pp.9-21, 2001.

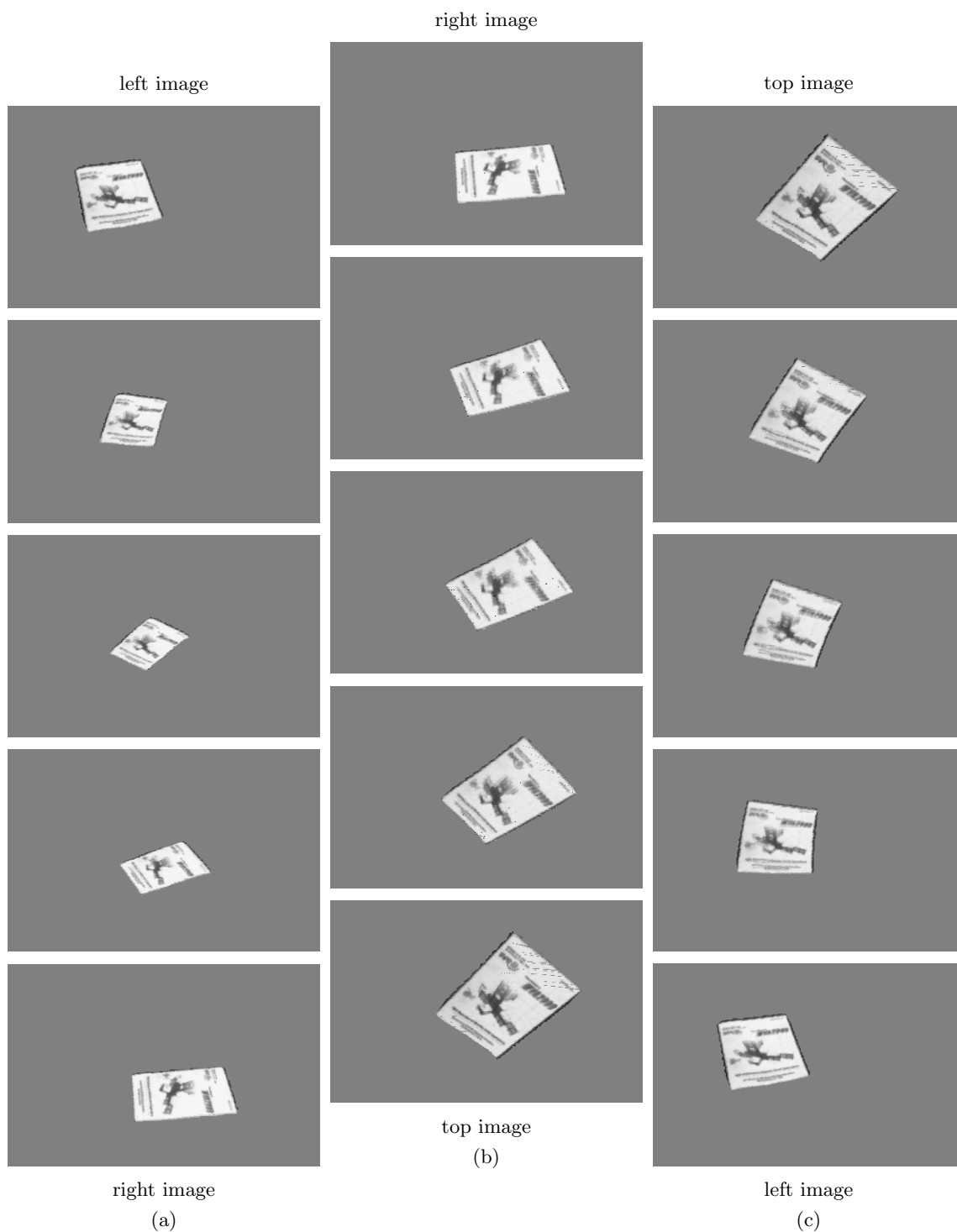


図 11 単純内挿による画像合成: 仮想視点の移動はそれぞれ (a) 左画像から右画像; (b) 右画像から上画像; (c) 上画像から左画像; (a) では仮想視点の位置が動くにつれて書籍の大きさが不自然に変化している。また, (c) では書籍の形状が不自然に歪んでいる。

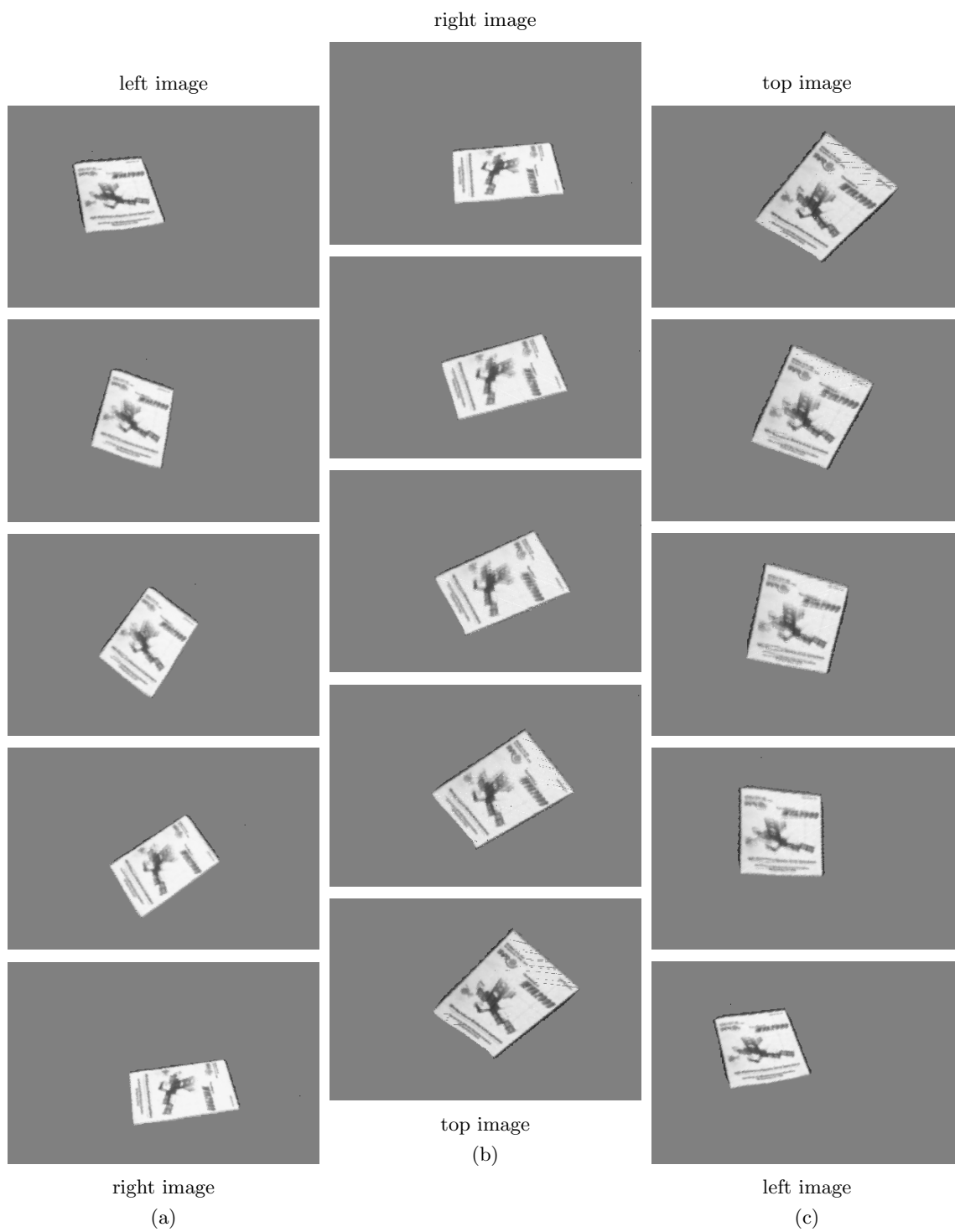


図 12 本手法による画像合成: 仮想視点の移動は, それぞれ (a) 左画像から右画像; (b) 右画像から上画像; (c) 上画像から左画像; 合成画像中において直線性が維持されているため, 書籍形状は平面らしく見える