

EPI 解析を利用したひずみのない複数ビデオカメラ画像の統合

三上 武志<sup>†</sup>      タンダ ウー<sup>†</sup>      小野晋太郎<sup>††</sup>      川崎 洋<sup>†</sup>  
 大沢 裕<sup>†</sup>      池内 克史<sup>†††</sup>

Distortion-Free Fusion of Multiple Video Camera Images Using EPI Analysis

Takeshi MIKAMI<sup>†</sup>, Thanda OO<sup>†</sup>, Shintaro ONO<sup>††</sup>, Hiroshi KAWASAKI<sup>†</sup>,  
 Yutaka OHSAWA<sup>†</sup>, and Katsushi IKEUCHI<sup>†††</sup>

あらまし エピポーラ平面画像 (Epipolar Plane Image, EPI) を用いて複数ビデオカメラの校正を行い、それらの画像をひずみなく統合する手法を提案する。一般にカメラの光学中心はカメラ内部に存在し、複数のカメラにおいてそれらを一点に集約することは物理的に困難である。このようなカメラ画像を統合する際は、対象の三次元形状情報などを利用しない限りひずみが生じる。本論文では、複数ビデオカメラを移動体に設置した条件のもとで、それらの光学中心を時空間内で一致させることでひずみなく画像統合を行う。光学中心の一致点は EPI マッチングにより自動的に算出する。この処理では各カメラ映像の時間的重複性を利用するため、撮影範囲の空間的重なりが少ない場合でも安定して対応関係が得られ、比較的自由にカメラを設置することが可能である。また、外部同期装置や測位装置などを用いずとも十分な品質を確保できるため、都市空間をはじめとした大規模な実空間のテクスチャ取得にも適している。

キーワード エピポーラ平面画像 (EPI), 時空間解析, 画像統合, 全方位画像

1. まえがき

都市空間に代表される大規模な対象のモデル化やそのための効率的なデータ取得方法の開発は、高度交通システム (ITS) やコンピュータグラフィックス (CG), コンピュータビジョン (CV) などの分野において重要な研究テーマとなってきた。

これまで大規模空間の効率的なデータ取得の対象としては、形状に関するものが主であった。例えば、航空写真から地図を自動生成する手法や衛星写真による数値標高モデル (DEM) 生成など、多くの手法が研究され、実用システムも多数開発されてきている。

一方で近年では、広域空間から取得したデータを

バーチャルリアリティ (VR) など、実際のシステムで利用しようとする試みも盛んになっており、景観シミュレーションや広域空間のデジタルアーカイブなど多くのアプリケーションが提案されている。これらのシステムを構築する際に重要となるのは、豊かな現実感の再現性である。計算機上で高い現実感を実現するためには形状情報のみに基づいたレンダリングでは不十分であり、テクスチャ画像の利用が不可欠である。このような理由から、広域空間のテクスチャ画像の取得に関する研究が盛んに行われるようになってきた。

広域空間の画像を効率良く計測する手法として、すべての方向を撮影可能なカメラを用いる手法が提案されている。代表的なものとして1台のカメラと曲面鏡を組み合わせたものがあるが、このようなシステムでは全方位のシーンを一画像として撮影するため、得られる画像の解像度が低くなる。高解像度の全方位画像を取得する方法としては、放射状に配置した複数カメラの画像を統合する手法 [1], [2] や、超多眼カメラを用いた手法 [3] が提案されている。しかし、これらの撮影系では個々のカメラの光学中心を一致させることが物理的に困難なことが多く、統合した画像にひずみが生じるほか、カメラ間の同期にも特別な機構が必要と

<sup>†</sup> 埼玉大学工学部, さいたま市  
 Faculty of Engineering, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama-shi, 338-8570 Japan  
<sup>††</sup> 東京大学情報理工学系研究科, 東京都  
 Graduate School of Information Science and Technology, The University of Tokyo, Institute of Industrial Science, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 Japan  
<sup>†††</sup> 東京大学情報学環, 東京都  
 Graduate School of Interdisciplinary Information Studies, The University of Tokyo, Institute of Industrial Science, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 Japan

なる。

本論文では、複数のビデオカメラを移動体に設置した条件のもとでそれらの光学中心を時空間内で一致させ、ひずみのない画像統合を行う。カメラのキャリブレーションには、エピポーラ平面画像 (EPI) 解析を利用する。カメラを設置する移動体はおおむね等速運動が可能であれば特に限定しないが、本論文では自動車の上部に搭載した場合を主に想定し、都市などの広域空間のテクスチャ画像を効率良く獲得することを目指す。

本論文の構成は以下のとおりである。2. では本手法の特徴を関連研究と併せて述べ、3. では本手法のアルゴリズムの詳細を説明する。4. ではひずみに関する誤差評価とカメラの配置に関する考察を行う。5. で実際に屋外空間において実験を行った結果を示し、6. で本論文をまとめる。

## 2. 提案手法の特徴と関連研究

### 2.1 全方位撮影系

代表的な全方位撮影系として、1台のカメラと曲面鏡を組み合わせたものがある。これはカメラの前に置いた曲面鏡により周囲の光線をレンズに集約することで全方位を撮影できるようにしたものであり、曲面として放物面 [4] や双曲面 [5] ~ [7]、円錐面 [8] などを用いるものが提案されている。このシステムでは1度の撮影で水平方向  $360^\circ$  すべての方向の画像を撮影できるため、複数のカメラを用いた場合とは異なり、時間的・空間的な整合性を保ちながら連続画像データを取得できる。反面、特殊な形状である曲面上に映った画像を透視投影画像などの一般的な形式に変換する追加処理が必要となるほか、方向により解像度が著しく異なり、画像全体の解像度も比較的低くなる問題点がある。

複数台のカメラを用いたシステムとして、天頂・底部を含む全周囲のカラー画像と距離画像をリアルタイムで取得できる全方向ステレオシステム (Stereo Omnidirectional System, SOS) [9], [10] が開発されている。このシステムは60個の CCD を3個1組とした20個のステレオカメラユニットとして正二十面体上に配置することで全方位にわたって一様な精度で計測することを可能としている。また、これを更に小型軽量化し移動体に搭載する miniSOS [11] も開発されている。しかし、これらのシステムは多数の CCD を使用するため高価であり、カメラの設置方法も特殊

である。

本研究では、一般的な民生用ビデオカメラ複数台から構成される撮影系により、大規模空間のテクスチャを簡易に取得する手法を提案する。

### 2.2 EPI とカメラキャリブレーション

複数のカメラ画像を統合するには、各カメラの外部パラメータが必要である。動画像から外部パラメータを推定したり、形状復元を行う研究はこれまでも多数行われてきた。EPI 解析 [12] は、カメラの動きが等速直線運動の場合に安定して撮影対象の三次元形状を推定可能な手法である。これは、カメラの時系列画像から生成される EPI と呼ばれる画像では、そのカメラが等速直線運動を行っている場合に (a) 特徴点の軌跡が直線エッジとなって現れ、(b) その傾きが特徴点の奥行きを表す、という性質を利用している。

従来、EPI 解析は主に対象の三次元情報を得るために利用されてきたが、我々は、同時に移動する複数のカメラから得られる EPI において、上記 (a) の性質に着目した。すなわち、同じ特徴点を同じ運動条件のもとで撮影したカメラでは、EPI 上において同一の軌跡が現れ、したがって EPI 同士でマッチングをとることができる。そのマッチング位置からは、カメラ間の外部パラメータ及びフレームずれに関する関係式を得ることができる。

また、EPI は空間と時間を二軸にとる時空間画像であるため、EPI 同士のマッチングは、空間軸方向の重なりが少ない場合でも時間軸方向の重なりが十分にあれば安定して働く。したがって、本手法では各カメラの撮影範囲に大きな空間的オーバーラップは必要なく、統合後の画像に隙間が生じない程度で十分である。このため、撮影領域に対する必要なカメラの台数が少なく済み、効率良く広い領域を撮影することが可能である。

## 3. ひずみのない画像統合と複数ビデオカメラのキャリブレーションの原理

### 3.1 時空間における光学中心の一致

画像統合とは、複数枚の画像を接合して1枚の画像にまとめることを意味する。一般に、カメラによって撮影された透視投影画像同士を統合する場合は、カメラの光学中心が一致している必要がある。例えば図1のようなカメラ配置により2本の円柱を撮影した場合、カメラ A, B の画像はひずみなく統合されるが、A, C の画像は一方の円柱に合わせて画像統合を行う

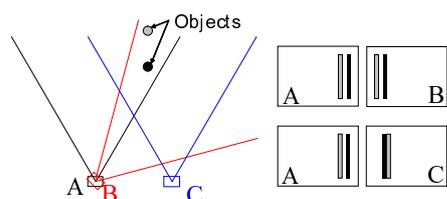


図 1 光学中心と画像統合におけるひずみ  
Fig. 1 Optical centers and distortions in mosaicing.

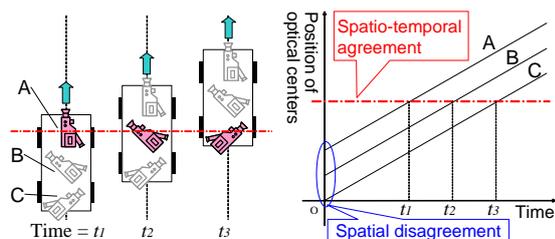


図 2 時空間における光学中心の一致  
Fig. 2 Spatio-temporal agreement of optical centers.

と他方の円柱が 2 個所に出現し、ひずみとなる。このような場合は、対象の三次元形状情報などを利用して補正を行う必要がある。通常、カメラの光学中心はカメラ内部に存在するため、複数のカメラにおいてそれらを一致させることは物理的に困難である。

この問題に対し、本研究では、複数のビデオカメラが移動体（車両）に設置された条件において、カメラの時間的な位置変化を考慮することで光学中心の一致を実現する。図 2 のようにカメラ A, B, C を配置し、カメラ A が時刻  $t_1$  に図の破線地点にあるものとする。簡単のため車両が等速直線運動をしているとすると、その後、時刻  $t_2$  にはカメラ B が、時刻  $t_3$  にはカメラ C が同じ地点を通過する。各カメラは車両上では異なる個所に設置され、光学中心は空間的には一致していない。しかし、時間方向の推移を同時に考慮することにより、異なった時刻において光学中心を一致させることができる。本論文ではこれを時空間における光学中心の一致と呼ぶこととする。各カメラを放射状に向かせて設置すれば、走行経路上の各地点において各方向の撮影画像が得られ、これらの画像からはひずみのない統合画像を合成することが可能である。各カメラは図 3(a), (b) のように進行方向に平行な直線上に並べて配置する。

以下ではまず、

- 条件  $C_1$  各光学中心は進行方向直線上に並んでいる  
条件  $C_2$  車両は直線運動をする

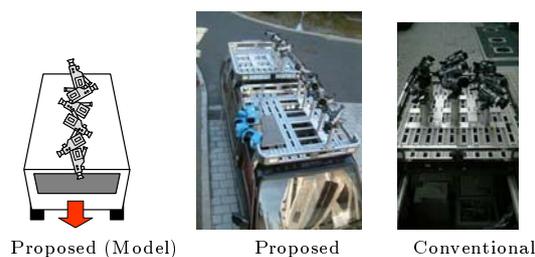


図 3 複数カメラの配置  
Fig. 3 Configuration of multiple cameras.

条件  $C_3$  車両は等速運動をする

という条件のもとで本手法の概論を述べ、後に実環境において運用する一般的な場合について述べる。カメラ及び対象物体の位置関係と画像統合におけるひずみの大きさについては 4. で評価を行う。

### 3.2 複数ビデオカメラの時空間キャリブレーション

#### 3.2.1 一般的定義

通常、カメラのキャリブレーションとは、カメラの位置・姿勢、及び画像面への射影に関するパラメータを求めることであるが、複数のカメラで動画像を撮影し、それらの動画像から画像統合を行う場合においては以下のように整理できる。

- 空間的キャリブレーション
  - 内部パラメータ（焦点距離、画像中心など）
  - 外部パラメータ（回転行列、並進ベクトル）
- 時間的キャリブレーション
  - 時間パラメータ（カメラ間の時間的対応関係）

空間的キャリブレーションは従来の定義と同様である。時間的キャリブレーションとは、各カメラ間における時間的な対応関係、すなわちカメラ 1 の時刻  $\tau$  における画像と、カメラ  $n$  の時刻  $\sigma$  における画像が対応するとき、その関係  $\sigma = \tau + \delta_n(\tau)$  を求めることを意味する<sup>(注1)</sup>。したがって、 $N$  台のビデオカメラのパラメータは、内部行列を  $A$ 、世界座標をカメラ座標に変換する回転行列、並進ベクトルを  ${}^w R, {}^w t$  として以下ようになる。

$$\begin{matrix}
 A_1, & {}^w R_1(\tau), & {}^w t_1(\tau), & \delta_1(\tau) \equiv 0, \\
 A_2, & {}^w R_2(\tau), & {}^w t_2(\tau), & \delta_2(\tau), \\
 \vdots & \vdots & \vdots & \vdots \\
 A_N, & {}^w R_N(\tau), & {}^w t_N(\tau), & \delta_N(\tau)
 \end{matrix} \quad (1)$$

(注1): この時刻は必ずしもグローバル時刻ではなく、カメラごとのローカル時刻でもよい。この場合はカメラ間の「時差」も  $\delta_n$  に含める。

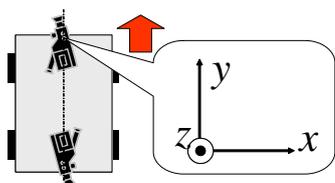


図4 基準座標系  
Fig. 4 Basic coordinate.

本論文では、これらのパラメータを複数カメラの時空間パラメータと呼び、これらを求めることをカメラの時空間キャリブレーションと呼ぶこととする。

### 3.2.2 本研究における時空間パラメータ

本研究においては図4のように第1のカメラの光学中心を原点とし、車両進行方向を $y$ 軸とする基準直座標系 $\Sigma$ をとり、この座標系(厳密には後述の $\Sigma'$ )における統合画像を車両の走行経路に沿った各地点で合成する。したがって、統合画像を合成するためには式(1)すべてのパラメータが必要なわけではない。

まず、回転行列は $\Sigma$ に対する行列 $R_n$ を求めれば必要十分である。この行列は時刻に依存しない。

次に並進ベクトルであるが、条件 $C_1, C_2$ のもとでは各カメラは配置間隔にかかわらず適当な時間後に第1のカメラと同じ位置を通過し、その時点での各撮影画像が統合画像を構成する。ゆえに、並進ベクトルは時間パラメータ $\delta_n(\tau)$ 内に集約されることになる。更に、条件 $C_3$ のもとでは $\delta_n(\tau)$ は定数 $\delta_n$ となる。

### 3.2.3 処理の流れ

以下では時空間パラメータ $A_n, R_n, \delta_n(\tau)$  ( $n = 1, 2, \dots, N$ )の推定処理の概略を述べる。内部パラメータ $A_n$ は、各カメラごとに植芝らのカメラキャリブレーション手法[13]を適用して推定する。回転パラメータ $R_n$ は、周囲の建造物形状から消失点を求めることで推定する。詳細は3.3に示す。これらの処理は、各カメラごとに画像列から適当な画像を選択し、1度だけ行えば十分である。

次に、回転パラメータを利用してすべての画像列に平面投影処理を施し、その画像列からカメラごとにEPIを作成する。時間パラメータ $\delta_n(\tau)$ はこれらのEPIのマッチングにより求まる。詳細は3.4に示す。

### 3.3 空間パラメータ(回転行列)の推定

カメラ姿勢の推定にはキャリブレーションパターンを利用するのが一般的である。小スケールの移動体の場合はボード上に印刷したものをいれればよいが、カ

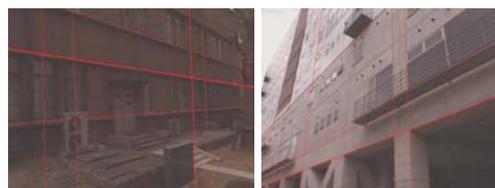


図5 平行直線群による消失点の導出  
Fig. 5 Detection of vanishing points using parallel lines.

メラを自動車上に設置する場合は巨大なボードが必要になる。そこで本研究ではH. Maitreらによる消失点を用いた導出手法[14]を踏襲し、屋外の建造物などのオブジェクトを利用する。これらのオブジェクトの輪郭は水平または垂直な平行直線群で構成されていることが多く、比較的容易に検出可能である。特に、都市空間においてはこのようなオブジェクトが多数存在する傾向が強く、回転行列を高い精度で検出可能である。以下ではこのキャリブレーション方法を具体的に説明する。なお、カメラは基準座標系上で固定されているため、この処理は全体を通してカメラごとに1度だけ行えば十分である。

1) 各画像列から建物の輪郭や窓枠などの水平・垂直な平行直線群が多く映っているシーンを選ぶ。以下では、これらの直線群が構成する座標系 $\Sigma'$ に対するカメラの姿勢が求まる。水平エッジと車の進行方向が平行である必要性は場合による。平行であれば $\Sigma$ と $\Sigma'$ が姿勢に関して一致し、完成後の統合画像と $\Sigma$ の関係<sup>(注2)</sup>が明らかになる。その必要がなければ平行でなくてもよい。統合自体はいずれの場合でも可能である。

2) Cannyオペレータによりエッジを検出し、ハフ変換により直線を抽出する。ノイズなどの影響により画像上での一線分に対するハフ空間上の投票は必ずしも一つのセルに集中しないため、ハフ空間上でk-means法によるクラスタリングを行い、その代表点に対応する直線を抽出結果とする。抽出された直線成分は垂直方向のものと水平方向のものに分類する。図5に直線抽出の例を示す。

3) 消失点を求める。抽出された水平・垂直方向の直線群はそれぞれ一点で交わり、その交点が消失点と呼ばれる。実際の画像ではノイズなどの影響により交点は唯一に定まらないため、すべての直線の組に対して交点の座標を求め、それらのうちすべての直線との

(注2): 画像中で車両正面方向に対応する座標など。

重み付き距離の総和が最小となる点を消失点とする．重み関数は式 (2) の Welsch ロバスト推定関数で定義し，遠く離れた直線による影響を軽減している．ここで  $d$  は交点と直線との距離， $c$  は定数である．

$$p(d) = \frac{c^2}{2} \left[ 1 - \exp \left\{ - \left( \frac{d}{c} \right)^2 \right\} \right] \quad (2)$$

4) 消失点から回転行列は以下のように計算される．まず，ベクトル  $r_1$  と  $r_2$  を次のように定義する．

$$r_1 = \begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix}, \quad r_2 = \begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix} \quad (3)$$

$(u_1, v_1), (u_2, v_2)$  は正規化画像座標における水平，垂直方向の消失点の位置である．したがって第 3 成分に焦点距離 1 を付加した  $r_1, r_2$  は，カメラ座標系において「光学中心から射影面上の消失点に向かうベクトル」を表す．これらのベクトルは  $\Sigma'$  の座標軸と平行になっている．したがって，水平，垂直方向の軸を  $\Sigma'(0, 1, 0), \Sigma'(0, 0, 1)$  ととれば，回転行列  $R$  に関して

$$R \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \hat{r}_1, \quad R \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \hat{r}_2 \quad (4)$$

が成り立つ．ここで  $\hat{\cdot}$  は正規化を表す．したがって，回転行列  $R$  は以下ようになる．

$$R = \left( \hat{r}_1 \times \hat{r}_2, \hat{r}_1, \hat{r}_2 \right) \quad (5)$$

### 3.4 時間パラメータの推定

#### 3.4.1 平面投影と EPI の生成

時間パラメータの推定には EPI を利用する．EPI とは，図 6 のように画像列をフレームごとに並べ，エピポーラ線に沿ってこれを切断した断面に現れる時空間画像である．

提案手法では，各カメラ画像列を直接に切断するのではなく，まず光学中心を同じ点に置き，前節で求めた  $R$  を利用して仮定の共通面上に投影した画像を求め，その画像から EPI を生成する．このような操作により，図 7 のように「ある二つのカメラが同じ位置にあるときに共有している光線」が EPI 上で同じパターンを描くようになる．ゆえに，逆にそのような EPI 上のパターンに関してマッチングをとることは，両カメラ間の時間パラメータを求めることと等価にな

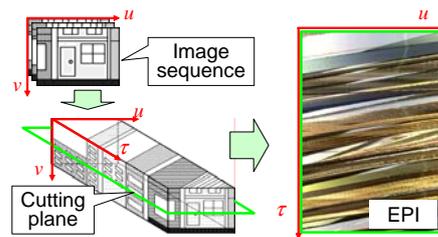


図 6 EPI と座標系の定義  
Fig. 6 Definition of EPI and coordinates system.

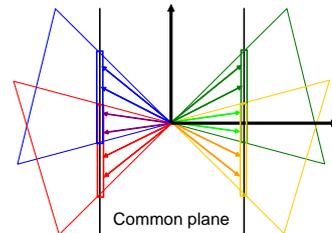
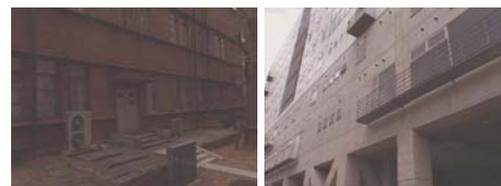


図 7 共通面上への投影  
Fig. 7 Projection to common plane.



(a) Original image



(b) Example of rectified image

図 8 平面投影の例  
Fig. 8 Examples of rectification.

る．なお，EPI 上に描かれるパターンは，オブジェクトのデプス値によって異なる傾きを描く性質を持っている．そのため，異なるデプス値をもつオブジェクトが多く撮影されているほど，マッチング処理は安定する傾向にあると考えられる．

我々の実験では投影する共通面は  $\Sigma$  上の平面  $x = \alpha$  とし，進行方向に対して右側を向いたカメラでは  $\alpha > 0$ ，左側を向いたカメラでは  $\alpha < 0$  とした．図 8 にその投影例を示す．

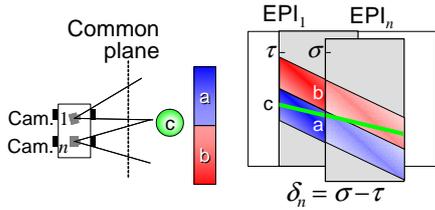


図 9 EPI マッチングによる時間パラメータの推定  
Fig. 9 Estimation of time parameter using EPI-EPI matching.

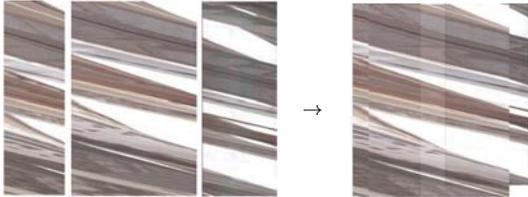


図 10 EPI マッチングの例  
Fig. 10 An example of EPI-matching result.

### 3. 4. 2 EPI 同士のマッチング

EPI 同士のマッチングを行い、時間パラメータを求める。なお、本手法では、マッチングのエネルギー関数に線形相関係数（付録参照）を利用しているため、カメラの個体差による色のばらつきはマッチングにほとんど影響しない。そのため、マッチングの際の色空間にはグレースケール空間を用いている。

図 9 のように、片方の EPI をずらしながら重なり部分の相関係数が最大となる点を二次元探索する。カメラ 1 の EPI とカメラ  $n$  の EPI が時刻  $\tau, \sigma$  でマッチするとき、時間パラメータは  $\delta_n(\tau) = \sigma - \tau$  として求められる。カメラ 1 と重なりのない EPI の場合は他のマッチング結果を通して計算する。

EPI の二次元マッチングは  $u$  軸方向の重なりが非常に少ない場合でも  $\tau$  軸方向の重なりが存在することにより安定して動作する。すなわちカメラの画角（撮影範囲）に大きな空間的オーバーラップをもたせておく必要がなく、カメラの設置は比較的自由に、厳密な調整を考慮することなく行うことができる。また、撮影領域に対して必要なカメラの台数が少なく済み、効率良く広い領域を撮影することが可能である。更に、カラーバランスのずれに対してもロバストであるという利点もある。図 10 に 3 台のカメラで撮影した画像を平面投影し得られた EPI において実際に二次元マッチングを行った結果を示す。複数の EPI が滑らかに結合していることが分かる。

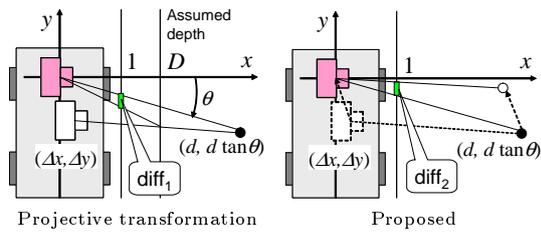


図 11 ひずみ量の解析  
Fig. 11 Analysis of distortion amount.

### 3. 5 画像統合

時空間パラメータが求まると各画像を統合することができる。各カメラの光学中心が  $\Sigma$  の原点に、画像平面上に時刻  $\tau + \delta_n(\tau)$  の画像が映っているものとして、それらを適当な共通面上に投影する。本実験ではこの共通面は  $\Sigma$  の原点を中心とした円筒、または  $x$  軸に垂直な平面としている。

### 4. ひずみ誤差の評価と自由運動への展開

本章では時空間キャリブレーションの有効性を示すため、ひずみ誤差の評価を行う。また、それらの結果を用いて 3. において仮定した  $C_1 \sim C_3$  の条件について考察し、移動体が自由運動を行った場合の適用性について論じる。

#### 4. 1 ひずみ量の解析

ひずみ量の解析は以下の二つについて行う。

(1) 従来のモザイクング（デプスを仮定した射影変換）による画像統合を行った場合のひずみ量

(2) 本手法による画像統合を行ったときに、光学中心が厳密には一致しないことにより生じるひずみ量

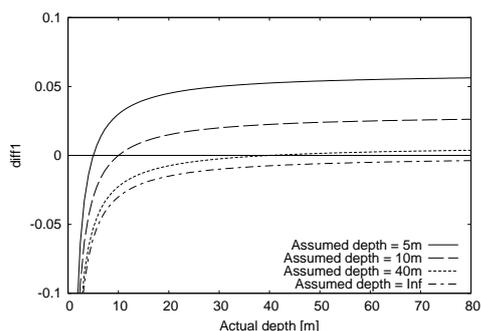
図 11 のように第一のカメラが原点に、他のカメラが  $(\Delta x, \Delta y)$  にあるとすると、それぞれの正規化画像座標上での位置ずれは図中の diff に相当し、以下のように表される。なお、diff<sub>1</sub> は従来の手法による位置ずれに diff<sub>2</sub> は本手法における位置ずれに相当する。

$$\text{diff}_1 = f(\Delta x, \Delta y) \left( \frac{d}{D} - 1 \right) \quad (6)$$

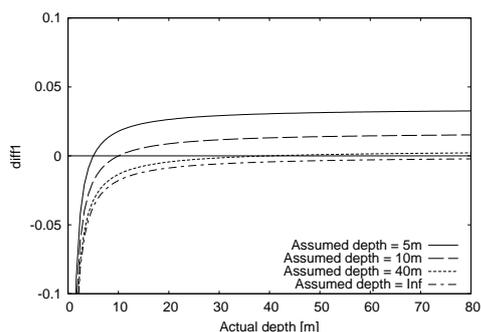
$$\text{diff}_2 = f(\Delta x, \Delta y) \quad (7)$$

$$f(\Delta x, \Delta y) = \frac{\Delta y - \Delta x \tan \theta}{d - \Delta x} \quad (8)$$

ここで  $d, \theta$  は注目点の位置を表すパラメータ、 $D$  は射影変換において仮定するデプスである。両者は式の上では係数だけの違いであるが、一般に  $d$  はシーン中で様々に変化する値をとり、diff<sub>1</sub> よりも diff<sub>2</sub> の方が



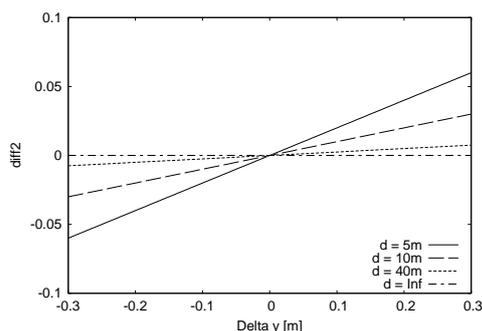
(a)  $(\Delta x, \Delta y) = (0, 0.3)$



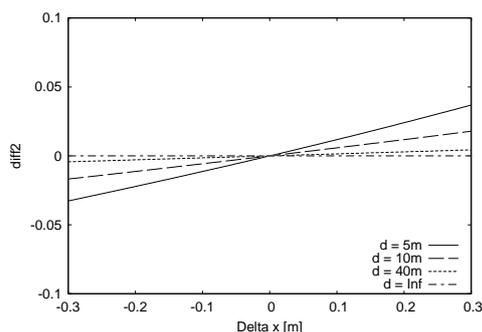
(b)  $(\Delta x, \Delta y) = (0.3, 0)$

図 12 ひずみ量の評価結果 (1)

Fig. 12 Evaluation result of distortion amount (1).



(a)  $\Delta x = 0$



(b)  $\Delta y = 0$

図 13 ひずみ量の評価結果 (2)

Fig. 13 Evaluation result of distortion amount (2).

$\Delta x, \Delta y$  の典型値は小さいことに注意されたい。

まず、カメラを従来手法である図 3 (c) の配置とし、ひずみ量を式 (6) により評価する。このような射影変換においてひずみ誤差を発生させるのは、仮定するデプス値  $D$  と実際のデプス値  $d$  の相違である。図 12 は、 $D$  及び  $d$  に対するひずみ量を計算したものである。カメラの配置間隔  $(\Delta x, \Delta y)$  は、一般的な民生用ビデオカメラの大きさを考慮して (a) で  $(0, 0.3\text{m})$ 、(b) で  $(0.3\text{m}, 0)$  とし、また  $\theta = -30^\circ$  としている。

次に、我々の提案手法におけるひずみ量を式 (7) により評価する。本手法においてひずみ誤差を発生させるのは、仮定  $C_1 \sim C_3$  が完全には成立しないことによる  $\Delta x, \Delta y$  の存在である。 $\Delta x = \Delta y = 0$  の場合はひずみ誤差は生じない。図 13 は  $\Delta x, \Delta y$  に対するひずみ量を計算したものである。なお、図 12、図 13 におけるひずみ量は、35mm のレンズを使用して水平解像度 640 ピクセルで撮影した写真のピクセル値に換算すると、 $\text{diff} = 0.01$  のときで約 6 ピクセル、 $\text{diff} = 0.05$  のときで約 30 ピクセルである。

これらの評価量は以下のことを示している。

- 射影変換による画像統合は、対象とするシーンがほぼ一定のデプスのみによって構成され、かつ、デプスの仮定値を適切に与えた場合に有効に機能する。
- 本手法による画像統合は、対象とするシーンの形状 (デプスの分布状況) に関係なく適用でき、また、光学中心のずれが小さければいくらかでも誤差を小さくすることができる。

#### 4.2 仮定条件の考察と自由運動への拡張

本節では、提案手法において仮定条件  $C_1 \sim C_3$  を緩和した場合のひずみ量について考察し、移動体が自由な運動をした場合の展開について論じる。ここでは移動体は自動車とする。

##### 4.2.1 直線配置条件 ( $C_1$ )

直線でない配置による水平方向のずれは図 11(b) の  $\Delta x$  で表される。垂直方向のずれは図示していないが、計算上は  $\Delta y$  と同じである。カメラの大きさと内部の光学中心の位置を考えると、 $\Delta x, \Delta y$  はたかだか 10cm 程度と見積られる。図 13 によれば、この影響による統合画像のひずみは近くの物体でも  $\text{diff}_2 = 0.02$  程度である。

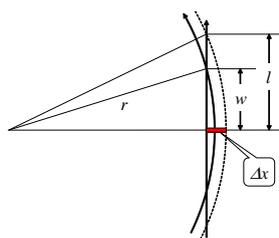


図 14 直線でない運動による水平方向のずれ  
Fig. 14 Horizontal difference caused by non-straight motion.

#### 4.2.2 直線運動条件 (C<sub>2</sub>)

直線でない運動による水平方向のずれは図 11(b) の  $\Delta x$  で表される。車が半径  $r$  の円運動をしたとすると、この大きさは図 14 により以下の式で表される。

$$\Delta x = \sqrt{r^2 + l^2 - w^2} - \sqrt{r^2 - w^2} \quad (9)$$

ここで  $2l$  はカメラを並べる長さ、 $2w$  は車のホイールベースである。  $2l = 2w = 3\text{m}$  とすると、 $\Delta x$  は一般国道において急カーブとされる  $r = 100\text{m}$  のときで 0.01 程度である。したがって、通常の道路を運転している状況では、この影響による統合画像のひずみはほとんど無視できるとしても問題ないと考えられる。

#### 4.2.3 等速運動条件 (C<sub>3</sub>)

等速でない運動に関しては、速度の変化を矩形近似、すなわち短区間ごとに等速運動を行っているのみとみなして対応することを考える。このとき実際の運動との間に生じるずれは図 11(b) の  $\Delta y$  に相当し、以下の式により見積もることができる。

$$\Delta y = \frac{1}{2} a \left( \frac{L}{v_0} \right)^2 \quad (10)$$

ここで  $a, v_0$  は車の加速度及び初速度、 $L$  は画像統合に使用するカメラの端から端までの距離である。  $L = 3\text{m}$  とすると、ずれ量は  $a = 5\text{km/h/s}, v_0 = 30\text{km/h}$  のときで  $\Delta y = 0.09\text{m}$  となり、図 13 によれば、このずれは近くの物体でも  $\text{diff}_2 = 0.02$  程度に抑えられる。一般に人が車を運転し、メータ読みにより等速を保つ場合でも生じる速度の変動は最大でも  $5\text{ km/h/s}$  程度と考えられる。したがって本条件もその程度の等速条件に緩和しても品質を保つことができる。

### 5. 実験

本手法の有効性を確認するために、従来のカメラ配置と提案手法のカメラ配置において実際に画像を撮影

表 1 カメラの仕様

Table 1 Specification of the camera.

Model No.	SONY VX-2000
Shutter speed	1/250 sec
F value	F6.8
Frame rate	29.97 frame/sec
Image size	720 × 480 pixel

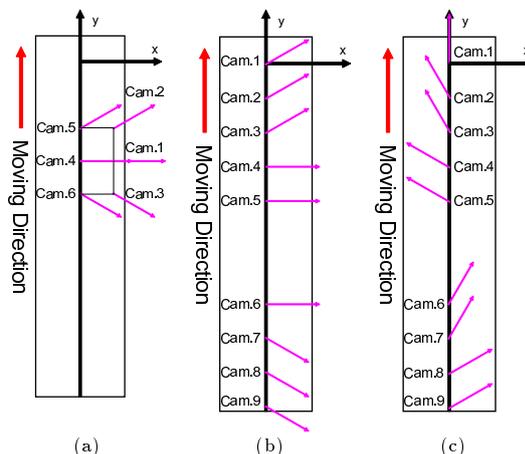


図 15 カメラのおおよその配置図

Fig. 15 Arrangement layout of cameras.

- (a) Conventional (6 cameras, right side)
- (b) Proposed (9 cameras, right side)
- (c) Proposed (9 cameras, multi-directional)

し、統合実験を行った。また、本手法を用いることで、前方 180° をカバーするようなパノラマ画像を生成した。更に、本手法を利用して生成したパノラマ画像を利用して、仮想自由視点画像の生成実験を行った。

#### 5.1 実験 (I)：複数カメラ画像の統合

本実験に使用したカメラの仕様を表 1 に記す。

まず、図 15、表 2、表 3 各 (b) の撮影条件のもとで道路を走行しながら撮影を行った。この画像列を提案手法によって統合した例を図 16 に示す。前方の電柱や後方の建物にも、対象のデプスによらず大きなひずみは発生していない。これは本手法によって光学中心を一致させた効果である。

これに対し、図 17 は、図 15、表 2、表 3 各 (a) の撮影条件のもとで撮影を行い、従来の射影変換によって画像統合を行った例である。この図では電柱が他の建物に対して近い位置にあるため、顕著なひずみが発生している。

このように、デプス値が大きく異なる物体 (図 16、図 17 における電柱と建物) が撮影されている画像を

表 2 カメラのおおよその配置条件  
Table 2 Arrangement condition of cameras.

- (a) Conventional (6 cameras, right side)
- (b) Proposed (9 cameras, right side)
- (c) Proposed (9 cameras, multi-directional)

	Camera No.	Position		Direction	
		$x(\text{cm})$	$y(\text{cm})$	Pan	Tilt
(a)	1	30	-90	0°	middle
	2	30	-60	30°	middle
	3	30	-120	-30°	middle
	4	0	-90	0°	high
	5	0	-60	30°	high
	6	0	-120	-30°	high
(b)	1	0	0	30°	low
	2	0	-30	30°	middle
	3	0	-60	30°	high
	4	0	-90	0°	low
	5	0	-120	0°	middle
	6	0	-213	0°	high
	7	0	-243	-30°	low
	8	0	-273	-30°	middle
	9	0	-309	-30°	high
(c)	1	0	0	90°	middle
	2	0	-30	120°	middle
	3	0	-60	120°	high
	4	0	-90	150°	middle
	5	0	-120	150°	high
	6	0	-213	60°	middle
	7	0	-243	60°	high
	8	0	-273	30°	middle
	9	0	-309	30°	high

表 3 撮影場所と走行速度  
Table 3 Capturing area and traveling speed.

- (a) Conventional (6 cameras, right side)
- (b) Proposed (9 cameras, right side)
- (c) Proposed (9 cameras, multi-directional)

	Area	Approx. speed
(a)	Urban area, Local street	25km/h
(b)	Urban area, Local street	15km/h
(c)	Urban area, Expressway	60km/h

対象として従来手法と本手法を適用したとき、光学中心を一致させている本手法では、いずれもひずみが生じていないが、それが行われていない従来手法では、同様のパラメータを用いたとき、片方にはひずみが生じている。

次に、提案手法により前方およそ 180° の範囲をカバーするようにカメラを配置して撮影を行った。撮影条件は図 15、表 2、表 3 各 (c) のとおりである。この画像列を仮想円筒上に投影してパノラマ画像列を生成した例が図 18 である。このように、本手法によれば特殊な装置を用いることなく簡易に光学中心の一致したパノラマ画像を生成でき、カメラの台数を増やせば全方位画像を生成することも可能である。カメラの画



図 16 本手法により光学中心を時空間的に一致させたカメラ画像の統合結果

Fig.16 Synthesized result of camera images with their optical centers agreed by our method.

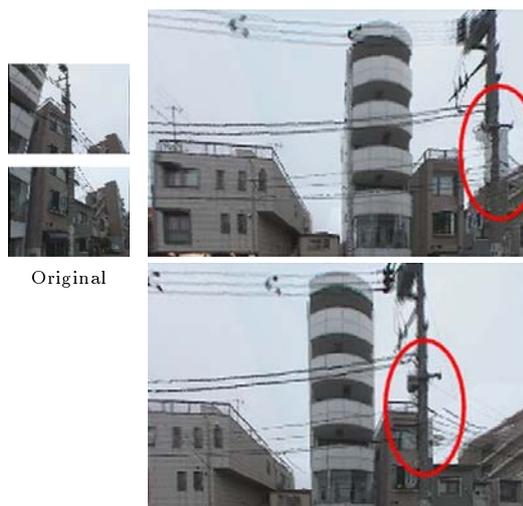


図 17 従来の射影変換によるカメラ画像の統合結果

Fig.17 Synthesized result of camera images by conventional projective transformation.

角を 60° 程度とすると、全方位画像の生成に必要なカメラの台数は、余裕を含めても 7 台ほどであり、カメラの大きさを考慮しても十分に通常のワゴン車などには直線上に並べることが可能である。

更に、図 19 はカーブ区間におけるパノラマ画像の生成結果である。ここでは速度変動が生じていることを勘案し、前後 1.5 秒分を含めた短区間で EPI マッチ

表 4 得られたカメラパラメータの数値例

Table 4 Examples of estimated camera parameters.

Camera pair $m, n$	Spatial parameter [pixel]			Temporal parameter $\delta_m - \delta_n$ [frame]		
	Estimated	Correct	Diff	Estimated	Correct	Diff
	1, 2	5	5	0	-54	-54
2, 4	15	15	0	80	80	0
1, 3	4	5	1	42	43	1
3, 5	3	3	0	38	39	1
1, 6	2	2	0	51	51	0
6, 8	1	1	0	22	22	0
1, 7	0	1	1	-63	-63	0
7, 9	1	1	0	97	96	-1

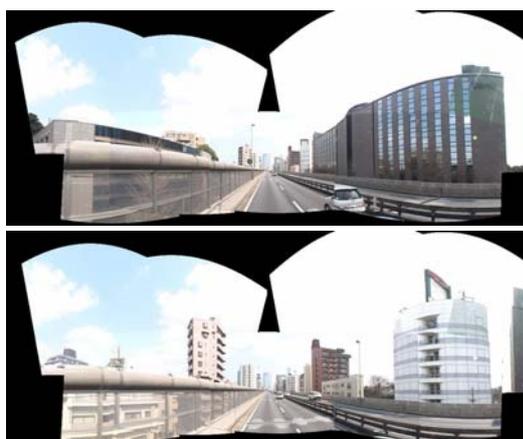


図 18 本手法により円筒上に生成した統合パノラマ画像  
Fig. 18 Synthesized result of cylindrical panorama image by our method.

ングを行って時間パラメータを求めている。等速直線運動でないカメラの動きによる光学中心のずれから生じるひずみは微小であり、一般的に等速を意識する程度の運転、及び通常のカーブ撮影された映像においても本手法を適用することは可能であるといえる。

参考として、EPI マッチングにより得られたカメラパラメータの数値例を表 4 に記載する。Spatial parameter はマッチングにおける図 6 の  $u$  軸方向の重なり量、Temporal parameter は対応する時刻（フレーム）の差である。それぞれ Estimated（推定値）は本手法を用いて自動的・数理的に得られた値、Correct（正解値）は手動によりマッチングを行った際に得られた値であり、Diff（誤差）はそれらの差をとったものである。これによれば、いずれも Diff がほぼ 0 になっており、高い精度で推定が行えていることが分かる。

## 5.2 実験 (II)：仮想自由視点画像の合成

図 16 で得られたパノラマ画像データに対し、Taka-



図 19 カーブ区間の統合パノラマ画像  
Fig. 19 Synthesized result of cylindrical panorama image in a curved zone.

hashi ら及び Shum らの画像ベースレンダリング (IBR) 手法 [2], [15] を用いて撮影経路外の仮想視点から見た画像を合成した。結果を図 20 に示す。同図 (a) ~ (c) は仮想視点から見た場合のパノラマ画像であり、いずれの画像においても、奥行位置にかかわらずひずみなく画像を生成できている。また、仮想視点の変化に伴って図中の木と背後の建物の位置関係・遮蔽状況が適切に変化している。更に、同図 (e), (f) は仮想視点を (a) と同じ位置に置いた場合の使用カメラの台数による合成結果の比較である。複数台のカメラを用いることで視野が広がり、IBR 手法を用いた場合でも死角のない自然な画像統合ができていることが分かる。

## 6. む す び

本論文では、EPI 解析による複数カメラのキャリブレーション手法を提案し、それを利用して実際に画像の統合を行った。

一般に、カメラの光学中心はカメラの内部にあるため、複数カメラの光学中心を一点に集めることは物理的に困難である。これに対して提案手法は、時空間、すなわち時間及び空間上でカメラ位置を同時に考慮することにより複数カメラの光学中心の一致を EPI マッチングで実現した。したがって、本手法を用いれば複数のビデオカメラを移動体（車両の屋根）に設置し、走行しながら撮影するだけで効率良く大規模シーンの画像を取得し、ひずみなく統合することができる。また、本手法では各カメラ映像の時間的重複性を利用するため、撮影範囲の空間的重なりが少なくとも安定して処理を行える。そのため、大規模な空間の取得を効率的に行うことができるという利点もある。

本手法の有効性を確認するため、実際の都市空間を対象として実験を行った。その結果、ひずみのないパノラマ画像を生成することに成功し、円筒上に進行方

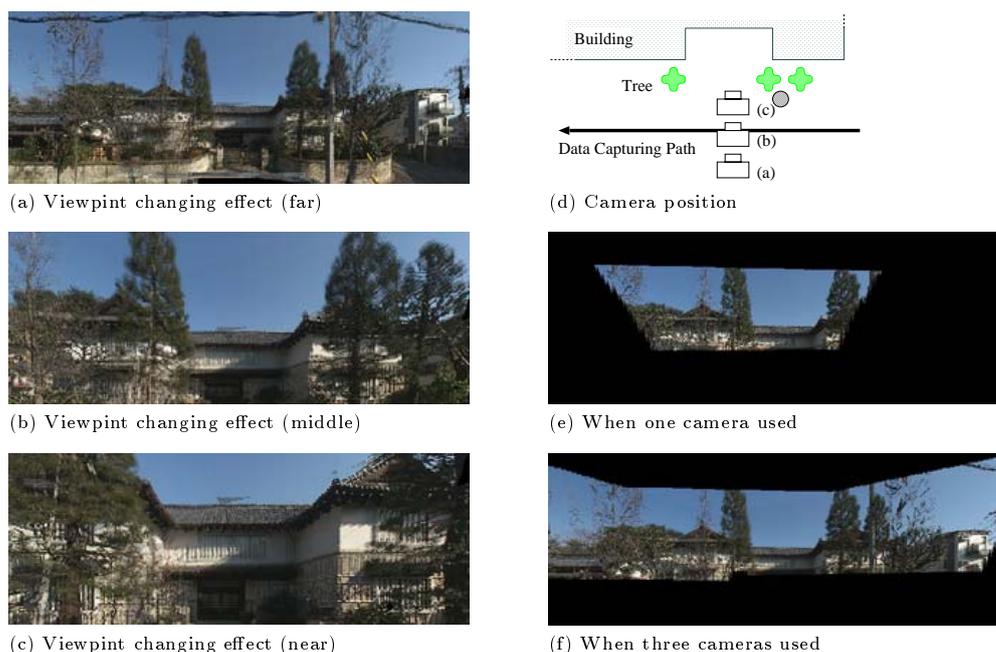


図 20 自由視点画像の生成  
Fig.20 Generation of free-viewpoint image.

向のパノラマ画像を生成することもできた。また、厳密な等速直線運動を行っていないと考えられる区間に対しても本手法を適用し、4.における誤差評価が視覚的にも信頼できるものであることが確認できた。更に、全方位画像の特性を利用して、IBRによる撮影地点以外の仮想視点からの画像を生成することにも成功した。

また、本研究はカメラの台数を増やすことで全方位画像の生成へと拡張されるので、今後の展望として、全方位画像を生成する際には各カメラごとにパラメータ推定誤差を分散させるなど、生成対象に応じたよりいっそうの最適化を行うことや、屋外大規模空間モデルの構築、表示システムの開発などが考えられる。

#### 文 献

- [1] M. Hirose and E. Takaaki: "Building a virtual world from the real world," International Symposium on Mixed Reality, pp. 183-197 (1999).
- [2] T. Takahashi, H. Kawasaki, K. Ikeuchi and M. Sakauchi: "Arbitrary view position and direction rendering for large-scale scenes," Vol. 2 Comput. Vis. Pattern Recognit., pp. 296-303 (2000).
- [3] 遠藤隆明, 谷川智洋, 広田光一, 廣瀬通孝: "超多眼カメラによる全天周画像の再構成," 情処学論, 43, SIG 11(CVIM 5), pp. 43-52 (2002).
- [4] S. K. Nayar: "Omnidirectional video camera," Comput. Vis. Pattern Recognit., pp. 482-488 (1997).
- [5] Y. Onoue, K. Yamasawa, H. Takemura and N. Yokoya: "Telepresence by realtime view-dependent image generation from omnidirectional video streams," Vol. 71 Comput. Vis. Image Understand., pp. 154-165 (1998).
- [6] Y. Yagi and S. Kawato: "Panorama scene analysis with conic projection," Int. Conf. on Intelligent Robots and Systems, pp. 181-187 (1990).
- [7] K. Yamasawa, Y. Yagi and M. Yachida: "New realtime omnidirectional image sensor with hyperboloidal mirror," 8th Scandinavian Conf. on Image Analysis, pp. 1381-1387 (1993).
- [8] J. Zheng and S. Tsuji: "Panoramic representation of scenes for route understanding," Int. Conf. on Pattern Recognition, pp. 161-167 (1990).
- [9] 棚橋英樹, 島田大輔, 山本和彦, 丹羽義典: "全方向ステレオシステム (sos) を用いた移動体追跡," View2000, pp. 67-72 (2000).
- [10] 山本和彦, 棚橋英樹, 桑島茂純, 丹羽義典: "実環境センシングのための全方向ステレオシステム (sos)," 電学論 (C) 電子・情報・システム部門誌, 121-C, pp. 876-881 (2001).
- [11] 佐藤雄隆, 王彩華, 棚橋英樹, 丹羽義典, 山本和彦, 桑島茂純: "移動体ビジョンを指向した小型全方向ステレオシステム (minisos) の開発," 画像センシングシンポジウム, pp. 311-316 (2003).
- [12] R. Bolles, H. Baker and D. Marimont: "Epipolar

plane image analysis: An approach to determining structure from motion,” Vol. 1 Int. J. of Computer Vision, pp. 7-55 (1987).

- [13] 植芝俊夫, 富田文明: “平面パターンを用いた複数カメラシステムのキャリブレーション,” 情処学論, 44, SIG 17(CVIM 8), pp. 89-99 (2003).
- [14] H. Maitre, E. Lutton and J. Lopez-krahe: “Determining vanishing points using hough transform,” IEEE Trans. Pattern Anal. Machine Intell, 16, pp. 430-438 (1994).
- [15] H. Shum and L. W. He: “Rendering with concentric mosaics,” SIGGRAPH '99 Proc. ACM, pp. 299-306 (1999).

## 付 録

### 画像の相関係数の計算式

3.4.2 において利用しているエネルギー関数としては、線形相関係数を用いている。なお、任意の二つの画像  $X, Y$  の座標  $(i, j)$  の画素値を  $x_{ij}, y_{ij}$ 、それらの平均値を  $\bar{x}, \bar{y}$  で示したときの相関係数  $r$  は

$$r = \frac{\sum_i \sum_j (x_{ij} - \bar{x})(y_{ij} - \bar{y})}{\sqrt{\sum_i \sum_j (x_{ij} - \bar{x})^2} \sqrt{\sum_i \sum_j (y_{ij} - \bar{y})^2}} \quad (\text{A.1})$$

であり、すべての画素値が同じ画像（すなわちすべての  $i, j$  において  $x_{ij} = y_{ij}$ ）においては 1 となる。

（平成 17 年 9 月 26 日受付，12 月 28 日再受付）



三上 武志（学生員）

2005 埼玉大・工・情報システム卒。現在，同大大学院・理工学研究科情報システム工学専攻博士前期課程在学中。コンピュータビジョンや映像解析に関する研究に従事。



タンダ ウー

1994 Yangon University, 理学部 卒。1998 University of Computer Studies, Yangon, 修士了。2002 埼玉大学大学院・理工学研究科情報システム工学専攻博士後期課程 入学。現在，埼玉大学大学院・理工学研究科情報システム工学専攻博士後期課程在学中。コンピュータビジョンや照明・反射解析に関する研究に従事。



小野晋太郎

2001 東大・工・電子情報 卒。2003 同大大学院・情報理工・電子情報学専攻修士了，2006 同博士了（情報理工学）。現在，同大生産技術研究所博士研究員。主に時空間情報解析，ITS に関する研究に従事。



川崎 洋（正員）

1994 京大・工・電気電子卒。2003 東大・工・電子情報工学 博士課程了。博士（工学）。現在，埼玉大・工・情報システム工学科助教授。主として広域空間のモデリング，三次元ビジョン，テクスチャ解析に関する研究に従事。情報処理学会，IEEE 各

会員。



大沢 裕（正員）

1976 信州大・工・電子 卒。1978 信州大大学院・工・電子 修士了。東大生研助手，埼玉大・工・助手，助教授などを経て，1998 より，埼玉大・工・情報システム工学科教授。工博。地理情報システム，時空間情報システムの研究に従事。信学会論文賞受賞。情報処理学会，映像情報メディア学会，日本 GIS 学会，ACM，各会員。



池内 克史（正員）

1973 京大・工・機械 卒。1978 東大大学院・工・情報工学 博士了。博士（工学）。MIT 人工知能研究所，電総研，CMU 計算機科学部を経て，1996 より東京大学生産技術研究所教授。現在，同大情報学環教授。人間の視覚機能，明るさ解析，物体認識，人間による組立作業の自動認識，文化財のデジタル保存，ITS などの研究に従事。論文賞（ICCV-90, CVPR-91, AIJ-92, ロボット学会誌-97, IEEE R&A 誌-98, 日本バーチャルリアリティ学会論文誌-99）等受賞。人工知能学会，日本ロボット学会，日本バーチャルリアリティ学会，OSA，IEEE 各会員（Fellow）。

**Abstract** We present Epipolar Plane Image (EPI) based calibration method of multiple video cameras, which enables distortion-free fusion of each image. Generally, since an optical center of a camera exists inside its body, it is physically impossible to coincide them at one point and distortions inevitably occurs when mosaicing them. In this paper, we propose to bring optical centers together in spatio-temporal field under condition that cameras are installed on top of a vehicle and capture images by running along the street. Our method is useful of capturing hi-resolution omnidirectional images without using special external synchronization devices nor positioning devices.

**Key words** epipolar plane image(EPI), spatiotemporal analysis, image mosaicing, omnidirectional image