

没入型ディスプレイを用いた全方向ステレオシステム (SOS) のカメラキャリブレーション

棚橋 英樹[†] 佐藤 淳[‡] 王 彩華[†] 丹羽 義典[†] 山本 和彦[§]

[†] (財) ソフトピアジャパン地域結集型共同研究推進室、科技団

[‡] 名古屋工業大学 電気情報工学科

[§] 岐阜大学 工学部 応用情報工学科

概要: 本論文では、我々が開発した全方向ステレオシステム (SOS) のカメラキャリブレーション手法について述べる。本システムは、3つのカメラからなるステレオユニットを正20面体の各面に配置することにより、システムを中心とした全ての方向のカラー画像とステレオ画像を同時に得ることが可能である。しかしながら本システムのキャリブレーションは、各ステレオユニットが異なる方向の画像を取得するため、各ユニット間のキャリブレーションが困難である。そのため、本システムを包含する没入型ディスプレイを用いて、パターン画像を投影し、その画像からシステムのキャリブレーションを行った。

A Calibration Method of Stereo Omni-directional System(SOS) Using a Fully Virtual Environment

Hideki Tanahashi[†], Jun Sato[‡], Caihua Wang[†], Yoshinori Niwa[†] and Kazuhiko Yamamoto[§]

[†]Office of Regional Intensive Research Project, Softopia Japan, JST

[‡] Department of Electrical and Computer Engineering, Nagoya Institute of Technology

[§] Department of Information Science, Faculty of Engineering, Gifu University

Abstract This paper proposes a calibration method of Stereo Omni-directional System (SOS), which we developed. In this system, the stereo units, which are composed of three cameras, are arranged on each plane of regular icosahedrons. This system can acquire all directional color and stereo pair images at the same time. However the calibration of this system is difficult, because the system acquired images from different viewpoints from the center of the system. For the calibration of this system, we used acquired pattern images using a fully virtual environment around the sensor.

1 はじめに

近年、コンピュータにおける画像処理能力の向上により、複数のカメラから得られる多視点動画を用いて、広範囲の環境を同時に取得し、3次元情報を取得する研究が盛んに行われている。多視点画像を用いることで、広範囲の画像が高解像度で取得でき、広い範囲の監視、環境の理解や臨場感を持った仮想環

境の提示等様々分野での応用に有効である。

こうした複数のカメラからの多視点画像を利用するシステムにおいては、全てのカメラ間の3次元位置・姿勢のキャリブレーションを正確に行う必要がある。複数のカメラからの多視点画像から3次元情報を得ることが可能なシステムは、大別すると2つに分けることができる。一つは、複数のカメラを対象とする環境の周囲に配置し、環境の外から中心方向の画像を取得するシステムである。金出ら [1] は、51個のカメラを半球ドーム上に配置し、これら画像から3次元

モデルの復元を行った。このシステムでは、対象とする空間に3次元位置が既知のマーカを配置し、各カメラ毎に得られた画像から独立にカメラのキャリブレーションを行った。この方法は各カメラ独立にキャリブレーションを行うため、各カメラパラメータ間で矛盾が生じ、得られる3次元情報に誤差として含まれる。こうした問題を解決するためにカメラ間のエピポラ幾何から”射影グリッド空間”を構成し、この空間において3次元形状の復元を提案している [2]。文献 [2] では、直線上に配列した LED を用いてカメラ間の相対関係を求めており、3次元空間座標と2次元画像座標の正確な関係を必要とはしないが、空間内の複数箇所でのキャリブレーション画像の取得が必要となる。もう一方は、複数のカメラをシステムを中心から外向きに配置し、その周りの環境の画像を取得するシステムである。川西ら [3] は、六角錐ミラーと各ミラー面にカメラを配置した全方位カメラを2つ用いてステレオカメラを構成することで、システムを中心とした360度パノラマステレオ画像を取得するシステムを開発した。このシステムでは、各カメラから得られる画像中のミラー領域の位置と大きさを等しくすることで各カメラの光軸方向の調整と、隣り合う画像の連続性と上下画像のエピポラ線が一致するようにミラーと2つの全方位カメラの姿勢の調整を行った。この手法は、カメラの調整が隣り合う画像間や上下のシステム間の相対的な調整であり、それぞれにキャリブレーション誤差が含まれている。この問題を解決するためにセンサを囲む大型円筒スクリーンに3次元位置の既知な格子パターンを投影し、それを撮像した画像からキャリブレーションを行う手法を提案した [4]。これらのキャリブレーション手法は、複数のカメラから得られる情報の統合と同時に正確な3次元情報を取得を目的としていた。

我々は、これまで観測点を中心とした広域かつ動的な環境情報の取得を目的とし、空間全ての方向のカラー画像とステレオ画像対を同時刻にリアルタイムで取得可能な全方位ステレオシステムを開発した。本システムは、3個のカメラからなるステレオユニットを正20面体に配置しているため、3次元情報を得るためのカメラキャリブレーションと各ユニットから得られる画像や3次元情報の統合のためのユニットキャリブレーションを独立に行うことが可能である。我々はこれまで、本システムを用いて、動的環境における複数のイベントの同時抽出 [6]、現実環境からの3次元仮想環境の構築 [7] や各方向の画像から球面統合画像の生成方法を提案 [8] を行ってきた。各ユニットから得られる画像や3次元情報の統合には、システ

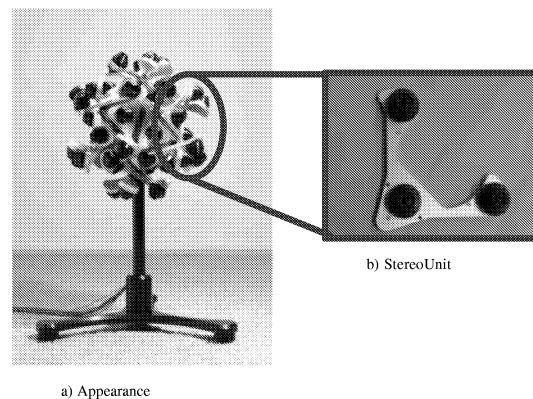


図 1: 全方位ステレオシステム (SOS)

ム設計時に用いたステレオユニットの配置パラメータを用いてきたが、ステレオユニットの取り付け時の誤差により、一部不連続が生じていた。そのため、本論文ではこのシステムを完全に包含するスクリーンで囲まれた空間を用いて、各ユニットの配置パラメータを推定し、得られたパラメータ初期値として、システム全体のパラメータの最適化により、システム全体のキャリブレーションを行ったので報告する。

2 全方位ステレオシステム (SOS)

全方位ステレオシステム (Stereo Omnidirectional System : 以下 SOS) [5] は、システムを中心とした3次元空間内の全方向のカラー画像と距離画像を同一時刻にリアルタイムで取得可能なシステムである。図 1a に本システムの外観を示す。本システムは、60個のカメラを用いて正20面体の各面上に3個のカメラ対からなるステレオユニット (図 1b) を配置した。これによってシステムを中心とした全ての方向における高解像度のカラー画像と3次元情報を同一時刻にリアルタイムで取得することが可能である。また、ステレオユニットを互いに隣り合うユニットのカメラの視野を遮らないように配置することにより、得られる3次元情報の精度を確保しながら、システム全体の大きさの小型化を図り、近接死角を少なくしつつ、システムから遠く離れた3次元情報を精度良く取得可能にした。表 1 にシステムの仕様を示す。

3 システムキャリブレーション

SOS は独立に3次元情報を取得可能なステレオユニットを各面に配置している。そのため、システム全

表 1: システム仕様

Image Sensor	1/3' CMOS Color Image Sensor
Effective Resolution	640 (H) * 480 (V)
Field of View	96.6 deg (H) * 71.9 deg (V)
Baseline Length	90mm
Sensor Size	Diameter 27cm
Sensor Weight	4.5kg

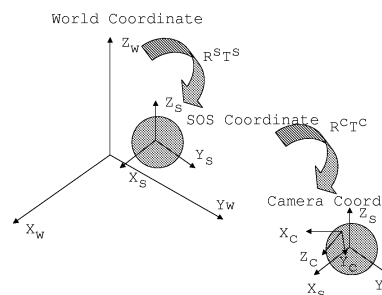


図 2: 世界座標とカメラ座標系の関係

体のキャリブレーションは、3次元の情報取得するためのステレオユニットにおけるカメラキャリブレーションと各ユニット間のキャリブレーションを独立に行うことが可能である。

図 2 にワールド座標とカメラ座標の関係を示す。ワールド座標とカメラ座標を関連づけるためには、ワールド座標における SOS システム座標の回転 R^sT^s 、並進 T^s と SOS システム座標における各カメラの回転 R_i^c 、並進 $T_i^c (i = 1, \dots, 20)$ の計 126 個のパラメータを求める必要がある。

以下、本節ではステレオユニット上の 3 個のカメラのキャリブレーションと各ステレオユニット間のキャリブレーション手法について述べる。

3.1 ステレオユニットキャリブレーション

ステレオユニット上の 3 個のカメラはアルミ板に固定されており、またカメラの焦点距離も固定しているため、一度、キャリブレーションを行えば正確な 3 次元情報の取得が可能である。本ユニットは、Tsai の手法 [9] をベースとしてキャリブレーションが行われている。これにより、各カメラ毎の内部パラメータ（レンズ歪み、焦点距離、カメラ中心、水平垂直のスケール比）と 3 個のカメラ間の位置関係のパラメータは正確に得られている。

得られたキャリブレーションデータを用いることにより、各ユニットからはレンズ歪みとカメラ間のアライメントずれの影響が除かれた補正画像が出力される。この補正画像の各ピクセルは、正方ピクセルをなしており、画像中心からの透視投影画像となっている。この校正精度は、640*480 の解像度において、最小二乗誤差で 0.1~0.3 ピクセルになっている [10]。

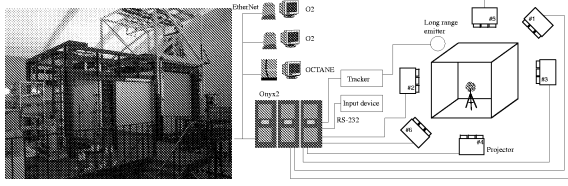
3.2 ユニット間キャリブレーション

前節で述べたように、ステレオユニット上の各カメラの内部パラメータは、各ユニットごとに正確にキャリブレーションが行われている。そのため、ワールド座標と各カメラ座標を関連づけるためには、ワールド座標における各カメラの外部パラメータ（並進、回転）を求めれば良い。これらパラメータを独立に求めることも考えられるが、それぞれに含まれる推定誤差の影響により各ユニット間で矛盾が生じる。

SOS においてはステレオユニット毎に精度良く加工されたアルミ板に正確に固定されており、システム中心から基準カメラ中心までの位置のずれは少ない。また、画像統合においては、並進のずれによる影響より、回転パラメータのずれの影響が大きい。そのため、SOS システム座標における各ユニットの回転とワールド座標と SOS システム座標との関係を求めることにより、全てのカメラ座標とワールド座標を関連づけることが可能になる。これにより、求めるパラメータは、SOS のシステム中心の並進、回転のパラメータ（6 個）と各ユニットの基準カメラの回転 ($3 * 20 = 60$ 個) の計 66 個となる。

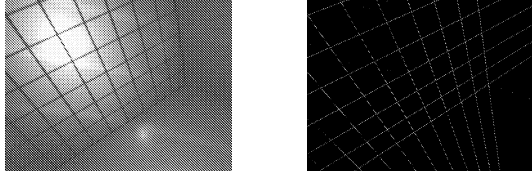
一般に複数のカメラ間のキャリブレーションを行うためには、3 次元形状が既知の校正器具または 3 次元位置の既知なマーカ等を撮影し校正する方法と、大きさも位置や姿勢もわからないような未知の対象物の画像から校正を行う方法（セルフ・キャリブレーション）がある [11]。

本論文では、SOS を中心とした 3 次元空間内に 3 次元位置の既知なマーカの画像を取得することで各カメラごとの外部パラメータを算出し、得られたパラメータを初期値として、システムの姿勢推定、各カメラの回転パラメータの最適化を行った。しかしながら、実空間の全方向に複数のマーカを精度良く配置することは困難なため、大きさと形状が既知の空間に



a) 概観図 b) システム構成図

図 3: 没入型 6 面ディスプレイ (COSMOS)



a) 入力画像 b) 直線抽出結果

図 4: 格子パターン画像例

繰り返しパターンを投影し、その画像を取得することでキャリブレーションを行った。キャリブレーションに用いた空間としては、図 3 に示すような没入型 6 面ディスプレイ [12] (以下 COSMOS と呼ぶ) を用いた。COSMOS は、1 辺が 3m のスクリーンで囲まれた立方空間で、この空間の外にあるプロジェクタから映像をスクリーンに投影することで完全な仮想空間を構築することができる。これら 6 枚のスクリーンの配置関係、大きさ及びプロジェクタの校正は、全てのスクリーンに投影したパターン画像を用いて正確にキャリブレーションが行われている。そのため、SOS を COSMOS のほぼ中央に設置して、1 面ごとにプロジェクタから格子パターンを投影し、各カメラごと取得した画像 (各 6 枚) から格子点の座標を求めて、その対応する 3 次元座標からキャリブレーションを行った。4a) に一つのユニットから得られる格子パターン画像の例を示す。なお、こうした校正されたプロジェクタを用いて、カメラキャリブレーションを行うことで、キャリブレーションパターンを特徴点が得やすいように簡単に変更できるといったメリットもある。

得られた画像からハフ変換を用いて直線群を抽出し、格子点座標を求めた。その格子点座標に 3 次元座標を対応づけた。この対応付けは人が行った。図 4b) にハフ変換で抽出した結果を示す。

以上より各カメラから得られた 2 次元画像上の点 $\mathbf{m}_i = [u, v]^T (i = 1, \dots, 20)$ とワールド座標における点 $\mathbf{X} = [X, Y, Z]^T$ から以下のように表現される。

$$\lambda \tilde{\mathbf{m}}_i = \mathbf{P}_i \tilde{\mathbf{X}} \quad (1)$$

ここで、ワールド座標における SOS 座標の並進を \mathbf{T}^s 、回転を \mathbf{R}^s とし、SOS 座標における各カメラ座標の並進、回転をそれぞれ \mathbf{T}_i^c 、 $\mathbf{R}_i^c (i = 1, \dots, 20)$ とすると、 \mathbf{P}_i は 3x4 の射影カメラ行列で、以下のように表せる。

$$\begin{aligned} \mathbf{P}_i &= \mathbf{A}_i \begin{bmatrix} \mathbf{R}_i^c & -\mathbf{R}_i^c \mathbf{T}_i^c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}^s & -\mathbf{R}^s \mathbf{T}^s \\ 0 & 1 \end{bmatrix} \\ &= \mathbf{A}_i \begin{bmatrix} \mathbf{R}_i^c \mathbf{R}^s & -\mathbf{R}_i^c \mathbf{R}^s (\mathbf{T}^s + \mathbf{R}^{s^{-1}} \mathbf{T}_i^c) \\ 0 & 1 \end{bmatrix} \\ &= \mathbf{A}_i \begin{bmatrix} \mathbf{R}_i & -\mathbf{R}_i \mathbf{T}_i \\ 0 & 1 \end{bmatrix} \quad (2) \end{aligned}$$

ただし

$$\mathbf{R}_i = \mathbf{R}_i^c \mathbf{R}^s \quad (3)$$

$$\mathbf{T}_i = \mathbf{T}^s + \mathbf{R}^{s^{-1}} \mathbf{T}_i^c \quad (4)$$

$$(i = 1, \dots, 20)$$

ここで、行列 \mathbf{A}_i はカメラの内部パラメータだけで構成される行列であり、3.1 節で述べたように、これらのパラメータは既に求められている。また、SOS 座標における各カメラの並進 (\mathbf{T}_i^c) は、先に述べたように設計パラメータを用いるため既知である。そのため、求めるパラメータは、ワールド座標における SOS 座標の並進 (\mathbf{T}^s)、回転 (\mathbf{R}^s)、SOS 座標における各カメラ座標の回転 (\mathbf{R}_i^c) の 66 個のパラメータとなる。そのため、式 1、式 2 より線形解法を用いて、ステップ 1 でこれら 66 個のパラメータの初期値を計算し、ステップ 2 でこの初期値をもとに 66 個のパラメータを最適化することで求めた。以下に各ステップにおける処理について具体的に述べる。

(ステップ 1)

1. 式 1 を用いて、 \mathbf{m}_i と \mathbf{X} から各カメラの \mathbf{P}_i を求める。
2. \mathbf{A}_i が既知であることにより、式 2 から、各カメラの \mathbf{R}_i と \mathbf{T}_i を求める
3. このままでは、誤差の影響により \mathbf{R}_i は回転行列とはならないため、特異値分解を用いて \mathbf{R}_i を正規直行化して回転行列 \mathbf{R}_i を求める。すなわち、 \mathbf{R}_i を特異値分解により $\mathbf{U}, \mathbf{W}, \mathbf{V}$ に分解し、 \mathbf{W} を単位行列に置き直して再び $\mathbf{R}_i = \mathbf{U}\mathbf{W}\mathbf{V}$ により合成する。
4. SOS の対称性を用いて、各カメラの \mathbf{R}_i と \mathbf{T}_i の平均値を取り、3 次元空間における SOS 座標の並進 \mathbf{T}^s 、回転 \mathbf{R}^s の初期値を求める。
5. $\mathbf{R}_i, \mathbf{R}^s$ から式 3 より各カメラの回転 \mathbf{R}_i^c の初期値を求める。

(ステップ 2)

ステップ1で求めたSOS座標の並進 \mathbf{T}^s と回転 \mathbf{R}^s 、および各カメラのSOS座標に対する回転 \mathbf{R}_i^c を初期値として、各カメラの画像 \mathbf{m}_i と3次元座標 \mathbf{X} の整合性が取れるよう $\mathbf{T}^s, \mathbf{R}^s, \mathbf{R}_i^c$ の66個のパラメータを同時に最適化する。

パラメータの最適化は、ステップ1で得られたパラメータ($\mathbf{T}^s, \mathbf{R}^s, \mathbf{R}_i^c (i = 1, \dots, 20)$)から各軸に対する回転角 $\theta_j^s, \theta_{i,j}^c (i = 1, \dots, 20, j = X, Y, Z)$ を求め、これを初期値として、カメラの画像 m_i と推定した P_i を用いて \mathbf{X} を投影した画像 \hat{m}_i の差 e_{ij} が最小となるよう次に示すように反復法で $T_j^s, \theta_j^s, \theta_{i,j}^c$ を変化させることにより行った。

$$T_j^s = T_j^s + k_{T^s} \frac{\delta e_{i,j}}{\sqrt{\sum_j \delta e_{i,j}^2}} \quad (5)$$

$$\theta_j^s = \theta_j^s + k_{\theta^s} \frac{\delta e_{i,j}}{\sqrt{\sum_j \delta e_{i,j}^2}} \quad (6)$$

$$\theta_{i,j}^c = \theta_{i,j}^c + k_{\theta^c} \frac{\delta e_{i,j}}{\sqrt{\sum_j \delta e_{i,j}^2}} \quad (7)$$

ここで、 $\delta e_{i,j}$ は、

$$\delta e_{i,j} = e_{i,j}(T_j^s, \theta_j^s, \theta_{i,j}^c) - e_{i,j}(T_j^s + k_{T^s}, \theta_j^s + k_{\theta^s}, \theta_{i,j}^c + k_{\theta^c}) \quad (8)$$

表現され、 i はカメラ番号(1~20)、 j は、X,Y,Z軸を指すインデックスである。また、 $k_{T^s}, k_{\theta^s}, k_{\theta^c}$ は、各パラメータごとの基準移動量である。式5~7による最適化を、画像座標と3次元座標との誤差が最小になるまで繰り返した。

4 実験結果

表2に本手法で得られた配置パラメータを用いてワールド座標における点を画像に投影し、画像座標との全カメラに対する平均誤差を示す。また、比較のため、CADで得られたパラメータを用いてワールド座標上の点を画像に投影した場合の結果を示す。また、図5に投影した画像例を示す。ここで○で表現された点は、画像上の格子点、×で表現された点は、画像に投影したワールド座標である。どちらのパラメータを用いた場合でも、カメラ全体での平均誤差はサブピクセル精度で得られているが、本手法により、高精度にキャリブレーションされているのがわかる。また、各カメラ独立に線形解法によりキャリブレーションを行った結果に対して、これらパラメータを用いて、SOS全体の最適化を行うことで、各カメラにおける推定誤差が小さくなっている。

表 2: 画像座標における平均誤差 (pixel)

最適化結果	0.2485
最適化前	0.3840
デザインパラメータ	0.7079

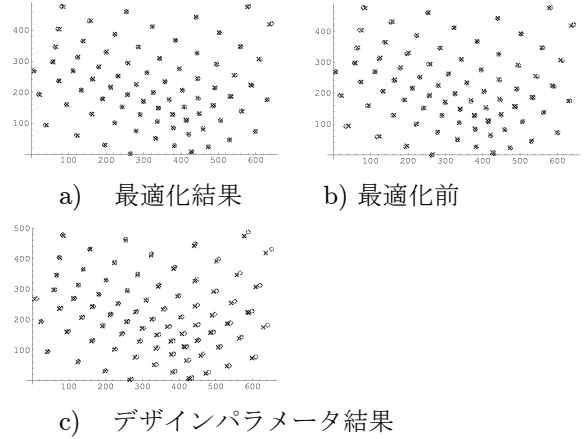
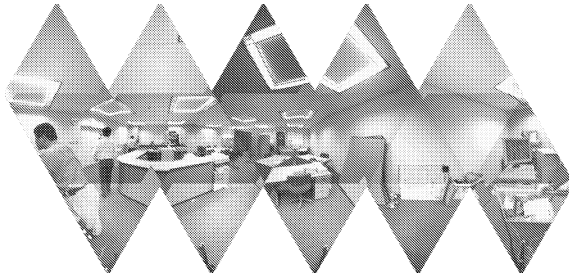


図 5: キャリブレーション結果

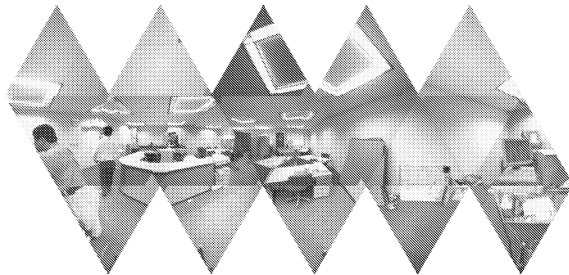
最後に得られた結果を用いて、SOSから得られる全カラー画像から全方向パノラマ画像を生成した。図6aは、取得した全カラー画像をカメラ視野角を用いて、システム中心から1.5mの距離で画像から3角形エリアを切り出して相対カメラ位置を用いて画像を補正し、作成した全方向パノラマ画像である。また、比較のため、図6bにデザインパラメータで同様な処理で作成した全方向パノラマ画像を示す。また、図7a,bに図6の部分拡大画像を示す。本手法により、画像間のつながりが滑らかに接続されていることがわかる。なお、図6aにおいて、一部不連続が見られるが、これはパノラマ画像を生成時にシステム中心から1.5mの距離で画像から3角形エリアを切り出しているためである。

5 むすび

本論文では、我々が開発した全方向ステレオシステムのカメラ間の配置パラメータのキャリブレーション手法について述べた。システムを中心とした全方向のカラー画像と3次元情報を取得可能なシステムであるため、このシステムを完全に包含するスクリーンで囲まれた空間に配置して、3次元位置が既知のパターン画像から各カメラ毎の配置パラメータを推定し、そのパラメータを初期値として全てのパラメータ

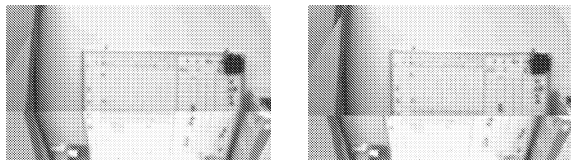


a) キャリブレーションパラメータ



b) デザインパラメータ

図 6: 全方向パノラマ画像



a) キャリブレーション後 b) キャリブレーション前

図 7: 図 6 の拡大画像

タの最適化を同時に行うことで、配置パラメータのキャリブレーションを行い、サブピクセル精度を実現した。

本論文では、3次元位置が既知のパターンを投影するために没入型ディスプレイを用いたが、今後、各カメラから得られる共通領域の画像特徴を用いたセルフキャリブレーションを行う予定である。

参考文献

- [1] T. Kanade, P. Rander and P. J. Narayanan: Virtualized Reality: Constructing Virtual Worlds from Real Scenes, *IEEE MultiMedia*, **4**, 1, pp.34-47 (1997)
- [2] H. Saito and T. Kanade: Shape Reconstruction in Projective Grid Space from Large Number of Images, *Proc. CVPR'99*, **II**, pp.49-54 (1999)

- [3] T. Kawanishi, K. Yamazawa, H. Iwasa, H. Takemura and N. Yokoya: Generation of High-resolution Stereo Panoramic Images by Omnidirectional Imaging Sensor Using Hexagonal Pyramidal Mirrors, *Proc. 14th ICPR*, **I**, pp.485-489 (1998)
- [4] 戎野聡一, 山澤一誠, 竹村治雄, 横矢直和: 円筒パターンを用いた全周ステレオ画像センサのキャリブレーション手法, *日本 VR 学会第 4 回大会論文集*, pp.211 - 212 (1999)
- [5] 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)*, **121-C**, 5, pp. 876 - 881 (2001)
- [6] 棚橋英樹, 島田大輔, 丹羽義典, 山本和彦: 全方向ステレオシステム (SOS) のステレオ情報を用いた空間イベントの取得, *情処学研報*, **CVIM125-14**, pp.103 - 110 (2001)
- [7] H. Tanahashi, D. Shimada, K. Yamamoto and Y. Niwa: Acquisition of Three-Dimensional Information in a Real Environment by Using the Stereo Omni-directional System(SOS), *Proc. 3rd 3DIM*, pp.365 - 371 (2001).
- [8] H. Tanahashi, C. Wang, K. Yamamoto and Y. Niwa: Live Events Accessing for Multiusers with Free Viewpoints using Stereo Omni-directional System, *Proc. PCM2001(2001)* (印刷中)
- [9] R. Y. Tsai: A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf TV cameras and lenses, *IEEE Journal of Robotics and Automation*, **3**, 4, pp.323-344 (1987)
- [10] 桑島茂純, 桑原裕之, 山本和彦, 棚橋英樹, 丹羽義典: 全方向ステレオカメラシステム, *O plus E*, bf 22, No.12, pp.1585 - 1591 (2000)
- [11] 佐藤淳: コンピュータビジョン—視覚の幾何学—, コロナ社
- [12] T. Yamada, M. Hirose and Y. Iida: Development of Complete Immersive Display: COSMOS, *Proc. VSMM'98*, **2**, pp.522 - 527 (1998)