

画像領域に適応したパノラマ映像生成技術の提案

磯貝 愛, 國田 豊, 木全 英明^{†1}

これまで、奥行きのあるシーンを対象に、高画質な3次元映像生成法である信頼度マッピング法を用いたパノラマ映像生成法を提案してきた。しかし、領域のすべての点を2台以上のカメラで撮影しなければならないため、カメラ台数に対して表現できる画角が狭く、合成処理に要する計算機負荷が大きいという問題があった。本稿では、画像間の一部のみが重なり合うように撮影した画像から、重なりがある領域と重なりがない領域に分割し、領域ごとに異なる画像処理を施すパノラマ映像生成法を提案する。提案手法では、重複領域には従来手法と同様に信頼度マッピング法を用いた3次元画像処理、非重複領域には重複領域の情報から推定した奥行き位置に画像を投影することにより、従来手法と比較して少ない計算機負荷で連続的なパノラマ映像を生成できることを示す。

Proposal of Panoramic Video Rendering Technology Adapting to Image Areas

MEGUMI ISOGAI, YUTAKA KUNITA, HIDEAKI KIMATA^{†1}

A panoramic video rendering scheme has been proposed that uses a probability mapping method to generate high-quality, roust 3-D images. In this method, the field angle is small and the amount of synthesis calculation is large because at least two cameras are needed to capture all points in a region. Our scheme handles image regions in an adaptive manner by dividing them into an overlap area and a non-overlap area. It uses a probability mapping method to process 3-D images in the overlap area, then projects the images into the non-overlap area in a position estimated from overlap area information. Experimental results indicate that our scheme provides continuous synthetic images and that the amount of synthesis calculation required is smaller than that in conventional methods.

1. はじめに

互いに重なるように撮影した複数の画像を合成し、解像度の高いパノラマ映像を生成する技術が研究されている。その多くは風景のような遠景を対象としたものであり、被写体が存在する位置に仮定した1枚の投影面に画像を貼り合わせることでパノラマ映像を生成する。しかし、仮定した投影面から離れた被写体が多重像となって表れるため、奥行きのあるシーンには適さないという問題がある。これに対して、複数の画像から被写体の奥行き情報を推定し、その位置情報に基づき画像を投影する方法が提案されている。しかし、奥行き推定が困難な被写体の場合、誤った位置に画像が投影され画質が低下するという問題がある。

これまで、奥行きのあるシーンでも連続的で解像度の高いパノラマ映像を得るために、高画質な3次元映像生成法である信頼度マッピング法¹⁾を用いたパノラマ映像生成技術を考案してきた²⁾。この手法は、被写体を多層平面で表現し、各平面における被写体表面が存在する信頼度を算出することにより、奥行き推定が困難な被写体でも良好な画質の画像を生成できるという特徴がある。しかし、被写体表面が存在する信頼度を算出するためには、表現したい領域の全ての点を2台以上のカメラで撮影していなければならない、隣接カメラ間で半分以上の重なり領域が必要とされる。そのため、カメラ台数に対して表現できる画角が狭く、3次元画像処理に要する計算機負荷が大きい。

そこで本稿では、全領域の被写体を正確に表現することから、画像間をつなぎ目なく連続的に表現することに重点を変更し、画像間の一部のみが重なった画像からパノラマ画像を合成する方法を提案する。提案手法では、重なりがある領域と重なりがない領域に分割し、重なりがある領域にはこれまでと同様に信頼度マッピング法を適用し、重なりがない領域には2次元画像処理を施す。ただし、奥行きのあるシーンを撮影しているため、誤った位置に画像を投影すると、重なりがある領域との境界でつなぎ目が発生する。そこで、重なり領域に存在する被写体表面の情報から、重なりのない領域の投影位置を推定することにより、領域間でつなぎ目のないパノラマ映像を生成可能とする。また、3次元画像処理が必要な領域を減らすことで、従来手法に対して処理の高速化を実現する。

2章では、関連するパノラマ映像生成技術について述べる。3章では、これまでに考案してきた信頼度マッピング法を用いたパノラマ映像生成技術について説明した後、4章では、

^{†1} 日本電信電話株式会社 NTT サイバースペース研究所
NTT Cyber Space Laboratories, NTT Corporation

本稿で提案する画像領域に応じたパノラマ映像生成技術について詳細を述べる。5章で提案手法の実験結果を示し、最後にまとめと今後の課題を述べる。

2. 関連研究

複数の画像からパノラマ映像を合成する方法は、大きく被写体の奥行き情報を利用しない方法と、奥行き情報を利用する方法に分けられる。奥行き情報を利用しない方法は、モザイクング手法と呼ばれ、画像に射影変換を施し重ね合わせることによって、パノラマ映像を生成する。モザイクング手法の中でも、カメラの投影中心と回転中心を一致させた状態で撮影した画像間には視差が生じないため、奥行きのあるシーンにおいても連続的な全方位映像を生成することができる³⁾。しかし、カメラの投影中心と回転中心を一致させるためには、高精度な位置合わせが必要である。また、1台のカメラを回転させながら撮影するため、同期のとれたパノラマ映像を取得することができない。

カメラの投影中心が一致していない場合でも、カメラから十分遠方の被写体は平面とみなすことができるため、被写体が存在する位置に投影面を仮定し、各画像を投影することでパノラマ画像を生成することができる⁴⁾。しかし、奥行きのあるシーンにおいては、仮定した投影面から離れた被写体が、パノラマ画像上で多重像となって表れる。

一方、奥行き情報を利用する方法として、複数の画像から推定した被写体の奥行き情報を表すデプスマップを用いて、被写体の奥行きに応じてパノラマ画像を生成する方法が提案されている⁴⁾。この手法は、奥行きのあるシーンにおいても、手前のものから奥のものまで連続的な画像を生成することができる。しかし、奥行き情報の推定にステレオマッチングを用いているため、被写体のテクスチャが乏しい場合やオクルージョンがある場合には、奥行き情報の推定が困難であり、推定誤りが生成画像の画質に影響を与える。

3. 信頼度マッピング法を用いたパノラマ映像生成法

3.1 概要

信頼度マッピング法を用いたパノラマ映像生成技術の概要を図1に示す。信頼度マッピング法とは、被写体を多層の平面レイヤによって表現し、各レイヤ $L_j (j \in J)$ を被写体表面の存在確率を表す信頼度マップ $\beta_j(\mathbf{m})$ と、色情報を表す色マップ $I_j(\mathbf{m})$ で構成する3次元映像生成処理技術である¹⁾。この技術を用いると視点 O_v から見たパノラマ画像 $I_v(\mathbf{m})$ は次式によって求められる。

$$I_v(\mathbf{m}) = \sum_{j \in J} \beta_j(\mathbf{m}) \cdot I_j(\mathbf{m}) \quad (1)$$

式(1)は、レイヤ上の2次元座標 $\mathbf{m} = [u, v]$ において、被写体表面が存在する信頼度を重み係数として色情報を描画することを表し、図1に示すように、被写体表面が存在する確率が高いと推定された点は鮮明に、低いと推定された点は不鮮明に表現される。デプスマップのように被写体の奥行きを一意に求めるのではなく、被写体表面が存在する信頼度として複数の値を持つことで、奥行き推定が困難な被写体でもロバストな画像を生成することができる。以降、信頼度マップおよび色マップの算出方法について述べる。

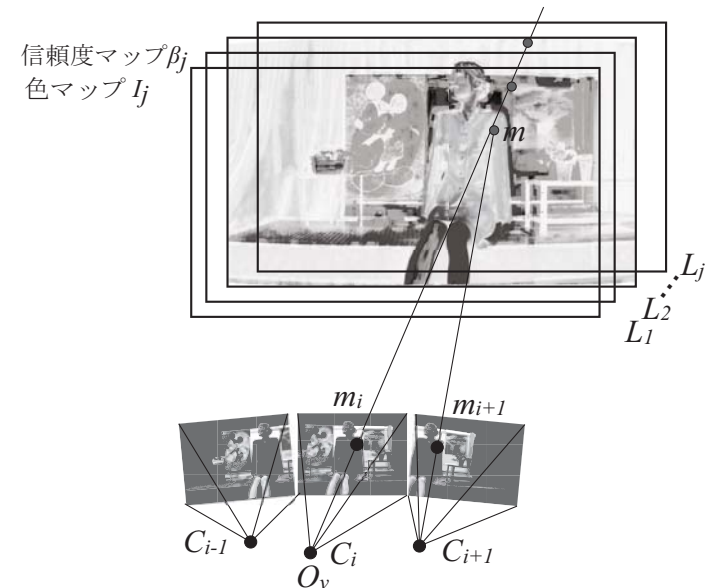


図1 信頼度マップを用いたパノラマ映像生成法の概要図

Fig.1 Conceptual diagram of panoramic video rendering method using probability maps

3.2 信頼度マップの算出

信頼度マップ $\beta_j(\mathbf{m})$ とは、レイヤ L_j 上の点 \mathbf{m} に被写体表面が存在する信頼度を表す関数であり、視点 O_v を通る直線上の点の信頼度の合計が1となるように設定する。

$$\sum_{j \in J} \beta_j(\mathbf{m}) = 1, \quad 0 \leq \beta_j(\mathbf{m}) \leq 1 \quad (2)$$

この信頼度マップ $\beta_j(\mathbf{m})$ は、レイヤ上の点 \mathbf{m} に対応する画像上の点の一致度を表わす $S_j(\mathbf{m})$ により、次式によって算出される。

$$\beta_j(\mathbf{m}) = \frac{S_j(\mathbf{m})}{\sum_{j \in J} S_j(\mathbf{m})} \quad (3)$$

ただし、一致度を表す $S_j(\mathbf{m})$ は、次式によって求められる。

$$S_j(\mathbf{m}) = \frac{1}{(\sigma_j^2(\mathbf{m}))^n + \epsilon} \quad (4)$$

ここで、 ϵ は 0 で割ることを避けるための微小値であり、 n は信頼度のピーク幅を調整するパラメータである。 $\sigma_j^2(\mathbf{m})$ は、対応点の近傍領域における色情報の相関関数であり、次式によって求められる。

$$\sigma_j^2(\mathbf{m}) = \frac{1}{NB} \sum_{i \in N} \sum_{\tilde{\mathbf{m}}_i \in B} \left\{ I_{C_i}(\tilde{\mathbf{m}}_i) - \frac{1}{N} \sum_{i \in N} I_{C_i}(\tilde{\mathbf{m}}_i) \right\}^2 \quad (5)$$

N はレイヤ上の点 \mathbf{m} に対応するカメラの台数、 B は近傍領域のブロックサイズを表す。 $\mathbf{m}_i = [u_i, v_i]$ は、レイヤ上の点 \mathbf{m} に対応するカメラの画像座標、 $\tilde{\mathbf{m}}_i$ は \mathbf{m}_i の近傍領域の画像座標を表す。

ここで、レイヤ上の点 \mathbf{m} に対応するカメラの画像座標 \mathbf{m}_i は、レイヤ L_j の座標をカメラ C_i の画像座標に変換するホモグラフィ行列 \mathbf{H}_i^j を用い、次式によって表わされる⁵⁾。

$$\begin{aligned} \tilde{\mathbf{m}}_i &= \mathbf{H}_i^j \tilde{\mathbf{m}} \\ &= \mathbf{A}_i \left(\mathbf{R}_i + \frac{\mathbf{t}_i \mathbf{n}^T}{d} \right) \mathbf{A}_v^{-1} \tilde{\mathbf{m}} \end{aligned} \quad (6)$$

$\mathbf{A}_i, \mathbf{A}_v$ はカメラの内部パラメータ、 $\mathbf{R}_i, \mathbf{t}_i$ はレイヤ座標系からカメラ座標系への回転行列と並進ベクトルを表し、事前のキャリブレーションにより求められるパラメータである。

3.3 色マップの算出

色マップ $I_j(\mathbf{m})$ は、各レイヤにおける被写体表面の色情報を表わすものであり、レイヤ上の点 \mathbf{m} に対応する画像の画素値 $I_{C_i}(\mathbf{m}_i)$ を重み付け加算することによって生成される。

$$I_j(\mathbf{m}) = \sum_{i \in N} w_{ij} I_{C_i}(\mathbf{m}_i) \quad (7)$$

重み係数 w_{ij} は、視点 O_v 、カメラ C_i の中心、レイヤ座標 \mathbf{m} との相対的な位置関係に

よって決められる。例えば、位置関係を角度によって表わした場合、重み係数は次式によって計算される。

$$w_{ij}(\mathbf{m}) = \frac{\frac{1}{1 - \cos \theta_{ij}}}{\sum_{i \in C_i} \frac{1}{1 - \cos \theta_{ij}}} \quad (8)$$

ここで、 θ_{ij} は、カメラ C_i の中心とレイヤ上の点 \mathbf{m} を結ぶ線と、全方位画像中心 O_v とレイヤ上の点 \mathbf{m} を結ぶ線がなす角を表し、なす角が小さいほどそのカメラが色マップに寄与する割合が高くなる。

4. 画像領域に適応したパノラマ映像生成法

これまでの手法では、表現したいシーン全体に対して被写体表面の信頼度を算出するために、すべての点が 2 台以上のカメラで撮影されていなければならない。そのため、カメラ台数に対して表現できる画角が狭く、3 次元画像処理に要する計算機負荷が大きい。

この問題を解決するためには、画像間の重なり領域を減らす必要があるが、重なり領域を減らしていくと重なりがない領域が生じる。以降、本稿では画像間の重なりがある領域を重複領域、重なりがない領域を非重複領域と呼ぶ。非重複領域は 1 台のカメラのみで撮影しているため、その領域に存在する被写体の信頼度を求めることができない。また、モザイクング手法のように 2 次元画像処理を施す場合も、そのカメラの情報だけでは被写体位置を推定できないため、誤った位置に画像が投影され重複領域と非重複領域との間につなぎ目が発生する。

そこで本稿では、パノラマ画像を重複領域と非重複領域に分割し、領域ごとに異なる画像処理を施すことで、被写体の奥行きに関わらず連続的に表現可能なパノラマ映像生成法を提案する。まず、重複領域については、従来通り信頼度マッピング法を適用する。その後、重複領域の信頼度情報から非重複領域の投影位置を推定し、2 次元画像処理を施す。以降、非重複領域の投影位置の推定方法について具体的に述べる。

4.1 投影位置推定法

画像間の一部のみが重なっている場合の信頼度マップを図 2 に示す。図 2 に示すように、画像間の重なりがある重複領域では、事前に各レイヤにおける被写体表面の信頼度が算出されている。非重複領域では、被写体が多層平面のどこに存在するか推定できないため、重複領域の情報から最適な投影面を 1 枚選択する。非重複領域は、その場所に応じて 2 種類に分けられる。1 つはパノラマ映像の両端に存在し、1 つの重複領域に隣接した領域である。

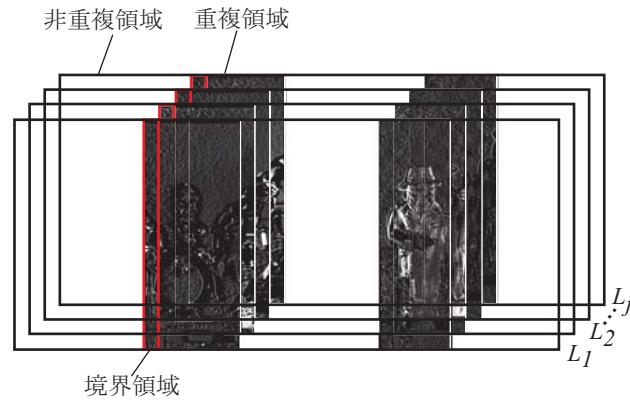


図2 画像間の一部のみが重複している信頼度マップ
Fig. 2 Probability maps generated from partially overlapped images

もう1つは重複領域の間に存在する領域である。

1つの重複領域に隣接した非重複領域では、重複領域と非重複領域に被写体がまたがるように存在しているため、その境界に存在する被写体の位置が非重複領域の投影位置となる可能性が高い。すなわち、被写体表面の信頼度が最大となるレイヤの位置が非重複領域の投影位置となる。そこで、被写体表面の信頼度が最大となるレイヤ l を次式により算出する。

$$l = \max_{j \in J} \sum_{m \in A} \beta_j(m) \quad (9)$$

ここで、 A は重複領域において非重複領域と接する領域であり、撮影対象に応じて調整可能なパラメータである。以降、この領域を境界領域と呼ぶ。そして、式(9)で求められたレイヤ l の非重複領域の信頼度の値を1、それ以外のレイヤの信頼度の値を0に設定する。

つぎに、重複領域の間に存在する非重複領域では、前者の領域と同様に被写体は境界領域にまたがるように被写体は存在するが、それぞれの境界領域で被写体の位置は異なる。そこで、非重複領域の被写体は、それぞれの境界領域に存在する被写体の中間位置に存在すると仮定する。すなわち、各境界領域の被写体表面の信頼度が最大となるレイヤの中間位置が、非重複領域の投影位置となる。式(9)により、それぞれの境界領域で信頼度情報が最大となるレイヤを求め、その中間位置となる非重複領域のレイヤの信頼度の値を1、それ以外のレイヤの信頼度の値を0に設定する。

このように提案手法では、重複領域の被写体表面の信頼度情報から、非重複領域の投影面

	信頼度マップ生成	投影位置推定	色マップ生成	レンダリング	合計
従来手法	34.3sec	-	6.2sec	4.2sec	44.7sec
提案手法	16.5sec	0.1sec	4.4sec	4.2sec	25.2sec

表1 各処理の所要時間
Table 1 Process time of each procedure

を推定することにより、奥行きのあるシーンでも境界領域でつなぎめのないパノラマ映像を得ることができる。また、被写体を多層平面で表現するという構成を維持することにより、これまでの合成手順を大きく変更することなくパノラマ画像を生成できるという特徴がある。

5. 実験

5.1 処理時間の計測

画像間の重複領域が半分以上(重複率50%)の従来手法と、画像間の重複領域が一部のみ(重複率30%)の提案手法との合成時間を比較するために各処理の所要時間を計測した。その結果を表1に示す。計測にはCPU2.8GHz、メモリ3GBを搭載したデスクトップPCを使用した。重複領域の割合を50%から30%に減らしたことにより、信頼度マップ生成処理が34.3secから16.5secに減少した。一方、提案手法で追加した投影位置推定はわずか0.1secの増加だったため、従来手法に対して約1.8倍の高速化を達成した。ただし、実時間性を求められるアプリケーションに適用するためには、更なる高速化が今後の課題である。

5.2 画質比較

次に、生成画像の性能を評価するために、3台のカメラで静止物体を撮影し、中央カメラを中心としたパノラマ画像を生成した。図3にカメラと被写体との位置関係を示す。被写体として4体の演奏者のミニチュアを用意し、中央カメラから被写体の中心までの距離が図3の位置となるように設置した。

実験に使用したカメラはPointGrey社のFleaで、焦点距離50mmのレンズが搭載されており、解像度UXGA(1600×1200)の画像を撮影することができる。カメラの間隔は50mmで、隣接カメラ間で画像領域の約30%が重複するように配置した。各カメラの撮影画像を図3の上部に示す。各画像の重複領域にはドラムとベース演奏者が存在する。

上述の実験環境で中央カメラを中心として、提案手法により生成した画像を図4、比較手法の画像を図5に示す。比較手法には、画像間の一部のみが重なった画像から生成可能なモザイク手法を採用した。

モザイク手法では、4体の被写体が配置された中央1575mmの位置を投影面と仮定

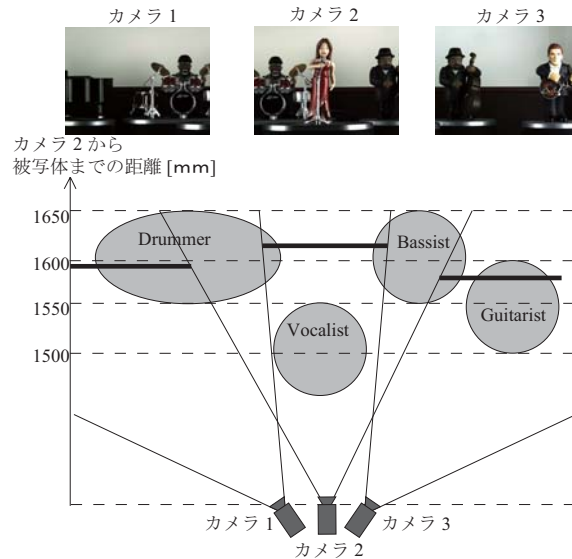


図3 実験環境
Fig. 3 Experimentation environment

し画像を生成した。その結果、重複領域に存在する被写体は多重像となり、すべての奥行き
の被写体を表現できていない。また、重複領域と非重複領域との境界においても、被写体の
エッジが一致しておらず、つなぎ目が発生している。

一方、提案手法では、重複領域については従来手法と同様に手前から奥まで連続的に表現
されている。また、提案手法により推定された非重複領域の投影面は図3の実線位置であ
り、境界領域に存在する被写体の位置が選択されている。しかし、カメラ2とカメラ3の
境界領域に存在するベースの下部は連続的に表現されているが、上部でつなぎ目が発生し
ている。このように、境界領域の被写体の信頼度情報から非重複領域に存在する被写体の位置
が推定できても、生成画像上で不連続な部分が発生することが明らかとなった。

6. 考 察

実験により明らかとなった境界領域で不連続な部分が発生する問題について考察を行った。
実験において、境界領域に存在するベース 2cm を表現するために必要なレイヤ数は 5 枚で

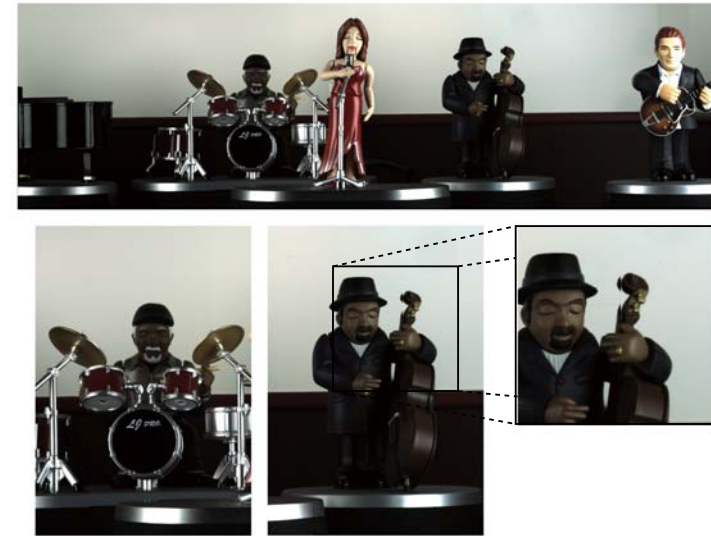


図4 上段：提案手法により生成したパノラマ画像
下段：拡大画像

Fig. 4 Upper figure: Panoramic image generated by the proposal technique
Lower figure: Enlarged images

あり、ベース全体の連続性を満たす投影面を選択することは困難である。図3では、複数の
レイヤの中から境界領域に占める割合が大きい物体が存在する投影位置が選択されている。
これは、境界領域に存在する被写体表面の信頼度の総和を基準としているためである。

そこで、投影面の位置を変更したときに生成画像にどのような影響を及ぼすか検証するた
めに、図4でつなぎ目が生じている部分が存在する位置に投影面を変更した。その結果を図
6に示す。図4でつなぎ目が生じていた部分のエッジが連続的に表現されており、境界領域
に被写体が複数の奥行きにわたって存在する場合には、部分ごとに最適な投影位置の選択が
必要であるといえる。今後は、境界領域に存在する被写体の奥行き幅を考慮した投影位置の
選択が課題である。

7. おわりに

本稿では、画像間の一部のみが重なった画像から重複領域と非重複領域に分割し、領域ご



図5 上段：モザイクキング手法により生成したパノラマ画像
下段：拡大画像

Fig.5 Upper figure:Panoramic image generated by the mosaicing technique
Lower figure:Enlarged images



図6 投影位置を変更した後の画像
Fig.6 Image after changing projection position

とに異なる処理を施すパノラマ映像生成法を提案した。実験により、画像間の重複領域を減らすことにより、従来手法に比べて合成処理が高速化されることを確認した。また、モザイクキング手法と比較して、提案手法は画像間の境界でつなぎ目のない連続的な映像を生成できることを示した。しかし、境界領域に存在する被写体の奥行きが広い場合、すべての被写体の連続性を満たす投影面の選択が困難なことが明らかとなった。今後は、境界領域において複数の奥行き位置に被写体が存在する場合への対応が課題である。

参考文献

- 1) 國田 豊, 上野雅浩, 田中敬二: 多層信頼度マップを用いた3次元映像の実時間生成システム, 映像情報メディア学会誌, Vol.60, No.7, pp.1102-1110 (2006).
- 2) 磯貝 愛, 國田 豊, 大谷佳光: 信頼度マッピング法を用いた全方位映像生成法, 映像メディア処理シンポジウム, Vol.12, pp.77-78 (2008).
- 3) T. Wada, T. Matsuyama: Appearance Sphere: Background Model for Pan-Tilt-Zoom Camera, Proc. IEEE ICPR, Vol.1, pp.718-722 (1996).
- 4) R.Szeliski: Video Mosaics for Virtual Environments, IEEE Computer Graphics and Application, Vol.16, No.2, pp.22-30 (1996).
- 5) 徐 剛, 辻 三郎: 3次元ビジョン, 共立出版(1998).