

被写体距離を考慮した Image-Based Rendering による 広域屋外環境のステレオ画像生成

堀 磨伊也[†] 神原 誠之[†] 横矢 直和[†]

[†] 奈良先端科学技術大学院大学 情報科学研究科

〒 630-0192 奈良県生駒市高山町 8916-5

E-mail: †{maiya-h,kanbara,yokoya}@is.naist.jp

あらまし 本報告では、広域屋外環境を自由経路で撮影された全方位動画画像から両眼ステレオ画像を生成する手法を提案する。両眼ステレオ画像生成は、撮影位置が計測された全方位画像群から Image-Based Rendering の一手法である光線空間法を用いて左右視点位置の画像を生成することにより行う。従来、生成画像において撮影位置が異なる画像を同時に使用したときに縦方向の視差を考慮しない場合には、生成画像において歪みが生じる問題があった。本手法では撮影位置が異なる同じ光線を持つ複数の画像からパターン類似度を用いて DP マッチングを行うことにより、簡易的な被写体距離を計算し、算出された距離を考慮したレンダリングを行うことにより歪みの軽減を行った。実験では、広域屋外環境の両眼ステレオ画像を生成した結果を示す。

キーワード Image-Based Rendering, ステレオ画像, 広域屋外環境, 全方位画像

Novel Stereoscopic View Generation by Image-Based Rendering Coordinated with Depth Information

Maiya HORI[†], Masayuki KANBARA[†], and Naokazu YOKOYA[†]

[†] Graduate School of Information Science, Nara Institute of Science and Technology

8916-5 Takayama, Ikoma, Nara, 630-0192 Japan

E-mail: †{maiya-h,kanbara,yokoya}@is.naist.jp

Abstract This paper describes a method of stereoscopic view generation by image-based rendering in wide outdoor environments. The stereoscopic view can be generated from an omni-directional image sequence with a light field rendering approach which generates a novel view image. The conventional methods of novel view generation have a problem such that the generated image is distorted because the image is composed of parts of several omni-directional images captured at different points. To overcome this problem, we have to consider the distances between the viewpoint and observed real objects in the rendering process. In the proposed method, in order to reduce the image distortion, a stereoscopic image is generated considering depth value estimated by dynamic programming (DP) matching using the images which are observed from different points and are captured a same ray in the real world. In experiments, the stereoscopic images in wide outdoor environments are generated and displayed.

Key words Image-Based Rendering, Stereoscopic Image, Outdoor Environment, Omni-directional Image

1. はじめに

遠隔地の情景をユーザに提示し、あたかもその場にいるのと同様の感覚を与える仮想現実感技術はテレプレゼンス [1] と呼ばれ、いかに遠隔地の情景を臨場感豊かに提示するかが課題となっている。遠隔地の情景を提示する手法として、多地点で撮影された画像から 3 次元復元によって仮想環境を構築・提示

する Model-Based Rendering (MBR) [2] と撮影環境を明示的に 3 次元復元することなく画像の変形・合成により仮想環境を構築・提示する Image-Based Rendering (IBR) [3] がある。

MBR では仮想空間内に存在する物体の 3 次元形状と反射特性など材質の光学特性を取得する必要があるが、屋外環境などの大規模なモデルの構築を自動化することは困難であった。一方 IBR では、光線空間内の光線の分布すべてを記録する必要があ

るが、形状や表面特性が複雑で MBR ではモデル化が困難な物体の場合にも実物体に近い見え方を再現することが可能であるため、広域屋外環境では IBR を用いた試みが多く行われている [4], [5]。IBR に基づく手法で遠隔地の情景をユーザに提示し、ユーザからの指示に従いインタラクティブに自由に視線方向を変更することができるシステム [6] が提案されている。このシステムでは、撮影に全方位型マルチカメラシステムを用いているため高解像度の画像を提示可能であるが、視点が撮影経路上に限定されるといった問題があった。

本研究では、屋外環境において撮影地点以外の視点映像を作成しユーザに両眼ステレオ画像を提示することにより立体視を可能とし、より臨場感の高いテレプレゼンスシステムの構築を目的とする。以下に撮影地点以外での視点映像を作り出す手法について概観する。文献 [7] は屋内環境において撮影された全方位画像から、撮影地点以外の視点映像を光線空間法 [8], [9] を用いて左右の視点位置の画像を生成する手法を提案している。しかし、全方位画像を曲面ミラーを用いた 1 台のカメラで撮影しているために解像度の不足や、視線方向による解像度の不均一などの問題があった。また想定する被写体距離が一定であり、撮影が等速直線移動で行わなければならないという制限があるため、屋外などの広域環境には適さなかった。一方、GPS を用いて撮影位置情報を取得することにより屋外環境において新視点画像を生成する手法 [10], [11] が提案されている。これらの手法では、複数のカメラの光軸が一点で交わるように円筒状に設置することによって全方位画像を取得しているために高解像度の画像を生成可能であるが、被写体距離の推定に Epipolar Plane Image (EPI) を用いているため撮影時にカメラ速度が一定である必要がある。また、自由経路で取得された全方位画像系列からの自由視点画像生成に対応しておらず、精度良く新視点画像を生成することは困難であった。

そこで本研究では、臨場感の高いテレプレゼンスシステムの構築を目的とし、広域屋外環境において自由経路で撮影された全方位動画と位置姿勢センサから得られる撮影位置・姿勢情報を用いて、ユーザが立体視を行いながら自由に視線変更を行うことのできるステレオ画像を生成する手法を提案する。従来、生成画像において被写体距離を考慮しない場合には縦視差の影響で歪みが生じることがあった [7]。提案手法は、従来法と同様に光線空間法を用いて生成するが、被写体距離を考慮してレンダリングを行うことにより生成画像の歪みの軽減を行う。被写体距離の推定はマルチベースラインステレオ [12] などの手法で高精度に行うことが可能である。一方、本手法では撮影位置が異なる同じ光線が撮影された 2 枚の画像においてパターン類似度を用いた DP マッチング [13], [14] によりエッジの対応を取り、センサから得られる撮影位置情報を用いることにより被写体距離の推定を行う。従来手法 [10], [11] では、撮影時にカメラの動作範囲に制限があったが、本手法では高精度な位置姿勢センサを搭載した車両を用いて撮影を行うことにより広域屋外環境を自由経路で撮影した画像を使用し、ステレオ画像を生成することが可能である。

以下、2 節では、提案するステレオ画像生成手法について詳

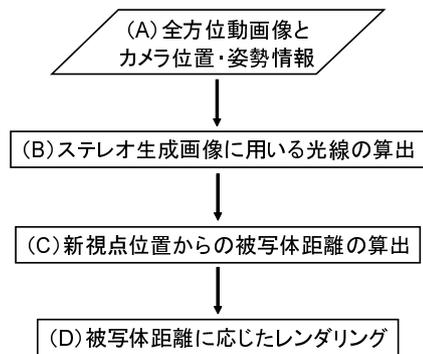


図 1 提案手法の手順



図 2 全方位カメラを搭載した撮影車両

述し、3 節では、提案手法を用いた実験として広域屋外環境の両眼ステレオ画像を生成した結果を示す。最後に 4 節で本稿のまとめと今後の課題について述べる。

2. 自由経路で撮影された全方位動画からの両眼ステレオ画像生成

2.1 提案手法の概要

本節では、屋外環境において全方位カメラ 2 台を眼間距離離して車両に固定し、移動しながら撮影した全方位画像系列を用い、撮影点以外でのステレオ画像を生成する手法について述べる。本手法の処理手順を図 1 に示す。

まず屋外環境で眼間距離離れた地点で撮影された全方位画像列と撮影時のカメラの位置・姿勢情報を取得する (A)。続いて、生成する新視点位置の計算を行い、生成に必要な光線を求め (B)、DP マッチングを行って被写体距離の推定を行う (C)。最後に被写体距離に応じてレンダリングを行うことにより (D)、縦視差の影響による画像の歪みが少ないステレオ画像を生成する。以下に各処理について詳細に述べる。

2.2 全方位画像とカメラ位置・姿勢情報の取得

本研究は広域屋外環境のテレプレゼンスを行うことを目的としているため、1 度の撮影で多数の画像を取得し、ステレオ画像生成に必要な光線情報を得る必要がある。提案手法では、全方位カメラ 2 台を図 2 に示すように眼間距離離して左右に固定し、移動しながら撮影を行うことにより図 3 に示すような撮影位置で全方位画像を取得する。2 台の全方位カメラを用いることにより、1 台のカメラで撮影した場合に困難であった撮影経路に平行な視線方向のステレオ画像生成を可能とする。また後述する被写体距離算出の際に必要な、撮影位置が異なりかつ同じ光線を撮影した画像の取得が可能となる。なお、それぞれのカメラの位置・姿勢情報は後述するセンサによって撮影時に同時に取得されているものとする。

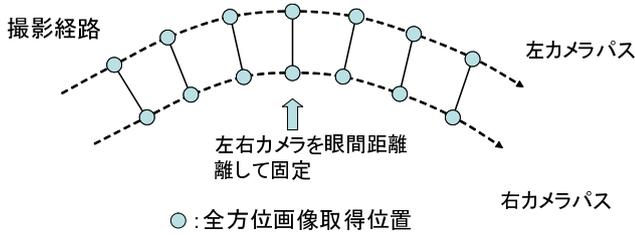


図3 全方位画像の取得位置（上から見た図）

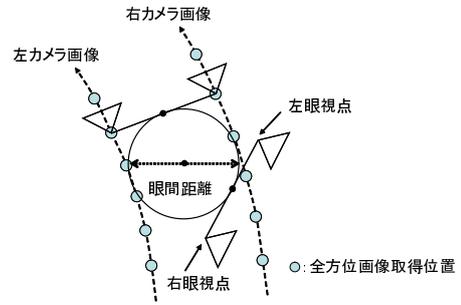


図4 2経路の全方位画像群からのステレオ画像生成（上から見た図）

2.3 画像生成に用いる光線の算出

2.3.1 生成画像の視点位置算出

本研究では、自由な見回しと立体視を可能とするために、図4のように左右眼間距離を直径とする円周上を左右の眼間中心が回転したときのステレオ画像を生成する。図4のように視点位置を定めることにより、進行方向のステレオ画像は、2台の全方位カメラの撮影画像の一部をそのまま利用することで生成可能である。また視点が変化した場合、センサにより計測された撮影時のカメラ位置情報を用いて、画像生成を行う視点位置の計算を行う。

2.3.2 新視点画像生成に用いる光線の選択

新視点画像生成に用いる光線を算出する手法を説明する。ここでは撮影された全方位動画の各フレームにおいて、撮影されたカメラの位置・姿勢情報が既知であるとする。本手法は、光が伝搬により減衰の変化がないとすると、新視点位置と生成する画像面が図5のような位置関係にあるときに光線APと光線QPは同じとする光線空間法の原理[8]に基づいている。例えば新視点Aでの画像上の点Pの画素は撮影経路上の点Qからの光線情報を用いて生成することができる。新視点を撮影経路周辺の光線が多く存在する場所に限定した場合、新視点画像を複数の全方位画像を用いて擬似的に生成することができる。しかし、全方位画像は移動経路上で離散的に存在しているため、光線と経路の接点に全方位画像が存在しない場合がある。その場合には、生成に必要な光線との距離が最も近い位置で撮影された全方位画像から生成に必要な光線と平行な光線情報を用いて画像生成を行う。生成時には、生成画像を幅1画素の縦方向のラインに分割し、それぞれの縦ラインにおいて生成に必要な光線に最も近い地点の全方位画像を用いて新視点画像の生成を行うことにより、図5に示すような縦の短冊に分割して画像を生成可能である。

本手法は、全方位カメラを2台用いて撮影を行っているため、画像生成に必要な光線が左右のカメラで撮影されたそれぞれの全方位画像に存在する場合がある。この場合は、新視点位置から近い位置で撮影された全方位画像を用いる。また左右のカメラを近接して固定しているため撮影された全方位画像にはもう一方のカメラ本体によるオクルージョンが発生する場所が存在する。この場合は、カメラ本体によるオクルージョンが発生していないもう一方のカメラで撮影された全方位画像を用いる。

2.4 新視点位置からの被写体距離の算出

上述の手法により生成された画像において、新視点位置からの被写体距離を一定としてレンダリングを行った場合には生成

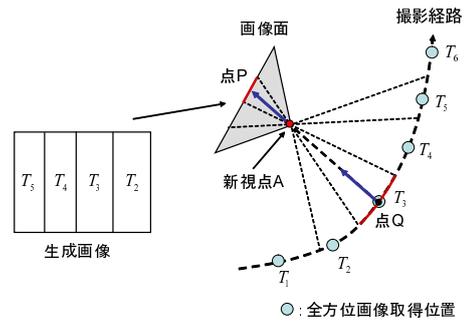


図5 新視点画像生成（上から見た図）

画像の隣り合う短冊間において縦方向の歪みが生じることがある。生成画像において歪みが発生した例を図6に示す。これは撮影位置と画像を生成する視点が異なるために生成画像に縦視差が生じるためである。そこで本手法では、被写体距離に応じたレンダリングを行うことにより歪みを軽減した画像を生成する。本手法では、撮影位置が異なり、同じ光線情報を持つ2枚の画像においてパターン類似度を用いたDPマッチングによりエッジの対応をとり、センサから得られる撮影位置情報を用いることにより被写体距離の推定を行う。

本手法では、画像取得は2台の全方位カメラを固定して行うために図7に示すように生成画像において生成に必要な光線が左右2台のカメラ画像に複数存在することになる。この2枚の画像からエッジを抽出し、以下の式(1)に示すDPマッチングにより各エッジの対応をとる。

$$g(L_i, R_j) = \min \begin{cases} g(L_{i-1}, R_j) + d(L_i, R_j) \\ g(L_{i-1}, R_{j-1}) + d(L_i, R_j) \\ g(L_i, R_{j-1}) + d(L_i, R_j) \end{cases}, \quad (1)$$

ここで $L_i (i = 1 \sim I)$ は左カメラで撮影された画像から抽出された I 本のエッジを、 $R_j (j = 1 \sim J)$ は右カメラで撮影された画像から抽出された J 本のエッジを表す。左カメラ画像と右カメラ画像間のエッジの特徴ベクトル間距離を $d(L_i, R_j)$ とし、その部分積 $g(L_i, R_j)$ を最小にする経路を算出することによって両エッジのマッチングを行う。なお特徴ベクトル間距離 $d(L_i, R_j)$ は両カメラ画像の画素同士の Sum of Squared Differences (SSD) を用いる。ただし DP マッチングの特徴としてマッチングを行う際に1本のエッジに複数のエッジが対応することがあるが、その場合は複数のエッジの中で SSD の値が最も小さいものを対応



図 6 生成画像における歪みの例

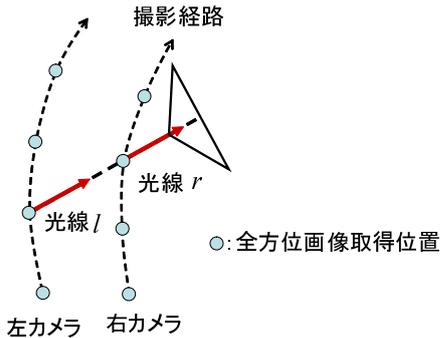


図 7 DP マッチングの際に使用する同じ光線情報を持つ画像（上から見た図）

エッジとする．SSD の計算はウィンドウサイズを縦 25 画素、横 25 画素で行った．なお、使用する全方位画像にカメラ本体のオクルージョンが発生した場合には上述の手法は適応できないが、オクルージョンが発生する光線方向では、歪みの影響が少ないために大きな問題は生じない．

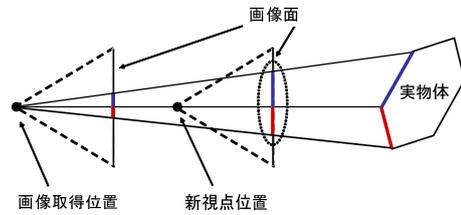
2.5 被写体距離を考慮したレンダリング

上述の手法で算出される被写体距離はエッジ上のみである．したがってエッジ以外の部分に関しては、算出されたエッジの被写体距離を用いて線形補間で算出する．図 8 に従来手法 [7] と提案手法のレンダリングの過程を示す．従来手法では被写体距離を考慮しないで撮影位置情報のみを用いてレンダリングを行っているために、実物体が画像中で正しいスケールで表現できない問題が生じる．一方、提案手法は新視点位置からの被写体距離を考慮してレンダリングを行っているために正しい大きさで表現され、生成画像の歪みを軽減することができる．

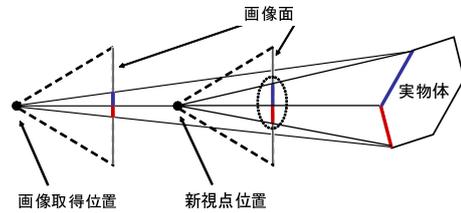
3. 実 験

屋外環境において曲線経路を含む自由経路で撮影した全方位動画を用いてステレオ画像を生成し、自由な見回しと立体視が可能であるか検証を行った．

実験では全方位型マルチカメラシステム 2 台と位置姿勢センサを搭載した車両（図 2）を自由経路で移動させながら全方位動画と撮影位置・姿勢情報を取得した．全方位型マルチカメラシステムは、Point Grey Research 社製の Ladybug2 を用い、図 2 に示すように 2 台を一定距離離して固定しながら撮影を行った．



(a) 従来手法



(b) 提案手法

図 8 レンダリング過程

Ladybug2 は横 768 画素、縦 1024 画素の画像を水平方向の 5 台と真上の 1 台の計 6 台の CCD カメラから同期して取得し、上方を含む全方向の約 75% 以上を 30fps の動画として取得することが可能である．カメラの位置・姿勢計測にはリアルタイムキネマティック GPS (RTK-GPS) と小型慣性航法装置 (INS) を統合したハイブリッドセンサを用いた．RTK-GPS は Nikon-Trimble 社製 LogPakII を用い、INS は Tokimec 社製 TISS-5-40 を用いた．このハイブリッドセンサは GPS の欠点である計測周期の低さと INS の欠点である計測誤差の蓄積をお互いに補うことにより表 1 に示すように高精度で高レートな計測を行うことが可能である [15]．計算機への入力画像として図 9 に示すような解像度が横 2048 画素、縦 1024 画素の全方位パノラマ画像を用いた．なお撮影は図 10 に示す曲線を含む自由経路で行った．

被写体距離の推定に用いた画像を図 11 に示す．図 11 の画像においてそれぞれの中央の縦ラインを通る光線情報を用いて被写体距離の推定を行った結果を図 12 に示す．図 12 に示されるように本手法で実環境に近い被写体距離が算出されていることを確認した．被写体距離に応じてレンダリングを行うことにより、従来生じていた短冊間の歪みが軽減されたことが図 13 より確認できる．なおステレオ画像生成は、PC (Pentium D 3.0GHz、メモリ 3.0GB) を使用した場合、1 枚の画像に約 5.8 秒を要した．生成されたステレオ画像の例を図 14 に示す．それぞれの画像の解像度は横 1024、縦 768 画素である．図 14 において、近い位置に存在する物体はそれよりも遠くにある物体と比較して視差が大きくなっており、立体視が可能である．オフラインで生成・蓄積された画像を提示する際には立体視を可能とし、

表 1 ハイブリッドセンサの性能

位置計測	
経度	± 3.3cm (± 3.0cm)
緯度	± 3.3cm (± 3.0cm)
計測レート	50Hz (1Hz)

() は RTK-GPS 単体使用時



(a) 左カメラ画像



(b) 右カメラ画像

図9 全方位パノラマ画像

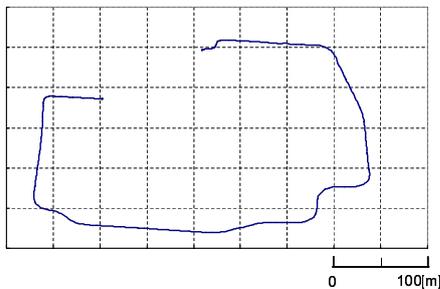


図10 撮影経路

ユーザの指示に従いインタラクティブに移動しながら自由な見回しが可能なことを確認した。

4. ま と め

本報告では、広域屋外環境において曲線経路を含む自由経路で撮影された全方位動画と RTK-GPS と INS を用いたハイブリッドセンサから得られる撮影位置・姿勢情報を用いて、任意方向のステレオ画像を生成する手法を提案した。従来手法 [7] では、全方位動画の撮影位置から物体までの仮定距離と実際の物体までの距離とのずれにより生成された両眼ステレオ画像に縦方向の歪みが生じた。そこで画像取得位置から物体までの距離を、同じ光線情報を持つ 2 枚の画像から算出し、新視点からの距離を考慮してレンダリングを行うことにより、生成画像における短冊間で歪みの軽減を行った。実験ではユーザにステレオ画像を提示することにより自由な見回しと立体視が可能であることを確認した。

提案手法では、両眼ステレオ画像生成に用いる画像は静的な環境に限定される問題があった。しかし広域屋外環境をテレプレゼンスの対象にした場合、車両や人間といった移動物体が存在する場合を考慮しなければならない。今後は、動的環境下で



(a) 左カメラ画像

(b) 右カメラ画像

図11 被写体距離推定に用いた画像

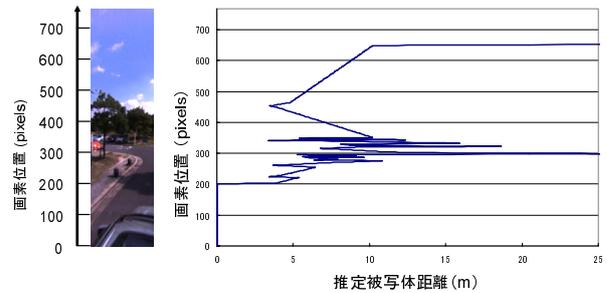


図12 被写体距離推定結果



(a) 被写体距離を一定としてレンダリングを行った画像

(b) 被写体距離を考慮してレンダリングを行った画像

図13 生成画像における歪みの軽減例

撮影された全方位画像からのステレオ画像生成が課題として挙げられる。

文 献

- [1] S. Moezzi, Ed.: Special Issue on Immersive Telepresence, IEEE MultiMedia, Vol.4, No.1, pp. 17-56, 1997.
- [2] T. Asai, M. Kanbara, and N. Yokoya: "3D Modeling of Outdoor Environments by Integrating Omnidirectional Range and Color Images," Proc. Int. Conf. on 3-D Digital Imaging and Modeling (3DIM), pp. 447-454, 2005.
- [3] E. Chen: "QuickTime VR -An Image-Based Approach to Virtual Environment Navigation," Proc. SIGGRAPH'95, ACM, pp. 29-38, 1995.
- [4] E. H. Adelson and J. R. Bergen: "The Plenoptic Function and the Elements of Early Vision," in M.Landy and J.Movshon., ed., Computer Models of Visual Processing, pp. 3-20, MIT Press. 1991.
- [5] L. McMillan and J. Bergen: "Plenoptic Modeling: An Image-Based Rendering System," Proc. SIGGRAPH'95, pp. 39-46, 1995.
- [6] S. Ikeda, T. Sato, M. Kanbara, and N. Yokoya: "Immersive Telepresence System with a Locomotion Interface Using High-Resolution Omnidirectional Videos," Proc. IAPR Conf. on Machine Vision Applications, pp. 602-605, 2005.



(i) 進行方向に対して左に 45 度回転した視線画像



(ii) 進行方向に対して左に 90 度回転した視線画像



(iii) 進行方向に対して左に 135 度回転した視線画像

(a) 左眼画像

(b) 右眼画像

図 14 ステレオ画像生成例

- [7] 山口, 山澤, 竹村, 横矢: “ 全方位動画像を用いた両眼ステレオ画像の実時間生成によるテレプレゼンス ”, 信学論 (D-II), Vol. J84-D-II, No. 9, pp. 2048-2057, 2001.
- [8] M. Levoy and P. Hanrahan: “ Light Field Rendering, ”Proc. SIGGRAPH'96, pp. 31-42, ACM, 1996.
- [9] S. Gortler, R. Grzeszczuk, R. Szeliski and M. Cohen: “ The Lumigraph, ”Proc. SIGGRAPH'96, pp. 43-54, ACM, 1996.
- [10] S. Ono, K. Ogawara, M. Kagesawa, H. Kawasaki, M. Onuki, K. Honda and K. Ikeuchi: “ Driving View Simulation Synthesizing Virtual Geometry and Real Images in an Experimental Mixed-Reality Traffic Space, ”Int. Sympo. on Mixed and Augmented Reality, pp. 214-215, 2005.
- [11] 高橋, 川崎, 池内, 坂内: “ 全方位画像を用いた広域環境の自由視点レンダリング ”, コンピュータビジョンとイメージメディア研究会 (CVIM), Vol. 42, No.SIG13(CVIM3), pp. 99-109, 2001.
- [12] 佐藤, 神原, 横矢, 竹村: “ マルチベースラインステレオ法を利用した動画像からの屋外環境の三次元モデル化 ”日本バーチャルリアリティ学会論文誌, Vol. 7, No. 2, pp. 275-282, 2002.
- [13] 大田, 山田: “ 動的計画法によるパターンマッチング ”, 情処論, Vol. 30, No. 9, pp. 1058-1066, 1989.
- [14] 磯道, 小川: “ 動的計画法によるパターン・マッチング ”, 情処論, Vol. 16, No. 1, pp. 15-22, 1975.
- [15] 神原, 横矢: “ RTK-GPS と慣性航法装置を併用したハイブリッドセンサによる屋外型拡張現実感システム ”, 画像の認識・理解シンポジウム (MIRU), pp. 933-938, 2005.