

複数経路で撮影された全方位動画像を用いた全周ステレオ画像生成

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 遠隔地の情景をあたかもその場にいるかのように提示する仮想現実感技術はテレプレゼンスと呼ばれ、臨場感を向上させることが要求されている。本報告では、臨場感を向上させるためのテレプレゼンス技術として立体視と自由な視線変更に注目し、それらを同時に実現するために全周ステレオ画像を生成する手法を提案する。従来、image-based rendering (IBR) により両眼ステレオ画像を生成可能であるが、全周のステレオ画像を生成するためには大量の光線情報が必要となるため、生成可能な視点位置が限定される問題があった。本研究では、複数経路で撮影された全方位動画像を入力とすることで、光線情報をより密に取得し、生成に必要な光線情報を光線の尤もらしさに基づくペナルティ関数を用いて適切に選択することにより、広範囲で全周ステレオ画像の生成を行う。実験では屋外環境の任意経路で撮影された実画像を用いて全周の両眼ステレオ画像の生成を行い、立体視と自由な視線変更が行えることを確認した。

キーワード テレプレゼンス, 立体視, ステレオ画像生成, image-based rendering

Omnidirectional Stereoscopic Image Generation Using Multiple Omnidirectional Image Sequences

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

[†] 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 proposes a novel way of making omnidirectional stereoscopic images for binocular vision and realization of arbitrary view. The stereoscopic images can be generated from omnidirectional image sequences by image-based rendering. In order to generate the novel stereoscopic view accurately by image-based rendering, much ray information is required because the novel view is composed of parts of several rays captured at different points. In our research, much ray information is acquired at many points by capturing omnidirectional image sequences in arbitrary multiple paths and appropriate ray information is used for generation of omnidirectional stereoscopic images. In experiments, omnidirectional stereoscopic images in outdoor environments were generated and displayed.

Key words telepresence, binocular vision, stereoscopic image generation, image-based rendering

1. はじめに

遠隔地の情景を利用者に提示し、あたかもその場にいるような没入感を与える仮想現実感技術はテレプレゼンス [1] と呼ばれ、盛んに研究が行われている。テレプレゼンスは、仮想旅行やゲームなどのエンターテインメント分野だけでなく、教育や医療分野などで応用が期待されている。テレプレゼンスにおいて遠隔地の情景を提示する手法の一つとして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提示する image-based rendering [2], [3] (以下, IBR) がある。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、形状が複雑な物体でも実物体と近い見え方を再現することが可能であるため、広域屋外環境では写実性の高い IBR を用いた試みが多く行われている [4], [5]。テレプレゼンスでは遠隔地の情景を臨場感豊かに提示することが重要となるため、写実性や没入感が高いこと、視線変更による自由な見回しが行えること、両眼立体視が可能であることなどが望まれる。

1 台のカメラから撮影された単眼画像を提示する場合、カメラを回転させることにより自由な見回しが可能であるが、カメラを回転させることにより画像提示を行うため、テレプレゼンスを行う際に時間遅延が発生する問題がある。遅延のない見回しを行う方法として、周囲 360

度の情景を一度に全方位カメラを用いて取得する方法 [6] がある。撮影された全方位画像から視線方向に応じた画像を提示することにより、ユーザは遅延なく見回しを行うことが可能である。しかし、撮影地点からの単眼画像しか表示できないために映像に奥行き感がなく、臨場感が低い問題がある。

一方、ステレオカメラを回転させながら撮影を行うことやステレオ画像を生成可能な特殊なカメラを用いること [7] により、ユーザは立体視を行うことが可能となるが、撮影地点でしか立体視が行えないため、ユーザが移動した場合に様々な視点位置で立体視を行うためには、膨大なデータを蓄積する必要がある。

そこで IBR を用いて撮影地点以外の視点画像を生成する手法 [8] を両眼の視点位置に適用することによりステレオ画像を生成し、立体視を行うことで臨場感を高める研究 [9] ~ [12] が行われている。しかし手法 [9] では、入力となる全方位画像を取得するカメラの撮影経路に制約があり、広域な屋外環境でステレオ画像を生成することが困難であった。手法 [10] は入力となる全方位画像を取得するカメラの撮影経路に制約はないが、生成可能な視点位置が画像撮影地点に限定される問題があった。手法 [11] は structure-from-motion 法によりカメラ位置・姿勢を推定することで屋外環境においてステレオ画像の生成を行う手法を提案しているが、生成される視点位置・視線方向が限定されていた。一方、手法 [12] は、高精度の位置姿勢センサを用いることにより屋外環境の任意経路で撮影された全方位画像から IBR を用いて全周のステレオ画像の生成を行っている。この手法では撮影経路上を移動しながら自由な見回しと立体視を目指しているが、これらを同時に実現するためには多数の光線情報が必要であり、入力となる全方位画像を眼間距離よりも短い間隔で取得する必要があった。そのため広域な屋外環境を対象とした場合に、移動車両などを用いて撮影することは現実的ではなかった。

本稿では、広域屋外環境で車両などを用いて移動撮影する場合に、同様の経路を複数回にわたり移動撮影することにより取得した全方位画像群を用いてステレオ画像を生成する手法を提案する。提案手法は従来手法 [12] と同様に、全方位画像群とそれらと同時に取得されたカメラ位置・姿勢情報を入力とし、自由な見回しと立体視を同時に実現するために IBR を用いて全周のステレオ画像を生成する。手法 [12] では視点位置が撮影経路上を移動した場合に、撮影時の画像取得間隔を短く設定する必要があったが、提案手法では、複数経路で撮影された全方位動画をとして、光線情報を密に取得し画像取得時の移動速度の制限を排除する。また生成時には取得した全方位画像群から選択される光線の尤もらしさに基づくペナルティ関数を用いることで生成に必要な光線情報を適切に選択することにより、広範囲で全周ステレオ画像の生成を行う。

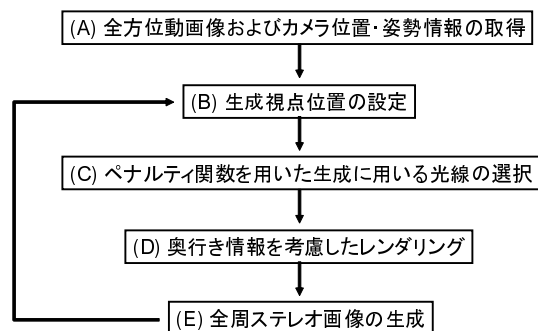


図 1 提案手法の処理手順

以下、2 節では、提案する複数の撮影経路で取得された全方位動画群から両眼ステレオ画像を生成する手法について詳述し、3 節では、提案手法を用いた実験として、屋外環境で取得された全方位動画群から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. 複数の任意経路で撮影された全方位画像群からのステレオ画像生成

2.1 両眼ステレオ画像生成手法の概要

本研究では複数の任意経路で撮影された全方位動画をを用いて、撮影された周辺領域において全周の両眼ステレオ画像を生成する手法を提案する。本項では IBR を用いて両眼ステレオ画像を生成する手法の概要を述べる。提案手法の処理手順を図 1 に示す。

まず屋外環境の複数の任意経路において全方位動画およびカメラ位置・姿勢情報の取得 (A) を行う。次にステレオ画像の生成を行う視点位置の設定 (B) および立体視を行うための光線の尤もらしさに基づくペナルティ関数を用いて生成に用いる光線の選択 (C) を行う。最後に被写体までの奥行き情報の推定を行い、奥行き情報を考慮したレンダリングを行う (D) ことにより、両眼ステレオ画像の生成を行う。次項以降にそれぞれの処理の詳細を述べる。

2.2 全方位動画およびカメラ位置・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、ステレオ画像生成に必要な光線情報を大量に得る必要がある。提案手法では、ステレオ画像生成に必要な光線情報を得るために全方位カメラを車両に固定して移動しながら撮影を行う。全方位カメラを用いることにより一度の撮影により全方向の光線情報を取得可能である。入力となる全方位画像の例を図 2 に示す。車両を移動しながら撮影する際に同様の経路を複数回にわたり移動撮影することにより、より多くの光線情報を取得することが可能である。ただし本研究では IBR を用いてステレオ画像を生成するため、入力画像において移動物体が存在しないという仮定と画像間で照明条件



図 2 入力となる全方位画像の例

が大きく変化しないという仮定が必要である．なお，ステレオ画像を生成する際に必要となる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る．

2.3 生成視点位置の設定およびペナルティ関数を用いたステレオ画像生成に用いる光線の選択

全方位画像が撮影された領域周辺に左右の視点位置を定め，それぞれの視点位置において IBR を用いて撮影されていない視点位置の画像の生成を行う．本研究では図 3 のように，左右の視点の中心が，眼間距離を直径とする円周上を回転した場合のステレオ画像を生成する．設定する視点位置は，ユーザが視線方向を回転させたときの視点位置を近似している．全方位画像が撮影された領域周辺において回転中心を移動させることにより広範囲において全方位のステレオ画像の生成を可能とする．

図 3 の例のように左右それぞれの視点画像が撮影時の画像群に存在しないために，撮影された画像群から生成を行う必要がある．従来手法 [12] では，撮影時の画像を見回し時の前方方向の視点画像として用いたため，生成する視点位置があらかじめ撮影された画像の撮影地点に依存する問題があった．また光線情報の不足のためステレオ画像の生成が困難な視点位置が存在した．本手法では，異なる経路で撮影された全方位画像群を用いることにより光線情報の不足を補い，撮影された画像群から適切な光線情報を選択することにより両眼ステレオ画像の生成を行う．これにより従来手法よりも広範囲で両眼ステレオ画像の生成が可能である．

本研究では撮影されていない位置の視点画像を生成するために IBR の一手法である light field rendering に基づいた生成に必要な光線の選択を行う．この手法は図 4 に示すように，光が伝搬により減衰がないという仮定に基づき，画像を生成する視点位置と生成する画像面が図 4 のような位置関係にあるときに画像が取得された位置を通る光線と画像を生成する視点位置を通る光線が同じ情報を持つという原理である．この原理に基づき新視点の画像を撮影された複数の全方位画像から再構成する．

本手法では，生成を行う際に生成画像を幅 1 画素の縦のスリットに分割し，それぞれのスリットごとに生成に用いる画像の選択を行う．生成に用いる画像は精度の良

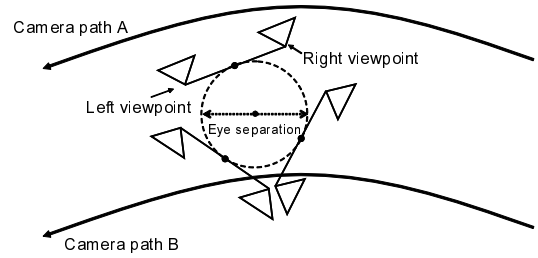


図 3 ステレオ画像を生成する視点位置の例

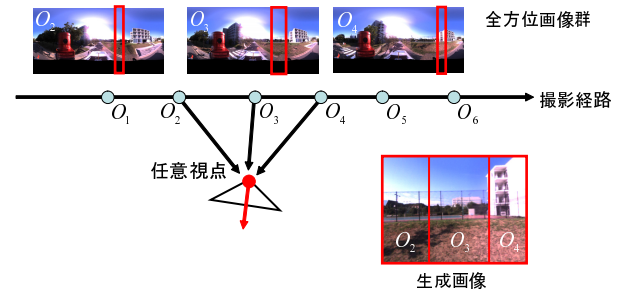


図 4 light field rendering の原理

い立体視を実現するために光線の尤もらしさに基づくペナルティ関数を定義し，そのペナルティ関数が最も小さくなるように画像を選択し生成を行う．以下にそれぞれの手法の詳細を述べる．

本手法では light field rendering の原理に基づき画像の生成を行うが，画像取得位置が離散的な位置であるため，生成に必要な光線が存在しない場合または複数存在する場合がある．本研究ではステレオ画像の生成に用いる光線の選択手法として，以下の式 (1) に示す光線の尤もらしさに基づくペナルティ関数を用いることにより適切な光線情報の選択を行い，視点画像の生成精度の向上を図る．

$$P = P_{ang} + wP_{pos}, \quad (1)$$

このペナルティ関数は，次に示すパラメータを持つ．

(1) 生成に必要な光線と生成に用いる画像が取得された位置との最短距離を表すペナルティ P_{ang}

(2) 画像の生成を行う視点位置と生成に用いる画像が取得された位置との距離を表すペナルティ P_{pos}

ここでペナルティ P_{ang} は，光線情報が存在しない場合に可能な限り生成に必要な光線と近い光線を選択することで生成精度の向上を図る目的がある．ペナルティ P_{pos} は，画像を生成する視点位置と近い位置で撮影された全方位画像を用いることで生成画像の解像度の均一化を図るとともに光学的な整合性を保つ目的がある．この 2 つのペナルティを用いることにより，生成画像においてスリット間の不連続を回避しつつ，高精度なステレオ画像生成を試みる． w は各ペナルティの大きさを決定する重みを表す．

各視点画像の各スリットにおいてこのペナルティ関数の値 P を最小とする光線情報を持つ全方位画像を生成に



図 5 被写体までの距離を一定と仮定して生成した視点画像

用いることにより，両眼ステレオ画像の生成を行い，立体視を実現する．

2.4 奥行き情報を考慮したレンダリング

前項の手法により選択された光線を用いて視点画像を生成する際に，視点位置からの被写体までの距離を一定としてレンダリングを行った場合には，図 5 に示すような生成画像の隣り合う短冊間において歪みが生じることがある．これは，各スリット間で生成に用いる全方位画像の取得位置が異なるために生じる視差の影響である．

そこで本手法では，奥行き情報を考慮し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，図 4 で示した生成画像の隣り合うスリット間で観測される対象物体の奥行きを推定を行う．生成画像における奥行き推定を行う位置の例を図 6 に示す．生成に用いる画像の取得位置が隣り合うスリット間で変化するのみ縦のスリットで奥行き情報の推定を行う．図 6 の例では O_1, O_2, O_3, O_4, O_5 の 5 地点で撮影された全方位画像を用いて新視点画像は構成され，生成に用いる画像が異なる 4 箇所のスリット間で奥行き推定を行う．奥行き推定を行う位置は縦のスリット上でエッジが検出された位置のみとし，その他の位置は算出された奥行き情報を用いて線形補間により算出する．歪みの影響が最も大きい画像の取得位置が変化するスリット位置で奥行きを推定を行うことにより写実性を保ちつつ生成画像における歪みの影響を軽減する．

奥行き推定は，撮影位置が異なり，同等な光線情報を持つ 2 枚の画像においてパターン類似度を用いたステレオマッチングにより同等の光線情報を持つ画素を対応付け，センサから得られる撮影位置情報を用いることにより行う．撮影位置が異なり同等な光線情報を持つ画像として，前節で定義したペナルティ関数 P の値が最も小さい地点の画像とその次に小さい地点の画像を用いる．センサによりあらかじめ計測される画像取得位置を用いる

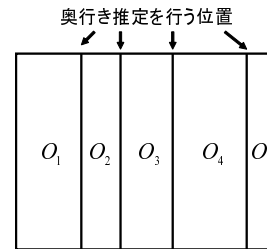


図 6 生成画像における奥行き推定を行う位置

ことにより，生成に用いる光線が画像上に投影された点の対応点は，もう一方の画像のエピポーラ線上に限定される．このエピポーラ線上でパターン類似度としてとして式 (2) に示す正規化相互相関を用いたマッチングを行うことにより対応点を求める．

$$R_{NCC} = \frac{\sum_{j=0}^{N-1} \sum_{i=0}^{N-1} I(i, j)T(i, j)}{\sqrt{\sum_{j=0}^{N-1} \sum_{i=0}^{N-1} I(i, j)^2 \times \sum_{j=0}^{N-1} \sum_{i=0}^{N-1} T(i, j)^2}} \quad (2)$$

ここで，生成に用いる光線が通過する画素位置を中心とするテンプレートの大きさを $M \times N$ ，テンプレートの位置 (i, j) における画素値を $T(i, j)$ ，テンプレートと重ね合わせた対象画像の画素値を $I(i, j)$ とする．エピポーラ線上でこの R_{NCC} の値が最も高い画素位置を対応点とし，この対応点とそれぞれの画像が撮影されたカメラ位置から奥行き情報の推定を行う．

推定された奥行き情報を用いて，生成する視点位置から対象物体を通る光線が，画像面に投影される位置を算出する奥行き情報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．

3. 実画像を用いた画像生成実験

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画像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．実画像を用いた実験として本研究では車両に搭載された全方位カメラを用いて動画画像の取得を行った．実験では全方位マルチカメラシステムと位置姿勢センサを搭載した車両を任意経路で複数回移動させながら全方位動画画像と撮影位置・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 を使い，車両に固定しながら撮影を行った．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 の動画画像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 を使い，INS

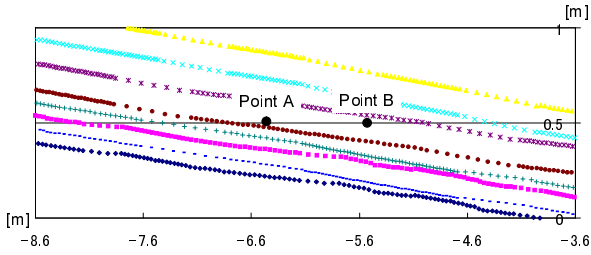


図 7 全方位画像の取得位置

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高精度で高レートな計測 (位置誤差 3.3cm, 計測レート 50Hz) を行うことが可能である [13]。計算機への入力画像として図 2 に示したような解像度が横 2048 画素, 縦 1024 画素の全方位パノラマ画像を用いた。なお, 撮影は屋外環境の曲線を含む任意経路で複数回行った。ハイブリッドセンサにより得られた全方位画像の取得経路の一部を図 7 に示す。複数の撮影経路で撮影することにより画像取得位置が広域になり光線が密に取得されていることが分かる。

本手法では生成画像の各スリットにおいてペナルティ関数 P を最小値とする位置の画像を用いて生成を行うが, 生成された画像における各スリットで算出されるペナルティ値の合計値 P_{sum} をその画像の生成精度の指標として算出する。

$$P_{sum} = \sum_{slit=0}^{width} P_{slit}, \quad (3)$$

ここで P_{slit} は各スリットでのペナルティ関数の最小値, $width$ はスリット数をそれぞれ表す。視線方向をある一方向に固定し, ステレオ画像を生成する場合に算出される各視点位置でのペナルティ合計値 P_{sum} をプロットした様子を図 8 に示す。生成位置周辺で光線情報が密に取得されている場合にはペナルティ合計値 P_{sum} が少ないことがわかる。それとは対照的に, 撮影経路の外の領域では光線情報が疎にしか取得されていないために, ペナルティ合計値 P_{sum} が高くなっている様子がわかる。本実験ではこのペナルティ合計値 P_{sum} が少ない場合には精度の良いステレオ画像の生成が可能だという仮定の下, P_{sum} がある一定値以下の領域で両眼ステレオ画像の生成を行う。

ペナルティ合計値 P_{sum} が少ない地点の例として地点 A および地点 B で両眼ステレオ画像を生成した結果例を図 9, 10 に示す。図 9 は, 地点 A において視線方向を 90 度ずつ変更して自由な見回しを行った結果である。また地点 B から地点 A に移動を行った際のステレオ画像の生成例を図 10 に示す。図 10 のステレオ画像生成例では視線方向を同方向に固定して移動を行っているため, 近

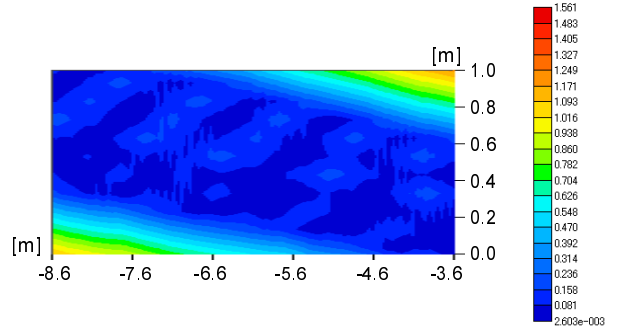


図 8 各視点位置におけるペナルティ合計値の分布

景を観察すると移動している様子が分かる。地点 A および地点 B でのステレオ画像生成例からペナルティ合計値 P_{sum} が少ない領域において各ステレオ画像を用いて立体視が可能であることを確認した。

4. まとめと今後の課題

本稿では, 屋外環境において任意経路で複数回にわたり入力画像の取得を行い, 取得した全方位画像群から IBR を用いて両眼ステレオ画像を生成する手法を提案した。実験において複数の任意経路で撮影された全方位画像群を用いることにより, 生成に用いることができる光線情報が増大し, 光線の尤もらしさに基づくペナルティ関数を用いて適切な光線情報を生成に用いることで立体視と自由な視線変更が可能なステレオ画像の生成を行った。また光線情報の増加に伴い, 広範囲な視点位置でステレオ画像が生成できることを確認した。

提案手法では, 異なる時刻に撮影された全方位画像を複数用いてステレオ画像を生成するために, 撮影された画像内に人や車両などの移動物体が存在する場合に生成したステレオ画像に違和感が生じる問題があった。そこで複数の地点で撮影された全方位画像から移動物体が存在しない領域を用いてステレオ画像を生成することが今後の課題として挙げられる。

謝辞 本研究の一部は科学研究費補助金 (基盤研究 (A), No.19200016) による。

文 献

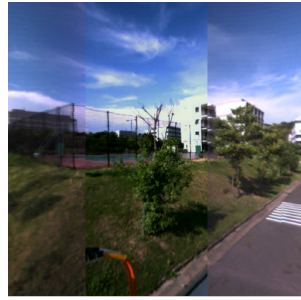
- [1] S. Moezzi, Ed.: Special Issue on Immersive Telepresence, IEEE MultiMedia, Vol. 4, No. 1, pp. 17–56, 1997.
- [2] L. McMillan and J. Bergen: “Plenoptic Modeling: An Image-Based Rendering System,” ACM Transactions on Graphics (SIGGRAPH 95), pp. 39–46, 1995.
- [3] J. I. Park and S. Inoue: “Image-Based Rendering from Multi-View Images,” Journal of ITE, Vol. 52, No.3, pp. 371–376, 1998.
- [4] N. Snavely, R. Garg, S. M. Seitz, and R. Szeliski: “Finding Paths through the World’s Photos,” ACM Transactions on Graphics (SIGGRAPH 2008), Vol. 27, No. 3, pp. 11–21, 2008.
- [5] G. Haan, J. Scheuer, R. Vries, F. Post: “Egocentric Navigation for Video Surveillance in 3D Virtual



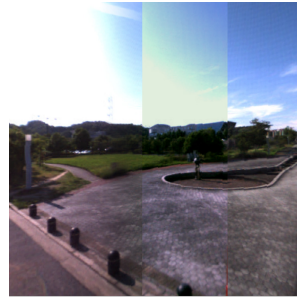
(i) 地点 A でのステレオ画像



(ii) (i) に対して 90 度左回転した方向のステレオ画像



(iii) (i) に対して 180 度左回転した方向のステレオ画像



(iv) (i) に対して 270 度左回転した方向のステレオ画像

図 9 見回しを行った場合の両眼ステレオ画像例 (左: 左目画像, 右: 右目画像)



(i) 地点 A でのステレオ画像



(ii) 地点 B でのステレオ画像 ((i) と同方向)

図 10 移動を行った際の両眼ステレオ画像例 (左: 左目画像, 右: 右目画像)

Environments,” IEEE Sympo. on 3D User Interfaces (3DUI), pp. 103-110, 2009.

[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.

[7] S. Peleg, M. Ben-Ezra, Y. Pritch: “Omnistereo: Panoramic Stereo Imaging,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 3, pp. 279-290, 2001.

[8] M. Levoy and P. Hanrahan: “Light Field Rendering,” ACM Transactions on Graphics (SIGGRAPH 96), pp. 31-42, 1996.

[9] K. Yamaguchi, K. Yamazawa, H. Takemura and N. Yokoya: “Real-Time Generation and Presentation of View-Dependent Binocular Stereo Images Using a Sequence of Omnidirectional Images,” Proc. 15th IAPR Int. Conf. on Pattern Recognition (ICPR2000),

Vol. IV, pp. 589-593, 2000.

[10] V. Vanijja and S. Horiguchi: “Omni-Directional Stereoscopic Images from One Omni-Directional Camera,” Jour. of VLSI Signal Processing Systems, Vol. 42, Issue 1, pp. 91-101, 2006.

[11] S. Knorr, M. Kunter and T. Sikora: “Super-Resolution Stereo- and Multi-View Synthesis from Monocular Video Sequences,” Int. Conf. 3-D Digital Imaging and Modeling (3DIM), pp. 55-64, 2007.

[12] M. Hori, M. Kanbara and N. Yokoya: “Novel Stereoscopic View Generation by Image-Based Rendering Coordinated with Depth Information,” Proc. 15th Scandinavian Conf. on Image Analysis (SCIA2007), pp. 193-202, 2007.

[13] 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 (MIRU2005) 講演論文集, pp. 933-938, 2005.