

無人飛行船に搭載された2台の全方位カメラを用いた不可視領域のない全天球HDRビデオの生成

大倉 史生^{*1} 神原 誠之^{*1} 横矢 直和^{*1}

Aerial HDR Video Generation of Full Spherical Views
Using Two Omnidirectional Camera Units

Fumio Okura^{*1}, Masayuki Kanbara^{*1} and Naokazu Yokoya^{*1}

Abstract – This paper describes a method for acquiring full spherical videos without missing areas using a pair of omnidirectional cameras mounted on the top and bottom of an unmanned airship. Omnidirectional videos are captured as multi-exposure images with various shutter speeds to realize high dynamic range (HDR) imaging. The generated videos are intended to be used for telepresence, augmented telepresence, and image-based lighting. In experiments, full spherical HDR videos are generated from image sequences captured from the sky of our campus.

Keywords : Omnidirectional image, High dynamic range, Telepresence, Augmented telepresence, Image based lighting

1 はじめに

本論文では、テレプレゼンス [1]、イメージベースドライティング (IBL) [2]、拡張テレプレゼンス [3] 等、品質の高い仮想環境の構築に要求される不可視領域のない全天球動画の生成手法を提案する。没入感の高い遠隔地の情景をユーザに提示する技術はテレプレゼンスと呼ばれ、多くの研究が行われている。特に、あらかじめ蓄積された遠隔地の情景を用いる蓄積再生型テレプレゼンスシステムとして、ウォークスルー [4] やエンタテインメント等を応用とする様々な例が見られる。特に全天球画像を用いた見回し可能な蓄積再生型テレプレゼンスは、近年では Google 社の Google Street View [5] などを通じて広く普及しており、当該研究分野の重要性が高まっている。また、遠隔地の情景をそのまま提示するだけでなく、カメラ等で取得した実環境の情景に仮想物体を重畳する拡張現実感 (Augmented Reality: AR) [6] と全方位テレプレゼンスを組み合わせることで、ユーザへの情報提示を効果的に行う拡張テレプレゼンス (Augmented Telepresence) [3] に関する研究が行われている。文献 [3] では、全方位テレプレゼンスとして蓄積された空撮全天球動画を用いてユーザの見回し動作を実現するだけでなく、仮想物体の写実的な合成を実現するため、全天球画像を光源環境として高品質な陰影付けを行う手法である IBL [2] を適用している。

本研究では、テレプレゼンスや IBL、拡張テレプレ

ゼンスのための全天球動画の作成を目的とし、全方位カメラを用いて実環境を撮影する際に生じる2つの問題点である (1) 不可視領域の存在と (2) ダイナミックレンジの不足を解決する空撮システムおよび動画生成手法を提案する。以降、全方位撮影における問題点および本研究における解決のためのアプローチを詳述する。

2 全方位撮影における問題点

2.1 不可視領域の存在

全方位カメラを用いて実環境を撮影する際、図 1(a), 1(b) に示すように、全方位カメラの視野外、もしくは全方位カメラを搭載する移動体により遮蔽される領域 (以下、不可視領域) が存在する。不可視領域の存在はテレプレゼンスにおいて没入感を損なう原因となり、さらに IBL においてはレンダリングに必要な光源環境の欠損によるレンダリング品質の低下につながる。

不可視領域の除去に関する従来研究は以下の2種類のアプローチに大別される。

(1) 推定される輝度値を用いる手法

空の輝度値を統計的にモデル化した天空光モデルを利用して空撮動画に含まれる不可視領域を補完する研究 [3] が行われている。文献 [3] では、飛行船の下部に搭載された全方位カメラ1台で空撮された画像の上方に存在する不可視領域を、不可視領域周辺に映りこんだ空の輝度値から天空光モデルのパラメータを推定し補完する。また、文献 [7] のような1枚の画像のみを用いた画像修復の手法を全天球画像に適用する手法も、本アプローチに含まれる。

^{*1}奈良先端科学技術大学院大学

^{*1}Nara Institute of Science and Technology (NAIST)

(2) 実測される輝度値を用いる手法

Kawai らは、1 台の全方位カメラで地上を移動しながら撮影された動画像を対象として、動画像中の他フレームから不可視領域およびその近傍との類似領域を探索することにより、地表面に存在する不可視領域を補完 [8] している。補完対象となる不可視領域に本来写りこむべき情景が動画像中の他フレームで撮影され、補完に用いられる。このような手法は、実際にカメラによって観測される輝度値を用いて不可視領域の除去を行なうアプローチである。

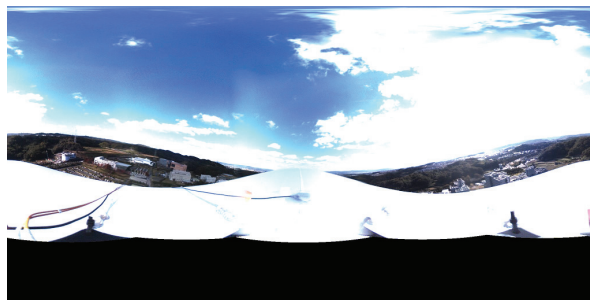
(1) のアプローチは不可視領域の真の輝度値を知ることが目的ではなく、これらの手法で不可視領域が修復された全天球画像を IBL の光源環境として用いた場合、実環境に生じる陰影との誤差が生じる。特に、最も強い光源である太陽が雲などに隠れているかどうか分からないため、太陽周辺の輝度推定の誤差が大きくなる可能性がある。また、文献 [3] の手法においては、雲が局所的に存在するような空のテクスチャを生成することができない。そのため空の散乱強度が一樣でない場合に、不可視領域外と大きく異なる見えのテクスチャを生成し、没入感を損なう原因となる。

このため本研究では、(2) のアプローチを採用し、飛行船上下に一对の全方位カメラを上向きおよび下向きに配置して各カメラの不可視領域を相互に補完するように撮影して不可視領域を除去することで、(1) のアプローチでは困難であった実環境の全天球の輝度値の観測を実現する。

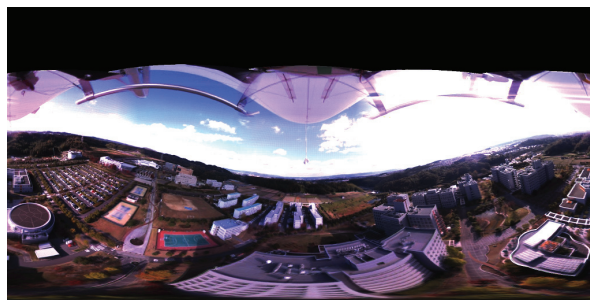
2.2 ダイナミックレンジの不足

IBL を用いた拡張テレプレゼンス等のために全天球画像を用いる場合、実空間の輝度値を正確に計測する必要がある。しかし、屋外での撮影の場合、太陽や空などの高い輝度を持つ領域を直接撮影するため、非常に高いダイナミックレンジが必要になる。屋外シーンでは太陽は最も主要な光源であり、陰影付けの際の仮想物体の見えの変化に大きく寄与する。しかし、晴れた日の空の最も暗い部分と最も明るい部分（太陽の中心）の輝度比はおよそ $1:2^{17}$ といわれ [9]、8bit のダイナミックレンジを持つ多くのカメラでは飽和なく撮影することが困難である。

撮影対象に対しカメラのダイナミックレンジが不足する場合、複数の異なる露出値で撮影されたローダイナミックレンジ (LDR) 画像 (多段階露出 LDR 画像) からハイダイナミックレンジ (HDR) 画像を生成する Debevec らの手法 [10] が広く用いられており、本研究でもこれを適用する。しかし、本研究のように移動撮影を行う場合、多段階露出 LDR 画像間で、撮影位置・姿勢の変化に起因する画像上での位置ずれが発生する。本研究では、カメラの移動量に対して撮影さ



(a) 飛行船上部のカメラから撮影される画像



(b) 飛行船下部のカメラから撮影される画像



(c) 不可視領域のない全天球画像

図 1 2 枚の全天球画像からの不可視領域のない HDR 画像の生成

Fig.1 Full spherical HDR image generation from two omnidirectional images

れる物体が遠い場合にも画像上の位置ずれへの影響が大きい、カメラの姿勢変化を撮影された画像群から推定することで、画像上での位置ずれを補正する。

多段階露出 LDR 画像の撮影において、各画像を撮影するためのシャッタースピードを決定する必要がある。しかし本研究で想定しているような全天球撮影においては、雲が移動し太陽を遮蔽するなど、時間によって周囲の光源強度が変化するため、シャッタースピードを固定して撮影し続けることは困難である。さらに、太陽を含む高輝度領域を撮影可能なシャッタースピードを決定したとき、地上等の低輝度領域で用いることのできる階調数が減少し、シーンの再現性が低下する可能性がある。そこで本研究では、光源環境の時間的変化に対応し、さらに低輝度領域に多くの階調を割り当てるため、実際に撮影された光源強度の分布から適応的にシャッタースピードを決定・更新する。

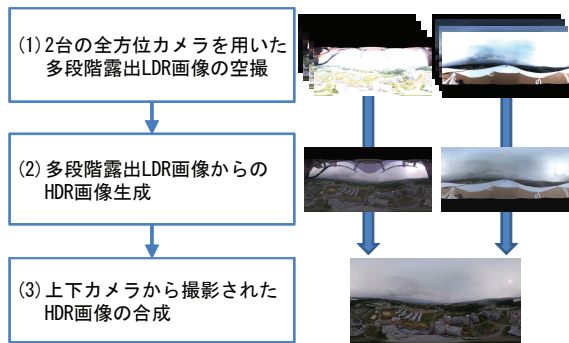


図2 提案手法の処理の流れ
Fig. 2 Flow of proposed aerial imaging

また、上向きの全方位カメラは太陽や空を直接観測するために、シャッタースピードを短く設定しても輝度値の飽和が発生する。そのため、太陽や空を飽和なく撮影する従来研究 [9] と同様に、上向きの全方位カメラに減光 (ND) フィルタを取り付ける。

以下では、本研究で提案するシステムを (1)2台の全方位カメラを用いた多段階露出 LDR 画像の空撮, (2)多段階露出 LDR 画像群からの HDR 画像生成, (3) 上下カメラから撮影された画像の合成, の3ステップに分けて詳述する。最後に、実際に本学キャンパス上空から撮影された画像群を用いて全天球 HDR 画像を生成する実験について述べ、その結果について考察する。

3 不可視領域のない全天球 HDR 画像の生成

3.1 概要

本研究では、飛行船の上下に取り付けられた2台の全方位カメラで撮影した画像を用いて、それぞれの不可視領域を補完するように合成し、不可視領域のない全天球画像を生成する。また、各全方位カメラのシャッタースピードを段階的に変化させて撮影された LDR 画像群を合成することで HDR 画像を生成する。

提案手法の流れを図2に示す。図2右は以下で述べる各処理で撮影・生成される画像の例である。

(1) 2台の全方位カメラを用いた多段階露出 LDR 画像の空撮

飛行船の上部に上向き、下部に下向きに、一対の全方位カメラ (上カメラ, 下カメラ) を取り付け、シャッタースピードを変化させながら動画を撮影する。飛行船の上方から撮影されるシーンには主に太陽や空が含まれ、本研究で用いる全方位カメラで設定可能な最短のシャッタースピードで撮影を行っても輝度値の飽和が起こるため、上カメラには減光フィルタ (ND フィルタ) を取り付ける。また、少ない撮影枚数で広いダイナミックレンジを計測するため、以前のフレームで撮影された画像のヒストグラムを用いてシャッタースピードを動的に自動設定する。

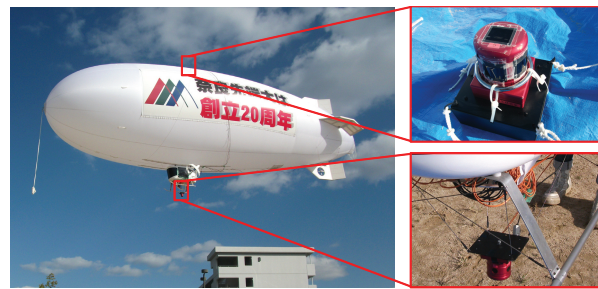


図3 無人飛行船の外観と撮影に用いた全方位カメラ
Fig. 3 Unmanned airship and omnidirectional cameras

(2) 多段階露出 LDR 画像からの HDR 画像生成

多段階露出 LDR 画像から上下カメラそれぞれの HDR 画像を生成する。多段階露出 LDR 画像間には、撮影時のカメラの移動および姿勢変化によって位置ずれが発生する。本研究ではその中でも画像上の位置ずれへの影響が大きいカメラの姿勢変化を推定し、画像補正を行う。

(3) 上下カメラから撮影された HDR 画像の合成

(2) で生成された2枚の HDR 画像を合成し、1枚の全天球 HDR 画像を生成する。本研究で用いる飛行船は形状変化が発生するため、2台の全方位カメラ間の相対的な位置関係を固定することは困難である。そのため、各カメラから撮影された画像の重複領域から相対的な姿勢変化を推定し、位置合わせを行う。また、ND フィルタの有無による色調の違いを補正する。

以下、上述の処理 (1)~(3) について詳述する。

3.2 2台の全方位カメラを用いた多段階露出 LDR 画像の空撮

2台の全方位カメラと飛行船を用いた空撮システム

本研究では、図3に示す全長12mのラジコン無人飛行船を用いて空撮を行う。飛行船には空撮画像を取得するための全方位マルチカメラシステム Ladybug2 (Point Grey Research 社製) を上向き、および下向きに計2台取り付け。Ladybug2の仕様を表1に示す。

表1 全方位マルチカメラ Ladybug2の仕様
Table 1 Specifications of omnidirectional multi-camera unit Ladybug2

項目	詳細
搭載カメラ数	6台
カメラ解像度	768×1024
カバーする視野	全天球視野の75%以上
搭載CDD	Sony ICX204AK
フレームレート	最大30fps
重量	1.1kg (バッテリー除く)
シャッタースピード 格納用レジスタ数	4 (多段階露出 LDR画像撮影用)

す. Ladybug2 にはシャッタースピードを格納できるレジスタが4つ用意されており, これらを切り替えながら撮影を行うことができる. よって, 本研究では4種類のシャッタースピードを切り替えながら撮影された, 露出の異なる4枚のLDR画像からHDR画像を生成する. 2台の全方位カメラは1台のPCに接続され, 各撮影画像にタイムスタンプが付与される. 本研究では, 使用カメラの絞りおよびシャッタースピードを最小に設定し撮影した際も輝度の飽和がみられたため, 主に太陽や空を撮影する上カメラには光量を $\frac{1}{100}$ に減衰するNDフィルタであるND 2.0(富士フィルム社製)を取り付ける. 飛行船上下の各カメラから撮影された画像は, 周辺減光の除去および幾何変換[11]により, それぞれ図1(a), 1(b)のようなパノラマ画像(解像度 2048×1024)に変換される. また, 拡張テレプレゼンスにおける仮想物体の位置合わせ等に用いる全天球画像の撮影位置情報を取得するため, 電子基準点網によるネットワーク補正を利用するDifferential GPSであるP4-GPS(日立造船株式会社製)を搭載する.

シャッタースピードの自動設定

光源環境の時間的変化に対応し, さらにシーンの低輝度領域に多くの階調を割り当てるために, 多段階露出LDR画像の撮影時に用いるシャッタースピードを, 撮影されたLDR画像群を用いて動的に更新する. 現在, 多段階露出LDR画像を撮影可能なカメラの一部にオートブラケット機能が存在し, シャッタースピードの自動設定が実装されている. これらのシステムでは一般的に, 周辺環境の光量からシャッタースピードを設定し, さらに露出を一定量変化させて複数の露出を決定する[12]. また, 撮影するシーンのダイナミックレンジからシャッタースピードを一意に求める研究[13]も存在する.

これに対し本研究では, 一定時間前のシーンの輝度分布が得られているため, 動画像中の以前のフレームで撮影されたシーンのヒストグラムを用い, 限られた撮影枚数で高いダイナミックレンジをカバーするために, (1)輝度値の飽和を抑え, (2)撮影される画像枚数が少ないことによる見た目への影響を低減するよう, シーンの輝度分布に対し適応的にシャッタースピードを変化させる. 本研究における実装では, シャッタースピード更新処理に1秒程度を要するため, 更新レートを数秒に1度程度としている.

ここで, 前フレームの撮影に用いたシャッタースピード群を短い順に $s_{old_1}, s_{old_2}, \dots, s_{old_n}$, 新たに設定されるシャッタースピード群を短い順に $s_{new_1}, s_{new_2}, \dots, s_{new_n}$ とする. 本研究では4枚の多段階露出LDR画像を用いるため, $n=4$ として実験を行った.

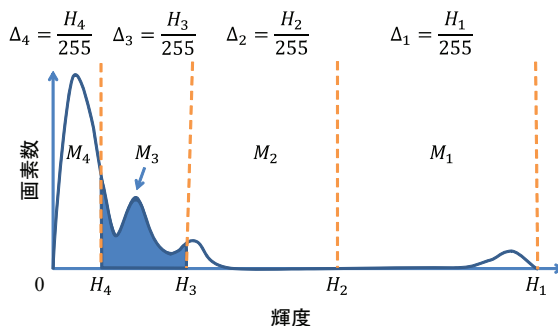


図4 HDRヒストグラムの模式図($n=4$ の例). Δ_i は量子化幅, H_i はシャッタースピード s_{new_i} で飽和なく撮影可能な最大輝度値を表す. M_3 はヒストグラム中で塗りつぶされた領域の画素数の合計であり, 他の M_i についても同様に定義される.

Fig.4 Diagram of HDR histogram (In case of $n=4$)

(1) 最も短いシャッタースピード s_{new_1} の決定

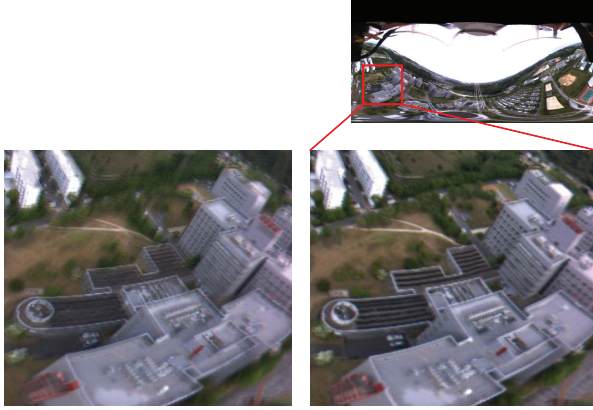
最も短いシャッタースピード s_{new_1} で撮影される画像は, シーン中で最も輝度の高い領域を撮影する. ダイナミックレンジを効率的に利用するため, s_{new_1} を, 輝度値の飽和を起こさない範囲で長く設定する. s_{old_1} で撮影されたLDR画像中で最大の輝度値を $L_1(s_{old_1})$ としたとき, 最も短いシャッタースピード s_{new_1} は以下のように決定される.

$$s_{new_1} = \begin{cases} 0.5s_{old_1} & (L_1 = 255) \\ \frac{255 + \theta_{sh}}{2L_1} s_{old_1} & (L_1 \leq \theta_{sh}) \\ s_{old_1} & (\text{otherwise}) \end{cases} \quad (1)$$

ただし, s_{new_1} の取りうる最長・最短シャッタースピードは, カメラの仕様により制限される. ここで, θ_{sh} はあらかじめユーザによって指定される閾値であり, レスポンスカーブ(CCDへの入射光量と画像間の関係)が線形である場合 s_{new_1} が $\theta_{sh} < L_1 < 255$ となるように設定される. 本研究における実装では, 経験的に $\theta_{sh} = 192$ を用いた. 光源環境の変化がない場合でも, s_{new_1} は1度の更新処理では収束せず, 数回の更新処理を要する場合がある.

(2) $s_{new_2}, \dots, s_{new_n}$ の決定

撮影される多段階露出LDR画像の枚数が少ない場合, HDR画像上での量子化幅が大きくなるために擬似エッジが発生する可能性がある. 本研究では, そのような見た目に対する影響を低減するために, 各LDR画像の量子化幅が生成されるHDR画像においてどの程度の大きさになるかを考慮し, HDR画像上での量子化幅が小さくなるようにシャッタースピードを決定する. そのために, 撮影されたLDR画像群を用いて, 図4に示すように, 生成されるHDR画像のヒストグラムを計算する. i 番目($i \geq 2$)に短いシャッタース



(a) 位置合わせ処理なし (b) 位置合わせ処理あり

図5 多段階露出LDR画像の位置合わせ処理の有無によるHDR画像の見えの変化

Fig. 5 HDR images with and without multi-exposure image alignment

ピード s_{new_i} について、 s_{new_i} で撮影されるLDR画像の最大輝度値255に対応するHDR画像上での輝度値を $H_i(s_{new_i})$ とする。ここで、 $M_i(s_{new_i})$ を以下のように定義する。

$$M_i = \sum_{k \in \mathcal{H}} j_k \quad (2)$$

ここで、 j_k は、輝度値 k を持つ画素数とする。また、 \mathcal{H} は以下のように定義される。

$$\mathcal{H} = \begin{cases} 0 & (i = n) \\ H_{i+1} & (\text{otherwise}) \end{cases} \quad (3)$$

カメラのレスポンスカーブが線形であると仮定したとき、 s_{new_i} により撮影されるLDR画像の量子化幅は、HDR画像のヒストグラム上で Δ_i として求められる。

$$\Delta_i = \frac{H_i}{255} \quad (4)$$

コスト関数 E_s は、各シャッタースピードに対応する量子化幅を、各シャッタースピードで撮影される画素がシーン中にどれだけ含まれるかを重みとした関数であり、 E_s を最小化することにより s_{new_i} を決定する。

$$E_s = \sum_i M_i \Delta_i \quad (5)$$

E_s はいくつかの局所解をもつ1次元関数となる。本研究では解空間を100分の1程度に間引いて探索し、最も小さい解の周囲を全探索する粗密探索により最小値を求めている。

3.3 多段階露出LDR画像からのHDR画像生成LDR画像群の位置合わせ

多段階露出LDR画像間には、撮影時のカメラの並進および姿勢変化によって、撮影された物体の画像上

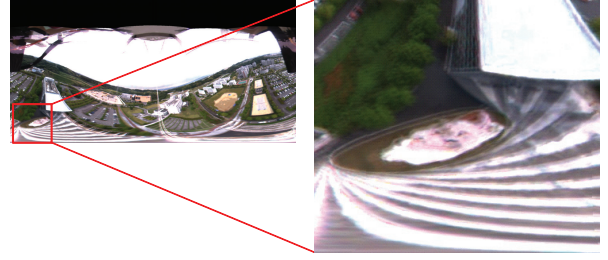


図6 カメラの並進による位置ずれ（姿勢変化による位置ずれの補正処理後）

Fig. 6 Remaining blur on HDR image due to position change of camera

での位置ずれが発生する。そのため、位置ずれを考慮せずに生成されたHDR画像では図5(a)のようにぼけが発生する。本研究ではその中でも、カメラの移動量に対して撮影される物体が遠い場合にも画像上の位置ずれへの影響が大きいカメラの姿勢変化を画像群から推定して補正を行うことによって、図5(b)のようにぼけを抑制する。

本研究で用いる無人飛行船の平均的な飛行状態（速度 $5m/s$ 、回転量 $30^\circ/s$ ）を仮定したときの、見回しテレプレゼンスシステムにおいてユーザに提示される透視投影画像上には、カメラの並進・姿勢変化によって表2に示すような位置ずれが発生する。飛行船の飛行高度を100mとしたとき、4枚の多段階露出LDR画像を撮影する間に、カメラの並進によって最大5.81画素、姿勢変化によって対象物体までの距離に関わらず60画素の位置ずれが発生する。本研究における位置合わせ処理は、このうちカメラの姿勢変化による画像上での位置ずれを解消することを目的とする。したがって、位置合わせ処理後もカメラの並進による位置ずれは残っており、移動量が多い場合、生成されたHDR画像に含まれる近景を中心に位置ずれ量は大きくなる（図6参照）。

位置合わせの流れを図7に示す。まず、図7(a)に示すように、多段階露出LDR画像群から白とび（本研究では輝度値が255と設定）、黒つぶれ（輝度値が16

表2 カメラの並進・姿勢変化による画像上での位置ずれ量

Table 2 Misalignment of LDR images due to change of position and orientation of the camera

	対象物体までの距離	1サイクルあたり視差	透視投影画像上の位置ずれ
並進	50m	1.43°	11.4画素
	100m	0.726°	5.81画素
	500m	0.143°	1.14画素
姿勢変化	n.a.	7.5°	60.0画素

※カメラの並進速度を $5m/s$ 、回転量を $30^\circ/s$ とし、垂直画角 45° 、 640×480 画素の透視投影画像上での位置ずれを評価。多段階露出画像4枚を撮影する時間（0.25秒）を1サイクルとした。姿勢変化による位置ずれ量は、対象物体までの距離に依存しない。

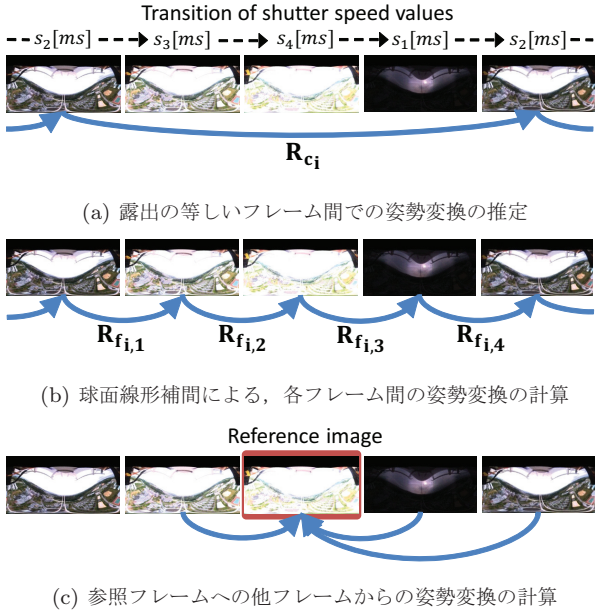


図7 多段階露出 LDR 画像の位置合わせの流れ
 Fig.7 Flow of multi-exposure image alignment

未満と設定) の最も少ない露出画像と、その近傍に存在する同露出で撮影された画像を選択し、2フレーム間の対応点を KLT トラッカ [14] により求める。得られた m 番目の特徴点および対応点を単位球面に投影したものをそれぞれ $\mathbf{p}_m, \mathbf{q}_m$ とすると、オイラー角の3パラメータで表される姿勢変換 \mathbf{R}_{c_i} は、 $\mathbf{p}_m, \mathbf{q}_m$ のユークリッド距離 $|\mathbf{p}_m, \mathbf{q}_m|$ の二乗和として定義されるエネルギー E_l を非線形最小化することにより推定される。

$$E_l = \sum_m |\mathbf{p}_m, \mathbf{q}_m|^2 \quad (6)$$

ここで、誤対応による推定誤差を軽減するため、RANSAC を用いて外れ値を除去する。

次に、図7(b)に示すように、 \mathbf{R}_{c_i} から、各フレーム間の姿勢変換パラメータ $\mathbf{R}_{f_i,j}$ を球面線形補間を用いて求める。

動画中の任意のフレームを参照フレームとし、図7(c)に示すように、 $\mathbf{R}_{f_i,j}$ を用いて参照フレームに対する他の露出画像からの姿勢変換を計算し変換を行うことで、参照フレームに位置合わせされた多段階露出 LDR 画像群を得る。

動画像全体に対して図7(a)および7(b)に示される処理を行い、さらに全フレームをそれぞれ参照フレームとして図7(c)の処理を行うことにより、動画像全体に対して位置合わせ処理を行うことができる。

ND フィルタによる光減衰を考慮した HDR 画像生成

本研究では、HDR 画像生成に多く用いられる Debevec らの手法 [10] のように、撮影時のシャッタースピードから各多段階露出 LDR 画像 I_l の輝度値に対す

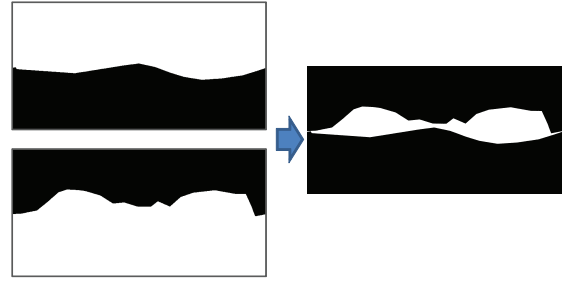


図8 マスク画像の作成：不可視領域が黒くマスクされている(左：上画像のマスク，左下：下画像のマスク，右：重複領域を白い領域として示す画像)

Fig.8 Mask images used for alignment between two cameras

る出力 HDR 画像の輝度値 I_h を求め、飛行船上下カメラからの2枚の HDR 画像を生成する。しかし、本研究では上カメラには ND フィルタを取り付けているため、ND フィルタの光の公称減衰量 η を用いて以下のように出力輝度値を補正する。

$$I_h = \begin{cases} \alpha \frac{I_l}{t} & (ND \text{ フィルタなし}) \\ \alpha \frac{I_l}{\eta t} & (ND \text{ フィルタあり}) \end{cases} \quad (7)$$

α はスケール係数であり、HDR 画像の輝度の絶対量を用いず相対量のみを求める場合には本質的な意味を持たないが、カメラ絞りや感度等の情報を用い適切に設定することで実際の物理量 $[cd/m^2]$ と対応付けることができる。各多段階露出 LDR 画像から計算された出力輝度値は、文献 [10] と同様に合成される。

3.4 上下カメラから撮影された HDR 画像の合成 上下カメラ間の姿勢推定および補正

飛行船上下に搭載された全方位カメラは、飛行船の細かな変形や揺れなどのため、相対的な位置関係を完全に固定することは難しい。そこで、各カメラの3自由度の相対的な姿勢変化を検出することで、全天球画像の姿勢補正を行う。両カメラの搭載位置は飛行船の直径である約 3.5m 程度離れているが、各カメラから撮影された全天球画像の重複部位は水平線に近く遠景が多いため、本研究では両カメラ間に発生する視差を無視する。

本研究では、同時刻で撮影された上下画像の対応関係および、前フレームで撮影された画像との対応関係を用いて以下のように姿勢変化を求め、補正する。

(1) マスク画像の作成

上カメラから撮影された画像(上画像)、下カメラから撮影された画像(下画像)ともにあらかじめ手動で不可視領域を示すマスクを作成しておく。全天球画像中の不可視領域は動画像全体を通じて大きく変化しないため、動画像中のある1フレームで作成したマスクを動画像全体に用いる。作成された上下カメラのマ



(a) ND フィルタなし (下画像) : 平均輝度値 $(R, G, B) = (93.2, 91.2, 100.6)$
 (b) ND フィルタあり (上画像) : 平均輝度値 $(R, G, B) = (83.8, 91.6, 99.0)$

図9 ND フィルタによる色調変化: 同じ露出に変換された上下画像の重複領域の一部
 Fig.9 Chromatic change due to ND filter

スクを用いて, 図8に示すように, 両画像に共通して情景が映り込む領域 (重複領域) を求める.

(2) 上カメラから下カメラへの相対的な姿勢変換の推定
 上画像の重複領域に含まれる特徴点に対する下画像の対応点を KLT トラッカ [14] を用いて検出する. 上下画像間のエネルギー E_s は, 3.3 における露出の等しいフレーム間の姿勢変換の推定手法と同様に求められる.

各フレームで独立に姿勢変換を求めた場合, マッチングの精度によっては求められる姿勢変換パラメータが時間的に細かく変化し出力される動画像に揺れが発生する可能性がある. 本研究では, 前フレームとの対応関係を用いることによりこの問題を解決する. n フレーム目の上画像の不可視でない領域に存在する特徴点 \mathbf{p}_n および, 1つ前フレームの対応点 \mathbf{p}_{n-1} を用いて, 以下のエネルギー E_t を定義する.

$$E_t = \sum_n |\mathbf{p}_n, \mathbf{R}\mathbf{f}_{n-1,n}\mathbf{p}_{n-1}|^2 \quad (8)$$

ここで, $\mathbf{R}\mathbf{f}_{n-1,n}$ は, 3.3 で求められた $n-1$ フレーム目から n フレーム目への姿勢変換である.

E_s および E_t の線形和としてエネルギー E_h を定義する. 3.3 と同様に E_h が最小となる姿勢変換パラメータを求め, 上カメラから下カメラへの姿勢変換とする.

$$E_h = E_s + \beta E_t \quad (9)$$

本研究では, 係数 $\beta = 1$ として実験を行い, 良好な結果を得た. ただし, 動画像の初期フレームについては E_t が定義できないため, E_s のみを最小化して姿勢変換パラメータを求める.

ND フィルタによる色調変化の補正

ND フィルタは一般的に可視光の全波長を均等に透過する減光フィルタとして用いられている. しかし, 厳密には全波長を均等に透過せず, ND フィルタを用いて撮影すると色調が変化する [9]. 本研究で用いる ND フィルタは, 図9に示すように赤成分をあまり透過しない. そこで, ND フィルタによる色調の違いを吸収するため, 姿勢補正済み上下画像の重複領域内の輝

度値を比較することで, 以下のように RGB の各チャンネルで独立に定義される色調パラメータを, 上画像の重複領域に含まれる画素の輝度値が下画像における対応画素の輝度値に最も近づくように推定し, 色調を補正する.

$$R'_{top}(\mathbf{x}) = \gamma_r R_{top}(\mathbf{x}) \quad (10)$$

$$G'_{top}(\mathbf{x}) = \gamma_g G_{top}(\mathbf{x}) \quad (11)$$

$$B'_{top}(\mathbf{x}) = \gamma_b B_{top}(\mathbf{x}) \quad (12)$$

ここで, \mathbf{x} を注目画素, $(R_{top}(\mathbf{x}), G_{top}(\mathbf{x}), B_{top}(\mathbf{x}))$ を補正前の上画像における \mathbf{x} の RGB それぞれの輝度値, $(R'_{top}(\mathbf{x}), G'_{top}(\mathbf{x}), B'_{top}(\mathbf{x}))$ を補正後の上画像における \mathbf{x} の輝度値とする.

$(R_{bot}(\mathbf{x}), G_{bot}(\mathbf{x}), B_{bot}(\mathbf{x}))$ を下画像に含まれる画素 \mathbf{x} の輝度値, A を重複領域, N_A を重複領域に含まれる画素数としたとき, γ_r は以下のように推定される. γ_g, γ_b についても同様に推定する.

$$\gamma_r = \frac{\sum_{\mathbf{x} \in A} \frac{R_{bot}(\mathbf{x})}{R_{top}(\mathbf{x})}}{N_A} \quad (13)$$

推定された色調パラメータを用いて, 上画像全体に対し式 (10)~(12) を適用することにより色調変換を行う.

補正済み上下画像の合成

最後に, 補正済みの上下画像を合成することで1枚の不可視領域のない全天球HDR画像を生成する. このとき, 重複領域に含まれる画素の輝度値は, 両画像をアルファブレンディングして求める. 上画像中の画素の輝度値 I_{top} , 下画像の同位置にある画素の輝度値 I_{bot} から, 最終出力となる画像の同位置の画素の輝度値 I_{full} は, 一般的なアルファブレンディングを用いて以下のように求められる.

$$I_{full} = \alpha_{br} I_{bot} + (1 - \alpha_{br}) I_{top} \quad (14)$$

重複領域の上端で $\alpha_{br} = 0$, 下端で $\alpha_{br} = 1$ となるように設定され, その間を線形に補間することで, 姿勢および色調が補正された上下画像を合成する.

4 飛行船からの空撮画像を用いた実験結果

4.1 静止画像を用いた実験結果

提案手法が実環境を忠実に反映したHDR画像を生成できることを確認するため, 実際にキャンパス上空から飛行船を用いて空撮を行い, 静止画像を用いて不可視領域のない全天球HDR画像を生成した. ここでは, 上下カメラ間の姿勢推定 (3.4) における連続フレーム間の制約である E_t を用いず, 動画像の初期フレームと同様に E_s のみを制約として姿勢変化の推定を行った. 撮影時, 飛行船の高度は地上からおおよそ130mであり, 約3m/sで移動していた. 各画像のシャッタースピー

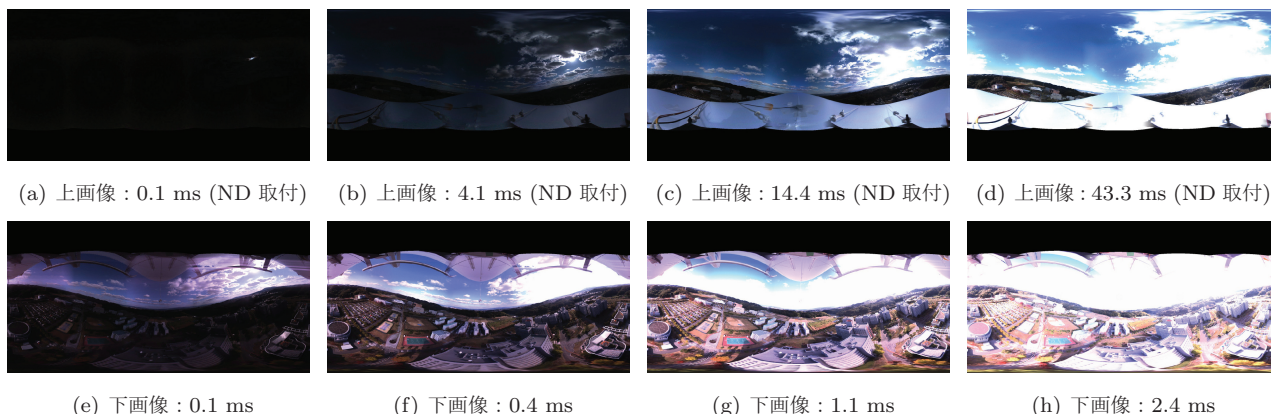


図 10 2 台の全方位カメラから撮影された多段階露出 LDR 画像

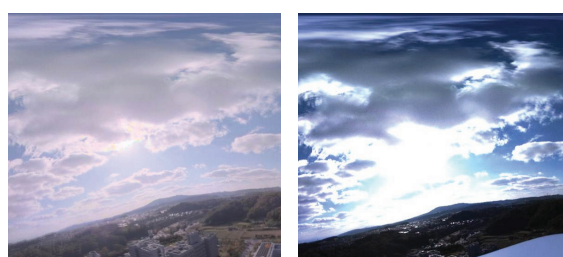
Fig. 10 Multi-exposure images captured from two omnidirectional cameras



(a) 図 10(e) と同露出 (b) 図 10(h) と同露出

図 11 多段階露出 LDR 画像と同露出になるよう変換された出力画像

Fig. 11 Full spherical images which intensity is cropped with exposures equal to multi-exposure images



(a) HDR 撮影 (図 12 の拡大図) (b) LDR 撮影 (図 10(c) の拡大図)

図 13 HDR 撮影と LDR 撮影の比較

Fig. 13 Closeups of spherical HDR and LDR image



図 12 Reinhard らの手法 [15] によりトーンマッピングされた出力画像

Fig. 12 Full spherical image generated by tone-mapping of Reinhard, et al. [15]

どと、撮影された多段階露出 LDR 画像例を図 10 に示す。撮影された画像群から生成された上下 HDR 画像が、先に示した図 1(a), 1(b) である。上画像と下画像の重複領域をアルファブレンディングして、図 1(c) に示した最終出力画像を生成した。

生成された HDR 画像 (図 1(c)) を図 10(e), 10(h) と同じ露出に輝度変換した画像を図 11 に示す。変換後の輝度は適切に表現され、実際に撮影された多段階露出 LDR 画像に近い見えに変換されており、生成された HDR 画像が実際に撮影された環境を忠実に反映している。Reinhard らのトーンマッピング手法 [15] を用いて、HDR 画像に含まれるテクスチャを保持したまま生成された LDR 画像を図 12 に示す。本研究で生

成された全天球 HDR 画像をトーンマッピングし、見直し可能な全天球テレプレゼンス等に用いることで、ユーザにあらゆる方向、あらゆる輝度の詳細なテクスチャを提示することが可能となる。

図 12 に含まれる空と地上の境界付近を拡大して図 13(a) に示す。上カメラで撮影された LDR 画像のうち、図 12 と最も輝度平均の近い図 10(c) の対応する領域の拡大図を図 13(b) に示す。空の領域においては、一般的な LDR 撮影で得られた図 13(b) の高輝度部分である太陽周辺に輝度の飽和が見られるが、HDR 画像をトーンマッピングした図 13(a) ではテクスチャが視認できる。一方、地上の領域においては、図 13(b) では図 13(a) と比較して明度が低下し、一部に黒つぶれが見られる。これらから、HDR 画像をトーンマッピングした画像は、従来の LDR 画像による表現と比較して視認性が向上していることが確認できる。

4.2 動画像を用いた実験結果

実際に、提案手法で拡張テレプレゼンス等で応用可能な HDR 動画像が生成できることを確認するため、本手法を全 500 フレームからなる動画像に適用した。飛行船の高度はおよそ 130m であり、移動速度は約 5 ~ 8m/s であった。生成された動画像は撮影時と同じ 16fps であり、処理時間は 1 フレームあたりおよそ 10

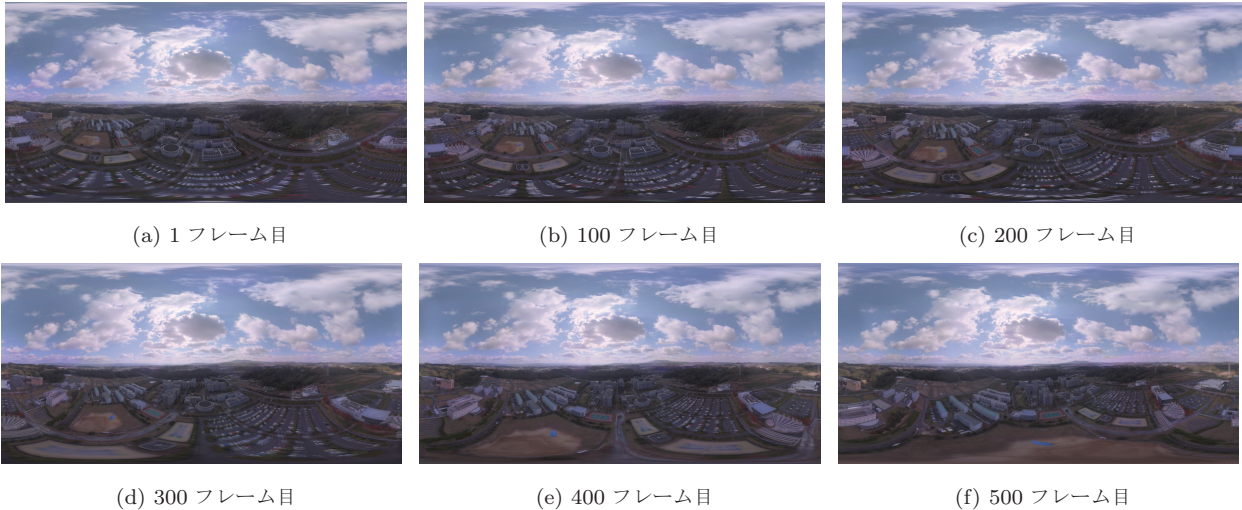


図14 動画像中の様々なフレームから生成された全天球HDR画像
Fig. 14 Full spherical HDR images generated from video frames

秒程度となった。下カメラから撮影された画像に対し全方位カメラを用いた structure-from-motion[16] を適用し得られた全方位カメラの位置姿勢を用いて、地平線が画像中の水平線と一致し、画像中心が常に同じ方角となるよう提案手法で生成されたHDR画像の姿勢を統一し、トーンマッピングした画像を図14に示す。本研究で得られた全方位画像についても、従来と同様にカメラの位置・姿勢を推定することが可能であり、提案手法で生成された全天球HDR動画像およびカメラの位置・姿勢情報は、拡張テレプレゼンス等において仮想物体を合成するなどの用途に用いることができる。

5 考察

5.1 天空光モデルを用いた補完による全天球画像生成手法[3]との比較

4.2において得られた実験結果と、天空光モデルを用いて全天球動画像の空に存在する不可視領域を補完する従来手法[3]によって得られる結果を比較する。手法[3]に4.2と同一の動画像を入力し、図14(b)と同じフレームの不可視領域を補完した結果を図15に示す。なお、図15には4.2の結果と同じパラメータでトーンマッピングを行なっている。

定性的な観点から、図15において補完された領域とそれ以外の領域のテクスチャが大きく異なり、テレプレゼンスシステムとしてユーザが見直しを行う際に没入感・臨場感低下の原因になると考えられる。

提案手法および文献[3]の手法による結果を用いたIBLによるライティングへの影響を調べるため、両手法によって出力される全天球画像の光源としての特性を定量的に考察する。ライティングにおいてキャストシャドウ等の表現に影響する画像中の最大輝度について、提案手法の最大輝度値を1としたとき、(提案手法

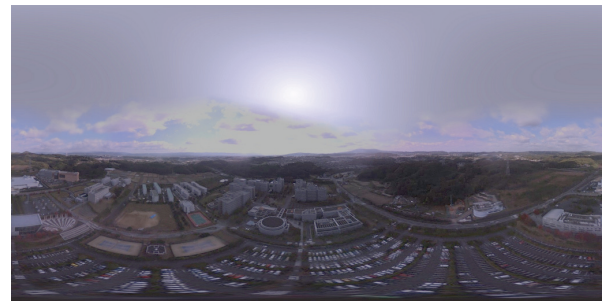


図15 従来手法[3]による不可視領域の補完結果
Fig. 15 Completion of missing area by previous method [3]

における最大輝度値) : (従来手法[3]における最大輝度値) = 1 : 0.10 となった。このとき、最大輝度を持つ画素の方角に 9.53° のずれがあった。また、両手法における空領域の平均輝度は、1 : 0.218 となった。これは、手法[3]によって生成された空を用いてライティングを行なった場合、実際の輝度の観測に基づく提案手法によるライティングと比較し、完全拡散反射を仮定できるような仮想物体の明るさは平均 $\frac{1}{4.586}$ 程度となることを意味する。実際にライティングを行った際に仮想物体の見えに著しい変化を引き起こすため、本実験環境において提案手法による全天球画像生成が有効であると考えられる。

しかし、仮想物体の合成における見えの違和感や満足度は光源環境の精度だけでなく、仮想物体表面に与えられた反射特性の精度や、仮想物体として用いる建造物等が実シーンとして用いる環境に存在しても自然なものかどうかなど、多くの要素に依存する。そのため、より自然な合成を実現するためには、人間の心理も含めた多角的な考察が必要である。

5.2 シャッタースピード自動設定についての考察

本研究におけるシャッタースピードの自動設定について考察を行う。

従来, 文献 [12] をはじめとする多くの手法において, ある露出を基準露出とし, さらに露出を一定量変化させることで多段階露出 LDR 画像撮影のためのシャッタースピードを決定している. ここでは, 3.2 で決定された, 輝度値の飽和を抑制するための最も短いシャッタースピードを s_{c1} [ms] とし, 露出時間を $2^2 = 4$ 倍ずつ増加させ s_{c2}, \dots, s_{c4} を決定する. なお, 露出時間の増加係数として, 多くのオートブラケットカメラにおいて 2^2 が用いられていることから, 本比較でもこれを用いる. ここで, 3.2 と同様に従来手法のシャッタースピード s_{ci} ($1 \leq i \leq 4$) において撮影可能な最大輝度を H_{ci} とし, 提案手法についても実際に 4.1 において上カメラのシャッタースピードとして設定された値を用いて同様に H_{pi} を定義する. シャッタースピード, 最大輝度, および各シャッタースピードで撮影した場合の量子化幅を表 3 にまとめる. 量子化幅は最大輝度に比例し, 本研究における実験と同様, 各シャッタースピードで撮影される画像が 256 階調を持つとして計算した. 表 3 より, 一般的に画素の輝度値が多く分布する低輝度領域において, 従来手法による量子化幅は提案手法の 6.7 倍となる.

実際に図 12 に示した全天球 HDR 画像を従来手法の量子化幅を用いて再量子化した. 図 12 および再量子化された HDR 画像の, 空と地上の境界付近を拡大して図 16 に示す. 図 16(c) およびその微分画像である図 16(d) より, 低輝度の領域を中心に擬似エッジが発生している.

以上より, 本研究におけるシャッタースピードの設定手法が, 従来手法における量子化幅の増大による見た目への影響を減少させることに有効であることが確認できる.

表 3 従来手法および提案手法のシャッタースピードと量子化幅
Table 3 Shutter speeds and quantization steps

	シャッター スピード [ms]	最大輝度	量子化幅
提 案 手 法	0.1	1000	3.92
	4.1	24.4	0.10
	14.4	6.94	0.03
	43.3	2.31	0.009
従 来 手 法	0.1	1000	3.92
	0.4	250	0.98
	1.6	62.5	0.25
	6.4	15.6	0.06

※最大輝度と量子化幅は相対量であり, 物理的な意味を持たない. また, 撮影される画像は 256 階調とする.

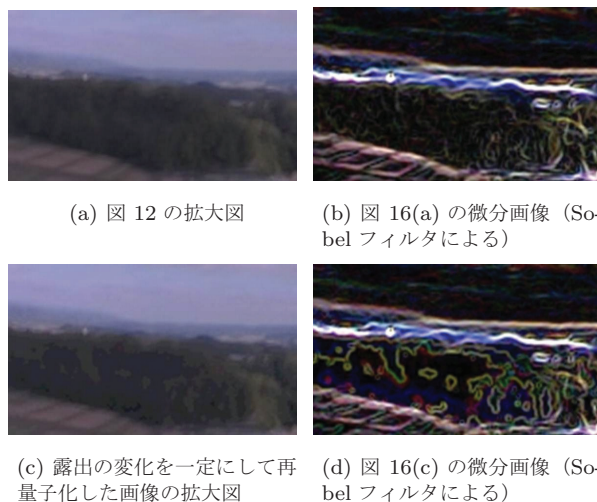


図 16 量子化幅の増大による見た目への影響
Fig. 16 Effect of quantization step to appearance of image

5.3 撮影可能なダイナミックレンジに関する考察

今回の実験では, ND フィルタを取り付けた上カメラで設定可能な最も短いシャッタースピード $s = 0.1$ [s] で撮影された画像においても, 太陽周辺の数画素に輝度の飽和が見られた. そこで, 本実験で生成された画像が輝度の飽和や黒つぶれなしに表現できるダイナミックレンジを計算し, 考察を行う. ここで, 最も露出の大きい画像において輝度値が 16 未満となる場合に黒つぶれ, 最も露出の小さい画像において輝度値が 255 となる場合に輝度値の飽和と定義する. カメラのレスポンスカーブを線形とし, 下カメラで黒つぶれなしに撮影可能な最小輝度を 1 とした場合, 図 10 に示したシャッタースピードにおけるダイナミックレンジを計算すると, 下カメラが $1 : 2^{8.58}$, 上カメラが $2^{2.47} : 2^{15.23}$ となる. 特に上カメラで撮影する空のダイナミックレンジは非常に高い (快晴時に $1 : 2^{17}$ 程度) ため, 本実験環境ではダイナミックレンジが不十分であった. 全画素を飽和なく撮影するためには, さらに強い減光を行う ND フィルタを取り付ける必要がある. また, 適切な ND フィルタの選定の際, 強い減光により黒つぶれする画素数が増加することによる, 見た目や画像補正の処理への影響も考慮すべきである.

6 まとめと今後の展望

本論文では, テレプレゼンスや IBL, 拡張テレプレゼンスに適用可能な, 不可視領域のない全天球 HDR 動画画像の生成手法を提案した. 提案手法では全方位カメラ 2 台を取り付けた飛行船を用いて HDR 空撮を行い, 上下カメラから撮影された全天球画像の姿勢補正および色調補正を行うことにより, 全天球 HDR 画像を生成した. 実際に本学キャンパス上空から撮影された動画画像を用いて全天球 HDR 画像および動画画像の

生成実験を行い、実際の環境を忠実に反映した全天球HDRビデオを生成できた。生成された全天球HDR画像群は、従来のLDR撮像と比較し視認性が向上した。また、提案手法における2台の全方位カメラによる不可視領域のない観測は、従来の天空光モデルによる不可視領域補完手法と比較し、仮想物体へのライティングに有効であることが確認された。今後、生成された全天球HDR動画像および4.2節で得たカメラ位置姿勢情報を用いて、IBL等による拡張テレプレゼンスシステムの構築を行う。

謝辞

本研究は、文部科学省特別経費「アンビエント環境知能研究創出事業」、日本学術振興会科学研究費補助金(基盤研究(A), No.23240024)、およびNAIST先端的研究連携事業による。

参考文献

- [1] J. Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, Vol. 42, No. 4, pp. 73–93, 1992.
- [2] P. Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *Proc. ACM SIGGRAPH'98*, pp. 189–198, 1998.
- [3] 大倉史生, 神原誠之, 横矢直和. 無人飛行船からの空撮全方位動画像を用いた蓄積再生型拡張テレプレゼンス. *日本バーチャルリアリティ学会論文誌*, Vol. 16, No. 2, pp. 127–138, 2011.
- [4] A. Lippman. Movie-maps: An application of the optical videodisc to computer graphics. In *Proc. ACM SIGGRAPH'80*, Vol. 14, pp. 32–42, 1980.
- [5] L. Vincent. Taking online maps down to street level. *IEEE Computer Magazine*, Vol. 40, No. 12, pp. 118–120, 2007.
- [6] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, Vol. 21, No. 6, pp. 34–47, 2001.
- [7] 河合紀彦, 佐藤智和, 横矢直和. テクスチャの明度変化と局所性を考慮したパターン類似度を用いたエネルギー最小化による画像修復. *電子情報通信学会論文誌(D)*, Vol. J91-D, No. 9, pp. 2293–2304, 2008.
- [8] N. Kawai, K. Machikita, T. Sato, and N. Yokoya. Video completion for generating omnidirectional video without invisible areas. *IPSJ Trans. on Computer Vision and Applications*, Vol. 2, pp. 200–213, 2010.
- [9] J. Stumpfel, C. Tchou, A. Jones, T. Hawkins, A. Wenger, and P. Debevec. Direct HDR capture of the sun and sky. In *Proc. Int'l Conf. on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa (AFRIGRAPH'04)*, pp. 145–149, 2004.
- [10] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *Proc. ACM SIGGRAPH'97*, pp. 369–378, 1997.
- [11] S. Ikeda, T. Sato, and N. Yokoya. High-resolution panoramic movie generation from video streams acquired by an omnidirectional multi-camera system. In *Proc. IEEE Int'l Conf. on Multisen-*

- sor Fusion and Integration for Intelligent System (MFI'03)*, pp. 155–160, 2003.
- [12] S. Kang, M. Uyttendaele, S. Winder, and R. Szeliski. High dynamic range video. In *Proc. ACM SIGGRAPH'03*, pp. 319–325, 2003.
- [13] M. Grossberg and S. Nayar. High dynamic range from multiple images: Which exposures to combine? In *Proc. ICCV Workshop on Color and Photometric Methods in Computer Vision (CPMCV)*, 2003.
- [14] J. Shi and C. Tomasi. Good features to track. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR'94)*, pp. 593–600, 1994.
- [15] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda. Photographic tone reproduction for digital images. In *Proc. ACM SIGGRAPH'02*, pp. 267–276, 2002.
- [16] T. Sato, S. Ikeda, and N. Yokoya. Extrinsic camera parameter recovery from multiple image sequences captured by an omni-directional multi-camera system. In *Proc. European Conf. on Computer Vision (ECCV'04)*, Vol. 2, pp. 326–340, 2004.

(2012年2月27日受付)

[著者紹介]

大倉 史生 (学生会員)



2009年立命館大学情報理工学部メディア情報学科中退。2011年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。現在、同大博士後期課程に在学中。複合現実感の研究に従事。電子情報通信学会、情報処理学会、IEEE、ACM各会員。

神原 誠之 (正会員)



2002年奈良先端科学技術大学院大学博士後期課程修了。同年同大情報科学研究科助教、2010年同大准教授、現在に至る。コンピュータビジョン、複合現実感の研究に従事。博士(工学)。2002年電子情報通信学会学術奨励賞受賞。FIT2005論文賞受賞。電子情報通信学会、情報処理学会、IEEE各会員。

横矢 直和 (正会員)



1974年大阪大学基礎工学部情報工学科卒。1979年同大大学院博士後期課程修了。工博。同年電子技術総合研究所入所。以来、画像処理ソフトウェア、画像データベース、コンピュータビジョンの研究に従事。1986～87年マッギル大・知能機械研究センター客員教授。1992年奈良先端科学技術大学院大学・情報科学センター教授。現在、同大情報科学研究科教授。1990年、2007年情報処理学会論文賞受賞。2005年情報処理学会フェロー、電子情報通信学会フェロー、2012年本会フェロー。情報処理学会、人工知能学会、日本認知科学会、映像情報メディア学会、IEEE、ACM SIGGRAPH各会員。