

修士論文

全方位動画像に対する欠損領域の修復による  
不可視領域のない全天球映像の生成

町北 幸大郎

2009年2月5日

奈良先端科学技術大学院大学  
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に  
修士(工学) 授与の要件として提出した修士論文である。

町北 幸大郎

審査委員：

横矢 直和 教授 (主指導教員)

木戸出 正継 教授 (副指導教員)

山澤 一誠 准教授 (副指導教員)

佐藤 智和 助教 (副指導教員)

# 全方位動画像に対する欠損領域の修復による 不可視領域のない全天球映像の生成\*

町北 幸大郎

## 内容梗概

遠隔地の実映像をユーザに提示し、あたかもその場所にいるような感覚を与えるテレプレゼンスは、娯楽、医療、教育などの分野への応用が期待されている。これらの分野ではユーザによる遠隔地の見回しを実現するために、全方位カメラによって撮影した全方位動画像を用いることが多い。しかし、一般的な全方位カメラには撮像系の構成上生じる死角が存在し、テレプレゼンスにおいて全方向の視野を完全に再現することができないため、臨場感が損なわれるという問題がある。そこで本論文では、死角により生じる欠損領域を修復し、不可視領域のない全天球映像を生成することで、臨場感の高いテレプレゼンスを実現する手法を提案する。従来、動画像上の欠損領域を修復する手法として、動画像中に存在する欠損領域周辺のテクスチャパターンと見え方が類似したパターンを事例として用いる手法が提案されているが、カメラが運動することによって欠損領域周辺のテクスチャの見え方が画像上で大きく変化する全方位動画像を対象として従来手法をそのまま用いることは難しい。また、従来手法では動画像上の全領域を事例として用いるため、実際にはその場に存在しないテクスチャパターンが生成され、違和感が生じる場合がある。これに対して、本研究では、欠損領域周辺における撮影対象の形状とカメラの運動を考慮し、テクスチャの変形補正および、事例として用いる画像領域の限定を行うことで全方位動画像に対する高品位な修復を行う。具体的には、全方位カメラの死角領域周辺の形状が平面で近似できると仮定

---

\* 奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 修士論文, NAIST-IS-MT0651114, 2009年2月5日.

し，あらかじめ Structure from motion 法によって推定された欠損領域周辺の自然特徴点の三次元位置に当てはめた平面に対して全方位画像を投影することで，カメラの運動による同一物体の見え方の変化を補償する．また，全方位カメラの位置・姿勢および，点群に当てはめた平面を用いて，欠損領域のテクスチャが他のフレームの画像上で存在する領域を特定し，事例として利用する画像領域を限定することで，実際には存在しないテクスチャが生成されることを防ぐ．実験では，全方位カメラにより撮影された動画像の欠損を修復し，不可視領域のない全天球映像を生成することで，提案手法の有効性を示す．

## キーワード

画像修復，全方位動画像，テレプレゼンス

# Generation of Omnidirectional Video without Invisible Area Using Image Inpainting\*

Kotaro Machikita

## Abstract

Telepresence that enables us to experience a remote site is expected to be used in a number of different fields such as entertainment, medical care and education. In these fields, an omnidirectional video captured with a moving omnidirectional camera is sometimes used. However, an ordinary omnidirectional camera cannot capture the entire direction of view due to the blind side. Thus, such an invisible part decreases realistic sensation in the telepresence system. In order to realize telepresence with high realistic sensation, this research aims at generating an omnidirectional video without invisible area by inpainting a missing region caused by the blind side. Although the conventional method that inpaints a missing region using the texture in a different part of the video has been proposed, it is difficult to obtain a good result when the appearance of the texture greatly changes due to camera motion. Moreover, because all the regions in a video are searched for the texture used for inpainting, inappropriate texture might be generated in the missing region. This thesis proposes a new method that successfully inpaints a missing region compensating the texture's distortion and limiting the searched area considering the camera motion and the shape of the scene around the missing region. In this research, by assuming the shape of the blind side of

---

\* Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT0651114, February 5, 2009.

the target scene is planar, the change in the appearance of the texture caused by the camera motion is compensated by projecting omnidirectional images onto the surface fitted to the 3-D positions of feature points acquired by structure from motion. Also, the quality of the inpainted image is improved by limiting the searching area using the estimated plane and camera motion. In the experiment, the effectiveness of the proposed method is demonstrated by inpainting missing regions in a real image sequence captured with an omnidirectional camera and generating an omnidirectional video without invisible area.

**Keywords:**

inpainting, omnidirectional image, telepresence

# 目次

1. はじめに	1
2. 動画像の欠損修復に関する関連研究および本研究の位置づけと方針	3
2.1 静止画像を対象とした欠損修復手法	3
2.2 動画像の欠損修復に関する従来研究	5
2.2.1 画像上の撮影対象の動きを用いない手法	5
2.2.2 画像上の撮影対象の動きを用いる手法	6
2.3 本研究の位置づけと方針	9
3. 欠損領域の修復による不可視領域のない全天球映像の生成	10
3.1 全天球映像生成手法の概要	10
3.2 全方位カメラの位置・姿勢と自然特徴点の三次元位置の推定	13
3.3 欠損領域周辺の形状の推定による平面投影画像の生成	15
3.4 エネルギー最小化による欠損領域の修復	19
3.4.1 エネルギー関数の定義	19
3.4.2 探索領域の限定	21
3.4.3 エネルギー最小化処理	22
3.5 修復画像を用いた全天球映像の生成	25
4. 実験	26
4.1 全方位動画像の欠損修復実験	26
4.1.1 入力情報の取得	26
4.1.2 平面投影画像に対する欠損修復	28
4.2 不可視領域のない全天球映像によるテレプレゼンスシステムの試作	35
5. まとめ	41
謝辞	42
参考文献	43

## 目 次

1	Criminisi ら [5] の手法を用いた画像修復の例 . . . . .	4
2	Wexler ら [10] の手法を用いた動画修復の例 . . . . .	6
3	Jia ら [18] の手法を用いた動画修復の例 . . . . .	8
4	福地ら [22] の手法を用いた画像修復の例 . . . . .	8
5	死角により生じた欠損領域を持つ全天球パノラマ画像の例 . . . . .	10
6	不可視領域のない全天球映像生成処理の流れ . . . . .	11
7	欠損領域周辺の特徴点の抽出 . . . . .	16
8	最小二乗法による平面の導出 . . . . .	16
9	平面投影画像の生成 . . . . .	17
10	画像上の各領域 . . . . .	20
11	探索領域の限定 . . . . .	23
12	全方位マルチカメラシステム Ladybug . . . . .	26
13	入力に用いた全方位動画の 1 フレーム . . . . .	27
14	推定されたカメラ位置・姿勢および自然特徴点の三次元位置 . . . . .	27
15	全方位動画の中間フレームに対して生成された平面投影画像 (第 91 フレーム) . . . . .	28
16	全方位動画から生成した平面投影画像 . . . . .	29
17	平面を利用した画素値の投影による修復結果と提案手法による修 復結果の比較 . . . . .	31
18	連続フレームに対する平面投影画像 . . . . .	33
19	連続フレームに対する修復結果 . . . . .	34
20	欠損領域を持つ全天球パノラマ画像 . . . . .	35
21	提案手法による平面投影画像の修復結果 . . . . .	36
22	修復された全天球パノラマ画像 . . . . .	36
23	全天球映像を用いた見回し例 1 . . . . .	37
24	全天球映像を用いた見回し例 2 . . . . .	38
25	全天球映像を用いた見回し例 3 . . . . .	39

## 1. はじめに

遠隔地の実映像をユーザに提示し、あたかもその場所にいるような感覚を与えるテレプレゼンスは、娯楽、医療、教育などの分野への応用が期待されている。これらの分野ではユーザが自由に遠隔地を見回すことで臨場感を高めるために、全方位カメラによって撮影した全方位動画像を用いることが多い [1, 2, 3]。しかし、一般的な全方位カメラには撮像系の構成上生じる死角が存在し、テレプレゼンスにおいて全方向の視野を完全に再現することができないため、臨場感が損なわれるという問題がある。そこで本論文では、死角により生じる欠損領域を修復し、不可視領域のない全天球映像を生成することで、臨場感の高いテレプレゼンスを実現する手法を提案する。

従来、画像中の欠損領域修復に関して、静止画像を対象とした欠損修復手法が多く提案されている [4, 5, 6, 7, 8, 9]。これらの手法には、輝度値の滑らかさを保存する手法や同一画像中の欠損領域以外のテクスチャを事例として利用する手法等が存在し、静止画像に対して違和感の少ない修復を実現している。また、これらの静止画像に対する手法を動画像の各フレームに対して適用することで、動画像中の欠損領域を修復することもできる。しかし、静止画を対象とする手法は、単一フレーム内の情報のみを用いて修復を行うため、単一フレームにおいて違和感なく欠損領域が修復された場合においても、各フレームで独立して実際には存在しないテクスチャが生成される。従って、静止画を対象とする手法によって修復された画像群を動画像として見た場合にはテクスチャが不連続に変化し、違和感が生じる場合が多い。このため、動画像の欠損修復手法においては、生成される画像の時間的な連続性を考慮して修復することが一般的である。これらの研究は、画像上の撮影対象の動きを用いない手法 [10, 11, 12, 13] と、動きを用いる手法 [14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24] に大別される。前者には、時空間における類似パターンを事例として用いて修復を行う手法や、画素値が滑らかに変化するという前提で修復を行う手法が存在する。また、後者は、撮影対象やカメラの動きに基づいて、修復に適したテクスチャを取得する領域を決定し、修復を行う。これら動画像を対象とした手法には、時間的に不連続なテクスチャが生成されにくいという特長がある。しかし、従来手法では、視点または撮影対象の移動

による撮影画像上でのテクスチャの変形を考慮していないため、カメラの移動に伴って撮影対象の見え方が大きく変化する全方位動画像を対象とした場合には、高品位な修復を実現することは難しい。

この問題に対して、本論文では全方位画像上における欠損領域周辺の撮影対象の形状と全方位カメラの運動を考慮し、テクスチャの変形補正および、修復の事例として用いる画像領域の限定を行うことで全方位動画像に対する高品位な修復を行う。具体的には、全方位カメラの死角に当たる領域に存在する物体の形状が平面で近似できると仮定し、あらかじめ Structure from motion 法によって推定された欠損領域周辺の自然特徴点の三次元位置に当てはめた平面に対して全方位画像を投影することで、カメラの運動による同一物体の見え方の変化を補償する。また、各フレームに対する全方位カメラの位置・姿勢情報と点群に当てはめた平面を用いて、カメラの死角に存在する平面領域が他のフレームの画像上で存在する領域を特定し、これによって事例として利用する画像領域を限定することで、実際には存在しないテクスチャが生成されることを防ぐとともに動画像上での不連続なテクスチャの変化を抑制する。

以降、2章では、動画像の欠損修復に関する関連研究および本研究の位置づけについて述べる。3章では、カメラの動き情報と死角領域の形状に基づいて欠損領域を修復することで不可視領域のない全天球映像の生成を行う提案手法について述べる。4章では、全方位動画像の欠損領域を修復し、不可視領域のない全天球映像によるテレプレゼンスを試作することで、提案手法の有効性を示す。最後に5章で、本論文のまとめと今後の課題について述べる。

## 2. 動画像の欠損修復に関する関連研究および本研究の位置づけと方針

動画像中の不要な物体を欠損として指定し、それを修復することで不要な物体を動画像中から違和感なく消去する動画像の欠損修復に関する研究が盛んに行われている。本章では、まず静止画像を対象とした欠損修復手法および、動画像の欠損修復に関する従来研究について述べ、最後に従来研究と本研究との位置づけについて述べる。

### 2.1 静止画像を対象とした欠損修復手法

静止画像を対象とした欠損修復手法は、同一画像中の欠損領域以外の情報を用いて欠損領域の修復を行う。この手法は、輝度値の連続性を用いる手法 [4, 9]、特徴空間での補間による手法 [8]、同一画像中のテクスチャを事例として用いる手法 [5, 6, 7] に大別でき、これらの手法を動画像の各フレームに対して適用することで、動画像中の欠損領域を修復することもできる。

輝度値の連続性を用いた手法 [4, 9] は、エッジのつながりと輝度値の連続性を考慮し、欠損領域内の画素値を繰り返し更新することで修復を行う。この手法は、引っかき傷やテロップのような小さい欠損領域に対しては違和感のない修復ができる。しかし、細かいテクスチャを欠損領域内に再現できないため、欠損領域周辺に複雑なテクスチャが存在する場合には、違和感が生じる場合が多い。一方、特徴空間を用いる手法 [8] は欠損領域を含む領域の特徴量を算出し、その特徴量に近い同一画像内の複数のテクスチャから分解した固有ベクトルを組み合わせ、欠損領域内にテクスチャを合成することで修復を行う。また、同一画像中のテクスチャを事例として用いる手法 [5, 6, 7] は、テクスチャの逐次的合成やテクスチャの欠損領域内での最適化により修復を行う。これらの手法は、図 1 に示すように、欠損領域内に複雑なテクスチャを生成することができるため、多くの静止画像に対して違和感のない修復を行うことができる。

これら静止画像を対象とした修復手法を用いれば、多くの静止画像に対して良好な修復結果を得ることができる。しかし、これらの手法はいずれも単一フレー



(a) 入力画像



(b) 出力画像

図 1 Criminisi ら [5] の手法を用いた画像修復の例

ム内の情報のみを用いて修復を行うため、現実には存在しないテクスチャが合成されており、動画像を対象とした場合には各フレームで全く異なるテクスチャが独立に生成されることが多い。従って、生成された画像群を動画像として見た場合には、欠損領域内のテクスチャが不連続に変化し、違和感が生じるという問題がある。

## 2.2 動画像の欠損修復に関する従来研究

本節では、動画像を対象とした欠損修復に関する従来研究を、画像上の撮影対象の動きを用いない手法、画像上での撮影対象の動きを用いる手法に分類し、それぞれの特徴と問題点について述べる。

### 2.2.1 画像上の撮影対象の動きを用いない手法

画像上の撮影対象の動きを用いない手法は、時空間における輝度値の連続性を考慮する手法 [13] と、動画像中の時空間のボリュームを事例として用いる手法 [10, 11, 12] に分類できる。時空間における輝度値の連続性を考慮する手法 [13] は、偏微分方程式を用いることで同一フレーム内およびフレーム間で輝度値が滑らかに変化するように欠損を修復する。しかし、この手法は原理的に複雑なテクスチャを欠損領域内に再現できないため、欠損領域が大きい場合には細かいテクスチャを欠損領域内に再現できず、違和感が生じてしまう。これに対して、動画像中の欠損領域以外の時空間ボリュームを事例として用いる手法 [10, 11, 12] は、欠損領域と動画像中のその他の領域において、画像空間と時間軸から成る局所的な 3 次元時空間ボリュームを定義し、欠損領域周辺と類似した時空間ボリュームを欠損領域にコピーまた類似した時空間ボリュームを用いて欠損領域内を最適化することで欠損を修復する。これらの手法は、図 2 に示すように、欠損領域内に複雑なテクスチャを再現でき、また時間軸におけるテクスチャの不連続も抑制されるため、違和感の小さい修復結果を得ることができる。しかし、これらの手法は、ボリューム間の類似度のみを尺度として欠損の修復を行うため、実際に不要物体の裏側に存在する現実物体のテクスチャを再現している訳ではない。従って、実際には存在しないテクスチャが欠損領域内に生成される場合がある。また、時空間ボリュームの変形は考慮されていないため、カメラの動きによる動画像上での物体の見え方の変化には対応できず、全方位カメラでのように、同一物体の画像上での見え方が視点位置によって大きく変化する場合には、良好な結果を得ることは難しい。



(a) 入力フレーム



(b) 出力フレーム

図 2 Wexler ら [10] の手法を用いた動画像修復の例

### 2.2.2 画像上の撮影対象の動きを用いる手法

画像上での撮影対象の動きを用いる手法は、複数フレーム間における画像上でのオブジェクトの動きまたはカメラの動きを推定し、欠損領域内での物体の動きを予測することで、修復に用いるテクスチャの領域を他フレームの画像上で限定することで修復を行う。これらの手法は、画像上でのオブジェクトの見かけの動きを推定し修復する手法 [14, 15, 16, 17, 18, 20, 21, 23]、カメラの位置・姿勢情報を用いて修復する手法 [22, 24] 分類できる。

#### [オブジェクトの見かけの動きを用いる手法]

オブジェクトの見かけの動きを推定し修復に用いる手法として、Jia ら [17, 18], Zhang ら [16], Garg ら [21] は、複数のフレーム間において、運動するオブジェクトを検出し、その動き情報を欠損領域内に補完することで修復を行う手法を提案している。図 3 に Jia ら [18] の手法を用いた画像修復の例を示す。この手法は、欠損領域に存在するテクスチャを再現することが可能である。しかし、この手法ではオブジェクトの運動をモデル化した上で画像を修復するため、その動きが周期運動もしくは単純な動きに限定されるといった制約がある。また、画像上の大半

の領域が画像上で静止していることを前提として運動物体を抽出する必要があり、カメラが運動する場合に利用することができない。これに対して、Matsushitaら [14] は、物体の陽な運動モデルを仮定せず、画像全体に対するオプティカルフローに基づいて修復を行うことで、カメラが運動する場合にも適用が可能な手法を提案している。しかし、2次元画像上での動きを用いる手法は、物体の見かけの動きを画像平面上で推定するため、奥行きが変化するような物体の動きに対応できないという問題がある。また、投影歪みによって発生する、視点位置の変化に伴うテクスチャの変形を考慮していないため、カメラが大きく動く条件で撮影された動画像に対しては上記の手法を適用することが難しい。

#### **[カメラの位置・姿勢情報を用いる手法]**

福地ら [22] は、パン・チルト固定カメラにおいて撮影された動画像を用いて、水滴やキズなどの視野妨害ノイズを除去する手法を提案した。具体的には、投影中心が固定された一定速度で回転する首振りカメラを用いて撮影を行い、得られた時空間断面画像上での軌跡に基づいて視野妨害ノイズを検出し、カメラの姿勢変化量に基づいて他フレームの画像上の対応する領域からテクスチャをコピーすることで画像修復手法を行っている。図4に福地ら [22] の手法を用いた画像修復の例を示す。また、Patwardhanら [24] は、カメラの動きがカメラの光軸方向に対して直行することを前提として、欠損修復を行う手法を提案した。しかし、前述した姿勢変化を考慮した手法を含め、自由なカメラの動きに対応できないという問題があり、移動撮影した全方位動画像に対する欠損修復に利用することは難しい。

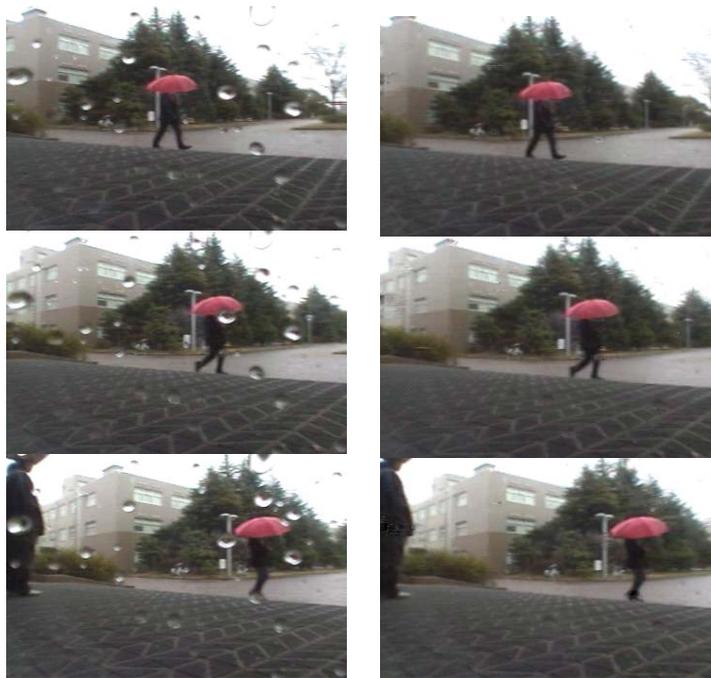


(a) 入力フレーム



(b) 出力フレーム

図 3 Jia ら [18] の手法を用いた動画像修復の例



(a) 入力フレーム

(b) 出力フレーム

図 4 福地ら [22] の手法を用いた画像修復の例

## 2.3 本研究の位置づけと方針

前節までに概観したように，動画像に適用可能な欠損領域の修復手法としてすでに多くの手法が提案されている．静止画像を対象とした手法は，動画像の各フレームに処理を適用することで動画像の欠損修復に利用することが可能であるが，単一フレーム内の情報のみを用いて修復を行うため，修復された画像群を動画像としてみた場合には，欠損領域のテクスチャが不連続に変化し，違和感が生じるという問題がある．これに対して，動画像上で時系列情報を用いて修復を行う手法が多数提案されている．画像上の撮影対象の動きを用いない手法は，テクスチャの不連続が生じにくく，違和感の小さい修復結果を得ることができるという特長を持つ．しかし，欠損領域が本来そこに存在しないテクスチャで修復される可能性がある．また，撮影対象の視点変化による投影歪みを考慮していないため，視点の大きく変化する全方位カメラで撮影された全方位動画像においては適用が難しい．撮影対象の動きやカメラの動きを用いる手法は，欠損となっている領域内に実際に存在すると考えられるテクスチャを用いて修復を行うため，欠損領域に存在するテクスチャを再現することが可能である．しかし，動きを用いない手法と同様に，視点位置の変化における物体の投影歪みを考慮していないため，全方位カメラで移動撮影された動画像に対して適用することは困難である．

本論文では，欠損領域周辺における撮影対象の形状とカメラの運動を考慮し，テクスチャの変形補正および，事例として用いる画像領域の限定を行い，欠損領域と限定した画像領域のテクスチャの類似度に基づくエネルギーの最小化による欠損領域の修復を行うことで，移動を伴って撮影された全方位動画像に対する高品位な修復を行う手法を提案する．本研究では，欠損領域周辺の形状が平面で近似できると仮定し，あらかじめ Structure from motion 法 [25] によって推定された欠損領域周辺の自然特徴点の三次元位置に当てはめた平面に対して全方位画像を投影することで，カメラの運動による同一物体の見え方の変化を補償する．また，全方位カメラの位置・姿勢および，点群に当てはめた平面を用いて，欠損領域のテクスチャが他のフレームの画像上で存在する領域を特定し，事例として利用する画像領域を限定することで，実際には存在しないテクスチャが生成されることを防ぐ．

### 3. 欠損領域の修復による不可視領域のない全天球映像の生成

#### 3.1 全天球映像生成手法の概要

全方位カメラで撮影された動画像を用いて作成された全天球映像には、図5に示すようにカメラの撮像系の構成上生じる死角による欠損領域が存在する。そこで本論文では、エネルギー最小化の枠組みを用いて欠損領域を修復することで、不可視領域のない全天球映像を生成する手法を提案する。なお、本研究では、全方位カメラの中でも複数のカメラユニットを放射状に配置することで一つのカメラブロックを構成し、一度に全方位を撮影する全方位型マルチカメラシステムを用いることを想定し、加えて死角領域は地面の方向に存在するものとする。また、欠損領域はあらかじめ手動で与えるものとする。提案手法の処理の流れを図6に示す。本研究では、まず Structure from Motion 法を用いることで、全方位カメラの位置・姿勢と自然特徴点の三次元位置を推定をする (a)。次に、カメラの位



図5 死角により生じた欠損領域を持つ全天球パノラマ画像の例

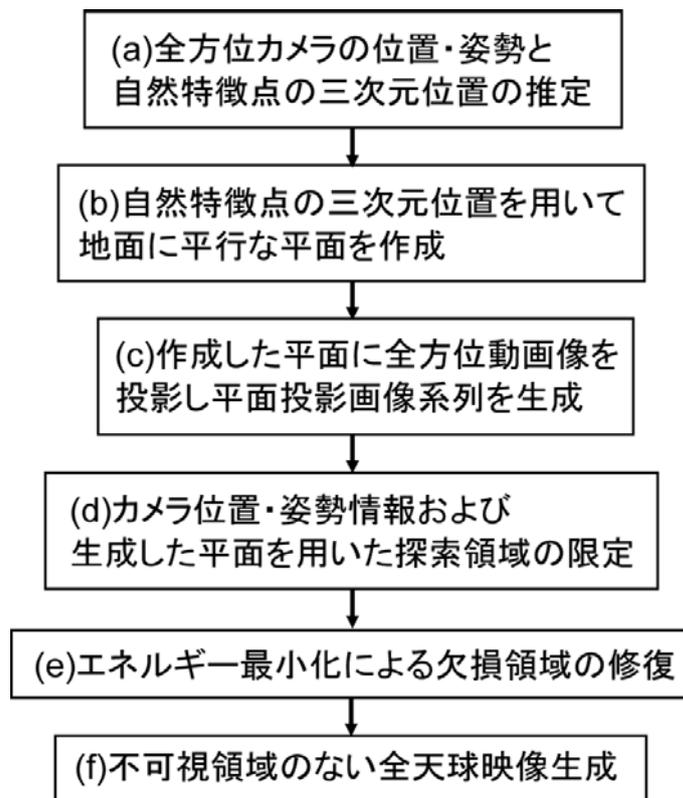


図 6 不可視領域のない全天球映像生成処理の流れ

置・姿勢情報と自然特徴点の三次元位置を用いてフレームごとに地面に対応する三次元点群に平面を当てはめ (b), 当てはめた平面に全方位動画像を投影することで平面投影画像系列を生成する (c). 続いて, カメラ位置・姿勢情報および生成した平面を用いることで修復に利用するテクスチャが存在するデータ領域を限定する (d). 最後に, 各フレームに対して順次, 欠損領域と限定したデータ領域のテクスチャの類似度に基づくエネルギーを最小化することにより欠損領域の修復を行い (e), 修復画像を全天球パノラマ画像に投影することで, 不可視領域のない全天球映像を生成する (f). 以上の処理を複数フレームに対して適用することで動画像の欠損修復を行う.

以下, 3.2 節では全方位カメラの位置・姿勢と自然特徴点の三次元位置の推定について述べ, 3.3 節で欠損領域周辺形状の推定による平面投影画像の生成について述べる. また, 3.4 節でエネルギー関数の定義および探索範囲の限定によるエネルギー最小化の手法について述べ, 3.5 節で修復画像を用いた全天球映像の生成について述べる.

### 3.2 全方位カメラの位置・姿勢と自然特徴点の三次元位置の推定

本研究では、全方位カメラの位置・姿勢および自然特徴点の三次元位置を、全方位カメラに対する Structure from Motion 法 [25] を用いて推定する。この手法では、自然特徴点を全方位動画像上で追跡することにより、撮影時の全方位カメラの外部パラメータと自然特徴点の三次元位置を推定する。以下に、手法 [25] による推定処理の手順を示す。

- 特徴点の追跡

Harris インタレストオペレータの出力が極大値を取る座標を自然特徴点の候補位置として利用し、テンプレートマッチングにより前フレームの各特徴点について、現フレームにおける対応点を探索する。

- カメラ位置・姿勢の推定

特徴点の画像上の二次元座標と、特徴点の世界座標における三次元座標を用いてカメラの外部パラメータを推定する。本ステップでは、前ステップにおいて検出された特徴点の各画像面への再投影誤差の和を最小化することで、カメラ位置・姿勢を推定する。

- 特徴点の三次元位置の推定

追跡されている全ての自然特徴点について、その特徴点が投影されている画像面への再投影誤差の和を最小化することで三次元座標を算出し、三次元位置をフレームごとに推定・更新する。

- 特徴点の追加と削除

特徴点の信頼度や誤差などの複数の評価尺度を用いて、自然特徴点の追加・削除を行う。

上記のステップを, 初期フレームから最終フレームまで繰り返すことで, 全方位動画のカメラ位置・姿勢および自然特徴点の三次元位置の初期値を推定する. 最後に, 動画全体での最適化処理 (bundle adjustment) を行うことで, 再投影誤差を最小化し, カメラ位置・姿勢と自然特徴点の三次元位置の推定結果に含まれる誤差を最小化する.

### 3.3 欠損領域周辺の形状の推定による平面投影画像の生成

一般的な全方位動画像においては、カメラの移動に伴って撮影対象の見え方が大きく変化するため、撮影されたそのままのテクスチャを用いて高品位な修復を行うことは難しい。そこで本研究では、全方位動画像上における欠損領域に当たる地面の形状が平面であると仮定し、地面の領域に存在すると考えられる特徴点に対して平面を当てはめた上で、その平面に全方位動画像を投影することで視点の変化による投影歪みを補償した平面投影画像群を作成する。

具体的には、まず全方位動画像に対する Structure from Motion 法によって得られた三次元点群から、図7に示すように、カメラを中心とした半径  $l$  の球内に存在し、かつ世界座標系における特徴点の高さ  $z$  が、 $(p < z < p + m)$  の範囲内に存在するものを、欠損領域周辺の地面領域に対応する点群として抽出する。次に、世界座標系における地面を表す平面の式を  $z = ax + by + c$  とおき、抽出された特徴点  $(x_i, y_i, z_i)$  と平面との  $z$  軸に沿った二乗距離の総和による以下のコスト関数  $L$  が最小となるように、最小二乗法を用いて平面の式のパラメータ  $(a, b, c)$  を決定する (図8参照)。

$$L = \sum_{i=1}^n (ax_i + by_i + c - z_i)^2 \quad (1)$$

ただし、 $n$  は抽出された特徴点の個数である。最後に、図9に示すように、生成した平面に対して全方位動画像を投影することで平面投影画像系列を作成する。またここでは、マルチカメラシステムにおける代表カメラの投影中心を通り鉛直下方向に向かう基底ベクトル方向の光線と生成した平面の交点を、平面投影画像の中心とすることで、欠損領域が常に平面投影画像の中心となるようにする。具体的には、代表カメラに対する第  $f$  フレームのカメラ座標系から世界座標系への変換行列が

$$\mathbf{M}_f = \begin{bmatrix} m_{f1} & m_{f2} & m_{f3} & m_{f4} \\ m_{f5} & m_{f6} & m_{f7} & m_{f8} \\ m_{f9} & m_{f10} & m_{f11} & m_{f12} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

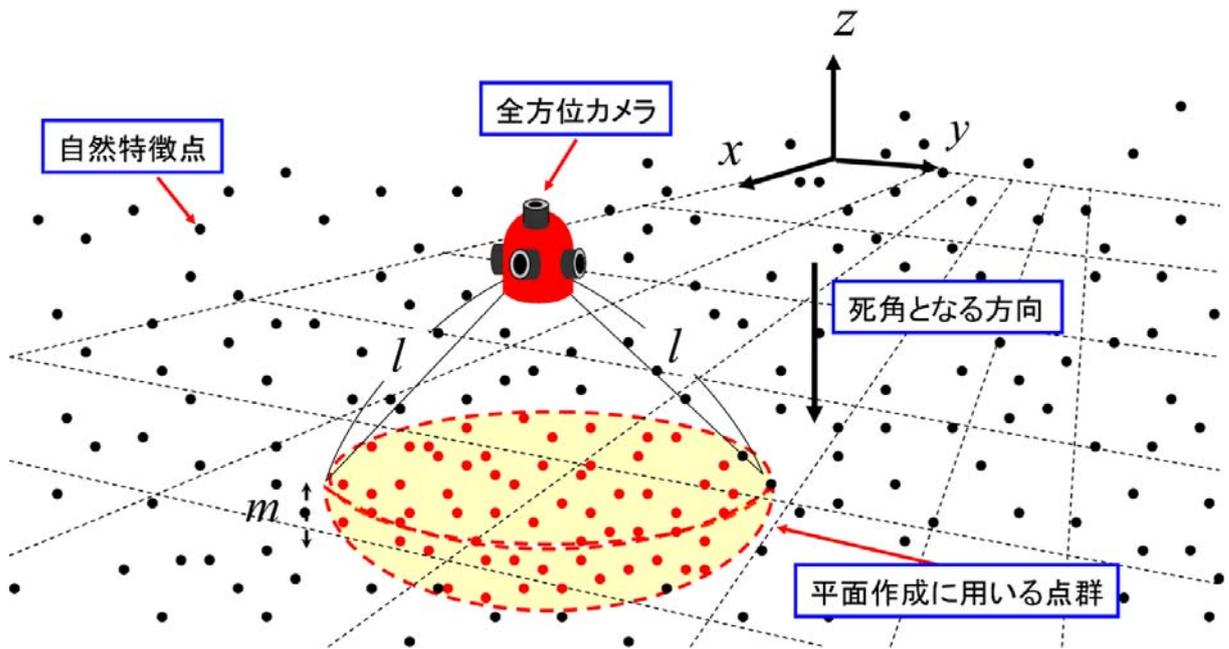


図 7 欠損領域周辺の特徴点の抽出

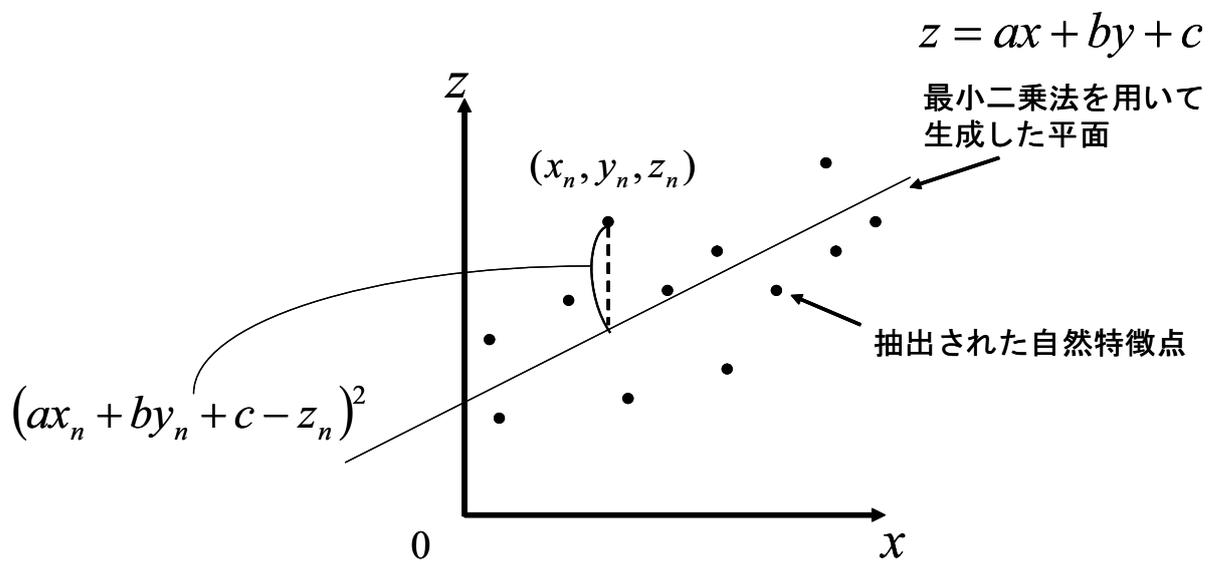


図 8 最小二乗法による平面の導出

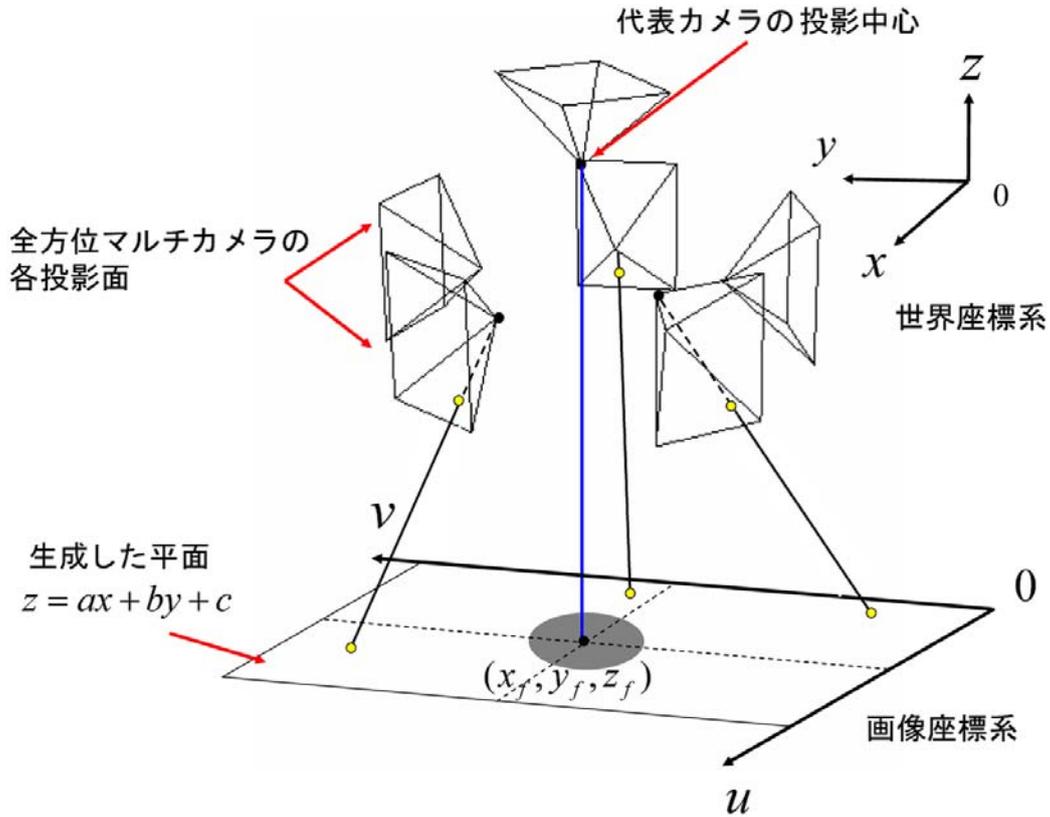


図 9 平面投影画像の生成

で表されるとき，平面投影画像の中心の世界座標  $(x_f, y_f, z_f)$  を以下の式から算出する。

$$\begin{bmatrix} x_f \\ y_f \\ z_f \end{bmatrix} = \begin{bmatrix} m_{f4} \\ m_{f8} \\ m_{f12} \end{bmatrix} + t \begin{bmatrix} m_{f2} \\ m_{f6} \\ m_{f10} \end{bmatrix} \quad (3)$$

ただし， $t$  は媒介変数である。式 (3) と第  $f$  フレームに関する平面の式  $z = a_f x + b_f y + c_f$  から  $t$  を消去すれば，以下の通り  $(x_f, y_f, z_f)$  を算出できる。

$$\begin{cases} x_f = m_{f4} + \frac{m_{f2}(a_f m_{f4} + b_f m_{f8} + c_f - m_{f12})}{m_{f10} - a_f m_{f2} - b_f m_{f6}} \\ y_f = m_{f8} + \frac{m_{f6}(a_f m_{f4} + b_f m_{f8} + c_f - m_{f12})}{m_{f10} - a_f m_{f2} - b_f m_{f6}} \\ z_f = a_f x_f + b_f y_f + c_f \end{cases} \quad (4)$$

また，修復に用いるテクスチャの回転を抑止するために，生成する画像平面の軸の向きを統一する．ここでは，世界座標系の基底ベクトルを  $(\mathbf{x}, \mathbf{y}, \mathbf{z})$  とし， $\mathbf{x}$ - $\mathbf{y}$  平面が重力方向  $\mathbf{z}$  に垂交する平面であるとしたとき，画像平面の基底ベクトル  $\mathbf{u}$  と世界座標系の基底ベクトル  $\mathbf{y}$  が直行するものとする．このとき，以下の式が成り立つ．

$$\mathbf{u} \cdot \mathbf{y} = 0 \quad (5)$$

式(4)と平面を表す式  $z = ax + by + c$  より，世界座標系における画像平面の基底ベクトル  $(\mathbf{u}, \mathbf{v})$  は，以下のように算出される．

$$\mathbf{u} = \begin{bmatrix} \frac{1}{\sqrt{1+a^2}} \\ 0 \\ \frac{a}{\sqrt{1+a^2}} \end{bmatrix} \quad (6)$$

$$\mathbf{v} = \begin{bmatrix} \frac{-ab}{\sqrt{(1+a^2)(1+a^2+b^2)}} \\ \frac{1+a^2}{\sqrt{(1+a^2)(1+a^2+b^2)}} \\ \frac{b}{\sqrt{(1+a^2)(1+a^2+b^2)}} \end{bmatrix} \quad (7)$$

このとき， $\mathbf{u}$ - $\mathbf{v}$  座標系における画素の座標  $(u, v)$  と世界座標系での三次元位置  $(x_{uv}, y_{uv}, z_{uv})$  の関係は，平面投影画像上での画素の大きさを  $p \times p$ ，生成する画像の解像度を  $S \times T$  画素としたとき，以下のように表される．

$$\begin{cases} x_{uv} = x_f + \frac{up}{\sqrt{1+a^2}} - \frac{Sp}{2\sqrt{1+a^2}} \\ y_{uv} = y_f + \frac{(1+a^2)vp}{\sqrt{(1+a^2)(1+a^2+b^2)}} - \frac{(1+a^2)Tp}{2\sqrt{(1+a^2)(1+a^2+b^2)}} \\ z_{uv} = ax_{uv} + by_{uv} + c \end{cases} \quad (8)$$

### 3.4 エネルギー最小化による欠損領域の修復

本研究では、前節で生成した欠損領域を持つ平面投影画像に対して、フレームごとに順次パターン類似度に基づくエネルギーを最小化することで欠損領域を修復する。以下ではエネルギー関数の定義、探索領域の限定、エネルギー最小化処理について順に述べる。

#### 3.4.1 エネルギー関数の定義

本研究では、図 10 に示す、修復対象フレームにおける欠損領域  $\Omega$  を含む領域  $\Omega'$  と、データ取得用フレームにおけるデータ領域  $\Phi_i$  のテクスチャの類似度に基づくエネルギー関数を修復に用いる。ここでは、欠損領域  $\Omega$  内の画素を中心とした一定サイズの正方ウインドウ  $W$  が一部でも含まれる画素の集合を  $\Omega'$  とする。修復に用いる欠損領域の尤もらしさに基づくエネルギー  $E$  は、領域  $\Omega'$  内の画素  $\mathbf{x}_i$  とそれに対応するデータ領域  $\Phi_i$  内の画素  $t(\mathbf{x}_i)$  周辺のパターン類似度 SSD (Sum of Squared Differences) の重み付き総和として以下のように定義する。

$$E = \sum_{\mathbf{x}_i \in \Omega'} w_{\mathbf{x}_i} SSD(\mathbf{x}_i, t(\mathbf{x}_i)) \quad (9)$$

$$SSD(\mathbf{x}_i, t(\mathbf{x}_i)) = \sum_{\mathbf{p} \in W} \{I_f(\mathbf{x}_i + \mathbf{p}) - I_k(t(\mathbf{x}_i) + \mathbf{p})\}^2 \quad (10)$$

ただし、 $I_i(\mathbf{x}_i)$  は第  $i$  フレームの画素  $\mathbf{x}_i$  の画素値を表す。また、 $f$  は欠損修復の対象となるフレームの番号を、 $k$  はデータ領域として用いる画像のフレーム番号を表す。ここでは重み  $w_{\mathbf{x}_i}$  として、欠損周辺の領域  $\Omega' \cap \bar{\Omega}$  では各画素の値が固定値となるため  $w_{\mathbf{x}_i} = 1$  を、欠損領域  $\Omega$  内では境界に近いほど画素値の信頼度が高くなるため、 $\Omega$  の境界からの距離  $d$ 、定数  $g$  を用いて  $w_{\mathbf{x}_i} = g^{-d}$  とする。本研究では、動画像上の一定の探索範囲内において、上記の SSD の総和  $E$  が最小となる類似パターン位置  $t(\mathbf{x}_i)$  を決定し、欠損領域内での画素値  $I_f(\mathbf{x}_i)$  を更新することで、欠損領域の修復を行う。

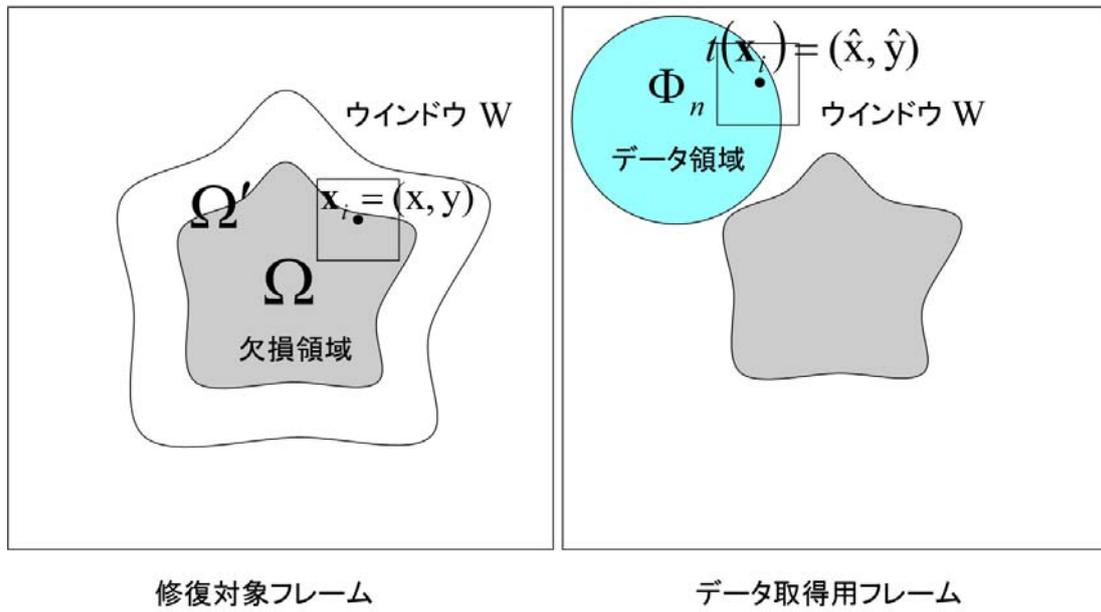


図 10 画像上の各領域

### 3.4.2 探索領域の限定

ここでは、次項で述べる画素ごとの類似パターン位置  $t(\mathbf{x}_i)$  を決定するプロセスにおいて、3.3節で述べた手法で推定された平面を用いることで、 $t(\mathbf{x}_i)$  の探索領域を限定する手法について述べる。図 11 に以下で述べる探索領域の限定の概要を示す。本研究では移動撮影を前提とするため修復の対象となるフレーム  $f$  において、欠損領域内に本来存在するべきテクスチャは、 $f$  から一定フレーム離れた第  $k(k = f + h)$  フレームの画像上に存在することが多い。また、第  $f$  フレームと第  $k$  フレームのカメラパラメータおよびそれぞれのフレームに対する地面の形状は既知である。そこで、本研究ではこれらの幾何学的な関係から第  $k$  フレームの平面投影画像上における類似パターン位置  $t(\mathbf{x}_i)$  の探索領域を限定する。

具体的にはまず、第  $f$  フレームに対して作成された平面投影画像上における欠損領域内の画素を、第  $k$  フレームにおけるカメラの画像上へ投影する。次に画素とカメラの投影中心を結ぶ光線と第  $k$  フレームにおける平面投影画像との交点を算出し、探索領域の中心とする。これを欠損領域  $\Omega$  内の全ての画素について行い、各画素ごとに探索領域を決定する。すなわち、第  $k$  フレームのカメラ座標系から世界座標系への変換行列が

$$\mathbf{M}_k = \begin{bmatrix} m_{k1} & m_{k2} & m_{k3} & m_{k4} \\ m_{k5} & m_{k6} & m_{k7} & m_{k8} \\ m_{k9} & m_{k10} & m_{k11} & m_{k12} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11)$$

で表されるとき、第  $f$  フレームにおける欠損領域上の画素  $\mathbf{x}_i(x_{fuv}, y_{fuv}, z_{fuv})$  と第  $k$  フレームのカメラの投影中心とを結ぶ直線は、以下のように表される。

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_{fuv} \\ y_{fuv} \\ z_{fuv} \end{bmatrix} + t \begin{bmatrix} m_{k4} - x_{fuv} \\ m_{k8} - y_{fuv} \\ m_{k12} - z_{fuv} \end{bmatrix} \quad (12)$$

ただし、 $t$  は媒介変数である。この直線と第  $k$  フレームの平面  $z = a_k x + b_k y + c_k$

との交点の座標  $(x_{kuv}, y_{kuv}, z_{kuv})$  は、以下のように求まる.

$$\begin{cases} x_{kuv} = \frac{(m_{k4}-x_{fuv})(m_{k4}a_k+m_{k8}b_k+c_k-m_{k12})}{m_{k12}-z_{fuv}-a_k m_{k4}+a_k x_{fuv}-b_k m_{k8}+b_k y_{fuv}} + m_{k4} \\ y_{kuv} = \frac{(m_{k8}-y_{fuv})(m_{k4}a_k+m_{k8}b_k+c_k-m_{k12})}{m_{k12}-z_{fuv}-a_k m_{k4}+a_k x_{fuv}-b_k m_{k8}+b_k y_{fuv}} + m_{k8} \\ z_{kuv} = a_k x_{kuv} + b_k y_{kuv} + c_k \end{cases} \quad (13)$$

式(8)より, 交点の三次元座標  $(x_{kuv}, y_{kuv}, z_{kuv})$  に対応する第  $k$  フレームの平面投影画像上での  $uv$  座標  $(u_k, v_k)$  は,

$$\begin{cases} u_k = \frac{\sqrt{1+a_k^2}(x_{kuv}-x_k)}{p} + \frac{S}{2} \\ v_k = \frac{\sqrt{1+a_k^2+b_k^2}}{p\sqrt{1+a_k^2}}(y_{kuv}-y_k) + \frac{T}{2} \end{cases} \quad (14)$$

となる. ただし

$$\begin{cases} x_k = m_{k4} + \frac{m_{k2}(a_k m_{k4} + b_k m_{k8} + c_k - m_{k12})}{m_{k10} - a_k m_{k2} - b_k m_{k6}} \\ y_k = m_{k8} + \frac{m_{k6}(a_k m_{k4} + b_k m_{k8} + c_k - m_{k12})}{m_{k10} - a_k m_{k2} - b_k m_{k6}} \end{cases} \quad (15)$$

である. このときの画素  $(u_k, v_k)$  を中心とした一定範囲  $W$  内の領域を次項で述べるエネルギー最小化処理で用いる類似パターンの探索範囲 (データ領域)  $\Phi_k$  とする.

### 3.4.3 エネルギー最小化処理

本研究では, Greedy Algorithm の枠組みを用いて, 式(1)で定義したエネルギー関数を最小化する [6]. ここでは, 全ての類似パターンの組  $(\mathbf{x}_i, t(\mathbf{x}_i))$  を固定すると, エネルギー  $E$  を欠損領域  $\Omega$  内の各画素で独立に扱えることに着目し,

(I) 類似パターン位置  $t(\mathbf{x}_i)$  の探索

(II) 欠損領域  $\Omega$  内の各画素値  $I_f(\mathbf{x}_i)$  の更新

をエネルギーが収束するまで繰り返すことで, エネルギーを最小化し, 欠損領域を修復する.

**プロセス (I):類似パターン位置  $t(\mathbf{x}_i)$  の探索**

3.4.2 項で述べた手法により欠損領域内の画素  $\mathbf{x}_i$  に対して限定された探索範囲  $\Phi_k$  上において, 以下の式を満たす SSD が最小となる位置を算出し, エネルギーの算

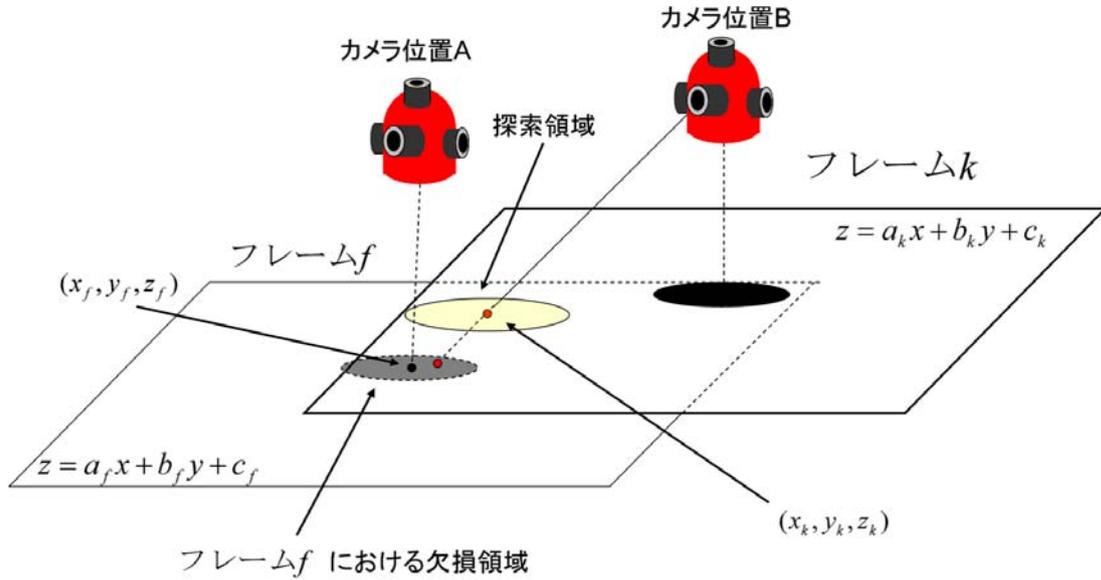


図 11 探索領域の限定

出に用いる類似パターンの位置  $t(\mathbf{x}_i)$  とする。

$$t(\mathbf{x}_i) = \underset{t(\mathbf{x}_i) \in \Phi_k}{\operatorname{argmin}} SSD(\mathbf{x}_i, t(\mathbf{x}_i)) \quad (16)$$

### プロセス (II): 欠損領域 $\Omega$ 内の各画素値 $I(\mathbf{x}_i)$ の更新

類似パターンの組を固定し、式 (1) で定義したエネルギー  $E$  を最小化する欠損領域内の画素値  $I_f(\mathbf{x}_i)$  を更新する。以下では、パターンの組を固定した場合の画素値  $I_f(\mathbf{x}_i)$  の算出法について述べる。ここではまず、エネルギー  $E$  を、欠損領域内の各画素の要素エネルギー  $E(\mathbf{x}_i)$  に分解する。更新対象となる欠損領域内  $\Omega$  内の画素の位置  $\mathbf{x}_i$  を中心とするウインドウ  $W$  内の任意の点を  $\mathbf{x}_i + \mathbf{p}$  ( $\mathbf{p} \in W$ ) とする。このとき、画素  $\mathbf{x}_i + \mathbf{p}$  を中心とするパターンに対してプロセス (I) で求まる類似パターンの位置は  $t(\mathbf{x}_i + \mathbf{p})$  であり、この類似パターン上において  $t(\mathbf{x}_i)$  と対応する画素の位置は  $f(\mathbf{x}_i + \mathbf{p}) - \mathbf{p}$  となる。ここで、注目画素  $\mathbf{x}_i$  に関する  $E$  の要素エネルギー  $E(\mathbf{x}_i)$  は、 $\mathbf{x}_i$  と  $f(\mathbf{x}_i + \mathbf{p}) - \mathbf{p}$  の関係から以下のように表すことがで

きる.

$$E(\mathbf{x}_i) = \sum_{\mathbf{p} \in W} w_{\mathbf{x}_i + \mathbf{p}} \{I_f(\mathbf{x}_i) - I_k(t(\mathbf{x}_i + \mathbf{p}) - \mathbf{p})\}^2 \quad (17)$$

このとき、欠損領域全体のエネルギー  $E$  と各画素における要素エネルギー  $E(\mathbf{x}_i)$  の関係は、以下のように表せる.

$$E = \sum_{i=1}^{N_\Omega} E(\mathbf{x}_i) + C \quad (18)$$

$C$  は、領域  $\bar{\Omega} \cap \Omega'$  内にある画素に関するエネルギーであり、ここでは類似パターン位置が固定されているため定数として扱える. ここで、要素エネルギー  $E(\mathbf{x}_i)$  の変数は  $I_f(\mathbf{x}_i)$  であるため、 $E$  を欠損領域内のある画素の画素値  $I_f(\mathbf{x}_i)$  で偏微分すれば、エネルギー  $E$  を最小化する  $I_f(\mathbf{x}_i)$  の必要条件は次式で表せる.

$$\frac{\partial E}{\partial I_f(\mathbf{x}_i)} = \sum_{j=1}^{N_\Omega} \frac{\partial E(\mathbf{x}_j)}{\partial I_f(\mathbf{x}_i)} = \frac{\partial E(\mathbf{x}_i)}{\partial I_f(\mathbf{x}_i)} = 0 \quad (19)$$

従って、 $E$  を最小化する画素値  $I(\mathbf{x}_i)$  は以下のように算出される.

$$I_f(\mathbf{x}_i) = \frac{\sum_{\mathbf{p} \in W} w_{\mathbf{x}_i + \mathbf{p}} I_k(t(\mathbf{x}_i + \mathbf{p}) - \mathbf{p})}{\sum_{\mathbf{p} \in W} w_{\mathbf{x}_i + \mathbf{p}}} \quad (20)$$

### 3.5 修復画像を用いた全天球映像の生成

欠損領域を持つ全天球パノラマ画像に修復された平面投影画像を投影することで、不可視領域のない全天球映像を生成する。具体的には、まず全天球パノラマ画像における欠損領域上の各点とカメラの投影中心を結ぶ光線について、修復に用いた平面  $z = a_f x + b_f y + c_f$  との交点の座標を導出し、その座標の画素値を全天球パノラマ画像上の画素にコピーすることで、欠損領域のない全天球パノラマ画像を生成する。

## 4. 実験

本章では、全方位動画像上の死角により生じる欠損領域を修復し、不可視領域のない全天球映像を生成することで、提案手法の有効性を明らかにする。ここではまず、提案手法によって作成する平面投影画像系列に対して欠損修復を行い、全方位カメラの死角となる領域に対して違和感の小さい修復を行えることを示す。次に、修復された画像系列を用いて生成した不可視領域のないテレプレゼンスシステムの試作例を示し、提案手法の有効性を示す。

### 4.1 全方位動画像の欠損修復実験

#### 4.1.1 入力情報の取得

本実験では、入力画像を取得するために Point Grey Research 社製の全方位型マルチカメラシステム Ladybug(図 12) を用い、これを頭部に装着した状態で徒歩で移動しながら図 13 に示す画像を含む 181 フレーム (1086 枚) の全方位動画像を取得した。Ladybug には水平方向に 5 個、上方向に 1 個の CCD カメラが外向きに配置されており、各カメラは  $768 \times 1024$  画素の解像度の画像を 15fps で取得可能である。ここではまず、3.2 で述べた Structure from Motion 法を全方位動画像に適用することで、図 14 に示す全方位カメラの位置と自然特徴点の三次元位



図 12 全方位マルチカメラシステム Ladybug

置を得た。画像上での欠損領域は、あらかじめ第一フレームの画像上で手動で与え、その範囲は全てのフレームで共通とした。



図 13 入力に用いた全方位動画の1フレーム

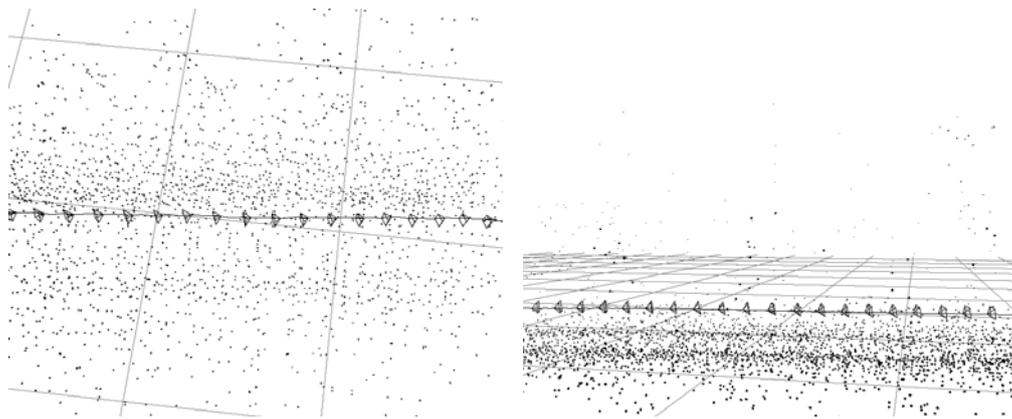


図 14 推定されたカメラ位置・姿勢および自然特徴点の三次元位置

#### 4.1.2 平面投影画像に対する欠損修復

ここでは、提案手法を用いて平面投影画像上における欠損領域を修復し、カメラの動きと撮影対象の形状情報のみから最適化を行わずに修復した画像との比較を行う。図 15 に、撮影された全方位動画画像の中間フレーム(第 91 フレーム)に対して、3.3 節で述べた手法を用いることで生成された平面投影画像を示す。また、図 16 に、その他のフレームについて、30 フレームごとの平面投影画像を示す。ここでは、画像上における 1 画素の世界座標系での長さを 20mm とし、各フレームに対して  $720 \times 720$  画素の平面投影画像を生成した。

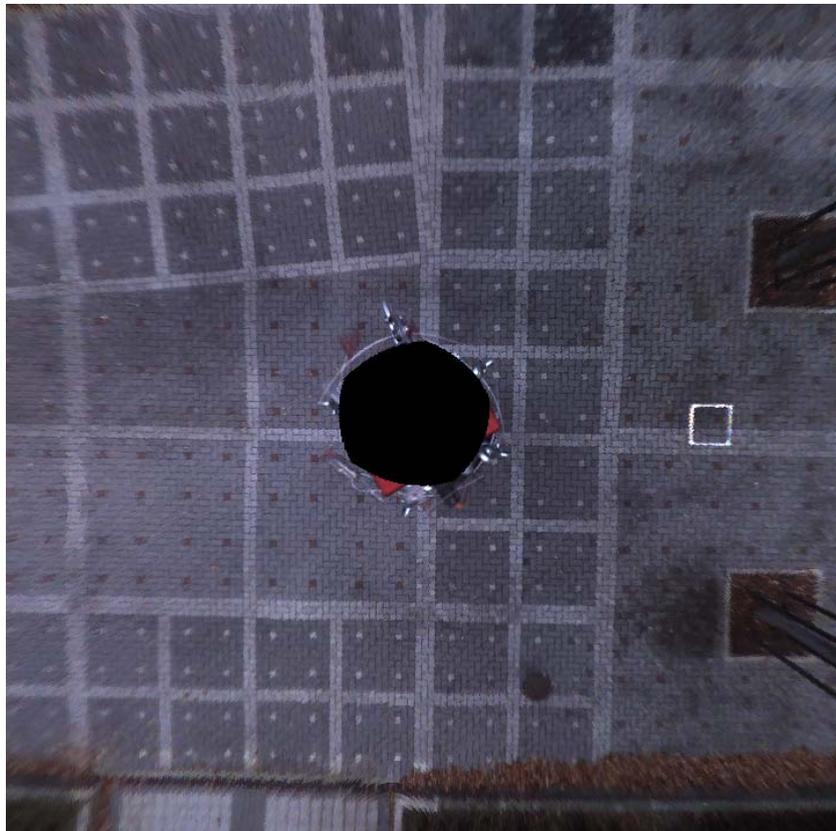


図 15 全方位動画画像の中間フレームに対して生成された平面投影画像 (第 91 フレーム)



(a) 1 フレーム目



(b) 31 フレーム目



(c) 61 フレーム目



(d) 121 フレーム目



(e) 151 フレーム目



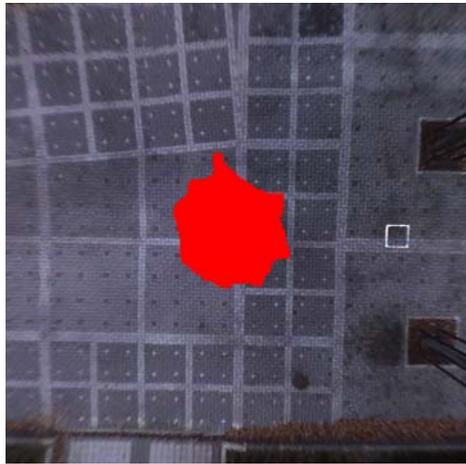
(f) 181 フレーム目

図 16 全方位動画像から生成した平面投影画像

平面投影画像の中心に表れた黒い円形の領域は，Ladybug の下側にカメラが存在しないために生じる機器構成上の死角である．ここで，生成された平面投影画像のタイル状のテクスチャは，平面投影画像上での位置によらず，均一な大きさとなっており，いずれのフレームにおいても，歪みのない修復に適したテクスチャが得られたことがわかる．ただし図 15 から，生成された平面投影画像上では，画像中心から離れるに従ってテクスチャの解像度が低下していることが分かる．

図 17 は，3.4 節で述べた手法を用いることで，第 91 フレームの画像上の欠損領域を，第 171 フレームの画像を用いて修復したものである．本実験では，エネルギー最小化による修復処理において，修復処理に用いるウインドウ  $W$  のサイズを  $41 \times 41$  画素，重み係数  $g$  を 1.2 とし，データ領域として用いる類似パターンの探索範囲  $\Phi$  は，欠損領域に対応する画素を中心とする  $11 \times 11$  画素の範囲とした．また，エネルギー関数  $E$  の収束判定は，反復におけるエネルギーの減少幅がエネルギーの大きさの 0.01% 以下となる場合とした．図 17(b) は，修復に利用したデータ領域  $\Phi$  が存在するフレームの平面投影画像を示しており，図中の  $\times$  印は，修復対象となる第 91 フレームの欠損領域に対応するデータ領域のおおよその位置を表している．また，同図中 (c) は，本実験で比較対象として用いた，幾何情報のみに基づく修復結果である．具体的には，提案手法において類似パターンの探索範囲  $\Phi$  の大きさを  $1 \times 1$  画素とすることで，類似パターンの探索を行わずに幾何情報のみから修復を行ったものである．同図中 (d) は，提案手法を適用して得られた修復結果画像である．

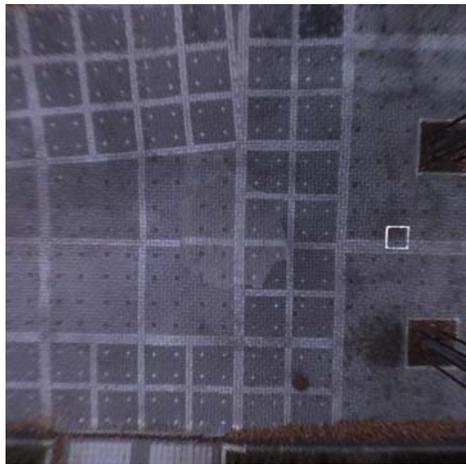
図 17(c) から，幾何情報のみから修復を行った場合には，欠損領域とその他の領域の境界部分において，テクスチャのずれが生じていることがわかる．この要因として，SFM 法によるカメラ位置・姿勢推定の誤差や，平面当てはめの誤差が考えられる．これに対して，提案手法を適用することで得られた修復結果画像 (図 17(d)) においては，類似パターンの探索により位置合わせのずれが最小化され，幾何学的なテクスチャのずれは見られない．ただし，提案手法によって得られた修復結果には，同一画像中における光学的な不連続が発生しており，違和感が生じている．この原因は，異なるフレームの画像上の物体に対する照明条件の違いであると考えられる．本実験では，全方位カメラやそのカメラを動かす人物



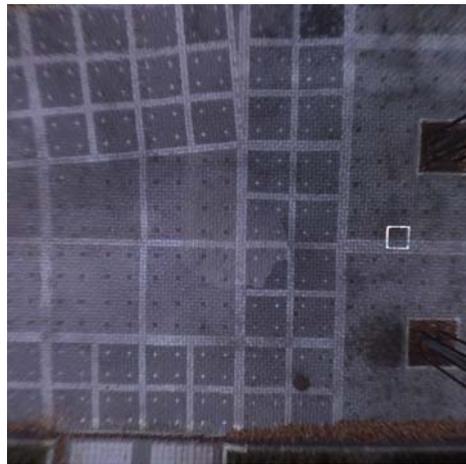
(a) 修復対象フレーム (第 91 フレーム)



(b) データ取得用フレーム (第 171 フレーム)



(c) 平面を利用した画素値の投影による修復結果画像



(d) 提案手法による修復結果画像

図 17 平面を利用した画素値の投影による修復結果と提案手法による修復結果の比較

の影響によって、欠損領域周辺の地面の明度が低下しており、他フレームの画像上に存在するデータ領域のテクスチャによる修復を行う提案手法では、欠損領域内の明度が周囲に対して相対的に明るくなる。これを解決するためには、従来提案されている明度変化を考慮したエネルギー項(文献 [6]) を本手法で用いるエネルギー関数に導入する必要がある。

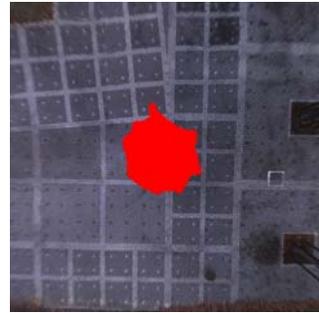
次に、連続するフレーム(第 105~113 フレーム)に対する平面投影画像を図 18 に、それらの画像に対して提案手法による修復処理を行った出力画像を図 19 に示す。図 19 より、提案手法を適用して得られた連続フレームに対する修復結果画像においては、テクスチャは滑らかに変化し、時間軸方向におけるテクスチャの不連続に起因する違和感は小さい。しかし、これらの画像上においても光学的なテクスチャの不連続が確認でき、違和感が生じている。



(a) 105 フレーム目



(b) 106 フレーム目



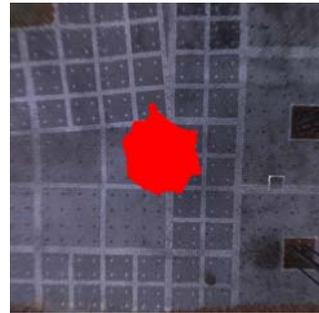
(c) 107 フレーム目



(d) 108 フレーム目



(e) 109 フレーム目



(f) 110 フレーム目



(g) 111 フレーム目



(h) 112 フレーム目



(i) 113 フレーム目

図 18 連続フレームに対する平面投影画像



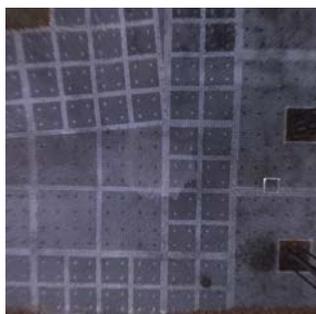
(a) 105 フレーム目



(b) 106 フレーム目



(c) 107 フレーム目



(d) 108 フレーム目



(e) 109 フレーム目



(f) 110 フレーム目



(g) 111 フレーム目



(h) 112 フレーム目



(i) 113 フレーム目

図 19 連続フレームに対する修復結果

## 4.2 不可視領域のない全天球映像によるテレプレゼンスシステムの試作

本実験では、前節の実験で得られた平面投影画像に対する修復結果を用い、全天球パノラマ画像上の欠損領域を補完することで、全周を自由に見回し可能なテレプレゼンスシステムを試作し、提案手法による修復の有効性を確認する。ここではまず、図 20 に示すような欠損領域を持つ全方位パノラマ画像の欠損部分に対する平面投影画像 (図 21) 上の画素を算出し、対応する画素の画素値を欠損領域内にコピーすることで、図 22 に示す不可視領域のない全天球パノラマ画像を作成した。次に、試作したシステムに図 22 に示した不可視領域のない全天球パノラマ画像を入力し、コントローラを用いて自由な見回しを行った。

図 23～25(a) に、ユーザへの提示に用いた画像の一例を示す。同図 (b) は、(a) と同一視野方向において、提案手法を用いなかった場合の欠損を含む画像である。提案手法を用いることで、欠損領域のない 360 度全ての方向の見回しが可能な臨場感の高いテレプレゼンスを実現することができた。また、(a), (b) の比較により、提案手法によりテレプレゼンスにおける臨場感が大幅に向上することがわかる。



図 20 欠損領域を持つ全天球パノラマ画像



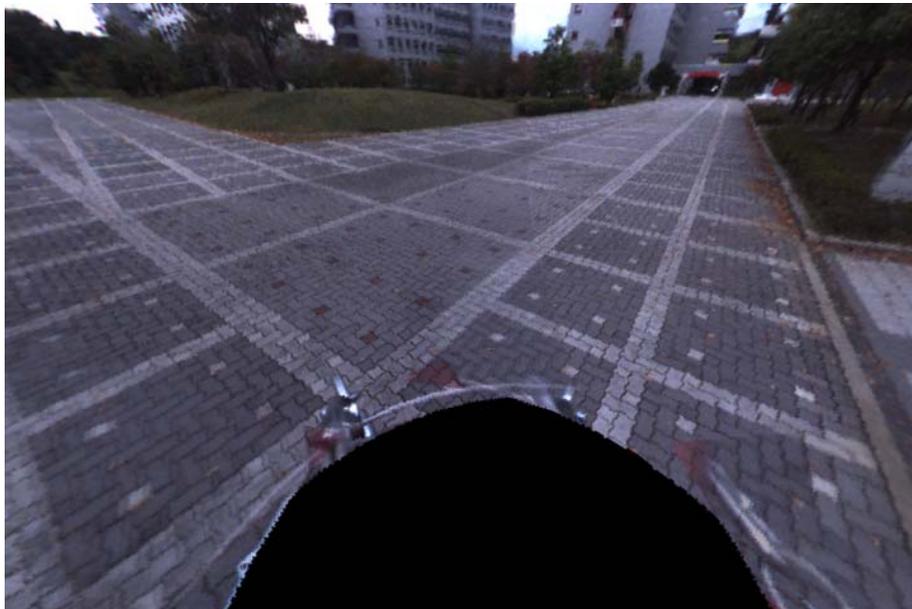
図 21 提案手法による平面投影画像の修復結果



図 22 修復された全天球パノラマ画像



(a) 欠損なし

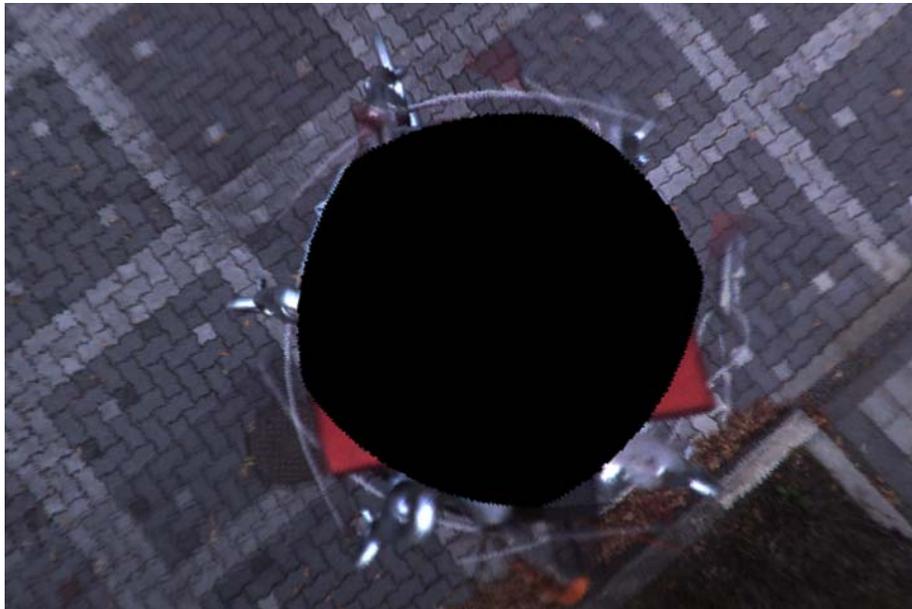


(b) 欠損あり

図 23 全天球映像を用いた見回し例 1



(a) 欠損なし



(b) 欠損あり

図 24 全天球映像を用いた見回し例 2



(a) 欠損なし



(b) 欠損あり

図 25 全天球映像を用いた見回し例 3

ただし、提案手法を用いた場合においても見直し画像上で欠損領域内のテクスチャがぼけていることがわかる。これは、データ領域として用いるテクスチャが、平面投影画像上での解像度が低い画像の周辺に存在することによって、修復されるテクスチャの解像度が低下することに起因している。平面投影画像上において、画像の中心から離れるに従って解像度が低下する問題は、カメラの構成上生じる問題であるため、これを解決するためには超解像手法を用いるなどの工夫が必要であると考えられる。また、平面投影画像上において、可能な限り欠損領域に近いテクスチャをデータ領域として選択することも有効であると考えられ、今後はデータ領域として用いるフレームを適切に選択するアルゴリズムを開発する必要がある。

## 5. まとめ

本論文では、全方位画像上における欠損領域周辺の撮影対象の形状と全方位カメラの運動を考慮し、平面投影画像を作成することで、テクスチャの変形補正および、修復の事例として用いる画像領域の限定を行い、不可視領域のない全天球映像を生成する手法を提案した。実験では、提案手法を用いることで、投影歪みのない平面投影画像を生成できることを示した。また、類似パターンの探索処理と最適化処理を行うことで、三次元復元や平面の当てはめ処理において生じる投影誤差の影響を抑制し、不連続なテクスチャの発生を抑止できることを示した。ただし、解像度や光度の違いにより、修復された領域とその他の部分の境界が目立つといった問題点も確認した。

今後の課題として、光学的なテクスチャの不連続を解消するために、修復処理におけるテクスチャの輝度値の変化を考慮する必要がある。また、修復された領域のテクスチャの解像度の低下を防ぐため、可能な限り欠損領域に近いテクスチャをデータ領域として用いるアルゴリズムの開発や、超解像手法の導入を検討する必要がある。また、提案手法の有効性を定量的に評価するためには、仮想空間に全方位カメラを配置することで撮影された動画像に対して修復を行い、修復された画像と、欠損方向を撮影した画像との比較を行うことなどの方法が必要である。

## 謝辞

本研究の全過程を通して、懇切なる御指導、御鞭撻を賜りました視覚情報メディア講座横矢直和教授に心より感謝いたします。本研究を進めるにあたり、有益な御助言、御指導を頂きました知能情報処理学講座木戸出正継教授に厚く御礼申し上げます。本研究を通じて、温かい御指導をして頂きました視覚情報メディア講座山澤一誠准教授に深く感謝いたします。そして、本研究を行うにあたり、多大なる御助言、御鞭撻を賜りました視覚情報メディア講座佐藤智和助教に深く御礼申し上げます。特に佐藤智和助教には、本論文の執筆、その他の論文の添削に至るまで細やかな御指導を頂きました。本研究に関する貴重な御助言や御指摘を頂きました視覚情報メディア講座河合紀彦氏に深く感謝いたします。研究室において物心両面において常に温かい御支援を頂きました視覚情報メディア講座高橋美央女史に厚く御礼申し上げます。最後に、研究活動だけでなく日々の生活においても大変お世話になりました視覚情報メディア講座の諸氏に深く感謝いたします。

## 参考文献

- [1] 池田, 佐藤, 横矢: “全方位型マルチカメラシステムを用いた高解像度な全天球パノラマ動画の生成とテレプレゼンスへの応用”, 日本バーチャルリアリティ学会論文誌, Vol. 8, No. 4, pp. 443–450, 2003.
- [2] 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, pp. 193–202, 2007.
- [3] 山口, 山澤, 竹村, 横矢: “全方位動画を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”, 電子情報通信学会論文誌, Vol. J84-D-II, No. 9, pp. 2048–2057, 2001.
- [4] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester: “Image Inpainting,” Proc. ACM SIGGRAPH2000, pp. 417–424, 2000.
- [5] A. Criminisi, P. Perez and K. Toyama: “Region Filling and Object Removal by Exemplar-Based Inpainting,” IEEE Trans. on Image Processing, Vol. 13, pp. 1200–1212, 2004.
- [6] 河合, 佐藤, 横矢: “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, Vol. J91-D, No. 9, pp. 2293–2304, 2008.
- [7] C. Allene and N. Paragios: “Image Renaissance Using Discrete Optimization,” Proc. IEEE Int. Conf. on Pattern Recognition, pp. 631–634, 2006.
- [8] 天野, 佐藤: “kBPLP法を用いた高次元非線形射影による画像補完”, 電子情報通信学会論文誌, Vol. J86-D-II, No. 4, pp. 525–534, 2003.
- [9] M. Bertalmio, A. L. Bertozzi and G. Sapiro: “Navier-Stokes, Fluid Dynamics, and Image and Video Inpainting,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Vol. 1, pp. 355–362, 2001.

- [10] Y. Wexler, E. Shechtman and M. Irani: “Space-Time Completion of Video,” *Trans. on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 3, pp. 463–476, 2007.
- [11] V. Cheung, B. Frey and N. Jojic: “Video epitomes,” *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 141–152, 2005.
- [12] R. Bornard, E. Lecan, L. Laborelli and J. Chenot: “Missing Data Correction in Still Images and Image Sequences,” *ACM Multimedia*, pp. 355–361, 2002.
- [13] 小畑, 守田: “3次元画像埋め込みによる3次元画像の修復”, *電子情報通信学会 技術研究報告*, PRMU 2004-41, pp. 49–54, 2004.
- [14] Y. Matsushita, E. Ofek, W. Ge, X. Tang and H. Shum: “Full-Frame Video Stabilization with Motion Inpainting,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 7, pp. 1150–1163, 2006.
- [15] A. Litvin, J. Konrad and W. Karl: “Probabilistic video stabilization using Kalman filtering and mosaicking,” *Proc. SPIE Electronic Imaging*, Vol. 5022, pp. 663–674, 2003.
- [16] Y. Zhang, J. Xiao and M. Shah: “Motion Layer Based Object Removal in Videos,” *Proc. IEEE Workshops on Application of Computer Vision*, Vol. 1, pp. 516–521, 2005.
- [17] J. Jia, T. Wu, Y. Tai and C. Tang: “Video Repairing: Inference of Foreground and Background Under Severe Occlusion,” *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Vol. 1, pp. 364–371, 2004.
- [18] J. Jia, Y. Tai, T. Wu and C. Tang: “Video Repairing under Variable Illumination Using Cyclic Motions,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 28 No. 5, pp. 832–839, 2006.

- [19] 阿部, 目黒, 川又: “移動物体検出と時間領域メジアンフィルタを用いた古いフィルム映像のプロッチ除去”, 電子情報通信学会論文誌, Vol. J88-A No.1, pp. 11–22, 2005.
- [20] Y. Shen, F. Lu, X. Cao and H. Foroosh: “Video Completion for Perspective Camera Under Constrained Motion,” Proc. IEEE Int. Conf. on Pattern Recognition, pp. 63–66, 2006.
- [21] K. Garg and S. K. Nayar: “Detection and Removal of Rain from Videos,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Vol. 1, pp. 528–535, 2004.
- [22] 福地, 山下, 金子, 三浦: “時空間画像処理による雨天時画像からの視野妨害ノイズ除去”, 映像情報メディア学会誌, Vol. 62, No. 5, pp. 771–777, 2008.
- [23] 山下, 蔵本, 金子: “複数カメラを用いた画像中の視野妨害ノイズ除去”, 電気学会論文誌 C, Vol. 127, pp. 480–488, 2007.
- [24] K. Patwardhan, G. Sapiro and M. Bertalmio: “Video Inpainting Under Constrained Camera Motion,” IEEE Trans. on Image Processing, Vol. 16, pp. 545–553, 2007.
- [25] 佐藤, 池田, 横矢: “複数動画からの全方位型マルチカメラシステムの位置・姿勢パラメータの推定”, 電子情報通信学会論文誌, Vol. J88-D-II No. 2, pp. 347–357, 2005.