

NAIST-IS-MT0951095

修士論文

視点移動可能なビデオテクスチャ

馬場口 豊

2011年2月3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。

馬場口 豊

審査委員：

横矢 直和 教授 (主指導教員)

木戸出 正繼 教授 (副指導教員)

山澤 一誠 准教授 (副指導教員)

視点移動可能なビデオテクスチャ*

馬場口 豊

内容梗概

動的環境を一定時間撮影したビデオ映像を適切なタイミングで繰り返し再生することで、違和感なく任意時間の動画像を合成する手法はビデオテクスチャと呼ばれ、仮想空間表現における従来の静止画によるテクスチャに代わる画像提示手法として多くの分野への応用が期待されている。ビデオテクスチャの生成手法として、従来、撮影した動画像中の物体の運動の連続性を考慮した視線方向が変更可能なビデオテクスチャを生成する手法が提案されているが、視点移動は考慮されていなかった。本論文では、離散的な位置で異なる時間に撮影した複数の動画像を入力とし、周期運動物体を対象とした映像の繰り返し再生を実現しつつ、映像の時間的・空間的な連続性を考慮した視点移動を行うことが可能なビデオテクスチャ生成手法を提案する。本研究では、まず、各地点で撮影した動画像に対して、それぞれ時間軸方向に画素値の中央値を求め背景画像を生成し、背景差分により各フレームにおける動物体領域画像を算出する。次に、各動画像中で時間的な連続性を考慮した類似フレーム探索を行い、各地点における映像の繰り返し再生を行う。隣接する視点位置で取得した動画像に対しても、時間的な連続性を考慮した類似フレーム探索を行い、類似フレーム間で映像を接続することにより、視点移動時の時間的な連続性を保持した映像合成を実現する。ただし、撮影位置の異なる画像間での類似フレーム探索には、対象の見え方の変化を考慮する必要がある。これに対して本研究では、周期運動物体を平面と仮定し、各視点で撮影された画像を射影変換することで対応する。また、隣接する視点において接続さ

*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-MT0951095, 2011年2月3日.

れた類似フレームの前後のフレームを用いて中間視点画像を生成することで，視点移動時の空間的な連続性を実現する．実験では，実環境における周期運動物体を対象とした動画像を入力として用い，提案手法によって視点移動可能なビデオテクスチャを実際に生成可能であることを示す．

キーワード

ビデオテクスチャ, 視点移動, 繰り返し再生, 中間視点画像生成

Viewpoint varying video textures*

Yutaka Babaguchi

Abstract

This thesis presents a method for generating viewpoint-varying video textures using asynchronous video sequences taken from different viewpoints. Video texture is a naturally looping video whose playback time can be infinity. As one of extensions for video textures, a method that enables users to change the line of sight considering the continuity of the movement of the object has been proposed. However, there is one limitation that users cannot change viewpoints. In this thesis, in order to make it possible to change viewpoints in video textures, video sequences taken from different viewpoints at different time are used as input data. The proposed method is constructed of two processes: (1) synthesis of naturally looping video for each viewpoint, (2) connection of frames between adjacent viewpoints considering temporal and spatial continuity. In this method, first, regions of moving objects are detected for each frame by the background subtraction. Next, for each viewpoint, pair of frames which have similar regions of moving objects are searched. By defining found pair of frames as a range of looping video, synthesis of naturally looping video for each viewpoint is achieved. Also for adjacent viewpoint, pair of frames which have similar region of moving objects are found considering homography transformation. This pair of frames ensures temporal continuity in connection of frames between adjacent viewpoints. Moreover, by generating images for interpolated viewpoints, spatial continuity of generated

*Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT0951095, February 3, 2011.

video is maintained. In experiment, we successfully generate viewpoint-varying video textures from real video sequences.

Keywords:

Video textures, changing viewpoints, looping videos, view interpolation

目次

1. はじめに	1
2. 仮想化現実空間の構築に関する関連研究と本研究の位置付け	4
2.1 IBRに基づく仮想化現実空間の構築に関する関連研究	4
2.1.1 全方位動画像の一部切り出しによる手法	5
2.1.2 画像群の変形に基づく手法	7
2.1.3 光線情報を用いる手法	8
2.1.4 映像の繰り返し再生による手法	9
2.1.5 画像変形に幾何形状を利用する手法	10
2.2 本研究の位置付けと方針	12
3. 視点移動可能なビデオテクスチャ	14
3.1 提案手法の概要	14
3.2 動画像中の類似フレームの探索	16
3.2.1 動物体領域の算出	16
3.2.2 各視点位置における時間的連続性を考慮した類似フレーム探索	16
3.2.3 隣接する視点位置に対する時間的連続性を考慮した類似フレーム探索	17
3.3 視点移動可能なビデオテクスチャの生成	18
3.3.1 ユーザの入力による視点位置の設定	18
3.3.2 中間視点画像の生成	19
3.3.3 中間視点位置における再生フレームの決定	22
4. 実験	24
4.1 単一の周期運動物体を対象とした視点移動可能なビデオテクスチャ	24
4.1.1 実験環境	24
4.1.2 実験の結果と考察	28
4.1.3 被験者による評価実験と考察	35

4.2	複数の周期運動物体を対象とした視点移動可能なビデオテクスチャ	39
4.2.1	実験環境	39
4.2.2	実験の結果と考察	42
5.	まとめ	48
	謝辞	49
	参考文献	50

目 次

1	Dynamosaicing により生成したパノラマ動画像の 1 フレーム [4] . . .	6
2	和田らのテレプレゼンスシステムの外観 [9]	6
3	View Morphing[11]	8
4	Town digitizing[13]	8
5	Video Texture の 1 フレーム [22]	10
6	Panoramic Video Texture の 1 フレーム [26]	11
7	セグメンテーションベースドステレオにより推定された奥行き画 像 (左) と自由視点画像例 (右) [28]	11
8	視点移動可能なビデオテクスチャ生成処理の流れ	15
9	類似フレーム間における繰り返し再生	17
10	撮影地点間を等距離と仮定した直線上における中間視点位置 . . .	19
11	背景画像生成処理の流れ ($q_i = 0.5$)	20
12	中間視点画像の生成 (隣接地点間を等速に視点移動した場合) . . .	21
13	中間視点画像生成処理の流れ ($q_i = 0.5$)	22
14	ビデオテクスチャの生成視点位置 (単一の周期運動物体)	25
15	各撮影位置における動画像の第 1 フレーム (単一の周期運動物体)	26
16	各撮影位置において生成された背景画像と手動で与えた対応点 (単 一の周期運動物体)	27
17	撮影地点 0 において算出された類似フレーム (単一の周期運動物体)	30
18	撮影地点 0 において算出された類似フレームに対応する動物体領 域画像 (単一の周期運動物体)	31
19	地点 (a) の前後のフレームの提示映像と対応する動物体画像と背景 画像	32
20	地点 (b) の前後のフレームの提示映像と対応する動物体画像と背景 画像	33
21	21 フレーム目の提示映像と拡大画像	34
22	繰り返し再生のない映像 (A-1) と繰り返し再生のある映像 (A-2) に 対する映像接続回数と回答人数の関係	37

23	提案手法を用いた視点移動映像 (B) に対する映像接続回数と回答 人数の関係	38
24	ビデオテクスチャの生成視点位置 (複数の周期運動物体)	40
25	各撮影位置における動画像の第 1 フレーム (複数の周期運動物体)	41
26	各撮影位置において生成された背景画像と手動で与えた対応点 (複 数の周期運動物体)	41
27	撮影地点 0 において算出された類似フレーム (複数の周期運動物体)	43
28	撮影地点 0 において算出された類似フレームに対応する動物体領 域画像 (複数の周期運動物体)	44
29	地点 (c) の前後のフレームの提示映像と対応する動物体画像と背景 画像	45
30	地点 (d) の前後のフレームの提示映像と対応する動物体画像と背景 画像	46
31	地点 (d) の 1 フレーム後の提示映像と拡大画像	47

表 目 次

1	従来の IBR に基づく仮想化現実空間の構築手法の特徴と提案手法 の位置付け	12
2	実験時のパラメータ	25
3	算出された各撮影位置における類似フレーム (m_p, n_p)	29
4	算出された各撮影位置における類似フレーム (m_p, n_p)	43

1. はじめに

現実空間の情報に基づいて構築された仮想空間は仮想化現実 (AV: Augmented Virtuality) 空間と呼ばれ、遠隔テレプレゼンスや景観のデジタルアーカイブ、エンターテインメント等の様々な分野への応用が期待されている。従来、予め取得した画像群を利用して、写実的な仮想空間を自動的に構築する試みが成されており、これらは三次元の幾何モデルを用いてシーンを描画するモデルベースドレンダリング (MBR) と、幾何モデルを直接用いずに実写画像群からシーンを描画するイメージベースドレンダリング (IBR) に大別できる。本研究では、現実環境を撮影した画像群を変形・合成することにより、一定の条件下において写実性の高い仮想空間を構築できる IBR に基づく手法に着目する。

仮想化現実空間の構築に用いられる IBR としては、以下のような手法が提案されている。

- 全方位動画像の一部切り出しによる手法 [1-9]
- 画像群の変形に基づく手法 [10-15]
- 光線情報を用いる手法 [16-21]
- 映像の繰り返し再生による手法 [22-26]
- 画像の変形に幾何形状を利用する手法 [27-29]

全方位動画像の一部切り出しによる手法は、離散的な位置で撮影された動画像を用いる手法 [1-5] と移動を伴って撮影された全方位動画像を用いる手法 [6-9] に大別できる。離散的な位置で撮影された動画像を用いる手法では、全方位カメラで撮影された動画像や回転を伴って撮影された動画像を用いることで、提示画像中の物体の運動を再現しつつ、視線方向の自由な変更を実現している。しかし、視点位置の変更が画像取得位置上に限定されているため、仮想環境における視点位置の移動は離散的なものとなり、滑らかな視点移動は実現されていない。移動を伴って撮影された全方位動画像を用いる手法は、ユーザに提示するフレームを変更することで、撮影経路上における連続的な視点移動が可能となる。しかし、視

点移動と環境中の物体の運動が連動してしまうため、視点移動時に物体の運動が正しく再現されないという問題がある。画像群の変形に基づく手法は、蓄積された画像群を直接変形することにより、画像取得時には撮影されなかった中間視点画像を生成する。基礎的な手法として、画像間に対応点を与え、モーフィング処理で補間画像を生成する手法 [10–12] が提案されているが、動的環境を対象とした場合、自動的に対応点を与えることは難しいという問題がある。対応点を明示的に与えることなく画像を変形することで中間視点画像を生成する手法 [13–15] も提案されているが、運動を再現できる時間が撮影時間によって制限されるため、物体の運動を任意時間再生することはできない。光線情報を用いる手法は、画像を光線情報の集合として捉え、光線を観測位置と方向によってモデル化することで、任意の視点・視線の画像を再現する。必要な光線情報を画像としてあらかじめ収集しておくことで、任意の視点・視線方向の画像を生成することができるが、動的環境を対象として必要な光線情報を収集することは難しいという問題がある。周期運動物体を一定時間撮影したビデオ映像を適切なタイミングで繰り返し再生することで、違和感のない任意時間の動画像を合成するビデオテクスチャ [22] と呼ばれる手法が、仮想空間表現における従来の静止画によるテクスチャに代わる画像提示手法として提案されている。現時点において、撮影した動画像中の物体の運動の連続性を考慮した、視線方向を変更可能とする拡張手法 [26] が提案されているが、視点移動は実現されていない。また、幾何形状を併用する手法は、現実環境の幾何形状を画像群の変形・合成時に利用することで、写実的な仮想環境を構築する。仮想環境における滑らかな視点移動と物体の運動の再現は撮影時間によって制限され、長時間の運動再現を行うことは、データ量、計算コストの点から現実的でない。

このように、従来、仮想化現実空間を構築するための様々な手法が提案されているが、動的環境を対象とした場合、入力が容易であり、また物体の運動再現を実現しつつ滑らかな視点移動を実現する手法は提案されていない。そこで、本論文では、動的環境を対象に、容易に取得可能な入力を用いて、物体の運動再現を実現しつつ滑らかな視点移動を実現する手法を提案する。提案手法は、動画像を違和感なく任意時間の動画像に合成するビデオテクスチャおよび画像群の変形に

よる中間視点画像生成手法を基礎として、離散的な位置で異なる時間に撮影した複数の動画像を入力とし、(1)周期運動物体を対象とした映像の繰り返し再生、(2)撮影地点間での類似フレームの接続による中間視点画像生成、を行うことで、視点移動可能なビデオテクスチャを実現する。具体的には、まず最初に、各撮影地点で撮影した動画像に対して、それぞれ時間軸方向に画素値の中央値を求め背景画像を生成し、背景差分により各フレームにおける動物体領域画像を算出する。次に、各動画像中で時間的な連続性を考慮した類似フレーム探索を行い、各撮影地点における映像の繰り返し再生を可能にする。隣接する視点位置で取得した動画像に対しても、時間的な連続性を考慮した類似フレーム探索を行い、類似フレーム間で映像を接続することにより、視点移動時の時間的な連続性を保持した映像合成を実現する。ただし、撮影位置の異なる画像間での類似フレーム探索には、対象の見え方の変化を考慮する必要がある。これに対して本研究では、周期運動物体を平面と仮定し、各視点で撮影された画像を射影変換することで対応する。また、隣接する視点において接続された類似フレームの前後のフレームを用いて中間視点画像を生成することで、視点移動時の空間的な連続性を実現する。

以下、2章では、仮想化現実空間の構築に関する従来研究と本研究の位置付けについて述べる。3章では、周期運動物体を対象とした映像の繰り返し再生と撮影地点間での類似フレームの接続による中間視点画像生成の手法について詳述する。4章では、提案手法によって視点移動可能なビデオテクスチャを生成した結果を示す。最後に5章では、まとめと今後の展望について述べる。

2. 仮想化現実空間の構築に関する関連研究と本研究の位置付け

予め取得した画像群を利用して、写実的な仮想空間を自動的に構築する手法は、三次元の幾何モデルを用いてシーンを描画するモデルベースレンダリング (MBR) と、幾何モデルを直接用いずに実写画像群からシーンを描画するイメージベースレンダリング (IBR) に大別できる。MBR は、現実世界の三次元形状や光学的特性をモデル化し、これに基づいて仮想世界の描画を行う手法であり、任意の視点に対してモデルを描画することができる。また光学的特性をモデル化した場合には、任意の照明条件を再現することも可能となる。しかし、複雑なシーンに対して正しく幾何形状を得ることは容易ではなく、写実的な画像を生成することが難しいという問題がある。これに対して IBR は、現実環境を撮影した画像群を変形・合成することにより、一定の条件下において写実性の高い仮想空間を構築することができる。本研究では、このような理由から IBR に着目する。以下では、IBR に基づく仮想化現実空間の構築手法について概観するとともに、本研究の位置付けを述べる。

2.1 IBR に基づく仮想化現実空間の構築に関する関連研究

IBR は、対象環境の三次元形状を陽に復元することなく、物体やシーンを撮影した画像群を変形・合成することにより写実性の高い仮想空間を構築する手法である。この手法では、対象物体の形状や反射特性を直接利用しないため、MBR では正確な幾何形状の推定が難しい複雑な形状を持つ木などの物体に対しても、写実的な仮想空間を描画することが可能である。以下では、IBR に基づく仮想化現実空間の構築手法を、全方位動画像の一部切り出しによる手法、画像群の変形に基づく手法、光線情報を用いる手法、映像の繰り返し再生による手法と、画像の変形に幾何形状を利用する手法に大別し、それぞれの特徴を述べる。

2.1.1 全方位動画像の一部切り出しによる手法

写実的な仮想環境を構築するための最も簡単な方法は、事前に撮影された現実環境の映像をユーザに対して提示することである。このような手法は、固定地点または離散的な位置で撮影された動画像を用いる手法や移動を伴って撮影された全方位動画像を用いる手法が提案されている。

固定地点または離散的な位置で撮影された動画像を用いる手法は、全方位カメラで撮影された動画像や回転を伴って撮影された動画像を用いることで、提示画像中の物体の運動を再現しつつ、自由な視線方向の変更を実現している。Chen の QuickTime VR [1] に代表される手法では、ユーザの入力に応じて全方位画像からの切り出しと幾何学的な補正を行った後に、これをユーザに提示することで遠隔地の情景を 360 度見回す感覚を再現している [2]。このような手法はウェブブラウザ上での観光案内や施設案内など、様々なところで利用されている。また、全方位カメラを用いることなく、回転を伴って撮影された動画像を用い、取得した動画像からパノラマ動画像を生成する手法が提案されている。Rav-Acha ら [3] は、画像の時間変化を空間的に捉えることができる時空間特徴を用いて動領域の時系列を操作することにより、パノラマ動画像を生成する *Dynamosaicing* を提案している (図 1)。Hermans ら [4] は、時空間特徴に基づいて配列したフレームから静的な背景領域、動的な背景領域、動的な前景領域を算出し、それらをユーザの視線方向に応じて変形することでパノラマ動画像を生成する。これら離散的な位置で撮影された動画像を用いる手法は、視点位置の変更が離散的な画像取得位置上に限定されるため、仮想環境における自由な視点位置の移動は実現されていない。

移動を伴って撮影された全方位動画像を用いる手法は、提示するフレームを変更することで、撮影経路上における連続的な視点移動を実現している。Onoe ら [5] は、全方位カメラによって移動撮影された映像をユーザの視線方向の変化に応じて提示することで、撮影経路上における連続的な視点移動と任意 (3 自由度) の視線方向の変更を実現する手法を提案している。Ikeda ら [6] は、撮影された全方位動画像の視点位置 (観測フレーム) と視線方向を変更することで、撮影経路上における 1 自由度の視点移動と任意 (3 自由度) の視線方向の変更を実



図 1 Dynamosaicing により生成したパノラマ動画の 1 フレーム [4]

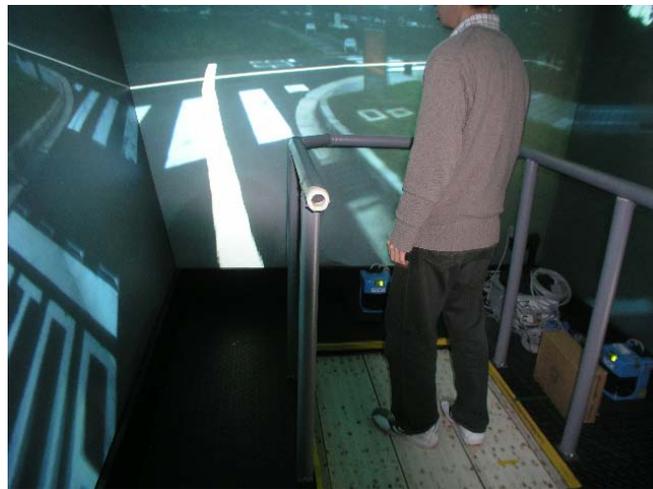


図 2 和田らのテレプレゼンスシステムの外観 [9]

現している．また，複数の経路上を撮影した全方位動画像を用い，複数の経路にわたり視点位置をインタラクティブに選択することが可能な手法が提案されている [7-9]．これらの手法では，HDR(High-Dynamic Range) 画像生成や映像の安定化 [7]，Structure-from-Motion による経路の自動生成 [8] や複数経路上の歩行移動 [9](図 2) といった拡張も行われている．しかし，これら移動撮影された動画像を用いる手法では，視点移動と環境中の物体の運動が連動してしまうため，視点移動時に動物体を正しく再現できないという問題がある．

2.1.2 画像群の変形に基づく手法

蓄積された画像群を直接変形することにより，画像取得時には撮影されなかった中間視点画像を生成する手法が提案されている．代表的な手法として，ワーピングやモーフィングによって中間視点画像を生成する手法 [10–12] がある．これらの手法は，入力となる複数の画像間で対応点を与え，中間視点位置において対応点がどの位置に現れるかを決定することで入力画像を変形させる．さらに，対応点の位置関係より入力画像の色をブレンディングすることで中間視点位置における各画素の色を求め，中間視点画像を生成する (図 3)．これらの手法では，対応点を与えることで，中間視点画像を生成し 2 視点間に限定した視点移動が実現できる．しかし，動的環境を対象とした場合，対応点を自動的にかつ正確に与えることが難しいという問題がある．

これに対して，対応点を明示的に与えることなく画像群を変形することで中間視点画像を生成し，仮想環境を構築する手法が提案されている [13–15]．離散的な位置で撮影された複数の全方位画像を用いる Town Digitizing [13](図 4) では，画像の拡大およびブレンディングを用い 2 枚の画像の取得位置の間を補間する画像を作成している．このシステムは静的環境を対象としており，動的環境は考慮されていない．亀井ら [14] は，動的環境を対象に移動撮影した映像を入力とし，ユーザが設定した仮想視点位置の映像を，動画像系列の画像を拡大・縮小と平行移動により合成する手法を提案している．この手法では，映像を取得した時間に限りユーザの撮影経路上の自由な視点移動と物体の運動再現を同時に実現しているが，物体の運動を任意時間再現し続けることができず，またユーザの視点移動が前後進のみに限定されている．密に設置された複数のカメラで非同期に撮影された動画像を用いる手法 [15] は，各画素のオプティカルフローに基づいて各視点上における動物体間の対応関係を算出し，その対応関係から算出された平面射影変換を用いて，シーン全体の画像変形モデルを作成することで，動的環境に対応した中間視点画像を生成する．これにより，仮想環境における物体の運動は再現されるが，密に設置された複数のカメラで対象を撮影する必要があり，データの取得コストが高いという問題がある．

これらの画像群の変形に基づく手法において，容易に取得できる入力であり，

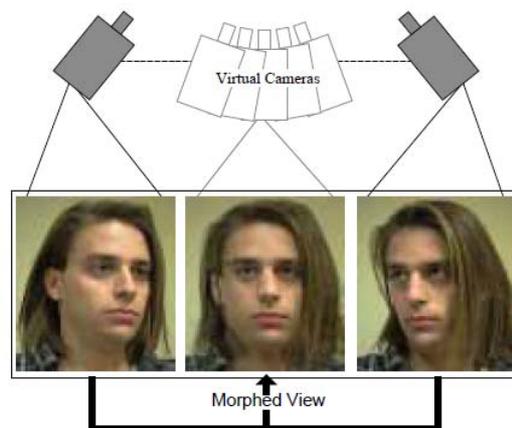


図 3 View Morphing[11]

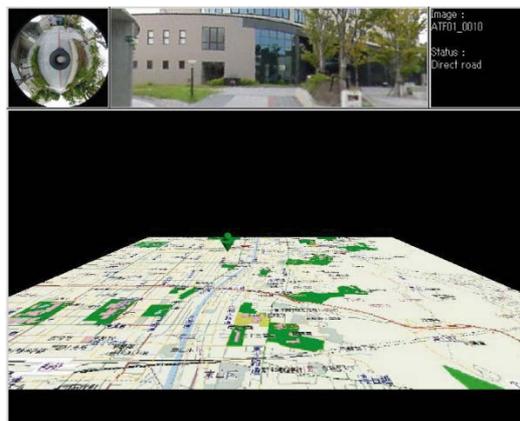


図 4 Town digitizing[13]

かつ動的環境を対象とした仮想環境におけるユーザの自由な視点移動と物体の運動再現を同時に実現する手法は提案されていない。

2.1.3 光線情報を用いる手法

画像を光線情報の集合として捉え、光線を観測位置と方向によってモデル化することで、任意の視点・視線の画像を再現する手法として、Plenoptic Function [16]、光線空間法 [17]、Light Field Rendering [18]、Lumigraph [19] 等が提案されている。Adelson ら [16] は、空間中のすべての光線を記述する関数として Plenoptic

Function を提案している . Plenoptic Function p は , 視点の座標 (x, y, z) , 光線の向き (θ, ϕ) , 波長 および時間 t を用いて , $p(x, y, z, \theta, \phi, \lambda, t)$ と表現する 7 次元関数として定義される . ただし , 7 次元のデータを全て記録するためには膨大な量の画像が必要となり , 現実的に全てのデータを収集することは難しい . これに対して Naemura ら [17] は , より低い次元の情報で空間を記述することで任意視点の画像を効率的に再構築する光線空間法と呼ばれる手法を提案している . 光線空間法は , 空間を伝搬する光線は減衰・変移しないと仮定することで , Plenoptic Function の 7 次元の光線情報から z 座標 , 波長 および時間 t を減らし , 情報量を 4 次元とすることで取得すべき画像の量を大幅に減らした . Levoy ら [18] および Gortler ら [19] は , 同時期に Light Field Rendering , Lumigraph という手法を提案した . これらの手法では , 光線空間法が平面 (x, y) および光線の向き (θ, ϕ) で 4 次元空間を表現したのに対して , この 4 次元を 2 つの平行した平面を横切る座標 (u, v) および (s, t) で表現することで , 効率的なレンダリングを可能にした . また , 全方位カメラを用いることで , 一度に多数の方向の光線情報を収集する手法も提案されている [20, 21] . 山口ら [20] は , 全方位画像センサ HyperOmni Vision を用いて直線上を移動しながら環境を撮影し , 任意の視線方向に対してステレオ視可能な画像を生成する手法を提案した . また , Takahashi ら [21] は同様に全方位画像センサを搭載した自動車で環境中を走行し , その移動経路付近の自由視点画像を生成する手法を提案している .

これらの光線情報を用いた手法の従来研究は , 主に静的シーンを対象としており , より多くの光線情報を必要とする動的環境を対象として , 画像の合成に必要な光線情報を収集することは難しい .

2.1.4 映像の繰り返し再生による手法

現実環境を違和感なく再現する仮想環境を構築するためには , 環境内の動物体が自然な運動を続けることが必要である . これに対して映像を繰り返し再生することによって任意時間の動画像を合成する手法が提案されている [22–26] . 映像の繰り返し再生による手法の原点は , Schödl らによるビデオテクスチャと呼ばれる手法 [22](図 5) である . この手法は , 周期運動物体を一定時間撮影した動画像から



図 5 Video Texture の 1 フレーム [22]

時間的な連続性を考慮した類似フレーム対を探索し，対応するフレーム間で映像を繰り返し再生する．さらに算出された対応フレーム間の類似度が低い場合，対応フレームに対してモーフィング処理によって中間画像を生成することで，繰り返し再生時の違和感を軽減している．また，映像中の対応フレームの遷移から取得される物体の運動を用いることで，動物体に対してユーザが指定した動きをさせることも可能としている [23, 24]．Kwatra ら [25] は，類似度の低い対応フレームに対しても，局所領域ごとに繰り返しフレームを算出することで，より違和感の小さいビデオテクスチャを生成する手法を提案した．Agarwala ら [26] は，カメラを回転させながら撮影した映像を違和感なく張り合わせることにより，物体の運動の連続性を保持しながら視線方向を変更可能とする拡張手法を提案した (図 6) ．

このように，映像の繰り返し再生によって物体の運動を再現しつつ任意時間の動画像を合成するビデオテクスチャの従来研究において，仮想空間における視線方向の変更を可能とする手法は提案されているが，視点移動は実現されていない．

2.1.5 画像変形に幾何形状を利用する手法

現実環境の幾何形状を画像群の変形・合成時に利用することで，写実的な仮想環境を構築する手法が提案されている [27–29]．Matusik ら [27] は，自由視点における画像面の各画素に対して，仮想カメラの光学中心とその画素を結ぶ直線が Visual Hull を貫通するか否かを判定することにより自由視点画像を生成する手法



図 6 Panoramic Video Texture の 1 フレーム [26]



図 7 セグメンテーションベースドステレオにより推定された奥行き画像（左）と自由視点画像例（右） [28]

を提案している．石川ら [28] は，複数の全方位カメラによって同期撮影された動画画像を基に，動的環境において自由視点画像を生成する手法を提案している．この手法では，まず，撮影環境を静的領域と動的領域に分離する．静的領域に対しては予め与えられた対応点から，環境の疎な 3 次元形状を復元し，3 次元形状を考慮したモーフィングにより高速な自由視点画像生成を行う．また，動的領域に対しては Visual Hull を用いた自由視点画像生成を行い，それぞれの領域の自由視点画像を合成することにより，環境全体の自由視点画像を生成している．Zitnick ら [29] の手法では，まずカメラアレイにより蓄積した多視点映像を基に，セグメンテーションベースドステレオによって，各カメラに対する奥行き画像を計算す

る．映像提示の際は，自由視点位置に隣接した二つのカメラの奥行き画像から二つの2.5次元モデルを生成し，それらを自由視点画像面に投影することで自由視点画像を生成する（図7）．

これらの手法は，対象環境の幾何形状を併用することで，動的環境における自由視点画像生成を実現しているが，幾何形状を取得するためには複数の同期撮影された動画像が必要であり，データの取得にかかるコストが大きいという問題がある．

2.2 本研究の位置付けと方針

前節までで述べたように，IBRに基づく仮想化現実空間の構築手法には，全方位動画像の一部切り出しによる手法，画像群の変形に基づく手法，光線情報を用いる手法，映像の繰り返し再生と，幾何形状を併用する手法による手法が存在する．これらの手法の特徴と提案手法の位置付けを表1にまとめる．固定地点または離散的な位置で撮影された全方位動画像の一部切り出しによる手法では，物体の運動再現については撮影時間のみ実現し，滑らかな視点移動については実現されていない．移動を伴って撮影された全方位動画像の一部切り出しによる手法，明示的に対応点を明示的に与える画像群の変形に基づく手法と，光線情報を用いる手法は，物体の運動再現を実現していないが，滑らかな視点移動については実現されている．明示的に対応点を与えない画像群の変形に基づく手法と，幾何形状を併用する手法は，物体の運動再現については撮影時間のみ実現し，滑らかな視点移動については実現されている．また幾何形状を併用する手法は，入力

表1 従来のIBRに基づく仮想化現実空間の構築手法の特徴と提案手法の位置付け

IBRに基づく仮想化現実空間の構築手法	入力	取得	物体の運動再現(任意時間)	視点移動
全方位動画像の一部切り出しによる手法	離散的な位置で撮影	○:容易	△:撮影時間のみ実現	×:不連続
	移動撮影	○:容易	×:未実現	○:滑らか(撮影経路上)
画像群の変形に基づく手法	明示的な対応点あり	○:容易	×:未実現	○:滑らか(撮影地点間)
	明示的な対応点なし	○:容易	△:撮影時間のみ実現	○:滑らか(撮影地点間)
光線情報を用いる手法		○:容易	×:未実現	○:滑らか(環境内)
映像の繰り返し再生による手法		○:容易	○:実現	×:未実現
画像変形に幾何形状を利用する手法		×:困難	△:撮影時間のみ実現	○:滑らか(環境内)
提案手法		○:容易	○:実現	○:滑らか(撮影地点間)

得が困難であるという問題がある．映像の繰り返し再生による手法は，物体の運動再現については実現し，滑らかな視点移動については実現されていない．このように，従来，仮想化現実空間を構築するための様々な手法が提案されているが，入力を容易に取得でき，また物体の運動再現を実現しつつ滑らかな視点移動を実現する手法は提案されていない．

そこで，本論文では，動的環境を対象に，容易に取得可能な入力を用いて，物体の運動再現を実現しつつ滑らかな視点移動を実現する手法を提案する．提案手法は，動画像を違和感なく任意時間の動画像に合成するビデオテクスチャおよび画像群の変形による中間視点画像生成手法を基礎として，離散的な位置で異なる時間に撮影した複数の動画像を入力とし，(1) 周期運動物体を対象とした映像の繰り返し再生，(2) 撮影地点間での類似フレームの接続による中間視点画像生成，を行うことで，視点移動可能なビデオテクスチャを実現する．具体的には，まず最初に，各撮影地点で撮影した動画像に対して，それぞれ時間軸方向に画素値の中央値を求め背景画像を生成し，背景差分により各フレームにおける動物体領域画像を算出する．次に，各動画像中で時間的連続性を考慮した類似フレーム探索を行い，各撮影地点における映像の繰り返し再生を可能にする．隣接する視点位置で取得した動画像に対しても，時間的連続性を考慮した類似フレーム探索を行い，類似フレーム間で映像を接続することにより，視点移動時の時間的な連続性を保持した映像合成を実現する．ただし，撮影位置の異なる画像間での類似フレーム探索には，対象の見え方の変化を考慮する必要がある．これに対して本研究では，周期運動物体を平面と仮定し，各視点で撮影された画像を射影変換することで対応する．また，隣接する視点において接続された類似フレームの前後のフレームを用いて中間視点画像を生成することで，視点移動時の空間的連続性を実現する．

3. 視点移動可能なビデオテクスチャ

3.1 提案手法の概要

本研究では、動画像を違和感なく任意時間の動画像に合成するビデオテクスチャおよび画像群の変形による中間視点画像生成手法を基礎として、離散的な位置で異なる時間に撮影した複数の動画像を入力とし、

- (1) 周期運動物体を対象とした映像の繰り返し再生
- (2) 撮影地点間での類似フレームの接続による中間視点画像生成

を行うことで、視点移動可能なビデオテクスチャを実現する。ただし、本研究では、直線経路上を等間隔で撮影し、すべてのカメラに共通の周期運動物体が写っていると仮定する。提案手法の処理の流れを図8に示す。本手法はオフラインでの動画像中の類似フレーム探索処理(A)とオンラインでのビデオテクスチャ生成処理(B)からなる。

オフラインでの動画像中の類似フレーム探索処理では、まず、離散的な位置で異なる時間に撮影された複数の動画像に対して、それぞれ時間軸方向に画素値の中央値を求め背景画像を生成し、背景差分により各フレームにおける動物体領域を算出する(A-1)。次に、各動画像中で時間的連続性を考慮した類似フレーム探索を行い、繰り返し再生の範囲を決定する(A-2)。隣接する視点位置で取得した動画像に対しても、時間的連続性を考慮した類似フレーム探索を行い、視点移動時の接続フレームを決定する(A-3)。

オンラインでのビデオテクスチャ生成処理では、ユーザによる視点位置の設定(B-1)、中間視点位置における再生フレームの決定(B-2)、設定された視点位置における中間視点画像の生成を行う(B-3)。ただし、再生フレームが(A-2)で算出した繰り返し再生範囲の最終フレームに到達した場合には、再生フレームの巻き戻しを行う。また、視点位置が切り替わる際には適切なフレームへの接続を行い、視点切り替えを行う(B-2)。これらの処理を繰り返すことで視点移動可能なビデオテクスチャを生成する。以下では、各処理について詳述する。

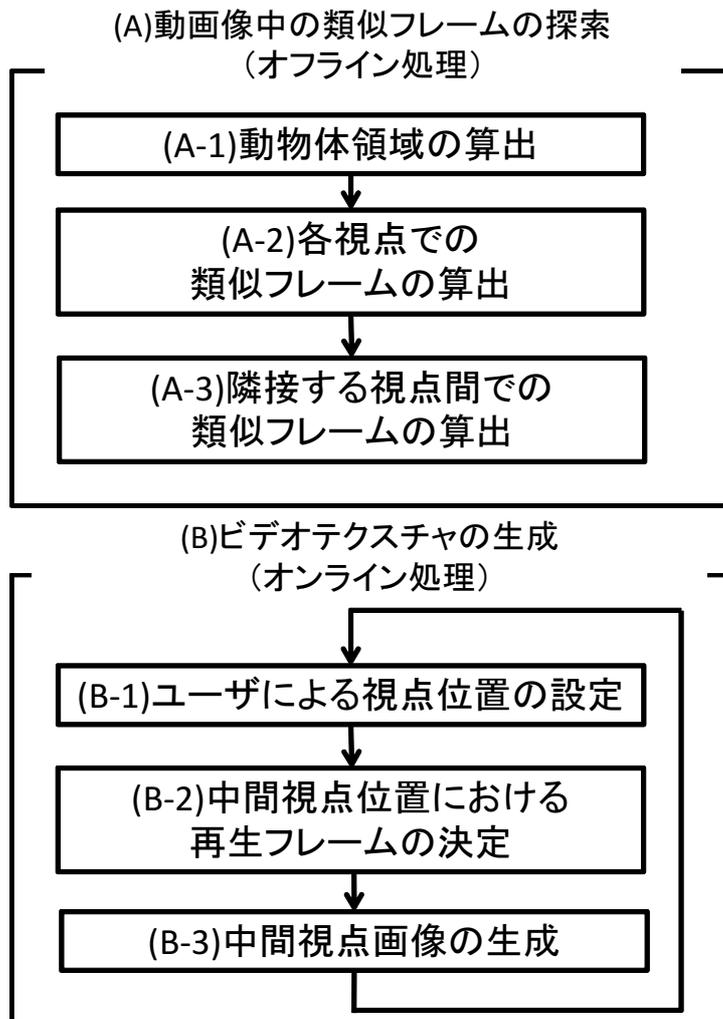


図 8 視点移動可能なビデオテクスチャ生成処理の流れ

3.2 動画像中の類似フレームの探索

3.2.1 動物体領域の算出

本研究では離散的な位置で非同期に一定時間固定して撮影された複数の動画像を入力として用いる．ここでは動物体を対象とした類似フレーム探索および中間視点画像生成で用いる動物体領域を算出する．まず，撮影時間の 50% 以上において背景が撮影されているという仮定の下，各地点で撮影された動画像 ($p = 0, 1, 2, \dots, M-1$) の各画素において時間軸方向に画素値の中央値を取るメディアンフィルタにより，背景画像 B_p を算出する．次に，背景差分により，各地点で撮影した動画像の各フレームにおける動物体領域画像を二値画像として得る．ただし，ここではノイズの影響を軽減するため，各地点の背景画像と各フレームの入力動画像に対しガウシアンフィルタを用いて平滑化処理を行った上で，背景差分を適用することにより，動画像 p のフレーム m における動物体領域画像 $I_p(m)$ を算出する．

3.2.2 各視点位置における時間的連続性を考慮した類似フレーム探索

提案手法では，オンライン処理においてユーザに対する映像提示が行われる間，周期運動物体の運動の連続性を保持した映像の繰り返し再生を行う (B-3)．本研究では，従来提案されている固定地点を対象とした映像の繰り返し再生による手法 [22] を用い，各地点における周期運動物体の運動を再現する．このために，まず，動物体領域画像を用いて，各動画像中で時間的連続性を考慮した類似フレーム探索を行う．ここで，時間的連続性とは時間的に連続するフレームを考慮することを表し，これにより運動の連続性を保持した繰り返し再生時の滑らかな映像接続を実現する．類似フレームの探索には，各動物体領域画像の排他的論理和領域の画素数を用い，これを最小にするフレーム対を探索する．具体的には，以下の式より，動画像 p における類似フレーム対 (m_p, n_p) を算出し，提示時の動画像の繰り返し再生の範囲を $(m_p + 1, n_p)$ とする (図 9)．

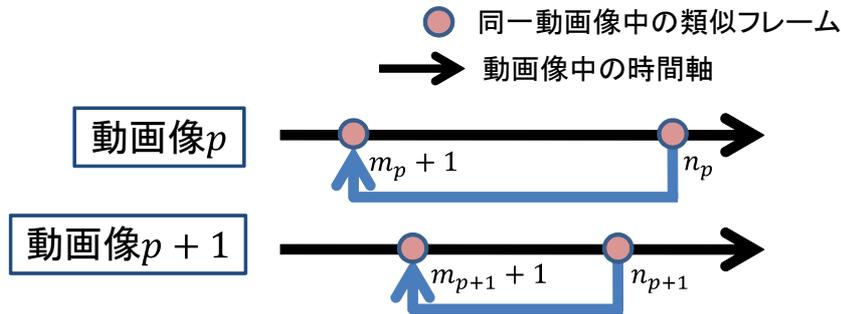


図 9 類似フレーム間における繰り返し再生

$$(m_p, n_p) = \operatorname{argmin}_{m, n} \left[\sum_{i=0}^{w-1} \text{XOR}(I_p(m+i), I_p(n+i)) \right] \quad (1)$$

ただし， $\text{XOR}(I_p(m), I_p(n))$ は，動物体領域画像 $I_p(m), I_p(n)$ の排他的論理和領域の画素数， w は時間的連続性を考慮するためのフレーム数を表す．また，一定の再生時間を確保するために $n_p - m_p \geq k$ (k は定数) とする．

3.2.3 隣接する視点位置に対する時間的連続性を考慮した類似フレーム探索

隣接する視点位置で取得した動画画像 ($p, p \pm 1$) に対して，時間的連続性を考慮した類似フレーム探索を行い，視点移動時に類似フレーム間で映像を滑らかに接続することにより，運動の連続性を保った視点移動を実現する．ただし，撮影位置の異なる画像間での類似フレーム探索には，対象の見え方の変化を考慮する必要がある．ここでは，周期運動物体を平面と仮定し，各視点で撮影された画像を射影変換することにより見え方の変化を補償する．具体的には，あらかじめ以下の式より動画画像 p の各フレーム s に対する動画画像 ($p \pm 1$) 上の類似フレーム $t_{(p \pm 1)}(s, p)$ を算出しておき，動画画像 p から動画画像 $p \pm 1$ への視点移動時には，動画画像 p のフレーム s を動画画像 $p \pm 1$ のフレーム $t_{(p \pm 1)}(s, p) + 1$ に接続し，再生する．

$$t_{(p \pm 1)}(s, p) = \operatorname{argmin}_t \left[\sum_{i=0}^{w-1} \text{XOR}(H_{(p, p \pm 1)}(I_p(s+i)), I_{(p \pm 1)}(t+i)) \right] \quad (2)$$

ただし, $H_{(p,p\pm 1)}(I_p)$ は, 動画像 $(p, p \pm 1)$ 間の射影変換パラメータを用い, 動物体領域画像 I_p を射影変換した画像を表す. ここで射影変換パラメータは, 各視点位置において背景画像中で周期運動物体と同一平面上の物体に対し手動で対応点を与えることにより算出する. また, 移動先の視点において繰り返し再生を可能とするために $t_{(p\pm 1)}(s, p) + 1 < n_{(p\pm 1)}$ とする.

3.3 視点移動可能なビデオテクスチャの生成

視点移動可能なビデオテクスチャの生成処理では, オフラインでの類似フレーム探索処理で求めた同一動画像中での繰り返し再生範囲および視点移動時の接続フレームを用い, ユーザの入力に応じた視点位置のビデオテクスチャをオンラインで生成する.

3.3.1 ユーザの入力による視点位置の設定

提案手法では, ユーザは入力動画像の撮影地点間を結ぶ経路上で視点位置を設定することができる. ただしオフラインで算出された類似フレームによる映像接続のみによる視点移動では, 提示映像における対象環境の空間的連続性は保持されない. そこで本研究では, 各撮影地点間の距離は同じであるという仮定の下, 初期撮影位置と最終撮影位置の間に中間視点を設定し, 中間視点の画像を合成する. 具体的には, 中間視点位置 q_i を $0 \leq q_i \leq M - 1$ の範囲の実数で指定する. ただし, M は撮影地点数, i は生成される動画像のフレーム番号を表す. ここで中間視点位置 q_i は, 初期撮影位置 0 と最終撮影位置 $M - 1$ 間を $q_i : (M - 1 - q_i)$ に内分する視点位置を表す (図 10). また, 中間視点位置 q_i における周期運動物体に対する仮想的な対応点の座標も同様に, 初期撮影位置 0 での対応点の座標と最終撮影位置 $M - 1$ の対応点の座標の間を $q_i : (M - 1 - q_i)$ に内分する座標となり, 中間視点画像の生成における画像の射影変換時にこれを用いる.

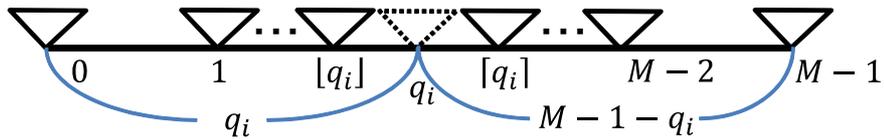
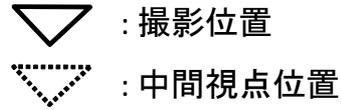


図 10 撮影地点間を等距離と仮定した直線上における中間視点位置

3.3.2 中間視点画像の生成

本研究では，動物体を平面を仮定し，画像の平面射影変換を用いることで，中間視点画像を生成する．ここで，中間視点 q_i の画像生成には撮影位置が最も近い 2 地点の動画像を用いる．(以下，これら 2 地点を床関数と天井関数を用いて $[q_i], [q_i]$ と表す) また，中間視点 q_i と撮影位置 $[q_i], [q_i]$ 間の射影変換パラメータは，撮影位置 $[q_i], [q_i]$ における対応点の座標と中間視点 q_i における仮想的な対応点の座標から算出される．本処理は，以下で述べる背景画像生成と動物体画像生成の処理から成る．

[背景画像の生成]

背景画像の生成には，中間視点 q_i に隣接する 2 視点 $[q_i], [q_i]$ における背景画像 $B_{[q_i]}, B_{[q_i]}$ を用いる．まず，背景画像 $B_{[q_i]}, B_{[q_i]}$ を，図 12(a) に示すように，平面射影変換により中間視点 q_i における画像に変形する．次に，視点移動時の映像接続における合成画像上の位置ずれによる違和感を軽減するため，変形後の二枚の背景画像をブレンディングし，中間視点 q_i における背景画像 B_{q_i} を生成する．ここで背景画像の混合率は視点位置に応じて，線形に変化させる．具体的には以下の式より，背景画像を生成する．また，図 11 に $q_i = 0.5$ の場合の背景画像生成の流れを示す．

$$B_{q_i} = ([q_i] - q_i)H_{([q_i], q_i)}(B_{[q_i]}) + (q_i - [q_i])H_{([q_i], q_i)}(B_{[q_i]}) \quad (3)$$

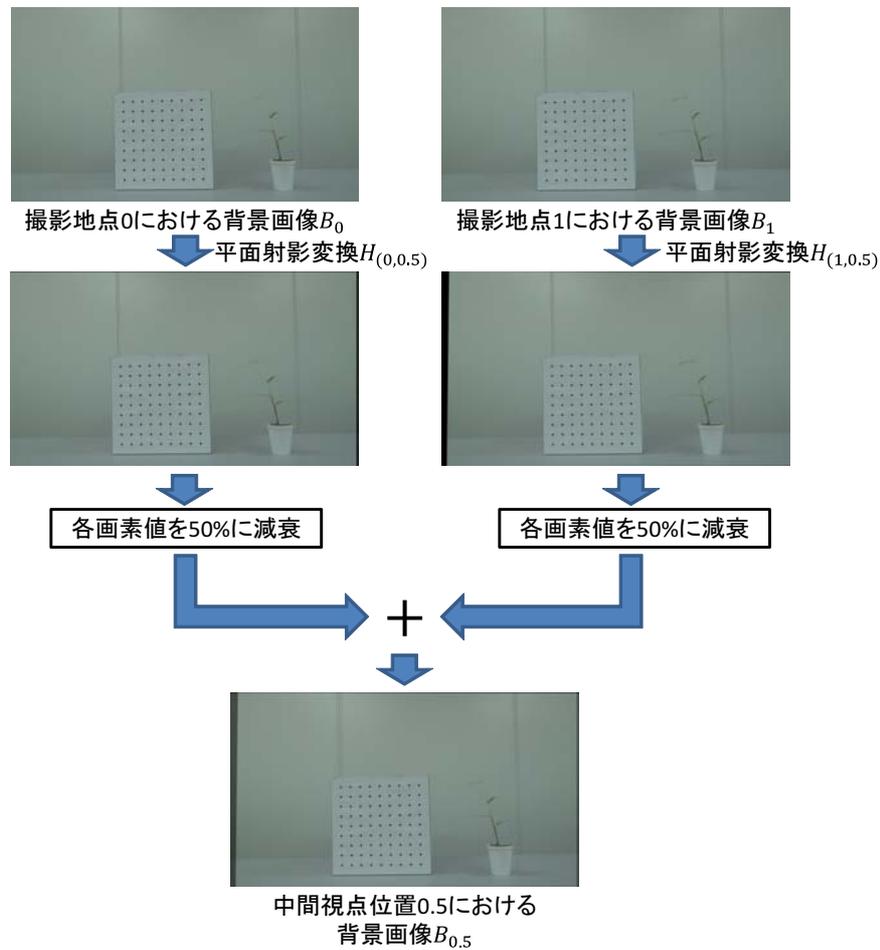


図 11 背景画像生成処理の流れ ($q_i = 0.5$)

[動物体画像の生成]

本研究では動物体領域を対象とした画像の位置合わせを行うため，合成画像上において背景画像上に生じるような視点間での大きな位置ずれは生じない．従って，ここではぼけや違和感の原因となるブレンディングを行わず，中間視点から最も近い視点位置 \bar{q}_i ($\lfloor q_i \rfloor$ または $\lceil q_i \rceil$) の動物体領域を，図 12(b) に示すように，射影変換により変形することで中間視点における動物体画像を生成する．具体的には，以下の式より動物体領域を変形する．

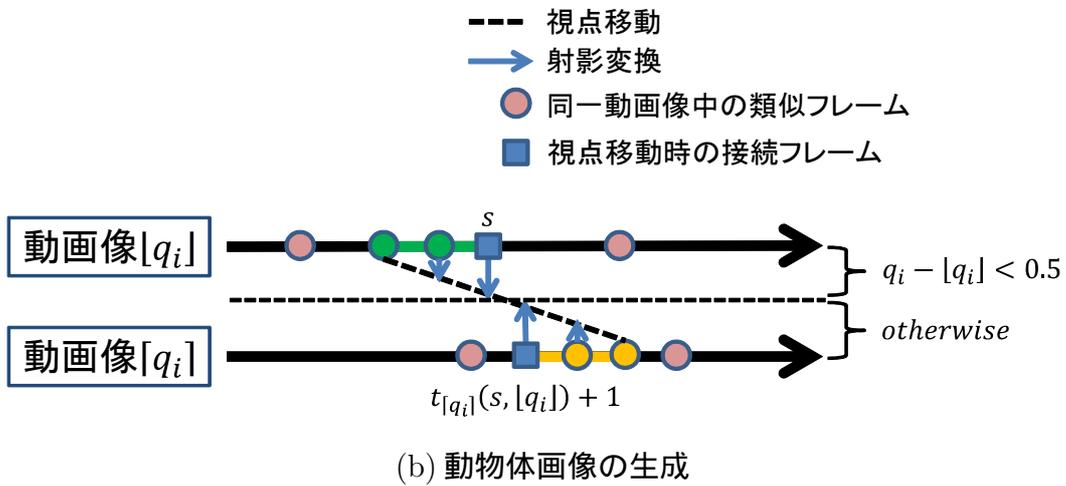
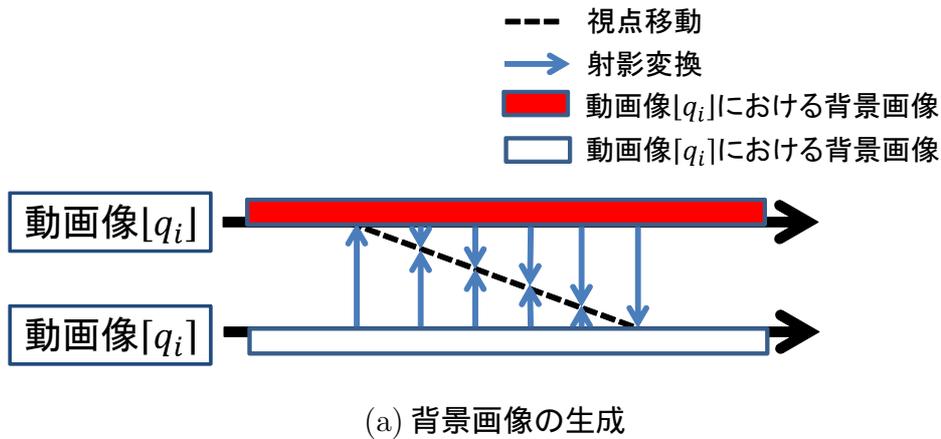


図 12 中間視点画像の生成 (隣接地点間を等速に視点移動した場合)

$$I_{q_i} = H_{(\bar{q}_i, q_i)} I_{\bar{q}_i}(j_i) \quad (4)$$

$$\bar{q}_i = \begin{cases} [q_i] & ; q_i - [q_i] < 0.5 \\ [q_i] & ; otherwise \end{cases} \quad (5)$$

ただし, j_i は動画像の再生に用いるフレーム番号であり, 次項で述べる手法により更新する.

以上の処理により, 算出された動物体画像の動物体領域 I_{q_i} を背景画像 B_{q_i} に書き込むことにより, 中間視点画像を生成する. また, 図 13 に $q_i = 0.5$ の場合の中

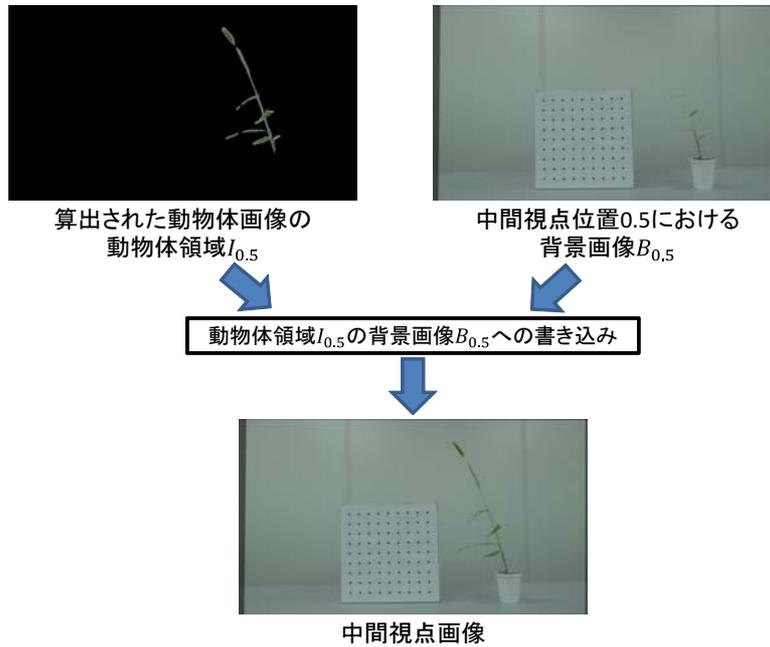


図 13 中間視点画像生成処理の流れ ($q_i = 0.5$)

間視点画像生成の流れを示す。

3.3.3 中間視点位置における再生フレームの決定

中間視点位置 q_i の合成に用いる撮影位置 \bar{q}_i のフレーム番号 j_i を以下のように決定する。

$$j_i = \begin{cases} m_{\bar{q}_i} + 1 & ; (\bar{q}_i = \bar{q}_{i-1}) \wedge (j_{i-1} = n_{\bar{q}_i}) \\ j_{i-1} + 1 & ; (\bar{q}_i = \bar{q}_{i-1}) \wedge (j_{i-1} < n_{\bar{q}_i}) \\ t_{\bar{q}_i}(j_{i-1}, \bar{q}_{i-1}) + 1 & ; otherwise \end{cases} \quad (6)$$

生成される動画の $i - 1$ フレーム目での \bar{q}_{i-1} と i フレーム目での \bar{q}_i が同一である場合，最も近い撮影位置 \bar{q}_i で算出された類似フレーム対 $(m_{\bar{q}_i}, n_{\bar{q}_i})$ に基づいて繰り返し再生することにより，ユーザ設定の視点位置に関わらず提示映像を任意時間再生し続けることが可能となる。また，視点位置が切り替わる場合，つまり

$i - 1$ フレーム目での $\overline{q_{i-1}}$ と i フレーム目での $\overline{q_i}$ が異なる場合，オフラインで算出された動画像 $\overline{q_{i-1}}$ の j_{i-1} フレーム目に対する接続フレーム $t_{\overline{q_i}}(j_{i-1}, \overline{q_{i-1}}) + 1$ へ映像接続を行い，視点切り替えを実現する．

4. 実験

本節では，離散的な位置で異なる時間に撮影した複数の動画像を用いて，単一の周期運動物体が存在すると見なせるシーン（以下，単一の周期運動物体と呼ぶ）および複数の周期運動物体が存在するシーン（以下，複数の周期運動物体と呼ぶ）を対象とし，視点移動可能なビデオテクスチャの生成結果を示す．

4.1 単一の周期運動物体を対象とした視点移動可能なビデオテクスチャ

4.1.1 実験環境

単一の周期運動物体として風で揺らした植物を対象とし，図 14 に示すような視点の等速直線運動を仮定した経路設定を行い，計 407 フレームの視点移動を伴うビデオテクスチャを生成した．撮影地点はおおむね直線上で，撮影地点間の距離はほぼ等距離となるように目視で確認し，15 地点で入力動画像（各撮影位置で 500 フレーム）を撮影した．また，画像上での対応点を正確に与えるために，環境内に目印の付いた箱を設置した．図 15 に入力として用いた各撮影地点における動画像の第 1 フレームを，表 2 に本実験のパラメータを示す．図 15 から，撮影位置が異なること，非同期に撮影されていることが確認できる．図 16 に各視点位置において生成された背景画像および背景画像上に手動で与えた対応点（赤い点）を示す．表 2 に実験時のパラメータを示す．

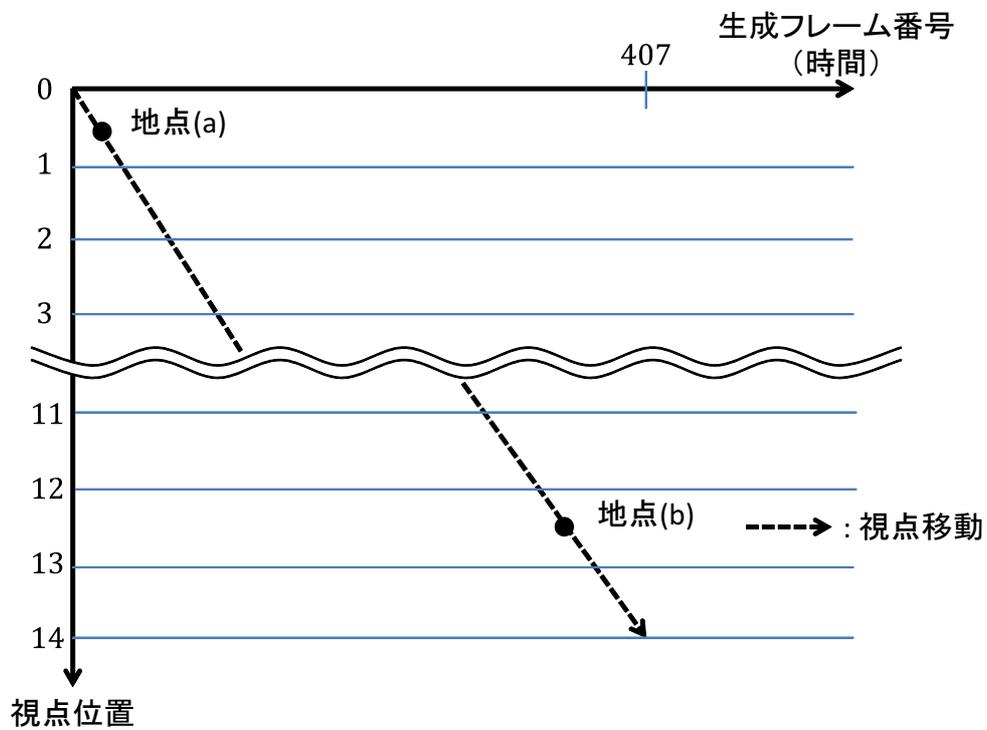


図 14 ビデオテクスチャの生成視点位置 (単一の周期運動物体)

表 2 実験時のパラメータ

入力動画像の解像度	横 640[pixel] 縦 360[pixel]
ガウシアンフィルタのサイズ	3 × 3
ウィンドウサイズ w	3
一定の再生時間を確保するための定数 k	100

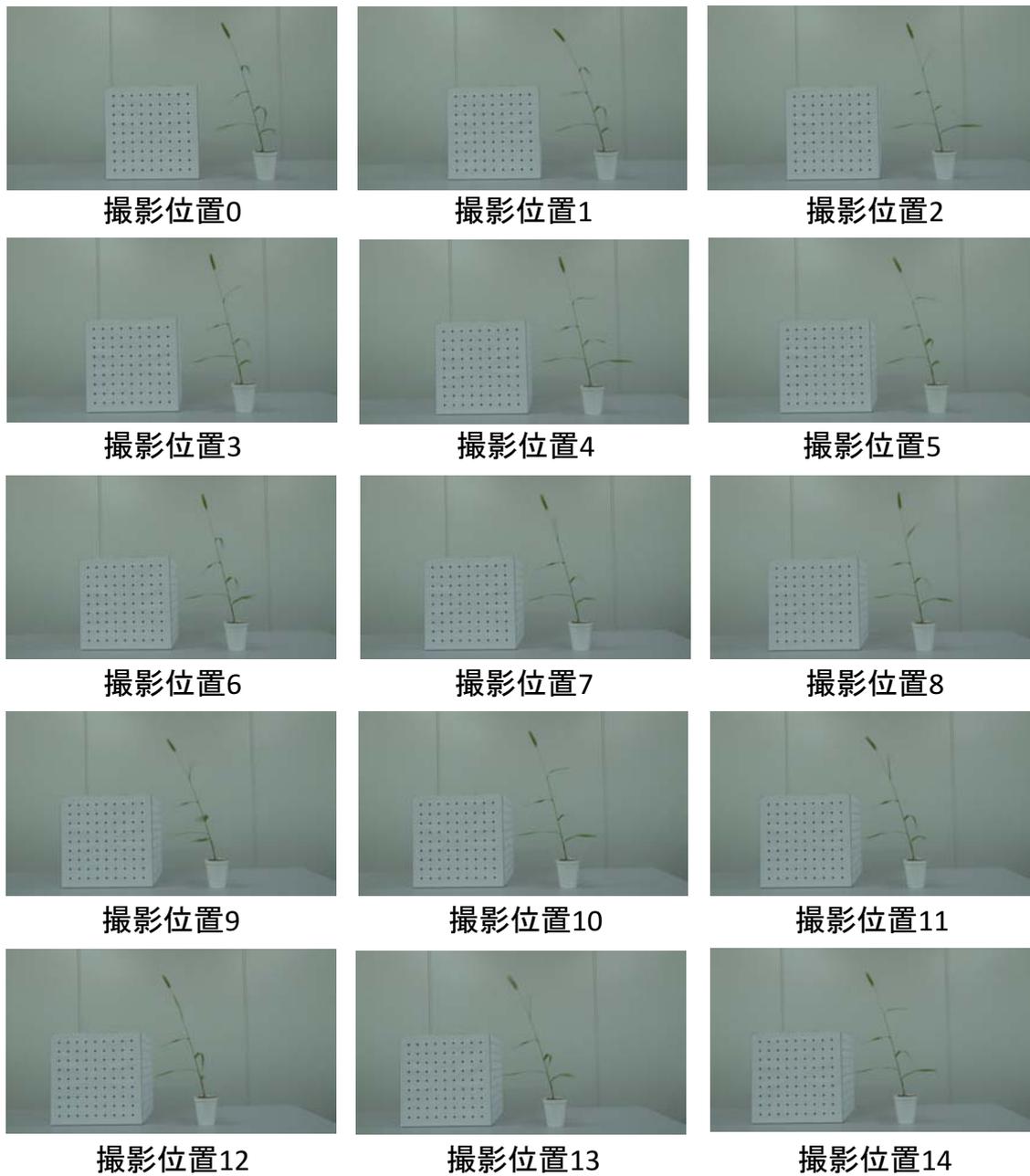


図 15 各撮影位置における動画像の第 1 フレーム (単一の周期運動物体)



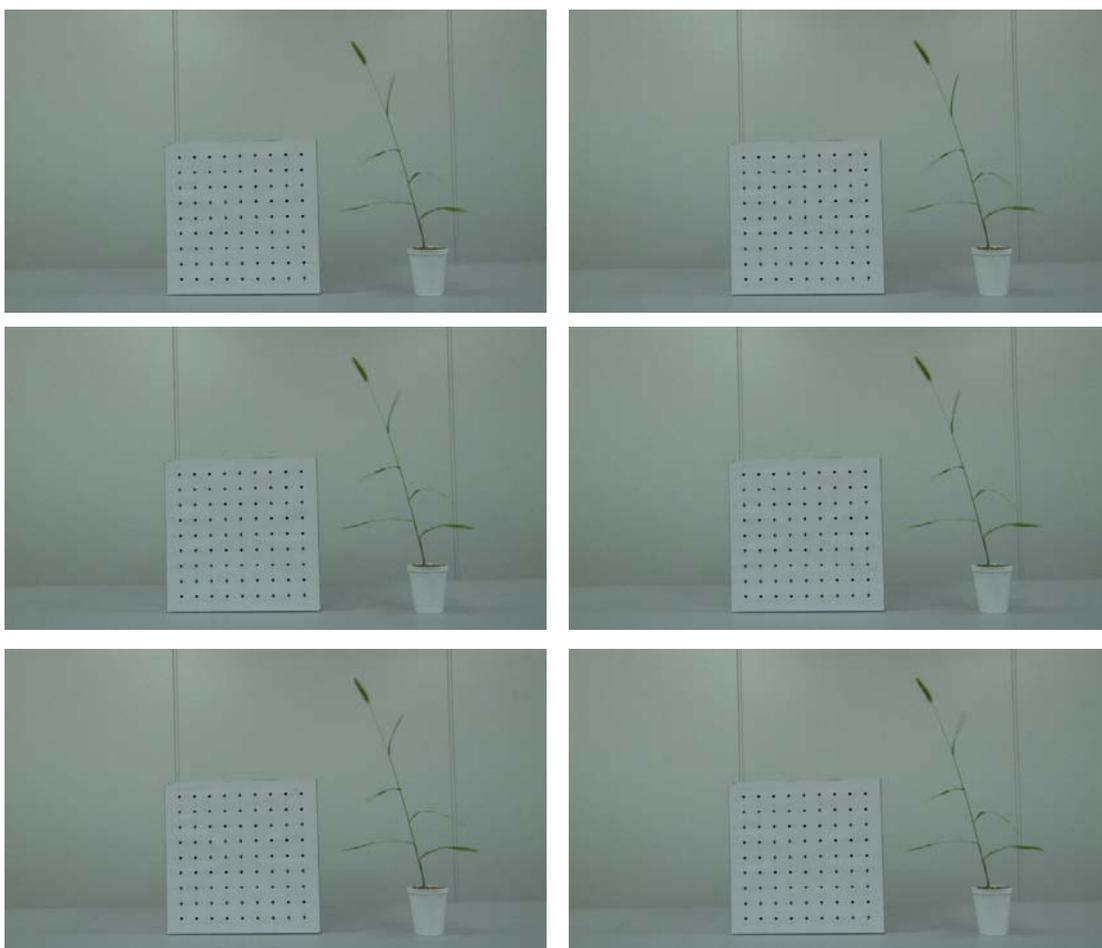
図 16 各撮影位置において生成された背景画像と手動で与えた対応点（単一の周期運動物体）

4.1.2 実験の結果と考察

表3に算出された各撮影位置における類似フレーム対 (m_p, n_p) を、図17,18に撮影地点0において算出された類似フレーム対と対応する動物体領域画像を示す。それぞれの撮影位置において運動の連続性が類似した接続フレームが算出され、提示映像では違和感のない繰り返し再生が実現できることを確認した。図14中に示した動物体画像が切り替わる地点(a)(b)の前後のフレームにおいて生成されたビデオテクスチャを図19,20に示す。図19,20より、運動の時間的・空間的連続性が保持された映像の接続が実現されていることが確認できる。ただし、一部のフレームにおいて、図21の拡大図に示すように、提示映像の背景のテクスチャが二重になるアーティファクト（実線の枠）や、対象物体の細い部分が表示されない箇所（破線の枠）が見られた。これは、動物体の近似に用いた平面から離れた領域では平面射影変換による誤差が大きく、背景画像生成時のレンディング処理においてアーティファクトが発生したことが原因と考えられる。また、対象物体の細い部分については平滑化処理の影響により動物体領域として抽出されず、提示映像において表示されなかったことが考えられる。

表 3 算出された各撮影位置における類似フレーム (m_p, n_p)

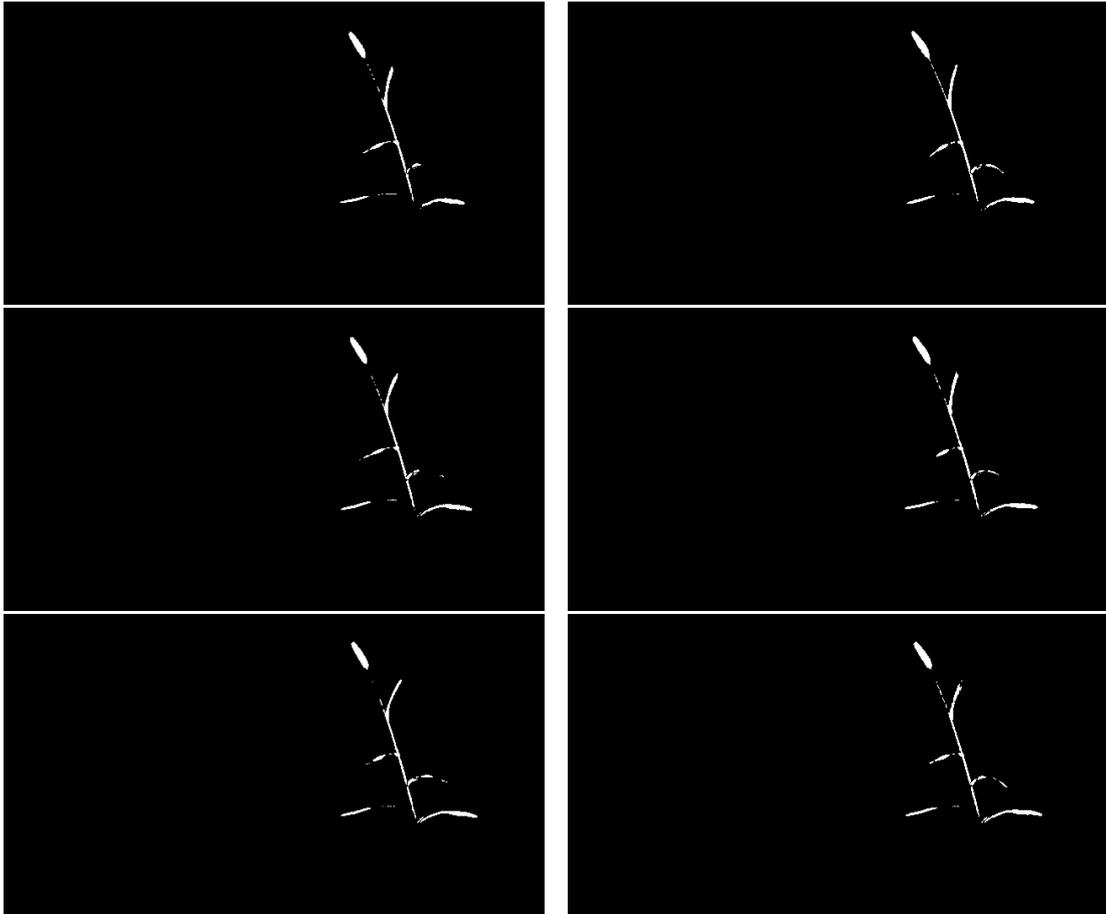
	(m_p, n_p)		(m_p, n_p)
撮影位置 0	(283,399)	撮影位置 8	(63,329)
撮影位置 1	(120,445)	撮影位置 9	(26,491)
撮影位置 2	(137,272)	撮影位置 10	(50,425)
撮影位置 3	(56,263)	撮影位置 11	(202,322)
撮影位置 4	(29,422)	撮影位置 12	(140,388)
撮影位置 5	(5,490)	撮影位置 13	(70,357)
撮影位置 6	(214,465)	撮影位置 14	(5,320)
撮影位置 7	(38,332)		



上から順に283,284,285フレーム目

上から順に399,400,401フレーム目

図 17 撮影地点 0 において算出された類似フレーム (単一の周期運動物体)



上から順に283,284,285フレーム目

上から順に399,400,401フレーム目

図 18 撮影地点 0 において算出された類似フレームに対応する動物体領域画像
(単一の周期運動物体)

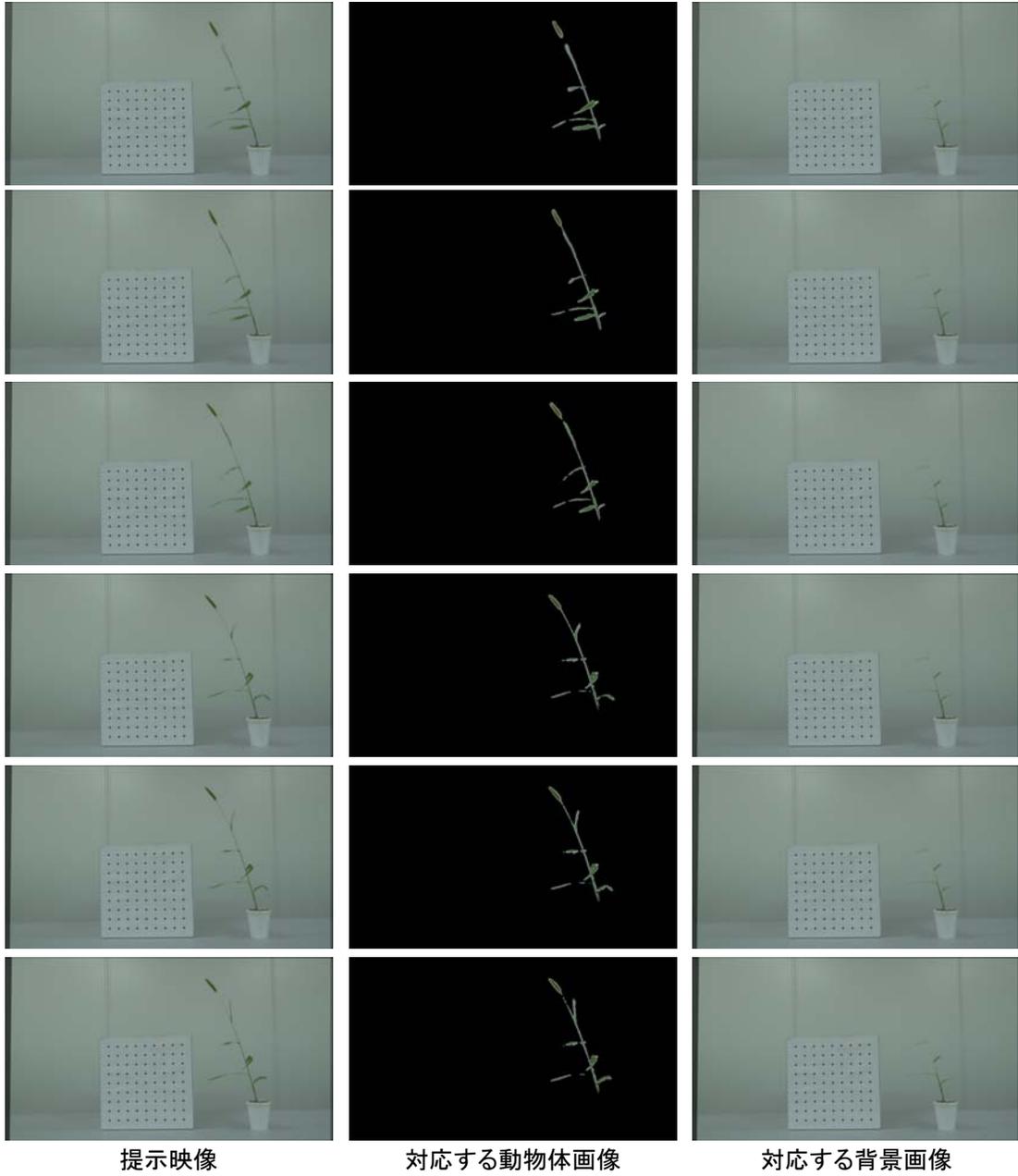


図 19 地点 (a) の前後のフレームの提示映像と対応する動物体画像と背景画像
(上から順に 13,14,15,16,17,18 フレーム目)

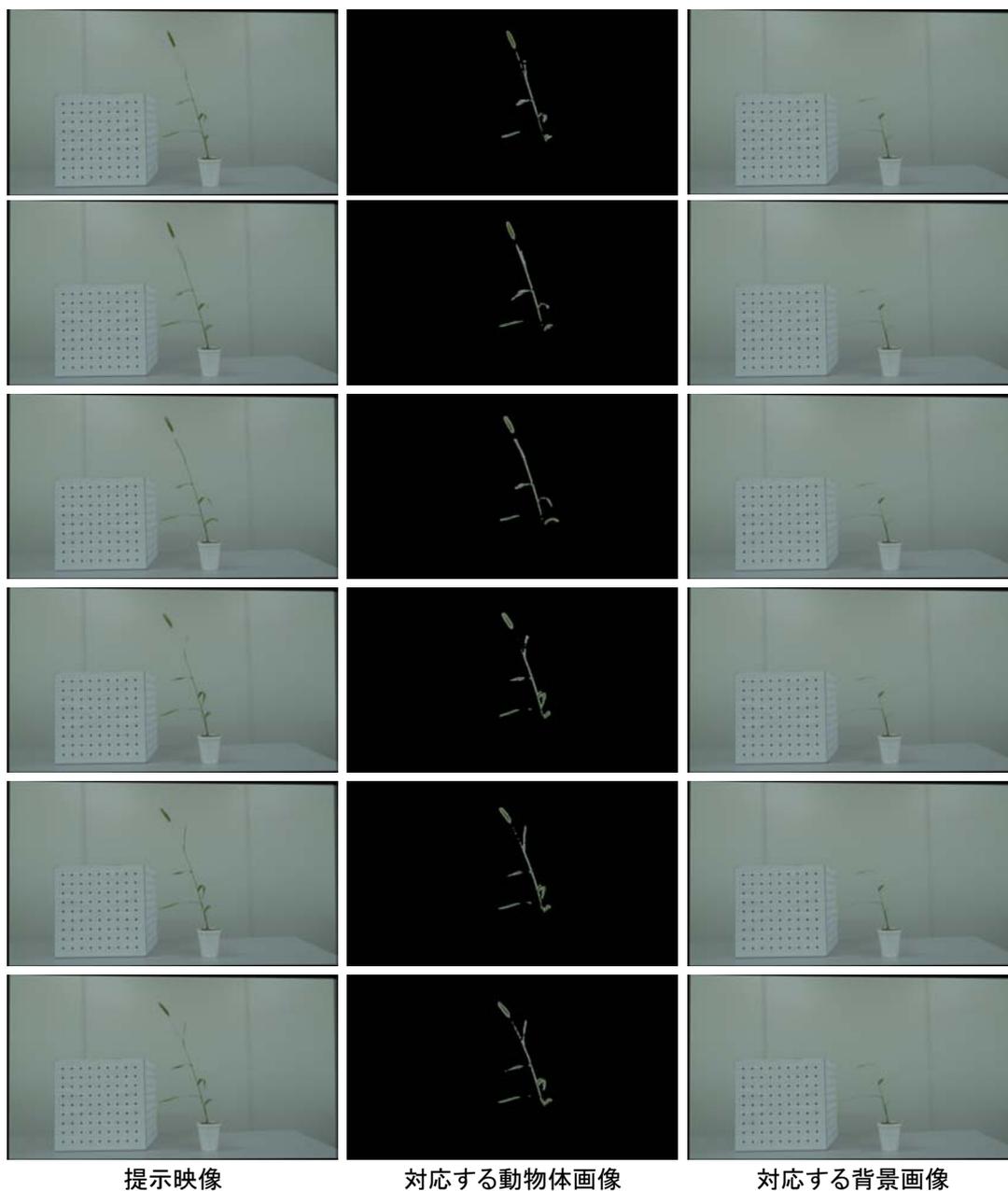


図 20 地点 (b) の前後のフレームの提示映像と対応する動物体画像と背景画像
 (上から順に 361,362,363,364,365,366 フレーム目)

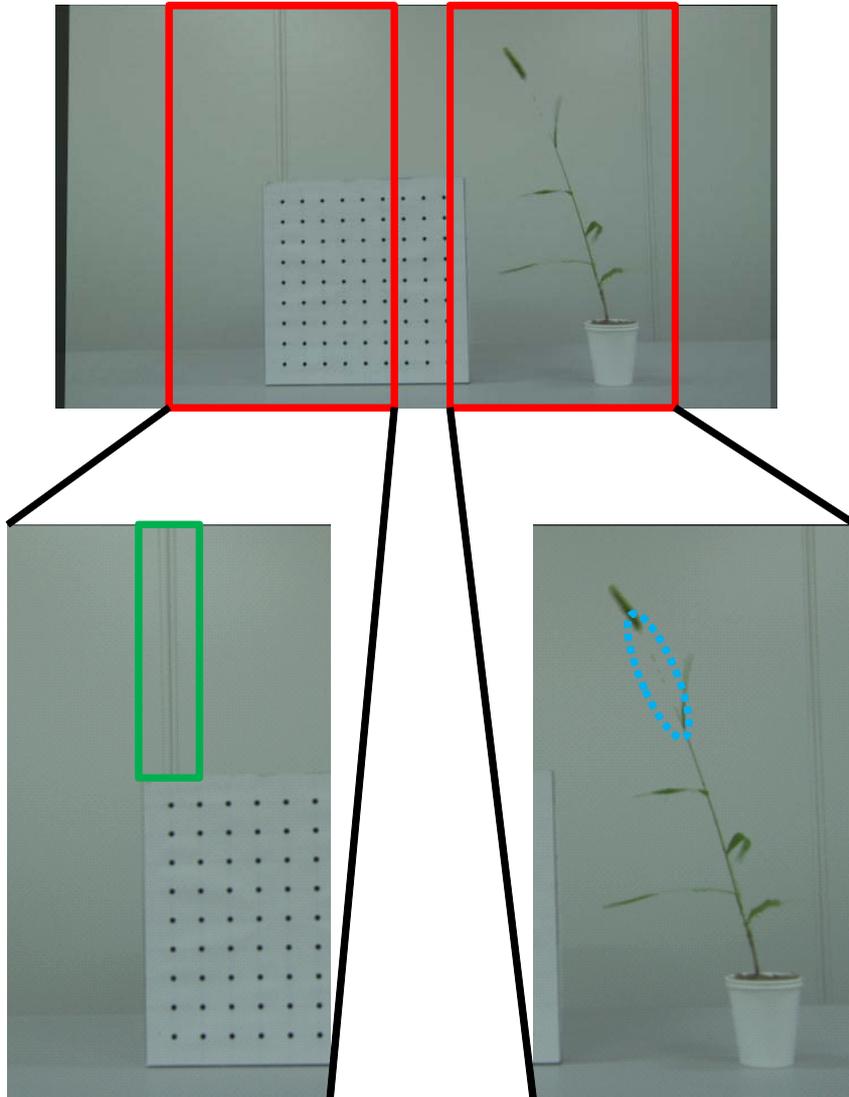


図 21 21 フレーム目の提示映像と拡大画像

4.1.3 被験者による評価実験と考察

本実験は、被験者により生成された映像を主観的に評価してもらうことで、提案手法の有効性を客観的に示すことを目的とする。

[評価方法]

16名の被験者に、本実験で用いる映像が一本以上の映像を接続して合成したものであることを説明し、提示映像における映像接続があるかどうかを判断してもらった。あると判断した場合にはその接続回数を回答してもらい、これを以下の3つの映像に対して3回繰り返してもらった。

- (A-1) 視点移動なし，繰り返し再生なし
- (A-2) 視点移動なし，繰り返し再生あり（2回接続）
- (B) 視点移動あり（14回接続），繰り返し再生なし

[評価結果と考察]

繰り返し再生のない映像(A-1)と繰り返し再生のある映像(A-2)に対する映像接続回数と回答人数の関係を図22に、提案手法を用いた視点移動映像(B)に対する映像接続回数と回答人数の関係を図23に示す。図22より、1回目、2回目の結果では、繰り返し再生のない映像(A-1)で接続があると答えた被験者が少数見られたが、繰り返し再生のある映像(A-2)に対しては全員が接続がないと回答した。また3回目の結果では、繰り返し再生のある映像(A-2)で接続があると答えた被験者が、1回目、2回目に比べ少数見られた。以上のことから、繰り返し再生のない映像(A-1)と繰り返し再生のある映像(A-2)に関して差を知覚することは難しく、違和感のない映像接続がされたと言える。また提案手法を用いた視点移動映像(B)に関しては、図23より、1回目で3回までしか映像接続を判断できなかった被験者は全員の88%であった。2回目は69%、3回目は57%と接続を判断できる被験者が増加したが、14回すべての映像接続を判断できたのは1名のみだった。以上のことから、提案手法を用いた視点移動映像(B)は映像接続に関する違和感を感じることは実際の接続回数よりも少なかったと言える。ただし、映像の

接続を感じた理由として、「茎が二重になる」、「壁にある直線が二重になる」という意見があり、動物体の近似に用いた平面から離れた領域におけるこのようなアーティファクトの軽減が必要であることが本実験からも確認することができた。

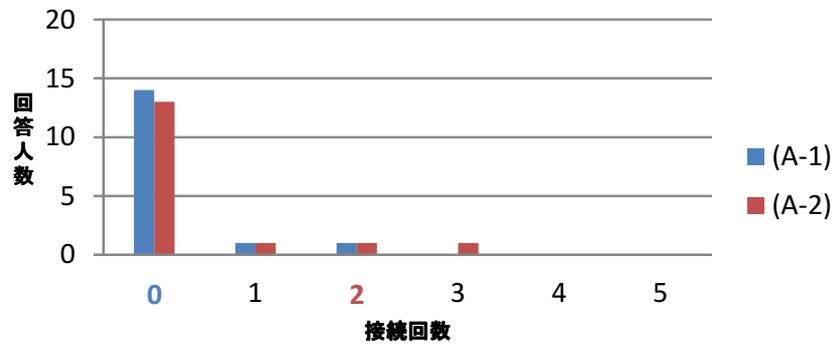
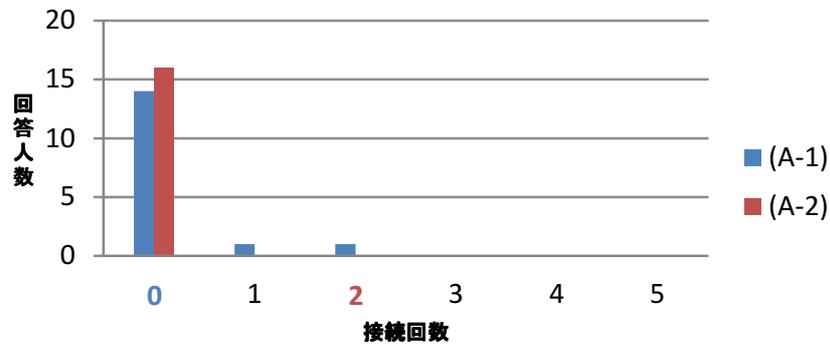
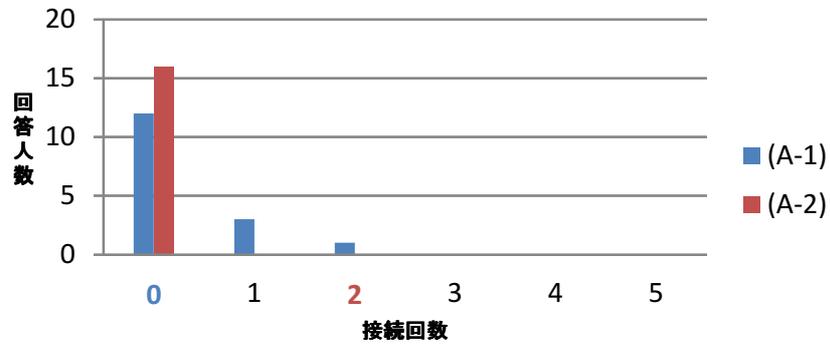


図 22 繰り返し再生のない映像 (A-1) と繰り返し再生のある映像 (A-2) に対する映像接続回数と回答人数の関係

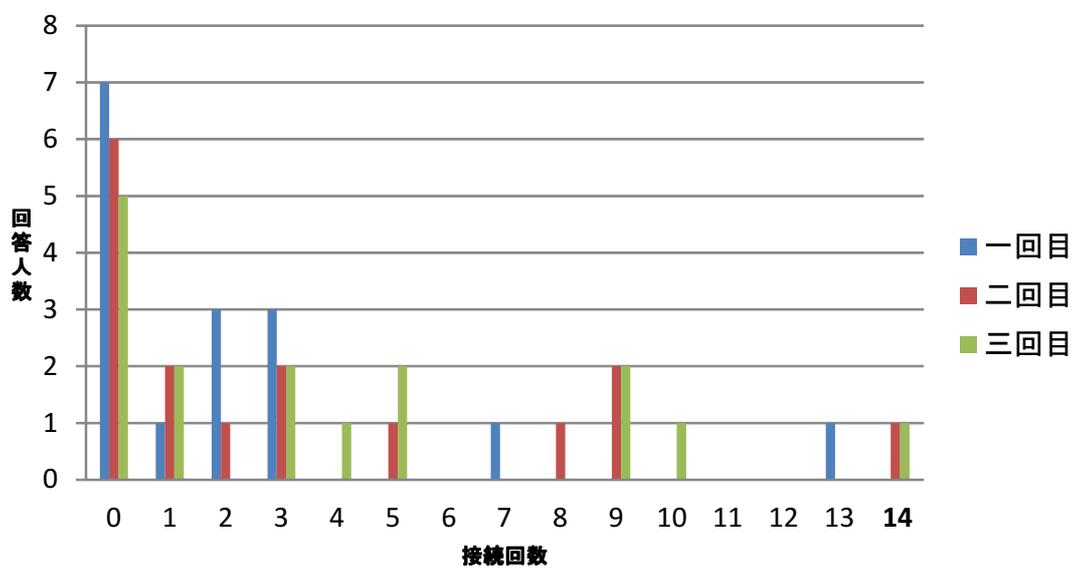


図 23 提案手法を用いた視点移動映像 (B) に対する映像接続回数と回答人数の関係

4.2 複数の周期運動物体を対象とした視点移動可能なビデオテクスチャ

4.2.1 実験環境

複数の周期運動物体として風に揺れる木を対象とし，図 24 に示すような視点の等速運動を仮定した経路設定を行い，計 357 フレームの視点移動を伴うビデオテクスチャを生成した．撮影地点はおおむね直線上で，撮影地点間の距離はほぼ等距離となるように目視で確認し，3 地点で入力動画像（各撮影位置で 500 フレーム）を撮影した．図 25 に入力として用いた各撮影地点における動画像の第 1 フレームを示す．図 26 に各視点位置において生成された背景画像および背景画像上に手動で与えた対応点（赤い点）を示す．本実験で用いたパラメータは 4.1 の実験と同一である．

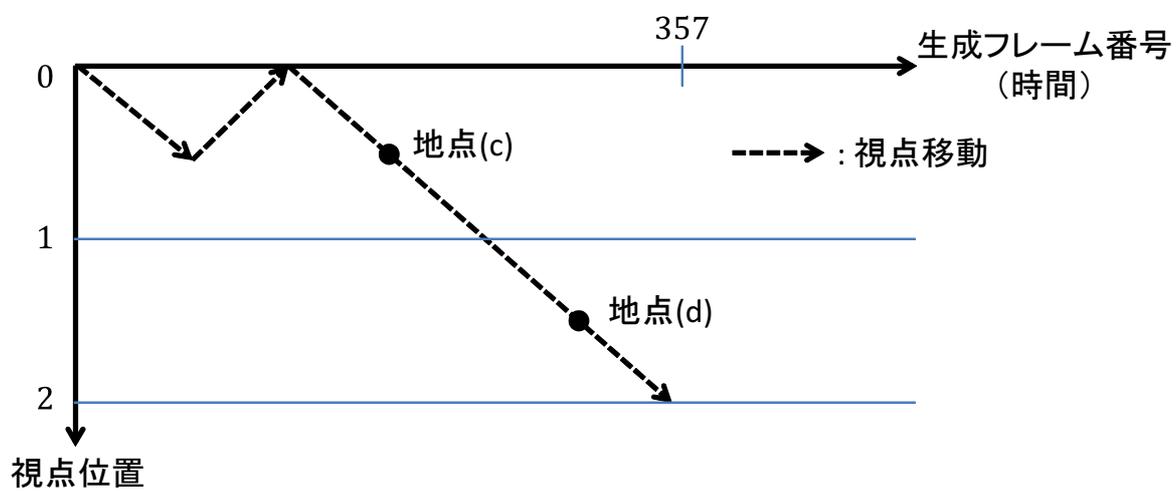


図 24 ビデオテクスチャの生成視点位置 (複数の周期運動物体)



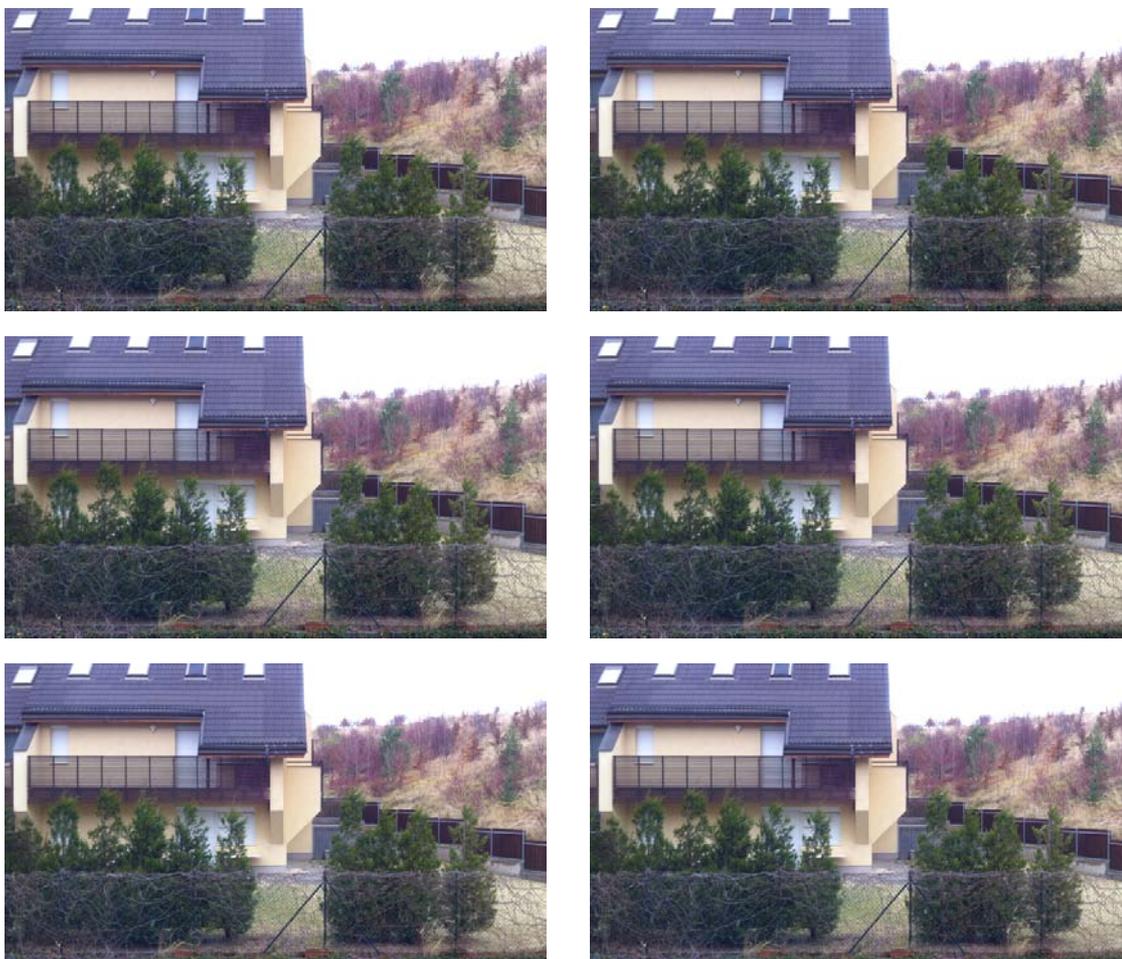
図 25 各撮影位置における動画像の第 1 フレーム（複数の周期運動物体）



図 26 各撮影位置において生成された背景画像と手動で与えた対応点（複数の周期運動物体）

4.2.2 実験の結果と考察

表 4 に算出された各撮影位置における類似フレーム対 (m_p, n_p) を，図 27,28 に撮影地点 0 において算出された類似フレーム対と対応する動物体領域画像を示す．提示映像において運動の連続性を考慮した繰り返し再生を実現できたが，単一の周期運動物体を対象とした場合に比べやや違和感が残った．これは，運動物体が複数になったことにより，運動の連続性が類似するフレームが出現する確率が減少したことが原因であると考えられる．図 29,30 に，図 24 中に示した動物体画像が切り替わる地点 (c)(d) の前後のフレームにおいて生成されたビデオテクスチャを示す．提示映像の地点 (c)(d) の前後においては、物体の運動と色調がやや不連続に切り替わることが確認された．これは，物体が撮影地点固有の動きをしたために，隣接する撮影地点における動画像中で類似するフレームが探索されなかったことや，撮影時に環境の照明条件が変化したことが原因と考えられる．また，単一の周期運動物体を対象とした場合と同様に，近似した平面から離れた領域では平面射影変換により誤差が生じるため，図 31 の拡大図に示すように，提示映像の建物部分（実線の枠）に違和感が生じた．さらに，運動が微小であった領域（破線の枠）にアーティファクトが発生した．これは，背景画像に運動物体が残存し，その結果，入力フレームの背景が動物体画像に書き込まれることが原因であると考えられる．提案手法では隣接地点の背景画像のブレンディングにより背景画像を生成するため，ブレンディングされていない動物体画像中の背景領域がアーティファクトとなる．



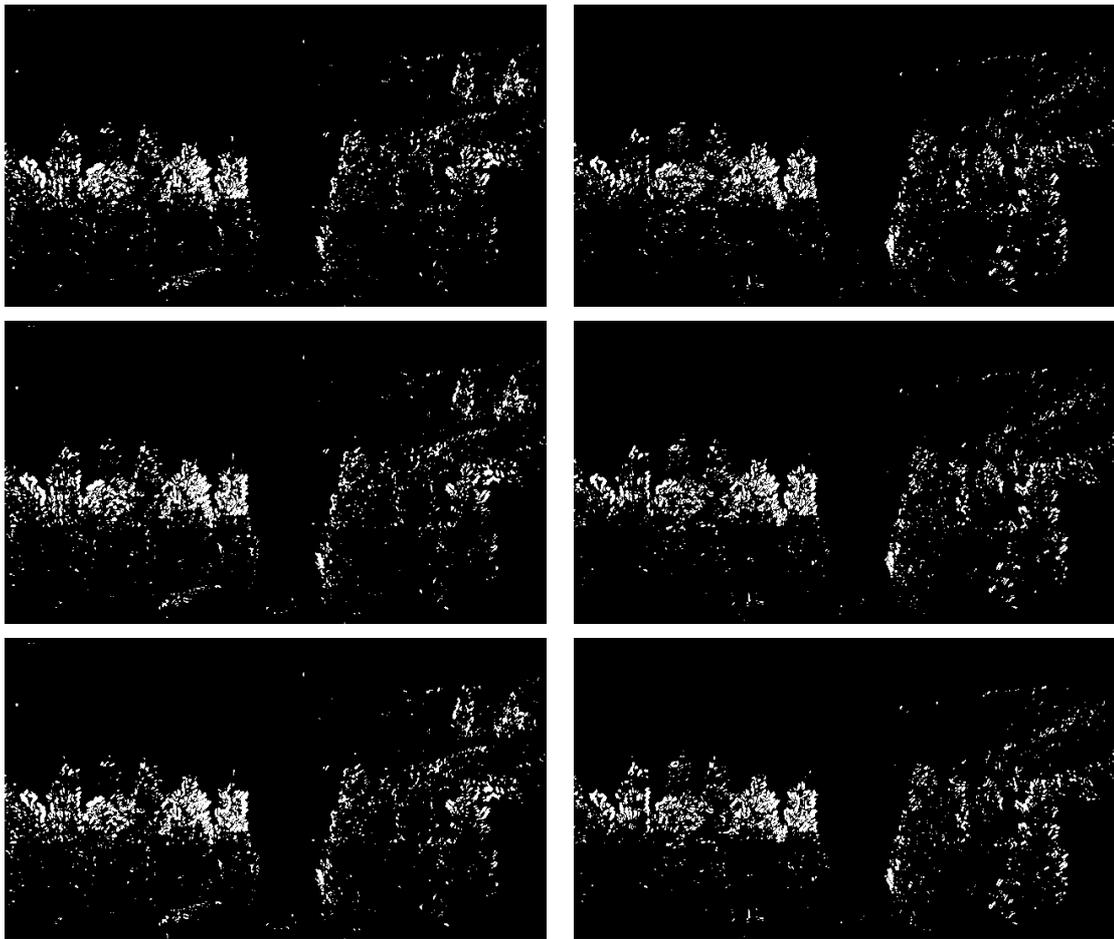
上から順に54,55,56フレーム目

上から順に231,232,233フレーム目

図 27 撮影地点 0 において算出された類似フレーム (複数の周期運動物体)

表 4 算出された各撮影位置における類似フレーム (m_p, n_p)

	(m_p, n_p)
撮影位置 0	(54,231)
撮影位置 1	(184,291)
撮影位置 2	(261,361)



上から順に54,55,56フレーム目

上から順に231,232,233フレーム目

図 28 撮影地点 0 において算出された類似フレームに対応する動物体領域画像
(複数の周期運動物体)

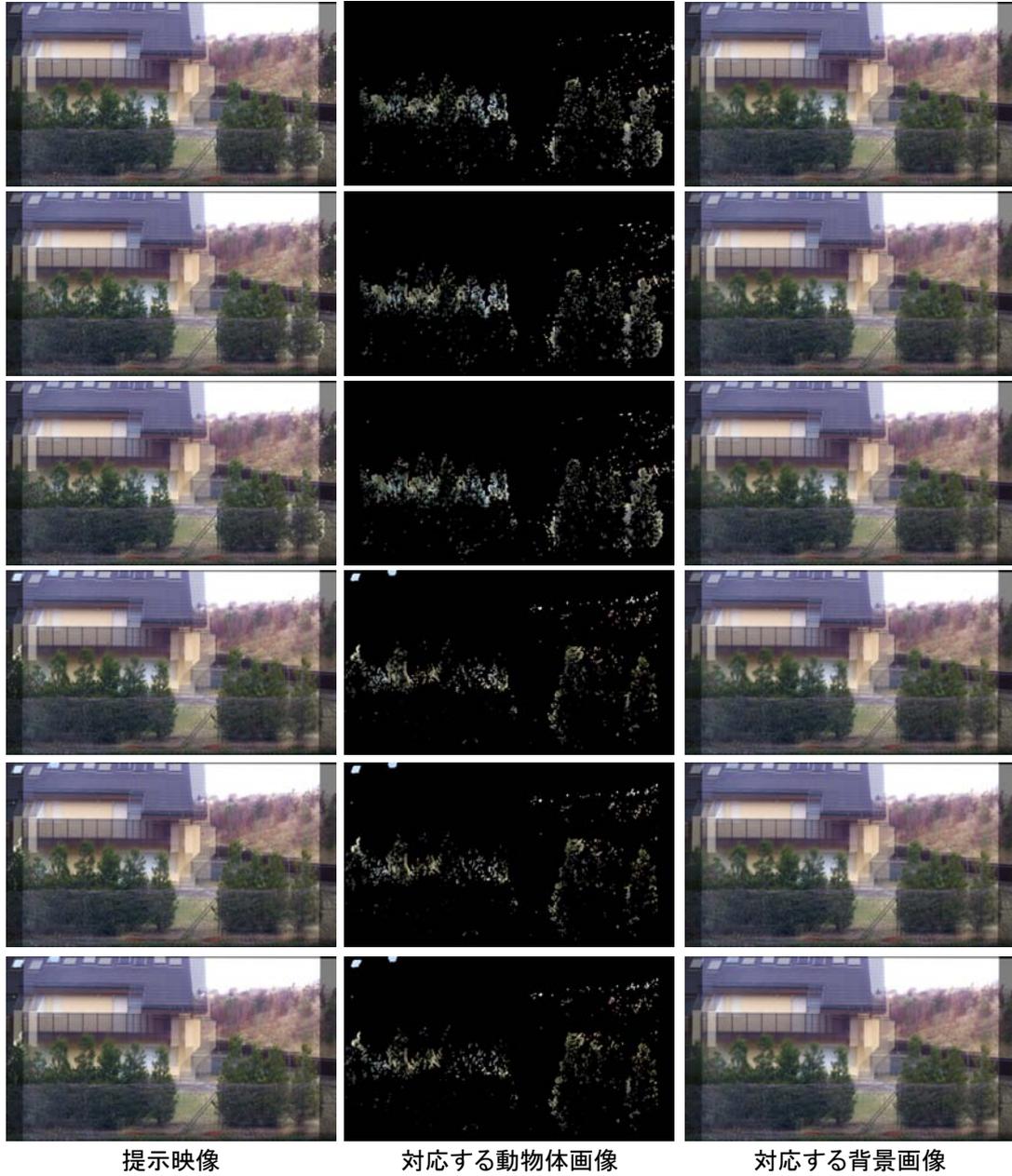


図 29 地点 (c) の前後のフレームの提示映像と対応する動物体画像と背景画像
 (上から順に 176,177,178,179,180,181 フレーム目)

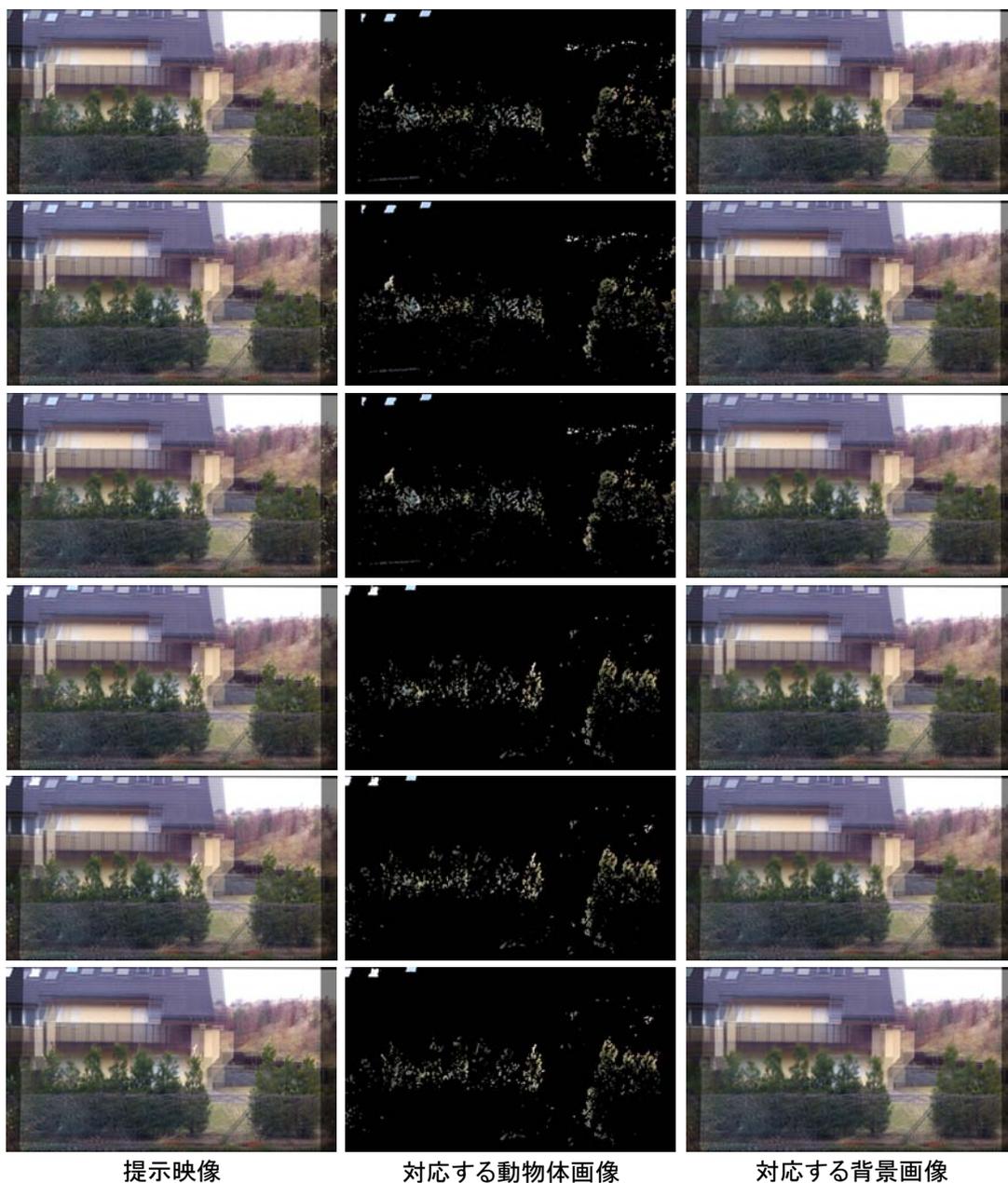


図 30 地点 (d) の前後のフレームの提示映像と対応する動物体画像と背景画像
 (上から順に 295,296,297,298,299,300 フレーム目)

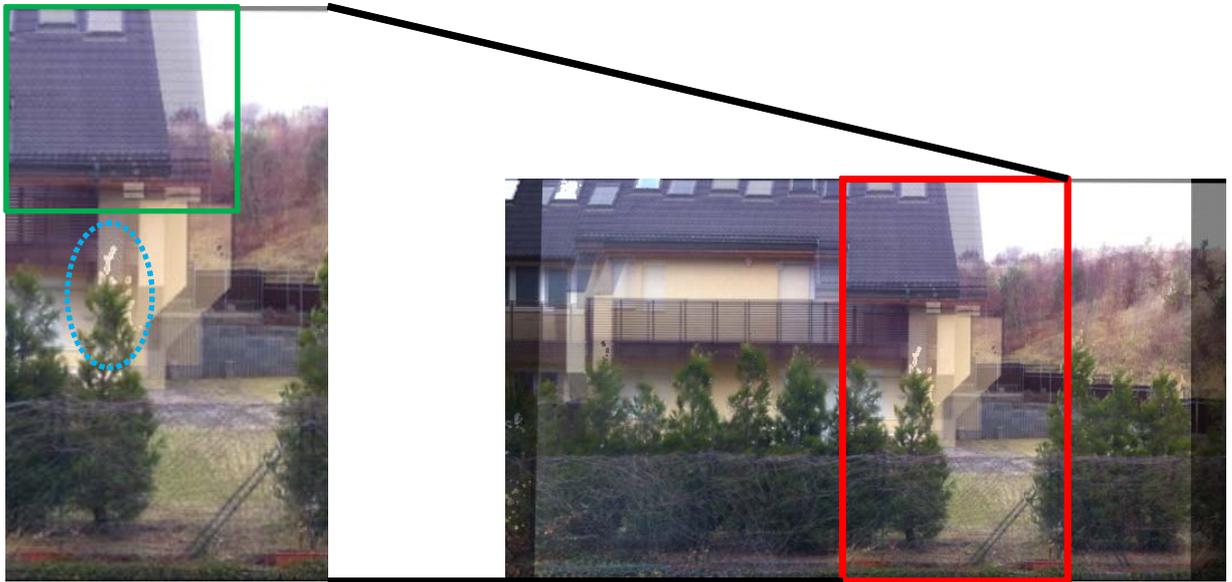


図 31 地点 (d) の 1 フレーム後の提示映像と拡大画像

5. まとめ

本論文では、動画像を違和感なく任意時間の動画像に合成するビデオテクスチャおよび画像群の変形により中間視点画像を生成する手法を基礎として、離散的な位置で異なる時間に撮影した複数の動画像を入力とし、(1) 周期運動物体を対象とした映像の繰り返し再生、(2) 撮影地点間での類似フレームの接続による中間視点画像生成、を行うことで、視点移動可能なビデオテクスチャを生成する手法を提案した。

実験では、離散的な位置で異なる時間に撮影した複数の動画像を用いて、単一運動物体および複数運動物体を対象とし、視点移動可能なビデオテクスチャを生成した。実験の結果、時間的・空間的連続性を保持した視点移動が可能なビデオテクスチャを生成できることを示した。また、16名の被験者による評価実験の結果、提案手法が単一の周期運動物体が存在すると見なせるシーンを対象とした繰り返し再生と視点移動において良好な結果が得られることが確認できた。このことから、本手法は、動的環境を対象とした仮想化現実空間の構築へ適用できると考えられる。ただし、提示映像において、対象物体に対し近似した平面から離れた領域や運動が微小であった領域にアーティファクトが発生し違和感が生じているため課題が残る。

今後、近似した平面から離れた領域におけるアーティファクトを解決するためには、画像上のすべての物体を一つの平面で近似するのではなく、物体ごとに異なる平面を近似し、画像群を変形することが考えられる。また、モーフィングを用い、形状の特徴を考慮した中間視点における背景画像を生成することも考えられる。

運動が微小であった領域におけるアーティファクトを解決するためには、背景の推定精度向上が必要であり、背景推定が困難な領域については、ポアソン画像合成 [30] を用いて領域の色調をシームレスに補正することが考えられる。

謝辞

本研究を進めるにあたり，その全過程において細やかな御指導，御鞭撻を頂いた視覚情報メディア講座 横矢 直和 教授に心より感謝致します．そして，本研究の遂行にあたり，有益な御助言，御鞭撻を頂いた知能情報処理学講座 木戸出 正 繼 教授に厚く御礼申し上げます．さらに，本研究の全過程を通して，始終暖かい御指導をして頂いた視覚情報メディア講座 山澤 一誠 准教授に深く感謝申し上げます．また，公私にわたり大変お世話になり，研究に関して的確な御助言を頂いた環境知能学講座 神原 誠之 准教授に熱く御礼申し上げます．研究活動の全過程を通して多大なる御助言，御指導を頂いた視覚情報メディア講座 佐藤 智和 助教に心より感謝致します．特に，佐藤 智和 助教には本研究のテーマの設定から本論文の執筆，その他発表論文の添削，発表練習に至るまで細やかな御指導，御鞭撻を頂きましたことを，重ねて御礼申し上げます．そして，研究の方向性について無数の御助言，御指導を賜った視覚情報メディア講座 糸 秀行氏に心より深く感謝致します．また，研究室での生活を支えて頂いた視覚情報メディア講座 事務補佐員 中村 美奈 女史に心より感謝致します．さらに，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った視覚情報メディア講座の皆さまに深く感謝致します．最後に，両親をはじめ，私の二年間の大学院生活に関わった全ての方々に感謝の意を表します．

参考文献

- [1] S. Chen: “Quicktime VR: an image-based approach to virtual environment navigation,” Proc. SIGGRAPH’95, pp. 29–38, 1995.
- [2] T. Pintaric, U. Neumann and A. Rizzo: “Immersive panoramic video,” Proc. Int. Conf. on Multimedia, pp. 493–494, 2000.
- [3] A. Rav-Acha, Y. Pritch, D. Lischinski and S. Peleg: “Dynamosaicing: Mosaicing of dynamic scenes,” IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 29, No. 10, pp. 1789–1801, 2007.
- [4] C. Hermans, C. Vanaken, T. Mertens, F. Van Reeth and P. Bekaert: “Augmented panoramic video,” Computer Graphics Forum, Vol. 27, No. 2, pp. 281–290, 2008.
- [5] Y. Onoe, K. Yamazawa, H. Takemura and N. Yokoya: “Telepresence by real-time view-dependent image generation from omnidirectional video streams,” Computer Vision and Image Understanding, Vol. 71, No. 2, pp. 154–165, 1998.
- [6] S. Ikeda, T. Sato, M. Kanbara and N. Yokoya: “Immersive telepresence system using high-resolution omnidirectional video with locomotion interface,” Proc. Int. Conf. on Artificial Reality and Telexistence, pp. 188–193, 2004.
- [7] M. Uyttendaele, A. Criminisi, S. Kang, S. Winder, R. Szeliski and R. Hartley: “Image-based interactive exploration of real-world environments,” IEEE Computer Graphics and Applications, Vol. 24, No. 3, pp. 52–63, 2004.
- [8] C. Taylor: “Videoplus: A method for capturing the structure and appearance of immersive environments,” IEEE Trans. Visualization and Computer Graphics, Vol. 8, No. 2, pp. 171–182, 2002.

- [9] 和田, 池田, 佐藤, 横矢: “複数経路上を歩行移動可能なテレプレゼンスシステムにおける違和感の低減に関する検討”, 日本バーチャルリアリティ学会 サイバースペースと仮想都市研究会研究報告, Vol. 11, No. 1, pp. 29–34, 2006.
- [10] S. Chen and L. Williams: “View interpolation for image synthesis,” Proc. SIGGRAPH’93, pp. 279–288, 1993.
- [11] S. Seitz and C. Dyer: “View morphing,” Proc. SIGGRAPH’96, pp. 21–30, 1996.
- [12] 遠藤, 片山, 田村, 廣瀬: “写実的な広域仮想空間構築のための画像補間手法”, 日本バーチャルリアリティ学会論文誌, Vol. 7, No. 2, pp. 185–192, 2002.
- [13] K. Kato, H. Ishiguro and M. Barth: “Town digitizing: Recording of street views by using ODVS & GPS,” Proc. IEEE Intelligent Vehicles Symp., pp. 183–188, 2001.
- [14] 亀井, 瀬尾: “Pace-the-scene movie : ビデオ映像による実環境の情景表現手法”, 電子情報通信学会論文誌 (D-II), Vol. J85-D-II, No. 2, pp. 604–612, 2002.
- [15] T. Stich, C. Linz, G. Albuquerque and M. Magnor: “View and time interpolation in image space,” Computer Graphics Forum, Vol. 27, No. 7, pp. 1781–1787, 2008.
- [16] E. Adelson and J. Bergen: “The plenoptic function and the elements of early vision,” Computational Models of Visual Processing, pp. 3–20, 1991.
- [17] T. Naemura, T. Takano, M. Kaneko and H. Harashima: “Ray-based creation of photo-realistic virtual world,” Proc. Int. Conf. on Virtual Systems and MultiMedia, pp. 59–68, 1997.
- [18] M. Levoy and P. Hanrahan: “Light field rendering,” Proc. SIGGRAPH’96, pp. 31–42, 1996.

- [19] S. Gortler, R. Grzeszczuk, R. Szeliski and M. Cohen: “The lumigraph,” Proc. SIGGRAPH’96, pp. 43–54, 1996.
- [20] 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”, 電子情報通信学会論文誌, Vol. 84, No. 9, pp. 2048–2057, 2001.
- [21] T. Takahashi, H. Kawasaki, K. Ikeuchi and M. Sakauchi: “Arbitrary view position and direction rendering for large-scale scenes,” Proc. Computer Vision and Pattern Recognition, Vol. 2, pp. 296–303, 2000.
- [22] A. Schödl, R. Szeliski, D. Salesin and I. Essa: “Video textures,” Proc. SIGGRAPH’00, pp. 489–498, 2000.
- [23] A. Schödl and I. Essa: “Controlled animation of video sprites,” Proc. ACM Symp. on Computer Animation, pp. 121–127, 2002.
- [24] M. Flagg, A. Nakazawa, Q. Zhang, S. Kang, Y. Ryu, I. Essa and J. Rehg: “Human video textures,” Proc. ACM Symp. on Interactive 3D Graphics and Games, pp. 199–206, 2009.
- [25] V. Kwatra, A. Schödl, I. Essa, G. Turk and A. Bobick: “Graphcut textures: Image and video synthesis using graph cuts,” ACM Trans. Graphics, Vol. 22, No. 3, pp. 277–286, 2003.
- [26] A. Agarwala, K. Zheng, C. Pal, M. Agrawala, M. Cohen, B. Curless, D. Salesin and R. Szeliski: “Panoramic video textures,” ACM Trans. Graphics, Vol. 24, No. 3, pp. 821–827, 2005.
- [27] W. Matusik, C. Buehler, R. Raskar, S. Gortler and L. McMillan: “Image-based visual hulls,” Proc. SIGGRAPH’00, pp. 369–374, 2000.
- [28] 石川, 山澤, 横矢: “複数の全方位動画像を用いた自由視点テレプレゼンス”, 映像情報メディア学会誌, Vol. 60, No. 10, pp. 1599–1610, 2006.

- [29] C. Zitnick, S. Kang, M. Uyttendaele, S. Winder and R. Szeliski: “High-quality video view interpolation using a layered representation,” *ACM Trans. Graphics*, Vol. 23, No. 3, pp. 600–608, 2004.
- [30] P. Pérez, M. Gangnet and A. Blake: “Poisson image editing,” *ACM Trans. Graphics*, Vol. 22, No. 3, pp. 313–318, 2003.