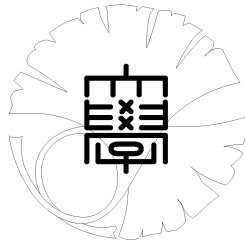


博 士 論 文

全方位画像を用いた実世界空間のモデリング



指導教官

坂内 正夫 教授

東京大学大学院工学系研究科 電子情報工学専攻

07121 川崎 洋

平成 14 年 12 月 16 日

目次

1	序論	1
1.1	本論文の背景	2
1.2	本論文の目的	4
1.3	本論文の構成	5
2	鏡面型全方位カメラによる広域空間の効率的獲得	9
2.1	はじめに	10
2.2	鏡面反射型全方位画像	11
2.2.1	双曲回転体型	11
2.2.2	放物回転体型	12
2.2.3	鏡面反射型全方位画像の特徴	14
2.3	鏡面反射型全方位画像の解像度の向上	15
2.3.1	撮影経路	15
2.3.2	関連研究	15
2.3.3	映像の超解像度化	18
2.3.4	障害物の除去	22
2.3.5	超解像度画像の画質について	24
2.4	実験	25
2.4.1	撮影経路	25
2.4.2	対象映像	26
2.4.3	結果	26
2.4.4	問題点	34
2.5	まとめ	34
3	複数台カメラによる広域空間の効率的獲得	42
3.1	はじめに	43
3.2	複数台カメラを用いた全方位画像	44
3.2.1	複数台カメラによる全方位画像	44
3.2.2	関連研究	45
3.3	複数カメラによる歪みの無い画像の生成	45
3.3.1	EPIを用いた光学中心の一致手法	48
3.4	実験	57
3.4.1	撮影経路	57

3.4.2	対象映像	57
3.4.3	結果	57
3.5	まとめ	58
4	画像の再構成による都市空間の生成	64
4.1	はじめに	65
4.2	関連研究	66
4.3	画像のサンプリング	67
4.3.1	実験車両	67
4.3.2	撮影経路	67
4.4	任意視点の生成	68
4.4.1	任意視点の生成アルゴリズム	68
4.4.2	視点合成が可能な領域	68
4.4.3	画像生成例	71
4.5	生成画像における問題点	71
4.5.1	横方向のひずみ	72
4.5.2	縦方向のひずみ	72
4.6	都市空間の特徴を利用した歪みの除去	73
4.6.1	カメラ・建物間の距離の推定	75
4.6.2	ひずみの補正方法	75
4.7	実験	76
4.7.1	実験方法	76
4.7.2	結果	77
4.8	まとめ	77
5	幾何形状を利用した都市空間の生成	82
5.1	はじめに	83
5.2	全方位画像からの幾何情報取得	84
5.2.1	時空間画像解析	84
5.2.2	EPI パターンの生成	88
5.2.3	2次元 DP マッチング	88
5.2.4	マッチング結果	91
5.2.5	3次元形状推定結果	93
5.3	任意視点の生成方法	94
5.3.1	表面光線空間データの生成	94
5.3.2	データの圧縮	96
5.3.3	レンダリング手法	97
5.4	実験	97
5.4.1	実験方法	97
5.4.2	画像合成結果	99
5.5	まとめ	106

6 結 論	107
6.1 本論文の主たる成果	108
6.2 今後の課題と展望	109
謝 辞	113

目 次

1.1	本論文の構成	6
2.1	全方位画像の例 (双曲面型)	10
2.2	双曲回転型全方位カメラの光学モデル	12
2.3	放物回転型全方位カメラの光学モデル	13
2.4	全方位カメラ時空間画像 (a) 放物回転体 (b) 双曲回転体	21
2.5	全方位カメラ時空間画像を縦に切断した例	22
2.6	全方位カメラ画像 EPI(a) 放物回転体 (b) 双曲回転体	23
2.7	超解像度化の過程	24
2.8	入力全方位画像例	26
2.9	オリジナル画像	27
2.10	バイリニア補間による画像	27
2.11	超解像度画像 (入力 20 枚, 繰り返し 15 回)	28
2.12	超解像度画像の PSNR	29
2.13	全方位カメラの入力画像例 (法物面型)	30
2.14	EPI 法による自動画像抽出例	31
2.15	ブロックマッチングによる位置合わせ	32
2.16	超解像度例:左 入力画像, 右 超解像度画像	33
2.17	超解像度画像例 2:上から順番に, 入力画像, 超解像度画像, アダプティブウィンドウを用いた超解像度画像	35
2.18	ノイズ除去例:上段, ノイズののった全方位画像, 左下, 透視変換画像, 右下, ノイズ除去を行った超解像度画像	36
2.19	EPI 上での特徴点の動き	41
2.20	奥行きの等しい等間隔の面が並ぶ経路上での EPI	41
3.1	円筒状に配置された複数カメラ	44
3.2	(a) 光学中心が一致していると仮定して透視変換画像を円筒座標系に投影 (b) ピクセル誤差	46
3.3	画像収集車	47
3.4	光学中心の一致していない複数カメラの統合例	47
3.5	光学中心を一致させるためのカメラ配置	49
3.6	実際のカメラ配置例 (1)	50
3.7	実際のカメラ配置例 (2)	51

3.8	EPIによる光学中心の一致	52
3.9	EPIのマッチング	53
3.10	画像の平行化	54
3.11	平行化の例(1)	55
3.12	平行化の例(2)	56
3.13	(a-f):撮影画像(レンズ歪み除去後)(f)統合画像	59
3.14	統合画像例(2)	60
3.15	車載し撮影した画像	60
3.16	時間軸方向を考慮しない位置あわせ	61
3.17	時間軸方向を考慮した位置あわせ	61
3.18	統合例1	62
3.19	統合例2	62
4.1	実世界での画像撮影用の車	68
4.2	画像の再構築方法	69
4.3	信濃町付近で撮影した画像を用いて作成した全方位画像	69
4.4	2次元上の合成可能範囲	70
4.5	(a):singular directionを含む場合の画像生成(b):singular directionを含まない場合の画像生成	71
4.6	仮想視点からの眺め	72
4.7	IBRによる歪み	73
4.8	縦方向のひずみ	74
4.9	3次元情報による補正方法	76
4.10	信濃町付近で撮影した画像を用いて作成した全方位画像	77
4.11	3次元情報による補正後(実画像).上図は撮影点よりも前にカメラを設置(図12と同じ位置).下図は撮影点とより後ろに仮想カメラを設置.	78
5.1	(a)全方位カメラ時空間画像(b)半径方向切断面	86
5.2	垂直方向切断面	87
5.3	全方位画像EPIからエッジ検出した結果	89
5.4	地図から作成したモデルEPI	90
5.5	モデルEPIから生成したパターン	91
5.6	2次元DPマッチングの模式図	92
5.7	2次元DPマッチング中の1次元サーチプレーン	92
5.8	2次元DPマッチングの結果	93
5.9	繰返し処理によるDPマッチングの結果の変化	94
5.10	奥行き情報を利用した3次元都市	95
5.11	獲得できた側面の例	95
5.12	奥行き情報を利用した3次元都市(2)	95
5.13	奥行き情報を利用した3次元都市(2)(鳥瞰図)	95
5.14	全方位カメラ時空間画像の半径方向切断面(PVI)	96

5.15 PVI 画像 (プレーン 5 での切断面)	97
5.16 PVI 画像 (プレーン 19 での切断面)	97
5.17 PVI 画像 (プレーン 39 での切断面)	97
5.18 PVI と EPI と表面空間光線の関係	98
5.19 表面光線空間データ	99
5.20 圧縮率	100
5.21 圧縮率	101
5.22 表面光線空間によるレンダリング	102
5.23 (a)(b)(c): 本手法による合成例, (d): 仮想カメラ位置, (e): 道路と建物の距離の 1 パラ メータのみを利用した合成画像例	103
5.24 オクルージョンの正確なレンダリング. 上から順にカメラが建物に向かって接近 . .	104
5.25 仮想カメラ位置	104
5.26 (a)(b)(c): 本手法により合成した画像 (d): カメラ位置 (e): 複数カメラからの入力画 像 (f): カメラを 3 台だけ使用した場合	105

表 目 次

2.1 超解像度化手法のまとめ	18
---------------------------	----

第 1 章

序論

近年，映像技術や，コンピュータグラフィックス (CG) の発展に伴ない，現実感の高い仮想空間を生成することが可能となってきた。しかしながら，これまで主流であった形状や表面特性モデルを用いた手法では，データ取得の困難性から大規模で複雑な環境を構築することに限界のあることが認識されつつある。

一方，大量の画像を蓄積し，これらの実画像を様々に加工・処理し，再構築することで撮影点以外での視点映像を作りだすイメージベースと呼ばれる手法が注目されている。この手法を用いれば形状や表面特性の取得が困難であるような場合でも写実性の高い CG を生成することが出来る。

これまで，イメージベース法の対象としては，比較的小さい対象が主であった。また，これらの研究の主眼は記述法であり，入力方法についてはあまり考察されていない。そこで本論文では，都市空間のように対象が非常に大きくかつ広域に分布する場合に，自由な視点からリアリティを持って景観を提示するイメージベースによる手法を提案する。また，その際に効率的な情報収集を実現するために，瞬時に全方向の画像が取り込める全方位カメラを利用したデータの取得方法を提案する。

さらに，従来イメージベース法で難しいとされてきた，歪みの除去や仮想物体の挿入などを，イメージベース法に幾何形状を付加することで解決可能であることを示す。

本章ではまず，本論文の位置付けを明確にするために，これらのキーワードを軸に本論文の背景を説明する。続いて，本論文の目標がそれらを統合することで実現されることを示す。最後に本論文の構成を説明する。

1.1 本論文の背景

実世界空間をモデリングし、計算機内に再現することは、コンピュータグラフィック (CG) やバーチャルリアリティ (VR) などの分野で重要なテーマである。特に、都市空間や遺跡などのような広域な空間を電子化し提示することが出来れば、景観シミュレーションやインターネット上での仮想ツアーなど、多くの応用が期待できる。最近では、高度交通システム (ITS) の分野においても、カーナビゲーションやショッピングガイドさらには都市交通計画の基礎技術として高い関心を集めている。

実世界空間のモデリング

これまで実世界空間を対象とし、モデリング・提示する試みが数多く行われてきた。特に、都市空間を対象とした試みは古くから行なわれている。Teller [1, 2] らは、全方位カメラで街全体をカバーするように撮影した画像群を用いて建物テクスチャを正確に切り出し、ポリゴンベースでの街の再現を試みている。横矢ら [3]、石黒ら [4]、広瀬ら [5] は、全方位カメラを車載し撮影した映像を用いて、撮影した経路上を自由に動くことが可能なシステムを提案している。横矢ら [3] は、光線空間法に基づいた考え方を用いることで、単眼全方位画像から立体視が可能なシステムへと拡張した。また、石黒らは自動走行車両により効率的に画像データを獲得するシステムを提案している。広瀬ら [5] は車載型撮影システムにより、実際に都市空間を撮影した画像を経路に沿ってインタラクティブに表示する VR システムを製作した。また、レンジセンサやカメラを搭載した車により、自動的に街をモデリングする計画も進められている [6-8]。

実世界空間のモデリングは、大きくデータの取得段階と提示段階の 2 つに分けて考えることが出来る。データ取得段階では、何らかの方法で実際に存在する建築物や都市の景観等を計算機で処理可能なデータ形式に変換し、保存する。通常では、このデータ形式は 3 次元形状や対象物の色などの表面特性である。データ提示段階では、これらの形状や光学情報を用いて、任意の光源方向における、任意の視線方向から見た都市空間画像を描画する。

形状と表面特性に基づいた実世界空間モデリング

このように、実世界空間のモデリングとは、実世界における対象物体の形状やテクスチャなどのデータを取得し、これを CG 等を用いて提示するシステムを一般的に指す。モデリングする空間が比較的狭い場合や対象物が複雑な構造を持たない場合、形状やテクスチャの取得は CAD や写真測量のツールを利用して主に人手によって行われる。しかし、実際の都市などの広域空間からデータを取得する場合、その対象があまりにも膨大であるため、人手でするには多くの困難がある。そこで、データ取得段階において何らかの手法で、現在の都市空間の形状や表面特性などを自動的に得られれば大きな省力化が図れる。

現実存在する都市などの広域な空間からモデルを得る場合、幾何形状データを得る部分と光学情報を得る部分に分割して考えることができる。そこで、それぞれの部分の自動化の試みについて述べる。

幾何形状の獲得　これまで幾何形状情報の取得法として航空写真によるもの [9-14]、動画像処理

によるもの [15, 16], レンジセンサによるもの [6, 8] などが提案されている。

航空写真による手法では, 画像処理によりエッジ抽出した結果を都市のモデルにあてはめたり [9, 10, 12], 2 枚以上の画像を用いたステレオ法により 3 次元の形状を求める手法 [11, 17] の他, それらを組み合わせる手法 [13, 14] など数多くの手法が提案されている。また, それら提案手法は実際に地図製作などで利用されるなど応用範囲も広く, 実装も進んでいる。しかし, 航空写真を利用しているため側面の情報は取得することが出来ない他, 広範囲を上空から撮影しているため細かい形状の取得が難しいなどの欠点がある。

動画像を用いた手法 [15, 16] では, 広範囲な都市などの空間を解像度の高いカメラやレンジセンサで計測し, それらデータを高い精度で統合することが, センサによる獲得可能な自己位置情報の精度限界や, データ間の同期の難しさ等のため, 現実的なレベルにはまだ達していない。最近では精度の良い 3 次元計測装置や GPS などのセンサも商品化されているが, 大規模な環境のモデリングにそれらの機器を適用するのは已然として難しい。

光学情報の獲得 光学情報取得に関しては, 大半は正面からのテクスチャを手で貼り付けるものが多い [18]。また, GPS などのセンサを用いてテクスチャ取得を自動化する試みも行われている [19]。最近, グラフィクスの分野では, 複数の入力画像列から対象となる物体上の点の色変化を解析し表面反射パラメータを求める手法が提案されている [20]。しかしこれらの手法は主に実験室内など理想的な環境で行なわれていることが多く, 野外などで計測したデータに適用することは難しい。

このように, これまで多くのモデリング手法が提案されているが, これら手法を実際の都市空間に利用しようとすると, 画像の解像度や 3 次元計測機器の精度限界により, 細かい形状や側面の情報を獲得することが困難である。また, 表面特性に関しても, 屋外であり光源環境が複雑なため, 正しい値を計測することが難しい。

見えに基づいた実世界空間モデリング

このようなモデル生成に伴う困難を回避する方法として, イメージベース法による仮想空間生成が考えられる [4, 21, 22]。イメージベース法は, 3 次元形状や表面特性などの情報を使用せずに, 予め蓄えられた画像のみを用いてレンダリング (提示) を実現する。このため, 対象物体の形状や反射特性を考慮する必要がなく, 複雑な形状や反射率を有する物体に対しても適用できる。従って, 簡便に実際に近い見えが生成できる。

見えに基づく表現方法

イメージベース法の中心概念として光線空間 [23] がある。物体をある方向から見た場合の見えというのは, その物体が存在する空間内において物体表面から視点位置に到達する光線の束により構成される。その分布を表現する空間を光線空間 [23] と呼ぶ。入力画像を元に, 物体上の各々の光線の明るさや色を光線の方向を示すパラメータに基づき光線空間内に記録する。新たな視点からの画像を表示 (レンダリング) する際には, 各画素の方向をパラメ

タ化し，そのパラメタに基づいて光線空間を参照し，そこで明るさや色を取り出し，これらを組み合わせることにより，あらたな視点における物体の画像を生成する．

光線空間を構成 (モデリング) するためには，対象物体をあらゆる方向から撮影した画像が必要である．光線空間内で光線の分布を効率良く記述するため，光線空間法 [24]，Light Field Rendering [25]，Lumigraph [26]，Concentric Mosaics [27] などが提案されている．

画像情報の獲得

上に述べた見えに基づく研究の主眼は記述法であり，入力方法についてはあまり考察されていない．すなわち，多数のカメラから得られる多数の画像列があるとして議論を進めている．これは対象とする物体が比較的小規模であったため，光線の取得法自体はそれほど検討する必要がなかったからである．一方，我々が目的とする都市空間は，対象が非常に大きくかつ広域に分布する．このため効率的な光線情報収集の検討を欠かすことができない．

このように，見えに基づく手法では，広い空間を現実感の高くユーザーに提示できる可能性がある一方で，実際にはそれを実現するために必要な莫大な光線空間データの獲得方法や，また，そのような巨大な光線空間をいかに効率的に処理し画像合成をするか，という課題が残されている．

1.2 本論文の目的

本論文では，このような観点から，広範囲な実世界空間を短時間で取得し，効率的に仮想空間を構築することを目的とする．これらを実現するためには以下のような課題が挙げられる．

- 広域な空間のデータの効率的な取得
- 取得した大規模なデータの整合性を得ること
- 取得したデータを用いて計算機内に広域な都市空間を効率的に再現すること
- 再現した都市空間内をユーザーが動き回ったり，新たな物体を設置するなどのインタラクティブな提示が可能であること

特に本論文では，景観シミュレーションやカーナビゲーションなどの幅広い分野で応用が見込まれる都市空間を広域空間の対象と設定し，上記の課題を解決することを目指す．

本論文では，上記の課題を解決するために全方位画像を利用する手法を提案する．全方位画像とは，視点を中心として 360 度全方向を撮影した画像のことである．この全方位画像を利用すれば，広範囲な実世界空間を短時間で取得することが出来るため，効率的に仮想空間を構築することが出来る．しかしながら，全方位画像には通常のカメラで撮影した画像にはない特徴や性質があり，単に計測しただけでは都市のモデリングに直接利用することは出来ない．そこで，本論文では，まず，全方位画像の取得方法や特徴について述べ，全方位カメラを利用して効率的に広域環境のデータを

獲得する手法を提案する。その際に、2種類の全方位カメラを用いることとし、それぞれのカメラの特徴や性質についての分析を行ない、全方位カメラに固有な特質を明らかにすることを旨とする。さらにカメラ画像を取り込む軌跡を直線とすることで情報収集の手間が非常に削減できることも示す。

また、撮影した全方位画像列から現実感の高い都市空間を再現することを試みる。これを実現するために、本論文では、撮影した全方位画像列から画像のみを用いることで簡単に現実世界を仮想空間中に再現することが可能な、イメージベース法を主に利用することとし、これを都市空間に適用するための手法を提案する。

イメージベース法は、物体の幾何形状や表面反射特性などの厳密なデータが無くても高いリアリティをもって画像を再現することが出来る点が注目されているが、固有な歪みやデータサイズが巨大であることなどの欠点も知られている。そこで、本論文では、イメージベース法に幾何情報を付加することで、単純なイメージベース法では発生する歪みを解消することができ、データサイズの削減が可能であることを示す。

また、本論文では、使用する幾何情報を、撮影した映像を処理することで獲得することを旨とする。提案する手法では、幾何形状を獲得するための特殊な機器が必要ないため、計測が大幅に簡略化され、実際の応用に際して大きな効果が期待される。

1.3 本論文の構成

本論文の構成は次の通りである。

- 第 1 章 序論
- 第 2 章 鏡面型全方位カメラによる広域空間の効率的獲得
- 第 3 章 複数台カメラによる広域空間の効率的獲得
- 第 4 章 画像の再構成による都市空間の生成
- 第 5 章 幾何形状を利用した都市空間の生成
- 第 6 章 結論

図 1.1 は各章の関係を表したものである。本論文の流れを簡単に説明すれば以下ようになる。

第 2 章では、都市空間をという広大な空間の画像データを、鏡面反射型の全方位カメラを用いることで効率よく取得する方法について述べる。全方位カメラとは、画像撮影地点において、水平方向の 360 度すべての景色を撮影可能なカメラのことである。鏡面反射型全方位カメラとは、回転体型の鏡面を用いることでこれを実現した全方位カメラである。また、鏡面反射型の全方位カメラには、画像解像度が低いという問題点があるが、これを解決する手法も併せて提案する。

第 3 章では、複数台のカメラを用いて、全方位画像を撮影する方法について述べる。複数台のカメラを用いてそれぞれ撮影した画像を統合すれば、解像度の高い全方位画像を作ることが出来る。一方、複数の画像をつなぎ合わせる際に画像の歪みが大きくなるという問題点がある。そこで、時空間画像解析を用いることで、このような複数台のカメラを統合する際の歪みを除去する手法を

提案する。

第 4 章では、都市空間という広域環境を、より自由に広範囲にわたり、リアリティを持って、イメージベース法を用いてレンダリングする手法を提案する。近年、大量の画像を蓄積し、これらの実画像を様々に加工・処理し、再構築することで撮影点以外での視点映像を現実感高く作りだすイメージベースと呼ばれる手法が注目されている。これまで、イメージベースの研究対象としては、比較的小さい対象が主であった。また、広域を対象とした研究であっても広範囲に自由視点の画像再構築を実現した例はない。そこで、本論文では、これを広域空間に拡張する手法を提案する。

第 5 章では、イメージベース法に幾何形状を付加し、都市を再現する手法を提案する。イメージベース法による都市空間の生成は、簡単な実装によりモデリングの困難さ無く、現実感の高い画像合成が出来るなどの利点があるものの、都市空間の形状を「道路からの距離」という一つのパラメータで表現しているため、歪みが正しく除去できない場合も多い。そこで、イメージベース手法に幾何形状を付加することで、これが歪みが軽減できることを示す。また、提案する手法により、データの圧縮やシンボリックな情報の付加等も実現できることを示す。

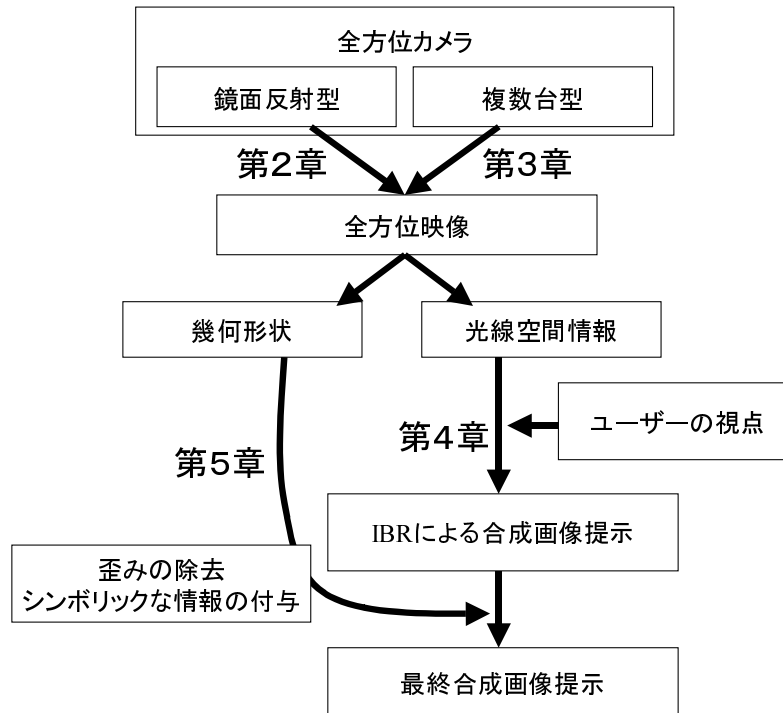


図 1.1 本論文の構成

【 参 考 文 献 】

- [1] S. Teller, M. Antone, M. Bosse, S. Coorg, M. Jethwa, and N. Master: “Calibrated, registered images of an extended urban area”, in *Computer Vision and Pattern Recognition* (2001).
- [2] Matthew Antone and Seth Teller: “Scalable, absolute position recovery for omni-directional image networks”, in *Computer Vision and Pattern Recognition* (2001).
- [3] 山口晃一郎, 山澤一誠, 竹村治雄, 横矢 直和: “全方位動画像を用いた任意方向両眼ステレオ画像の実時間生成と提示”, 電子情報通信学会技術報告, PRMU99-159 (Nov. 1999).
- [4] 加藤浩仁, 石黒浩, マシューバース: “Town digitizing -全方位カメラと gps による実時間モデリング”, 情報処理学会研究報告, CVIM-125, pp. 111-118 (Jan. 2001).
- [5] 広瀬通孝, 遠藤隆明: “Building a virtual world from the real world”, in *International Symposium on Mixed Reality*, pp. 183-197 (Mar. 1999).
- [6] Huijing Zhao and Ryosuke Shibasaki: “Reconstruction textured urban 3d model by fusing ground-based laser range image and ccd image”, in *MVA*, pp. 232-237 (Nov. 1998).
- [7] H. Zhao and R Shibasaki: “Reconstructing urban 3d model using vehicle-borne laser range scanners”, in *Third International Conference on 3D Digital Imaging and Modeling*, pp. – (May 2000).
- [8] C. Fruh and A. Zakhor: “3d model generation for cities using aerial photographs and ground level laser scans”, in *Computer Vision and Pattern Recognition*, Vol. 2, pp. 31-38 (2001).
- [9] Yukiko Ogawa, Kazuaki Iwamura, and Shigeru Kakumoto: “A map-based approach for extracting object information from aerial images”, in *MVA*, pp. 220-223 (Nov. 1998).
- [10] Katsuyuki Kamei, Wayne Hoy, Takashi Tamada, and Kazuo Seo: “Modeling of urban scenes by aerial photographs and simply reconstructed buildings”, in *MVA*, pp. 224-227 (Nov. 1998).
- [11] Z C Shi and Ryosuke Shibasaki: “Automated building extraction from digital stereo imagery”, *Automatic Extraction of Man-Made Objects from Aerial And Space Images*, pp. 119-128 (May 1997).
- [12] 相良毅, 大沢裕, 坂内正夫: “マルチメディア地図情報システムの構想-航空写真とデジタル地図の融合-”, 第 6 回機能図形シンポジウム講演論文集, pp. 27-32 (Apr. 1995).
- [13] Nevatia R and Price Keith E: “Automatic and interactive modeling of buildings in urban environments from aerial images”, in *IEEE International Conf. on Image Processing*, Vol. III, pp. 525-528 (2002).
- [14] Kim Zu Whan, Huertas Andres, and Nevatia Ramakant: “Automatic description of buildings with complex rooftops from multiple images”, in *Computer Vision and Pattern Recognition*, Vol. II, pp. 272-279 (Dec. 2001).

- [15] 納富幹人, 小澤史朗, 全炳東: “移動体観測による都市モデルの構築”, 電子情報通信学会論文誌, Vol.**J81-D-II**, pp. 872–879 (May 1998).
- [16] Zhigang Zhu, Guanyon Xu, and Xueyin Lin: “Panoramic EPI generation and analysis of video from a moving platform with vibration”, in *Computer Vision and Pattern Recognition99*, pp. 531–537 (1999).
- [17] “EROS DEM data”, <http://edc.usgs.gov/geodata/>.
- [18] P E Debevec, C J Taylar, , and J Malik: “Modeling and rendering architecture from photographs: A hybrid geometry-and image-based approach”, in *Proceedings of ACM SIGGRAPH 96*, pp. 189–198 (1996).
- [19] Shigang Li: “Localization along routes based upon panoramic view and gps in large-scaled outdoor environments”, in *Asian Conference on Computer Vision*, Vol. 1, pp. 11–16 (2000).
- [20] Y Sato, M Wheeler, , and K Ikeuchi: “Object shape and reflectance modeling from observation”, in *Proceedings of ACM SIGGRAPH 97*, pp. 379–387 (1997).
- [21] A Lippman: “Movie-maps. an application of the optical videodisc to computer graphics”, in *Proceedings of ACM SIGGRAPH '80*, pp. 32–43 (1990).
- [22] S E Chen: “Quicktime vr - an image-based approach to virtual enviroment navigation”, in *Proceedings of ACM SIGGRAPH'95*, pp. 29–38 (1995).
- [23] E H Adelson and J Bergen: “The plenoptic function and the elements of early visi on”, in *MIT Press Cambridge, MA*, pp. 3–20 (1991).
- [24] T Naemura and H Harashima et al.: “Ray-based creation of photo-realistic virtual world”, in *Virtual Reality and MultiMedia (VSMM'97)*, pp. 59–68 (1997).
- [25] Marc Levoy and P Hanrahan: “Light field rendering”, *ACM SIGGRAPH*, pp. 31–42 (1996).
- [26] S J Gortler, R Grzeszczuk, R Szeliski, and M F Cohen: “The lumigraph”, *ACM SIGGRAPH*, pp. 43–54 (1996).
- [27] Heung-Yeung Shum and Li-Wei-He: “Rendering with concentric mosaics”, in *Proceedings of ACM SIGGRAPH*, pp. 299–306 (1999).

第 2 章

鏡面型全方位カメラによる広域空間の効率的 獲得

都市空間をイメージベース法で記述するためには、膨大な入力画像が必要であり、そのため、どのようにして効率よく画像を取得するかが重要な課題である。

本論文では、この問題を解決するため、入力画像として全方位画像を使用する。全方位画像は画像撮影地点から全方位を撮影しているため、画像撮影地点において、水平方向の 360 度すべての光線を持っている。このような全方位画像を撮影するデバイスとして鏡面を用いた全方位カメラがある。

本章では、この鏡面反射型の全方位カメラを用いて効率的に全方位データを計測する手法について説明する。

2.1 はじめに

本章では、効率的に広域空間の情報を獲得するために鏡面型の全方位カメラを用いる手法について述べる。鏡面型の全方位カメラは、一度の撮影で水平方向の 360 度すべての方向の映像を撮影できる利点がある反面、鏡に写った景色を撮影するため、通常のカメラで撮影した画像にはない特徴や性質がある(図 2.1 に撮影した画像例を示す)。そのため、都市のモデリング等に利用するためには、鏡面型全方位カメラの特性を分析しなければならない。また、鏡面型全方位カメラは、全方位を一つの画像に収めるため、必然的に解像度が低くなるという問題点もある。

以下では、まず、鏡面型の全方位カメラの特性について紹介し、鏡面全方位カメラで撮影した画像から通常の画像に戻す方法について述べる。また、鏡面全方位カメラの問題点である低解像度について、これを復元する手法を提案し、実験により提案手法が有効であることを示す。最後に、鏡面型全方位カメラを用いたデータの取得方法について検討しまとめる。

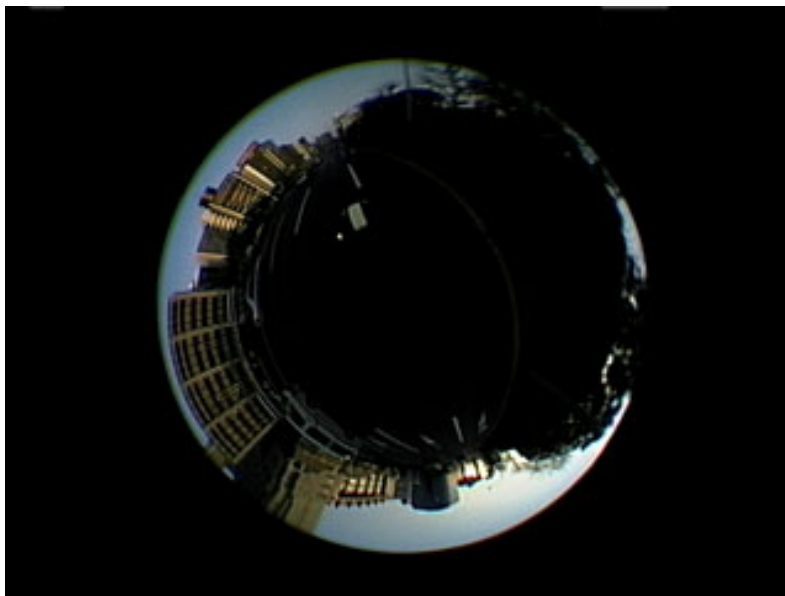


図 2.1 全方位画像の例(双曲面型)

2.2 鏡面反射型全方位画像

全方位画像を簡単に撮影する手段として鏡面反射型の全方位カメラ [1-5] の利用が考えられる。ここで鏡面反射型の全方位カメラとは、通常のカメラの前に球面や方物回転面または双曲回転面等の鏡をおき、これに周辺の状態を反射させ撮影する事で、全方位画像を得られるようにしたものである。この鏡面反射型全方位カメラで撮影した画像を画像変換処理することで透視変換画像を作ることができる。以下に、本研究で用いた 2 種類の鏡面反射型全方位カメラの特徴について述べる。

2.2.1 双曲回転体型

双曲回転体の鏡面に周囲を投影し撮影する全方位カメラ [1, 6, 7] は、双曲面の性質として二つの焦点を持つ。この双曲回転型の鏡面は、図 2.2 に示された通り、双曲面の焦点の性質から、一つの焦点に向かう光線が回転体鏡面で反射されてもう一つの焦点に集まるという特徴を持つ。このため片方の焦点にカメラを置くと、あたかももう一つの焦点から全周囲を撮影したかのような映像を撮影することが出来る。

双曲回転型の全方位カメラの形状を表す式および、全方位カメラ画像から透視変換された画像を得るための計算式は以下のようなものである。ここで、 X, Y, Z は 3 次元空間中の座標値を表しており、 x, y は、これが 2 次元画像面に投影された時のピクセル座標を表している。

$$\frac{X^2 + Y^2}{a^2} - \frac{Z^2}{b^2} = -1 \quad (Z > 0) \quad (2.1)$$

$$(c = \sqrt{a^2 + b^2})$$

$$\text{Focal point} : (0, 0, +c) \quad (0, 0, -c)$$

(a, b, c are parameter of hyperbolical mirror face)

$$Z = \sqrt{X^2 + Y^2} \tan \alpha + c \quad (2.2)$$

$$\alpha = \arctan \frac{(b^2 + c^2) \sin \gamma - 2bc}{(b^2 - c^2 \cos \gamma)} \quad (2.3)$$

$$\gamma = \arctan \frac{f}{\sqrt{x^2 + y^2}} \quad (2.4)$$

$$\theta = \arctan \frac{Y}{X} = \arctan \frac{y}{x} \quad (2.5)$$

α : an elevation between upper focus and p

γ : an elevation between lower focus and p

θ : an azimuth angle

$$x = X \times f \times \frac{b^2 - c^2}{(b^2 + c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}} \quad (2.6)$$

$$y = Y \times f \times \frac{b^2 - c^2}{(b^2 + c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}} \quad (2.7)$$

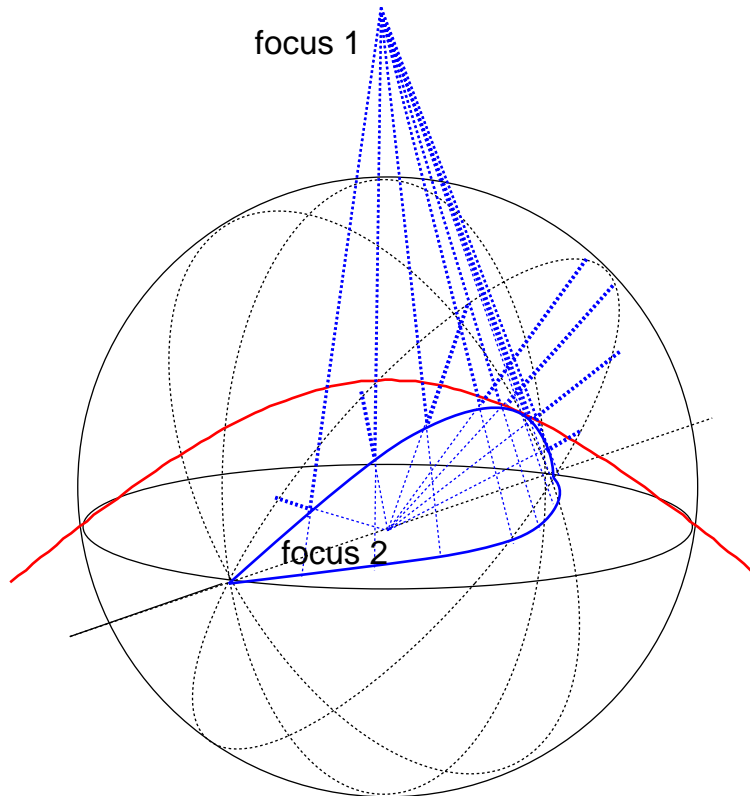


図 2.2 双曲回転型全方位カメラの光学モデル

このように，三角関数を含まない式 (2.6) と (2.7) の計算式により，3次元空間上の点を2次元平面の座標に変換することが出来る．

2.2.2 放物回転体型

放物回転体の鏡面に周囲を投影し撮影する全方位カメラ [8] は，放物面の性質として一つの焦点を持つ．従って，カメラの焦点を放物回転体鏡面の焦点に合わせることで，周囲の環境を正射影した画像を撮影することが出来る (図 2.3)．これにより，撮影画像から透視変換画像などへの画像変換を容易に実現することが可能である．以下に放物回転型の全方位カメラの形状を表す式と，放物回転型の全方位カメラから透視変換された画像を得るための計算式を示す． X, Y, Z は3次元空間中の座標値を表しており， x, y は，これが2次元画像面に投影された時のピクセル座標値である．

$$\frac{h^2 - r^2}{2h} = Z \quad (2.8)$$

$$(r^2 = X^2 + Y^2)$$

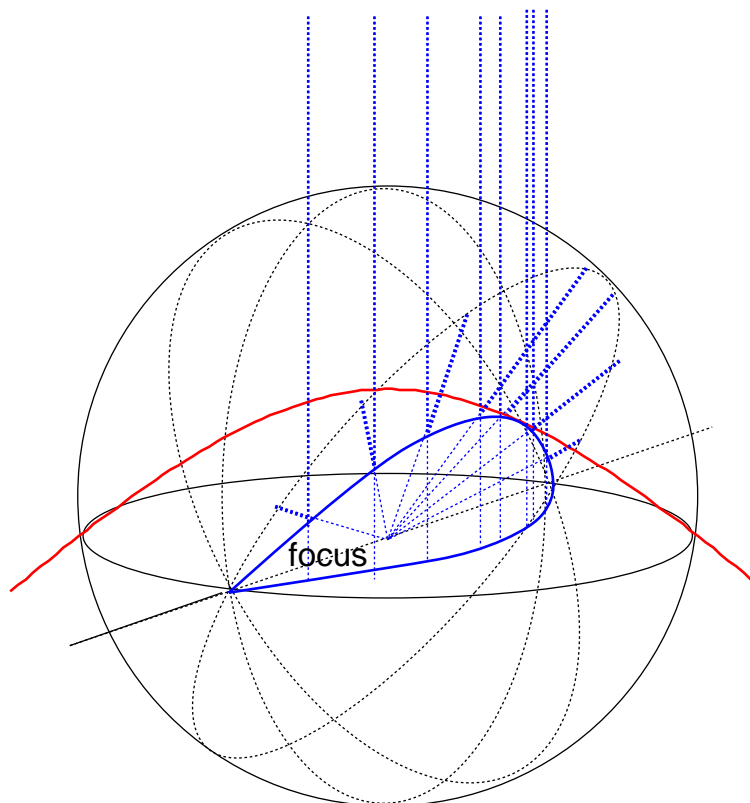


図 2.3 放物回転型全方位カメラの光学モデル

Focal point : $(0, 0, 0)$

(h is a parameter of parabolical mirror face)

$$\theta = \arccos \frac{Z}{\sqrt{X^2 + Y^2 + Z^2}} \quad (2.9)$$

$$\phi = \arctan \frac{Y}{X} \quad (2.10)$$

$$\rho = \frac{h}{(1 + \cos \theta)} \quad (2.11)$$

θ : a polar angle

ϕ : an azimuthal angle

ρ : a distance between focus and p

$$x = \rho \sin \theta \cos \phi$$

(2.12)

$$y = \rho \sin \theta \sin \phi$$

(2.13)

このように、式 (2.12) と (2.13) の計算式により、3 次元空間上の点を 2 次元平面の座標に変換することが出来る。全方位画像から透視変換画像へ変換する際には、式 (2.12) と (2.13) を用いて、予め、全方位カメラ面からピンホールカメラ面への変換を計算して変換テーブルに保存しておけば、実際の処理の際には、変換処理がピクセルからピクセルへマッピングで済むため、高速な処理が可能となる。

2.2.3 鏡面反射型全方位画像の特徴

時空間の整合性

鏡面反射型全方位カメラは、周囲 360 度の景観を 1 度の撮影で獲得できるため、時間的、空間的な整合性を保ちながら、連続的な画像データを計測することが出来る。野外環境等では非常に短い時間であっても天候などが大きく変化するため、広域な環境のデータを、光源環境等の時間的な整合性を維持しながら取得することが難しい。このような場合、全方位カメラは極めて適したデバイスである。また、鏡面反射型全方位カメラは、全周囲を 1 つの CCD カメラで計測するため、空間的な整合性を常に保つことが出来るが、これは複数台のカメラを用いて画像を統合する手法で実現することは容易ではない。

画像変換処理

2.2.1 節及び 2.2.2 節で述べたように、簡便な式で画像変換が可能であるため、任意方向の透視変換画像を容易に生成することが出来る。そのため、透視画像変換を実時間でこなうことが出来、アクティブカメラのようにカメラ自体を動かさなくても、注視点の透視画像を実時間で提示することも可能となり、ビデオ会議システム [8] などで利用することが出来る。また、HMD と組み合わせて、自由な方向を見ることが出来る VR システムでの応用 [1] も可能である。

解像度

鏡面反射型全方位カメラは、周囲 360 度の撮影が 1 度で済むため、時間的、空間的な整合性が容易に得られることや透視変換画像への変換が簡便である等の利点がある反面、全方位を 1 枚の画像中に写し込むため、画像解像度が著しく低いという問題がある。本論文では、このような解像度の低い画像を複数枚利用し解像度を向上させる手法で、この問題を解決する。解像度を向上させる手法の詳細について次節で述べる。

2.3 鏡面反射型全方位画像の解像度の向上

鏡面反射型全方位カメラは、全方位を 1 枚の画像中に写しこむため、画像解像度が著しく低いという問題点がある。本節では、このような解像度の低い画像を複数枚利用することで、解像度を向上する手法を提案する。また、本手法を応用することで、解像度の向上だけでなく、電柱や街路樹、歩行者などの障害物を除去することが可能であることも示す。以下では、前提となる撮影経路について述べ、続けて関連研究を紹介し、最後に本論文における画像の向上手法を提案しまとめる。

2.3.1 撮影経路

本論文は、広域空間の効率的なデータ獲得を目的の一つとしている。全方位カメラを用いることで、部屋の内部全体のように、ある程度広域な環境を効率良く獲得することが可能ではあるものの、都市空間のように非常に広域な空間に対しては、全方位カメラを使用するだけでは十分ではない。

そこで、本論文では、広域な空間を効率良くサンプルするため、既知の軌道上、都市空間においては公道上で、カメラを車載して走りながら映像を全方位カメラでサンプルする。また、議論を簡単にするため直線道路を等速で走行しながら都市の全方位画像を撮影したものと仮定する。もちろん、車線変更や曲がりくねった道路を走ったため直線以外の軌跡となったとしても、GPS やジャイロセンサ等で各々の撮像位置が既知ならば同じ議論が展開できるため、この仮定により提案手法が一般性を失うことは無い。

従って、本論文における撮影経路の前提条件は以下のようなものである。

- カメラは車の屋根に水平に設置してある。(全方位画像は車の周り 360 度を撮影することが出来る)
- 車は等速、直線運動をする。
- 車の位置、速度などの状況はセンサにより既知である。

2.3.2 関連研究

画像を解像度を向上させる試みはこれまで多くの研究がなされてきた。このような画像の解像度向上手法を超解像度化と呼ぶ。超解像度化は、一般に $y = Hx + b$ (y : 観測値, H : linear degradation operator, x : 実際の対象物, b : ノイズ) において、観測値 y から x を求めるという逆問題の一種である。

これまでのアプローチとしては主に、1 枚の画像から復元する試みと、複数枚画像を利用する手法とに大きく分けられる。前者はノイズ除去や経験的な知識等を利用した画像復元なども含み、本論文における超解像度化とは異なるため軽く触れるに留める。以下にそれぞれの研究についてまとめる。

一枚からの復元

代表的な手法としては以下のようなものがある。

- ・ ローパスフィルター等によるノイズ除去
- ・ 最近傍値を用いた補間
- ・ バイリニア補間
- ・ スプライン補間

これらを組み合わせたり，経験的な知識を加えることで画像の超解像度化を行う．基本的には人間の目が画像の不連続（境界線等やノイズ）に対して敏感であることから，スムージングがベースとなる．またこれら手法を，超解像度化とは呼ばずに単に画像の復元・修復ととらえることも多い．このようなスムージングに対して，逆に境界などをシャープに復元することによる超解像度化の試みとしては，

- ・ 画像のローカルな統計量を利用することで，繰り返し処理による超解像度化を試みる方法 [9]
- ・ マルコフネットワークにより予め画像の特徴などを学習しておき，これにベイズ推定を組み合わせる超解像度化を試みる方法 [10]

などが挙げられる．

複数枚からの超高解像度化

対象とする複数画像に動きがあるか無いかで大きく二つに分けることができる．また動きのある画像についても，カメラは静止しており撮影対象が動いている場合と，カメラ自身が動いている場合の二つに分けることができる．

静止した連続画像

ベイズ推定等の統計的な手法により，サブピクセル値を推定することが主な手法となる．その際に対象物体の 3 次元形状を考慮 [11] したり，さらに 3 次元形状そのものも精密に復元 [12] するなどの研究も行われている．また，このような統計的な手法のみに頼らないものとしては，

- ・ RGB それぞれのチャンネルの，画像内，画像間の相関を用いる方法 [13]
- ・ 複数のカメラで撮影した画像をブロックマッチングにより registration することで超解像度化 [14].
- ・ [14] に加えて時間方向にも画像系列を統合 [15] する手法
- ・ イメージワーピングを用いる方法 [16]
- ・ 画像そのものでなくエッジを利用する方法 [17]

等がある．

対象物体に動きのある連続画像

ビデオカメラを固定した状態で、動きのある物体を撮影した映像からの超解像度化を行う研究を指す。基本的な手法としてはオプティカルフローやブロックマッチング等により対象物体の動きを推定し、その結果を使って超解像度化を行う。これまでの主な研究としては、

- ・ ブロックマッチングによる動きベクトル推定 [18]
- ・ カルマンフィルターを用いる方法 [19]
- ・ 各カラーチャンネルで動きベクトルを推定することで、ブロックマッチングによる動きベクトルの精度向上を利用した超解像度化 [20]
- ・ 映像を空間方向にはバイリニア補間し、時間方向には動き補償しながら繰り返し処理を行う手法 [21]
- ・ 人間の顔のように、複雑な形状をした物体が自由に動いている映像からの超解像度化 [22]

等が挙げられる。

カメラに動きのある連続画像

本論文のように、カメラが動く場合には、カメラのエゴモーション等の推定が必要になり、比較的小さい対象物の動き推定のみによる手法よりも困難になることが多い。そのため、これまでの研究としては、

- ・ 開口時間が 0 でないことによるモーションブラーを、動き補償により解消することによる超解像度化 [23].
- ・ 背景も巨大なオブジェクトとみなすことでカメラモーションにも対応した超解像度化 [24]

等の研究にとどまっている。比較的カメラの動きが大きい後者であっても、動きのある物体をそれぞれセグメンテーションしてトラッキングすることで実現しており、カメラの動きがさらに大きい本研究において、これを適用することは難しい。

その他

一般的な画像の空間周波数などを考慮した画像の超解像度化の理論的な分析としては [25] に詳しい。また、以下のような研究も試みられている。

- ・ バックプロジェクション法による超解像度化 [24]
- ・ モザイクングでは重なりあう部分が多いにもかかわらず、これら画像を単に位置合わせにしか利用していないことが多い。そこで、これら画像系列をサブピクセルベースで位置合わせすることで、モザイクングと超解像度化を同時に行う研究 [26]

また、全方位画像を用いた超解像度化の試みも行われている [27]。この研究においてはリアルタイム処理を目指している点が注目される。ここではカメラ自体は固定されている。これに対して、本研究ではカメラ自身が動きながら撮影した映像を対象としている点、また、イメージベース法のための映像データの構築が目的である点などが異なっている。

入力画像が一枚	入力画像が複数枚		
	対象物体は静止		対象物体の動きあり (カメラは静止)
	カメラは静止	カメラの動きあり	
Hong et al. [9] Freeman and Pasztor [10]	Cheeseman et al. [11] Smelyanskiy et al. [12] Igarashi et al. [14] Nakazawa et al. [15] Chiang and Boulton [16] Chiang and Boulton [17]	Patti et al. [23] Irani and Peleg [24]	Schultz and Stevenson [18] Elad and Feuer [19] Tom and Katsaggelos [20] Hong et al. [21] Baker and Kanade [22]

表 2.1 超解像度化手法のまとめ

2.3.3 映像の超解像度化

関連研究で述べたようにこれまで多くの超解像度化の手法が提案されてきた．本研究においては対象がビデオ映像であるため，複数枚の連続画像を利用することが出来る．しかしながら，本研究ではカメラ自身が大きく移動するため，これまで主に試みられてきた超解像度化手法をそのまま適用することは出来ない．そこで，本論文では時空間解析を応用してカメラのエゴモーションを正確に推定することで，大きな動きの中でもピクセル単位での初期位置合わせを実現し，これによる超解像度化を試みる．具体的には，画像の解像度は本来，レンズの開口関数や空間周波数応答等の光学的作用に，ピクセル密度や AD 変換等の電気的作用が複雑に絡んでいるものと考えられるが，これを一つの点拡がり関数で表現することで，この関数をカーネルとするバックプロジェクション法により，画像誤差を最小化することによる超解像度化を試みる．本研究においてバックプロジェクション法を用いた理由としては，抽象化したカーネルでも安定して収束する性質や，カメラ自身が移動するためベイズ推定のような統計的な手法を使うことが難しい場合でも，誤差最小化により超解像度化を実現することが出来るためである．

以下ではバックプロジェクション法について述べた後，本論文におけるカーネルの設計方法について説明する．また，本論文では動きのある映像を対象としているため，カメラのエゴモーションの推定及びこれら映像の位置合わせが必須である．これら手法についても述べる．

バックプロジェクション法

本研究で用いるバックプロジェクション法は文献 [24] を主に参考に行っている．具体的な手法としては，得られた超解像度画像と実際の画像系列との誤差を評価関数を用いて表現し，これを最小化するように繰り返し処理を行うというものである．個々の繰り返し処理は，2段階のステップを踏む．ステップ 1 は複数画像からの超解像度画像の生成であり，ステップ 2 はその超解像度画像を再び低解像度化する画像処理である．ステップ 1 の処理における超解像度画像の生成は，ステップ 2 の誤差をその前の繰り返し処理で生成された超解像度画像にバックプロジェクションすることで行

われる．以下にその過程を示す．

$$f^{(n+1)} = f^{(n)} + \frac{1}{K} \sum_{k=1}^K H_k^{-1} ((g_k - g_k^{(n)}) * p) \quad (2.14)$$

ここで K は入力画像の数であり， $f^{(n)}$ は n 番目に生成した超解像度画像， g_k は k 枚目のオリジナル入力画像， $g_k^{(n)}$ は $f^{(n)}$ から再生成した低解像度画像である．また， H_k^{-1} は後に述べる移動マトリクス H_k から計算され， p は後述のバックプロジェクションカーネルである．

また，誤差評価関数はステップ 2 の処理で得られた低解像度画像と，実際の入力画像系列とを用いて以下の式により定義される．

$$e^{(n)} = \sqrt{\frac{1}{K} \sum_{k=1}^K \|g_k - g_k^{(n)}\|_2^2} \quad (2.15)$$

この誤差評価関数が最小値に収束した時の画像を最終的な超解像度画像とする．評価関数の最小化は，繰り返し処理の度に画像系列の中から順に一枚ずつ（第 k 画像とする）サブピクセル単位で xy 方向に平行移動（マトリクス H_k ）することで行われる．また，効率よく収束させるためにニュートン法を用いる．

カーネル設計

バックプロジェクション法は一種の逆問題の解法であるため，カーネルの設計次第では，解が不安定になったり，現実とは異なる画像に収束してしまう．文献 [24] では，超解像度画像が収束するためのカーネルの条件について分析している．また，本研究においては通常のカメラと異なり，全方位カメラを使用しているため，全方位カメラに適したカーネルの設計が必要である．全方位カメラ特有の光学特性についての分析は [27] に詳しい．

ここではこれらカーネルに関する分析及び全方位カメラ特有の光学特性を考慮した上で，全方位カメラに適したカーネルを設計する．理論的には，カーネルの設計には，全方位カメラの MTF，CCD によるローパスフィルタとしての振る舞い，および俯角による周波数要求の変化の 3 つを特に考慮する必要があると考えられる．しかしながら，これら全ての要素を厳密に反映させることは現実的には困難であり，また，文献 [24] で解析されている通り，超解像度画像が収束するためのカーネルは，ただ一つだけ存在している訳ではなくある程度自由度がある．そこで，本研究においては，これらの複数の要素を組み合わせたものを，1 つの点拡がり関数で表現することで具体的な処理を実現した．実際には，CCD によるローパスフィルタとしての効果を点拡がり関数で表し，この関数の拡がり度合いを俯角により変化させている．具体的には J_1 を第 1 種第 1 次のベッセル関数， F をフーリエ変換とする次式で表される $p(l)$ をカーネルとして用いた．ここで， d は俯角により変化する値であり，全方位画像の中心に近いほど大きく離れるほど小さな値をとる．また， l は注目画素からの距離であり， k は係数である．

$$p(l) = F \left(\left[\frac{2J_1(kdl)}{kdl} \right]^2 \right) \quad (2.16)$$

エゴモーションの推定

本論文においては、カメラが固定でないため、エゴモーションの推定が必要である。また、バックプロジェクション法において超解像度画像を収束させるためには、前処理として入力画像系列をピクセル単位で初期位置合わせしておく必要がある。

本論文では、2.3.1節で述べたように、撮影経路を等速直線と仮定しているため、エゴモーションの推定のための初期位置あわせに、EPI による時空間解析を用いることが出来る (EPI については、第 2 章付録参照)。これにより、効率よく高精度な初期位置の推定が可能である。また、初期位置あわせ後、高精度なエゴモーション推定には、階層ブロックマッチングを利用する。これは、本論文の対象映像がビデオ映像であるため、超解像度化に必要な画像枚数を取得するための時間は微小時間ですみ、カメラの動きを車の進行方向及び上下の振動だけと仮定し、回転及び伸縮の影響はほとんどないものとして差し支えないためである。

以下では、エゴモーションの推定方法を、『初期位置あわせ』、および、『初期位置あわせ後の推定』に分けて、それぞれ具体的に説明する。

『初期位置合わせ』

時空間画像の解析手法として EPI 法が知られている [28, 29]。この EPI 法とは画像から 3 次元情報を獲得する為の解析手法である。EPI 法は、時空間画像の断面 (Epipolar Plane Image) 上では画像間の対応点が連続することを利用して、現実の画像で得ることが困難な対応点の取得を直線の検出に置き換えることで、比較的安定した 3 次元情報の獲得を可能にする手法である。(EPI については、第 2 章付録参照)

本論文では対象映像として都市を車載カメラから撮影した映像を用いているため、以下に述べる理由から EPI 解析に適している。

1. 道路に沿って直進運動する車から撮影しているため、微小時間においては一定速度、撮影経路を直線と仮定することが出来る
2. 一般に建物は道路に沿って平行に建っているため、EPI 上で直線の軌跡を描く
3. 建物は平面で構成されていることが多く、これを拘束条件とすることでより安定かつ現実的な解析が出来る
4. 地図などの情報とマッチングすることで建物の同定が可能である

ここでは、EPI 解析を用いることで都市の建物群を撮影したビデオ映像から 3 次元情報を獲得し、建物を同定すると同時にカメラ自身の初期位置推定を試みる。実際には、建物の形状や位置が正確には分からないため、ここで得られるカメラの初期位置は相対的なものである。

全方位カメラ画像から EPI を作成するために、全方位カメラ画像から時空間画像を生成する。この時空間画像は全方位カメラ画像を時間軸にそって積層することによって得ることが出来る (図 2.4)。通常の EPI はこの時空間画像を画像の x 軸に平行な直線 (エピポーラ線)

で切断した断面図として得ることが出来る．しかし，全方位カメラ画像におけるエピソード線は楕円または双曲線を描くため，通常カメラと同じ EPI を得るためには，このようなエピソード曲線に沿って時空間画像を切断しなければならず困難を伴う．本論文では，全方位時空間画像を単純に縦方向に切断することで得られる疑似的な EPI を使用することで，カメラの初期位置の推定を試みる．(疑似的な EPI についての詳細は，第 5 章で述べる) 図 2.5 に縦方向に切断した例を，図 2.6 に得られた EPI を示す．切断面において，通常の EPI 平面同様の直線軌跡が描かれているのが分かる．

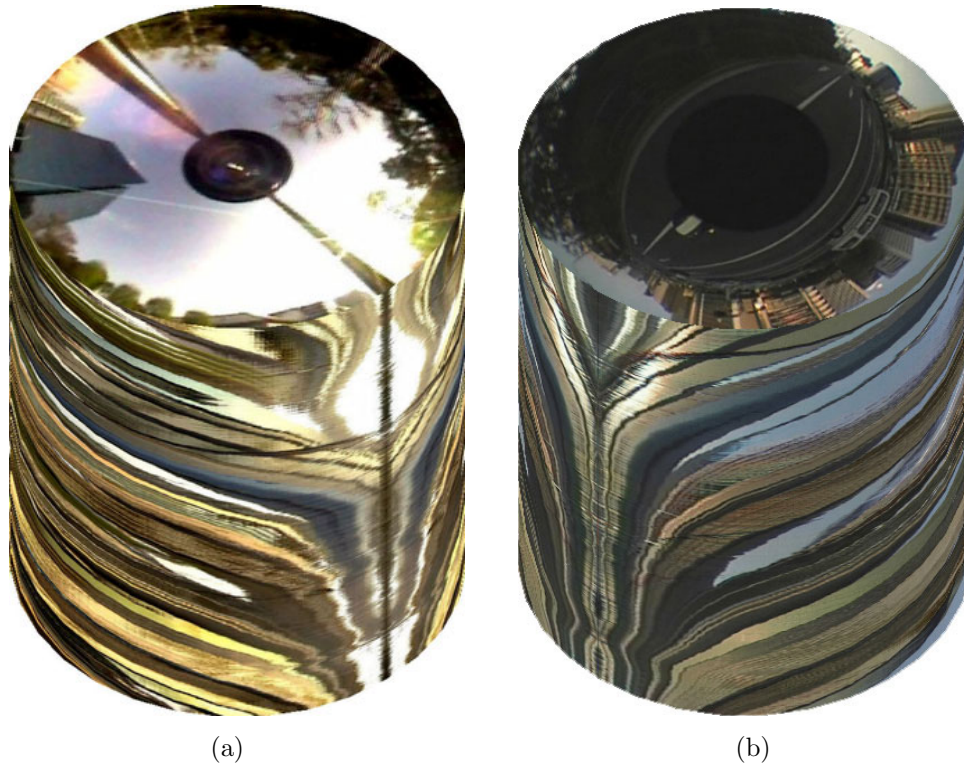


図 2.4 全方位カメラ時空間画像 (a) 放物回転体 (b) 双曲回転体

この EPI 上から直線を検出し，傾きを得ることで建物の 3 次元形状の取得とカメラの初期位置の推定が可能である．つまり，EPI 上にある『帯』は，それぞれ一つの建物の側面を表現しているため，EPI 上で検出した直線を利用してこの帯を切り出すことが出来れば，これは同一の建物だけの画像系列を透視変換して映像から抽出したことに他ならない(第 2 章付録)．EPI から『帯』を抽出する手法の詳細については，第 5 章で述べる．

『初期位置合わせ後の，階層ブロックマッチングによるエゴモーションの推定』

本論文においては対象映像として，車載カメラから撮影したビデオ映像を利用しているため，超解像度化に必要な画像枚数を取得するための時間は微小時間ですみ，カメラの動きを車の進行方向及び上下の振動だけと仮定し，回転及び伸縮の影響はほとんどないものとして

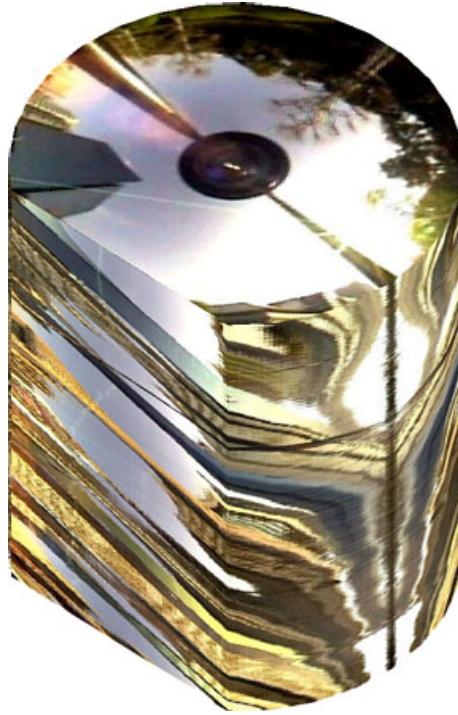


図 2.5 全方位カメラ時空間画像を縦に切断した例

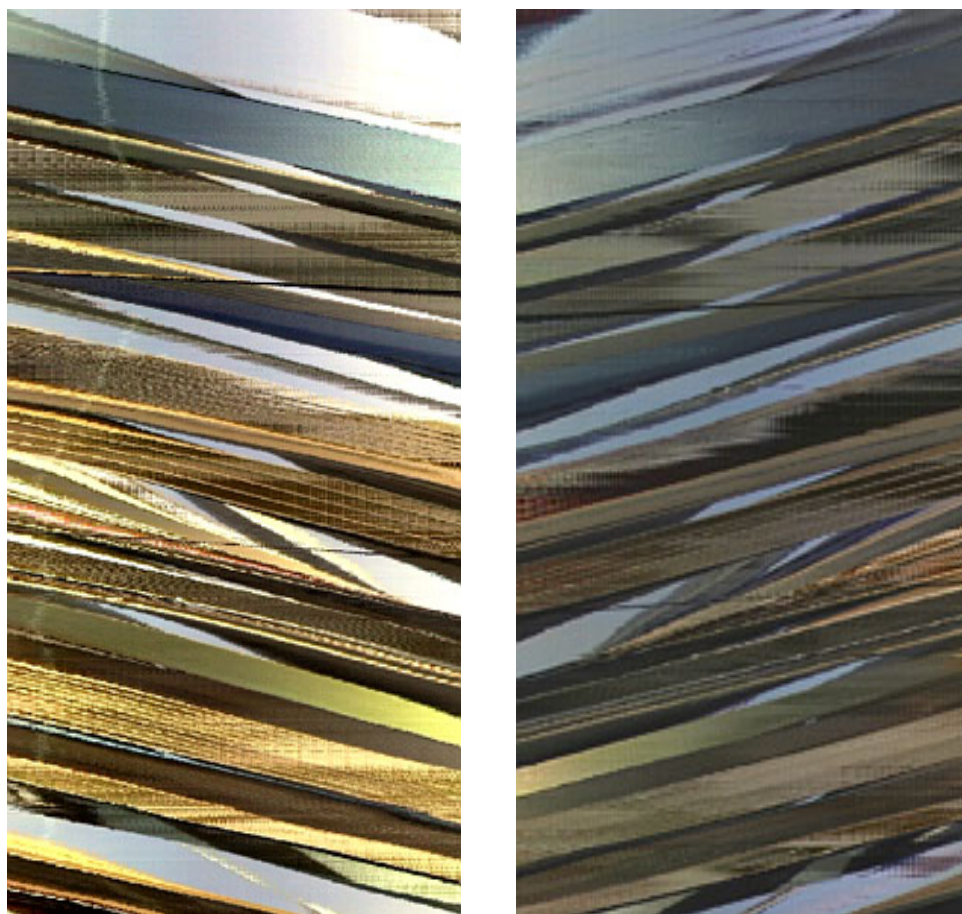
差し支えない。このため、最終的な位置あわせにブロックマッチングによる 2 次元探索を利用することが出来る。

既に入力画像系列がピクセル単位で初期位置合わせされているため、ブロックマッチングを用いた 2 次元探索によるサブピクセルベースでの位置合わせを行うことで、超解像度化を実現することが出来る。超解像度化の過程を図示すると図 2.7 のようである。評価関数 (2.15) が最小となるまで、入力画像をサブピクセル単位で微小移動させながら、繰り返し処理を行なう。

また、超解像度化を安定して行うためには、数十枚の入力画像系列が必要なため、単純なサブピクセルベースでの 2 次元探索は通常コストの高いものとなる。そこで、本論文では階層ブロックマッチングを適用し、解を収束させるためにニュートン法を用いることで高速な処理を実現する。

2.3.4 障害物の除去

本手法を統計的手法と組み合わせることで、画像の超解像度化と同時に障害物の除去が可能である。バックプロジェクション法においては得られた画像と個々の画像との誤差を最小にするようにピクセル値を決定するため、この時の評価関数 (2.15) に閾値を設定し、閾値以上にはずれたピクセル値を除外し処理をすることで障害物やノイズの除去が可能となる。しかし、初期画像が単純なピ



(a)

(b)

図 2.6 全方位カメラ画像 EPI(a) 放物回転体 (b) 双曲回転体

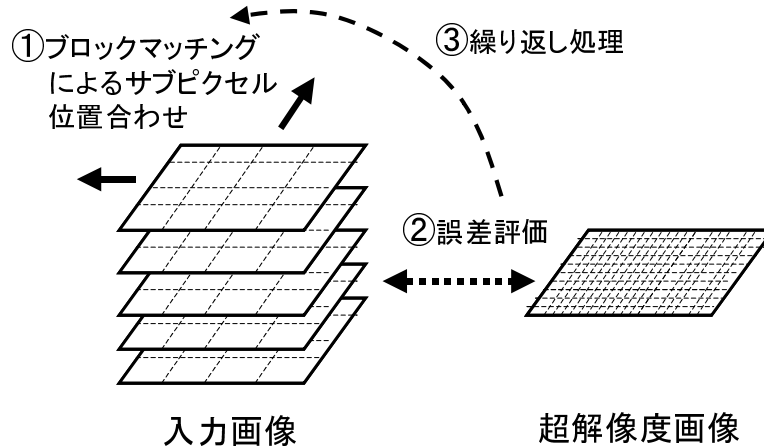


図 2.7 超解像度化の過程

クセルの平均値であれば、閾値を設定するだけでは誤ったピクセル値に収束してしまうため、本研究では投票によりピークを形成するピクセル値を採用し超解像度画像を生成している。

このような障害物の除去は、個々の建物映像を取得する際に、電柱や街路樹等を除去するのに有効である。

2.3.5 超解像度画像の画質について

これまで述べた超解像度化手法により、画像の解像度を無限に向上できる訳ではないため、実際に撮像系を設計するためには、最終的に必要な画質と、その画質を得るためにどの程度高密度な CCD カメラが必要であるかを予め分析しておく必要がある。以下では、本研究で使用した複数枚の画像をサブピクセル単位で微小位置合わせすることによる超解像度化手法が、結果的にどの程度まで画質を向上することが可能であるかを分析を行なった。また、実際の画質の向上を数値的に評価するためにシミュレーション実験も併せて行なった。

画質向上の理論的な限界

複数枚の入力画像から超解像度化を実現する手法について、これまで多くの理論的な解析が行われて来た。前述した通り、低解像度化の原因は一つではなく、光学的なばけであったり、またはピクセルが大きいことによるローパスフィルターとしてのふるまいであったりと複雑であるため、解析するには多くの困難を伴う。論文 [30] は、これら複雑な条件を整理することで、超解像度化の理論的な限界に対しての解析を行なった。特に微小位置ずれによる超解像度化は、得られる画像の解像度がもとの画像の解像度の M 倍であるとし、また CCD の大きさを S_i^2 とすると、 $(M \cdot S_i^2)$ の値が増えるにつれて画像の復元が急速に不安定となることを示した。その結果、ノイズ等の無い理想的な微小位置ずれした画像列を無限に必要な枚数使用出来ることを前提にした超解像度化においては、出力画像の解像度が入力画像の 8~16 倍を越えた時点で、新たに画像を何枚加えたとしても新

しい情報を得ることが出来なくなり、これ以上の解像度の向上が望めなくなることを明らかにした。

これは本手法においても同様であり、本研究における超解像度化の限界であると考えられる。実際の条件はさらに厳しいため、超解像度化の限界はさらに一段と低いものになる。そこで次節において、シミュレーションにより得た全方位画像を使用した実験により、実際にどの程度超解像度化が可能なのか評価を行なった。

本手法における実際の画質向上の評価

全方位カメラの光学過程をシミュレートして作成した全方位画像 (図 2.8) を用いて、本手法により超解像度化した画像の画質向上を数値的に評価した。全方位画像をシミュレートするために使用したオリジナル画像が図 2.9 である。この全方位画像からバイリニア補間した画像例を図 2.10 に、超解像度化画像例を図 2.11 に示す。入力全方位画像は実際の撮影とほぼ同じ条件にするために、全方位カメラの回転体ミラーが 400×400 ピクセルの画像にちょうど納まるように解像度を調整した。図 2.10、図 2.11 を比較すると、画像の超解像度化を視覚的に確認することができる。また、使用する画像の枚数や、繰り返し処理の回数を変化させながら超解像度化を行ない、オリジナル画像との差を PSNR (Peak Signal to Noise Ratio) を用いて評価した。PSNR は画像の S/N 比であり、値が大きい程オリジナル画像との差が小さいことになる。結果を図 2.12 に示す。バイリニア補間した画像の PSNR は入力画像 20 枚の平均値をプロットしている。このため、入力画像の枚数に関わらず一定値であり、実際の値は 15.12 である。これに対し、超解像度画像はバイリニア補間画像に対して PSNR が 3~4 程度上回っており、繰り返し処理 15 回、画像枚数 20 枚において最大値 19.86 に収束している。この PSNR の差は解像度に換算すると $4 * 4$ 倍程度に相当する。また、図 2.12 において、繰り返し処理が 10 回の時に、画像枚数が 20 枚を越えた時点でかえって PSNR 値が減少しているのは、画像枚数が多くなったため、少ない繰り返し処理では十分に収束しきれていないためと思われる。

2.4 実験

提案した手法の有効性を検討するために実際に撮影した映像を用いて実験を行った。以下にその概要と結果を述べる。

2.4.1 撮影経路

撮影経路としては、2.3.1 節で述べた条件に合うように、国道 256 号線のうち、直線に近いセグメント (東京都・赤坂付近) を等速で走行するものとする。しかし、実際には車両を理想的に等速直線させることは困難であるため、撮影したデータを GPS、ジャイロ等のデータおよび画像処理により補正したものを使用する。本実験に使用した映像の車速は約 50km/h 程度であった。

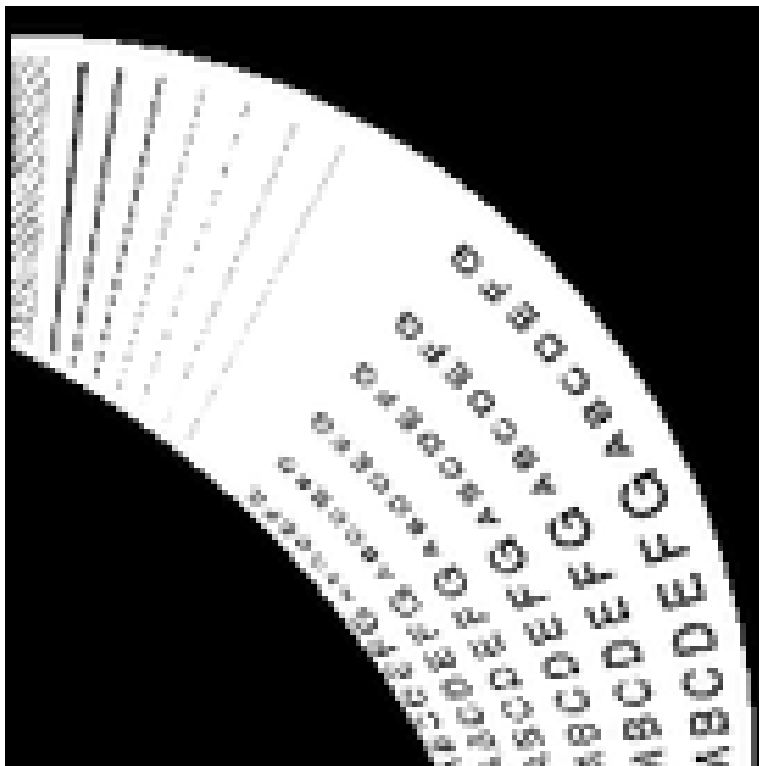


図 2.8 入力全方位画像例

2.4.2 対象映像

撮影した映像は NTSC 信号に変換された後、VCR に保存される。このインターレース信号をバイリニア補間により 640x480 ピクセルの画像系列として計算機に保存し、実験に使用した。使用した画像の例を図 2.13 に示す。また、実際の超解像度化の際には、最も画質が良いと思われる画像を最大 20 枚選択し繰り返し処理を行うように実装した。

2.4.3 結果

EPI 法を用いた初期位置合わせ

EPI 法により得られた建物の 3 次元形状とカメラのエゴモーションを利用することで、全方位画像から特定の建物のみを抽出した例を図 2.14 に示す。カメラが大きく動いているにも関わらず同じ建物が正しく抽出できていることが確認できる。

階層ブロックマッチングによる位置合わせ

このようにして得られた画像の系列を階層ブロックマッチングし、位置合わせした結果を図 2.15 に示す。このようにして EPI を用いてエゴモーションを推定し、位置合わせされた画像の系列をパッ

A B C D E F G A B C D E F
A B C D E F G A B C D E F
A B C D E F G A B C D E F
A B C D E F G A B C D E F
A B C D E F G A B C D E F
A B C D E F G A B C D E F

図 2.9 オリジナル画像

A B C D E F G A B C D E F
A B C D E F G A B C D E F
A B C D E F G A B C D E F
A B C D E F G A B C D E F
A B C D E F G A B C D E F
A B C D E F G A B C D E F

図 2.10 バイリニア補間による画像



図 2.11 超解像度画像 (入力 20 枚, 繰り返し 15 回)

クブプロジェクション法により超解像度化する。

超解像度の例

このようにして初期位置合わせされた画像の系列をバックプロジェクション法により超解像度化した画像の例を図 2.16 に示す。左側が入力画像であり、右側がそれを超解像度化したものである。左側の入力画像でははっきりしない窓枠の縦の直線が、右側の画像上の建物全面において視認できるなど、超解像度化されていることが分かる。

図 2.17 においては、入力画像が空を多く含むため、初期位置合わせとして階層マッチングを用いただけでは平均画像 (図中段) のようになってしまっている。そこで、アダプティブウィンドウを用いてマッチングした結果、超解像度化出来ているのが確認できる (図下段)。

これらの超解像度画像から、仰角による周波数応答を考慮したカーネルを使用した結果、ある程度まで仰角が変化しても、画像の超解像度化が行われていることが分かる。しかし、仰角がさらに大きな部分 (建物の上方) に関しては解像度の向上があまり見られない。これは、全方位カメラは中心部分 (透視変換画像では上方部分) においては、ピクセル密度が周辺部分に対して著しく低いことに加えて、通常ピクセル密度の高い周辺部分の MTF が最大となるように設計されているため、上方部分の情報欠損が大きすぎるためと考えられる。

障害物の除去

また、提案した手法による障害物の除去をシミュレーション画像を用いて実験した。入力画像として、連続する 36 フレームの画像に CG の木を人工的に障害物として加え、この障害物の除去を試みた。木は車速と建物の位置関係を考慮し、建物から一定の距離にあると仮定して計算される位置に添付した。障害物を加えた入力画像例を図 2.18 上段に、超解像度化した画像例を図 2.18 右下に示

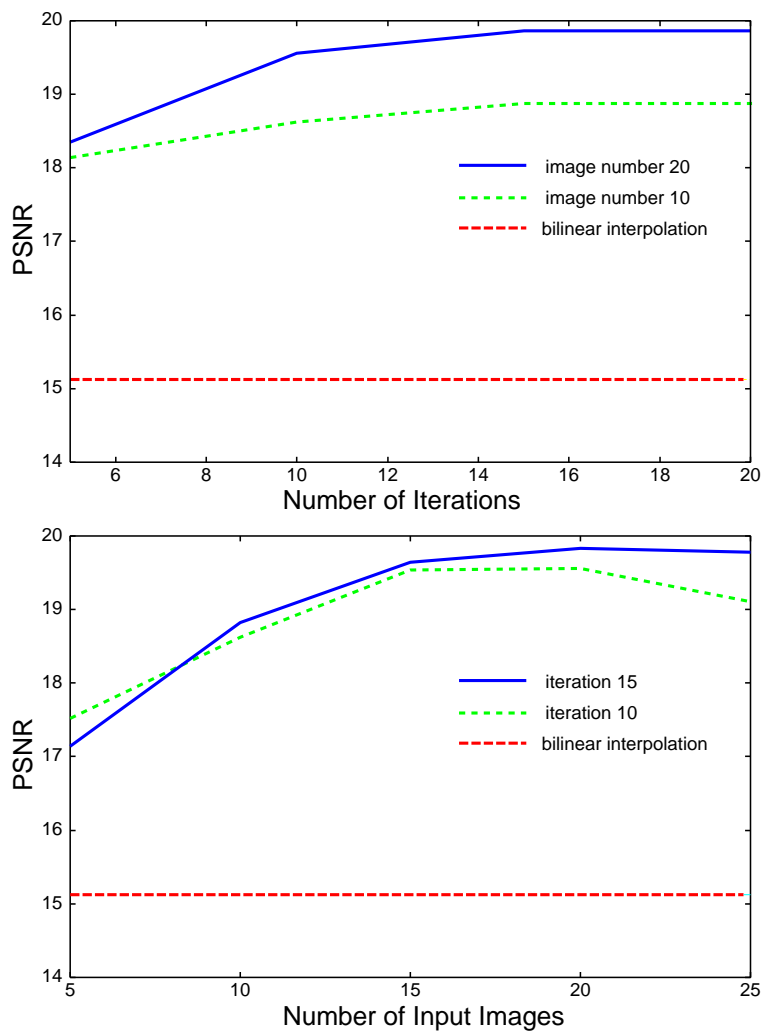


図 2.12 超解像度画像の PSNR

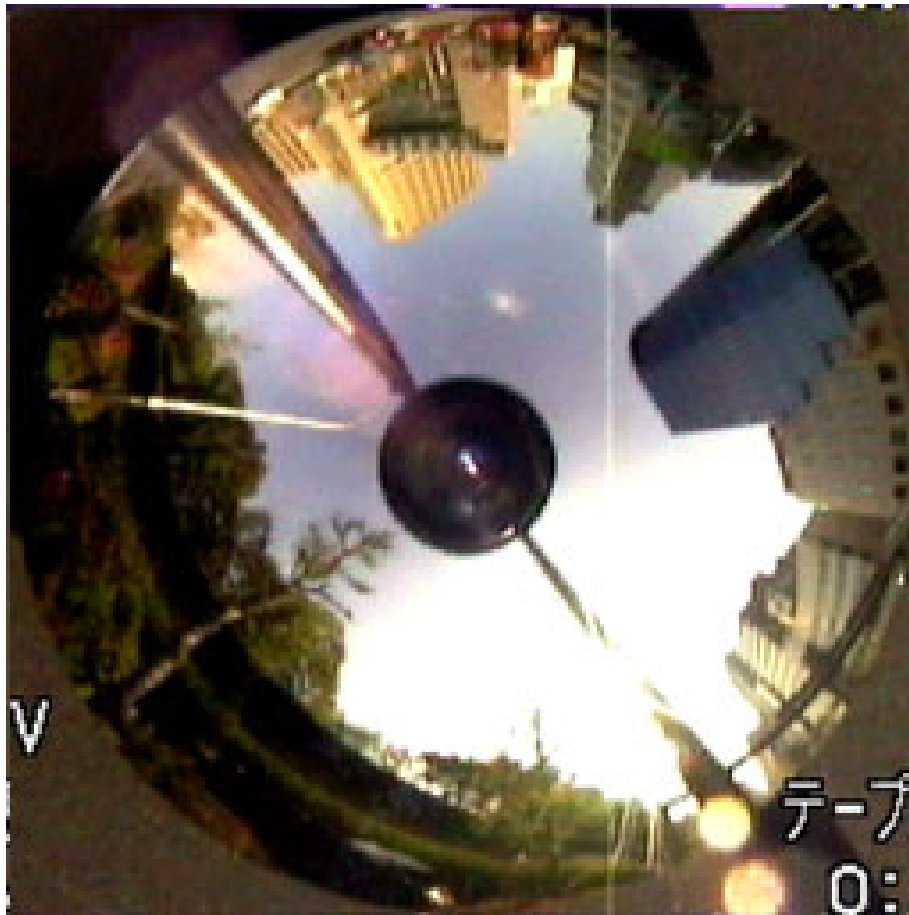


図 2.13 全方位カメラの入力画像例 (法物面型)

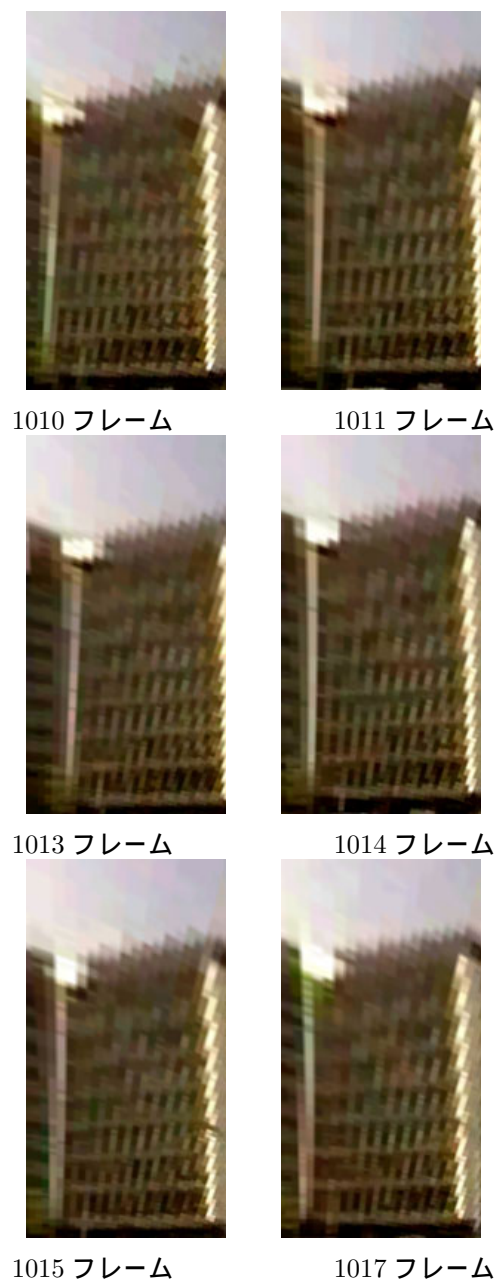


図 2.14 EPI 法による自動画像抽出例

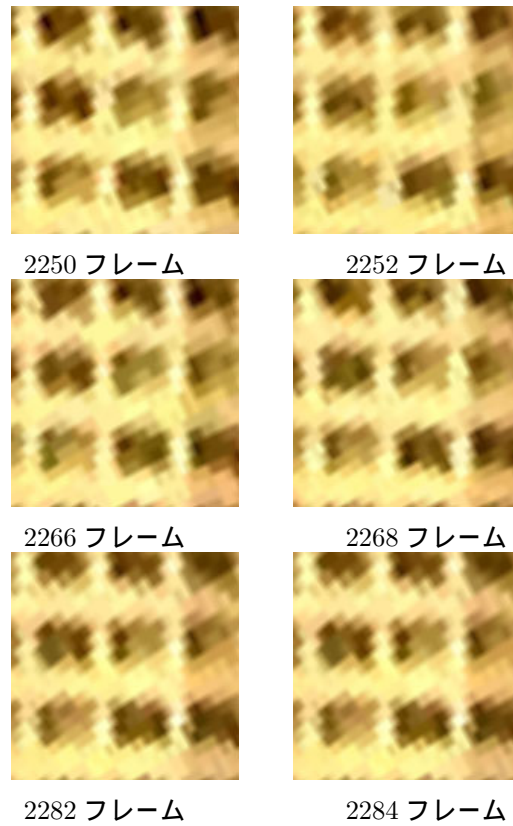


図 2.15 ブロックマッチングによる位置合わせ

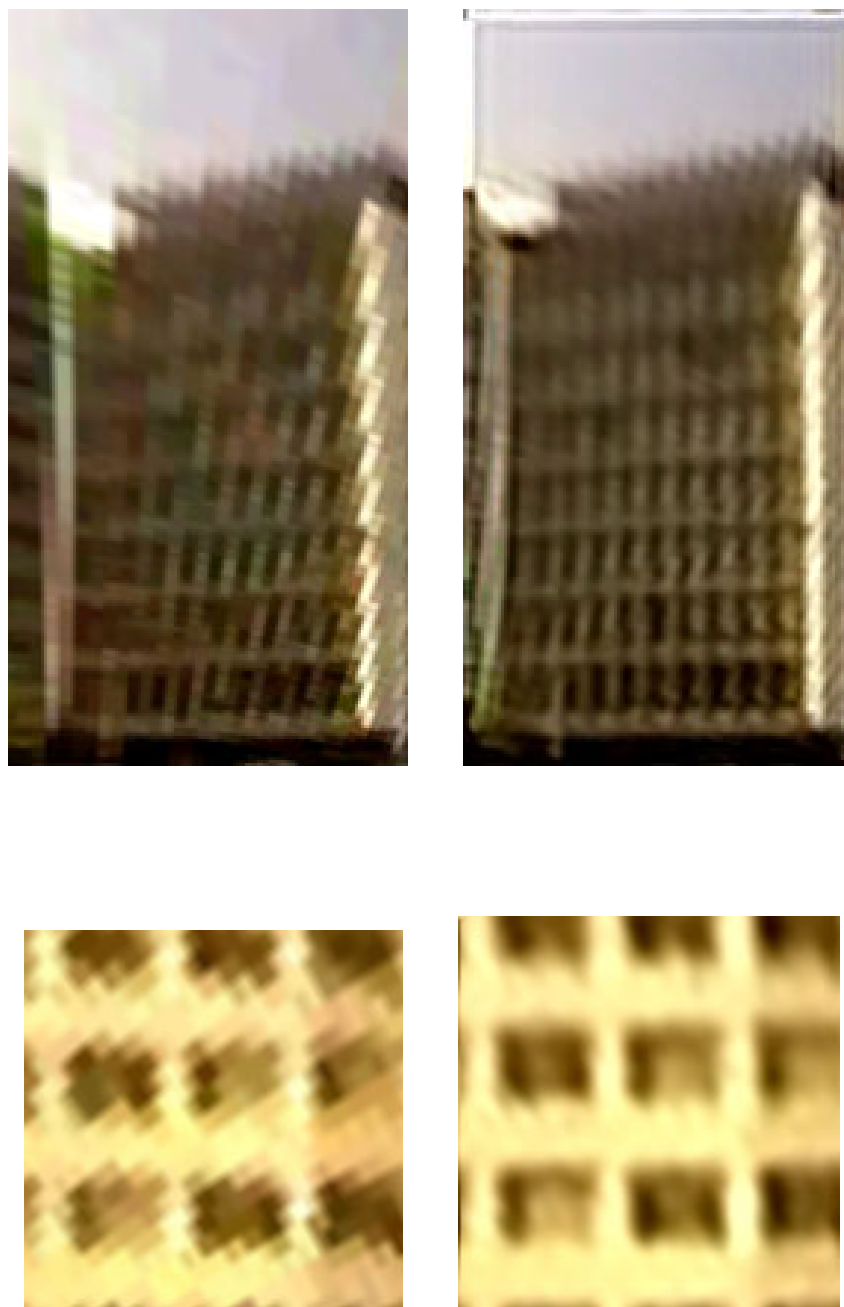


图 2.16 超解像度例:左 入力画像,右 超解像度画像

す．また，図左下の画像は上段の入力画像から EPI マッチングにより建物一つだけを抽出し透視変換した例である．これらの図より障害物が除去されて超解像度化されていることが確認できる．

2.4.4 問題点

本章で述べた手法により，鏡面反射型全方位カメラで全方位映像を撮影する際に問題となる低解像度を向上させることが出来た．しかしながら，画質の向上は 4×4 倍程度が限界であることも併せて示された．現在一般に入手可能な CCD 密度は数千 * 数千ピクセル程度であるが，ビデオカメラ用で考えた場合には，最大でもその数分の一程度であるのが現状である．従って，写実性の高い画像を合成するには，現状の鏡面反射型全方位のビデオカメラでは不十分であると言える．

また，十分に高解像度のビデオカメラ用の CCD が一般に利用可能になったとしても，超解像度化可能なレベルに十分に高い精度を持った鏡面や光学系の製作が簡単ではない等の問題点もある．

2.5 まとめ

本章では，動きのあるカメラから撮影した映像の超解像度化手法を提案し，これを全方位カメラに適用するためのカーネルの設計方法および位置合わせの手法について述べた．また，本手法を実際の映像に適用し実験した結果，画像を超解像度化することが出来た．

本研究は，これまでの超解像度化の研究が主に対象としてきた映像と異なり，カメラ自身が大きく動きながら撮影した映像を対象としている点，また，全方位カメラ画像に超解像度化を試みた点が新しい．また，本手法を統計的な手法と組み合わせることで，超解像度化と同時に障害物の除去にも大きな効果が得られることが確認できた．

今後の課題としては，未だに十分な解像度が得られていない建物上方部分の超解像度化および，建物映像データベースとして利用するために，さらなる画質の改善等が挙げられる．また，現状の CCD 密度では，現実感の高い仮想空間の構築には不十分であることも分かった．

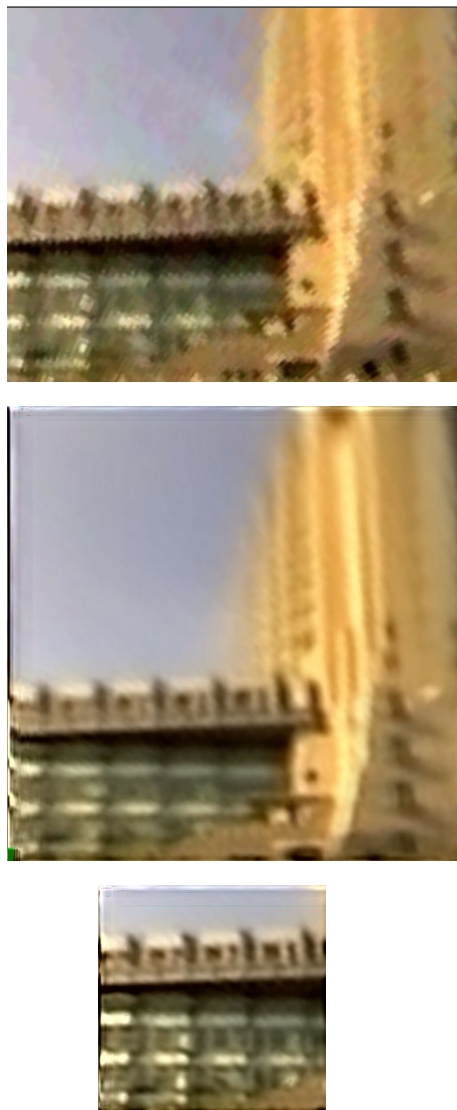


図 2.17 超解像度画像例 2:上から順番に,入力画像,超解像度画像,アダプティブウィンドウを用いた超解像度画像

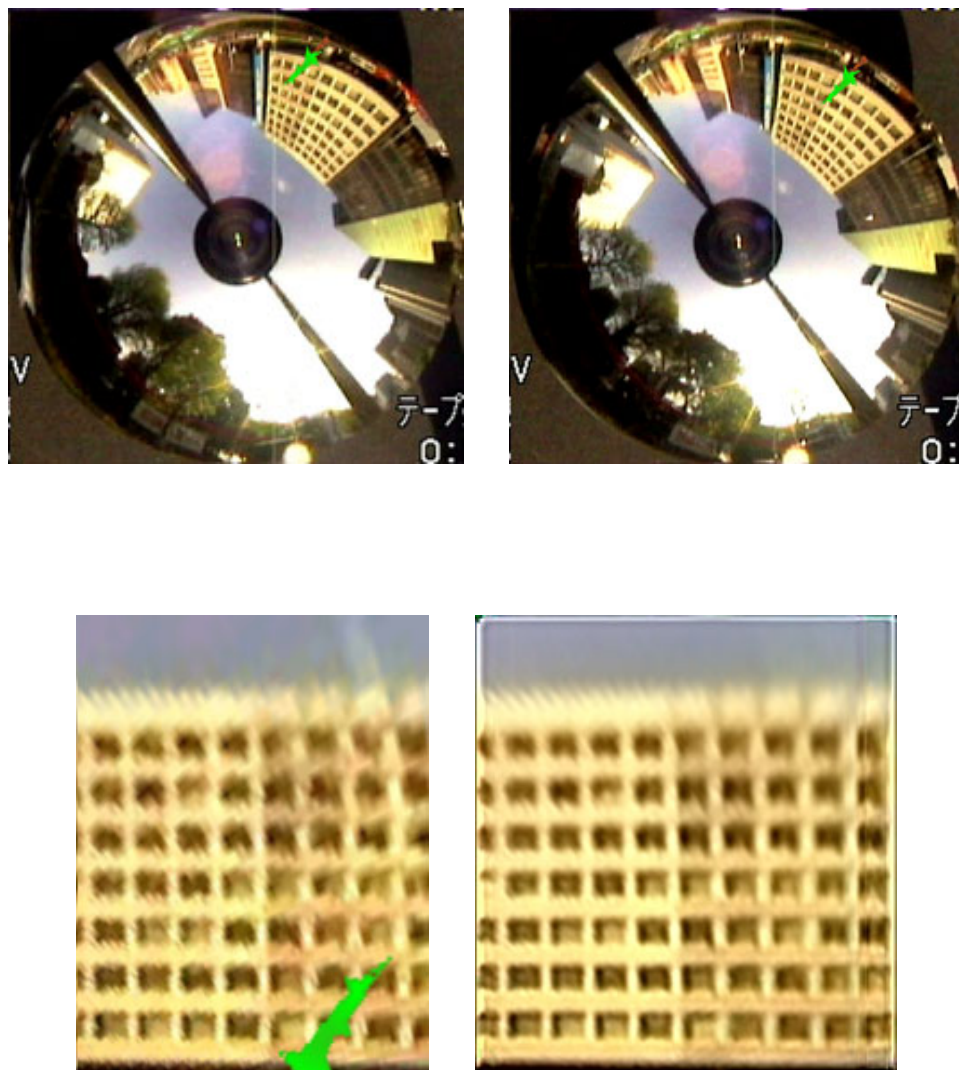


図 2.18 ノイズ除去例:上段, ノイズののった全方位画像, 左下, 透視変換画像, 右下, ノイズ除去を行った超解像度画像

【 参 考 文 献 】

- [1] Y. Onoue, K. Yamasawa, 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, Aug. 1998.
- [2] Y. Yagi and S. Kawato, “Panorama scene analysis with conic projection,” *Proc. Int. Conf. on Intelligent Robots and Systems*, pp. 181–187, July 1990.
- [3] K. Yamasawa, Y. Yagi and M. Yachida, “New real-time omnidirectional image sensor with hyperboloidal mirror,” *Proc. 8th Scandinavian Conf. on Image Analysis*, pp. 1381–1387, May 1993.
- [4] J. Y. Zheng and S. Tsuji, “Panoramic representation of scenes for route understanding,” *International Conference on Pattern Recognition*, pp. 161–167, June 1990.
- [5] S. K. Nayar, “Omnidirectional video camera,” *Proceedings of Computer Vision and Pattern Recognition*, 1997.
- [6] 尾上良雄, 山澤一誠, 横矢直和, 竹村治雄, “全方位画像からの視線追従型実時間画像生成によるテレプレゼンス,” *電子情報通信学会技術報告*, PRMU97-20, pp. 151–157, May 1997.
- [7] 山澤一誠, 八木康史, 谷内田正彦, “移動ロボットナビゲーションのための全方位視覚系,” *電子情報通信学会論文誌*, vol. J79-D-II, no. 5, pp. 698–707, May 1996.
- [8] S. Nayar, “Catadioptric omnidirectional video camera,” *Computer Vision and Pattern Recognition*, pp. 482–488, June 1997.
- [9] M.-C. Hong, T. Stathaki and A. K. Katsaggelos, “Iterative regularized image restoration using local constraints,” *Proc. IEEE Workshop on Nonlinear Signal and Image Processing*, vol. 1, Mackinac Island, MI, pp. 145–148, Sept. 1997.
- [10] W. T. Freeman and E. C. Pasztor, “Learning low-level vision,” *IEEE International Conf. on Image Processing*, oct 1999.
- [11] P. Cheeseman, B. Kanefsky, R. Hanson and J. Stutz, “Super-resolved surface reconstruction from multiple images,” *Technical Report FIA-93-02*, NASA Ames Research Center, Artificial Intelligence Branch, Feb. 1993.
- [12] V. N. Smelyanskiy, P. Cheeseman, D. A. Maluf and R. D. Morris, “Bayesian super-resolved surface reconstruction from images,” *Computer Vision and Pattern Recognition*, vol. 2, pp. 375–382, June 2000.
- [13] B. C. S. Tom, K. T. Lay and A. K. Katsaggelos, “Multi-channel image identification and restoration using the expectation-maximization algorithm,” *Optical Engineering*, vol. 35, no. 1, pp. 241–254, Jan. 1996.

- [14] T. Igarashi, T. Komatsu, K. Aizawa and T. Saito, "Very high resolution imaging scheme with multiple different-aperture cameras," *Signal Processing: Image Communication*, vol. 5, pp. 511–526, 1993.
- [15] Y. Nakazawa, T. Saito, T. Komatsu, T. Sekimori and K. Aizawa, "Two approaches for image-processing based high resolution image acquisition," *IEEE International Conf. on Image Processing*, vol. 3, Nov. 1994.
- [16] M.-C. Chiang and T. E. Boult, "Efficient image warping and super-resolution," *Proceedings of the Third IEEE Workshop on Applications of Computer Vision*, pp. 56–61, Dec. 1996.
- [17] M.-C. Chiang and T. E. Boult, "Local blur estimation and super-resolution," *CVPR*, 1997.
- [18] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *Transactions on Image Processing*, vol. 5, no. 6, pp. 996–1011, June 1996.
- [19] M. Elad and A. Feuer, "Super-resolution reconstruction of continuous image sequence," *IEEE International Conf. on Image Processing*, Oct. 1999.
- [20] B. C. Tom and A. K. Katsaggelos, "Resolution enhancement of color video," *Proc. European Conf. on Signal Processing, Italy*, pp. 145–148, Sept. 1996.
- [21] M.-C. Hong, T. Stathaki and A. K. Katsaggelos, "An iterative weighted regularized algorithm for improving the resolution of video sequences," *IEEE International Conf. on Image Processing*, vol. II, Santa Barbara, CA, pp. 474–477, Oct. 1997.
- [22] S. Baker and T. Kanade, "Super-Resolution optical flow," *CMU-RI-TR-99-36 Technical Report*, 2000.
- [23] A. J. Patti, M. I. Sezan and A. M. Tekalp, "Super-resolution video reconstruction with arbitrary sampling lattices and non-zero aperture time," *Transactions on Image Processing*, vol. 6, no. 8, pp. 1064–1076, Aug. 1997.
- [24] M. Irani and S. Peleg, "Motion analysis for image enhancement: Resolution, occlusion, and transparency," *J. on Visual Communications and Image Representation*, vol. 4, no. 4, pp. 324–335, Dec. 1993.
- [25] T. S. Huang and R. Y. Tsai, "Multi-frame image restoration and registration," *Computer Vision and Image Processing*, vol. 1, pp. 317–339, 1984.
- [26] A. Zomet and S. Peleg, "Applying super-resolution to panoramic mosaics," *IEEE Workshop on Applications of Computer Vision*, Oct. 1998.
- [27] 長原一, 八木康史, 谷内田正彦, "全方位時空間画像からの超解像度化," *情報処理学会研究報告*, 00–CVIM–121, pp. 97–104, March 2000.
- [28] 山本正信, "連続ステレオ画像からの3次元情報の抽出," *電子情報通信学会論文誌 (D)*, vol. J-69, no. 11, pp. 1631–1638, Nov. 1986.

-
- [29] R. Bolles, H. Baker and D. Marimont, “Epipolar plane image analysis:an approach to determining structure from motion,” *Int.J.of Computer Vision*, vol. 1, pp. 7–55, 1987.
- [30] S. Baker and T. Kanade, “Limits on super-resolution and how to break them,” *Computer Vision and Pattern Recognition*, vol. 2, pp. 372–379, June 2000.

付録: EPI 解析

等速直線運動する移動体に、カメラを進行方向に対して鉛直横向きに設置して撮影した連続画像を、時間軸方向に積層することで、時空間ポリゴンを作ることが出来る。この、時空間ポリゴンをエピポーラ線に平行に切断したものが EPI(Epipolar Plane Image) である。EPI 座標系を (u, z) で表すものとし、実世界中の存在するオブジェクト (特徴点) の位置を (X, Y) 、カメラの移動速度を v とすると、

$$u = f \frac{Y}{X} \quad (2.17)$$

である。これを時間で微分すると、

$$\dot{x} = \frac{f}{X} \dot{Y} \quad (2.18)$$

$$= \frac{f}{X} v \quad (2.19)$$

となり、EPI 上の傾きがデプス値 (X) を表していることが分かる。なお、図中の u 軸は画像の水平方向 (epipolar 方向) を z 軸は時間方向を表している。

よって、EPI 上での特徴点の動きは図 2.19 に表される直線状の軌跡を描く。また、カメラに対するオブジェクトの奥行き値が大きく変化せず、表面のテクスチャが一様な場合には、オブジェクトは EPI 上で『帯』状の軌跡を描く。さらに、複数のオブジェクトが、カメラの進行方向に対して平行に存在しているとすると、オブジェクトの奥行きに大きな変化がないため、図 2.20 のように、それぞれのオブジェクトが作る帯は重なること無くほぼ平行な軌跡を描くことになる。

市街地の建物の形状はこの仮定が良く当てはまるため、市街地を撮影した画像から生成される EPI は、一般に図 2.20 のような画像となる。

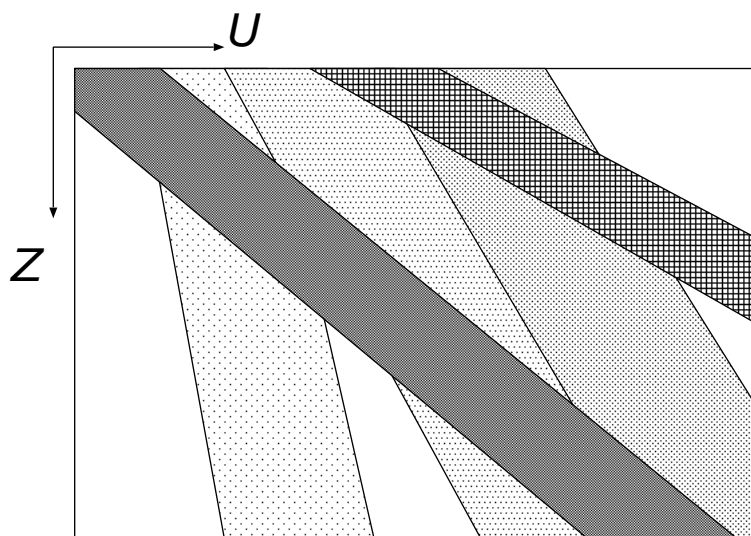


図 2.19 EPI 上での特徴点の動き

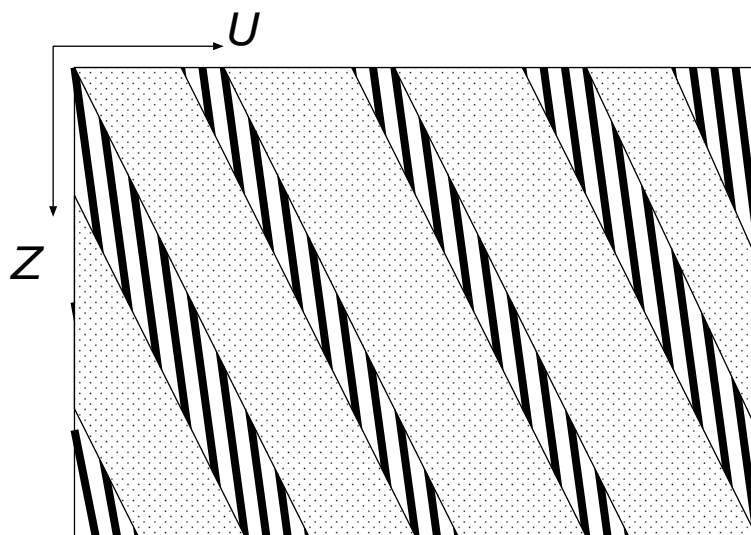


図 2.20 奥行きが等しい等間隔の面が並ぶ経路上での EPI

第 3 章

複数台カメラによる広域空間の効率的獲得

都市空間をイメージベース法で記述するためには、膨大な入力画像が必要であり、そのため、どのようにして効率よく画像を取得するかが重要である。この問題を解決するため、入力画像として全方位画像を使用する。

前章では、鏡面全方位カメラを用いる手法について述べた。一方、このような全方位画像を撮影するデバイスとして複数台のカメラを用いる手法がある。複数台のカメラを用いることで、解像度の高い全方位画像を作ることが出来る半面、複数の画像をつなぎ合わせる際に画像の歪みが大きくなることもある。

本章では、このような複数台のカメラを用いて効率的に歪みの少ない全方位データを獲得する手法について説明する。

3.1 はじめに

本章では、効率的に広域空間の情報を獲得するために複数台のカメラを組み合わせることで全方位画像を得る手法を提案する。前章において、鏡面型の全方位カメラによる全方位画像の取得方法および、鏡面型全方位カメラの問題点である低解像度の向上方法を提案した。しかし、現在生産させている CCD の密度では、現実感の高い都市のモデリングに十分なレベルの解像度を得ることが難しい。そこで、十分に高い解像度を持った複数台の通常のカメラで全方向をそれぞれ撮影し、得られた画像群を統合すれば解像度の問題を解決することが出来る。

自由に撮影した複数の画像を歪み無く統合するためには、通常、対象物体の 3 次元幾何形状が必要となる。一方、カメラの光学中心が一致するように拘束し撮影すれば、ホモグラフィ変換により、3 次元幾何形状が未知であっても歪み無く画像を統合することが可能である [1]。しかし、通常のカメラでは光学中心はカメラの内部にあるため、光学中心が一致するように撮影することは難しい。また、光学中心がなるべく近接するように撮影したとしても、カメラの位置関係を正確に知る必要があり、これを実現することは容易ではない。

以下では、まず、複数台のカメラを用いて全方位画像を生成するこれまでの研究について紹介する。続いて、複数台のカメラを用いて光学中心の一致した全方位画像を生成する手法を提案する。提案する手法は、光学中心が一致するように配置することが空間的に不可能な場合に、これを時間的に拡張することで実現する。そこで、時空間内で光学中心の一致を実現するための原理、カメラ配置および計測方法、具体的な手法について述べる。また、実際に画像を統合するためには、精密な複数台カメラの位置関係が必要であるが、計測した画像データに対して時空間画像解析を用いることで、精度良く安定して位置関係を得る手法についても提案する。

3.2 複数台カメラを用いた全方位画像

前章で述べたように、鏡面反射型全方位カメラは、周囲 360 度の撮影が 1 度で済むなど簡単である他、透視画像への変換が容易であるなどの利点がある反面、全方位を 1 枚の画像中に写しこむため、画像解像度が著しく低いという問題がある。本論文では、第 2.3 で述べたように、アルゴリズムを用いた超解像度化手法によりこの問題を軽減することに成功したが、2.4.4 で述べたように、現在の CCD では未だ、現実感の高い仮想空間の構築には限界がある。本章では、複数台のカメラで同時に撮影した画像を繋ぎ合わせることで、解像度の高い全方位画像を生成する手法を提案する。

3.2.1 複数台カメラによる全方位画像

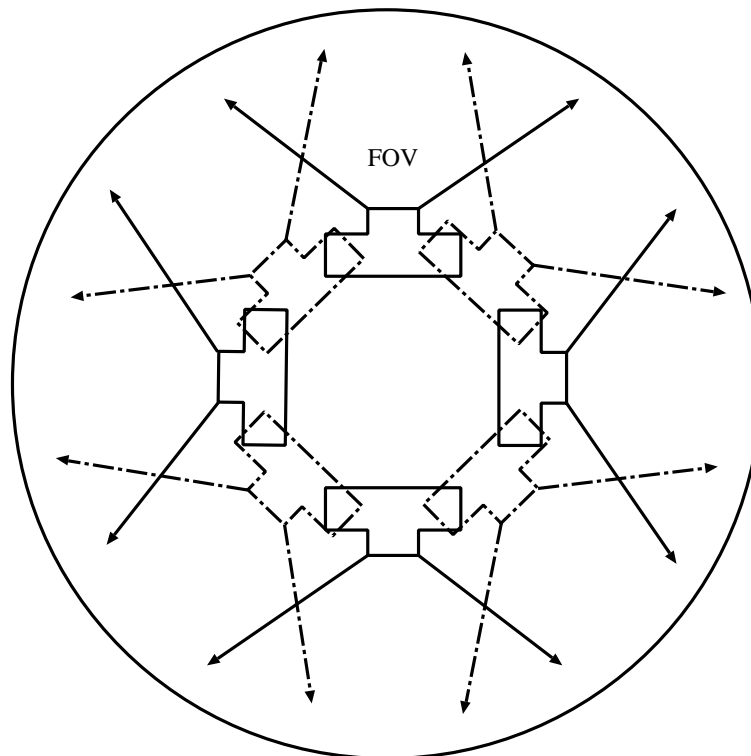


図 3.1 円筒状に配置された複数カメラ

画像精度を落とさず全方位画像を撮影するハードウェア的な手法として、図 3.1 の様に幾つかのカメラを、それらすべてのカメラの光軸が 1 点で交わる様に円筒状に並べることを考える。各々で撮影した透視変換画像を円筒座標系に投影することで、全方位のパノラマ画像を得ることができる [2]。

しかし、撮影したそれぞれの透視変換画像を円筒座標系に投影する際、図 3.2-(a) のように、円筒座標系とカメラの光学中心との間は位置が離れているため、光学中心が一致していると仮定した

計算では、カメラがピンホールカメラであると仮定した場合、

$$\theta_{diff} = \arctan\left(\frac{z}{z-d}\tan\theta\right) \quad (3.1)$$

で計算される誤差が生じる。一般に広角と言われる範囲に入る画角 65 度のビデオカメラを使用し、これを幅 640 ピクセルの画像として取り込み円筒座標に変換することを考慮し、式 (1) で計算される誤差をピクセル換算しプロットしたものが、図 3.2-(b) である。カメラと円筒座標系の光学中心の距離 d に対して、円筒座標系の光学中心と建物との距離 z は通常十分に離れていれば、誤差は数ピクセル以内に収まるものの、カメラの近くにある物体に対しては急激に誤差が拡大することが分かる。その結果、統合した画像上においては、画像の継目に物体が 2 重に描かれるなどの影響が出る (図 3.4)。物体の幾何形状を利用するなどして、これら歪みを補正することも可能ではあるが、広大な空間内にある物体の幾何形状を知ることは現実には難しい。

本論文では、映像の取得手段としてビデオ映像を用いて、光学中心を空間的に一致させることが出来ない場合において、時空間的においてこれを実現することを考える。

3.2.2 関連研究

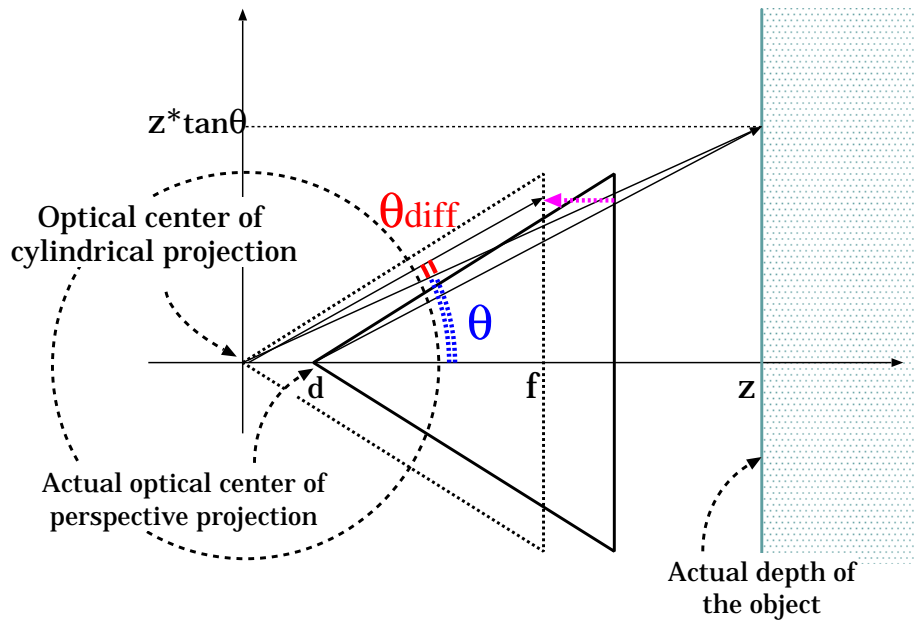
複数の画像を統合する手法はこれまで多くの手法が提案され多くの実装があるが、射影変換により複数のカメラ画像を歪み無く統合する手法 [1, 3] としてイメージモザイクという手法が知られている。これを、ビデオで撮影した映像に適用し、次々と映像を繋ぎ合わせて行くことで大きな画像に統合する研究 [4, 5] も行なわれている。この時、ビデオカメラの位置がほぼ同じであれば、歪みの目立たない統合画像を生成することが可能であるが、近傍フレーム間でのみ画像を繋ぎ合わせて行く手法であるため、撮影範囲が広がるにつれ徐々に誤差が蓄積していく問題点がある。

また、複数のカメラを用いて撮影した画像を統合する試みも多く行なわれている。SOS システム [6] のように、複数の CCD カメラにより周囲全てが撮影可能なシステムが提案されている。他にも、小型の CCD を組み合わせて周囲全てを計測するシステムも市販されている [7]。

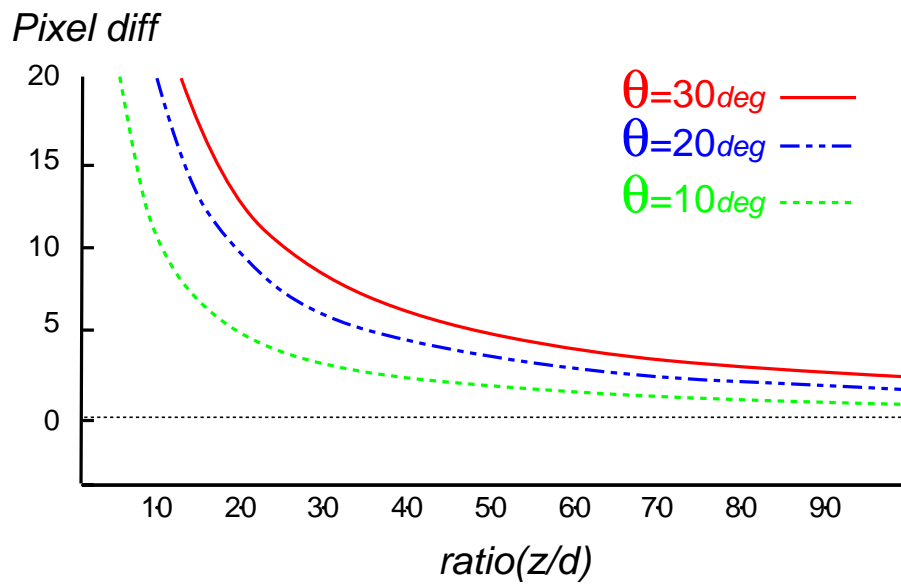
通常のビデオカメラを複数台利用して都市などの広域空間のモデリングに利用した例としては、広瀬らのシステム [8] がある。このシステム (図 3.3) においては、それぞれのカメラの光学中心が大きく離れているため、画像における歪みが大きい (図 3.4)。

3.3 複数カメラによる歪みの無い画像の生成

複数カメラを利用して撮影した画像を 1 つに統合することは、通常のカメラでは光学中心がカメラの内部にあるためこれらを一致させることが難しく、その結果、奥行き情報などの補正なしで歪みの無い画像を得ることが困難であった。本論文は、光学中心がカメラの内部にあるような通常のカメラを利用して、光学中心を一致させる装置およびその手法を提案する。提案する手法を利用す



(a)



(b)

図 3.2 (a) 光学中心が一致していると仮定して透視変換画像を円筒座標系に投影 (b) ピクセル誤差



図 3.3 画像収集車



図 3.4 光学中心の一致していない複数カメラの統合例

れば、通常のカメラを利用して大規模な環境のテクスチャ画像を、簡単に歪み無く取得することが出来る。

3.3.1 EPI を用いた光学中心の一致手法

本章では、映像の取得手段として通常ビデオカメラ複数台を車に設置し撮影することを前提としている。通常ビデオカメラを利用したのでは、光学中心を空間的に一致させることが出来ない。そこで、時空間的においてこれを実現することを考える。特に、時空間内において効率的に安定して光学中心を一致させるために、EPI を用いた手法を提案する。

カメラの配置

ビデオカメラは、図 3.5 の如く、進行方向に対して一直線上に並ぶように配置する。カメラを向ける方向に関しては、計測したいエリアがカバー出来るようであれば特に制限はない。これは、取得された映像に対して EPI によるキャリブレーションが適用できるため、正確にカメラを配置する必要がないためである。配置に関する要件をまとめると以下のようである。

- 撮影領域全体がカメラに収まる（オーバーラップは必ずしも必要ではない）
- 図 3.5 のように、カメラ中心が移動方向と平行に直線に沿って並べられていることが望ましい（理由については次節で述べる）

本論文では、2 通りの配置により実験を行った。一つは、理想的には、カメラは直線上に並べられている必要があるが、機材の制限上、全てのカメラを一列に並べられないこともあり得るため、3 台を直線上に配列したものを、3 列設置したものである（図 3.6）。もう一つは、理想的に 9 つ全てのカメラを一列に配置したものであり、図 3.7 のような配置となっている。

光学中心一致の原理

ここでは、時空間的に光学中心を一致させる原理について述べる。まず、図 3.8 のように、時間 T_1 において、先頭にあるビデオカメラがある地点にあるとする。車は走行しているため、時間 T_2 の時には図 3.8 中央のような位置に車は進行している。この時、中央にあるビデオカメラは、時間 T_1 において先頭のカメラが位置していた場所と同じ位置にある。同様に、時間 T_3 においては、最後尾のカメラが同じ位置にある。このように、時間をずらすことで、複数のカメラであっても、あたかも同じ位置で周囲の環境を撮影したかのような画像を取得することが出来る。また、このようにして得られた画像は、撮影している位置が同じとなるため、カメラがどのような方角を向いていてもカメラの光学中心は一致している。そこで、こうした画像を用いれば、複数のカメラであっても光学中心が一致しており、その結果、奥行き値などの形状の情報なしで歪のない画像に統合することが出来る。これを実現するためには、それぞれのカメラ n が同じ地点を通った時間 T_n が必要となるが、以下にその手法を述べる。

時間 T_n の取得方法

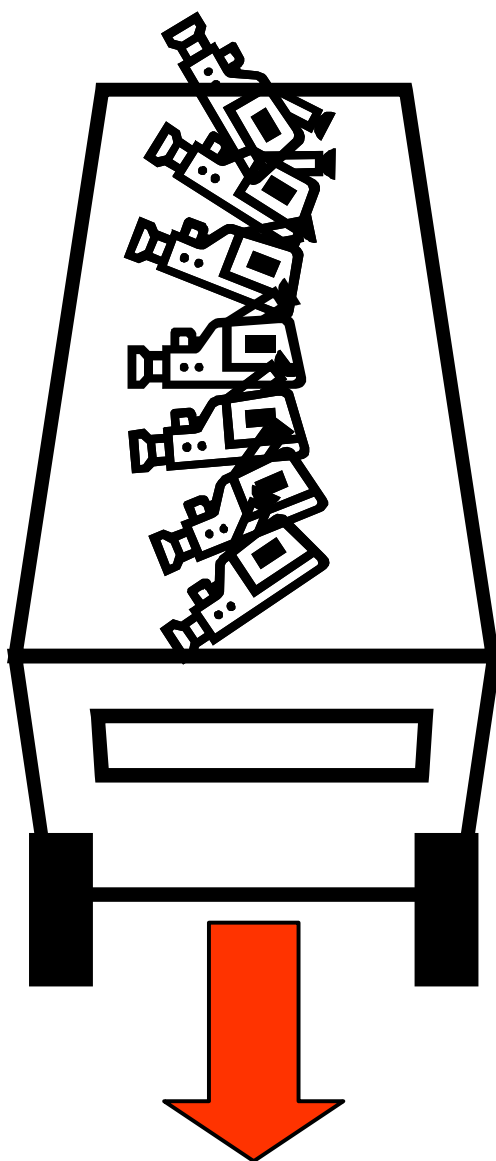


図 3.5 光学中心を一致させるためのカメラ配置



図 3.6 実際のカメラ配置例 (1)



図 3.7 実際のカメラ配置例 (2)

GPS やジャイロセンサなどを用いて直接 T_n を取得することが出来る．しかし、精度の良い位置データが得られるGPS やジャイロセンサは高価であり、また、高層ビル群の間などGPS データが得られないこともある．そこで、これらセンサを用いない手法について述べる．

等速直線運動をしているカメラで進行方向に対して垂直な方向を撮影した映像からは、EPI [9, 10] を作ることが出来る (EPI については、第 2 章付録参照)．この EPI 画像上では、特徴点が直線パターンとして検出される．そこで、複数のカメラからそれぞれの EPI 画像を生成し、その画像間の直線パターンをマッチングすれば、 T_n を取得することが出来る．これを図 3.9 を用いて説明する．まず、簡単のため画像の同期が出来ているものと仮定すると、2 台のカメラの EPI はそれぞれ、同じ画像座標系内で、図 3.9(i) となる．そこで、それぞれの EPI がなめらかに重なるように横方向にのみ動かすと図 3.9(ii) のようになる．しかし、この場合では光学中心は一致していないため、奥行きが違う建物などがあると、画像が歪む．そこで、図 3.9(iii) のように、縦、横それぞれに動かすことで、光学中心を正確に一致させることが出来る．図中、 Z で表されているのが、同期のためのパラメタである．このため本手法では、画像をあらかじめ同期させて撮影する必要がない．以上の手法では、車の等

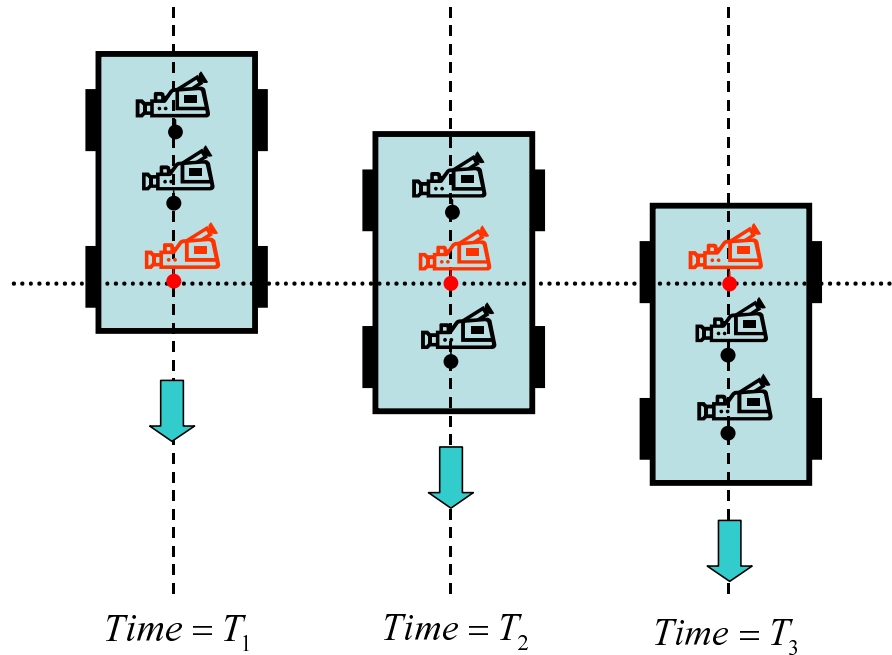


図 3.8 EPI による光学中心の一致

速直線運動を仮定しているが、等速直線運動をしていない場合であっても、車が急激に速度や進行方向を変化させることはないため、簡単な速度センサや、精度の低いGPSやジャイロセンサを用いて補正することが出来る。また、車速は画像処理により、推定することも出来る。

画像の平行化

設置したビデオカメラは全方位をカバーするようにそれぞれ異なった方向を向いている。EPIを作るためには、条件として『カメラは進行方向に対して垂直横向き』である必要がある。そのため、進行方向に対して垂直横向きとなっていないカメラで撮影した画像には平行化処理を行わなければならない。

1. 画像中の FOE の検出

通常、都市空間における建物やその他のオブジェクトは地面に対して水平・垂直な二種類の直線で構成されている。これらの直線は画像処理によって精度よく検出することができるため、画像の平行化にはこれらの直線を利用することが出来る。ただし、実際のシーン画像にはさまざまなノイズが含まれている他、平行化に必要なパラメータの推定はカメラを設置した後ただ1度だけ行なえば良いため、本論文では、水平・垂直な二種類の直線の選択は、最終的には人手で行なうものとした。ここで選択する直線は、キャニーオペレータによって検出されたものである。

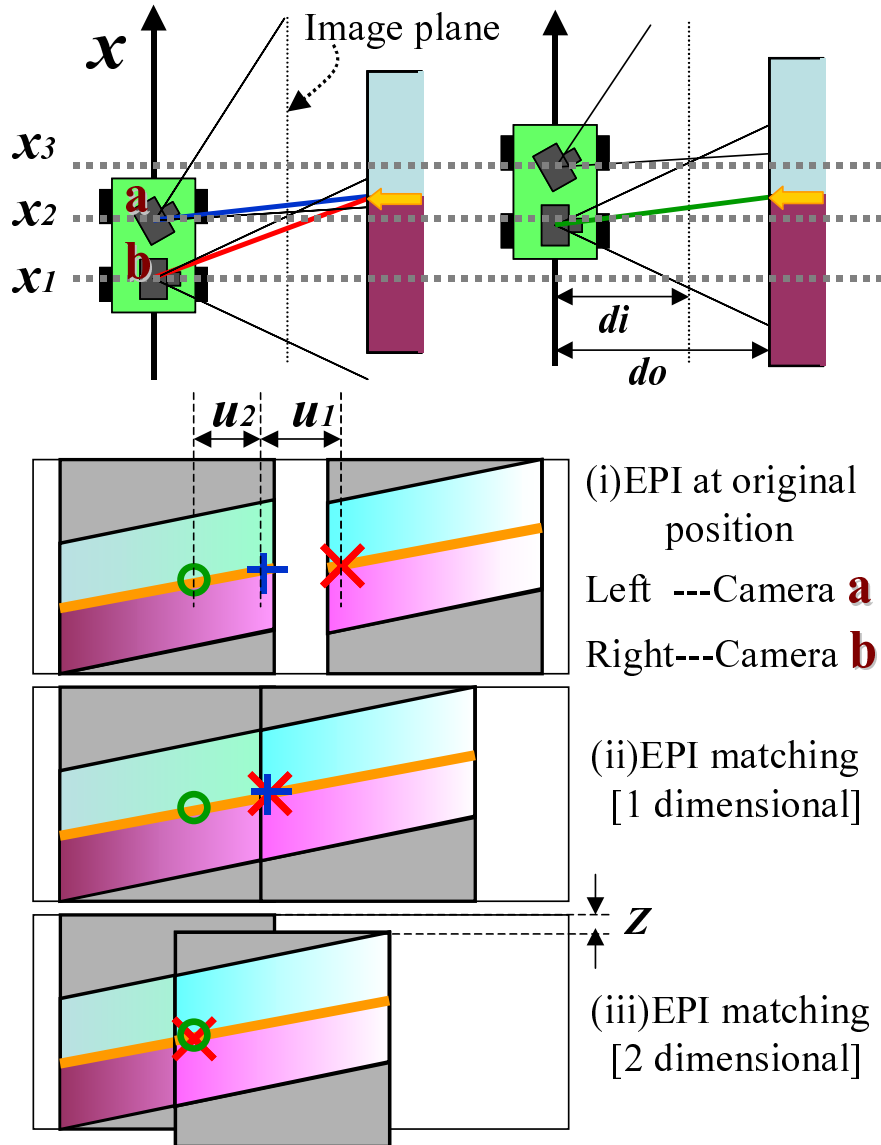


図 3.9 EPI のマッチング

直線を何本か選択した後、交点の座標を求める。この交点は Focus Of Expansion (FOE) と呼ばれる (図 3.10)。画像中の FOE 点は、水平方向・垂直方向の二つが存在する。理論的には FOE 点に向かう直線はすべて一点で交わるが、実際にはノイズや解像度の問題から一点で交わることはない。本論文では、直線を選択を手で行なっているため、大きな外れ値が無いことから、これらの交点の平均値を FOE の座標とした。

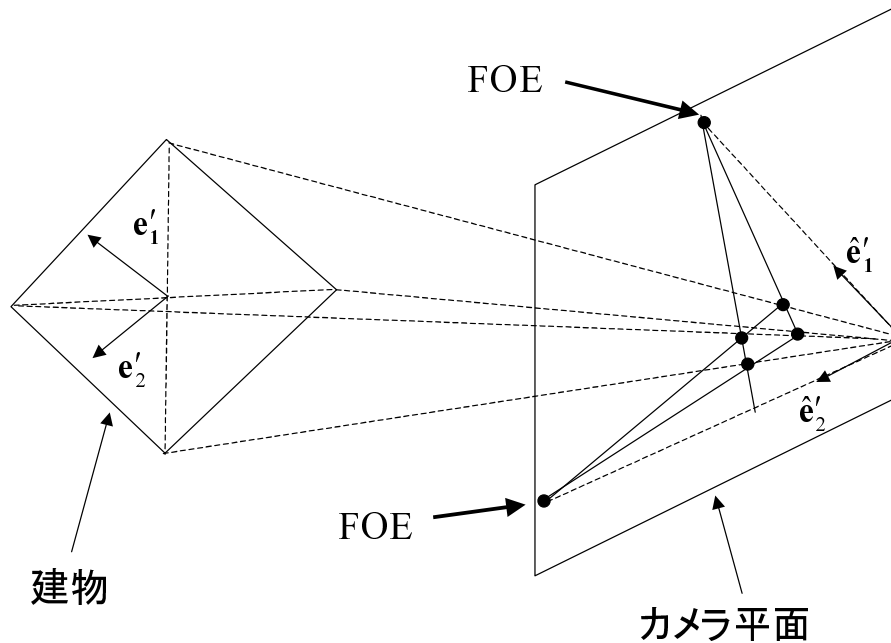


図 3.10 画像の平行化

2. 回転行列と焦点距離の推定 FOE 点を利用すると、回転行列と焦点距離は以下のように求められる。まず、 r_1 と r_2 を以下のように定義する:

$$r_i = \begin{pmatrix} u_i \\ v_i \\ f \end{pmatrix} \quad (i = 1, 2) \quad (3.2)$$

ここで (u_i, v_i) は画像の FOE 点の座標、 f はそれぞれのカメラの焦点距離 (ピクセル値) である。これらのベクトルは実世界中での水平・垂直方向にそれぞれ平行となる [11]。よって r_1 と r_2 は直行し、内積は 0 となる。したがって f は以下のように求められる。

$$f = \sqrt{-(u_1, v_1) \cdot (u_2, v_2)} \quad (3.3)$$

r_1 と r_2 を正規化したものを e_1 と e_2 とすると、 e_3 は以下のように計算される。

$$[e_3 = e_1 \times e_2] \quad (3.4)$$

よって、カメラの回転行列 R は以下ようになる。

$$R = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} \quad (3.5)$$

推定した回転行列 R および焦点距離を用いることで、射影変換により画像の平行化を実現出来る。図 3.11 および 図 3.12 に平行した画像例を示す。画像が正しく平行化出来ていることが分かる。



図 3.11 平行化の例 (1)

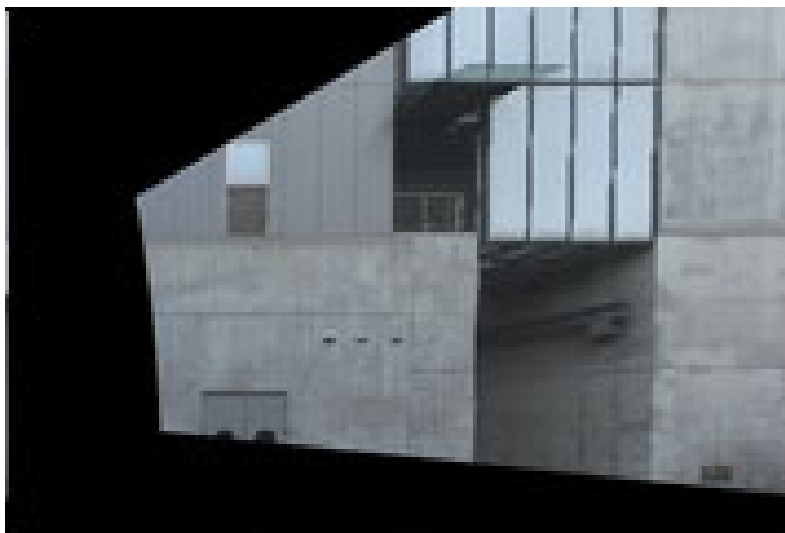
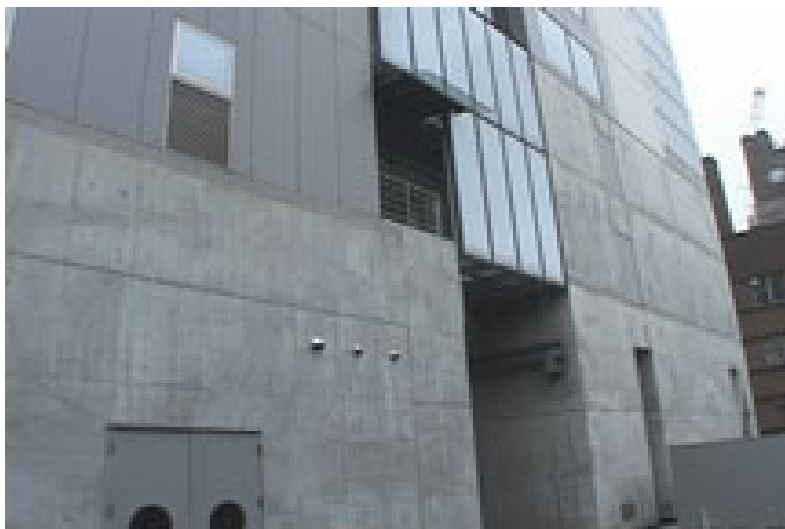


図 3.12 平行化の例 (2)

3.4 実 験

提案した手法の有効性を検討するために実際に撮影した映像を用いて実験を行った．以下にその概要と結果を述べる．

3.4.1 撮影経路

撮影経路は，2.3.1節で述べたように，東京大学生産技術研究所の付近で，直線に近い道路を選択した．撮影の際には等速運転を行なうが，実際には車両を理想的に等速直線運動させることは困難であるため，最終的に使用する画像は，撮影したデータを GPS，ジャイロ等のデータおよび画像処理により補正したものである．本実験に使用した映像の車速は約 30km/h であった．

3.4.2 対象映像

本実験では，複数台のカメラを車載し直線に近い道路に沿って撮影した建物映像を対象とした．実験に使用した車両は，図 3.6 および図 3.7 である．撮影した映像は DV 信号に変換された後，VCR に保存される．このインターレース信号をバイリニア補間により 720x480 ピクセルの画像系列として計算機に保存し，実験に使用した．

3.4.3 結果

3 台 * 複数列カメラ配置

図 3.13-(g) に画像の統合例を示す．提案した手法により，6 台のカメラ映像が歪み無く統合されていることが分かる．また，図 3.14 に別の統合例を示す．この場合には，手前側にある電柱が 2 重に写っていることが分かる．原因として，3 台 * 複数列のカメラ配置では，全てのカメラの光学中心を一致させることが出来ないため，奥行き異なる対象の場合，統合の際に歪みを除去することが出来ないためと考えられる．

9 台 1 列カメラ配置

図 3.15 に撮影した画像例を示す．図 3.16 および図 3.17 は統合処理中の画像である．光学中心の一致を考慮せずに統合した場合，図 3.16 のように，奥にある建物が正しく統合されているのに対し，手前側の電柱が歪んでいることが分かる．これに対し光学中心を考慮した場合，図 3.17 のように，奥にある建物も，手前にある電柱も歪み無く統合出来ていることが分かる．また，図 3.18 および図 3.19 に統合した結果画像例を示す．提案した手法により，9 台のカメラ映像が歪み無く統合されていることが分かる．

3.5 まとめ

本章で提案した複数のカメラを組み合わせる全方位画像を生成する方法により、広い範囲の画像を歪み無く構築することが出来る。これまでの手法と異なる点としては以下が挙げられる。

1. 小型 CCD などにより光学中心が一致するように作られた特殊なカメラではなく、通常のビデオカメラを用いていること
2. そのようなカメラでは光学中心はカメラ内部にあるため、通常では光学中心を一致させることが出来ないが、本手法では、これを実現できること
3. そのための EPI に基づいた効率的な同期手法（本文中では T_n の推定にあたる）を提案
4. 同期のために、高価で特殊な GPS，ジャイロなどのセンサを用いる必要がないこと
5. 光学中心が一致しているため、奥行き値などの形状の情報を利用することなく、大規模な環境のテクスチャ画像を、簡易に歪み無く生成することが出来ること

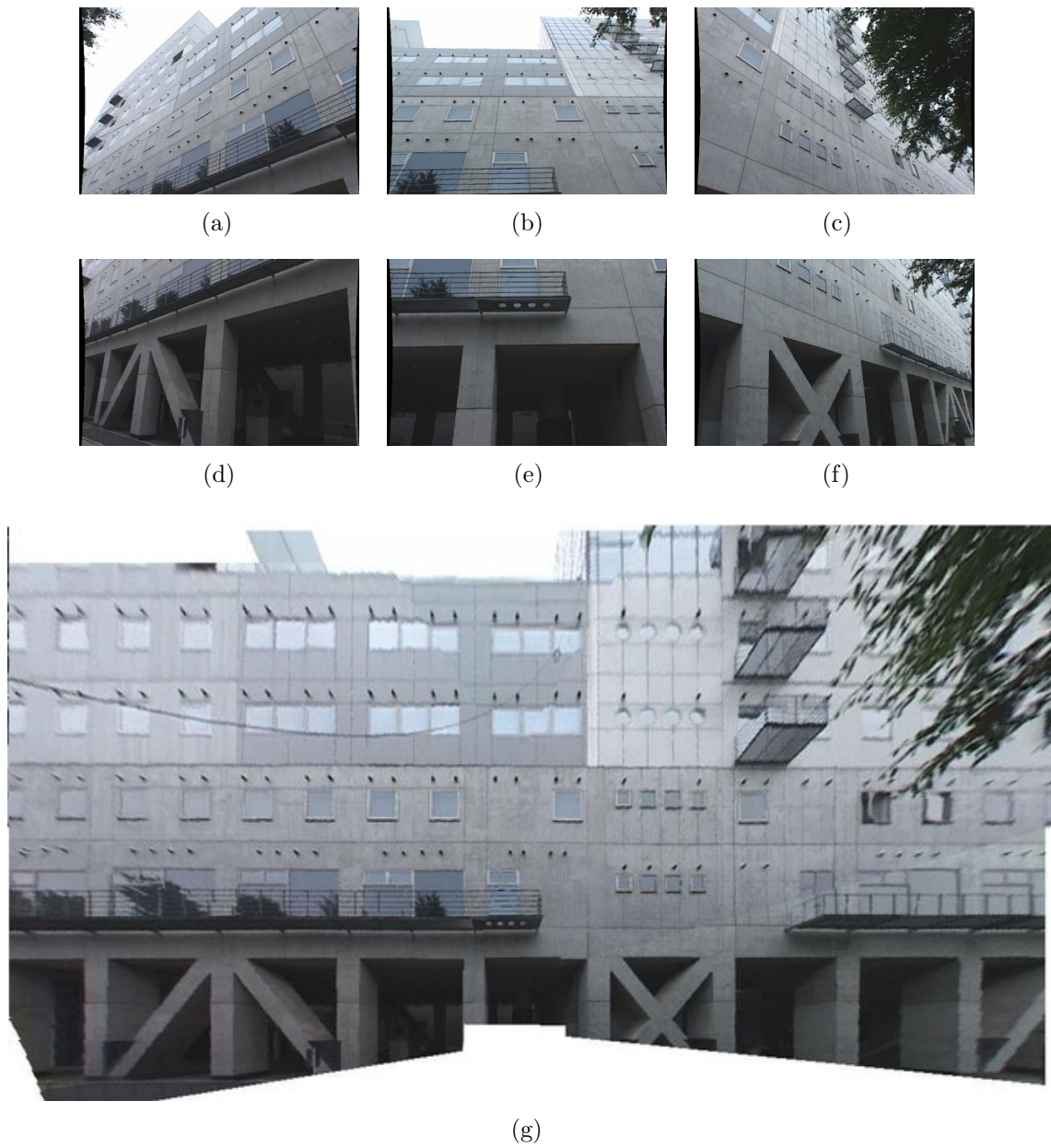


図 3.13 (a - f):撮影画像 (レンズ歪み除去後) (f) 統合画像



図 3.14 統合画像例 (2)



図 3.15 車載し撮影した画像



図 3.16 時間軸方向を考慮しない位置あわせ



図 3.17 時間軸方向を考慮した位置あわせ



図 3.18 統合例 1



図 3.19 統合例 2

【 参 考 文 献 】

- [1] O. Faugeras., “Three-Dimensional Computer Vision - A Geometric Viewpoint,” Artificial intelligence. M.I.T. Press Cambridge, MA, 1993.
- [2] M. Hirose and E. Takaaki, “Building a virtual world from the real world,” Proceedings of International Symposium on Mixed Reality, pp. 183–197, March 1999.
- [3] 清水慶行, 太田直哉, 金谷健一, “信頼性評価を備えた最適な射影変換の計算プログラム,” 情報処理学会研究報告, CVIM-111, pp. 33–40, May 1998.
- [4] R. Szeliski, “Video mosaics for virtual environment,” IEEE Comput., pp. 22–30, 1996.
- [5] 千葉直樹, 蚊野浩, 美濃導彦, 安田昌司, “画像特徴に基づくイメージモザイク,” 電子情報通信学会論文誌 (D-II), vol. J82-D-II, no. 10, pp. 1581–1589, 1999.
- [6] 山本和彦, 棚橋英樹, 桑島茂純, 丹羽義典, “実環境センシングのための全方向ステレオシステム,” 電子学会論文誌 C 電子・情報・システム部門誌, vol. 121-C, no. 5, pp. 876–881, 2001.
- [7] “View PLUS,” <http://www.viewplus.co.jp/>.
- [8] 広瀬通孝, 遠藤隆明, “Building a virtual world from the real world,” International Symposium on Mixed Reality, pp. 183–197, March 1999.
- [9] 山本正信, “連続ステレオ画像からの 3 次元情報の抽出,” 電子情報通信学会論文誌 (D), vol. J-69, no. 11, pp. 1631–1638, Nov. 1986.
- [10] R. Bolles, H. Baker and D. Marimont, “Epipolar plane image analysis:an approach to determining structure from motion,” Int.J.of Computer Vision, vol. 1, pp. 7–55, 1987.
- [11] S. D. Ma, “A self-calibration technique for active visio systems,” IEEE Trans. RA, vol. 10, pp. 7–25, 1993.

第 4 章

画像の再構成による都市空間の生成

近年、映像技術や、コンピュータグラフィックス (CG) の発展に伴ない、現実感の高い仮想空間を生成することが可能となっている。しかしながら、これまで主流であった幾何形状に基づいたポリゴンベースによるレンダリングでは、高性能なハードウェアを使用したとしても大規模で複雑な環境を構築することには限界のあることが認識されつつある。

一方、大量の画像を蓄積し、これらの実画像を様々に加工・処理し、再構築することで撮影点以外での視点映像を作り出すイメージベースと呼ばれる手法が注目されている。この手法を用いればジオメトリベースでは困難であった CG 画像を再現出来る可能性がある。

これまで、イメージベースの研究対象としては、比較的小さい対象が主であった。また、広域を対象とした研究であっても広範囲に自由視点の画像再構築を実現した例はない。本章では、イメージベースによる手法を用いて、より自由に、広範囲にわたり、都市空間という広域環境をリアリティを持ってレンダリングする手法を提案する。

4.1 はじめに

近年、映像技術の発展に伴って、高品質な映像を一般的に見ることが出来るようになってきた。このため、仮想空間を構築する際に要求される映像の質や技術は非常に高いものとなってきている。コンピュータグラフィックス (CG) は、現実感の高い仮想空間を視覚的に生成するための一般的な手法であり、従来から幅広い分野で研究がなされてきたが、近年では、バーチャルリアリティや GIS, ITS などにおける技術のベースとしても盛んに研究が行われている。

最近の CG では、計算機の能力向上やアルゴリズムの進歩により、大規模かつ精密な三次元幾何形状モデルを扱うことで、相当にリアリティの高い精細な映像の生成も可能となっている。しかしながら、このような幾何形状に基づいたモデルベースによるレンダリングでは、高性能なハードウェアを使用したとしても大規模な対象の場合には限界のあることが認識されてきている。

一方、大量の画像を蓄積し、これらの実写画像を様々に加工・処理し、再構築することで撮影点以外の視点での映像を作りだすイメージベースと呼ばれる手法が注目されている。この手法を用いればジオメトリベースでは解決できなかった画像を再現出来る可能性がある。

以上のようなコンピュータグラフィックス (CG) の研究に関する流れの中で、近年、都市のような大規模な仮想空間を構築しようという試みがなされている。しかし、これまでの研究のほとんどは、地図などモデルを用いたジオメトリベースによるものであり、再現される仮想空間は現実感の乏しいものであった。最近では、これに実テクスチャを貼るなどして、よりリアリティを高める研究も行なわれているが、十分に現実感のあるレベルには達していない。

これに対して、イメージベースを用いた仮想空間の構築が古くから行なわれている。例えば、パノラマ画像などは視点位置が限定されるが、イメージベースの仮想空間構築の一例と言える。イメージベースであるため、ジオメトリベースでは再現できないような特殊な光源効果や幾何形状を再現することができる。

本章では、このイメージベースによる仮想環境構築をさらに進めて、光線空間という概念を用いることで、より自由に、広範囲にわたり、都市空間という広域環境をリアリティを持ってレンダリングすることを試みる。

4.2 関連研究

近年、イメージベースド・レンダリング (IBR) と呼ばれるアプローチが注目されている。すなわち、モデルを推定せずに、物体やシーンの見え方そのものを入力画像として記録しておき、これらを再合成することで新しい画像を生成するという手法である。この手法では対象物体の形状や反射特性を考慮する必要がないため、先に指摘したモデル化の困難な物体に対しても適用することができる。

IBR の代表的な研究例

1. Plenoptic Function [1]…空間中に存在する光線は 7 次元で全て表現することが出来る [1]。すなわち、視点の座標 (x, y, z) 、光線の向き (θ_x, θ_y) 、波長 λ および時間 t である。この 7 次元のデータを用いることで空間中の全ての光線を記述することが可能であるとされた。
2. 光線空間法 [2]…実際に、空間を光線で記述することで任意視点の画像を効率的に再構築する手法として提案された。空間を伝搬する光線は減衰・変移しないものとするので、7 次元の光線情報を 4 次元に減らし、取得する画像を大幅に減らした点が特徴である。ここでいう 4 次元とは先程の 7 次元から z 座標、波長 λ および時間 t を差し引いたものである。画像を再構築する際には逆に光線を任意平面上にプロットすれば良い。
3. Lumigraph [3]…先に述べた光線空間法が平面 (x, y) 及び光線の向き (θ_x, θ_y) で 4 次元空間を表現したのに対して、この論文においては、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。ここでは、このような 4 次元空間を Lumigraph と呼んでいる。この論文では、データ圧縮のために物体の 3 次元形状情報を利用している点が注目される。
4. Light Field Rendering [4]…この研究においても光線を Lumigraph 同様 (u, v) 、 (s, t) 平面を用いた 4 次元で光線空間を表現している。この論文では、「近い所を通る光線は連続的に変化するため相関が高い」という特徴を利用して光線空間の圧縮を試みている。

現在のイメージベースト法の研究動向

このように IBR においてはモデルを利用しないため、形状や表面特性が複雑でモデル化が困難な物体の場合でも、光線を十分密に記録することで再現することが可能である。しかし一方で、データ量は、物体の形状と反射パラメータなどに限定される通常のポリゴンベースの手法に対して、IBR では前述のように光線空間内の光線の分布を全て記録するために非常に大きい。そこで、最近では、3 次元形状を IBR に利用することで、画像データを大幅に削減する研究が試みられている。

論文 [5] では、イメージベースとモデルベースとの関係はデータの大きさと表現する画像品質とトレードオフの関係にあるとした分析を行なっている。また、論文 [6] では、シミュレーションによる定量的な分析により、画像品質を落さない最低必要なサンプリング枚数の推定が試みられている。このように、現在のイメージベースの研究はデータの圧縮と密接に関わって進められている。

都市などの広域空間を対象とした研究

このようなイメージベースの研究に関する流れの中で、近年、都市のような大規模な仮想空間を構築しようという試みがなされてきている。これは、一つにはジオメトリベースによる限界から、必然的にイメージベースによる試みが多く行なわれるようになってきたためとも考えられる。広瀬らは車載型撮影システムにより、実際に都市空間を撮影した画像を元に仮想空間の生成を試みている [7]。この手法はイメージベースの考え方に基づいているが、画像生成に関しては光線空間などの複雑な手法は用いず、全ての経路における全周画像データをディスクに保存しておくことで、レンダリングの際には必要な画像を選択し視線方向の部分を再提示するというものである。石黒らも、映像の取得に全方位カメラを用いて同様のシステムを構築している [8] が、広瀬らと同様に光線の再構築などは行っていない。また、Teller らも、パノラマ画像にジオメトリを加えたユニークな街のモデリングを行なっているが、光線空間を積極的に利用したものではない [9, 10]。また、光線空間を利用した手法として、山口らが、本章で紹介する手法とほぼ同じ手法による画像生成手法を提案している [11]。しかし、本論文が、任意視点からの単一画像を生成しようとしているのに対し、山口らは経路から眼間距離だけ離れた対の画像を生成することで両眼ステレオ視を実現しようとする等目的が異なる。

また、本論文で提案する手法は、スリットを用いている点で論文 [12] と類似した部分もあるが、撮影経路が直線であることや広域環境を対象としているなどの違いがある。

4.3 画像のサンプリング

4.3.1 実験車両

実際の画像取得の際には、図 4.1 の車を使用する。この車には全方位画像が取得可能なカメラの他に、differential GPS、ジャイロセンサー、スピードセンサー等が搭載されている。また光線空間法にあるように各光線を記述する必要があるため、車載されている D-GPS などを使って各全方位画像の撮影した地点を画像と共に記録できる。

4.3.2 撮影経路

撮影経路は、2.3.1 節で定義したように、既知の軌道上で全方位画像をサンプルする場合を議論する。また、議論を簡単にするため直線道路を走行しながら都市の全方位画像を撮影したと仮定している。本論文では、GPS やジャイロセンサ等の設置された車両を用いて画像を計測しているため、一般の軌跡であっても、各々の撮像位置を記録したデータを使って補正することで同じ議論が展開できる。



図 4.1 実世界での画像撮影用の車

4.4 任意視点の生成

4.4.1 任意視点の生成アルゴリズム

第 2 章および第 3 章で得られた全方位画像列を用いて任意視点からの任意方向の画像を生成する。図 4.2 の様に、 Y 軸上を車で走りながら全方位画像を撮影し、点 P からの視界 (仮想的には図の実線) をつくることを考える。光線空間の考え方によれば、点 P まわりの R_s から R_e までの光線が必要となる。ここで図の θ_i について考える。 Y 軸上の任意の点で全方位画像を撮影しているので点 C_1 においても全方位画像が撮影されており、その全方位画像には点 C_1 まわりの全方向の光線が含まれていると考えられる。すなわち点 C_1 における全方位画像には図の θ_i なる光線が含まれていると考えられるので、この光線に関しては図 4.2 右側にある様に、点 C_1 で撮影した全方位画像から θ_i に相当するスリットを切り出すことになる。このようにして撮影した全方位画像列から、各光線に対応するスリットを切り出し、それらを並べ合わせることで、点 P からの視界を作成することができる。

4.4.2 視点合成が可能な領域

2 次元上の合成可能領域

ここでは 2 次元上、すなわち地平面上において、どの程度の範囲内で新しい視界を生成することが出来るかの検討をおこなう。図 4.4 において、 xy 平面を水平面と考え、またその xy 平面上におい

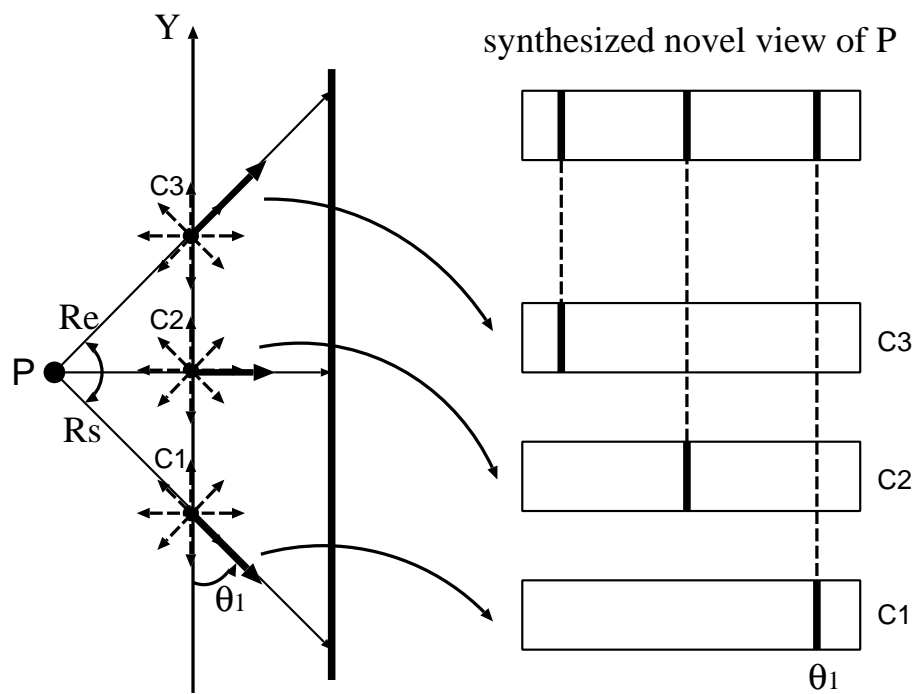


図 4.2 画像の再構築方法

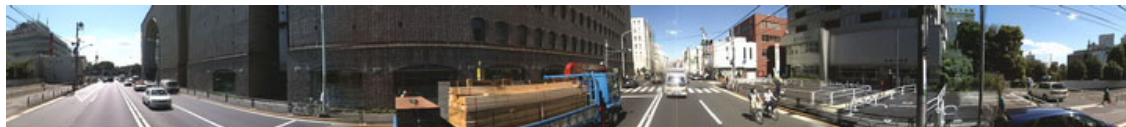


図 4.3 信濃町付近で撮影した画像を用いて作成した全方位画像

てカメラが点 C_0 から C_n までの間を全方位画像を撮影しながら直線上を移動するものとする．再構築できる視線の方向は場所により異なるが，図の斜線部に示される，視角 FOV を円周角とする 2 つの扇形の内部となる [13]．現実には，全ての視線方向の再構築が出来なければ，システムとして不十分であるため，自由に移動可能な範囲は経路近くのものに限定される．しかし，本研究の目的が都市における広域仮想空間の生成であり，画像の取得は道路上を走行することで行ない，新しく生成する画像は，主に道路幅内からの視点のみであることを考えると，この範囲でも十分に目的を満たすことができる．

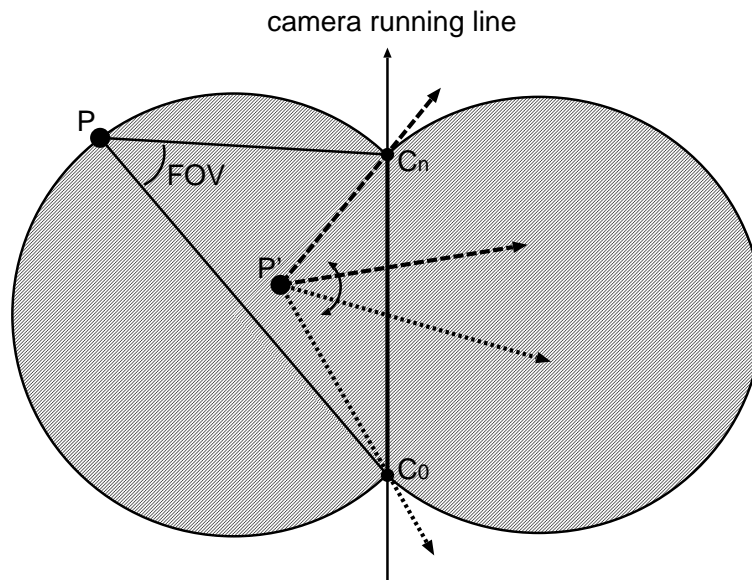


図 4.4 2次元上の合成可能範囲

進行方向の光線

上述の視界範囲の分析において問題となるのが，「singular direction」を含む視線方向，すなわち，画像を撮影した走行路に平行な光線を含んでいる場合である．

図 4.5-(b) に示すように，必要な光線がすべて撮影した全方位画像に含まれている場合には，画像生成時には，これらの光線を切り出し再構成するだけで良い．これに対して，singular direction を含む場合 (図 4.5-(a)) には，スリットを切り出してくる全方位画像について，点 A から前方の全方位画像列と点 B から後方の全方位画像列の 2 つの画像列から切り出してくることになる．つまりこの図 4.5-(a) では，生成する画像の右側は点 A より前方で撮影された全方位画像からのスリットで構成され，左側の画像は点 B より後方で撮影された全方位画像からのスリットで構成される．この部分の不連続性により，生成される画像において，singular direction を横切る部分では合成画像に違和感が生じる可能性がある．

また，singular direction に平行な光線は存在しないので，画像を生成する際にはこの光線を補完しなくてはならない．必要な光線に最も近い光線によって代替することで，補間することが考え

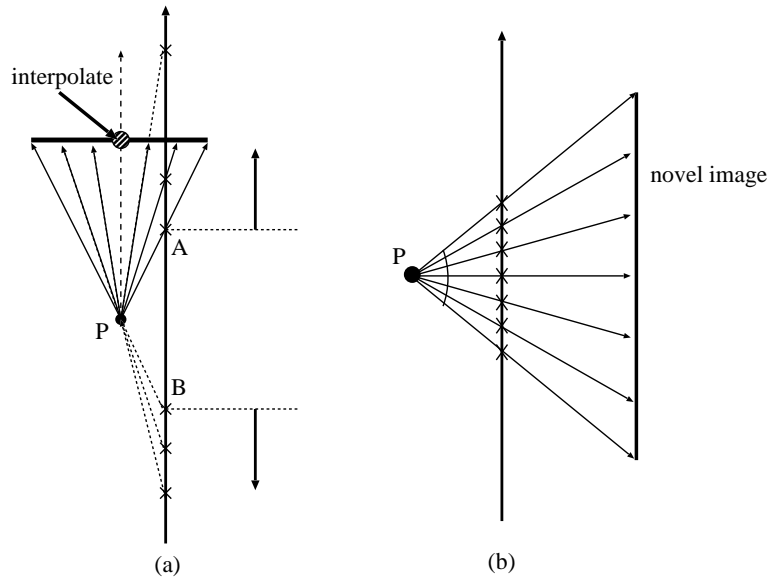


図 4.5 (a):singular direction を含む場合の画像生成 (b):singular direction を含まない場合の画像生成

られるが、境界部分においては、幾何学的な歪みやカラーの不連続等が起こる．この幾何学的な歪みは、画像の取得を車載カメラで行なっているため、離れた地点でサンプリングした画像を同じ画像内で多数使用する進行方向において、車の微小な揺れ等が蓄積されるため生じる．また、カラーの不連続は、離れた地点でサンプリングした画像を同じ画像内で多数使用することで生じる．しかし、singular direction には大きな視差がないため、このような補間による歪みが生じても、実際には合成画像に大きな違和感が生じない．

4.4.3 画像生成例

実画像を用いて画像を生成した例を示す．全方位画像は第 3 章の図 3.1 のカメラシステムを実験車両の屋根の上に車載し撮影した．

このようにして得られた全方位画像を使用して、新しい視点からの画像を生成した．図 4.6 が結果の画像である．新しい視点からの画像を生成出来ていることが分かる．しかし、画像全体としての大きな誤りは無いものの、スリットとスリットの継目が不連続であったり、直線であるはずの道路が曲線になっているなどの歪みが観察される．

4.5 生成画像における問題点

光線空間を記述するためのデータが十分であれば、蓄えられているデータから該当する光線を選択するだけのため画像上に歪みは発生しない．しかし実際には Plenoptic function にある 7 次元光



図 4.6 仮想視点からの眺め

線を全て記録しておくことは不可能であり、実際に画像を生成する場合は 4 次元, 3 次元と次元を落とすことでレンダリングを可能にしている。このように IBR による仮想空間の生成は、次元を落とすことによる歪み、すなわち光線データの欠落による歪みが生じる。特に、ジオメトリ情報を全く使わずに仮想空間を生成した場合は、図 4.6 のように画像中の歪みが顕著となることがある。

4.5.1 横方向のひずみ

再構築画像をレンダリングする際に、幅 1 pixel の縦のラインをそれぞれ 1 枚のサンプリング画像から切り出すことが出来れば横方向の歪みは存在しない。しかし、現実のシーンにおいて、特に広域環境においてこれを実現することは困難である。このため、実際のオブジェクトの位置関係によってはダブルイメージやゴーストと呼ばれる IBR 特有の歪みが生じることになる(図 4.7)。これらの歪みを、ジオメトリ情報を付加することで、回避する研究が行なわれている [14]。本論文では、都市の最構築が目的であり、都市においては横方向のカメラと対象物体との距離に差があまり無く、また、ビデオレートが十分に高いため、実際には、横方向の歪みはあまり発生しない。

4.5.2 縦方向のひずみ

また、同様に光線が無いことによる縦方向の歪みも指摘されている [11-13]。ここでは、オブジェクトの位置や幾何形状が分かっているならば、歪みを除去することが可能である。図 4.8 において、直線 Y 上を全方位画像を取得しながら走行し、点 P からの視界を生成することを考える。この図において、点 P から点 A_0 の方向の光線は全方位画像中の光線 A_0A で代用される。同様に点 P から点 B_0

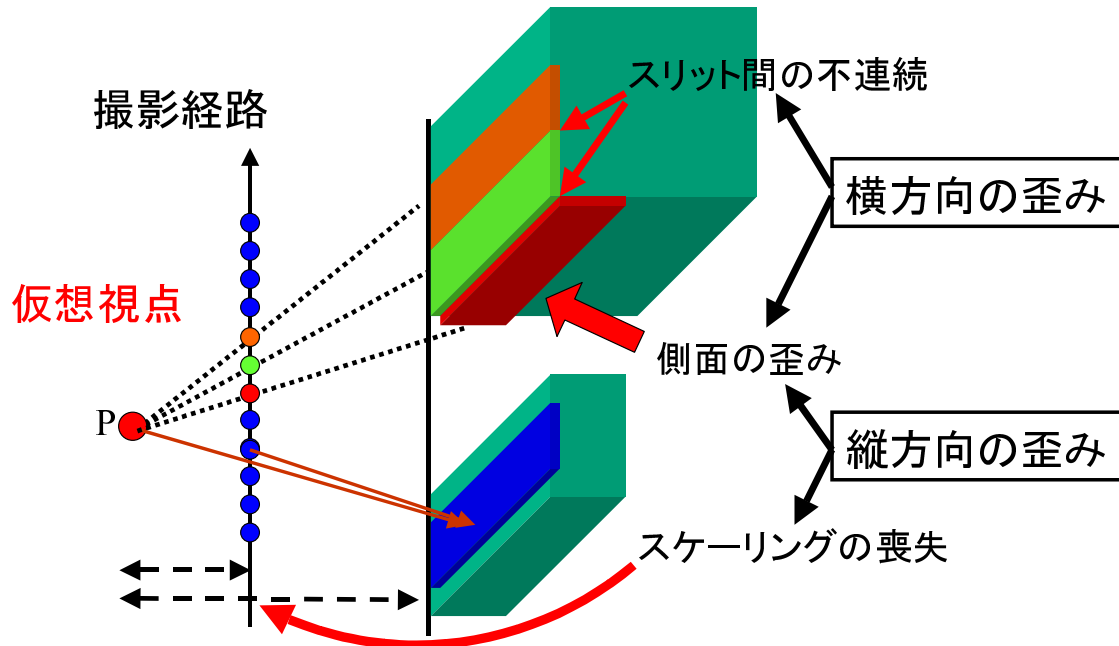


図 4.7 IBR による歪み

の方向の光線は全方位画像中の光線 B_0B で代用される。もし、建物の位置や形状の情報既知ならば、新しい視点からの光線 PA を全方位画像中の光線 A_0A で代用する際に、縦方向に正しく伸縮させることが出来れば縦方向の歪みの除去が可能である。以下では、この縦方向の歪みを除去する方法について述べる。

4.6 都市空間の特徴を利用した歪みの除去

一般的にビルなどの建造物は直方体をしていることが多い。これは特に都市において、主要道路に面している大型の建造物等に関して、よく当てはまる。加えて、そのようなある程度幅の広い道路に面しているビルは、多くの場合その道路に面している壁が、隣のビルと並んでいることが多い。そこで、このようなビル群に関しては以下のような特徴を仮定できる。

1. 建造物の多くは直方体に近い形であり、道路に面している面は平面に近い形状をしていることが多い。
2. 特に幹線道路などの広い道路に面している建造物は、その道路に面している面に関して、隣接する建造物のそれと同じ平面上にあることが多い。

都市の建造物の多くに見られる、これらの特徴を用いて、イメージ・ベースト・レンダリングによる歪みの除去を試みる。そのために以下のような都市のモデルを設定する。

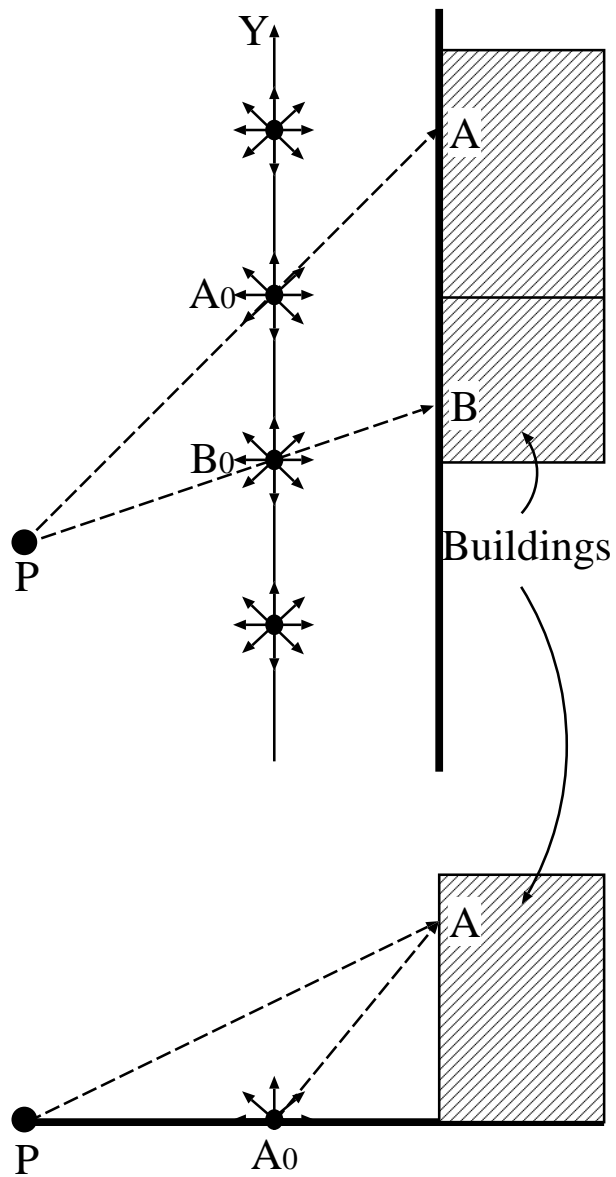


図 4.8 縦方向のひずみ

設定 道路に面している多くの建築物に対して，それらの道路に面している平面は，1 つの平面に乗っていると近似することができる．

4.6.1 カメラ・建物間の距離の推定

上記の設定は，都市の形状を一つのパラメータ（カメラ・建物間の距離）で表現したものである．そこで，全方位画像を撮影した地点から，ビルなどの建築物の道路に面している面までの距離を求める．本研究では GPS などのセンサにより，撮影位置がある程度分かっているため，映像中の特徴点をトラッキングするだけで建物の奥行き推定が可能である．今回は都市に特化した EPI 手法 [15] を用いてこの奥行き距離の推定を行なう．EPI (Epipolar Plane Image) とは，進行方向に対して垂直横向きに設置したカメラを，直進運動させながら映像を撮影し，この映像を時間軸方向に積層し，エピポーラ面と平行な平面で切断して得られる画像のことである（第 2 章付録参照）．EPI 上では，奥行き情報が傾きとして現れるため，この傾きを調べることで物体の奥行きを推定することが出来る．特に都市を撮影した映像は，建物境界などにおいて強いエッジが出るため，これを EPI 上で比較的容易に検出でき，安定した奥行き推定が可能である．また，通常 EPI 法ではカメラの速度が一定であることが条件となっているが，実際の車速は一定ではないため，これを補正するために GPS と速度センサのデータを用いている．論文 [15] では，建物の微妙な凹凸を検出するためにさらに複雑な処理も行なわれているが，今回はこれらばらつきのある奥行きを一つの代表値で表すため，処理は簡単である．実際の処理では，EPI 法で得られる奥行き情報のうち，最も頻度の高かったものを代表値とした．

4.6.2 ひずみの補正方法

図 4.9 の「panoramic image capturing line」を全方位画像を撮影する走行路とする．ここでは全方位画像として円筒座標系に投影された全方位画像を考える．この条件のもとで点 P から見た視界を合成することを考える．このとき生成される画像中に全方位画像 I_n から斜線部のスリットが含まれているとする．ここではこのスリットをどのように補正するかについて具体的に説明する．

まず点 I_n から撮影した画像を点 P からの視界として用いるのでその距離に応じた補正を行う必要がある．すなわち合成される画像の横方向の各ピクセル列に対して，縦方向に次のような補正を行う．

$$prop1 = I_n E / PE \quad (4.1)$$

その後，このままでは生成された画像は円筒座標系に投影された画像になっているため，これを通常の透視変換画像にするために，次のような補正を縦方向に行う．

$$prop2 = 1 / \cos \alpha \quad (4.2)$$

ただし生成される画像と，撮影された円筒座標系に投影された全方位画像の焦点距離は等しいとする．また， α は，視線方向と PE とのなす角である．

以上の 2 つの補正を行い，縦方向のひずみ補正を行う．すなわち生成される画像の横方向の各ピクセルに対して，対応する全方位から切り出してきたスリットと推定したカメラ・建物間の距離を用いて，以下のような補正を縦方向に行うことで，縦方向のひずみを軽減する．

$$\begin{aligned}
 prop &= prop1 \times prop2 \\
 &= \frac{I_n E / PE}{\cos \alpha}
 \end{aligned}
 \tag{4.3}$$

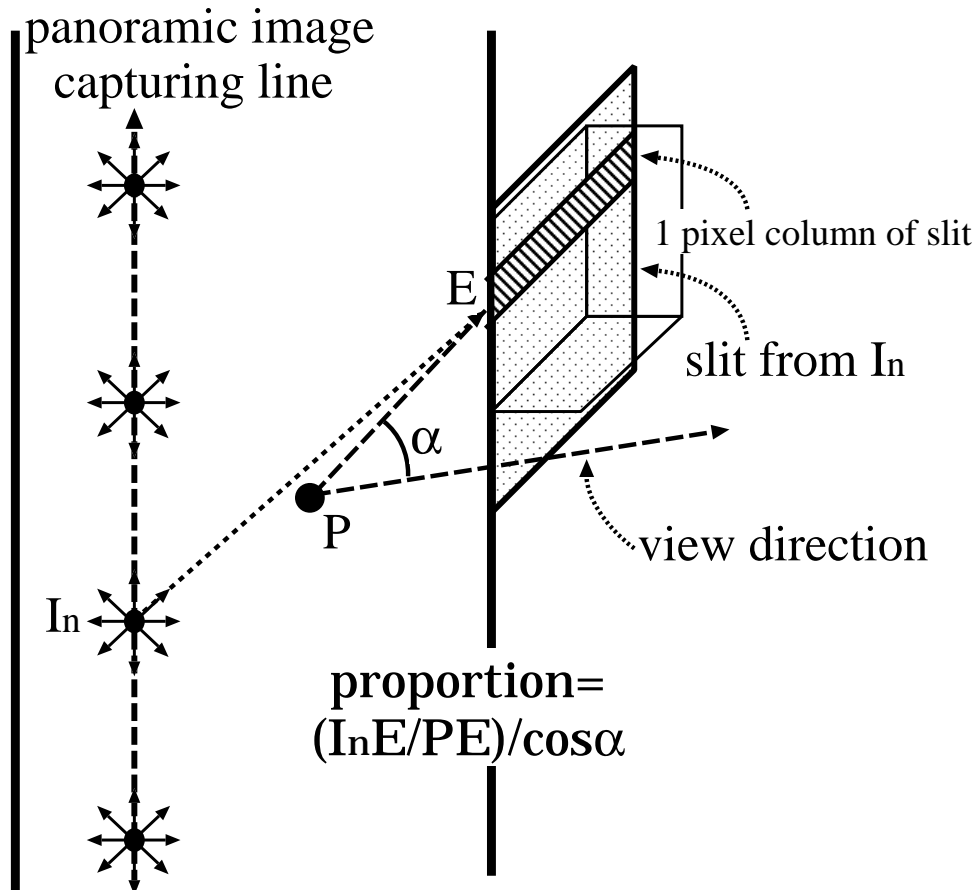


図 4.9 3次元情報による補正方法

4.7 実 験

4.7.1 実験方法

次に実画像において実験を行った結果を示す．全方位画像は3.2.1節の図3.1のように車載型撮影システムの車の屋根の上にカメラを配置するシステムで作成した．すなわち，カメラの光軸が出来るだけ1点で交わるように配置したカメラで撮影した透視変換画像を，円筒座標系に投影すること

で全方位画像を生成した。この際使用したカメラは SONY 社製の DXC-9000 であり、レンズは富士写真光機社製の S12x5BRM-38 を使用した。スペックは、水平画角が $65^{\circ}14'$ 、垂直画角が $51^{\circ}17'$ 、対角画角が $77^{\circ}19'$ である。また Trimble 社の 7400 型 GPS を用いて画像を撮影した地点を記録した。また使用した GPS の精度は $\pm 3\text{cm}$ (カタログ値) である。以上のようなシステムを用いて秒間 30 フレームで撮影しそれらを用いて全方位画像を生成した。具体的には JR 信濃町駅付近から JR 四谷駅付近までの外苑東通りを走行し、図 4.10 のような全方位画像を生成した。



図 4.10 信濃町付近で撮影した画像を用いて作成した全方位画像

4.7.2 結果

図 4.6 に対してカメラ・建物間の距離を用いて補正をおこなった画像が図 4.11 である。この図においては図 4.6 で認められたひずみが大幅に軽減されて、建物や窓のエッジ、また歩道と道路の境目が連続的につながっていることがわかる。これらの結果画像の比較により、3 次元情報による補正が、縦方向のひずみに対して大きな効果のあることがわかる。なお、本節の生成画像において使用した計算機のスペックは、CPU が UltraSPARC-II (400MHz) で、メモリが 1536MB である。また画像生成時間は、各画像とも約 0.5 秒から約 1 秒であった。また、図 4.11 下図において電柱が 2 重に見えるが、これはパノラマ画像生成時に 3.2.1 で述べた効果により出来た歪みである。

4.8 まとめ

本章では全方位画像を用いた仮想空間の新しい生成方法を提案した。この手法は、撮影した全方位画像をもとにして任意地点から任意方向を向いた視界を作ることができ、特に街などの広域空間を生成することに適している。また手法自体も、道路をただ 1 回走りながら全方位画像列を撮影し、その中から光線空間の考えに基づいて、新しい視点に対応した適切なスリットを切り出して合成するだけという、非常にシンプルなものである。そのためにリアルタイムレンダリングも可能である。

しかしながら、都市を『道路からの距離』という一つのパラメータで表現しているため、歪が正しく除去できない場合がある。特に建物の側面や、交差点などは、このようなモデルではジオメトリを正しく表現することが出来ない。

またイメージベース法は画像にもとづくため建物、道路等のシンボリックな情報が入っていない。仮想都市空間の応用としては、単なるドライビングシミュレータの画像生成だけでなく、インタラクティブな 3D ナビゲーションシステム等への応用も考えられる。こういったシステムでは画像の各部分が何らかの幾何情報としてまとめられ、その上に属性のようなものが付属している必要がある。



図 4.11 3次元情報による補正後(実画像)。上図は撮影点よりも前にカメラを設置(図12と同じ位置)。下図は撮影点とより後ろに仮想カメラを設置。

さらに、仮想空間の範囲をが広がればイメージベース法はデータが膨大なものとなり、実現が困難になることが予想される。

そこで、次章では、これらの問題点を解決するために、このイメージベース法で得られた画像に幾何情報を付加することを検討する。

【 参 考 文 献 】

- [1] E H Adelson and J Bergen: “The plenoptic function and the elements of early vision”, in *MIT Press Cambridge, MA*, pp. 3–20 (1991).
- [2] 苗村健, 原島博: “Ray-based creation of photo-realistic virtual world”, in *Virtual Reality and MultiMedia*, pp. 59–68 (1997).
- [3] S J Gortler, R Grzeszczuk, R Szeliski, and M F Cohen: “The lumigraph”, *ACM SIGGRAPH*, pp. 43–54 (1996).
- [4] Marc Levoy and P Hanrahan: “Light field rendering”, *ACM SIGGRAPH*, pp. 31–42 (1996).
- [5] Sing Bing Kang, Richard Szeliski, and P Anandan: “The geometry-image representation tradeoff for rendering”, in *IEEE International Conf. on Image Processing*, Vol. 2, pp. 13–16, Vancouver, BC, Canada (Sept. 2000).
- [6] Jinxiang Chai, Xin Tong, Shing-Chow Chan, and Heung-Yeung Shum: “Plenoptic sampling”, *ACM SIGGRAPH*, pp. 307–318 (July 2000).
- [7] 広瀬通孝, 遠藤隆明: “Building a virtual world from the real world”, in *International Symposium on Mixed Reality*, pp. 183–197 (Mar. 1999).
- [8] 加藤浩仁, 石黒浩, マシューバース: “Town digitizing -全方位カメラと gps による実時間モデリング-”, 情報処理学会研究報告, CVIM-125, pp. 111–118 (Jan. 2001).
- [9] S. Teller, M. Antone, M. Bosse, S. Coorg, M. Jethwa, and N. Master: “Calibrated, registered images of an extended urban area”, in *Computer Vision and Pattern Recognition* (2001).
- [10] Matthew Antone and Seth Teller: “Scalable, absolute position recovery for omni-directional image networks”, in *Computer Vision and Pattern Recognition* (2001).
- [11] 山口晃一郎, 山澤一誠, 竹村治雄, 横矢 直和: “全方位動画像を用いた任意方向両眼ステレオ画像の実時間生成と提示”, 電子情報通信学会技術報告, PRMU99-159 (Nov. 1999).
- [12] Heung-Yeung Shum and Li-Wei-He: “Rendering with concentric mosaics”, *ACM SIGGRAPH*, pp. 299–306 (1999).
- [13] Takuji Takahashi, Hiroshi Kawasaki, Katsushi Ikeuchi, and Masao Sakauchi: “Arbitrary view position and direction rendering for large-scale scenes”, in *Computer Vision and Pattern Recognition*, Vol. 2, pp. 296–303 (June 2000).
- [14] B. Heigl, R. Koch, M. Pollefeys, J. Denzler, and L. Van Gool: “Plenoptic modeling and rendering from image sequence taken by a hand-held camera”, in *German society for Pattern Recognition*, pp. 94–101 (Sept. 1999).

- [15] Hiroshi Kawasaki, Tomoyuki Yatabe, Katsushi Ikeuchi, and Masao Sakauchi: “Construction of a 3D city map using EPI analysis and DP matching”, in *Asian Conference on Computer Vision*, Vol. 2, pp. 1149–1155 (Jan. 2000).

第 5 章

幾何形状を利用した都市空間の生成

前章におけるイメージベース法による都市空間の生成手法は、簡単な実装により、現実感の高い画像合成が出来るという利点があるものの、都市空間の形状を「道路からの距離」という一つのパラメータで表現しているため、歪みが正しく除去できない場合も多い。特に建物の側面や、交差点などは、このようなモデルではジオメトリを正しく表現することが出来ないため、歪みが顕著となる。また、幾何形状を持たないため、再現した仮想都市にシンボリックな情報を付加することも困難である。さらに、仮想空間の範囲が広がればデータは膨大なものとなり、提案した手法を単純に広域に拡大することは現実的ではない。

そこで、本章ではイメージベース手法に幾何形状を付加することで、これらの問題点の解決を試みる。

5.1 はじめに

イメージベース法により生成された都市空間は、実装が簡単であり、モデリングの困難さが無いなどの利点があるものの、都市空間の形状を「道路からの距離」という一つのパラメータで表現しているため、多くの問題点がある。

一つには、歪みが正しく除去できない場合がしばしばあることが挙げられる。特に建物の側面や、交差点などは、このようなモデルではジオメトリを正しく表現することが出来ない。また、モデルを持たないため、再現した仮想都市にシンボリックな情報を付加することも困難である。さらに、仮想空間の範囲が広がればデータは膨大なものとなり、イメージベース法を広域に拡大することは、ディスクなどの記憶容量の制限により実現が難しくなる。

一方、最近ではイメージベース法に幾何形状を付加することで、歪みやデータサイズなどの問題点を解決する試みが行なわれている。例えば、光線を記述するのに、物体の表面の2次元平面のみに注目すれば、大幅にデータを削減することが出来る(表面表面光線空間法 [1-4])。また、少ないサンプリングでも幾何形状を利用することで歪みの少ない画像を再現する試みも行なわれている [5]。

そこで、本章ではイメージベース手法に幾何形状を付加することで、前章における問題点の解決を試みる。ただし、これまでの『イメージベース手法に幾何形状を付加する研究』においては、対象物体が比較的小さかったため、幾何形状の取得にレンジセンサなどの使用が可能であった。これに対して本論文の対象は広域な都市空間であり、そのように巨大な空間の形状をレンジセンサで計測することは現実には難しい。そこで、全方位画像から画像処理により幾何情報を取得することを試みる。

以下では、まず、全方位画像から幾何情報を取得する手法を述べ、続いて獲得した幾何形状をイメージベース法に適用する手法を提案する。最後に実験を行ない評価する。

5.2 全方位画像からの幾何情報取得

前章と同じセットアップで全方位画像列を収集し、この画像列から都市の幾何情報を得ることを考える。その際、3次元情報を獲得する方法として、ビデオ映像から得られる多くの画像列を利用することで安定して3次元情報が得られる EPI 解析 [6, 7] を利用する。本来、実世界における EPI 解析は、振動、ノイズ、光源環境の変化などの影響を受けるため、正確な3次元情報を得ることが難しい。また、撮影経路は直線の必要性がある等、拘束条件も多い。そこで本手法では、形状モデルとしてデジタル地図を使用できることを利用して、モデル(地図データ)及び実画像それぞれから EPI を生成し、EPI 同士を直接マッチングさせることで、従来の手法よりも拘束の少ないロバストな解析を実現する。また、マッチングには種々の手法があるが、本論文においては、以下に述べる理由から DP マッチングが最適であると考え利用する。

- パターンの要素が変化した時などにおける対応が柔軟であること
- 本手法においてはパターンの長さの変化は多く起こるものの、順序が変わることはほとんどない
- 高速化の検討なども多くなされている他、実装自体も簡便であること

実際に DP マッチングを適用するためには、それぞれの EPI からマッチングに適したパターンを生成してマッチングを行う。画像は2次元であるため、画像同士のマッチングは、2次元 DP マッチングとなる。具体的な2次元 DP マッチングの処理の流れについては5.2.3に述べる。

3次元情報取得の処理の基本的な流れは以下のようである。

1. 全方位カメラから EPI を生成
2. デジタル地図を用いて同じ撮影区間におけるモデル EPI を生成
3. それぞれの EPI 同士をマッチングすることで、ビデオ映像から3次元情報を獲得

以下では、まず、全方位カメラから EPI を作成する手法について述べ、続いてデジタル地図からモデルとなる EPI を作成する方法について述べる。次に、これら2つの EPI のマッチング手法について述べたあと、最後に実際のビデオ映像にこれらを適用し、3次元形状を推定した結果を述べる。

5.2.1 時空間画像解析

本論文では、計測データとしてビデオカメラで撮影した連続画像群を用いるため、分析の際に時空間画像解析を利用することが出来る。既に、第2章および第3章において、EPI を用いた時空間画像解析手法を提案しているが、本章においても EPI を用いた時空間画像解析を行なう。EPI 上では、第2章付録で紹介したように、空間の奥行き情報が『傾き』となって現れる。また、建物などのように同じ奥行きとテクスチャを持った物体は『帯』を形成する。そこで、EPI 上から『傾き』と『帯』を検出できれば、建物などの幾何形状を得ることが出来る。以下では具体的な『傾き』と『帯』を検出し幾何形状を推定する手法を述べる。

鏡面型全方位カメラの時空間画像解析

全方位カメラによる時空間画像は全方位画像を時間軸方向に積層することによって得ることが出来る (図 5.1 (a))。通常の EPI は、この時空間画像を画像上のエピポール線 (通常の画像では水平線、オム二画像では流線形) に沿った切断面として得ることが出来る。しかし、この流線の形状は回転体ミラーに固有のものであり、かつ、このような流線形に沿って時空間画像を切断することは現実的には困難を伴うことが多い。そこで、本稿ではこのオム二時空間画像を直線で切断することを検討する。

図 5.1 (b) に、半径方向に切断した例を示す。図から明らかなように半径方向にはパノラマ画像が出現する。これはオム二画像の各半径方向の切断線上の画像が通常のカメラの縦方向の切断線上の画像に対応し、これらを時系列方向につなぎ合わせればパノラマ画像になることに対応する。

縦方向に切断した例を図 5.2 に示す。ただし、時空間画像の座標系を (x, y, t) で表すものとし、進行方向をオム二画像の鉛直上方向 (y 軸上向) とする。回転体ミラーにおいては垂直な構造物はミラーの中心に向かって写り込むため、実世界中の (X_0, Y_0) にある特徴物体は、オム二画像上では、直線

$$y = \frac{Y_0}{X_0} \cdot x \quad (5.1)$$

上にある。ここで、カメラが一定速度 v で動いているとすると、 $x = k$ のときの特徴物体の y 座標は、

$$y = k \frac{Y_0}{X_0} - k \frac{v}{X_0} \cdot t \quad (5.2)$$

となり、この特徴点の時間微分は、

$$\dot{y} = -k \frac{v}{X_0} \quad (5.3)$$

である。従って、縦方向の切断面 ($x = k$ における $y - t$ 平面) においては軌跡が直線となり、またこの直線の傾きが奥行き情報 (X_0) を表していることが分かる。

複数台カメラの時空間画像解析

複数台カメラの場合は、画像を全方位画像に統合する際に既に EPI を生成しているため、直接これを利用することが出来る。時空間ポリゴンの座標系を (x, y, t) で表すものとし、進行方向を鉛直上方向 (y 軸上向) とする。この時空間ポリゴンを水平方向に切断したものが EPI である。

実世界中のオブジェクトの位置を (X, Y) 、カメラの移動速度を v とすると、

$$x = f \frac{Y}{X} \quad (5.4)$$

である。これを時間で微分すると、

$$\dot{x} = \frac{f}{X} \dot{Y} \quad (5.5)$$

$$= \frac{f}{X} v \quad (5.6)$$

となり、EPI 上の傾きがデプス値 (X) を表していることが分かる。

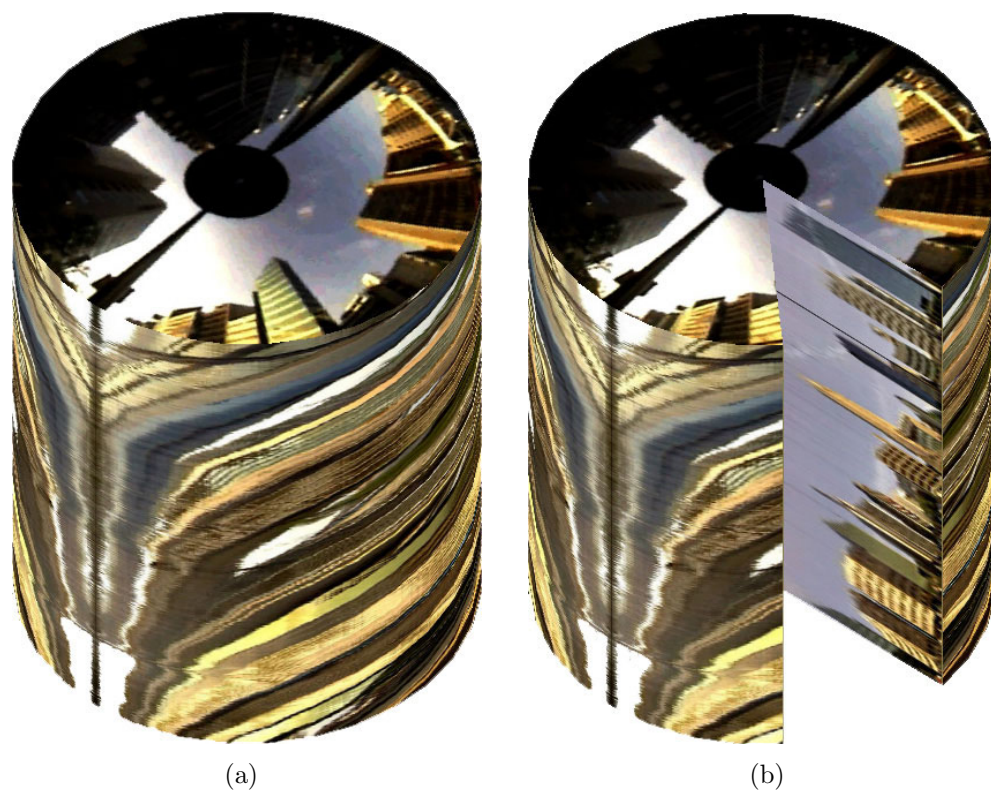


図 5.1 (a) 全方位カメラ時空間画像 (b) 半径方向切断面

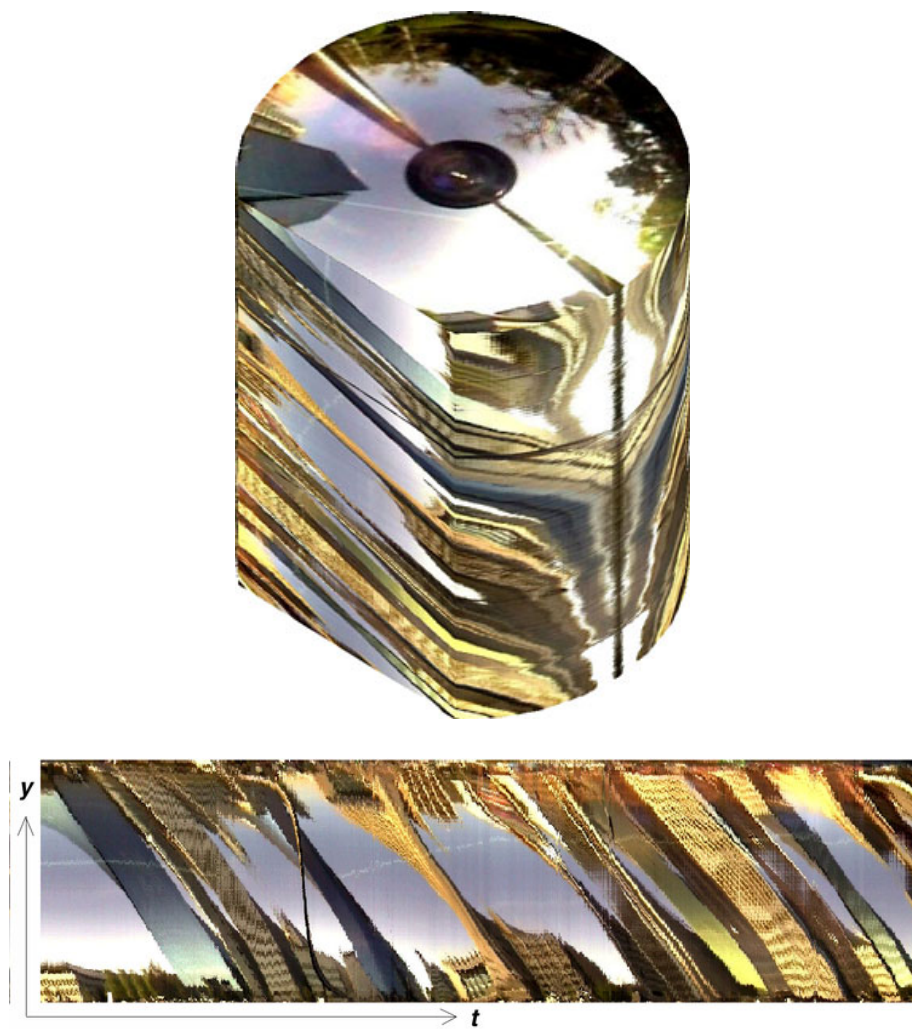


図 5.2 垂直方向切断面

5.2.2 EPI パターンの生成

本論文におけるマッチングでは DP マッチングを利用するため、EPI からエッジ成分を抽出し、これをもとにパターンを生成しなければならない。以下に全方位画像 EPI からのパターンの生成方法について述べる。

全方位画像 EPI からのパターンを生成する手順は以下の 2 つの手順からなる。

1. EPI からエッジを検出
2. 2 次元 DP マッチングで使用するプレーン数に応じて、EPI エッジ画像をスリット状に分割し、分割したスリット上にエッジが乗っていた場合にそれをパターンとする

また、手順 (1) における全方位画像 EPI からエッジを検出する処理は具体的には以下のものである。

- i. Canny アルゴリズムによりエッジを検出
- ii. ある閾値以内にあるエッジが近接している場合には同一のエッジとみなし、一線分にまとめる。

以上の処理により全方位画像 EPI から検出したエッジの例を図 5.3 に示す。

モデル EPI

同様にモデル EPI からもパターンを生成することが出来る。デジタル地図上を仮想的なカメラが等速直線をしたと仮定して自動生成したモデル EPI の例を図 5.4 に示す。

こうして生成したモデル EPI に対して、前述の手順 (2) の処理を行なうことで、マッチング用のパターンを生成することが出来る (図 5.5)。

5.2.3 2 次元 DP マッチング

本研究では画像である EPI パターン同士のマッチングを実現するためにエッジの連続性を拘束条件として 2 次元の DP マッチングを行う。

具体的には各 EPI を図 5.6 に示してあるように横方向にスリット状に切り分けることで、図 5.7 に示されるようなマッチングプレーンを作成する。そして、各マッチングプレーンにおいて通常の 1 次元 DP マッチングを行い、その結果得られるパスが隣接するプレーン間において連続するということを拘束条件として繰り返し処理を行った。

実際にマッチングをするにあたり、次のことが問題となる。

- 実画像 EPI にはノイズやテクスチャなどの影響により、必要でないエッジが多く検出される
- このため、エッジがある、ないという単純なマッチングでは、曖昧さが大きく、正解ではない解に収束することが頻繁に起こる。



図 5.3 全方位画像 EPI からエッジ検出した結果

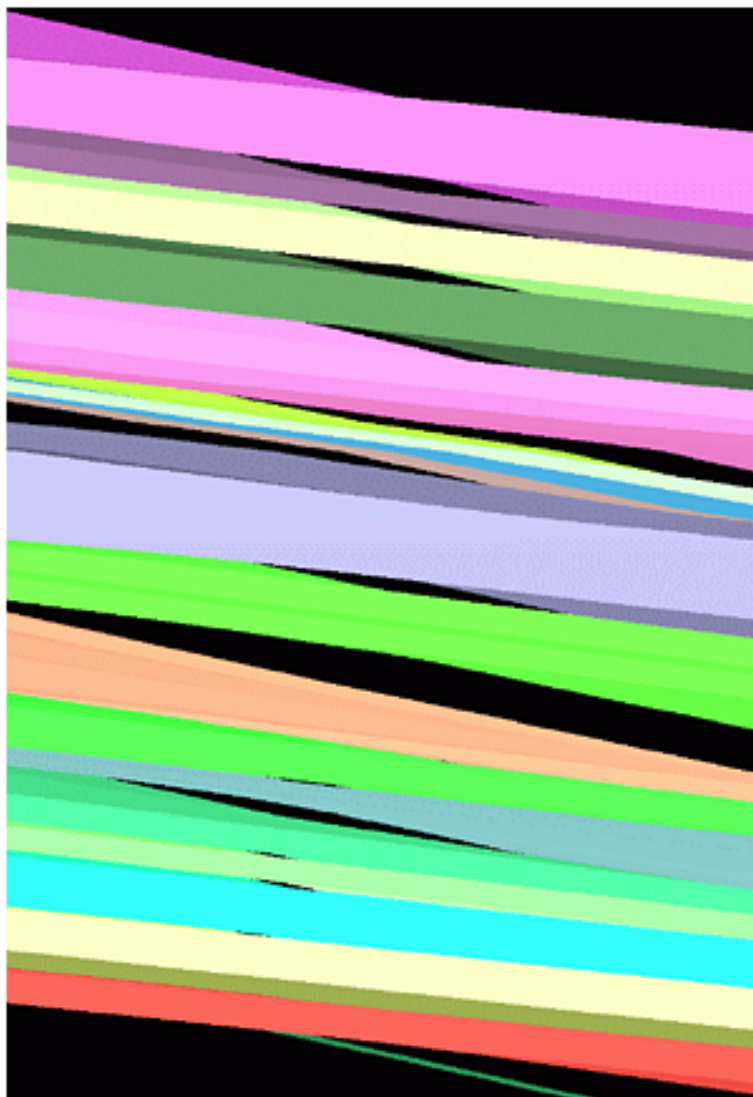


図 5.4 地図から作成したモデル EPI

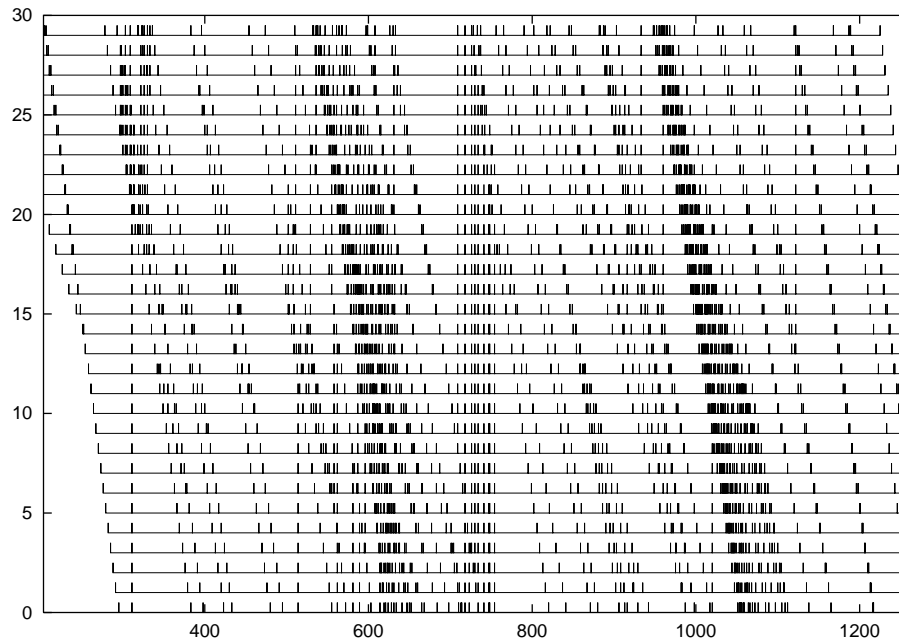


図 5.5 モデル EPI から生成したパターン

- また、地図データから作成した EPI は、地図データに存在する誤差や推定した経路が実際に走る経路とは微妙に位置が異なっていることなどの影響により、全方位カメラから得られる実際の EPI とはパターンが大きく異なることがある。

そこで、これらの問題を解決するためにマッチングに利用するパターンの種類を増やすことを検討する。本論文においてはマッチングの精度向上に高い効果が期待できると考えられる、空と建物の境界の 2 つを利用するものとした。いずれの要素も、モデル EPI からは簡単に取得することが可能である。しかし、実世界映像を用いた全方位画像 EPI からそれらのパターンを得ることは簡単ではない。特に建物の境界を得ることは困難である。そのため、全方位カメラから作成した EPI に関しては比較的容易に安定して取得可能な空のみをマッチングの要素とした。

全方位画像 EPI から空の要素を得るには、空を表す色ヒストグラムを対象画像列中から平均値として得、これを閾値とし全方位画像 EPI 上において領域拡大法によりセグメンテーションすることで容易に得ることが出来る。

5.2.4 マッチング結果

5.2.2 で生成したモデル EPI と、全方位画像 EPI のパターンを、5.2.3 で述べた DP マッチングを用いてマッチングした。マッチングにおいて必要となる入力には始点と終点の他、今回は第 4 章で求めた EPI 上の代表的な傾きの 3 つである。傾きについては連続性の拘束をより安定させるために用いている。つまり、図 5.6 を見れば明らかな通り、マッチングパスが EPI の傾き方向に連続するこ

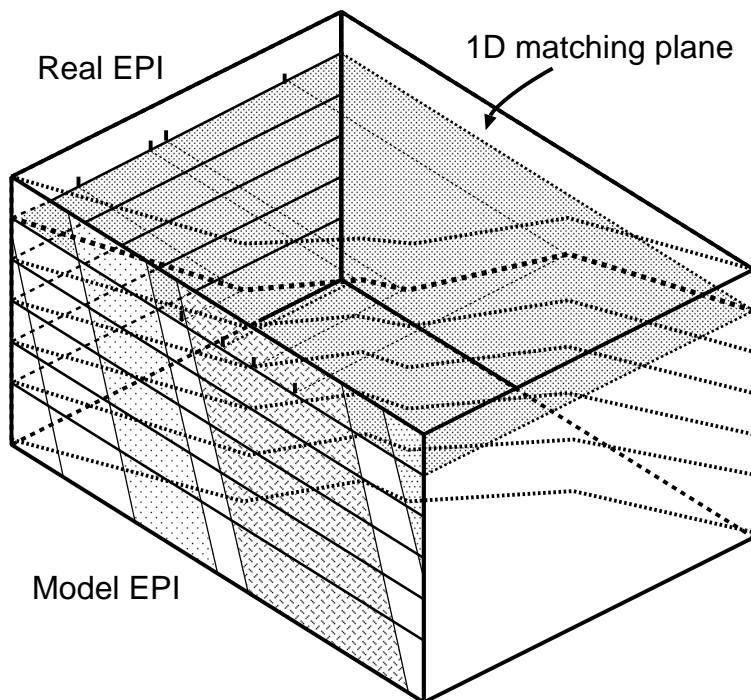


図 5.6 2次元 DP マッチングの様式図

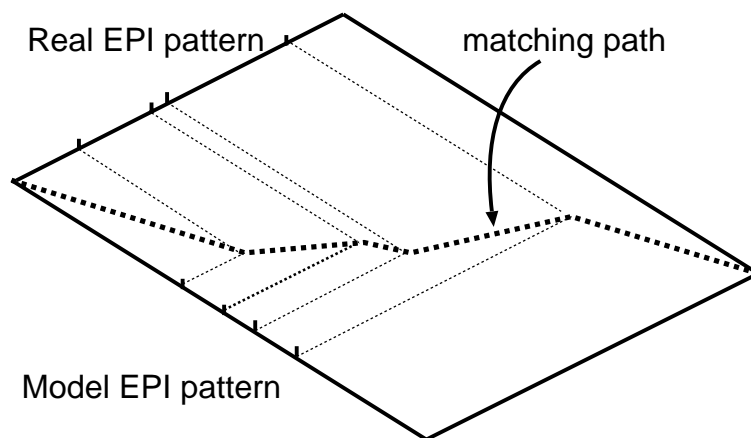


図 5.7 2次元 DP マッチング中の1次元サーチプレーン

とを利用している。

マッチングの結果を図 5.8 に示す。各線分がそれぞれのプレーン毎のマッチングの経路を表している。本実験においては、40 のプレーンに分割しているため、線分の数は合計 40 本となっている。また、線分が全体として曲面を形成しているが、これは隣り合うプレーン同士においてマッチングパスが連続するという拘束条件を用いて処理を行った結果である。

図 5.9 に例として第 20 番目のプレーンを示す。図 5.9 における縦軸は映像から生成されたエッジパノラマ画像の座標であり、横軸は地図から生成された境界パターンの座標である。ここで示される経路とはそれぞれの座標の対応点を示している。したがって、もし、両者の座標に誤差が存在しないならば経路は図中で始点と終点を結ぶ直線となる。しかし、実際にはパノラマを作成する際や地図から境界パターンを生成する際に誤差が発生するため、実際の経路は直線とはならない。図中実線で示してあるのが人手で得た正解の経路である。また破線で示しているのが本手法による DP マッチングの結果得られた経路である。2 本ある破線はそれぞれ繰り返し処理をしたもの、しなかったものである。繰り返し処理によりマッチング結果が正解のパスに近付いているのが分かる。

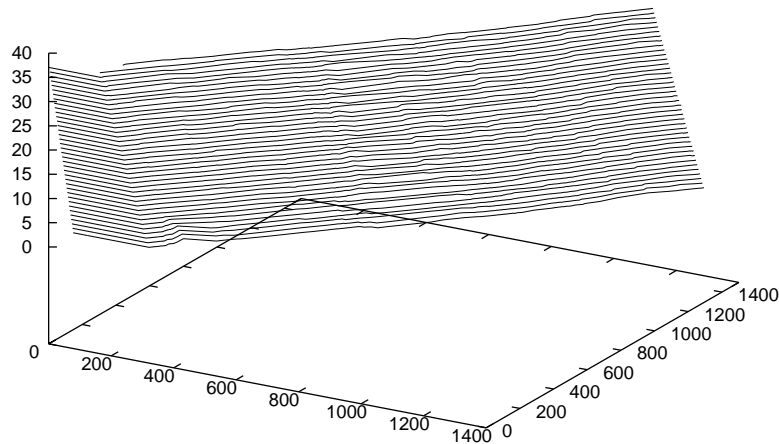


図 5.8 2次元 DP マッチングの結果

5.2.5 3次元形状推定結果

推定した3次元形状にテクスチャを貼ったものを図 5.10 ~ 図 5.13 に示す。正しく3次元情報が推定できているため、建物に正しくテクスチャを貼ることが出来ていることが分かる。また、図 5.11 を見ると、建物側面にも正しくテクスチャが貼れていることが分かる。しかし、このようにポリゴ

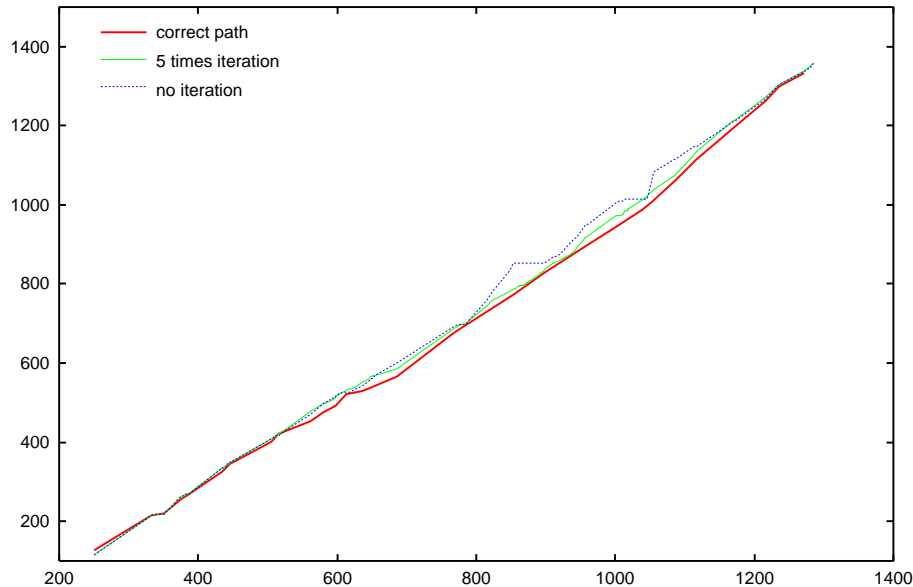


図 5.9 繰り返し処理による DP マッチングの結果の変化

ンに単にテクスチャマップしただけでは極めて現実感に乏しい。そこで、次節では第 4 章で述べたイメージベース法を用いることで現実感の高い画像の合成を行なう。

5.3 任意視点の生成方法

得られた幾何形状を用いてイメージ法により仮想 3 次元都市を構築する手法について述べる。本手法は、表面光線空間法 [2-4] に基づいている。前章で述べた光線空間法が、形状とは無関係なパラメタを用いて光線を定義していたのに対し、表面光線空間法は、物体の表面から光線が出ているものとし、物体表面上で光線空間を定義している点異なる。こうすることにより、通常的光線空間法では、ある空間を定義するために、その空間全てを横切る光線が必要であるのに対し、表面光線空間法では、その空間内にある物体の表面部分の光線のみを取得すれば済む。

5.3.1 表面光線空間データの生成

表面光線空間データを構築するためには、形状に光線を再投影すれば良い。撮影した全方位画像を効率良く形状に投影するために、時空間画像解析を用いる。時空間画像をエピポーラ線に平行に切断したものが EPI であったが、これをエピポーラ線に垂直に切断するとパノラマ画像 (Panoramic View Image) が出現する。また、切断する場所によって複数の PVI 画像が得られる (図 5.15 ~ 図 5.17)。これら複数の PVI 画像は、同じ景観をを異なった方向から見ることで得られるパノラマ画像に他ならない。そこで、この複数の PVI 画像を EPI 解析結果を用いて微小移動させることで、効率

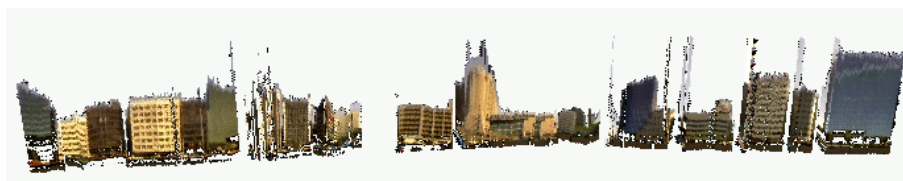


図 5.10 奥行き情報を利用した 3 次元都市

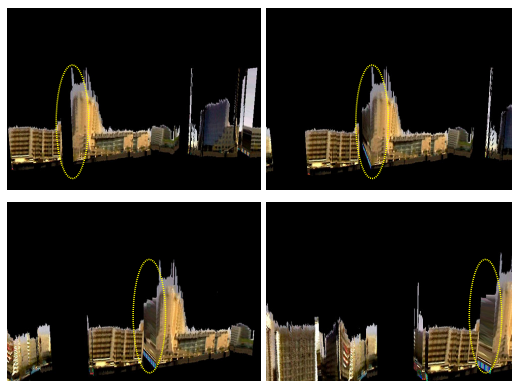


図 5.11 獲得できた側面の例

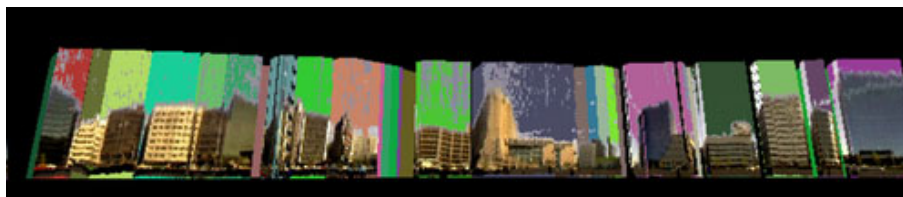


図 5.12 奥行き情報を利用した 3 次元都市 (2)

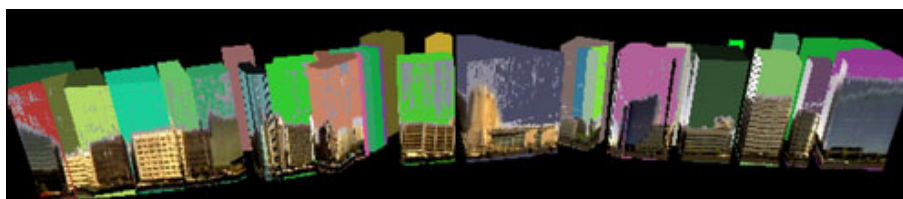


図 5.13 奥行き情報を利用した 3 次元都市 (2)(鳥瞰図)

良く表面光線空間を得ることが出来る (図 5.18) .

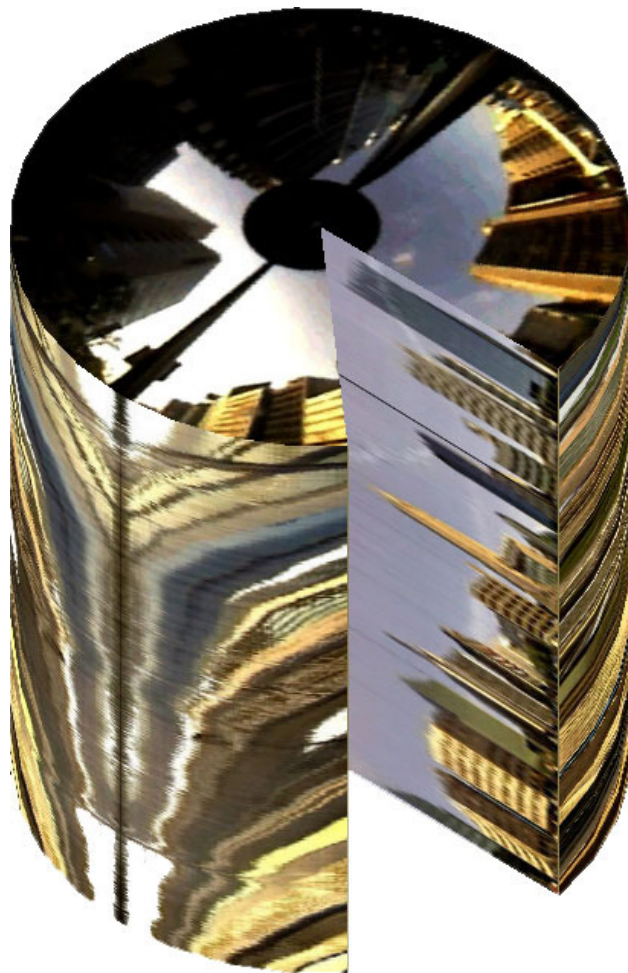


図 5.14 全方位カメラ時空間画像の半径方向切断面 (PVI)

また，個別の建物ごとに PVI を切り分けることで建物ごとの表面光線空間データデータを作成することが出来る．切り分けは，幾何形状を得る際の EPI 上での DP マッチングにより得た対応結果を，PVI に投影することで自動的に行われる．さらに，PVI を建物の側面に射影することで，側面方向の表面光線空間データも構築することが出来る．表面光線空間データの構造を図 5.19 に示す．

5.3.2 データの圧縮

得られた表面光線空間データは，物体表面の同じ場所をあらゆる方向から見た場合の見えを集めたものである．従って，見る方向によって見え方の変化の少ない物体では，表面光線空間データはほとんど変化がないテクスチャの集合となる．一方，見る方向により大きく見えが変化する物体や，



図 5.15 PVI 画像 (プレーン 5 での切断面)



図 5.16 PVI 画像 (プレーン 19 での切断面)



図 5.17 PVI 画像 (プレーン 39 での切断面)

ある視線角度の場合にだけ、急激に見えが変化するような物体もある。このように、物体によって差が有るものの、同じ場所を撮影したデータであるため、一般に高い類似性を持っている。このような性質のデータを効率良く圧縮する手法として、固有空間法がある。

本論文では、表面光線空間データを $16 * 16$ のブロックに分割し、圧縮を行なった。また、圧縮による画質劣化は 1% 以下とした。圧縮結果を図 5.20 および図 5.21 に示す。 $x - y$ 軸は、画像の $u - v$ 座標軸に相当し、 z 軸が画質劣化を 1% 以下を達成するために必要な画像の枚数を示している。図 5.20 は、入力画像枚数が 80 枚、図 5.21 は 40 枚である。本手法で推定した形状は実際の建物の概形であることや、撮影環境が野外であることなどから、圧縮率はあまり高くないが、平均すれば、約 20% 近い圧縮を実現出来ている。

5.3.3 レンダリング手法

画像を再構築する際は、第 4 章と同様、スリットを選択し張り合わせるだけで良い。ただし、表面光線空間であるため、本手法においては、建物など物体の表面に視線方向に応じたスリットを張り付けることになる (図 5.22)。こうすることで、建物の側面なども歪み無く描画することが出来る。

5.4 実 験

5.4.1 実験方法

実画像において実験を行った結果を示す。全方位画像は 3.2.1 節の図 3.1 のように車載型撮影システムの車の屋根の上にカメラを配置するシステムで作成したものと、3.2 節の図 3.5-(b) により撮影したものを使用した。ビデオカメラはいずれも、SONY 社製の DXC-9000 であり、ワイドコンバージョンレンズを用いて広角画像としている。

撮影経路は、2.3.1 節で述べたように直線に近い道路を選択し、等速で走行しながら撮影した。実

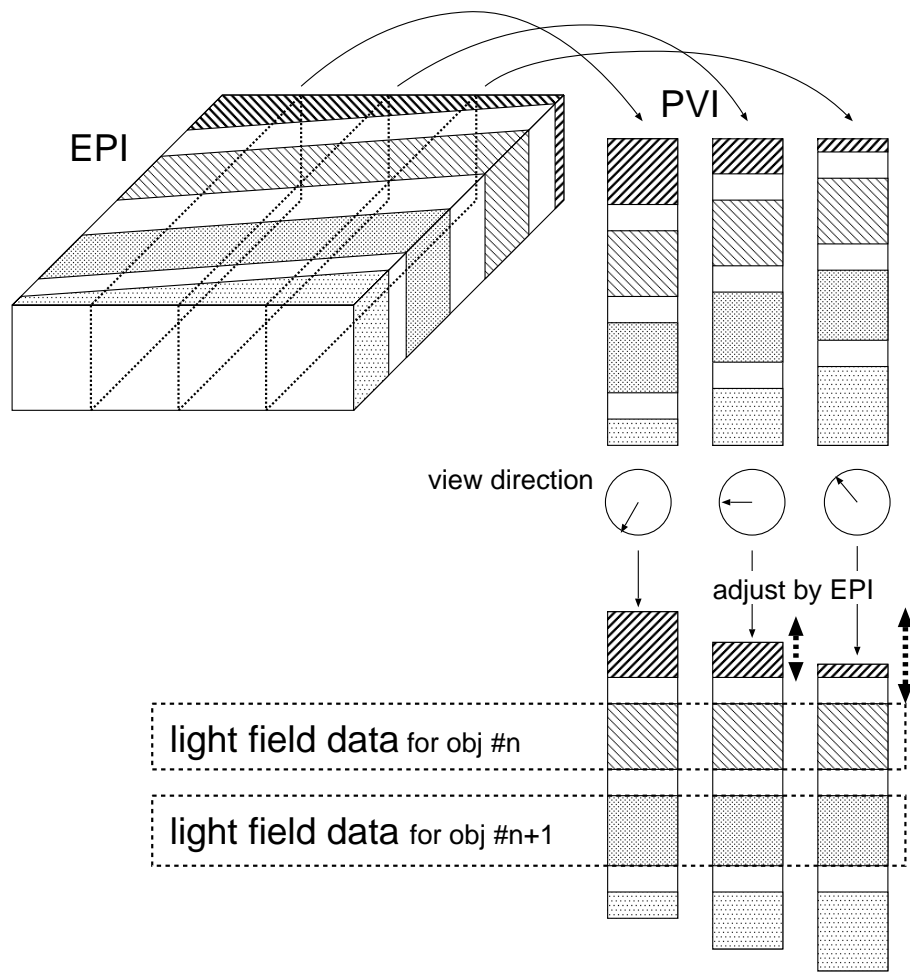


図 5.18 PVI と EPI と表面空間光線の関係

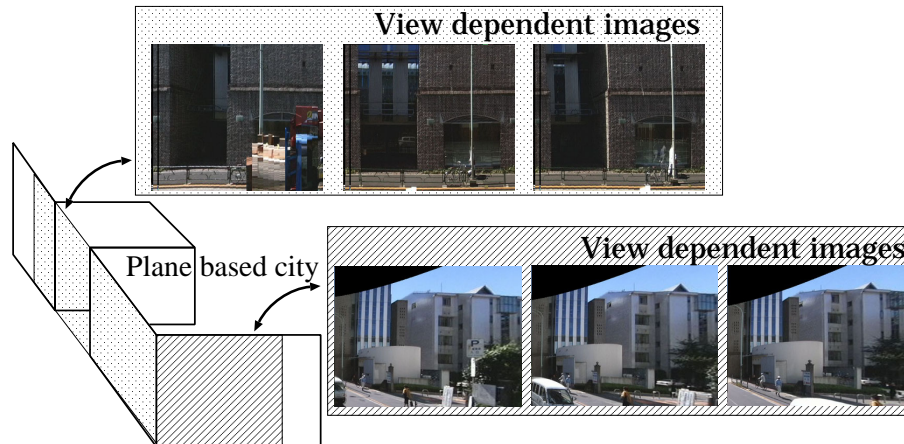


図 5.19 表面光線空間データ

際には車両を理想的に等速直線運動させることは困難であるため、撮影したデータを GPS、ジャイロ等のデータおよび画像処理により補正し、最終的に利用した。

5.4.2 画像合成結果

図 5.23 に、画像を合成した例を示す。(e) は、前章の手法による合成画像である。側面部分に、不連続が発生しているのが分かる。これに対して、ジオメトリを考慮した場合、(a) ~ (c) のように、歪みのない正しい合成画像が描画出来ている。

また、図 5.24 は、建物からの距離が異なる仮想カメラを想定し (図 5.25(d) の (a)(b)(c)) 画像をレンダリングしたものである。本章で提案した幾何形状の取得方法では、電柱や、建物のへこみなどのような細かい形状を取得することが出来ない。しかし、イメージベース法を用いているため、視点位置の変化に応じて、電柱や、建物のへこみ等が正しく描くことが出来ている。

次に、6 台のカメラを用いた結果を示す。新しい視点や建物などのオブジェクトは図 5.26(d) のように位置しているものとする。図 5.26(a)(b)(c) が仮想的なカメラ位置・方向から見た場合の一連の合成画像である。図 5.26(f) は (a) と同じ位置にあるが、3 台のカメラのみを使用して画像合成したものである。図 (a)(b)(c) の合成画像上において、ほとんど歪みが見られないことが分かる。また、視点によって図中の木の位置が移り、同時に窓が建物によって遮蔽されているなど、正しくレンダリング出来ていることが分かる。また、第 3 章の手法により、より多くのカメラ (6 台) を歪み無く統合することが出来た結果、3 台のみしか使用しない場合には画角の制限より画像の上下に描くことが出来ない部分があったが、これを正しくレンダリングすることが出来た。



PVI 画像

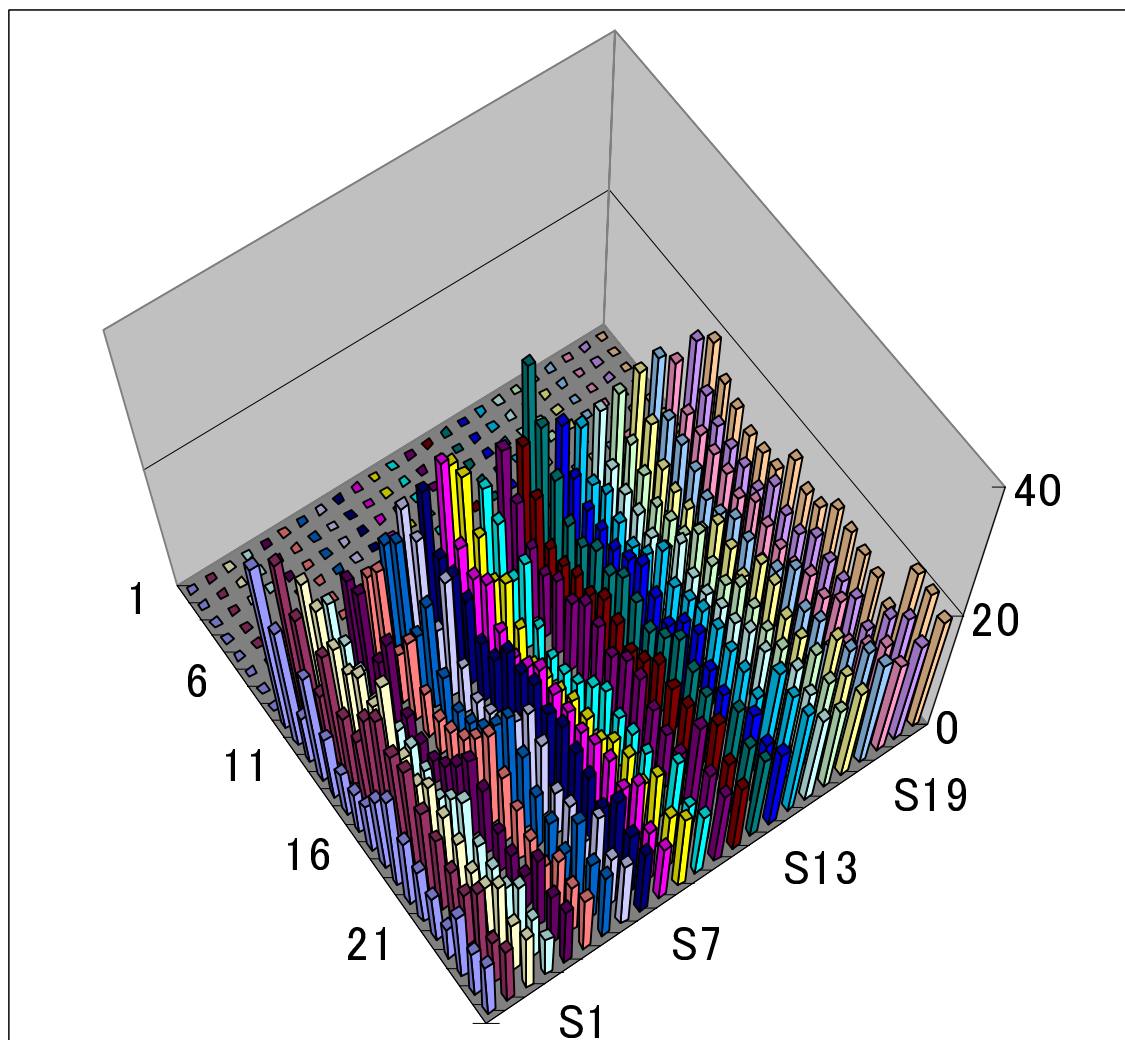


図 5.20 圧縮率



異なる視点位置からの透視変換画像

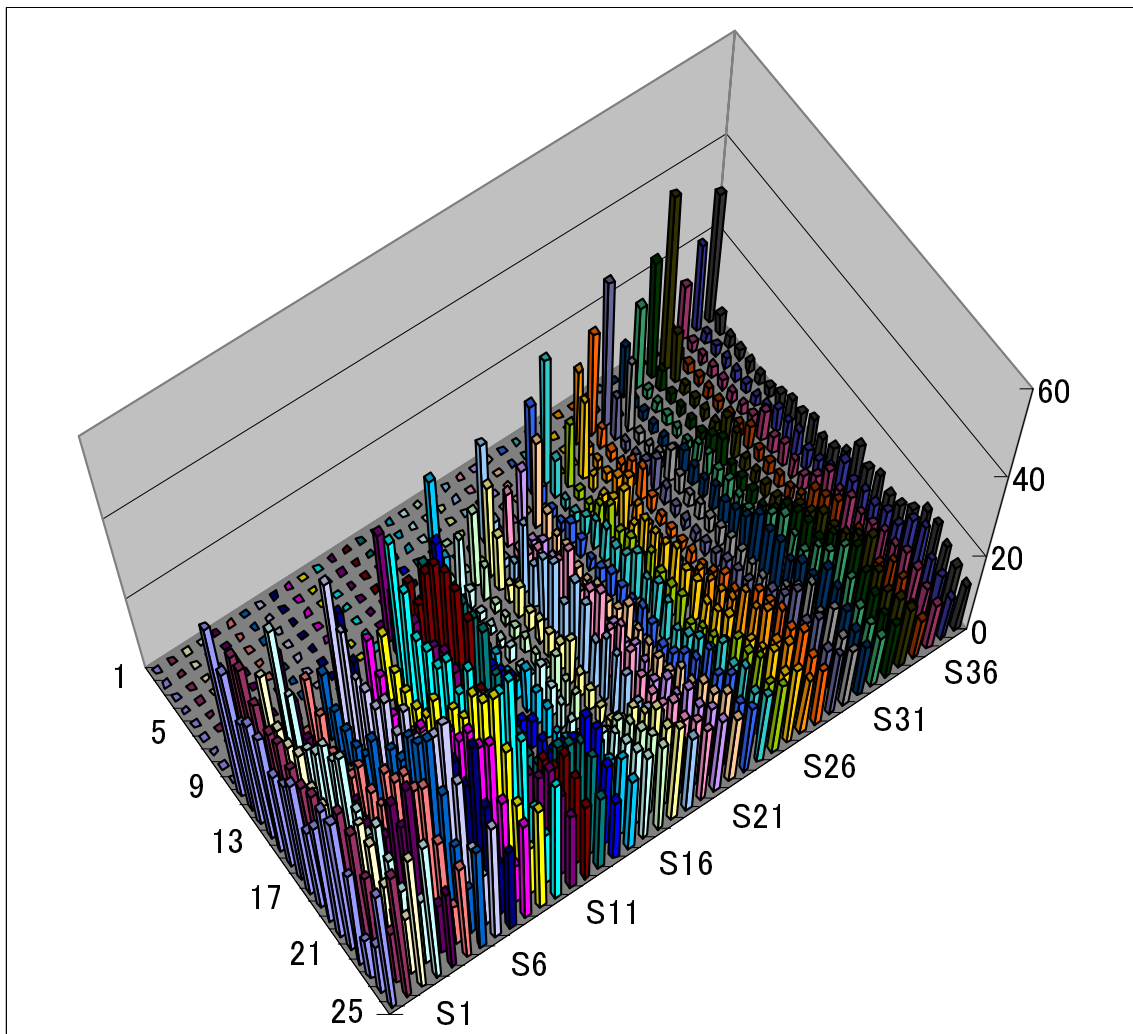


図 5.21 圧縮率

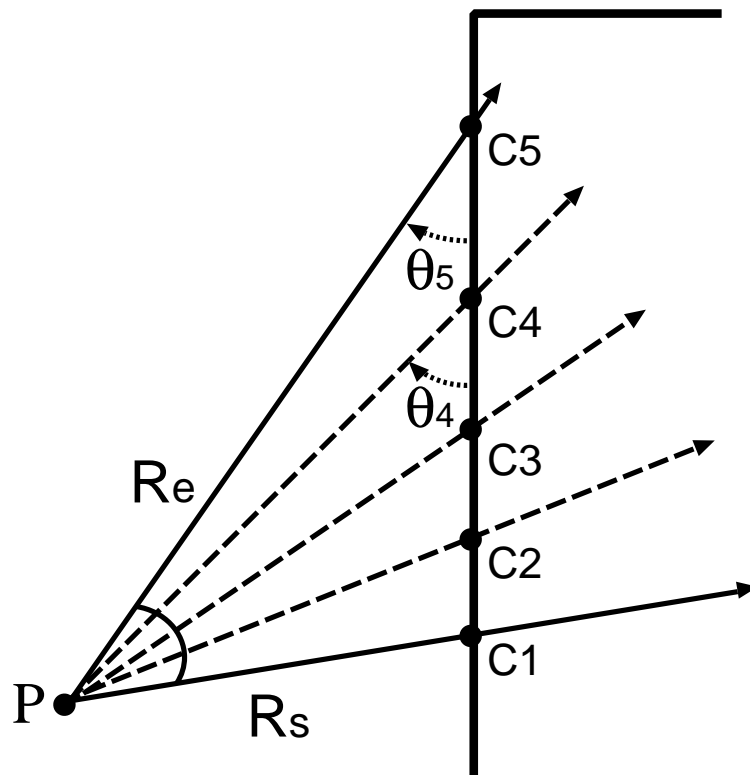


図 5.22 表面光線空間によるレンダリング

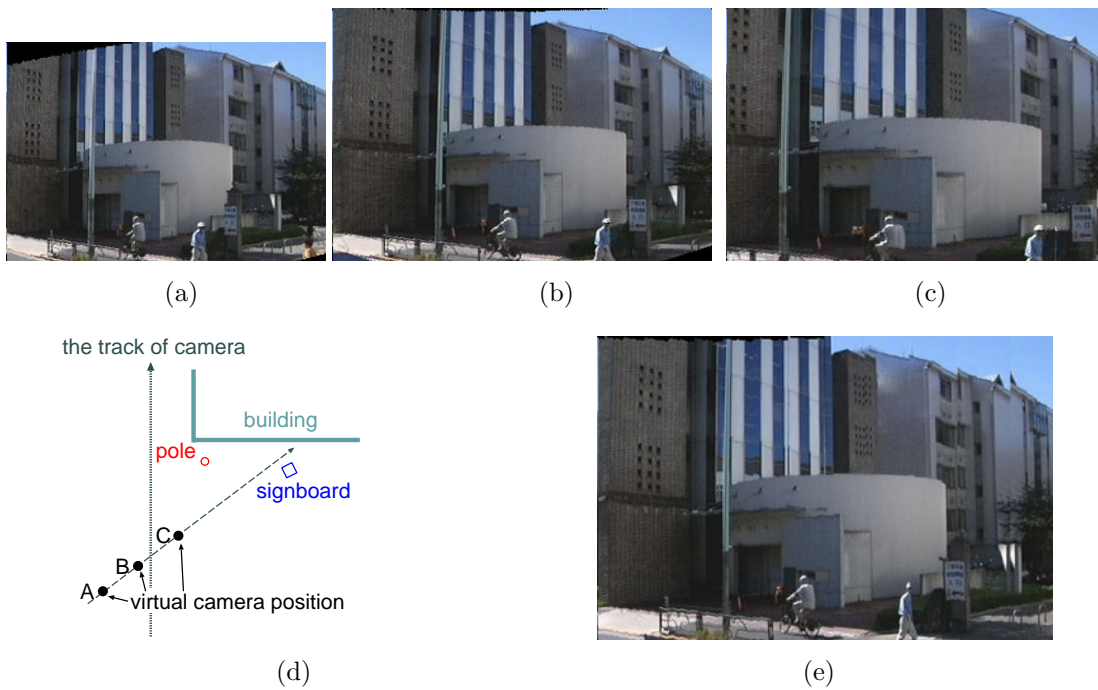


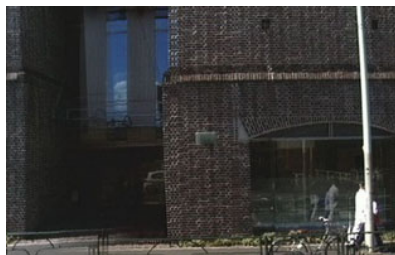
図 5.23 (a)(b)(c): 本手法による合成例,(d):仮想カメラ位置. (e): 道路と建物の距離の 1 パラメータのみを利用した合成画像例



(a)



(b)



(c)

図 5.24 オクルージョンの正確なレンダリング. 上から順にカメラが建物に向かって接近

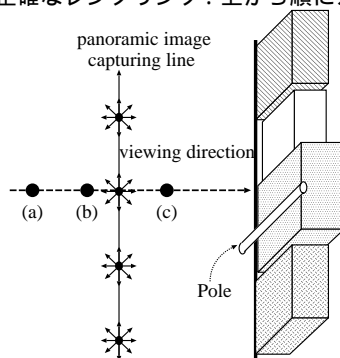


図 5.25 仮想カメラ位置

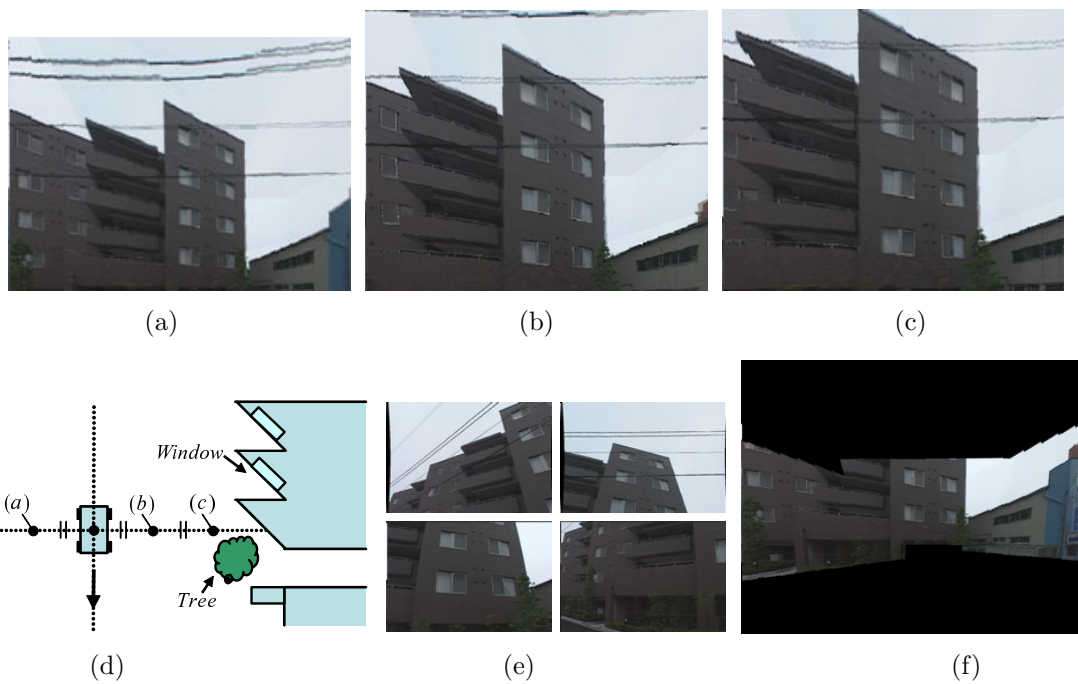


図 5.26 (a)(b)(c): 本手法により合成した画像 (d): カメラ位置 (e): 複数カメラからの入力画像 (f): カメラを 3 台だけ使用した場合

5.5 まとめ

本章では幾何情報を用いた都市空間のモデリング手法を提案した。その際、全方位画像から幾何情報を推定するための時空間解析手法について述べた。また、鏡面反射型全方位画像に特殊な変換を施すこと無く、パノラマ画像や奥行き情報の獲得が可能となる EPI に類似した切断面を得る手法を示した。

画像合成に関しては、表面光線空間法に基づいた歪みの無い合成手法、及びこれを広域環境に拡張する手法を提案した。また、時空間画像解析を利用した効率的な表面光線空間データの構築手法について述べた。さらに、得られた表面光線空間データが、固有空間法などの手法により、効率良く圧縮出来ることも示した。

提案した手法を実際のデータに適用したところ、歪みの無い現実感の高い 3 次元都市を構築することが出来た。

【 参 考 文 献 】

- [1] Jinxiang Chai, Xin Tong, Shing-Chow Chan, and Heung-Yeung Shum: “Plenoptic sampling”, *ACM SIGGRAPH*, pp. 307–318 (July 2000).
- [2] Nishino: “Eigen-texture method”, *IEICE Trans. (In Japanese)*, Vol.**J82-A**, No. 8, pp. 1351–1359 (1999).
- [3] W C. Chen, R. Grzeszczuk, and J Y. Bouguet: “Hardware-accelerated visualization of surface light fields”, in *SIGGRAPH 2001 Course*, Computer Graphics Proceedings, Annual Conference Series ACM, ACM Press / ACM SIGGRAPH (2001).
- [4] Daniel Wood, Daniel Azuma, Wyvern Aldinger, Brian Curless, Tom Duchamp, David Salesin, , and Werner Stuetzle: “Surface light fields for 3d photography”, in *ACM SIGGRAPH* (July 2000).
- [5] Kari Pulli, Michael Cohen, Tom Duchamp, Hugues Hoppe, Linda Shapiro, and Werner Stuetzle: “View-based rendering”, in *Visualizing Real Objects from Scanned Range and Color Data Proceedings of 8th Eurographics Workshop on Rendering*, St. Etienne, France (June 1997).
- [6] 山本正信: “連続ステレオ画像からの 3 次元情報の抽出”, 電子情報通信学会論文誌 (D), Vol.**J-69**, No. 11, pp. 1631–1638 (Nov. 1986).
- [7] R Bolles, H Baker, and D Marimont: “Epipolar plane image analysis: an approach to determining structure from motion”, *Int.J.of Computer Vision*, Vol. **1**, pp. 7–55 (1987).

第 6 章

結 論

本論文は、全方位カメラを用いることで、効率良く広域な実世界空間をサンプリングし、現実感の高い仮想空間の構築の実現する手法について論じたものである。最後に、本論文の主たる成果についてまとめ、次いで今後の課題と展望について触れ、全体の結論とする。

6.1 本論文の主たる成果

本論文では、全方位画像を利用した都市空間作成の手法について述べた。主に、全方位画像の取得方法とその画質向上、イメージベース法による都市空間の作成方法、およびイメージベース法と幾何情報を組み合わせる手法について述べた。これらの手法において、全方位画像とその時空間画像解析を利用することで効率的な画像取得が図れ、その結果、自由視点から見た合成画像を歪みなく簡便に生成できることがわかった。以下では、本論文の主たる成果についてまとめる。

広域な空間のデータの効率的な取得

2種類の全方位カメラを用いて広域な空間を効率的に計測する手法を述べた(第2章及び第3章)。

全方位カメラはその種類や形態により、それぞれ光学的に特殊な性質がある。本論文では、鏡面反射型全方位カメラおよび複数台のカメラを用いた全方位カメラを用いて実験を行なった。

そこで、まず、鏡面反射型全方位カメラについてその光学特性を明らかにし、全方位映像から任意方向の透視変換画像を生成する方法について述べた。さらに、鏡面反射型全方位カメラの問題点である低解像度を解決するための超解像度手法を提案した。

次に、複数台カメラを用いて歪みの無い全方位画像を生成する手法を提案した。複数台のカメラで撮影した画像群を歪み無く統合するためには、光学中心が一致している必要がある。通常のカメラのように光学中心がカメラ内部にある場合、空間的に光学中心を一致させることは難しい。そこで、本論文では、時空間へ拡張することで光学中心の一致を実現する手法を提案した。

また、実際に計測したデータを利用するためには、計測したデータがその他のセンサと同期している必要がある。本論文では、EPIによる時空間画像解析を用いることで、大規模なデータの整合性をセンサを用いずに得ることを実現した。

全方位画像データを用いて計算機内に広域な都市空間を現実感高く再現

このようにして得られた全方位データを用いて、広域な都市空間を現実感高く計算内に再現する手法を提案した(第4章)。

これまで主であった、ポリゴンにテクスチャを貼るというジオメトリベースの手法により、都市のように広域な空間を現実感高く再現することは、形状や表面テクスチャの取得などのデータ取得の面からも、膨大な量のポリゴンを描画することのハードウェア的な限界などからも難しい。そこで本論文では、形状などのモデルが無くても現実感高い画像合成が可能であるイメージベース法による広域空間の再現を試みた。イメージベース法はこれまで主に小さな物体を対象としており、都市のように広域な空間に適用した例は少なかった。本論文は全方位画像を用いることで、これを広域空間に拡張することに成功した。実験の結果、提案した手法により簡単な実装で広域な都市空間が表現可能であることが示された。

イメージベース法に幾何形状を付加することで、イメージベース法の種々の問題を解決

単にイメージベース法を広域な空間に拡張したのでは、歪みが発生することが分かっている。また、データも膨大になる他、シンボリックな情報を付加することも難しい。そこで、本論文では、イメージベース法に幾何形状を付加することで、種々の問題を解決出来ることを示した(第5章)。

幾何形状の取得に際しては、時空間画像解析(EPI)を用いることで、計測した画像そのものを利用する手法を提案した。これにより、多くの場合、新たなセンサを用意しなくても本手法を適用することが出来る。また、EPIとDPマッチングを利用した本手法は、精度の良い形状データを安定して推定することが可能である。

提案した手法を実際のデータに適用したところ、歪みの無い現実感の高い3次元都市を構築することが出来た。

6.2 今後の課題と展望

ここでは、今後の課題と展望について述べる。なお、本論文で示した手法に関する課題は各章のむすびに示してある。

今後の課題としては、イメージベース法で必要となる膨大な画像データをさらに圧縮することや、経路が直線でない場合の実際の対処方法などが挙げられる。また、周囲の環境を一度に取得できるという全方位画像の光学的特性を生かすには、ダイナミックレンジの改善も重要な課題であろう。さらに、野外であるため影や反射などの影響を避けられないが、これらを除去することも実際には必須である。

また、複数台カメラから全方位画像を生成する際に、設備的な制限からこれまで3台までしかカメラを直線上に並べることが出来なかったが、この台数を増やしてより広域な空間を歪み無く統合した画像を生成することも課題の一つである。

今後の展望としては、上述の課題を克服し、モデルベース、イメージベース両手法の利点を取り入れながら、大規模な都市空間を、実際に計算機内に光学的、空間的、時間的整合性を保ちつつ構築する予定である。

発表論文

■ 学会誌論文

- [1] 川崎洋, 谷田部智之, 池内克史, 坂内正夫: “実世界映像の自動構造化によるインタラクティブ利用”, 電子情報通信学会論文誌 (D-II), **J82-D-II**, 10, pp. 1561–1571 (1999).
- [2] 谷田部智之, 川崎洋, 坂内正夫: “地図とリンクした実写映像のインタラクティブ操作”, 映像情報メディア学会誌, Vol. **53**, No. 10, pp. 1430–1438 (Oct. 1999).
- [3] 川崎洋, 池内克史, 坂内正夫: “時空間画像解析を用いた全方位カメラ映像の超解像度化”, 電子情報通信学会論文誌 (D-II), **J84-D-II**, 8, pp. 1891–1902 (2001).
- [4] 高橋拓二, 川崎洋, 池内克史, 坂内正夫: “全方位画像を用いた広域環境の自由視点レンダリング”, 情報処理学会論文誌 CVIM 3, **42**, SIG13, pp. 99–109 (2001).
- [5] 池内克史, 坂内正夫, 川崎洋, 高橋拓二, 村尾真洋, 佐藤いまり, 甲斐郁代: “全方位画像による仮想都市空間の生成”, 情報処理学会論文誌 CVIM 3, **42**, SIG13, pp. 49–58 (2001).
- [6] 川崎洋, 池内克史, 坂内正夫: “全方位画像と都市モデルを用いた広域都市空間の再構築手法” (投稿準備中).

■ 国際会議

- [7] Hiroshi Kawasaki, Tomoyuki Yatabe, Katsushi Ikeuchi, and Masao Sakauchi: “Construction of a 3D city map from real-world video and digital maps”, in *International Workshop on Urban Multi-Media/3D Mapping* (Sept. 1999).
- [8] Tomoyuki Yatabe, Hiroshi Kawasaki, and Masao Sakauchi: “Construction of a 3D City map from Real-world Video and Digital maps for Interactive Video System”, in *Proceedings of Third IAPR International Workshop on Graphics Recognition (GREC'99)*, pp. 67–74, Jaipur, India (Sept. 1999).
- [9] Hiroshi Kawasaki, Tomoyuki Yatabe, Katsushi Ikeuchi, and Masao Sakauchi: “Automatic modeling of a 3D city map from real-world video”, in *ACM Multimedia* (Oct. 1999).
- [10] Hiroshi Kawasaki, Tomoyuki Yatabe, Katsushi Ikeuchi, and Masao Sakauchi: “Construction of a 3D city map using EPI analysis and DP matching”, in *Asian Conference on Computer Vision*, Vol. 2, pp. 1149–1155 (Jan. 2000).
- [11] Hiroshi Kawasaki, Katsushi Ikeuchi, and Masao Sakauchi: “Spatio-Temporal analysis of omni image”, in *Computer Vision and Pattern Recognition*, Vol. 2, pp. 577–584 (June 2000).

- [12] Takuji Takahashi, Hiroshi Kawasaki, Katsushi Ikeuchi, and Masao Sakauchi: “Arbitrary view position and direction rendering for large-scale scenes”, in *Computer Vision and Pattern Recognition*, Vol. 2, pp. 296–303 (June 2000).
- [13] Hiroshi Kawasaki, Katsushi Ikeuchi, and Masao Sakauchi: “Automatic 3D city construction system using omni camera”, in *International Conference on Multimedia and Expo 2000*, New York City, NY, USA (July 2000).
- [14] Takuji Takahashi, Hiroshi Kawasaki, Katsushi Ikeuchi, and Masao Sakauchi: “Virtual driving system with real-world image”, in *7th World Congress on Intelligent Transportation Systems*, Turin, Italy (Nov. 2000).
- [15] Takuji Takahashi, Hiroshi Kawasaki, Katsushi Ikeuchi, and Masao Sakauchi: “Expanding possible view point of virtual environment using panoramic images”, in *International Conference on Pattern Recognition, IAPR 15th*, Vol. 4, pp. 468–472, Barcelona, Spain (Sept. 2000).
- [16] Hiroshi KAWASAKI, Katsushi IKEUCHI, and Masao SAKAUCHI: “Epi analysis of omni-camera image”, in *International Conference on Pattern Recognition, IAPR 15th*, Vol. 1, pp. 379–383, Barcelona, Spain (Sept. 2000).
- [17] Hiroshi Kawasaki, Michihiro Muro, Katsushi Ikeuchi, and Masao Sakauchi: “Enhanced navigation system with real images and real-time information”, in *8th World Congress on Intelligent Transport Systems* (Oct. 2001).
- [18] Hiroshi Kawasaki, Katsushi Ikeuchi, and Masao Sakauchi: “Light field rendering for large-scale scenes”, in *Computer Vision and Pattern Recognition*, Vol. 2, pp. 64–71, Kauai, Hawaii, USA (Dec. 2001).
- [19] Hiroshi Kawasaki, Katsushi Ikeuchi, and Masao Sakauchi: “Arbitrary view synthesis of real-world environment”, in *Asian Conference on Computer Vision*, Melbourne, Australia (Jan. 2002).

■ 学会研究会

- [20] 川崎洋, 谷田部智之, 池内克史, 坂内正夫: “実世界映像の自動構造化およびこれを用いた 3 次元都市の自動生成”, 信学技報, PRMU99–39, IE99–15, MVE99–35 (July 1999).
- [21] 川崎洋, 谷田部智之, 池内克史, 坂内正夫: “Omni ビデオカメラからの 3 次元都市の自動生成”, 情処研報, CVIM–119, pp. 25–32 (Nov. 1999).
- [22] Takuji Takahashi, Hiroshi Kawasaki, Katsushi Ikeuchi, and Masao Sakauchi: “Rendering with panoramic images”, 情処研報, CVIM–119, pp. 33–40 (Nov. 1999).

- [23] 川崎洋, 池内克史, 坂内正夫: “車載全方位カメラ映像からの建物画像の超解像度化”, 情処研報, CVIM-125, pp. 95-102 (Jan. 2001).

■ 国内大会

- [24] 川崎洋, 谷田部智之, 坂内正夫: “ネットワークを用いた対話型映像情報システムにおける効率的オブジェクト記述法”, 1998 電子情報通信学会ソサイエティ大会, D-II-89 (Dec. 1998).
- [25] 川崎洋, 池内克史, 坂内正夫: “Omni カメラによる時空間画像解析”, 2000 電子情報通信学会総合大会, pp. D-12-146 (Mar. 2000).

謝 辞

本研究に際して、終始丁寧な指導を賜りました坂内正夫教授に深く感謝の意を表します。研究室に配属されて以来、研究に対する姿勢や目標設定などの研究に関する有益な御助言を頂くと共に、生活面に関しても親身になって相談にのっていただきました。

また、毎週のように打ち合せをして頂き、研究に関して適時に適切な助言・指導を頂いた池内克史教授に心から感謝いたします。研究の内容だけでなく、研究の進め方や方向性について大変勉強になりました。

また、研究室において数々の便宜を図って下さった、佐藤 秀 技官をはじめとする職員の方々、秘書の皆様には、大変お世話になり、深く感謝いたします。

研究室の先輩である、学術情報センターの佐藤 真一 助教授、孟 洋 助手、NTT サイバースペース研究所の佐藤 隆 博士、三菱総合研究所の谷田部 智之 博士、CMU の森山 剛 博士には、研究に関して、研究室内に限らずさまざまな議論をさせて頂きました。

研究室の諸先輩、後輩、アルバイトの方々、池内研究室、佐藤研究室の方々には本研究全般にわたって有益な御助言を頂くと同時に、生活面においても種々の御教示を頂きました。特に、高橋 拓二氏、小野 慎太郎氏、松下剛士氏、中島 章氏には、本研究の遂行にあたり熱心な御討論を頂きました。ここに深く感謝いたします。また、広島市立大学の古川 亮 助手には、研究を遂行する上で大変お世話になりました。

本論文は筆者個人のみ力でできたものではなく、上記の方々以外にも電気系専攻同期の方々をはじめとする多数の方々から御教示や御協力を頂きました。最後に、昼夜を問わず研究生生活を共にしたこれらの方々、そして学生生活をあらゆる面で支えてくれた妻、娘、両親、兄弟達に感謝致します。

2002 年 12 月 16 日