

博士論文

車両周辺監視用車載カメラシステムのための
仮想視点映像生成に関する研究

登 一生

2019年 3月 15日

奈良先端科学技術大学院大学
情報科学研究科

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
博士(工学) 授与の要件として提出した学位論文である。

登 一生

審査委員：

萩田紀博教授	(主指導教員)
向川康博教授	(副指導教員)
神原誠之准教授	(副指導教員)
浮田宗伯教授	(豊田工業大学)

車両周辺監視用車載カメラシステムのための

仮想視点映像生成に関する研究*

登 一生

内容梗概

本研究は、車に設置した複数のカメラで撮影した映像から、車の周りの物の有無やその距離や方向を把握しやすい映像を生成できる、車両周辺監視用車載カメラシステムのための仮想視点映像生成方法の開発を目的とする。

この仮想視点映像生成には、車の周りの3次元形状情報が必要であるが、駐車場や道路には、透明物や鏡面反射物、低反射物が存在するため、既存の3次元形状推定法ではこの3次元形状を正しく推定できない。そこで、本研究では、3次元形状を推定する代わりに、複数の近似的な3次元形状モデルを用いた仮想視点映像生成法を提案する。駐車場の車庫入れを想定した3次元シーンを用いたシミュレーション実験と、リアルタイム映像生成ハードウェアを搭載した実験車両による路上評価により、近似3次元形状モデルを用いることで、駐車場や道路の仮想視点映像を生成できることを示す。

次に、本提案法だけでは、180度以上の広視野で映像を生成すると生成した像が歪むために、運転者が物の距離や方向の把握が難しくなる問題を解決することを試みる。この問題は従来の仮想視点映像生成法でも指摘されていた。そこで、仮想視点映像生成用のカメラモデルとして、像は歪まないが180度以上の広視野にできない透視射影と、視野は広いが像が歪む等距離射影との2つを組み合わせた複合射影モデルとを併用する仮想視点映像生成法を提案する。駐車シーンを用いたシミュレーション実験および客観的評価指標により、車の近くは透視射影モデルで像の歪みがなく、遠くは等距離射影モデルで180度以上の広視野の仮想視点映像を生成できる

* 奈良先端科学技術大学院大学 情報科学研究科 博士論文, 2019年3月15日

ことを示す.

しかしながら, 近似 3 次元形状モデル, 複合射影モデルを併用しても, 仮想視点映像中に, 解像度が大きく低下する領域が発生して, 低解像度領域中の物の有無を把握しにくくなる. そこで, 本研究では, 上記のモデルを併用した上で, この問題を解決するカメラ映像の高解像度化法を提案する. この方法は, まず, カラー映像うちの 1 色を高解像度かつ低フレームレートで撮像することで, 高解像度情報を含む映像を, データ量を増やすことなく取得する. 次に, フレーム間の動きと高解像度映像を同時に推定する動き補償型超解像処理によって, カラー映像を高解像度化する. シミュレーション実験により, 仮想視点映像中の低解像度領域の解像度を改善し, 物の有無を把握しやすくなることを示す.

これらの成果により, 車両周辺監視用車載カメラシステムを実現するために, 車に設置した複数のカメラで撮影した映像から, 車の周りの物の有無やその距離や方向を把握しやすい映像を生成できる仮想視点映像生成方法が確立することを結論づける.

キーワード

自由視点映像生成, 車載カメラ, 色画像, 超解像

Studies on Generation of Virtual Viewpoint Video for Vehicle Camera System Monitoring Vehicle Surroundings※

Kunio Nobori

Abstract

The purpose of this research is to develop a method generating virtual viewpoint video, for a vehicle camera system monitoring the surroundings of a vehicle, which can generate virtual viewpoint videos easy to grasp the presence of objects around the vehicle and the distance and direction of the objects, from videos taken with multiple cameras installed on the vehicle.

In order to generate such virtual viewpoint videos, three-dimensional (3D) shape information around the vehicle is necessary. However, since there are transparent objects, specular reflection objects and low reflective objects on parking lots and roads, it is not possible to estimate their 3D shape with existing 3D shape measurement methods. Therefore, in this research, instead of estimating the 3D shape, we propose a virtual viewpoint view generation method using multiple approximate 3D shape models. Simulation experiment using 3D parking scene shows that virtual viewpoint video of parking lot and road can be generated by using approximate 3D shape model.

Next, with the above proposed method alone, when a video is generated with a wide field of view of 180 degrees or more, since the generated video is distorted, there arises a problem that it becomes difficult for the driver to grasp the distance and direction of the object. This problem was also pointed out in the conventional virtual viewpoint image generation method. Therefore, this research attempts to solve this problem. We propose a composite projection model, as a camera model for generating a virtual viewpoint image, that combines two, a perspective projection that does not distort the image but can not be a wide field of view of 180 degrees or more, and an equidistant projection

※Doctoral Dissertation, Graduate School of Information Science, Nara Institute of Science and Technology, March 15, 2019.

where the field of view is wide but the image is distorted. Based on the simulation experiment using the parking scene and the objective evaluation index, it shows that there is no distortion of the image in the perspective projection model in the vicinity of the vehicle and in the distance can generate the virtual viewpoint video with the wide field of view of 180 degrees or more in the equidistant projection model.

Even when the approximate 3D shape model and the composite projection model are used together, a region with significantly lower resolution occurs in the virtual viewpoint video, making it difficult to grasp the presence of objects in the low resolution region. Therefore, in this research, we propose a method to improve the resolution of camera image, which solves this problem by using the above model together. In this method, first, by imaging one color at high resolution and low frame rate, video including high resolution information are acquired without increasing the amount of data. Next, the resolution of the color video is improved by motion compensation type super-resolution processing that simultaneously estimates motion between frames and high resolution image. Simulation experiment shows that the resolution of the low resolution region in the virtual viewpoint video can be improved and grasped easily.

Thus, our proposed virtual viewpoint video generation method is capable of generating videos that are easy to grasp the presence of objects around the vehicle and its distance and direction, from the videos captured by the plurality of cameras installed on the vehicle.

Keywords:

Free viewpoint video, vehicle camera system, color image, super-resolution.

目次

1. 序論	1
1.1. 研究の背景	1
1.2. 本研究の目的	3
1.3. 関連研究とその課題	4
1.3.1. 自由視点映像の生成	4
1.3.2. 自由視点映像の課題	9
1.4. 本研究の方針と本論文の構成	12
2. 近似3次元形状モデルを用いた車載カメラシステムのための仮想視点映像生成手法 15	
2.1. 序言	15
2.2. 近似3次元形状モデルを用いた仮想視点映像生成	18
2.2.1. 手法概要	18
2.2.2. 仮想視点映像生成手順	18
2.2.3. 近似3次元形状モデル	22
2.2.4. リアルタイム仮想視点映像生成ハードウェア	23
2.3. 評価実験	25
2.3.1. 実験車両のシステム構成	25
2.3.2. リアルタイムシステムによる実験結果	26
2.3.3. 考察	29
2.4. 結言	34
3. 複合射影カメラモデルを用いた車載カメラシステムのための低歪み全周囲画像生成手法 35	
3.1. 序言	35
3.2. 関連研究	37
3.3. 複合射影カメラモデルを用いた全周囲画像生成	38
3.3.1. 全周囲画像生成手法の概要	39
3.3.2. 複合射影カメラモデル	40
3.3.3. 半円柱形状モデル	41
3.3.4. 設計指針	41

3.4.	評価実験	42
3.4.1.	カメラモデルの比較評価	42
3.4.2.	形状モデルの比較評価	46
3.4.3.	カメラモデルと形状モデルの客観評価	46
3.4.4.	駐車シーンの実画像での評価	50
3.4.5.	考察	54
3.5.	結言	54
4.	フレーム間動きと生成画像の同時推定による動き補償型高解像度化手法	56
4.1.	序言	56
4.2.	高解像度高フレームレート動画生成	58
4.2.1.	色ごとに時空間解像度が異なる動画の撮像	59
4.2.2.	動きと生成画像の同時推定による高解像度高フレームレート動画の生成	60
4.3.	評価実験	66
4.3.1.	シミュレーションによる検証方法	66
4.3.2.	提案手法による動きの推定精度の向上	69
4.3.3.	提案手法による生成画像の画質向上	71
4.3.4.	提案手法による解像度の向上および折り返しの復元	82
4.3.5.	考察	85
4.4.	結言	85
5.	提案手法を用いた車両周辺監視用車載カメラシステムの評価実験	86
5.1.	評価方法	86
5.2.	評価実験結果	89
5.2.1.	仮想視点映像の低解像度領域の高解像度化	89
5.2.2.	データ伝送遅延	99
5.2.3.	3つの手法により生成した仮想視点映像	100
5.2.4.	考察	101
5.3.	結言	107
6.	結論	108
6.1.	本論文のまとめ	108
6.2.	今後の展望	110

謝辭.....	111
參考文獻.....	112
研究業績.....	122

图目录

Figure 1 Conventional in-vehicle camera systems.	2
Figure 2 In-vehicle camera system using virtual viewpoint video.	3
Figure 3 Classification of methods for generating free viewpoint video.	4
Figure 4 The virtualized reality[8].	6
Figure 5 Camera configurations for free-viewpoint video.	8
Figure 6 An example of distorted free viewpoint video with a wide field of view.	10
Figure 7 Example of partial low resolution in free viewpoint video.	11
Figure 8 Overview of thesis content and the proposed system.	14
Figure 9 Approximate 3D shape models of vehicle’s surroundings.	17
Figure 10 Overview of the proposed method.	19
Figure 11 Geometry model for virtual view generation.	20
Figure 12 Areas taken with multiple cameras.	22
Figure 13 Configurations of approximate 3D models and virtual viewpoints.	23
Figure 14 Block diagram of the image synthesis hardware.	24
Figure 15 Configuration of the experimental vehicle with 4 cameras.	25
Figure 16 Generated Virtual Viewpoint Video.	28
Figure 17 Generated Virtual Viewpoint Video from different viewpoints.	29
Figure 18 Example of virtual viewpoint video with a large difference between approximate 3D shape model and actual shape.	31
Figure 19 Generated images with different method in overlapping area.	33
Figure 20 A configuration of vehicle camera system and examples of generated images. .	36
Figure 21 3D shape models of the proposed method.	41
Figure 22 Camera images of parking scene used for experiment.	43
Figure 23 Images generated by (a) perspective projection, (b) equidistance projection and (c) composite projection (proposed) using the plane shape model.	45
Figure 24 Images generated by composite projection using (a) a plane, (b) a half ball and (c) a halfpipe shape models.	47
Figure 25 Surround view images generated using the real images of the parking scene. ...	50

Figure 26 An example of the advantage of wide field of view.	51
Figure 27 Generated images of a pedestrian near the vehicle.	52
Figure 28 Generated images in various parking scenes (proposed method).	53
Figure 29 3-image-sensor camera.	59
Figure 30 Outline of proposed method.	60
Figure 31 Block diagram of simultaneous estimation of motion and high resolution images.	65
Figure 32 Flowchart for evaluating results of simulations.	67
Figure 33 ITE video sequences for evaluation.	68
Figure 34 RMSE of estimated motions.	71
Figure 35 CPSNRs for ITE video sequences ($N_p=N_t=16$).	73
Figure 36 CPSNR of different enlargement ratio.	74
Figure 37 Reconstructed and input images of "Canal" ($N_p=N_t=16$).	75
Figure 38 Reconstructed and input images of "Yachting" ($N_p=N_t=16$).	76
Figure 39 Reconstructed and input images of "Driving" ($N_p=N_t=16$).	77
Figure 40 Street-view video sequences for evaluation, captured by vehicle mounted cameras.	79
Figure 41 Reconstructed images and input images.	80
Figure 42 CPSNR of different enlargement ratio of street view sequences.	81
Figure 43 CPSNRs of generated high resolution images by color cross-correlation.	83
Figure 44 Up-sampled images using color cross-correlation.	84
Figure 45 Flowchart of evaluation procedure.	87
Figure 46 Test video sequences for evaluation.	88
Figure 47 Reconstructed video from same amount of pixel data ($N_p=N_t=16$).	91
Figure 48 Examples of reconstructed image ($N_p=N_t=16$).	92
Figure 49 Resolution of reconstructed video.	93
Figure 50 Generated virtual viewpoint video.	97
Figure 51 Resolution of virtual viewpoint videos	98
Figure 52 Data Transfer Delay of the Proposed Method	99
Figure 53 Virtual viewpoint video generated by the three proposed methods.	103
Figure 54 Virtual viewpoint videos generated by the three proposed methods.	104

Figure 55 Virtual viewpoint images of “Side by side parking lot“.....	105
Figure 56 Virtual viewpoint images of “Parallel parking lot“.....	106

表目次

Table 1 Characteristics of methods for generating free viewpoint video	5
Table 2 Values of camera parameters used in the experiment.	43
Table 3 Values of shape model used in the experiment.	48
Table 4 Objective evaluation results of camera and shape models.	49
Table 5 Characteristics of ITE video sequences.	69
Table 6 Camera parameters of vehicle mounted cameras and virtual viewpoint.	89
Table 7 Number of pixels and frame rate of input and reconstructed videos.	89
Table 8 Spatial resolution of reconstructed videos.	93

1. 序論

1.1. 研究の背景

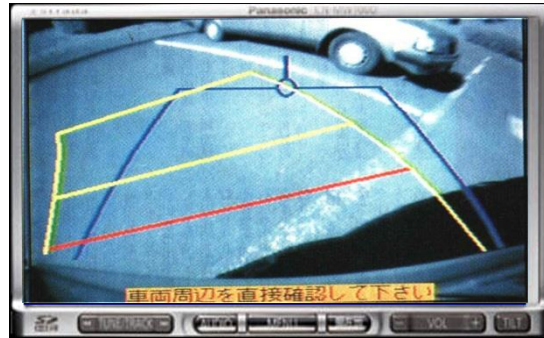
運転者の安全な運転を支援することを目的として、車の周りの映像を提示する車両周辺監視用の車載カメラシステムが実用化されている。このシステムを用いると、運転者から直接見ることができない車の周りの死角領域にある障害物や白線などの監視対象物を、モニタを通じて見ることができるようになるため、障害物との接触事故の回避などの安全な運転の支援に有効である。[1][2][3]

このような車載カメラシステムの典型例として、車の後方や前後左右にカメラを設置して車の周りの状況を撮影し、撮影した映像を車内に設置したディスプレイで表示するシステムがある(Figure 1)。これらの車載カメラシステムは、撮像機能、映像生成機能、表示機能の3つの機能で構成される。

撮像機能は、車の周りの映像を撮影する機能であり、運転者が直接見ることができない車の周りの死角領域や、運転者が前方を向いているときに視野外となる領域を含む映像を撮影する。通常的車載用カメラのほか、夜間の撮影に有効な赤外線カメラや、広範囲の撮影に有効な魚眼カメラを用いることで、さまざまな環境下での映像を撮像できる。車の周りの死角領域を全て撮影するためには、車体による遮蔽があるため、複数のカメラが必要である。一方、カメラを設置でき、かつ、死角を撮影できる場所は限られているため、限られた台数で死角領域の映像を撮影することが問題となる。

映像生成機能は、撮像した映像から運転者が理解し易い映像を生成する機能である。映像生成機能に関しては、複数のカメラで撮影された複数の映像のうちの1つを選択して出力するシステムは(Figure 1a)、運転者が車の全周囲の死角領域を同時に見ることができないという問題がある。これに対し、複数のカメラで撮影された映像を並べて表示するシステムは(Figure 1b)、運転者が車の全周囲を同時に見ることができる。しかし、異なる視点位置と方向で撮影された複数の映像を並べただけの映像から、映像中の監視対象物までの距離や方向を、運転者が把握することが困難であるという問題がある。

表示機能とは、生成した映像を車室内のディスプレイなどに表示することで、ド



(a) A single rear camera system



(b) A multi-camera system

Figure 1 Conventional in-vehicle camera systems.

ライバに見せる機能である。表示機能に関しては、カーナビゲーションや車載用のTVなどを利用することができる。

以上のような背景から、本研究では、車に設置した複数のカメラで撮影した映像から、車の周りにある障害物や白線などの監視対象物の有無とその距離や方向の把握が容易な映像を生成できる、車両周辺監視用車載カメラシステムのシステムの実現を目指す。

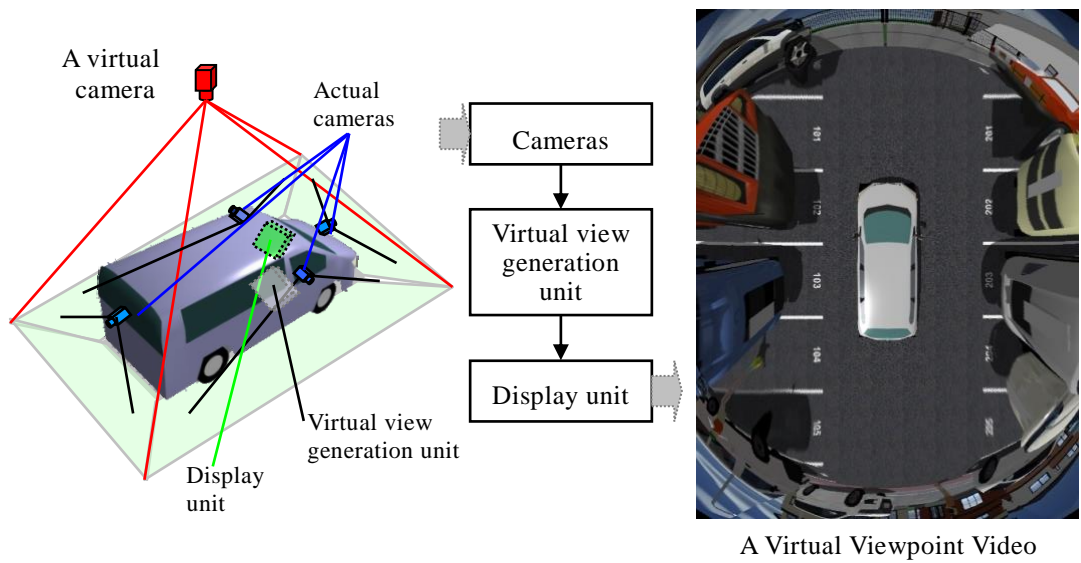


Figure 2 In-vehicle camera system using virtual viewpoint video.

1.2. 本研究の目的

本研究では、車に設置した複数のカメラで撮影した映像から、車の周りにある障害物や白線などの監視対象物の有無や、監視対象物までの距離や方向の把握が容易な映像を生成できる、仮想視点映像生成方式の確立を目的とする。

確立した方式によって、監視対象物の有無やその距離や方向の把握が容易な全周囲の映像を運転者に提示する車両周辺監視用の車載カメラシステム(Figure 2)の実現を目指す。

上記の目的のためには、(i)駐車場や道路などに存在するさまざまな被写体に対応可能であること、(ii)監視対象物の有無やその距離や方向の把握の妨げとなる原因となるアーチファクトが少ないこと、(iii)大がかりな装置を用いずに誰もが簡便に利用できるシステムであること、が重要である。

1.3. 関連研究とその課題

1.3.1. 自由視点映像の生成

カメラで撮影した映像から仮想的な視点からの映像を生成する方法は、自由視点映像生成を目的とした研究領域において多くの方法が提案されている。自由視点映像を生成する方法は、被写体の3次元形状を用いるモデルベース手法と、多視点映像を用いるイメージベース手法の2つに大別される(Figure 3)。モデルベース手法とイメージベース手法のそれぞれの特徴を Table 1 に示す。

イメージベース手法 (Image-based Rendering):

イメージベース手法は、被写体を多視点で撮影した映像から、被写体の3次元形状を用いることなく、自由視点映像を生成する手法である[4][5][6][7]。複数の多視点映像から1映像を選択して幾何変形することで映像を生成する方法から、1画素ごとに選択して映像を生成する方法が提案されている。

Light Field Rendering[4]は、カメラで撮影した画像の1画素ごとの画素値とその画素への入射光の3次元空間中の直線(光線)とを対応付ける。そして、構築した多数のカメラ画像の光線群と、新しい視点の各画素を対応づけることで映像を生成する。大量のカメラを用いることで、実写に近い高い画質の映像を生成できる。

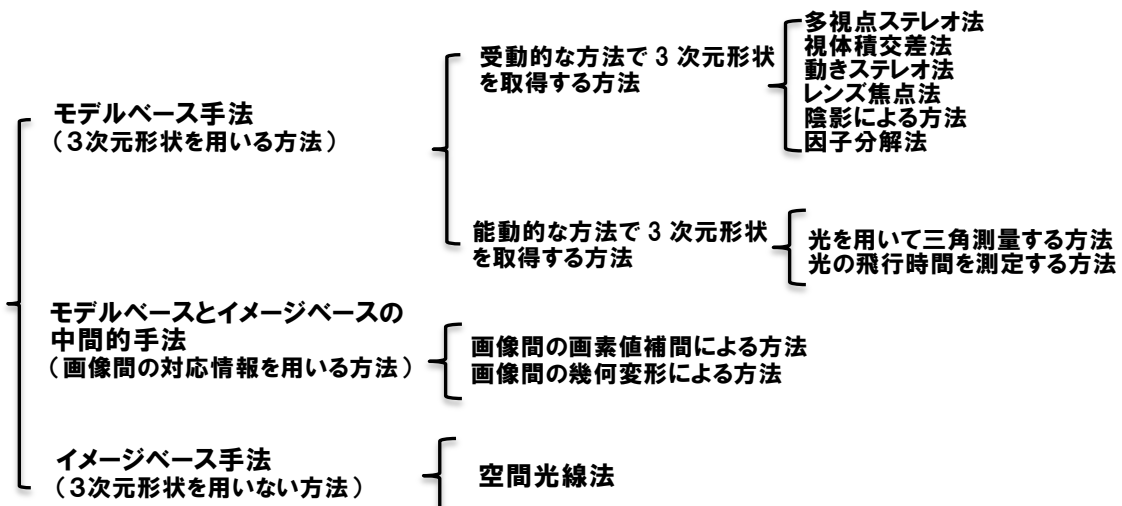


Figure 3 Classification of methods for generating free viewpoint video.

Table 1 Characteristics of methods for generating free viewpoint video

	Model-based method	Image-based method
Procedure	Estimate 3D shape of objects from multi-viewpoint video, and generate free viewpoint video using 3D shape and multi-viewpoint video by CG techniques.	Construct light-ray model (3D lines and color) from multi-viewpoint video, and generate free viewpoint video by reconstructing a set of light-rays at new viewpoint.
Features	High degree of freedom of viewpoint position.	High image quality when using dense camera array.
Challenges	Error in 3D shape estimation. Computational complexity.	Restrictions on the position of the viewpoint. Vast amount of video data.

一方、この方法は複数のカメラ映像の選択や補間などにより、新しい視点の映像を生成する手法であることから、カメラから離れた位置に新しい視点を置くと、生成画像の画質が低下する課題がある。言い換えると、映像生成の視点位置は、カメラに囲まれる位置に限定される課題がある。また、多数のカメラで構成されるカメラアレイの大きさや、膨大な映像データ量も、適用するシステムによっては大きな課題となる(Table 1)。

モデルベース手法 (Model-based Rendering) :

モデルベース手法は、被写体の3次元形状を用いて、自由視点映像を生成する手法である[8][9][10][11][12][13][14]。この手法は、映像から被写体の3次元形状を推定する処理と、推定した3次元形状と映像を用いて自由視点映像を生成する処理の2つの処理で構成される。

モデルベース手法の代表的な手法として、多視点映像から映像中の被写体の3次元形状を推定し、自由視点映像を生成する方法が提案されている。Kanade[8]らは、ドーム状に配置した多数のカメラ(Figure 4a)を用いて被写体の多視点画像を撮影し、画像間の対応付けによってドーム内の被写体の3次元形状の情報である、Depth mapを推定する(Figure 4c)。そしてこのDepth mapと被写体の画像を用いて、新たな視点の画像を生成する(Figure 4d)。

モデルベース手法は、被写体の3次元形状を利用できるため、映像生成のための視点位置や方向の自由度が高い利点がある。一方、推定した3次元形状の誤りや誤差に起因する、自由視点映像の画質低下が課題になる(Table 1)。

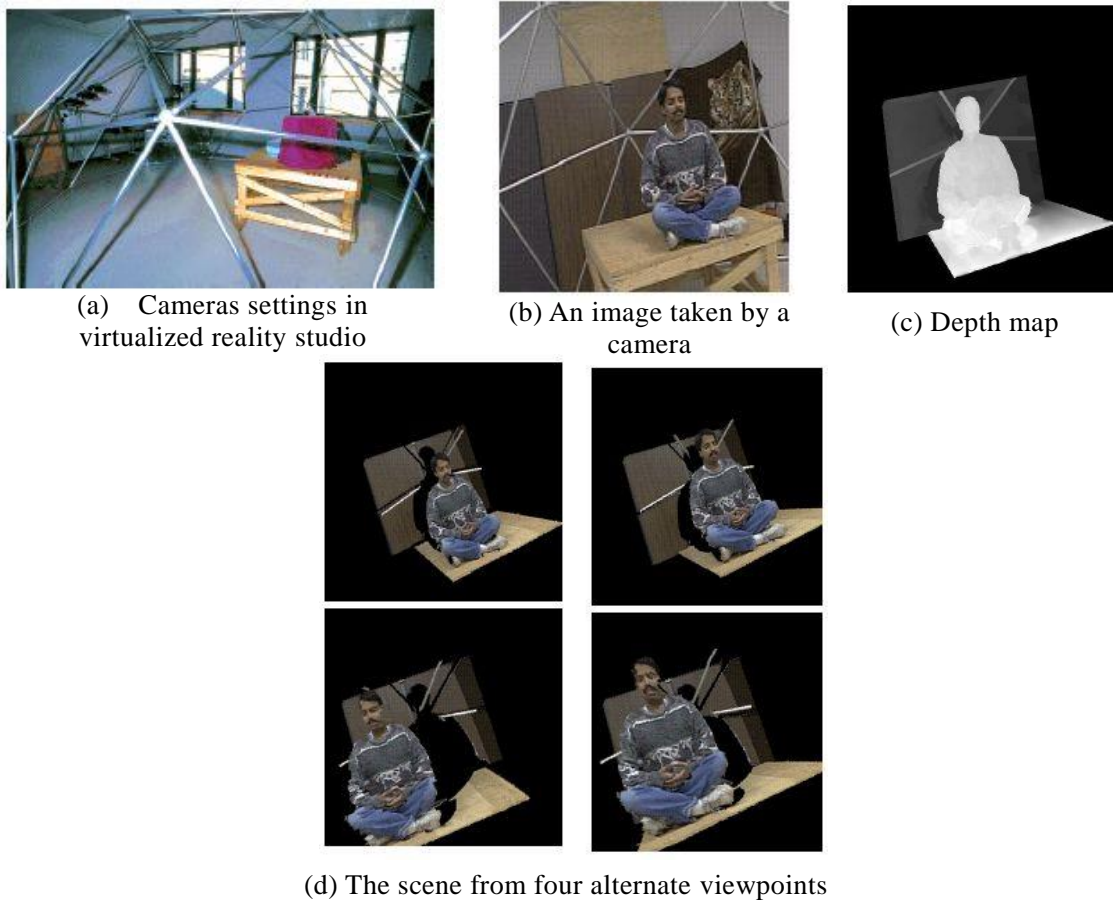


Figure 4 The virtualized reality[8].

上述したモデルベース手法とイメージベース手法の2つの手法は、明確な境界はなく、中間的な手法も提案されている[15].

以降、モデルベース手法に適用可能な、従来の3次元形状推定手法および3次元計測手法を示す.

3次元形状推定手法，3次元計測手法：

多視点映像からの3次元形状の推定は、Computer Visionの研究領域での重要な課題の1つでもあり、多くの手法が提案されている[16][17][18][19][20]. これらの3次元形状推定手法は、映像間の対応付けに用いる特徴量として、テクスチャ、特徴点、陰影、輪郭、等を用いた手法が提案されている.

テクスチャを用いる手法は、画像から局所特徴量[21][22]を算出し、画像間で特徴量

が似ている点や領域を対応付け、事前の校正により取得したカメラ配置の情報を用いて、被写体の3次元形状を推定する。これらの手法は、テクスチャがない被写体や、ガラスなどの透明物、金属などの鏡面反射物、低反射物の3次元形状を推定できないという課題がある。この課題に対し、陰影や輪郭を用いる手法が提案されている。

陰影を用いる方法は、被写体にテクスチャがなく、光源と被写体の材質が既知と仮定できる場合に、被写体表面の輝度値と被写体の法線との関係を用いて、被写体の3次元形状を推定する[23][24]。この方法はテクスチャのない被写体であっても、3次元形状を推定できる。しかしながら、透明物、鏡面反射物、低反射物に適用できない課題がある。また、屋外などの光源が変動する環境や、テクスチャのある被写体に適用できない課題がある。

視体積交差法は、多視点映像中の同一被写体の輪郭を用いて、各視点を頂点、輪郭を断面とする錐体の共通部分を求め、これを被写体の3次元形状とする手法である[9][16][25][26]。この方法は、テクスチャのない被写体や、透明物、鏡面反射物、低反射物であっても、輪郭を抽出することができれば、3次元形状を推定できる。しかしながら、カメラ台数が少ない場合や、カメラの位置に偏りがある場合は、3次元形状の誤差が大きい課題がある。

前述したテクスチャ、特徴点、陰影、輪郭をそれぞれ用いる手法のほか、下記のように、2つ以上を組み合わせることで各手法の欠点を補う方式も提案されている。
テクスチャ+陰影：テクスチャを用いた3次元形状推定方法において、テクスチャがない領域の3次元形状を推定できない課題を解決する手法として、陰影を組み合わせることでテクスチャのない領域の3次元形状を取得する手法[27]。

輪郭+テクスチャ：視体積交差法において、カメラ台数が少ない場合に3次元形状の誤差が大きい課題を解決する方法として、テクスチャを組み合わせることでより誤差の小さい3次元形状を取得する手法[28][29]。

輪郭+陰影：視体積交差法において、カメラ台数が少ない場合に詳細な3次元形状を取得できない課題を解決する方法として、陰影を組み合わせることで詳細な3次元形状を取得する手法[30][31]。

以上に示した、多視点映像から3次元形状を推定する手法には、透明物や鏡面反射物、低反射物の3次元形状を推定でき、かつ、カメラ台数が少ない場合に適用できる手法がなかった。

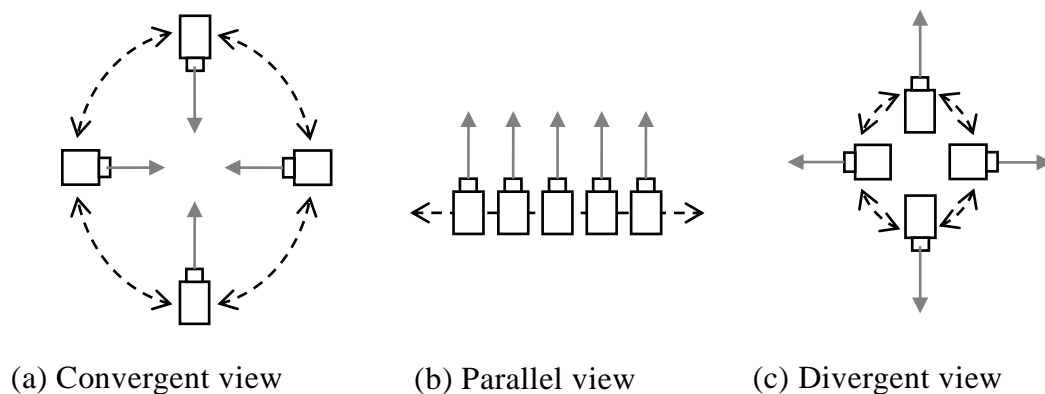


Figure 5 Camera configurations for free-viewpoint video.

被写体の3次元形状を取得する方法は、上述した多視点映像から3次元形状を推定する手法のほかに、光の飛行時間を計測することで、画像と距離情報を同時に取得するRGB-Dカメラを用いる方法も提案されている[32]. RGB-Dカメラを用いると、遮蔽のない距離情報を取得することができる。しかしながら、透明物や鏡面反射物、低反射物の3次元形状を推定できないという課題があった。また、RGB-Dカメラは直射日光下では光量不足により3次元形状を計測できない課題や、カメラの近くに被写体がある場合に、光源の多重反射によって遠方の3次元形状に大きな誤差が生じる課題もある。

カメラ配置：

自由視点映像生成に利用されるカメラ配置の典型例を Figure 5a に示す。自由視点映像のためのカメラ配置は、Convergent, Parallel, Divergent の3つに大別される[33]. 上述した自由視点映像生成方法は、映像を取得するカメラと合成する視点の位置や向きが近いほど、生成される画像の劣化が少ない。そのため被写体に対する視点の位置によって、適切なカメラ配置が異なる。被写体を360度囲むような視点から観察する場合はConvergent view (Figure 5a), 被写体を左右に移動した視点から観察する場合はParallel view(Figure 5b), カメラの周りを取り囲む被写体を観察する場合はDivergent view (Figure 5c)が、それぞれ適切なカメラ配置と言える。

一方、本論文で扱う車載カメラシステムは、車体に設置したカメラで撮影した、車の周りの自由視点映像を生成するためDivergent view (Figure 5c)のカメラ配置となる。

1.3.2. 自由視点映像の課題

従来の自由視点映像生成方式を，車両周辺監視用の車載カメラシステムに適用する場合の課題を示す．

イメージベースの自由視点映像生成は，撮影視点の内挿関係にある視点位置の映像しか生成できないため，Divergent view のカメラ配置，かつ，限られた台数のカメラで撮影した映像から，自由な視点位置からの映像を生成することが困難である．

そこで，本論文ではモデルベースの自由視点映像生成手法を，車載カメラシステムへ適用することを試みる．モデルベースの自由視点映像生成方法を車載カメラシステムに適用する場合の根本的な課題は，システムが稼働する駐車場や道路に，透明物，鏡面反射物，低反射物などの3次元形状を推定できない被写体が存在することである．

また，モデルベースの自由視点映像生成方法によって生成した映像には，理想的な映像には存在しないアーチファクトが発生する場合がある．このアーチファクトの代表的な原因を下記に示す．

- ① 像の幾何的歪み：生成した映像中の被写体像の幾何的な歪みのアーチファクト．撮影カメラの校正誤差，被写体の3次元形状の誤差，および，広視野映像を生成する際のカメラモデルに起因する．3次元形状の誤差は，モデルベース手法における被写体の3次元形状の推定の誤差や誤り，イメージベース手法における光線情報の対応付けの誤差や誤りなどが原因．カメラモデルについては，従来のモデルでは歪みのない180度を超える広視野画像は原理的に生成不可能であることが原因．
- ② 解像度の非均一：生成した映像中の位置によって，解像度が異なるアーチファクト．自由視点映像を生成する際の仮想的な視点と被写体の位置関係は，実際に被写体を撮影するカメラと被写体との位置関係と異なる．そのため自由視点映像中の位置によって，被写体とカメラおよび被写体と仮想視点との距離や向きに大きな差があることが原因．
- ③ 輝度や色の不連続：複数のカメラの境界部で像が不連続になるアーチファクト．複数カメラの感度特性の違いや，その違いを補正するカラーバランスなどのパラメタの誤差が原因．

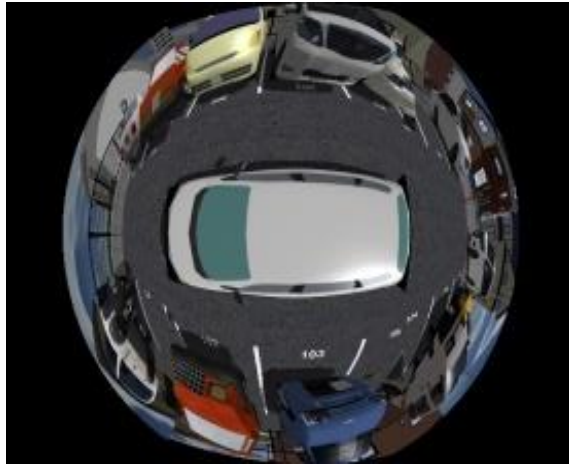


Figure 6 An example of distorted free viewpoint video with a wide field of view.

上記のうち、①像の幾何的歪みの原因である、カメラの校正誤差については、カメラ校正に関する研究[34]や、未校正カメラでの自由視点映像生成手法[13]を適用することで解決できる。3次元形状推定の誤差や誤りは、自由視点映像生成やコンピュータビジョンの主題として取り組まれている。しかしながら、駐車場や道路などに存在するすべての被写体の3次元形状を誤りなく推定できる手法は未だにない。また、広視野映像を生成する際のカメラモデルに起因するアーチファクトの例をFigure 6に示す。カメラモデルについては、コンピュータグラフィックスの主題として取り組まれているが、未だに解決されていない。

②の解像度が非均一になるアーチファクトについては、従来の研究では議論されてこなかった。この自由視点映像生成における、車載カメラと仮想視点の配置によって生じるアーチファクトの例をFigure 7に示す。図から、車の後方の離れた領域、および、車両後方の左右の領域に、解像度が局所的に低下していることがわかる。この図のような非均一な解像度は、運転者の視認性の低下に繋がる可能性がある。

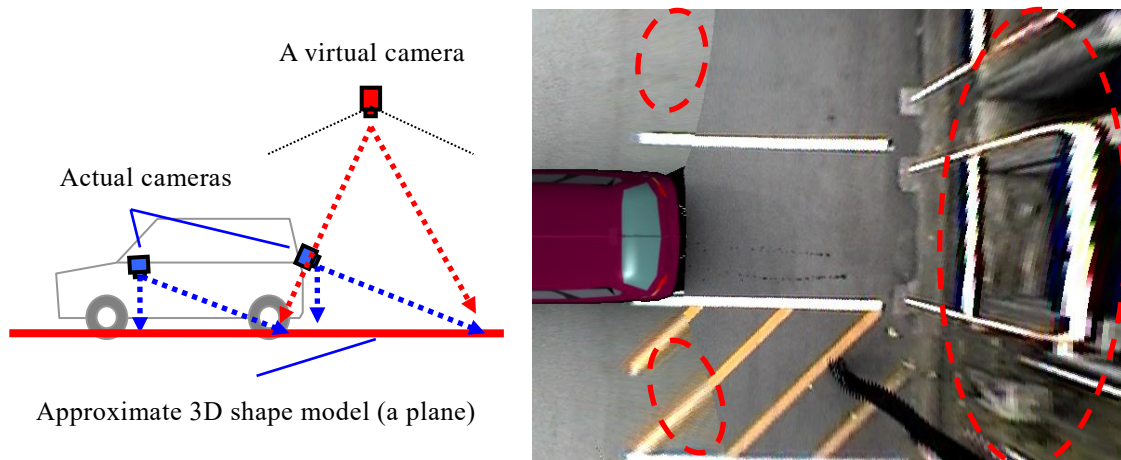


Figure 7 Example of partial low resolution in free viewpoint video.

このような解像度低下に対する1つの解決策は、高画素数のカメラを使用することである。しかし、高画素数のカメラは、カメラや画像データを伝送するシステムのコストが増大するという課題を伴う。また、撮像素子の1画素あたりの面積が小さくなるため、感度が低下するという別の課題も発生する。このコスト増大や感度低下の課題は、車両周辺監視システムにおいては、重要な課題となる。

③の輝度やカラーバランスの不一致は、カメラ間の色輝度補正の処理を適用することで解決できる。

以上のことから、モデルベースの自由視点映像生成方式を車両周辺監視用の車載カメラシステムに適用した場合の課題を示す。

Problem 1: モデルベースの自由視点映像生成方法を車載カメラシステムに適用する場合の根本的な課題は、システムが稼働する駐車場や道路には、透明物、鏡面反射物、低反射物などの3次元形状を推定できない被写体が存在することである。そのため、被写体の3次元形状の推定する従来のモデルベース手法では、自由視点映像を生成できないことである。

Problem 2: 自由視点映像生成方法により180度以上の広い視野角の映像を生成する場合の課題は、魚眼レンズの広視野角カメラモデルを用いて生成した映像におい

て、像が歪み、路面上の直線が画像上では曲線になることである。その結果、画像上での車と物との距離が実際の距離が比例しなくなり、車と物との距離や方向がわかりにくいことである。

Problem 3: 自由視点映像生成方法を、同一領域を撮影するカメラ台数が少なく、かつ、車両の上空から俯瞰する視点など、実カメラから離れた位置に仮想視点がある場合に適用する場合の課題は、仮想視点画像中に局所的な低解像度領域が発生することである。

1.4. 本研究の方針と本論文の構成

本研究の方針と本論文の構成を示す(Figure 8)。

第1章では、本研究の背景と目的を述べた。車両周辺監視システムの構成について概観し、研究課題と位置づけを述べ、関連研究を紹介する。

第2章では、本研究の前提となる、Problem 1を解決し、3次元形状を推定する代わりに、複数の近似的な3次元形状モデルを用いた仮想視点映像生成法を提案する。駐車場の車庫入れを想定した3次元シーンを用いたシミュレーション実験にと、リアルタイム仮想視点映像生成ハードウェアを搭載した実験車両による路上評価により、近似3次元形状モデルを用いることで、駐車場や道路の仮想視点映像を生成できることを示す[35][36][37][38]。

第3章では、第2章の提案法だけでは、180度以上の広視野で映像を生成すると生成した像が歪むために、運転者が物の距離や方向の把握が難しくなる問題(Problem 2)を解決することを試みる。この問題は従来の仮想視点映像生成法でも指摘されていた。そこで、仮想視点映像生成用のカメラモデルとして、像は歪まないが180度以上の広視野にできない透視射影と、視野は広いが像が歪む等距離射影との2つを組み合わせた複合射影モデルとを併用する仮想視点映像生成法を提案する。駐車シーンを用いたシミュレーション実験および客観的評価指標により、車の近くは透視射影モデルで像の歪みがなく、遠くは等距離射影モデルで180度以上の広視野の仮想視点映像を生成できることを示す[39][40]。

第4章では、近似3次元形状モデル、複合射影モデルを併用しても、仮想視点映像中に、解像度が大きく低下する領域が発生して、低解像度領域中の物の有無を把

握しにくくなる問題 (Problem 3) を解決するカメラ映像の高解像度化法を提案する。この方法は、まず、カラー映像うちの 1 色を高解像度かつ低フレームレートで撮像することで、高解像度情報を含む映像を、データ量を増やすことなく取得する。次に、フレーム間の動きと高解像度映像を同時に推定する動き補償型超解像処理によって、カラー映像を高解像度化する。シミュレーション実験により、仮想視点映像中の低解像度領域の解像度を改善し、把握しやすくなることを示す [41][42][43][44][45][46]。

第 5 章では、提案した 3 つの手法が車両周辺監視システムに適用可能であることを示すため、提案手法で生成される仮想視点映像の画質（解像度）の評価と、データ伝送遅延の評価を行った結果を示す。CG 画像を用いたシミュレーション実験により、提案法の有効性を示す。

第 6 章では、本研究で得られた成果により、車両周辺監視用車載カメラシステムを実現するために、車に設置した複数のカメラで撮影した映像から、車の周りの物の有無やその距離や方向を把握しやすい映像を生成できる仮想視点映像生成方法が確立することを結論づけるとともに、今後の展開について述べる。

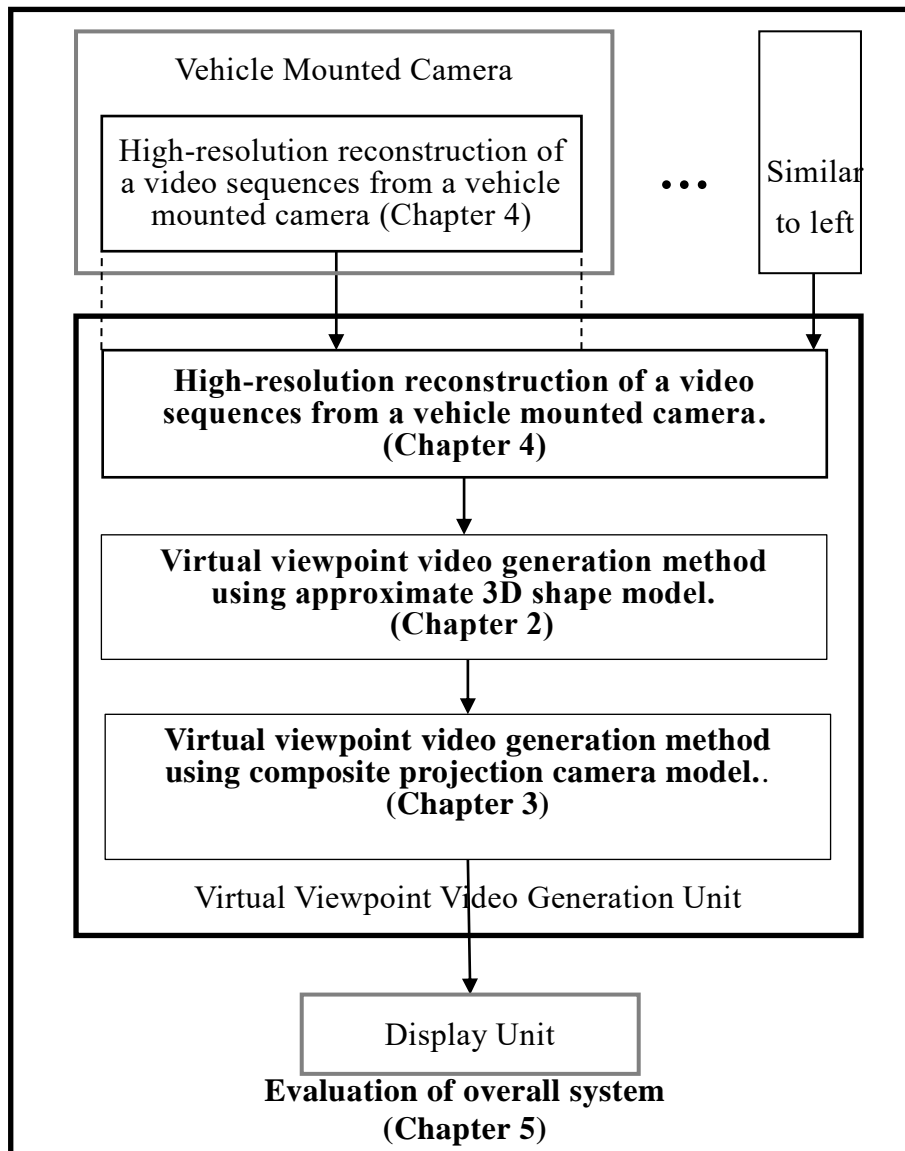


Figure 8 Overview of thesis content and the proposed system.

2. 近似3次元形状モデルを用いた車載カメラシステムのための仮想視点映像生成手法

この章では、車両周辺監視用の車載カメラシステムに適用することができる、ロバストな仮想視点映像生成方法について提案する。本提案の基本のアイデアは、運転者の安全な運転を支援するために、車の周りを上空の1視点から俯瞰した映像を運転者に提示することである。実際にはカメラを置くことができない、車の上空の仮想的な視点から撮影した映像、すなわち、仮想視点映像を生成するためには、物体の3次元形状情報が必要である。しかしながら、駐車場や道路などの実環境における物体の3次元形状の推定は、コンピュータビジョン分野における重大な問題として現在も数多くの研究が行われている。そこで提案法は、被写体の3次元形状を推定する代わりに、事前に作成した近似的な3次元形状モデルを用いて仮想視点映像を生成する。この近似3次元モデルは予め、複数個決定されており、運転操作に基づいて1つが選択される。そのため、3次元形状の推定が困難な被写体が存在する駐車場や道路等においても、仮想視点映像を生成することができる。実験では、実際の駐車場や道路などで撮影した映像を用いたシミュレーション実験と、リアルタイム仮想視点映像生成ハードウェアを搭載した実験車両による路上評価により、3次元形状を推定できない被写体が存在する場合でも、仮想視点映像を生成できることを示す[35][36][37][38]。

2.1. 序言

車両周辺監視用の車載カメラシステムが普及している[47][57][58]。このシステムは、駐車時、車の後退、車両の発進、またはブラインドコーナーを通過するなどの低速での運転シーンにおいて、車の周りの映像を運転者に提示することで、運転者の安全な運転の支援することができる。周りの物や他の車との接触を防止するために、システムによって提示される映像は、(i)死角がないこと、(ii)車の周りの物の有無や、車と物との距離や物の方向がわかり易いこと、(iii)運転シーンに応じた視点や視野範囲の映像を提示できること、が重要である。

従来の車両周辺監視用車載カメラシステムの1つであるリアビューカメラシステ

ム[47]は、1990年代後半から利用可能であった。このシステムは、車体の後部に設置したカメラによって、車の後方を撮影し、撮影した映像を車室内に設置したディスプレイに表示することで、運転者に車の後方の映像を提示する (Figure 1a)。しかし、このシステムでは、車の前や左右に、運転者からが見ることができない死角が残る。

車の全周囲の死角のない映像を提供するために、車の前部、後部、左右に4台のカメラを設置し、これらのカメラで撮影した車の前後左右の4つの映像を並べてディスプレイに表示する、マルチカメラシステムが提案されている (Figure 1b)。しかし、このシステムで表示される、車の前後左右の異なる視点位置と撮影方向で撮影した4つの映像から、運転者が、映像中の物の方向や距離を把握することが困難であることがある。

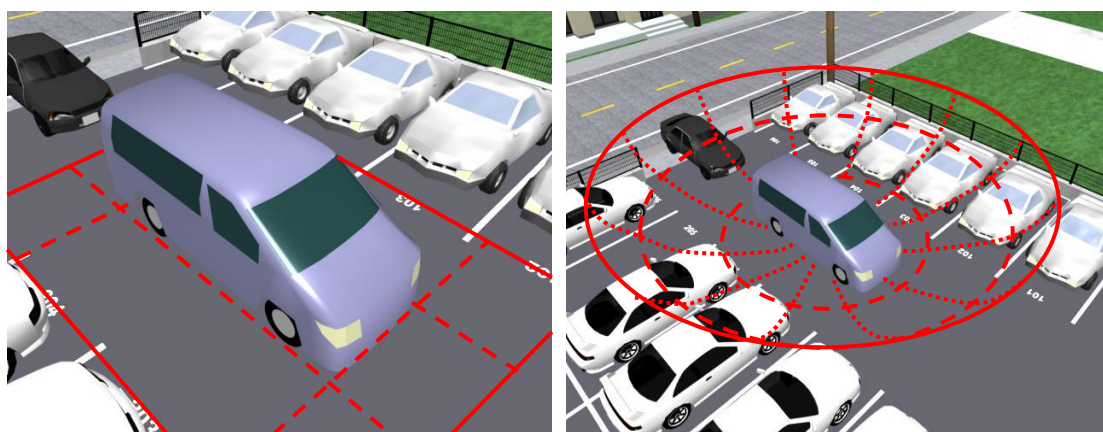
1視点からの360度視野の映像を撮影できる、全方位カメラ[48]が提案されている。しかしながら、1台の全方位カメラでは、車体によって遮蔽されるため、車の全周囲を死角なく撮影することは不可能である。

一方、運転者が低速で運転している間に事故を回避するのに役立つ自動駐車システムも提案されている[49]。このシステムは、GPS、位置センサ、ステアリングセンサを利用して、自動的に駐車することができる。自動駐車システムを使用する場合においても、事故を避ける責任は運転者にあることから、車の周りを死角なく、かつ、わかりやすく運転者に提示するシステムが必要である。また、これらの自動駐車システムでは、駐車できないシーンがあることや、コストが高いことも課題である。

上記した問題を解決するために、車両周辺監視用の車載カメラシステムに、仮想視点生成法を適用した。提案法の基本的なアイデアは、車の周りを上方の1視点から俯瞰した仮想視点映像を運転者に提供することである。この映像は、車の周りを死角がなく、車の周りの物の有無や、車と物との距離や物の方向がわかり易いことから、運転者の安全な運転を支援できる。

第1章で説明したように、これまでも仮想視点映像生成のための様々な方法が提案されている。これらの方法は、モデルに基づく方法と画像に基づく方法に大別される。仮想視点の位置が実際のカメラから離れていても歪みの少ない画像を生成できるため、モデルベースの手法を採用した。

モデルベースの自由視点映像生成方法は、以下の2つのステップからなる。



(a) A plane model for near surroundings (b) A spherical model for far surroundings

Figure 9 Approximate 3D shape models of vehicle's surroundings.

ステップ 1 : シーン内の物体の 3 次元形状を取得する.

ステップ 2 : カメラ画像を 3 次元形状に投影するテクスチャマッピング技術によって仮想視点映像を生成する.

しかし、駐車場や道路などには、透明物や鏡面反射物、低反射物などの物体が存在し、既存の 3 次元推定方法では、これらの物体の 3 次元形状を推定することはほとんど不可能であった。そのため、従来の自由視点映像生成法を車載カメラシステムに適用することは困難であった。

この問題を解決するために、車の周りの形状を近似した、近似 3 次元形状モデルを用いて仮想視点映像を生成するロバストな方法を提案する。提案法は、車両の周囲の 3 次元形状を、平面、平面の組み合わせ、または半球のような単純な形状モデルによって近似する。例えば、典型的な駐車スペースにおける車の周りの 3 次元形状は、平面モデル Figure 9a によって近似する。また、車両の遠方の 3 次元形状は、球状モデル Figure 9b によって近似する。これら複数の近似 3 次元モデルを予め用意しておき、運転状況に基づいて 1 つのモデルを選択する。そして、選択した近似 3 次元形状を用いて仮想視点映像を生成する。

本章の構成は以下の通りである。まず、2.2 節では、近似 3 次元形状モデルを用いて仮想視点映像を生成する提案法の詳細を説明する。2.3 節では、提案法をリアルタイムシステムで評価した結果を、2.4 節では、結言と残課題について述べる。

2.2. 近似 3 次元形状モデルを用いた仮想視点映像生成

本節では，提案手法を詳細に提示し，路上評価のために開発されたリアルタイムシステムについて説明する．

2.2.1. 手法概要

提案手法は，モデルベースの自由視点映像生成アプローチに基づいて仮想視点映像を生成する．従来のモデルベース手法が被写体の 3 次元形状を推定する代わりに，提案法は，複数の近似 3 次元形状モデルをあらかじめ用意しておき，運転シーンに応じて近似 3 次元形状モデルの 1 つを選択することによって，被写体の 3 次元形状を得る．車両周辺監視用の車載カメラシステムの利用シーンは，駐車，車両の後退，車両の発進，またはブラインドコーナーの通過等，限られたカテゴリ数の運転シーンに分類することができる．そこで各運転シーンにおける，典型的な車の周りの 3 次元形状を，近似 3 次元形状モデルとしてあらかじめ用意しておく．これらの運転シーンは，GPS 信号，ギアシフト，速度センサ，および運転者によって操作される他のスイッチなどのセンサによって判定することができる．したがって，本手法は，運転シーンから想定される車の周りの 3 次元形状に近い，近似 3 次元形状モデルを選択して用いることで，仮想視点映像を生成する．

2.2.2. 仮想視点映像生成手順

近似 3 次元形状モデルを用いた仮想視点映像の生成手順を Figure 10 に示す．車載カメラの内部および外部パラメータは，事前に校正しておく．また，近似 3 次元形状モデルや仮想カメラは，想定される運転シーンに応じて予め複数設計しておき，動作時には，センサ信号や運転手の操作等に基づいて，近似 3 次元形状モデルと視点の構成の 1 つのセットを選択する．そのうえで，以下の手順で仮想視点映像を生成する．

Step1: 車の前後左右の 4 カ所に設置された車載カメラによって車の周りの画像を撮影する．

Step2: 車載カメラの画像を，車の周りの 3 次元形状を表す形状モデルに投影する．

Step3: 仮想カメラから見た形状モデル上に投影された画像を生成する．

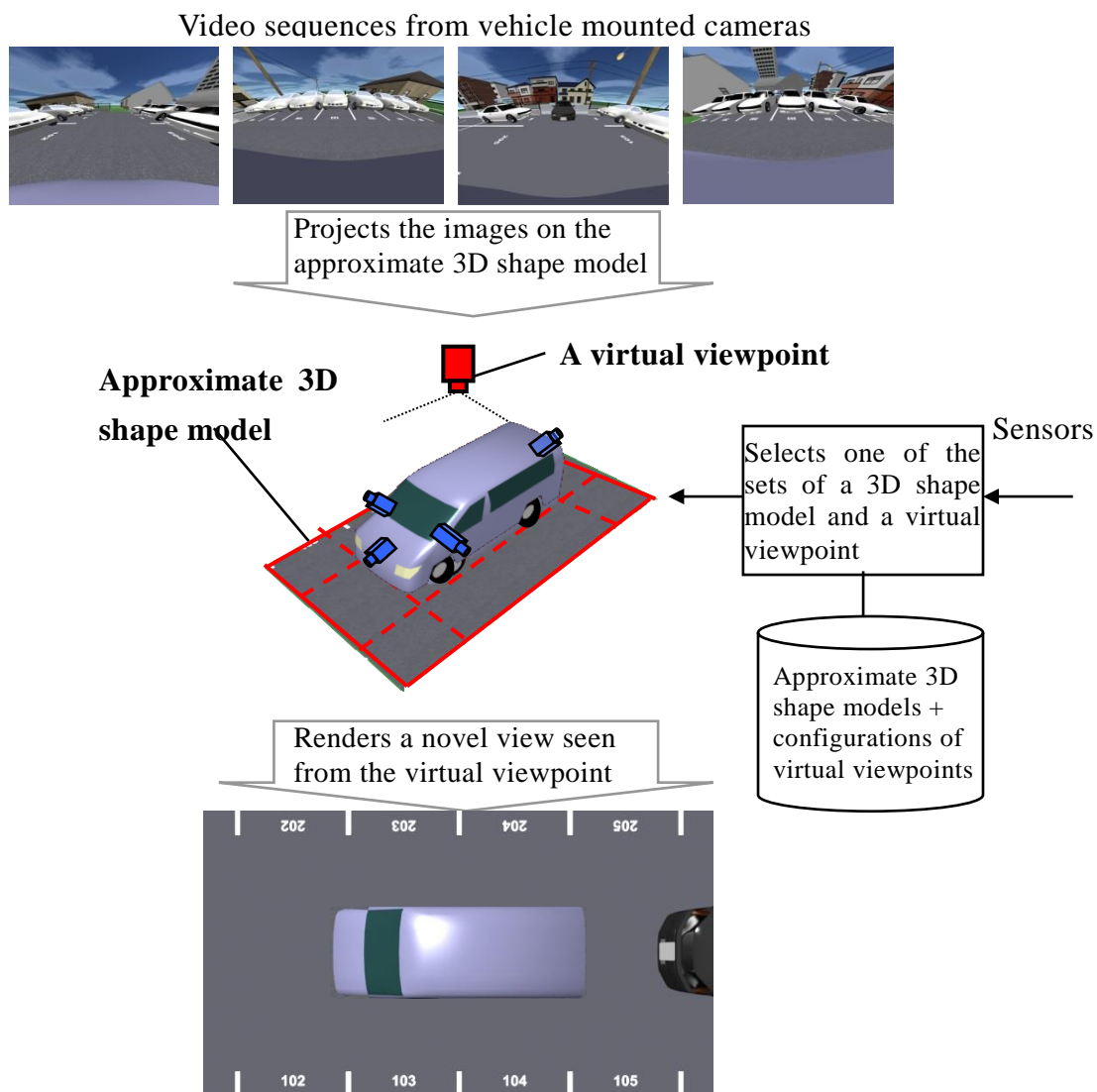


Figure 10 Overview of the proposed method.

車を基準とした座標系と、車載カメラ、仮想カメラ、3次元形状モデルの配置を Figure 11 に示す。仮想カメラから見た車の周りを俯瞰する画像を生成するためには、仮想カメラの画素座標 (u_v, v_v) に対応する、車載カメラの画素座標 (u_c, v_c) を求めればよい。車載カメラ画像の画素座標 (u_c, v_c) は、全て形状モデル上に存在する（車両座標系

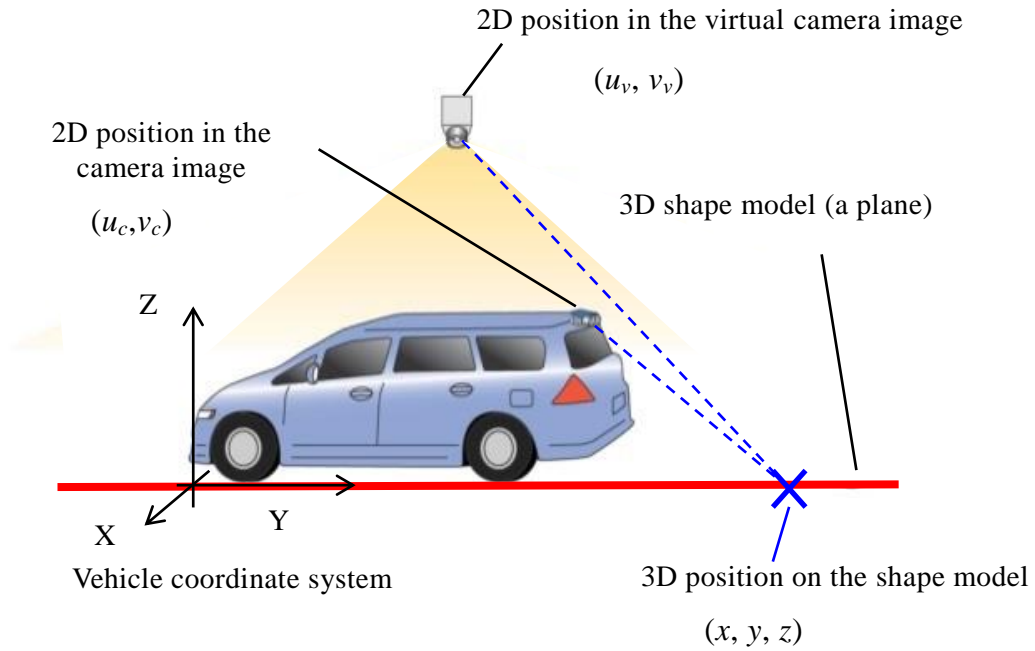


Figure 11 Geometry model for virtual view generation.

の座標値 (x, y, z) と仮定すると、式(2-1)が成り立つ。そのため、仮想カメラの画素座標 (u_v, v_v) に対応する、車載カメラの画素座標 (u_c, v_c) を求めることができる。

$$\begin{aligned}
 (u_v, v_v) &= P_v(x_v, y_v, z_v) \\
 (x_v, y_v, z_v, 1)^T &= \mathbf{F}_v(x, y, z, 1)^T \\
 (u_c, v_c) &= P_c(x_c, y_c, z_c) \\
 (x_c, y_c, z_c, 1)^T &= \mathbf{F}_c(x, y, z, 1)^T \\
 \mathbf{F}_v &= \begin{pmatrix} \mathbf{R}_v & \mathbf{T}_v \\ 0 & 1 \end{pmatrix}, \mathbf{F}_c = \begin{pmatrix} \mathbf{R}_c & \mathbf{T}_c \\ 0 & 1 \end{pmatrix}
 \end{aligned} \tag{2-1}$$

ここで P_v, P_c は、それぞれ仮想カメラと車載カメラのカメラモデル(内部パラメタ)を表す射影関数であり、各カメラ座標系の3次元座標と画素座標の関係を表す。 $\mathbf{F}_v, \mathbf{F}_c$ は、それぞれ仮想カメラと車載カメラの車両座標系における位置姿勢(外部パラメタ)を表す3次元座標変換であり、3次元の回転移動 $\mathbf{R}_v, \mathbf{R}_c$ と並進移動 $\mathbf{T}_v, \mathbf{T}_c$ を用いて、車両座標系の座標値と各カメラ座標系の座標値との関係を表す。 $(x_v, y_v, z_v), (x_c, y_c, z_c)$ は、それぞれ車両座標系の座標値 (x, y, z) の仮想カメラ座標系と車載カメラ座標系の3次元座標値である。車載カメラの射影関数 P_c および座標変換 \mathbf{F}_c は、カメ

ラ校正法により事前に計測しておく。仮想カメラの射影関数 P_v と座標変換 \mathbf{F}_v は、設計値であり、設計者が全周囲画像の構図を決定し、その視野範囲や視点位置を基に、事前に設計しておく。仮想カメラが透視射影の場合は式(2-2)、車載カメラが透視射影の場合は式(2-3)で、それぞれ表すことができる。本章では、特に断りが無い仮想カメラは透視射影カメラモデルとする。

$$(u_v, v_v) = P_v(x_v, y_v, z_v) = \left(\frac{f_p}{z_v} x_v, \frac{f_p}{z_v} y_v \right) \quad (2-2)$$

$$(u_c, v_c) = P_c(x_c, y_c, z_c) = (r_c \cos \varphi_c, r_c \sin \varphi_c) \quad (2-3)$$

$$r_c = f_e \theta_c$$

$$(\theta_c, \varphi_c) = \left(\tan^{-1} \left(\frac{\sqrt{x_c^2 + y_c^2}}{z_c} \right), \tan^{-1} \left(\frac{y_c}{x_c} \right) \right)$$

車載カメラの画素座標 (u_c, v_c) に対応する形状モデル上の点 (x, y, z) は、画素座標 (u_c, v_c) に対応する車載カメラの視線（直線）と、形状モデルとの交点計算により求める。交点が複数ある場合は、車載カメラに最も近い点を用いる。さらに、式(2-1)を用いて、形状モデル上の点 (x, y, z) に対応する仮想カメラの画素座標 (u_v, v_v) を求める。

以上により、車載カメラの画素座標 (u_c, v_c) と、仮想カメラの画素座標 (u_v, v_v) との対応関係を得ることができる。

なお、車載カメラを車の前後左右に設置した場合、車の前左、前右、後左、後右の4方向の領域は、2台の車載カメラで重複して撮影されるため、仮想カメラの画素座標 (u_v, v_v) に、2台の車載カメラの画素に対応する（Figure 12）。この重複撮影領域の仮想視点映像生成方法として、以下の3つの方法からシーンに応じて選択する。

- (i) 左または右の車載カメラ映像を使用する
- (ii) 前または後の車載カメラ映像を使用する
- (iii) 2台の車載カメラ映像を半透明合成（画素値を平均化）する

以降、特に断りが無い限り、(i)左または右の車載カメラ映像を使用する。

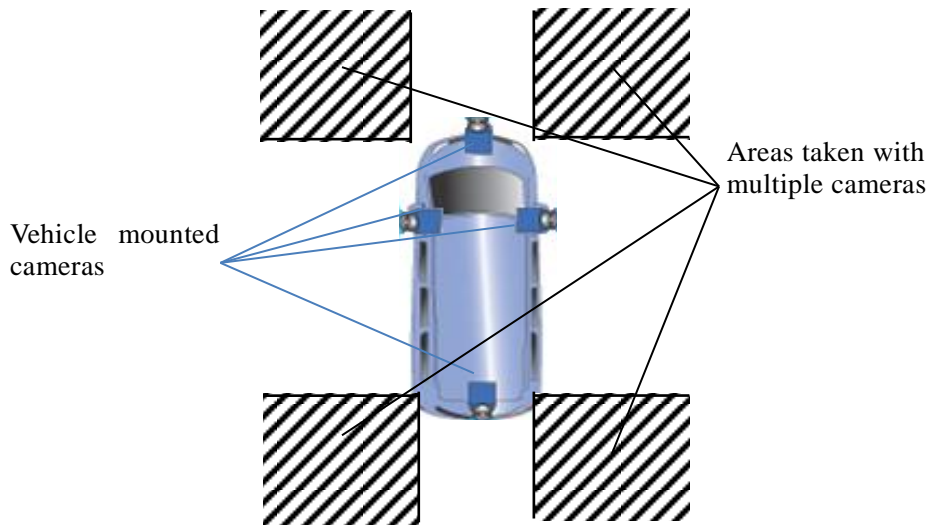


Figure 12 Areas taken with multiple cameras.

2.2.3. 近似 3 次元形状モデル

平面，平面の組み合わせ，または半球のような単純な幾何学的モデルによって表される車両の周囲の近似 3 次元形状モデルを Figure 13 に示す。

1 平面モデル Figure 13a は，車両の周辺が典型的には路面平面であるため，駐車場の周辺の 3 次元形状を近似するのに適している．球面モデル Figure 13b は，車両の周囲に物体が存在するため，特に車両の発進時に，駐車スペースの遠方の 3 次元形状を近似するのに適している．

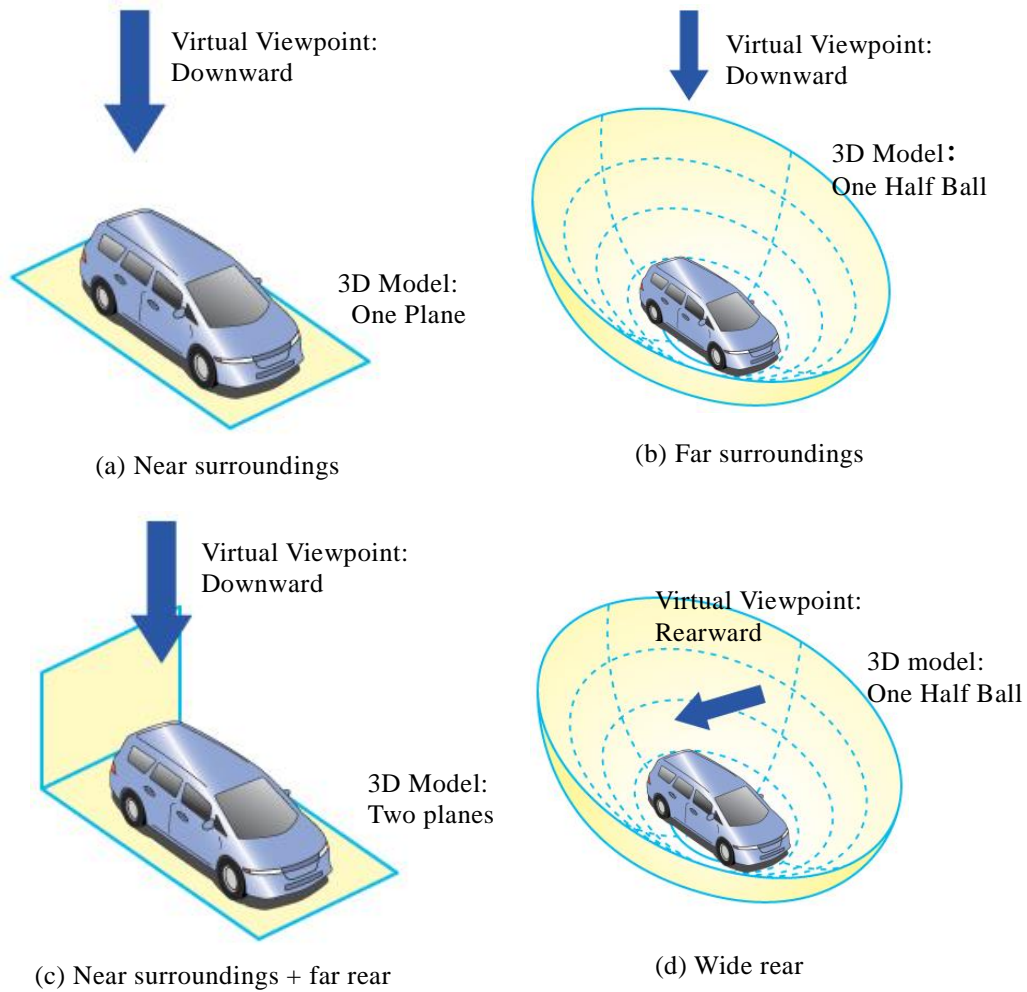


Figure 13 Configurations of approximate 3D models and virtual viewpoints.

2.2.4. リアルタイム仮想視点映像生成ハードウェア

路上評価のためのリアルタイム仮想視点映像生成ハードウェアを開発した(Figure 14). このハードウェアは、最大8台のカメラ映像を入力して、リアルタイムで仮想視点映像を生成することができる。提案法は、予め定めた近似3次元形状モデルを用いるため、入力となるカメラ画像から仮想視点映像を生成するための幾何変換を事前に計算しておくことができるため、実行時の計算コストを削減することができる。本ハードウェアには、生成画像内のカメラ境界における不連続性のアーチファクトを低減するための、半透明合成および輝度調整の機能も有している。

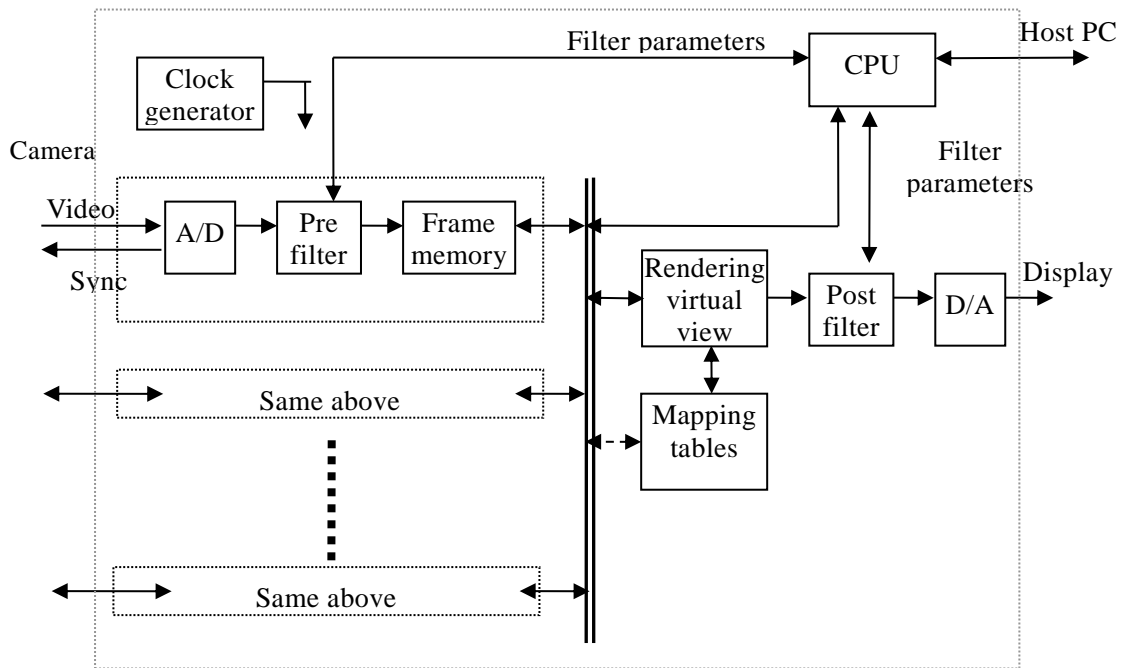


Figure 14 Block diagram of the image synthesis hardware.

半透明合成：

仮想視点映像におけるカメラ境界の不連続を低減するために、2カメラの境界近傍の領域において、半透明合成する手法を採用した。境界近傍の適切な幅で、2つのカメラ画像を混合する比率を徐々に変えることで、境界の不連続性を低減する。

輝度調整：

同様に、仮想視点映像におけるカメラ境界における不連続を低減するために、各カメラの輝度を調整する。各画像は、境界付近での平均輝度と輝度分布が一致するように調整する。

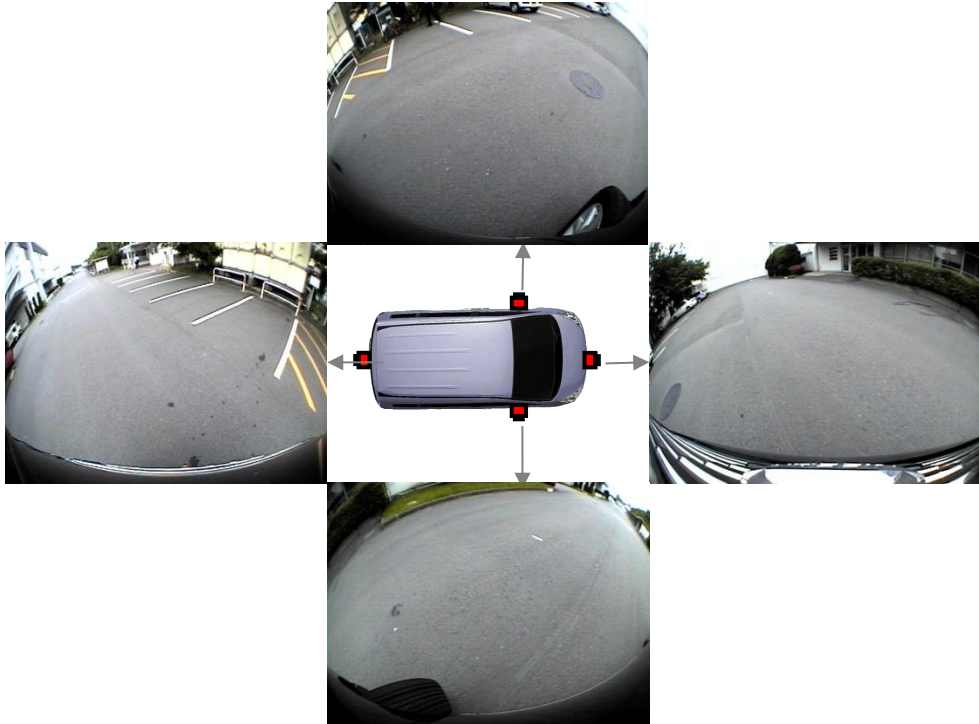


Figure 15 Configuration of the experimental vehicle with 4 cameras.

2.3. 評価実験

提案手法の有効性を検証するために、実時間で仮想視点映像を生成できる実験車両を用いて評価実験を行った。

2.3.1. 実験車両のシステム構成

実験のため、車種（セダン、ワンボックス、アルミバン）およびカメラ台数（4～8台）の異なる、5台の実験車両を構築した。実験車両のカメラ構成と撮影画像の一例を Figure 15 に示す。各カメラのレンズは、水平 160 度および垂直 110 度の広い FOV レンズを使用した。

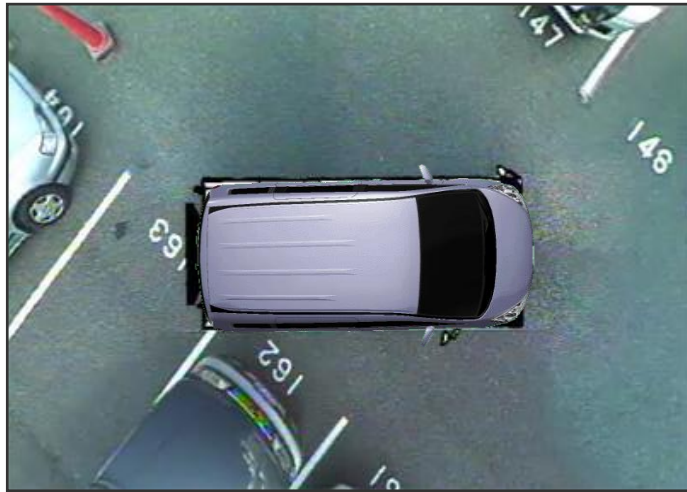
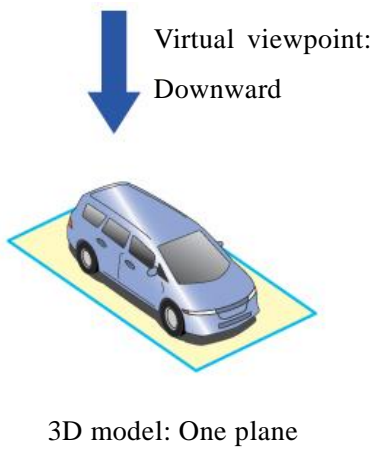
2.3.2. リアルタイムシステムによる実験結果

リアルタイムハードウェアを搭載した実験車両を用いて路上実験を行った。Figure 16 に、近似 3 次元形状モデルと視点位置、生成された仮想視点映像の例を示す。システムは、運転者によって車両のエンジンが始動されると Figure 16b の半球の近似 3 次元形状モデルを、運転者によってギヤが後退設定されると Figure 16a の 1 平面の近似 3 次元形状モデルを、運転者のスイッチによる選択で Figure 16c または Figure 16d の近似 3 次元形状モデルを、それぞれ選択して映像を生成する。

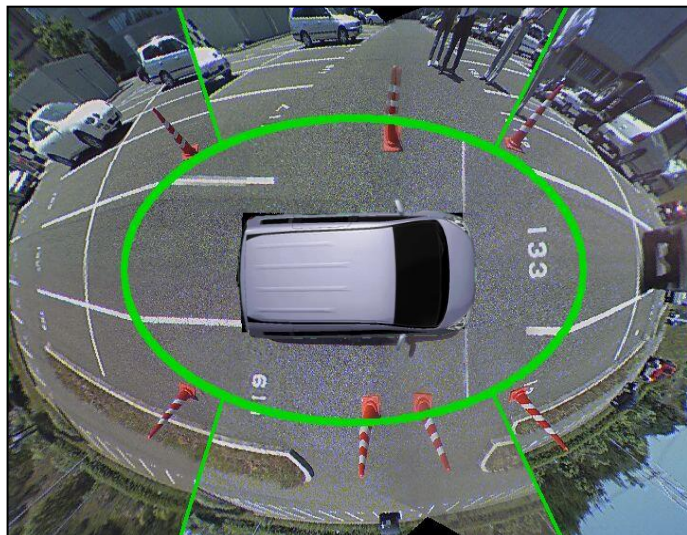
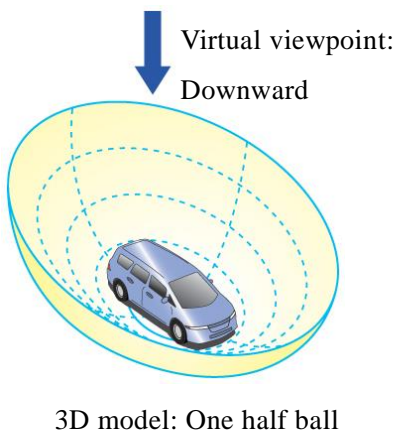
Figure 16a は、典型的な駐車場または道路の近くの周辺を近似するのに適した、1 つの平面を使用して生成された映像を示す。Figure 16b は、一般的に遠方の 3 次元形状を近似するのに適した、球面モデルを用いて生成された映像を示す。この映像は、車の遠くを含む全周囲の物の有無の把握に適している。Figure 16c は、道路の近くの周辺および遠方の後方を近似するのに適した 2 つの平面モデルを使用して生成された映像を示す。Figure 16d は、ブラインドコーナーで後方または前方周囲を近似するのに適した、傾斜した視点と球形モデルを使用して生成された画像を示す。

これらの実験車両を用いた実験結果により、近似 3 次元形状モデルを用いて仮想視点映像を生成する本手法は、既存の手法では 3 次元形状の推定が困難な、透明物、鏡面反射物、低反射物が存在する駐車場や道路においても、仮想視点映像を生成できるという点で、ロバストであることを示した。

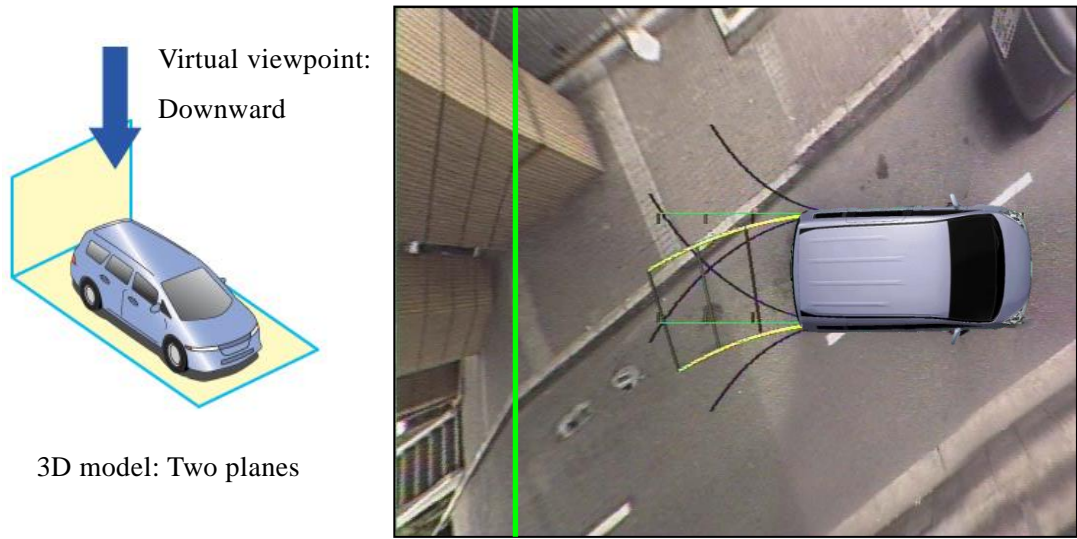
提案手法により生成した、車の周りを上空の 1 つの視点から俯瞰した映像(Figure 16a-c)は、車の近くに死角がないため、物の有無を把握し易い。さらに、車の周りを平面の近似 3 次元形状モデルを用いた映像(Figure 16a,c)は、路面上の直線が映像上でも直線であることと、実空間中の距離と映像上の距離が比例するため、車と物との距離や物の方向がわかり易い、という点で、従来の車載カメラシステムより優れている。



(a) Near surroundings view



(b) Far surroundings view



(c) Near surroundings and far rear view



(d) Wide rear view

Figure 16 Generated Virtual Viewpoint Video.

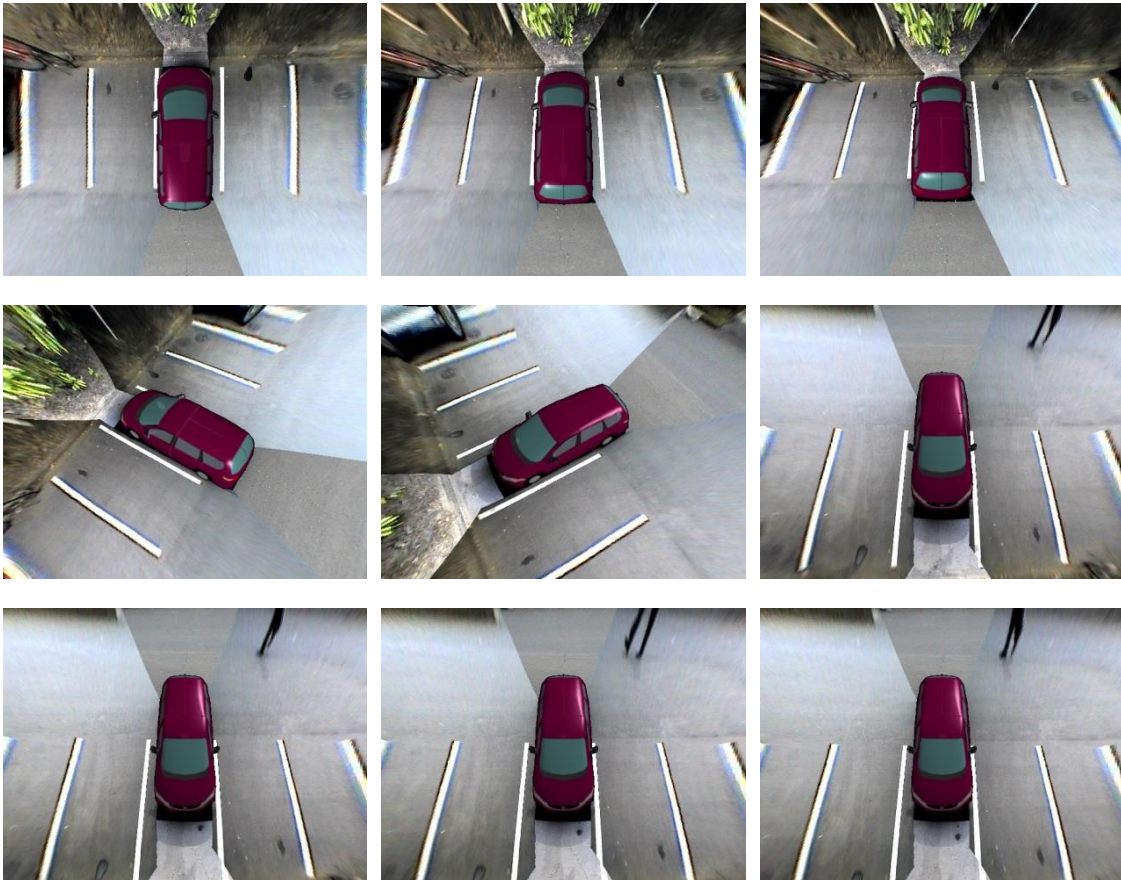


Figure 17 Generated Virtual Viewpoint Video from different viewpoints

This image sequences are generated by a software simulator using video sequences captured and recorded by using experimental vehicle.

2.3.3. 考察

本章で提案した近似3次元形状モデルを用いた仮想視点映像生成手法の効果は、2.3.2節の実車評価において、ガラス等の透明物、金属等の鏡面反射物、タイヤ等の低反射物などの3次元形状の推定ができない被写体が存在する駐車場や道路で、仮想視点映像を生成できることとして確認できた。平面形状モデルを用いて生成した映像は、死角がなく、車の周りの物の有無がわかり易い。また、路面上の直線が画像上で直線となり、実空間中の距離が画像上の距離と比例することから、車と物との距離や方向がわかり易い。半球形状モデルを用いて生成した映像は、視野が広く、遠くの物の有無がわかり易い。

提案法で用いる近似 3 次元形状モデルと実際の被写体の 3 次元形状とは、差異がある場合がある。この 3 次元形状の差異に起因して、提案法で生成した仮想視点映像に以下の 2 つの影響が生じる

- ① 3 次元形状の差異がある被写体について、仮想視点映像上の被写体の位置と、仮想カメラの視点から実際に撮影される映像上の位置に、差異が生じる。
- ② ①の仮想視点映像上の位置の差異は、車載カメラの位置によって異なる。そのため、複数の車載カメラの生成した仮想視点映像中のカメラ境界部において、像が不連続になる。

前者①に関して、近似 3 次元形状モデルと路面の 3 次元形状との間に大きな差異がある場合の影響を評価した。Figure 18 は、近似 3 次元形状モデルとして路面位置の 1 つの平面モデルを用いた場合の仮想視点映像の例である。この平面モデルを用いた仮想視点映像では、平面モデル表面上の位置を基準に、車載カメラ位置を中心とした放射線状に、平面モデルから高いほど車載カメラから遠くに、平面モデルより低いほど車載カメラの近くに表示される。Figure 18a は、駐車スペースの後方に高さ 2m の壁があるシーンであり、Figure 18b は、駐車スペースの後方に 3m の段差があり、段差の後方に別の駐車場があるシーンである。

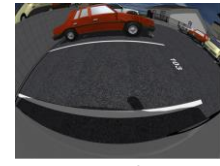
Figure 18c は、駐車スペースの後方に高さ 2m の壁があるシーン(Figure 18a)における仮想視点映像である。この映像から、駐車スペースの後方に壁があることや、路面と壁面の境界位置を確認できる。Figure 18d は、駐車スペースの後方に 3m の段差があるシーン (Figure 18b) における仮想視点映像である。この映像から、駐車スペース後方に段差があることや、路面と段差の境界位置を確認できる。一方で、段差の下にある駐車場までの距離を、Figure 18d の仮想視点映像から正確に把握することは困難であるが、車両周辺監視用の車載カメラシステムの用途においては、車を中心とした路面上の走行可能な範囲にある、壁や段差の位置を把握できればよく、段差の下にある駐車場までの距離を把握が困難であったとしても実用上の問題は無い。

以上の Figure 18 の例から、提案法は、近似 3 次元形状モデルと路面の 3 次元形状との間に大きな差異がある場合であっても、車両周辺監視用の車載カメラシステムとして、実用上の問題のない仮想視点映像を生成できることを示した。

後者②の、仮想視点映像中のカメラ境界部において、像が不連続になる影響については、2.2.2 節で示した、複数の車載カメラで重複撮影領域の映像生成方法とあわ



Front



Left



Rear

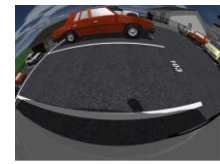


Right

(a) Parking scene with a wall of 2m in height behind the parking space.



Front



Left



Rear



Right

(b) Parking scene with a cliff of 3m in height behind the parking space.



(c) Generated virtual viewpoint video with a wall of 2m in height behind the parking space.



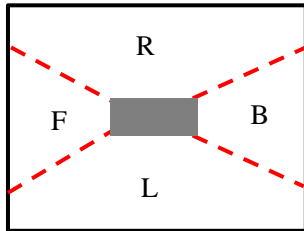
(d) Generated virtual viewpoint video with a cliff of 3m in height behind the parking space.

Figure 18 Example of virtual viewpoint video with a large difference between approximate 3D shape model and actual shape.

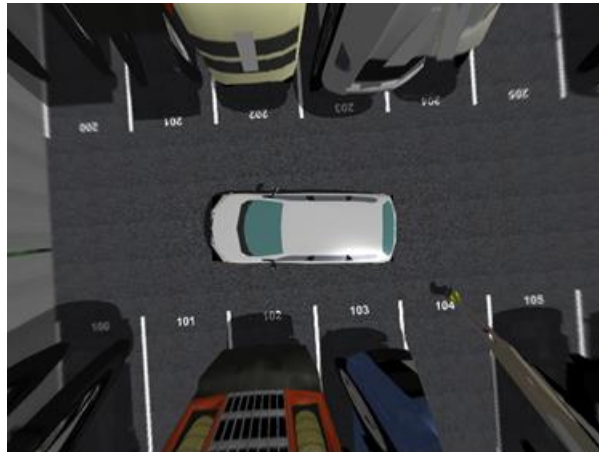
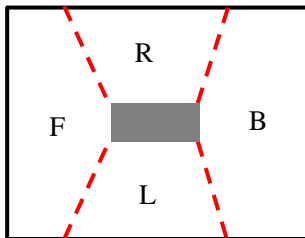
せてその影響を評価した。Figure 19 は、後カメラと左カメラの重複撮影領域に人がいるシーンにおいて、近似 3 次元形状モデルとして路面位置の 1 つの平面モデルを用いて生成した仮想視点映像である。Figure 19a-c は、重複撮影領域における画素値の決定方法が異なり、Figure 19a および b は重複撮影領域をそれぞれ左右および前後のカメラの映像を用いて生成した仮想視点映像であり、Figure 19c は 2 カメラの映像の画素値を平均化した値を用いて生成した仮想視点映像である。

Figure 19a-c は、いずれも、路面および路面上の白線や駐車スペースの数値、影、など、路面上の被写体は、仮想視点映像中で同じ位置にある。また、車の後方の歩行者についても、平面モデルと接する足元は、仮想視点映像中で同じ位置になっている。そのためいずれの方法で生成した仮想視点映像であっても、歩行者の足元の有無や、その距離や方向を把握できる。一方、Figure 19b では、歩行者の上半身まで表示されているのに対し、Figure 19a では、足元しか表示されておらず、歩行者の表示範囲に差がある。これは、歩行者の 3 次元形状が、近似 3 次元形状モデルである平面モデルと一致しないことに起因し、後カメラと左カメラの 2 つのカメラで撮影した映像から、それぞれ仮想視点映像を作成した場合に、平面から高さのある歩行者が、それぞれの車載カメラを中心とした放射線状に、カメラから遠くに離れた映像になり、この 2 つの映像から仮想視点映像を生成すると、歩行者が 2 つのカメラ映像の境界部にある場合に、足元より上の部分が仮想視点映像上に表示されなくなるためである。Figure 19a のように、仮想視点映像中のカメラ境界部において、被写体の一部が表示されない現象は、Figure 19b に示したように、カメラ境界部の位置が異なる仮想視点映像生成手法を用いることによって、歩行者の表示される範囲が増加するという点で改善できる場合がある。また、Figure 19c のように、重複撮影領域の 2 つのカメラ映像の画素値を平均化した値を用いて仮想視点生成することでも、改善できる場合がある。

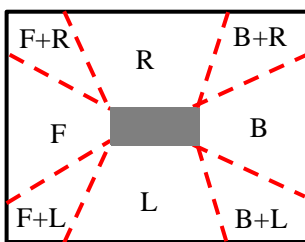
このほか、Figure 19a と Figure 19b を一定時間間隔で切り替える表示方法によっても改善できる。また、仮想視点映像中のカメラ境界部に境界線を描画することで、仮想視点映像のカメラ境界部が不連続であることと、足元しか表示されていない場合に、その原因は、カメラ境界が不連続であることを、運転者に把握させる効果が期待できる。



(a) An image generated using left and right camera images in overlapping area.



(b) An image generated using front and back camera images in overlapping area.



(c) An image generated by averaging in overlapping area

Figure 19 Generated images with different method in overlapping area.

2.4. 結言

本章では，車載カメラで撮影された映像から仮想視点映像を生成するロバストな仮想視点映像生成手法について説明した．駐車場や道路などの実環境で仮想視点を生成するために，提案法は，近似3次元形状モデルを用いて仮想視点を生成する．近似3次元形状モデルは予め決定されているので，3次元形状を推定できない被写体がある場合でも仮想視点映像を生成できる．実験結果では，提案法が，ガラス等の透明物や金属等の鏡面反射物，タイヤ等の低反射物などの，3次元形状を推定できない被写体が存在する駐車場や道路において，仮想視点映像を生成できることを示した．また，提案法によって生成した上空の1視点から俯瞰した映像は，死角がなく，車の周りの物の有無がわかりやすいこと，車と物との距離や物の方向がわかり易いことを示した．提案手法を用いた車載カメラシステムにおいて，近似3次元形状モデルおよび仮想視点の位置は，運転手に運転操作に応じて選択されるので，駐車，車両の後退，車両の発進またはブラインドコーナーを通過するなどの運転シーンに適した視野範囲や視点の仮想視点映像を，運転者に提供することができる．

しかしながら，第1章で述べたように，複数の車載カメラ映像から画像処理によって生成された仮想視点映像は，180度以上の広視野角映像における像の歪み (Figure 6)，および，仮想視点映像における局所的な低解像度 (Figure 7) の2つアーチファクト発生課題があり，これらの課題に対処する研究はなかった．そこで，3章，4章では仮想視点映像生成におけるこれらの2つの課題の解決に取り組む．

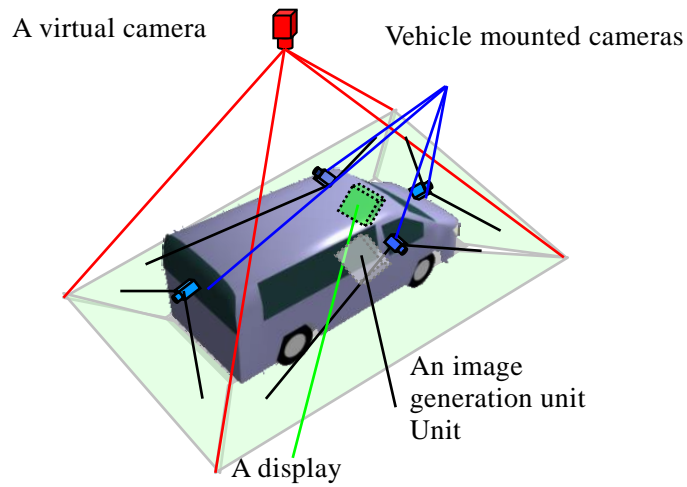
3. 複合射影カメラモデルを用いた車載カメラシステムのための低歪み全周囲画像生成手法

本章では，車載カメラシステムのための低歪み全周囲画像生成手法を提案する．駐車時の運転支援のためには，視野が広く，かつ，車と近くの物との距離や方向を把握しやすい画像が望ましい．しかし，広視野の全周囲画像を生成する従来法は，魚眼レンズのカメラモデルや半球状の形状モデルを用いるため，像が歪み，路面上の直線が画像上では曲線になり，車と物との距離と画像上の距離が比例しなくなる．その結果，車と物との距離や方向がわかりにくいという課題があった．提案法では，画像生成のカメラモデルとして，透視射影と等距離射影の2つの射影モデルを組み合わせた複合射影モデルを用いる．これにより，車の近くは透視射影を用いることで歪みが少ない，遠くは等距離射影を用いることで視野が広い，画像を生成することができる．さらに提案法では，車の周りの形状モデルとして，半円柱形状モデルを，複合射影カメラモデルと組み合わせて用いることで，車の近くの歪みを増やすことなく，水平線より上を含む広い視野の画像を生成することができる．実験では，駐車シーンの画像から，視野が広く，車の近くの歪みが少ない全周囲画像を生成できることを示す[39][40]．

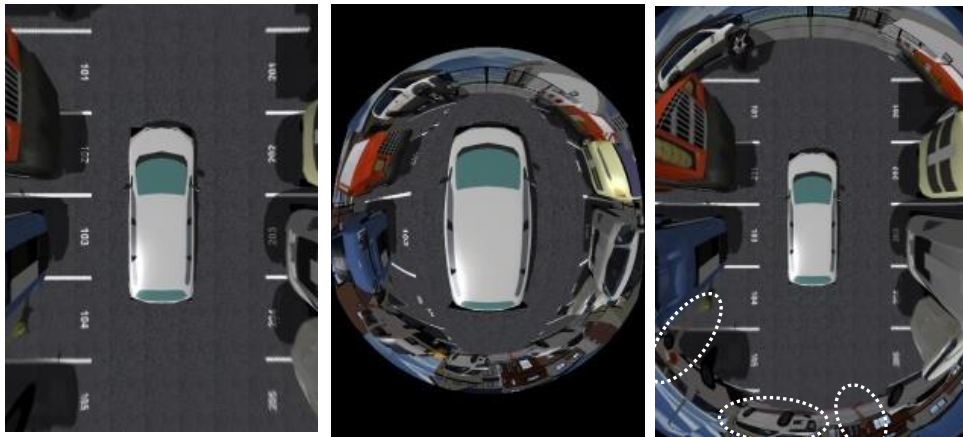
3.1. 序言

駐車時の事故防止や運転者の負荷軽減を目的として，運転者の視認を支援する車載カメラシステムが数多く提案され[35][36][51]-[61]，一部は実用化されている．これら車載カメラシステムは，車体に設置した車載カメラで車の周りを撮影し，撮影した画像を編集加工して，車室内に設置したディスプレイに表示する．このシステムを用いると，運転者は，直接見ることができない位置にある歩行者や車などを，ディスプレイに表示された画像を介して見ることができる．そのため，死角領域にある物との接触事故を防止する効果が期待できる．

駐車時の事故防止や運転者の負荷軽減を実現する別のシステムとして，自動制御によって運転者の運転操作を支援する自動駐車システムが提案されている



(a) A typical configuration of vehicle camera systems



(b) Perspective projection on a plane

(c) Equidistance projection on a half ball model

(d) Composite projection on a halfpipe (proposed)

Figure 20 A configuration of vehicle camera system and examples of generated images.

(a) 4 台の車載カメラ画像を用いて、車の周りを俯瞰する画像を生成。(b) 透視射影。視野が狭く遠くの物が映らない。(c) 等距離射影。直線が画像上で曲線になる。(d) 複合射影 (提案法) では、視野が広く遠くの物が映り、車の近くは直線が維持される。

[62],[63]. 駐車位置までの経路判断と運転操作とをシステムが実行するため、運転者の負荷を軽減できる。しかしながら、システムによる自動運転操作中も、事故防止

のために運転者は車の周りを見ておく必要がある。そのため、自動駐車システムを搭載した車においても、運転者が車の周りを見ることが出来る車載カメラシステムは有効である。

車載カメラシステムの代表例として、車の後部に設置した 1 台のカメラで車の後方を撮影し、その画像を運転者に提示するシステムがある。しかし、このシステムでは、運転者は車の左右の死角領域を見ることができない。そこで近年では、Figure 20a のように、複数台のカメラ(vehicle mounted cameras)を車に設置して車の周りを死角なく撮影し、撮影した画像から車の周りを上空の仮想カメラ(a virtual camera) から俯瞰した全周囲画像を生成し(an image generation unit), 生成した画像を車内のディスプレイ(a display)に表示して運転者に見せる車載カメラシステムが実用化されている [35][36][57][58].

これらの全周囲画像を生成する車載カメラシステムのための画像生成手法として、画像生成のカメラモデルと車の周りの形状モデルに、それぞれ透視射影カメラモデルと平面形状モデルを用いる手法[35][36][57]や、等距離射影カメラモデルと半球形状モデルを用いる手法[36][58]が提案されている。前者の手法で生成した画像は、Figure 20b のように、車の周りの死角領域を含む全周囲の画像を見ることが出来るように、車と物の位置がわかり易い。しかし、視野が狭く、原理的に水平線より上を表示できない。一方、後者の手法で生成した画像は、Figure 20c のように、広い視野の画像を生成することができる。しかし、路面上の直線が画像上では曲線になり、車と物との実際の距離と画像上の距離が比例しない。その結果、車と物との距離や方向がわかりにくい。

以上の背景から、本論文では、Figure 20(d)に示したように、水平線より上を含む広い視野と、路面上の直線が画像上で直線になる、低歪みを、同時に満たす画像を生成する手法を提案する。

3.2. 関連研究

全周囲画像を生成する手法は、超広角カメラを用いる方法、3次元モデルを用いた画像再構成手法、の2つに大別できる。

全周囲画像を生成する代表的な手法は、1台で360度の水平視野を撮影できる超広角カメラを用いる手法がある[56]。しかし、車体近傍の全周囲を1台のカメラでは

撮影することができない。

複数のカメラで撮影した画像に，コンピュータビジョン技術を適用することで，撮影画像から被写体の3次元形状を算出し，得られた3次元形状と撮影画像にコンピュータグラフィックスの技術を適用することで，車の周りを任意視点から見た画像を生成する手法が提案されている[51]．この手法は，ガラスなどの光の透過率が高い被写体や，金属などの反射率が高い被写体は，3次元形状の算出ができず，これらの被写体の画像を生成できない。

複数のカメラで撮影した画像間の対応関係を推定し，推定した対応点で画像が接合するように画像を変形することで，複数の画像をつなぎ合わせた画像を生成する手法が提案されている[52][53]．この手法は，3次元形状に誤差が含まれる場合でも，2画像の接合部のずれがない画像を生成できる．しかし，ガラスなどの光の透過率が高い被写体などは，2画像間の対応関係を正しく推定できないため，画像を生成できない。

3次元形状を推定する代わりに，車の周りの形状を近似した形状モデルを用いる手法が提案されている．画像生成のカメラモデルと車の周りの形状モデルに，それぞれ透視射影カメラモデルと平面形状モデルを用いる手法や，等距離射影カメラモデルと半球形状モデルを用いる手法が提案されている[35][36][57][58]．前述したように，前者は，視野が狭く，原理的に水平線より上を表示できない．後者は，直線が画像上では曲線になり，車と物との実際の距離と画像上の距離が比例しなくなる．全周囲画像生成において，広い視野の画像と，直線が歪まない画像の生成を目的として，数多くの手法が提案されているが，2つを両立した手法はない．被写体の3次元形状に着目した手法[35][36][51][52][53][57][58]のほか，画像の歪みを補正する手法[59],[60]や，ガイド線を描画する方法[61]も提案されているが，本質的な解決策は提案されていない。

以上の背景から，本論文では，Figure 20 (d)に示したように，広い視野と，直線が画像上で直線となる，低歪みを，同時に満たす画像を生成する手法を提案する．

3.3. 複合射影カメラモデルを用いた全周囲画像生成

本論文では，車の周りの形状を近似した形状モデルを用いる全周囲画像生成手法において，仮想カメラのモデルとして複合射影カメラモデルを，車の周りの形状モ

デルとして半円柱形状モデルを、それぞれ用いる全周囲画像生成手法を提案する。複合射影カメラモデルは、透視射影と等距離射影の2つを組み合わせたカメラモデルである。入射角が所定値より小さい画像中央付近の領域は透視射影、その周辺は等距離射影とすることで、画像中央付近は直線が画像上で直線になり、かつ、180度以上の広い視野の画像を生成できる。

半円柱形状モデルは、平面と半球を組み合わせた形状モデルである。車の近くの平面と、その周り半球とを組み合わせた半円柱形状にすることで、車の近くでは直線が画像上で直線になり、かつ、水平線より上の視野を含む画像を生成できる。

3.3.1. 全周囲画像生成手法の概要

本論文で前提とする全周囲画像の生成手法は、仮想カメラのカメラモデルと近似3次元形状モデル（以降、形状モデル）を除いて2.2.2節と同様である。

車の周辺の形状モデルは、[35][36][57][58]で用いられている、平面形状モデル、半球形状モデルのほか、提案法では、後述する、半円柱形状モデルを用いる。仮想カメラの射影関数 P_v は、透視射影カメラモデル、等距離射影カメラモデル、のほか、提案法では、後述する複合射影カメラモデルを用いる。透視射影カメラモデルおよび等距離射影カメラモデルの、画素座標 (u_v, v_v) とカメラ座標系の座標値 (x_v, y_v, z_v) との関係を示す。

透視射影カメラモデル：

$$(u_v, v_v) = (r_p \cos \varphi, r_p \sin \varphi) \quad (3-1)$$

$$r_p = f_p \tan \theta$$

$$(\theta, \varphi) = \left(\tan^{-1} \left(\frac{\sqrt{x_v^2 + y_v^2}}{z_v} \right), \tan^{-1} \left(\frac{y_v}{x_v} \right) \right)$$

等距離射影カメラモデル：

$$(u_v, v_v) = (r_e \cos \varphi, r_e \sin \varphi) \quad (3-2)$$

$$r_e = f_e \theta$$

$$(\theta, \varphi) = \left(\tan^{-1} \left(\frac{\sqrt{x_v^2 + y_v^2}}{z_v} \right), \tan^{-1} \left(\frac{y_v}{x_v} \right) \right)$$

式(3-1)、式(3-2)において、 r_p, r_e は像高であり、画素座標系における光軸中心からの

距離である． (θ, φ) はカメラ座標系における，点 (x_v, y_v, z_v) の方向を表す2つの角度であり， θ はZ軸（光軸）とのなす角度， φ はZ軸中心の回転角である． f_p, f_e はそれぞれ透視射影，等距離射影における焦点距離である．以降の説明では，Z軸(光軸)との角度 θ を入射角と呼ぶ．

3.3.2. 複合射影カメラモデル

本論文で提案する複合射影カメラモデルについて説明する．複合射影カメラモデルにおける，入射角 θ と像高 r_c の関係を，式(3-3)に示す．

複合射影カメラモデル（提案法）

$$(u_v, v_v) = (r_c \cos \varphi, r_c \sin \varphi) \quad (3-3)$$

$$r_c = \begin{cases} f_p \tan \theta & \text{if } \theta \leq \theta_c \\ f_e(\theta - \theta_c) + f_p \tan \theta_c & \text{otherwise} \end{cases}$$

$$(\theta, \varphi) = (\tan^{-1}\left(\frac{\sqrt{x_v^2 + y_v^2}}{z_v}\right), \tan^{-1}\left(\frac{y_v}{x_v}\right))$$

f_p, f_e はそれぞれ透視射影，等距離射影における焦点距離である． θ_c は，2つの射影モデルの境界の入射角である．式(3-3)は，入射角 θ が θ_c 以下の場合に透視射影，入射角 θ が θ_c より大きい場合に等距離射影であることを示す． f_p, f_e および θ_c は，任意に設定できる．

提案法では，複合射影カメラモデルを，前節で説明した全周囲画像生成の仮想カメラに用いる．複合射影カメラモデルの仮想カメラを，車の上空から見下ろす位置に配置した場合，入射角 θ が θ_c 以内の車の近傍は，透視射影によって画像を生成するため，3次元空間中の直線は画像上でも直線になる．また，入射角 θ が θ_c より大きく車から離れた範囲は，等距離射影によって画像を生成するため，180度以上の広い視野の画像を生成できる．

複合射影カメラモデルで生成した画像は，透視射影と等距離射影の2つの射影モデルで，入射角 θ と像高 r_c の比率（拡大率）が異なる．そのため，2つの射影モデルの境界 θ_c において拡大率が不連続であることによる，像の歪みが生じる．この像の歪みは，射影モデルの境界 θ_c における，2つの射影モデルの像高 r_c を入射角 θ で微分した値が近いほど，小さくなる．すなわち， $f_p / \cos 2\theta_c$ と f_e との差が小さくなるよう， f_p, f_e, θ_c を設計することで，歪みを低減することができる．

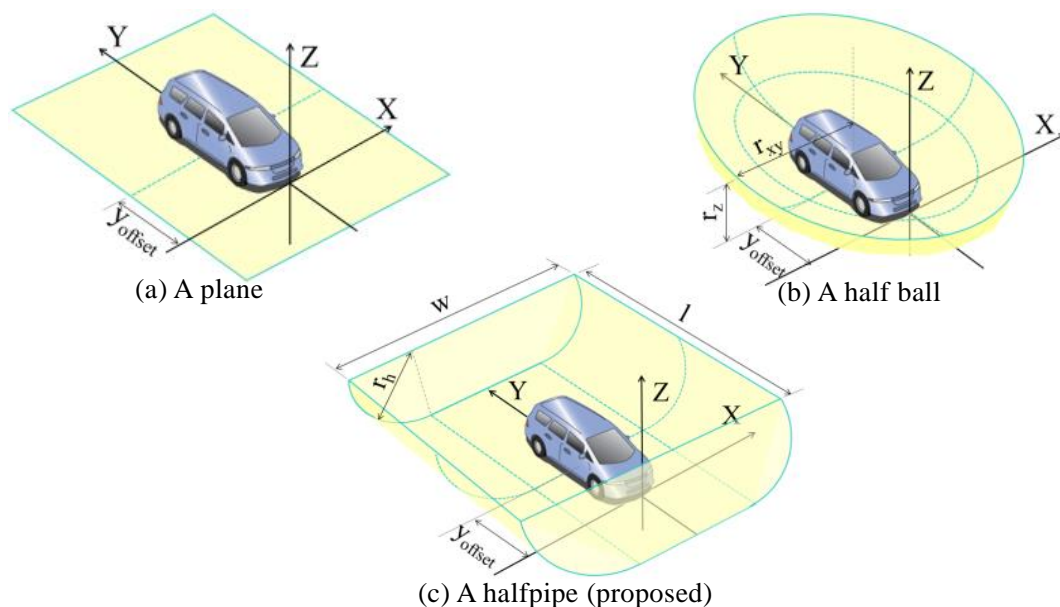


Figure 21 3D shape models of the proposed method.

3.3.3. 半円柱形状モデル

半円柱の形状モデルを Figure 21c に示す．車の前後方向に沿った半円柱であり，半円柱の底面は平面である．半円柱形状モデルの形状と車両座標系における位置は，幅 w ，長さ l ，円柱部半径 r_h ，Y 軸方向のオフセット y_{offset} で指定する．車の周囲を所定の高さで覆うため，水平線より上を含む広い視野の画像を生成できる．

従来の平面形状モデルを Figure 21a, 半球形状モデルを Figure 21b に示す．平面形状モデルは $z=0$ (一定) の平面である．半球形状モデルの形状と車両座標系における位置は，水平方向半径 r_{xy} , 垂直方向半径 r_z , Y 軸方向のオフセット y_{offset} で指定する．平面形状モデルは，仮想カメラの視野の水平線より下にしか平面が無いため，水平線より上を含む画像を生成できないという課題がある．半球形状モデルを用いると，水平線より上の視野を含む画像を生成できる．しかし，半球形状モデルは，直線が画像上で曲線になるという課題がある．これに対し，半円柱形状モデルでは，2つの課題が発生しない．

3.3.4. 設計指針

複合射影カメラモデルと半円柱形状モデルのパラメタの設計指針を以下に示す．

複合射影カメラモデルのパラメタ θ_c, f_p, f_e : 全周囲画像上で歪なく表示したい視野範囲を決定し、その範囲が透視射影に入るように θ_c を決める。例えば、車の前後に近接する駐車枠を歪なく表示したい場合、車の中心から車前後端までの長さ 2m、駐車枠長さ 5m、とすると、車の中心から半径 7m 以内の視野範囲を透視射影で生成すればよい。このとき仮想カメラの高さを 2m とすると、 $\theta_c = \tan^{-1}(7/2)$ が決まる。次に、全周囲画像上での透視射影の範囲を、像高 r で与えることで、式(3-3)から、焦点距離 f_p が求まる。最後に、全周囲画像の最大視野の入射角 θ と、それに対応する像高 r_c を与えることで、 f_e が求まる。なお、 θ_c での拡縮率の違いによる歪みを低減する場合、 f_p, f_e の一方を上述の方法で決め、他方を $f_p / \cos 2\theta_c = f_e$ から求める。

半円柱形状モデルのパラメタ w, l, r_h : 下記の(i)-(iv)をできるだけ満たすように、パラメタを決める。

- (i) 全周囲画像上で歪なく表示したい範囲が、半円柱形状モデルの底面の平面の幅($w-2r_h$)に入る。
- (ii) 半円柱形状モデルの幅 w は、車の周りの物の存在する幅とできるだけ一致。
- (iii) 円柱部の半径 r_h は、仮想カメラの高さ以上。
- (iv) 長さ l は、幅 w より十分大きい値を与える。

上記指針により各パラメタを設計したのち、設計値とその前後の値を用いて生成した合成画像の主観評価に基づいて、より良い設計値を求める。

3.4. 評価実験

本節では、シミュレーションによる評価実験の結果について述べる。

3.4.1. カメラモデルの比較評価

提案法の複合射影カメラモデルを用いて生成した全周囲画像と、従来法のカメラモデルを用いて生成した全周囲画像とを比較評価した。従来法のカメラモデルとして、透視射影、等距離射影の2つを用いた。

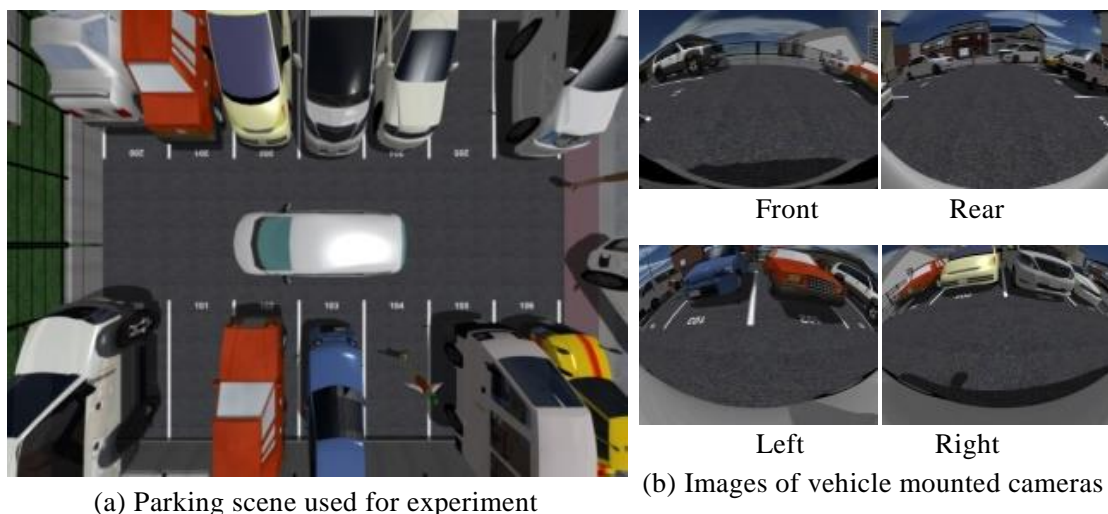


Figure 22 Camera images of parking scene used for experiment.

Table 2 Values of camera parameters used in the experiment.

表 2 カメラモデルのパラメタと実験で用いた数値

カメラモデル	パラメタと実験で用いた値
共通	画素数:640×480 [pixel] 画素サイズ:0.01×0.01 [mm]
透視射影	焦点距離 : $f_p = 0.8$ [mm] カメラ位置: $(x, y, z) = (0.0, 2.5, 2)$ [m]
等距離射影	焦点距離 : $f_e = 1.53$ [mm] カメラ位置: $(x, y, z) = (0.0, 2.5, 2)$ [m]
複合射影	焦点距離 : $f_p = 0.8$ [mm] 焦点距離 : $f_e = 1.78$ [mm] 角度閾値 : $\theta_c = 74$ [degree] カメラ位置 : $(x, y, z) = (0.0, 2.5, 2)$ [m]

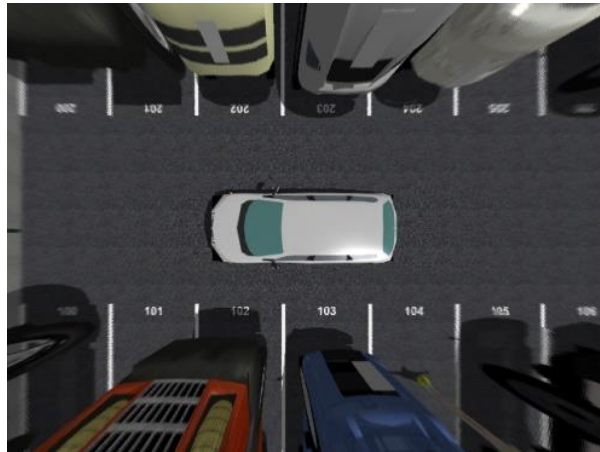
本節の実験には、コンピュータグラフィックス(CG)で生成した、車の前後左右に設置した4台のカメラの画像を用いた (Figure 22). カメラの設置位置は、前部はフロントグリル、左右はドアミラー下部、後部はナンバープレートとした。隣り合う

カメラによって重複撮影された領域の画像生成には、左右のカメラを前後のカメラより優先的に用いた。カメラモデルの比較を容易にするため、形状モデルは、Figure 21a の平面モデルを用いた。

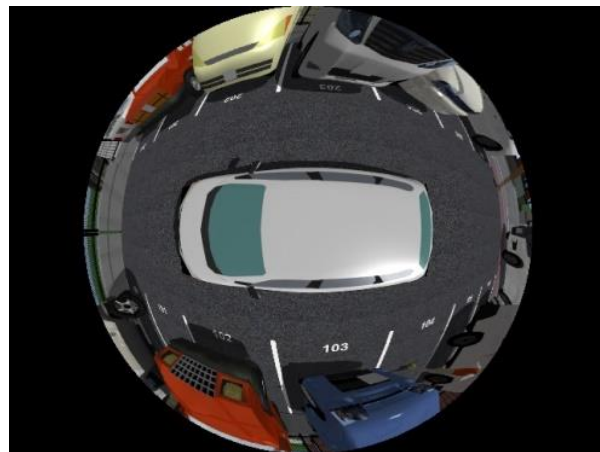
3つのカメラモデルの主要パラメタと実験で用いた数値を Table 2 に示す。透視射影カメラモデルは車両の前後方向の視野が 16m となるように。等距離射影カメラモデルは車の左右方向の視野角が 180 度となるように。それぞれ設定した。複合射影カメラモデルは、透視射影の焦点距離 f_p は同じで、車の前後方向 14m より外側が等距離射影、水平視野角が 180 度となるように、角度閾値 θ_c と焦点距離 f_e を設定した。

透視射影、等距離射影、複合射影の3つのカメラモデルを用いて生成した全周囲画像を Figure 23 に示す。透視射影カメラモデルを用いて生成した Figure 23a では、車の全周囲を見ることができうるうえに、車と物の位置がわかり易い。しかし、視野が狭く、遠方を確認することができない。等距離射影カメラモデルを用いて生成した Figure 23b は、視野が広く、遠方を確認できる。しかし、直線が画像上では曲線になっている。そのため、車と物との実際の距離と画像上の距離が比例していない。その結果、車と物との距離や方向がわかりにくい。

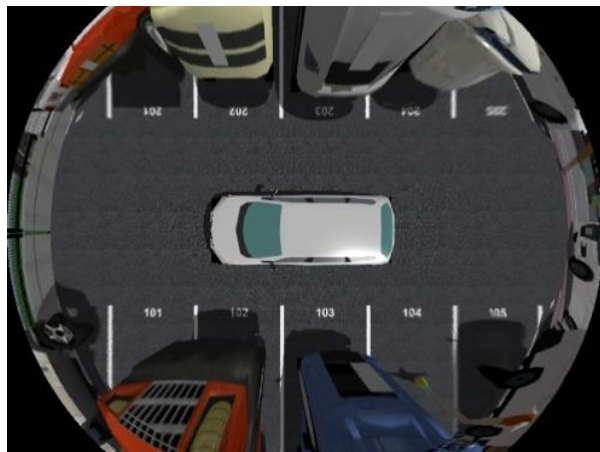
これに対し、提案法の複合射影カメラモデルを用いて生成した Figure 23c は、視野が広く、かつ、車の近くでは直線が画像上でも直線になっている。そのため、車と物との距離や方向がわかり易い。以上のことから、提案法の複合射影カメラモデルを用いると、広い視野と、直線が歪まない画像を、同時に満たす画像を生成できることを確認できた。



(a) Perspective projection on a plane



(b) Equidistance projection on a plane



(c) Composite projection (proposed) on a plane

Figure 23 Images generated by (a) perspective projection, (b) equidistance projection and (c) composite projection (proposed) using the plane shape model.

3.4.2. 形状モデルの比較評価

提案法の半円柱形状モデルを用いて生成した全周囲画像と、従来法の形状モデルを用いて生成した全周囲画像を、比較評価した。従来法の形状モデルとして、平面と半球を用いた。カメラモデルは、複合射影カメラモデルを用いた。3つの形状モデルの主要パラメータと、実験で用いた数値を Table 3 に示す。Table 3 の各パラメータは、3.3.4 節の設計指針により決定した。

平面、半球、半円柱の3つの形状モデルを用いて生成した全周囲画像を Figure 24 に示す。平面形状モデルで生成した画像 (Figure 24a) は、水平線より下までの視野しかない。半球形状モデルで生成した画像 (Figure 24b) は、直線が画像上では曲線になり、車と物との距離や方向がわかりにくい。これに対し、半円柱形状モデルを用いて生成した画像 (Figure 24c) では、水平線より上の視野を含み、遠方の車や人の全体 (図中の白破線) を画像で確認することができる。そのため、車や人の有無がわかり易い。さらに、車の近くの直線は画像上でも直線になるため、車と物との距離や方向がわかり易い。

以上のことから、複合射影カメラモデルと半円柱形状モデルとを組み合わせることで、水平線以上の広い視野、かつ、直線が歪まない全周囲画像を、生成できることを確認できた。

3.4.3. カメラモデルと形状モデルの客観評価

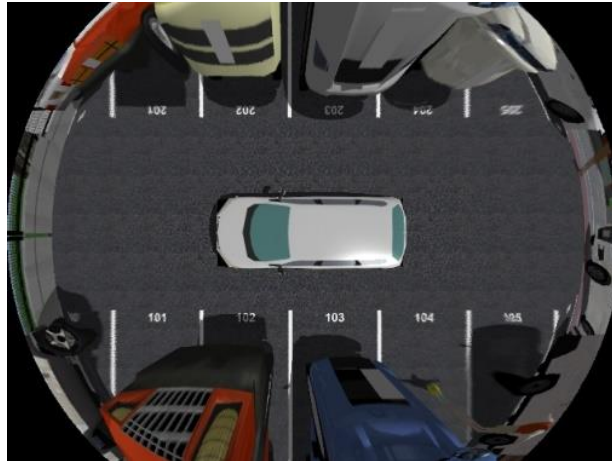
前2節で示した、提案法と従来法で生成した全周囲画像を、客観指標により比較評価する。

駐車操作の支援には、車の周りの物の有無や、物の方向や距離を正しく把握できることが望ましい。これらを表す指標として、視野の広さと、直線の歪み、距離の歪み、をそれぞれ表す、下記の3つの評価指標を用いた。

直線歪率：3次元空間で直線の、全周囲画像上での直線からの歪み度合い。値が0に近いほど、画像上でも直線に近い。

距離歪率：3次元空間で、距離が等しい2組の2点間の、全周囲画像上で距離比の歪み度合い。値が0に近いほど、3次元空間と画像上の距離比に近い。

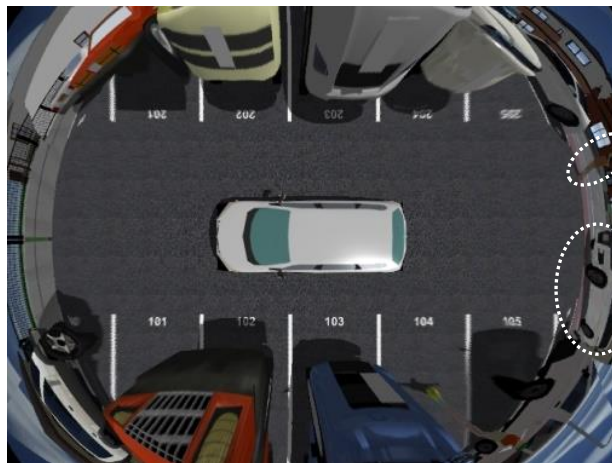
非表示画素率：入力のカメラ画像の全有効画素のうち、全周囲画像の表示範囲外の画素の比率。値が0に近いほど、全周囲画像で表示される範囲が広い。



(a) Composite projection on a plane



(b) Composite projection on a half ball



(c) Composite projection on a halfpipe

Figure 24 Images generated by composite projection using (a) a plane, (b) a half ball and (c) a halfpipe shape models.

Table 3 Values of shape model used in the experiment.

表 3 形状モデルのパラメタと実験で用いた数値

形状モデル	パラメタと実験で用いた値
全モデル共通	Y 方向オフセット： $y_{offset} = 2.5$ [m]
半球	水平方向半径： $r_{xy} = 10.0$ [m] 垂直方向半径： $r_z = 3.0$ [m]
半円柱	左右方向幅： $w = 16.0$ [m] 前後方向長さ： $l = 48.0$ [m] 円柱部半径： $r_h = 2.0$ [m]

直線歪率と距離歪率は、下記の手順で算出した。

Step1: 全周囲画像上のランダムな位置の 2 点を結び、直線を求める。

Step2: 直線上の 2 点に対応する被写体上の 3 次元座標を、CG 画像を作成する際の 3 次元モデルから算出し、3 次元空間の中点を算出。この中点を仮想カメラに投影した画素座標を評価点とする。

Step3: 評価点と画像上の直線との距離を、直線の長さで割った値を直線歪率とする。

Step4: 評価点と、直線の端点 2 点とをそれぞれ結んだ 2 つの距離をもとに、大きい方の距離を 2 つの距離の和で割った値と、0.5 との差を、距離歪率とする。

全周囲画像の、直線の歪みや距離の歪みは、画像の場所や、評価する方向によって、偏りがある。上述の算出法では、多数の評価点を、画像上のランダムな場所に配置して、統計処理する手法とした。ここでは、10,000 本の直線について、各指標の平均値を算出した。直線歪率と距離歪率は、評価に用いる直線の 2 つの端点が、車の中心から 7m 以内の場合（表中”内の以内”）と、全周囲画像の全範囲（表中”全範囲”）の 2 通りを算出した。非表示画素率は、母数を水平線より上を映した画素とした場合（表中”水平以上”）と、カメラ画像の全範囲（表中の”全範囲”）の 2 通りを算出した。

透視射影、等距離射影、複合射影の 3 つのカメラモデルと平面形状モデルを組み

Table 4 Objective evaluation results of camera and shape models.

各評価指標で最良の手法を赤字で表示.

カメラモデル, 形状モデル	図番	直線歪率		距離歪率		非表示画素率	
		全範囲	7m 内	全範囲	7m 内	全範囲	水平以上
透視射影, 平面	Figure 23(a)	0.00	0.00	0.00	0.00	0.33	1.00
等距離射影, 平面	Figure 23(b)	0.06	0.06	0.25	0.11	0.24	1.00
複合射影, 平面	Figure 23(c) Figure 24(a)	0.03	0.00	0.11	0.00	0.28	1.00
複合射影, 半球	Figure 24(b)	0.07	0.03	0.09	0.03	0.17	0.60
複合射影, 半円柱	Figure 24(c)	0.05	0.00	0.13	0.00	0.17	0.56

合わせて生成した全周囲画像 (Figure 23a-c), および, 等距離射影カメラモデルと, 平面形状, 半球形状, 半円柱形状の3つの形状モデルを組み合わせそれぞれ生成した, 全周囲画像 (Figure 24a-c) について, 3つの評価指標を算出した結果を Table 4 に示す.

透視射影, 等距離射影, 複合射影の3つのカメラモデルと平面形状モデルを組み合わせた結果 (Table 4, 1-3 行目) から, 7m 以内の直線歪率および距離歪率において, 提案法の複合射影と透視射影の評価値は, いずれの値も 0.00 と, ほとんど歪みが無いのに対し, 等距離射影は 0.06, 0.11 と歪みがある. 一方, 非表示画素率は, 等距離射影 0.24, 複合射影 0.28, 透視投影は 0.33 の順で値が小さく, 視野が広い. これらのことから, 提案法の複合射影カメラモデルは, 平面形状モデルを用いる全周囲合成において, 車の近くの直線歪みと距離歪みが無く, かつ, 視野が広いという点で, 透視射影や等距離射影より優れていると言える.

複合射影カメラモデルを, 平面, 半球, 半円柱の3つの形状モデルと組み合わせ生成した全周囲画像の評価結果 (Table 4 の 3-5 行) から, 水平線以上の非表示画素率において, 平面の 1.0 に対し, 半球と半円柱形は, それぞれ 0.60, 0.56 小さい. 全範囲の非表示画素率においても, 半球と半円柱形は 0.17 と小さく, とともに視野が広いことがわかる. 一方, 7m 以内の直線歪率および距離歪率では, 半球はいずれも

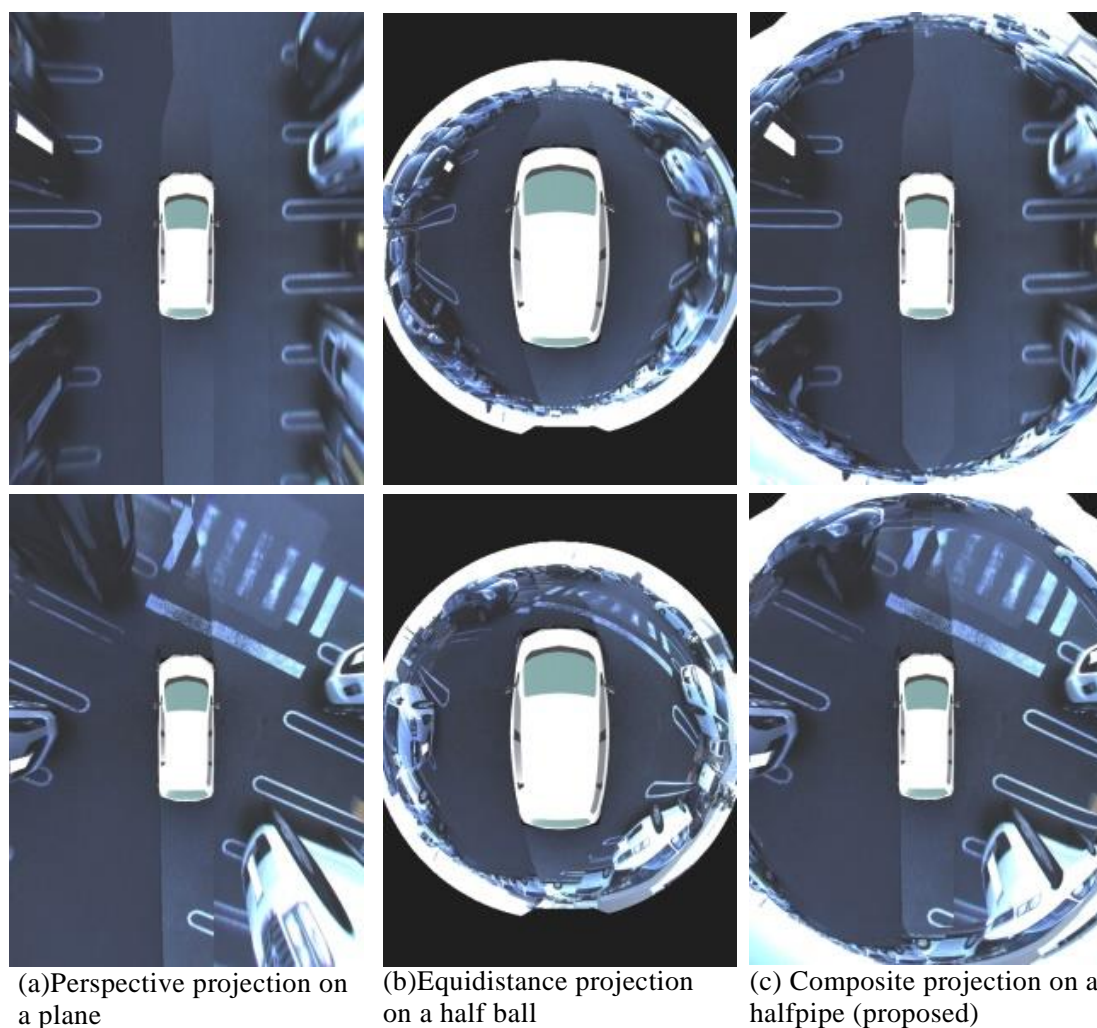
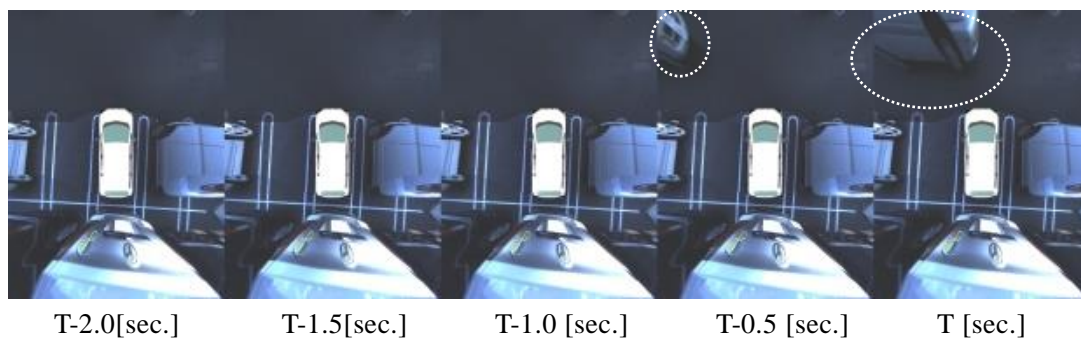


Figure 25 Surround view images generated using the real images of the parking scene.

0.03 と歪を含むのに対し、半円柱では 0.00 と歪がほとんど無い。これらのことから、半円柱形状モデルは、複合射影カメラモデルと組み合わせることで、車の近くの直線歪率と距離歪率を上げることなく、水平以上および全範囲の非表示画素率が小さい、すなわち視野が広い全周囲画像を生成できる、効果があると言える。

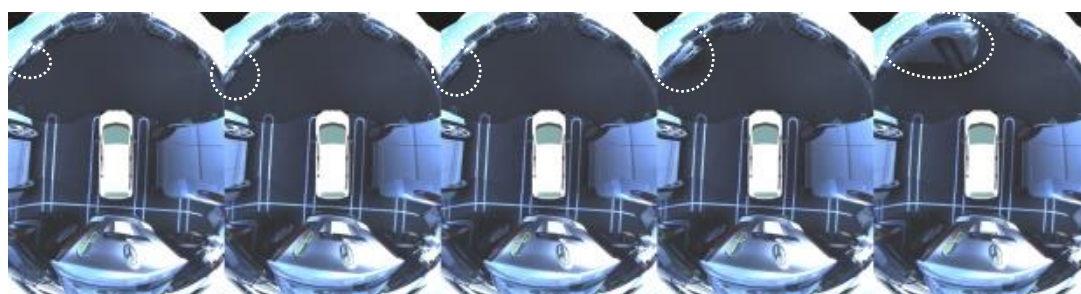
3.4.4. 駐車シーンの実画像での評価

提案法による、広い視野と、直線が歪まない全周囲画像を生成できる効果を確認するため、車載カメラで撮影録画した駐車シーンの実画像を用いて、シミュレーションにより画像生成を行った。比較対象として以下の2つの従来法を用いた。



T-2.0[sec.] T-1.5[sec.] T-1.0 [sec.] T-0.5 [sec.] T [sec.]

(a) Perspective projection



T-2.0[sec.] T-1.5[sec.] T-1.0 [sec.] T-0.5 [sec.] T [sec.]

(b) Composite projection (proposed)

Figure 26 An example of the advantage of wide field of view.

従来法 1 : 透視射影カメラモデルと平面形状モデルを用いる全周囲画像生成手法 [57][35][36]

従来法 2 : 等距離射影カメラモデルと半球形状モデルを用いる全周囲画像生成手法 [58][36]

提案法及び従来法 1, 2 の 3 つの手法で生成した, 駐車シーンの全周囲画像を Figure 25 に示す. 提案法の複合射影カメラモデルと半円柱形状モデルを用いて生成した画像 (Figure 25c) は, 従来法 1 で生成した画像 (Figure 25a) より視野が広く, 従来法 2 で生成した画像 (Figure 25b) より, 車の近くの歪が小さく, 車の近くの直線は画像上でも直線になる. そのため, 車と物との距離や方向がわかり易い.



(a) Perspective projection on a plane



(b) Composite projection on a halfpipe (proposed)

Figure 27 Generated images of a pedestrian near the vehicle.

提案法の広い視野の効果を示す別の例を，Figure 26 に示す．Figure 26a, b は，車の前方左側より別の車が接近してくるシーンにおいて，それぞれ従来法 1，提案法で生成した 0.5 秒間隔の画像である．いずれも隣り合うカメラによって重複撮影された領域の画像生成には，前後のカメラ画像を用いた．接近車両の移動速度は約 24[km/h]であった．提案法で生成した画像は，従来法で生成した画像と比べて，1.5 秒早く接近車を表示できる．このことは，提案法で生成した画像の視野が広いことの効果であると言える．



(a) Side by side parking lot



(b) Parallel parking lot



(c) Parking lot adjacent to the store

Figure 28 Generated images in various parking scenes (proposed method).

車の近くを歩行者が通過するシーンにおいて、従来法1と提案法で生成した全周囲画像を Figure 27 に示す。従来法1で生成した画像 (Figure 27a) は、歩行者の足元しか映っていないのに対し、提案法で生成した (Figure 27b) は上半身まで映っている。このことから、提案法で生成した全周囲画像は、従来法1で生成した全周囲画像と比べて、車の近くの歩行者の確認し易いと言える。

さまざまな駐車シーンにおいて、提案法により生成した全周囲画像を Figure 28 に示す。Figure 28 a-c は、それぞれ並列駐車、縦列駐車、店舗前での駐車における、駐

車開始から完了までの画像である。図から、提案法で生成した全周囲画像は、車の近くの駐車枠の直線の歪みが無く、かつ、視野が広く、駐車場が面する道上の遠方を含むことがわかる。

3.4.5. 考察

提案法における複合射影カメラモデルの効果は、3.4.1 節と 3.4.4 節の実験および 3.4.3 節の客観評価において、広い視野の画像を生成でき、かつ、車の周りの直線が画像上でも直線である全周囲画像を生成できることを確認できた。同様に半円柱形状モデルの効果は、3.4.2 節と 3.4.4 節の実験および 3.4.3 節の客観評価において、複合射影モデルの効果を増やうことなく水平線より上を含む視野の広い画像を生成できることと、半球モデルのように直線が曲線にならないこととして、確認できた。提案法の直線が画像上で直線となる効果は、表 Table 2, Table 3 に示したパラメタに限らず、複合射影モデルの透視射影内 ($\theta \leq \theta_c$)、かつ、半円柱形状モデル底部の平面内の領域で、得ることができる。同様に、提案法の視野が広い効果は、複合射影モデルの視野角が 180 度以上、かつ、半円柱モデルの高さが仮想カメラ位置より高い場合に、得ることができる。一方で、複合射影カメラモデルは、2つ射影モデルの像の拡大率が異なるために、境界 θ_c で折り曲がった像になる欠点がある。この欠点は、複合射影カメラモデルのパラメタ f_e を $f_p/\cos 2\theta_c$ に近づけることで改善できるが、 f_e , f_p , θ_c の設計の自由度が低下する、別の課題が生じる。

3.5. 結言

本章では、水平線より上を含む広い視野と、路面上の直線が画像上で直線になる、低歪みを、同時に満たす全周囲画像を生成する手法を提案した。提案法は、画像生成のカメラモデルとして、透視射影と等距離射影の2つを組み合わせた複合射影モデルを用いる。これにより、車の近くは透視射影を用いることで、直線の歪みが無く、遠くは等距離射影を用いることで視野が広い、画像を生成することができる。さらに提案法は、車の周りの形状として、従来の平面や半球状曲面の代わりに、半円柱状の形状モデルを、複合射影カメラモデルと組み合わせて用いることで、水平線より上を含む、視野が広い画像を生成できる。

実験では、実験車で撮影した様々なシーンの画像から、全周囲の水平線より上を

含み，車の近くでは，路面上の直線が画像上でも直線になり，かつ，車と物との実際の距離と画像上の距離が比例する，歪の少ない画像を生成できることを確認した．提案法で生成した全周囲画像は，視野が広く，車と物との実際の距離と画像上の距離が比例することから，遠方の物を確認し易く，車と物との距離や方向がわかり易いことが期待できる．提案法は，駐車時の運転支援のみならず，後退時や狭路での右左折を伴う走行などにおいても有効であると考えられる．このような駐車以外の運転支援における有効性についても，今後検証を進めたい．

4. フレーム間動きと生成画像の同時推定による動き補償型高解像度化手法

本章では，仮想視点映像中の局所的な低解像度領域の解像度を向上させることを目的として，車載カメラ映像を高解像度化する方法について提案する．高解像度の動画を，データ伝送量を増やすことなく生成するために，提案手法では，カラー画像の3色のうちの1色を高解像度低フレームレートで撮影可能な3板撮像センサカメラを想定する．そして，得られた動画を入力として，フレーム間の動きを利用して高フレームレート化する処理と，色間の局所的なテクスチャの相関を利用して高解像度化する処理とを行う，動き補償型の高解像度化手法により，高解像度高フレームレートの動画を生成する．その際，フレーム間の動きの推定と高解像度画像の生成を同時に行うことで，動きの推定精度の向上と生成画像の誤差低減を図る．シミュレーション実験により，動きと生成画像の同時推定が，フレーム間の動きの推定精度の向上と生成画像の誤差低減の2つの点で有効であることを示す[41][42][43][44][45][46]．

4.1. 序言

撮像素子の多画素化が進み，近年では，デジタルシネマや超高精細放送のための高解像度のビデオカメラの開発が盛んに行われている[64][65]．さらに高臨場感を目指した広視野映像やインテグラル立体映像[66]，監視映像のためには，より一層の多画素化が必要とされている．今後，撮像素子の多画素化が進んだ場合，解像度とフレームレートの積が画素読み出し速度の上限を越える動画を撮影できない点が重要な課題となる．本研究は，この課題を解決するために，単位時間当たりの読み出し画素数が一定の動画から，解像度とフレームレートを向上した動画を生成する手法の開発を目的とする．

動画の解像度やフレームレートを向上する画像処理手法として，超解像手法やフレームレート変換手法が提案されている．

超解像手法は，再構成型と学習型に大別できる．再構成型の超解像手法は，サブ

ピクセルの位置ずれを含む複数の低解像度画像を位置あわせし、低解像度画像に折り返しとして含まれる高周波成分を復元することで高解像度画像を生成する[67][68][69][70]。しかし、この手法は、静止領域の解像度を向上することができない。また、低解像度画像間の位置合わせ精度に起因して解像度の向上に限界があることが知られている[69][71][72]。学習型の超解像手法は、低解像度画像と高解像度画像の関係を学習しておき、低解像度画像から高解像度画像を推定する[73][74][75]。学習型は、被写体のカテゴリを限定した場合に、解像度の向上に有効である。しかし、不特定多数のカテゴリを含む画像では、低解像度画像に対する高解像度画像の対応関係を一意に定めることが難しく、良好な結果が得られない。また、解像度の差が大きいほどこの傾向は強くなるため、解像度の大幅な向上は困難である。

フレームレート変換手法は、低フレームレートの入力動画像の中間フレームを生成することで、フレームレートを向上した動画像を生成する。代表的な手法として、フレームレート間の動きを推定し、内挿した動きと画像を用いて中間フレームを生成する手法がある。しかし、この手法は、非直線の動きを伴う領域では、内挿した動きと真の動きとの誤差が大きく、生成した動画像と真の動画像と誤差が大きい。

以上のように、超解像手法やフレームレート変換手法は、静止領域や、非直線動き領域、不特定多数のカテゴリの被写体に対して、高解像度化や高フレームレート化の効果が期待できないという課題があった。

これらの課題に対し、マルチセンサカメラ等で解像度とフレームレートが異なる複数の動画像を撮影し、画像処理によって高解像度高フレームレートの動画像を生成する手法が提案されている[76][77][78][79][80][81]。渡邊ら[77]は、同一視野かつ同一視点の、低解像度高フレームレートの動画像と高解像度低フレームレートの動画像を撮影し、2つの動画像から得たフレーム間の動きと高周波成分を用いてウェーブレット領域で動き補償を行うことで、高解像度高フレームレートの動画像を生成する手法を提案している。今川ら[81]は、高解像度高フレームレート動画像の生成と露光量の確保の両立を目的として、色ごとに露光時間、解像度、フレームレートが異なる動画像から、高解像度高フレームレートの動画像を生成する、再構成型の手法を提案している。この手法では、高解像度低フレームレートの動画像を長時間露光により撮影する。そして、長時間露光によって生じる動きぶれを含む動画像から、画像復元の枠組みによって動きぶれを低減した高解像度高フレームレートの動画像を生成する。

これらの解像度とフレームレートが異なる複数の動画像を用いる従来法は、静止領域や非直線動き領域、不特定多数のカテゴリの被写体に対しても高解像度化や高フレームレート化の効果が期待できる。しかし、低解像度の動画像からフレーム間の動きを推定しているため、動きの推定精度が低い。これは、動きの量子化誤差が大きいことと、エリアシングが動画像に含まれていることに起因する。その結果、精度の低い動きを用いて生成した動画像は、真の動きを用いて生成した動画像に対して誤差が大きいという課題があった。また、再構成型の手法は、高解像度高フレームレートの全画素値の算出することから、動き補償型の手法と比較して、計算負荷が非常に大きいことも課題であった。

本論文では、色ごとに解像度とフレームレートが異なる動画像に対し、フレーム間の動きと生成画像の同時推定を適用することで、動きの推定精度を向上し、生成画像の誤差を低減した高解像度高フレームレートのカラー動画像を生成する、動き補償型の高解像度化手法を提案する。提案法は、同一露光時間で撮影した色成分（RGB）ごとに解像度とフレームレートの異なる動画像を入力として、フレーム間の動きに基づいた中間フレーム生成によってフレームレートを向上する処理と、色間の局所的なテクスチャの相関を用いて解像度を向上する処理とを組み合わせることで、高解像度高フレームレートの動画像を生成する。提案法では、エリアシングの影響を低減して精度の高いフレーム間動きを推定するために、高解像度化した生成画像においてフレーム間の画素値の差を最小化する位置変化として、フレーム間の動きを定義する。そして、これを解くために、動きと生成画像の同時推定手法を適用する。提案法は、小領域ごとに、フレーム間の動きの2変数を推定する手法であることから、全画素値を推定する再構成型手法より計算負荷が小さいという効果も期待できる。

以降、4.2節では、本論文で想定する3板撮像センサカメラと、高解像度高フレームレート動画像を生成する提案手法を説明する。4.3節ではシミュレーションによる本手法の評価結果を説明し、4.4節でまとめと今後の展開を述べる。

4.2. 高解像度高フレームレート動画像生成

色(RGB)ごとに解像度とフレームレートが異なる動画像から、高解像度高フレームレートのカラー動画像を生成する手法を提案する。以降、4.2.1節では色ごとに解像

度とフレームレートが異なる動画像の撮像方法について、4.2.2 節では動きと生成画像の同時推定による高解像度高フレームレート動画像の生成方法について、順に説明する。

4.2.1. 色ごとに時空間解像度が異なる動画像の撮像

RGB ごとに解像度とフレームレートが異なる動画像を、同一視野かつ同一視点から撮像するために、Figure 29 の構成の 3 板撮像センサカメラを想定する。レンズからの入射光は色分解プリズムによって RGB の 3 色に分解され、3 つの撮像素子に投影される。各撮像素子では、高解像度低フレームレートの緑色(G)の動画像 G_{HL} と、低解像度高フレームレートの赤色(R)の動画像 R_{LH} 及び青色(B)の動画像 B_{LH} を撮影する。各動画像の露光時間は同じである。撮像素子面上で低解像度動画像の 1 画素の大きさは、高解像度動画像の $N_x \times N_y$ 画素の大きさと一致する。したがって、低解像度の R_{LH} , B_{LH} には、 $N_x \times N_y$ 画素の平均化によるエリアシングが含まれる。低フレームレートの G_{HL} と、高フレームレートの R_{LH} , B_{LH} とは、一定のフレーム数 N_f の間隔で同期して撮影する。この同期したフレームをキーフレーム、それ以外を中間フレームと呼ぶ。色間の空間的な対応と、キーフレームと中間フレームの時間的な対応を利用することで、解像度とフレームレートが異なる動画像を統合することができる。以降、画素の位置や大きさを表す値は、特に断りが無い限り、高解像度動画像の画素座標値を基準とする。

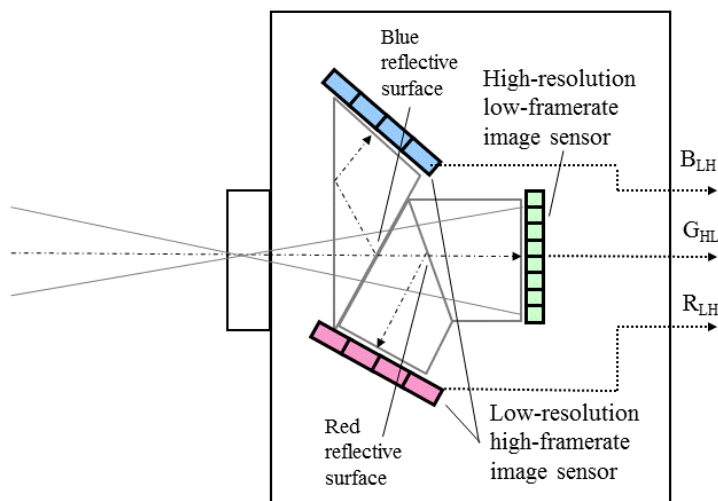


Figure 29 3-image-sensor camera.

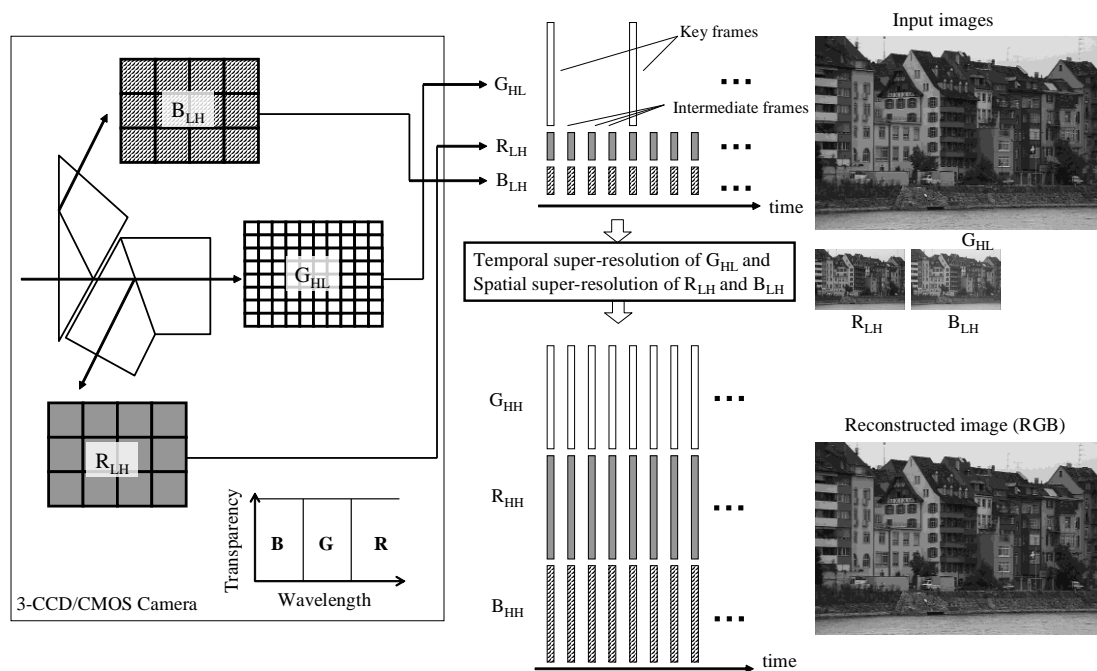


Figure 30 Outline of proposed method.

4.2.2. 動きと生成画像の同時推定による高解像度高フレームレート動画の生成

提案法では、RGBごとに解像度とフレームレートが異なる動画を入力として、空間と時間の情報を相互に補完することで高解像度化と高フレームレート化を図る。具体的には、高解像度低フレームレートの G_{HL} と、低解像度高フレームレートの R_{LH} 、 B_{LH} を入力として、 G と同じ高解像度かつ R 、 B と同じ高フレームレートである動画 R_{HH} 、 G_{HH} 、 B_{HH} を生成することを目的とする (Figure 30)。入力動画 R_{LH} 、 G_{HL} 、 B_{LH} には、 R 、 B の空間的な高域成分の情報や、 G の時間的な変化の情報が、明には含まれていない。そのため高解像度高フレームレートの動画 R_{HH} 、 G_{HH} 、 B_{HH} を生成するためには、これらの動画において成立する時間的、空間的なモデルと、フレーム間の動きの情報が必要である。

本手法では、解像度の異なる色間の空間的なモデルとして、局所領域において低域成分の相関が強い場合、高域成分の相関も強い、という仮定 (局所色相関モデル)

を用いる。また、動画像の時間的な変化のモデルとして、中間フレームの小領域はキーフレームの小領域の並進運動によって近似できる、という仮定（並進動きモデル）を用いる。

これまでの多くの研究は、フレーム間の動きを事前に推定し、その動きに基づいて動画像を生成する[76][77][79][80][81]。この解法は、既存の動き推定手法でフレーム間の動きを高い精度で推定できる場合に、妥当かつ効率的である。しかし、既存の動き推定手法では、本手法で扱う解像度が低くエリアシングが含まれる動画像から、高い精度で動きを推定するのは困難である。

そこで本手法では、フレーム間の動きの推定精度を向上するために、生成する高解像度高フレームレートの動画像において、フレーム間の画素値の差を最小化する位置変化として、フレーム間の動きを定義する。そして、生成画像において定義されたフレーム間の動きを解くために、動きと生成画像を同時に推定する。生成画像は解像度が高いことと、エリアシングが低減されることから、従来の低解像度フレームから推定した動きに対して、動きの推定精度が向上する効果が期待できる。さらに、精度を向上した動きに基づいて生成した動画像は、正解動画像との誤差が低減する効果が期待できる。

本手法と類似するアプローチとして、再構成型の超解像手法において、複数の低解像度画像間の位置と生成画像とを撮像モデルを介して関連づけ、同時最適化により推定する手法が提案されている[82][83]。しかし、これらの手法は、エリアシングが含まれる低解像度画像間の画素値の差を最小にする位置変化として動きを定義しているため、動きの推定精度を向上する効果は限定的である。これに対し提案法は、エリアシングが含まれない高解像度の生成画像間の画素値の差を最小化する位置変化として、フレーム間の動きを定義することで、動きの推定精度の向上を図る。

色間の相関を用いて解像度の向上を図るアプローチとして、カラーフィルタ画像のデモザイキング手法が数多く提案されている[84][85][86]。しかし、これらの手法は、色数が3(RGB)の場合に3倍を越える画素数の向上ができない。これに対し提案法では、時間方向の相関を積極的に利用することで、入力動画像の4倍以上の画素数のカラー動画像の生成を目指す。

以降、提案法のフレーム間の動きと2つのモデルについて定式化し、同時推定手法の詳細を示す。

フレーム間の動き推定精度の向上

フレーム間の動きの推定精度を向上するために、生成する高解像度高フレームレートの動画像において、フレーム間の画素値の差を最小化する位置変化として、フレーム間の動きを定義する。フレーム t 、位置 (x,y) のキーフレーム tk に対する動き $m(x,y,t)$ と、生成する高解像度高フレームレートの動画像 R_{HH} の関係式を式(4-1)に示す。

$$m(x, y, t) = \arg \min_{dx, dy} (\varepsilon_{x,y,t}(dx, dy)) \quad (4-1)$$

$$\varepsilon_{x,y,t}(dx, dy) = \sum (R_{HH}(x, y, t) - R_{HH}(x+dx, y+dy, tk))^2$$

ここで $R_{HH}(x,y,t)$ は、動画像 R_{HH} のフレーム t 、位置 (x,y) の画素値である。 $\varepsilon_{x,y,t}(dx,dy)$ は、 R_{HH} のフレーム t 上の領域 r と、そこから (dx,dy) 移動したキーフレーム tk 上の領域間の、画素値の差の二乗和 SSD(Sum of Squared Differences)である。フレーム間の動き $m(x,y,t)$ は、この SSD を最小にする動き (dx,dy) である。このフレーム間の動きは、 R,G,B に共通とする。式(4-1)から、 $m(x,y,t)$ を求めるためには、 R_{HH} が必要である。生成画像 R_{HH} と入力動画像の関係を次に示す。

局所色相関モデルによる高解像度化

解像度が異なる R,G,B 動画像間の空間的なモデルとして、局所色相関モデルを用いる。この局所色相関モデルにおける、入力となる低解像度の R_{LH} 、 B_{LH} と生成する高解像度の R_{HH} 、 B_{HH} との関係を式(4-2)(4-3)に示す。同一時刻のフレームのみを扱うため、時間項は省略する

$$R_{HH}(x, y) = (1 + c_{RG}(x_L, y_L) dG_{HH}(x, y)) R_{LH}(x_L, y_L) \quad (4-2)$$

$$B_{HH}(x, y) = (1 + c_{BG}(x_L, y_L) dG_{HH}(x, y)) B_{LH}(x_L, y_L) \quad (4-3)$$

$$c_{RG}(x_L, y_L) = \frac{\sum_{v=y_L-1}^{y_L+1} \sum_{u=x_L-1}^{x_L+1} (G_{LH}(u, v) - \bar{G}_{LH})(R_{LH}(u, v) - \bar{R}_{LH})}{\sqrt{\sum_{v=y_L-1}^{y_L+1} \sum_{u=x_L-1}^{x_L+1} (G_{LH}(u, v) - \bar{G}_{LH})^2 \sum_{v=y_L-1}^{y_L+1} \sum_{u=x_L-1}^{x_L+1} (R_{LH}(u, v) - \bar{R}_{LH})^2}}$$

$$c_{BG}(x_L, y_L) = \frac{\sum_{v=y_L-1}^{y_L+1} \sum_{u=x_L-1}^{x_L+1} (G_{LH}(u, v) - \bar{G}_{LH})(B_{LH}(u, v) - \bar{B}_{LH})}{\sqrt{\sum_{v=y_L-1}^{y_L+1} \sum_{u=x_L-1}^{x_L+1} (G_{LH}(u, v) - \bar{G}_{LH})^2 \sum_{v=y_L-1}^{y_L+1} \sum_{u=x_L-1}^{x_L+1} (B_{LH}(u, v) - \bar{B}_{LH})^2}}$$

$$dG_{HH}(x, y) = \frac{G_{HH}(x, y) - G_{LH}(x_L, y_L)}{G_{LH}(x_L, y_L)}$$

ここで c_{RG} は、低解像度の R_{LH} と G_{LH} の局所領域での相関係数である。 (x_L, y_L) は低解像度動画像の位置であり $(x_L, y_L) = (x/NX, y/NY)$ となる。 \bar{G}_{LH} , \bar{R}_{LH} は低解像度動画像の局所領域での平均値であり、低解像度動画像の3低解画素の領域を局所領域とする。低解像度の G_{LH} は、 R_{LH} , B_{LH} の撮像モデルと同様に、高解像度の G_{HH} の $Nx \times Ny$ 画素の平均値である。また、 dG_{HH} は G_{HH} の高域成分である。

式(4-2)において、 G_{HH} から G_{LH} の変換は、撮像系における R_{HH} から R_{LH} への低解像度化と等価である。また、式(4-2)で生成される R_{HH} は、 R と G の相関 c_{RG} が1に近いほど、入力動画像 R_{LH} に G の高域成分 dG_{HH} と R_{LH} との積が重畳される。ここで $(1+dG_{HH})$ は低解像度の G_{LH} から高解像度の G_{HH} への逆変換であることから、 R と G の相関が1のとき、式(4-2)は R_{LH} から R_{HH} への完全な逆変換となる。その結果、平均化によるエリアシングが含まれない生成画像 R_{HH} を得ることができる。生成画像 B_{HH} の式(4-3)も同様である。式(4-2)(4-3)から、 R_{HH} , B_{HH} を求めるには G_{HH} が必要である。生成画像 G_{HH} と入力動画像の関係を次に示す。

並進動きモデルによる高フレームレート化

動画像の時間的な変化のモデルとして、並進動きモデルを用いる。この並進動きモデルにおける、入力となる低フレームレートの G_{HL} と生成する高フレームレートの G_{HH} との関係を、式(4-4)に示す。

$$G_{HH}(x, y, t) = G_{HL}(x + dx, y + dy, tk) \quad (4-4)$$

ここで $G_{HH}(x,y,t)$ は、動画像 G_{HH} のフレーム t 、位置 (x,y) の画素値である。 (dx,dy) は、フレーム t 、位置 (x,y) のキーフレーム tk に対する動きであり、式(4-1)の動き $m(x,y,t)$ と同じである。

動きと生成画像の同時推定

生成すべき高解像度高フレームレートの動画像 R_{HH} , G_{HH} , B_{HH} と、入力動画像 R_{LH} , G_{LH} , B_{LH} との関係を、式(4-1)~(4-4)に示した。式(4-1)~(4-4) 及び図 2 から、 R_{HH} , B_{HH} を求めるには G_{HH} が必要であり、 G_{HH} を求めるには R_{HH} と動き $m(x,y,t)$ が必要である。そのため、生成画像 R_{HH} , G_{HH} , B_{HH} と動き $m(x,y,t)$ は、入力動画像から直接算出することができない。そこで本手法では、以下の手順による探索手法により、動き $m(x,y,t)$ と、 R_{HH} , G_{HH} を同時に推定する。

Step 1 所定の探索範囲内で動きの候補 (dx, dy) を与える。

Step 2 動き候補 (dx,dy) に対応する G_{HH} , R_{HH} を、それぞれ式(4-2), 式(4-3)より算出し、式(4-1)より生成画像での画素値の差 ε を算出する。

Step 3 探索範囲内で Step1~2 を繰り返し、 ε が最小となる動き (dx, dy) をフレーム間の動き $m(x,y,t)$ とする。そして、同時に生成される $R_{HH}(x,y,t)$, $G_{HH}(x,y,t)$ を生成画像とする。

Step 4 推定した $G_{HH}(x,y,t)$ と B_{LH} から $B_{HH}(x,y,t)$ を算出する。

以上のように提案手法では、フレーム間の動きを、高解像度の生成画像 R_{HH} で定義し、この動きと生成画像 R_{HH} , G_{HH} , B_{HH} を同時推定によって求める。生成画像 R_{HH} は解像度が高いこととエリアシングが低減されることから、従来の低解像度動画像から推定した動きに対して、動きの推定精度が向上する効果が期待できる。さらに、精度が向上した動きに基づいた生成画像は、正解動画像との誤差が低減し、画質が向上する効果が期待できる。

提案手法の効果と、入力動画像の特性との間には、式(4-1)~(4-4)から、次の関係があると考えられる。

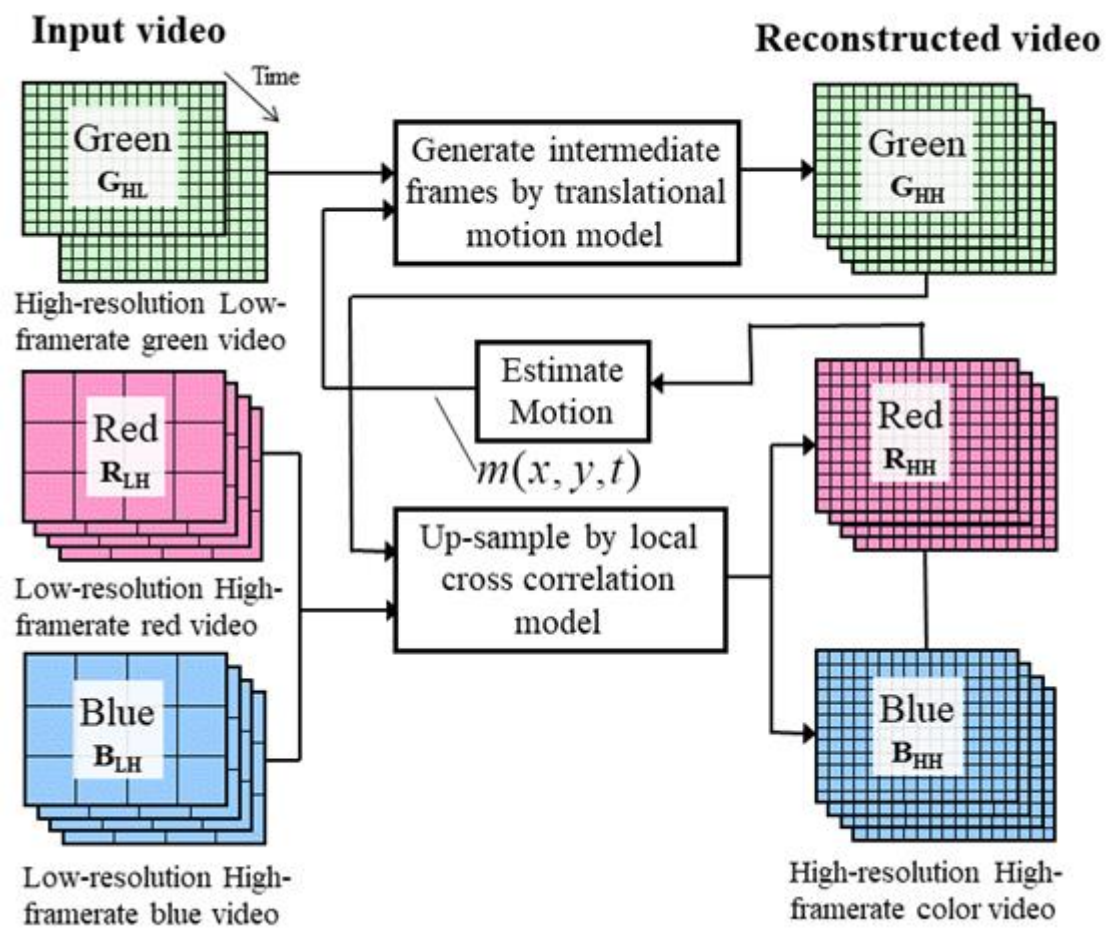


Figure 31 Block diagram of simultaneous estimation of motion and high resolution images.

- ① R, G 間の相関が高いほど, また, G の高域成分が多いほど, 動きの推定精度と生成画像の画質が向上する効果が高くなる.
- ② R, G 間の相関が低いほど, また, G の高域成分が少ないほど, 動きの推定精度と生成画像の画質を向上する効果が低くなる.

①は, 式(4-2)の生成画像 R_{HH} の高解像度化とエリアシングの低減により, 及び式(4-1)の動き $m(x,y,t)$ の精度向上が期待できるためである. ②は, 式(4-2)の生成画像 R_{HH} が入力動画像 R_{LH} に近づき, 式(4-1)の動き $m(x,y,t)$ が入力動画像 R_{LH} から推定し

た動きに近づくためである。

4.3. 評価実験

本章では、シミュレーションによる評価実験の結果について述べる。以降、4.3.1節でシミュレーションによる検証方法について説明し、4.3.2節で提案手法による動きの推定精度評価について、4.3.3節で生成画像の画質評価について、4.3.4節で局所色相関による高解像度化と折り返し低減の効果について、順に説明する。

4.3.1. シミュレーションによる検証方法

シミュレーション実験による、動きの推定精度及び生成画像の定量的画質の検証方法について説明する。まず、高解像度高フレームレートの正解動画像 R_{true} , G_{true} , B_{true} を用意する。次に、この正解動画像を用いて、 G_{true} から Nt フレーム間隔で抜き出すことで G_{HL} 動画像を作成し、 R_{true} , B_{true} を $Nx \times Ny$ 画素ごとに平均化してダウンサンプルすることで R_{LH} , B_{LH} 動画像を作成する。そして、3つの動画像 G_{HL} , R_{LH} , B_{LH} を入力動画像として、各手法で正解動画像と同じ解像度とフレームレートの動画像 G_{HH} , R_{HH} , B_{HH} を生成する。最後に、生成画像と正解動画像との比較により画質評価を行う (Figure 32)。低解像度画像に対する高解像度画像の画素数の比率 $Np (=Nx \times Ny)$ を空間拡大率、低フレームレート動画像に対する高フレームレート動画像のフレームレートの比率 Nt を時間拡大率と呼ぶ。なお、時間拡大率 Nt と空間拡大率 Np は、独立に設定し得るが、本実験では、各色の入力動画像の総画素数が等しくなるよう時間拡大率と空間拡大率は同じ $Nt=Np$ とし、拡大率と呼ぶ。

生成画像の評価指標には、式(4-5)の CPSNR(Color Peak Signal to Noise Ratio)を用いる。 N は画像の画素数である。

$$CPSNR = 10 \cdot \log_{10} \frac{255^2}{CMSE} \quad (4-5)$$

$$CMSE = \sum \frac{(R_{HH} - R_{true})^2 + (G_{HH} - G_{true})^2 + (B_{HH} - B_{true})^2}{3N}$$

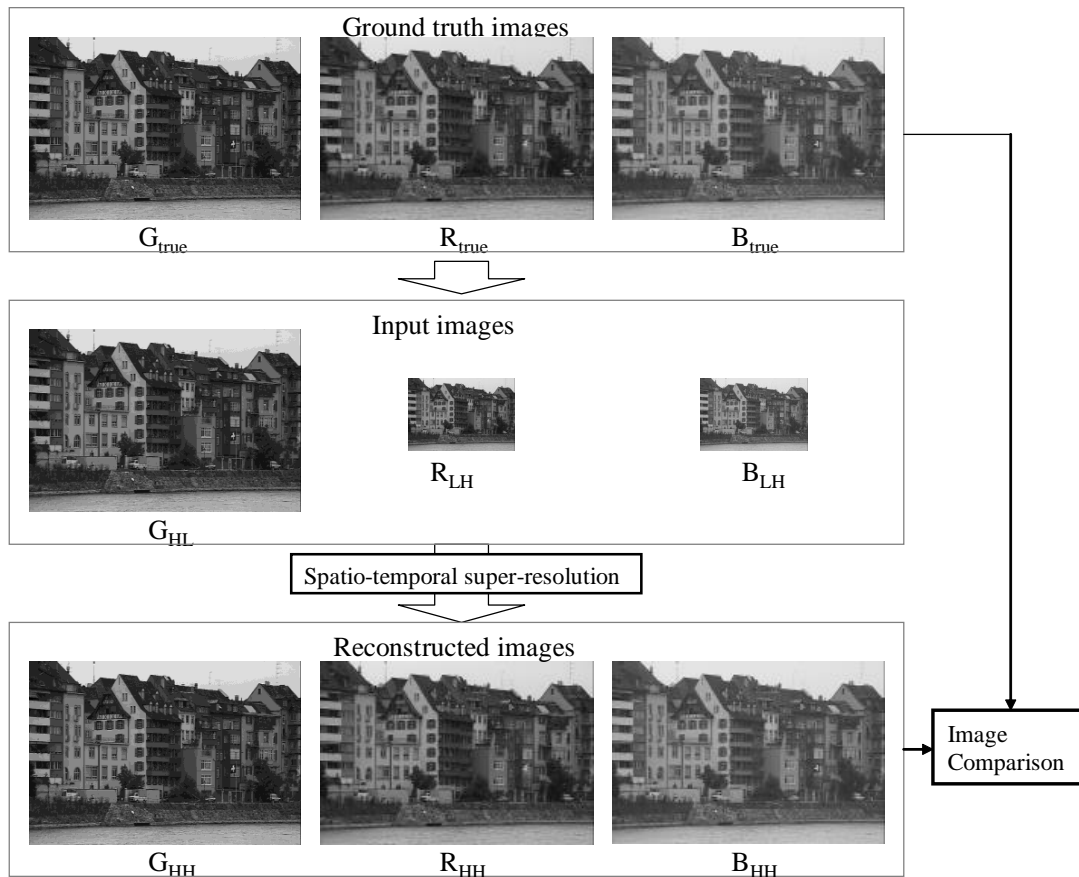


Figure 32 Flowchart for evaluating results of simulations.

実験には、(社)映像情報メディア学会監修の標準動画画像集[87]を RGB 色空間に変換したものを正解画像として用いた。評価用動画画像を Figure 33 に示す。実験に使用した正解動画画像のフレーム数は、90 フレーム前後で時間拡大率 Nt の整数倍 + 1 となるように設定した。

各動画画像の特性を示す指標として、正解動画画像 R_{true} , G_{true} 間の相関係数と、縮小動画画像の CPSNR を Table 5 に示す。縮小動画画像の CPSNR は、それぞれ Bi-cubic 法により、縦横 1/4 倍に縮小した後、縦横 4 倍に拡大し、もとの動画画像に対して算出した。Table 5 において、相関係数の絶対値が大きくなるほど、色間の相関が高く、相関係数が 0 に近づくほど色間の相関が低い。また、縮小画像の CPSNR が大きいほど高域成分が少なく、CPSNR が小さいほど高域成分が多いことを表す。



(a) River



(b) Church



(c) Canal



(d) Driving



(e) Buddhist



(f) Swinging



(g) Japanese



(h) Flash



(i) Airplane



(j) Yachting



(k) Cognac



(l) Square

Figure 33 ITE video sequences for evaluation.

Table 5 Characteristics of ITE video sequences.

Test video sequences.	Correlation coefficients between R and G.	CPSNR of 1/16 reduced images [dB]
River (Wintertime River)	0.994	24.7
Church	0.994	21.4
Canal (Buildings along the Canal)	0.988	23.4
Driving	0.982	25.9
Buddhist (Buddhist Images)	0.980	27.3
Swinging	0.973	26.7
Japanese (Japanese Room)	0.965	25.8
Flash (Flash Photography)	0.936	32.1
Airplane (Airplane Landing)	0.935	31.1
Yachting	0.893	26.0
Cognac (Cognac and Fruit)	0.847	28.6
Square (Walk through the Square)	0.823	24.8

4.3.2. 提案手法による動きの推定精度の向上

動きと生成画像の同時推定による動きの推定精度の向上の効果を評価した。動きが既知の正解動画像として、標準動画像の1フレーム（原画像）と、これをX,Y方向にそれぞれ0-3画素平行移動した16フレームの合成画像を正解動画像として用いる。最初のフレームをキーフレーム、残りを中間フレームとして生成した17フレームを入力動画像として用いる。拡大率は16倍($N_t=N_p=16, N_x=N_y=4$)とした。比較対象として、以下の2つの方法を用いる。

Motion estimation from R_{LH} (MEL): 低解像度の R_{LH} から動きを推定する方法

Motion estimation from bi-cubic R_{LH} (MEB): 低解像度の R_{LH} を Bi-cubic 補間拡大した動画像から動きを推定する方法

実験では、中間フレーム上に設定した複数の小領域を基準に、キーフレームとの動きを各手法で推定し、正解動きとの平均二乗誤差(RMSE)を算出した。小領域は、画像上に等間隔に384箇所(X方向24箇所×Y方向16箇所)設定した。各手法とも共通して、探索範囲はX,Y方向にそれぞれ±16画素、小領域のサイズは12×12画素とした。探索範囲は、正解動きを含む十分な範囲を用いた。小領域のサイズは、

実験により複数のサイズ(12×12, 20×20, 28×28, 36×36)で画像を生成した結果, 生成画像の主観画質が最も良い 12×12 を用いた. 推定の精度は, パラボラフィッティングによって実数精度で算出した.

動きの推定精度の評価結果を Figure 34 に示す. 提案法(PM)の動きの誤差は, 低解像度の R_{LH} から動きを推定する MEL 法の動きの誤差の 7~86%, Bi-cubic 補間拡大画像から動きを推定する MEB 法の動きの誤差の 11~112%であり, 1 動画像を除いて MEL 法, MEB 法の動きの誤差より小さい. これは MEL 法, MEB 法がエリスシングを含む低解像度動画像から動きを推定しているのに対し, 提案法ではエリスシングを低減した生成画像から動きを推定しているためである. 一方, Cognac の 1 動画像のみ, 提案法の動きの誤差は, MEB 法より大きい. これは, Cognac の色間の相関が低く, かつ, 高域成分が少ないためと考えられる (Table 5). 提案法で推定した動きは, 4.2.2 節で述べたように, 色間の相関が低いほど, また, 高域成分が少ないほど, MEL 法で推定した動きに近くなる. 低解像度動画像から動きを推定する MEL 法は, 低解像度動画像を Bi-cubic 補間拡大した動画像から動きを推定する MEB 法より, 動きの誤差が大きい. その結果, Cognac における提案法の動きの誤差は, MEB 法より大きくなる.

提案法の動きの誤差を低減する効果と動画像の特性との関係を明らかにするため, Table 5 の色間の相関及び高域成分の量と, Figure 34 の MEL 法の動きの誤差に対する提案法の動きの誤差の比率 (誤差低減率) との相関係数を示す. R,G 間の相関係数 (正のみ) と誤差低減率との相関係数は -0.694 であり, 負の相関がある. このことは, 色間の相関が高いほど, 動きの誤差を低減する効果が高いことを示す. 縮小画像の CPSNR と誤差低減率との相関係数 0.419 であり, 正の相関がある. このことは, 高域成分が多いほど, 動きの誤差を低減する効果が高くなる傾向があることを示す.

以上のことから, 提案法は, 色間の相関が高い, 又は, 高域成分の多い動画像において, 低解像度動画像から動きを推定する従来法に対して, 動きの推定精度を向上できることが確認できた.

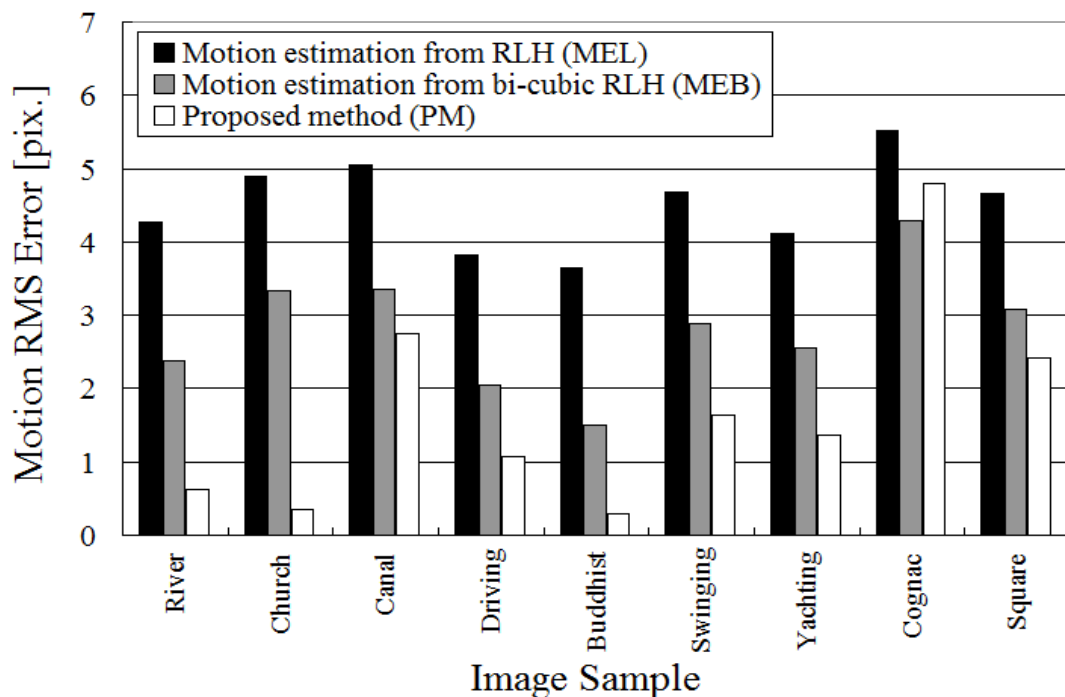


Figure 34 RMSE of estimated motions.

4.3.3. 提案手法による生成画像の画質向上

提案法によって生成される高解像度高フレームレートのカラー動画の画質を評価した。比較対象として、以下の4つの方法を用いる。

Frame-rate conversion (FC) : 線形内挿と動き補償を組み合わせたフレームレート変換手法。高解像度低フレームレートの動画像 R_{HL} , G_{HL} , B_{HL} を入力として、線形内挿により中間フレームを生成し、動き補償により中間フレームを高解像度化 [79] 。

Super resolution (SR) : 再構成型の超解像手法。低解像度高フレームレートの動画像 R_{LH} , G_{LH} , B_{LH} を入力として、フレーム間の動きを用いて高域成分を復元することで高解像度化 [68]。

SR of color-separated video (SRC) : 色ごとに露光時間と時空間解像度が異なる動画像を用いる超解像手法。低解像度高フレームレートの R_{LH} , B_{LH} と長時間露光した高解像度低フレームレートの G_{HL} を入力とし、画像復元の枠組みにより高解像度高

フレームレート化[81].

Proposed method without simultaneous estimation (MEB): 提案法と同じ入力に対し、動きと生成画像の同時推定をしない方法. R_{LH} , G_{HL} , B_{LH} を入力として、低解像度の動画像 R_{LH} を Bi-cubic 補間拡大した動画像からフレーム間の動きを推定 (4.3.2 節の MEB と同じ). 式(4-2)–(4-4)で高解像度高フレームレート化.

実験では、12 の標準動画像を正解動画像として、入力動画像を生成し、これを用いて各手法によって生成した動画像 R_{HH} , G_{HH} , B_{HH} の全フレームの CPSNR を算出した. 入力動画像の総画素数と拡大率は、全ての手法で同じである.

各手法による拡大率が 16 倍 ($N_p=N_t=16$) の生成画像の CPSNR を Figure 35 に示す. 提案法(PM)の生成画像の CPSNR は、全ての 12 動画像においてフレームレート変換手法 (FC) より 6.8~15.5dB 高い. また、提案法の CPSNR は、12 動画像中 11 において再構成型の超解像手法(SR)より 1.3~9.1dB 高く、色ごとに露光時間と時空間解像度が異なる動画像の超解像手法 (SRC) より 0.5~7.0dB 高い. また、提案法の CPSNR は、12 動画像中 9 において、低解像度の R_{LH} を Bi-cubic 補間拡大した動画像から動きを推定する手法(MEB)より 0.5~2.2dB 高い. 一方、Cognac, Square, Flash の 3 動画像においては、提案法の CPSNR は、MEB 法より 0.04~0.3dB 低い. Cognac と Square は R,G 間の相関係数がそれぞれ 0.847, 0.823 と低く、Flash は 1/16 縮小画像の CPSNR が 32.1[dB] と高域成分が少ない (Table 5). 前節及び 4.2.2 節で示したように、色間の相関が低い場合、又は高域成分が少ない場合、提案法の生成画像 R_{HH} は低解像度の入力動画像 R_{LH} に近づき、動きの推定精度が MEB 法より低くなる場合があるためと考えられる.

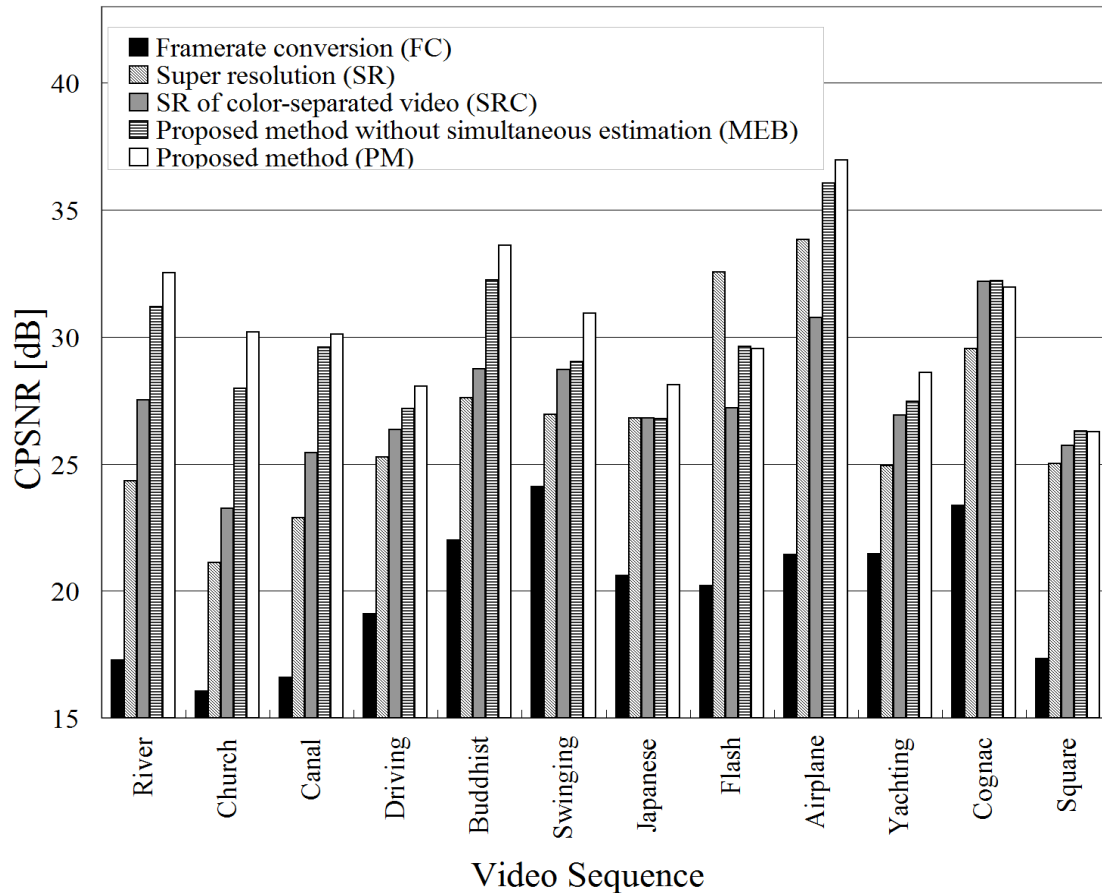


Figure 35 CPSNRs for ITE video sequences ($N_p=N_t=16$).

拡大率を 4~16 倍 ($N_t=N_p=4, 6, 9, 16$) とした場合の, 12 動画像をあわせた CPSNR を Figure 36 に示す. いずれの手法も拡大率が大きくなるに従い CPSNR は低くなる. 提案法は, 拡大率の増加にともなう CPSNR の低下が最も少なく, 拡大率 6 倍以上で他の 4 手法より CPSNR が 0.2~10.6dB 高い. また提案法は, 拡大率 6 倍以上で, 同時推定を行わない MEB 法より CPSNR が 0.2~0.9dB 高い. 一方, 拡大率 4 倍では MEB 法の CPSNR が 0.4dB 高い. この原因は, 4.2.2 節で述べたと同様に, 拡大率が小さいと G の高域成分が 0 に近くなり, 生成画像 R_{HH} は入力動画像 R_{LH} を最近傍法により拡大した動画像に近くなる. その結果, 入力動画像 R_{LH} を Bicubic 法により補間拡大した動画像から動きを推定する MEB より, 提案法の CPSNR が低くなると考えられる.

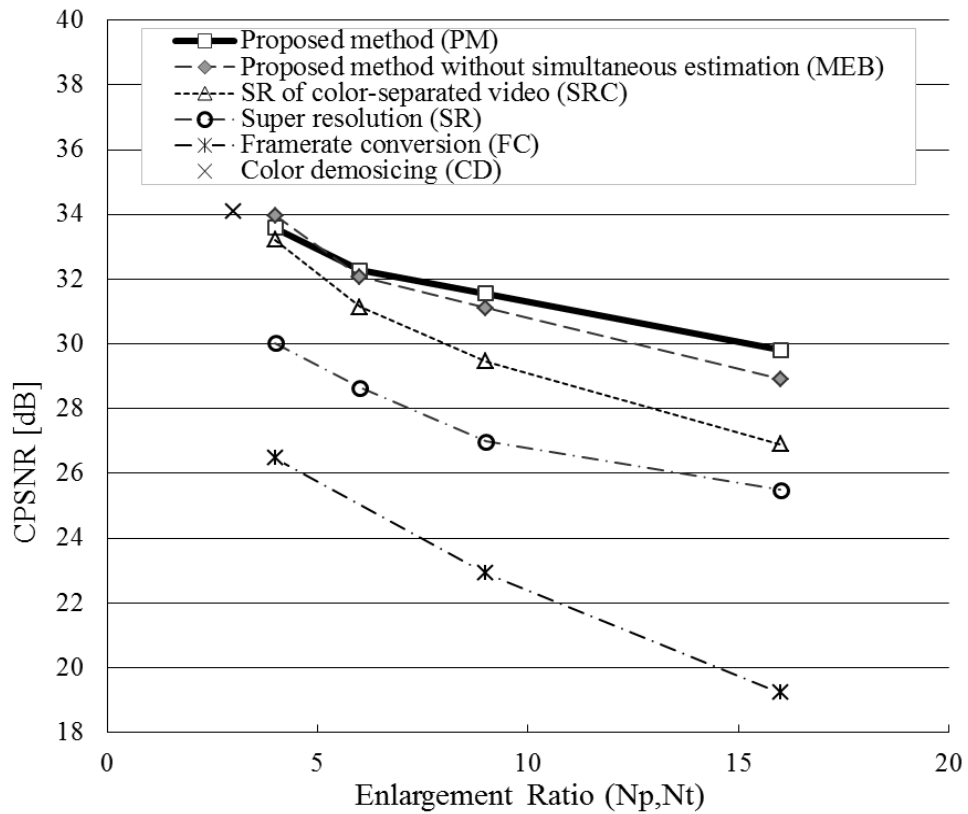


Figure 36 CPSNR of different enlargement ratio.

提案法及び SRC 法の生成画像の例を Figure 37~Figure 39 に示す. いずれも拡大率は 16 倍 ($N_t=N_p=16$) で, キーフレームから 8 フレーム目である. 提案法による生成画像 R_{HH} (下段右) は, 入力動画像 R_{LH} (上段右) から高解像度化されていることがわかる. SRC 法[81]による生成画像 (中段 2 像) に対して, 提案法による生成画像 R_{HH} , G_{HH} (下段 2 画像) は, 主観的な解像感が高く鮮明であることが確認できる. これは, SRC 法は長時間露光によって動領域に生じた動きブレにより高域成分が低減するのに対し, 提案法は, 動きぶれによる高域成分の低減がないためである.

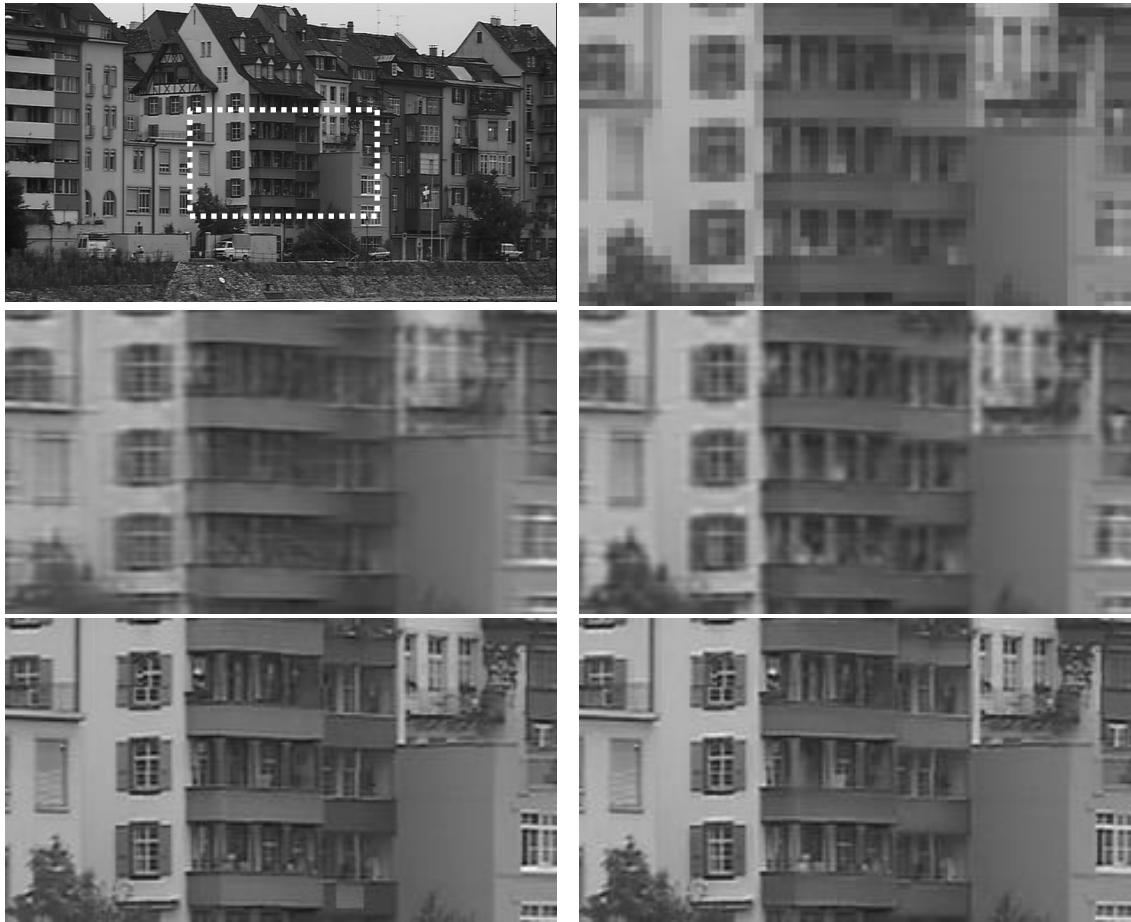


Figure 37 Reconstructed and input images of "Canal" ($N_p=N_t=16$).

Upper left: Ground truth image (gray), Upper right: Input image (R_{LH}),

Middle left: SR of color-separated video (green), Middle right: SR of color-separated video (red)

Lower left: Proposed method (G_{HH}), Lower right: Proposed method (R_{HH})



Figure 38 Reconstructed and input images of "Yachting" ($N_p=N_t=16$).

Upper left: Ground truth image (gray), Upper right: Input image (R_{LH}),

Middle left: SR of color-separated video (green), Middle right: SR of color-separated video (red)

Lower left: Proposed method (G_{HH}), Lower right: Proposed method (R_{HH})

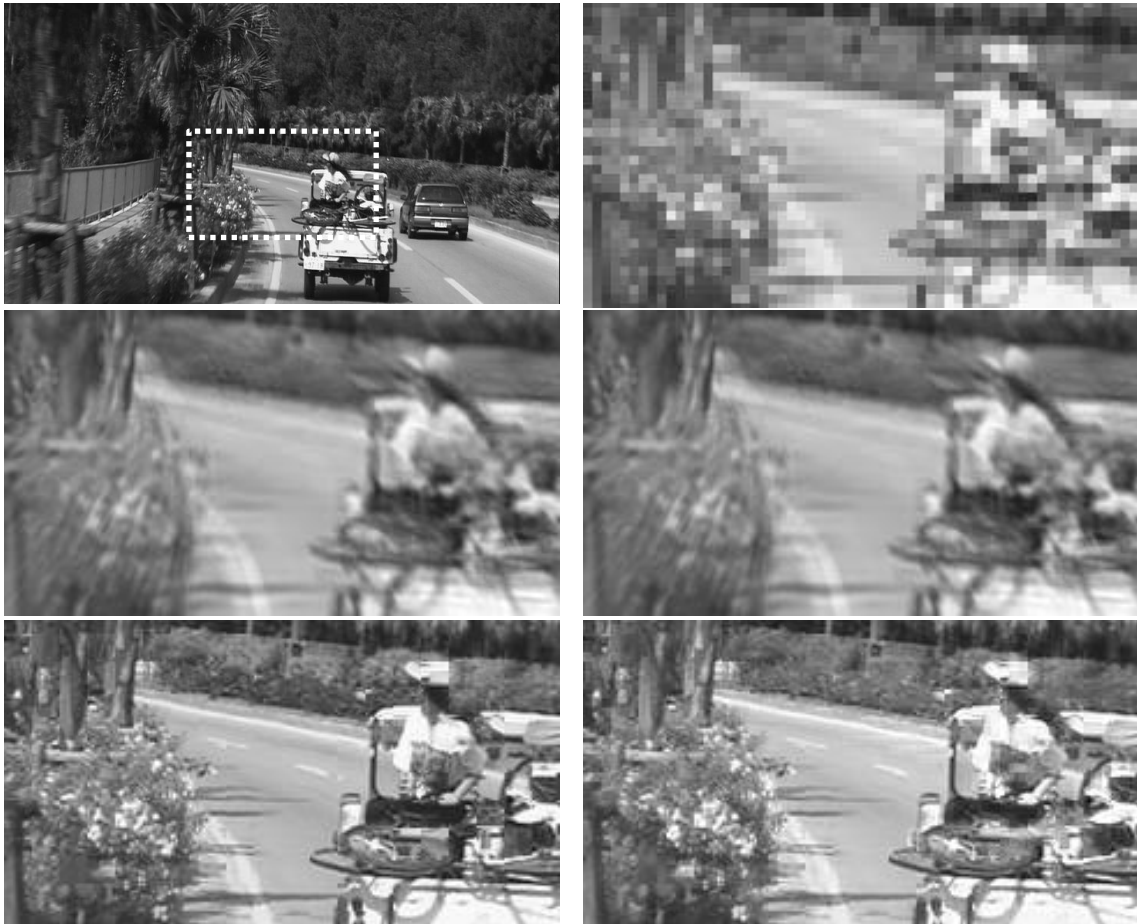


Figure 39 Reconstructed and input images of "Driving" ($N_p=N_t=16$).

Upper left: Ground truth image (gray), Upper right: Input image (R_{LH}),

Middle left: SR of color-separated video (green), Middle right: SR of color-separated video (red)

Lower left: Proposed method (G_{HH}), Lower right: Proposed method (R_{HH})

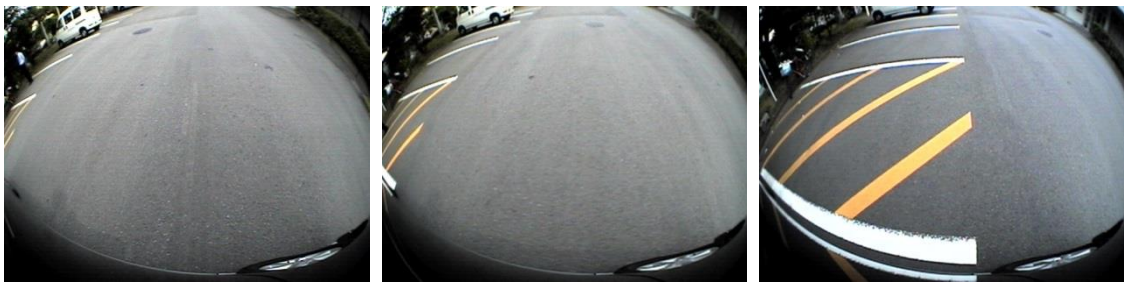
次の実験では、車の前後左右に設置した 4 台の車載カメラで撮影した駐車場シーンの動画像 (Figure 40) を正解動画像として用いて、入力動画像を生成し、これらの入力動画像を用いて各方法で生成した R_{HH} , G_{HH} , B_{HH} 動画像について CPSNR を計算した。

入力動画像の空間拡大率と時間拡大率はともに $16(Np=16, Nx=4, Ny=4, Nt=16)$ とした。提案手法により生成した入力画像 R_{LH} と中間フレーム R_{HH} の例を Figure 41 に示す。提案法による生成画像 R_{HH} (Figure 41b) は、入力動画像 R_{LH} (Figure 41a) から高解像度化されていることがわかる。

拡大率を 4~16 倍 ($Np=Nt=4, 6, 9, 16$) とした場合の、4 動画像をあわせた CPSNR 値を Figure 42 に示す。提案法(PM)は、拡大率 4~16 倍において、低解像度の R_{LH} から動きを推定する MEL 法よりも、CPSNR が高い。一方、提案法(PM)は、Bi-cubic 補間拡大画像から動きを推定する MEB 法と比較して、拡大率 16 倍 ($Np=Nt=16$) ではわずかに CPSNR が高いが、拡大率 4~9 倍では CPSNR ほぼ同じであった。これは、高周波成分がこれらのシーケンスに多く含まれていないためと考えられる。



(a) front



(b) left



(c) right



(d) Rear

Figure 40 Street-view video sequences for evaluation, captured by vehicle mounted cameras.



(a) Input image R_{LH} ($N_p=16$)

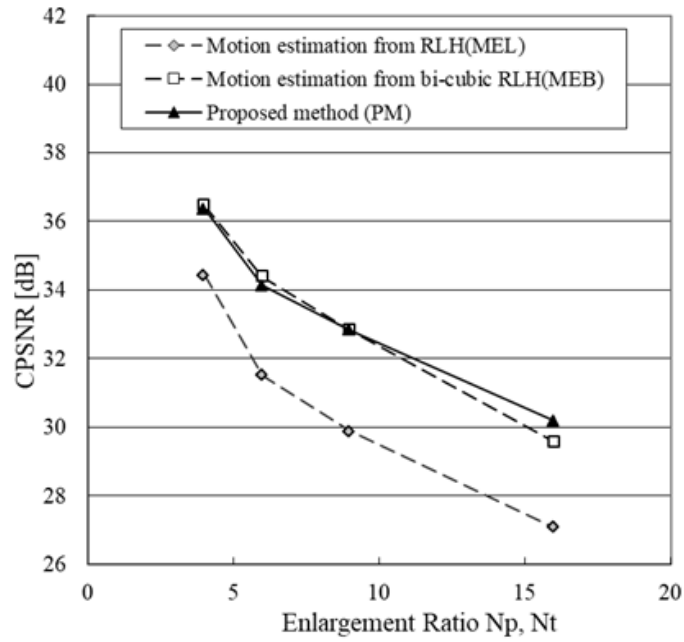


(b) Reconstructed image R_{HH}

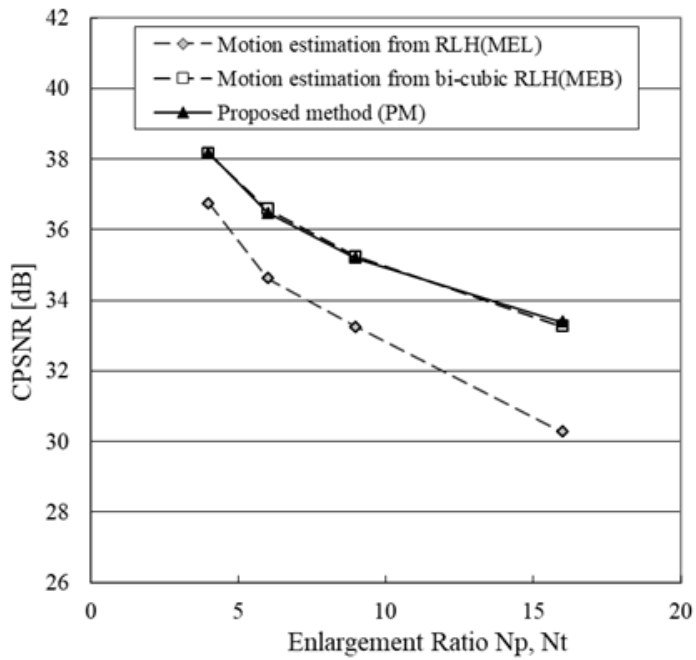


(c) Reconstructed color image ($N_p=N_t=16$)

Figure 41 Reconstructed images and input images.



(a) "Rear" sequence



(b) "Front" sequence

Figure 42 CPSNR of different enlargement ratio of street view sequences.

4.3.4. 提案手法による解像度の向上および折り返しの復元

提案法の局所色相関モデルによる解像度の向上の効果および折り返し復元効果を評価した。

ここでは、動き推定の影響を排除しながら、局所色相関モデルの効果を評価するため、 R_{LH} , G_{HL} , B_{LH} のキーフレームを入力として、式(4-2)および(4-3)により R_{HH} および B_{HH} を生成し、 G_{HL} と組み合わせた1フレームの R_{HH} , G_{HH} , B_{HH} を評価した。比較のため、 R_{LH} および B_{LH} のキーフレームを、最近傍法およびバイキュービック法により補間拡大し、 G_{HL} と組み合わせた画像を評価した。

各手法による拡大率16倍($N_p=N_t=16$)の生成画像のCPSNRをFigure 43に示す。局所色相関モデルを用いる提案法は、すべての評価画像において、最近傍法およびバイキュービック法よりもCPSNRが高い。提案法の局所色相関モデルは、全評価シーンのCPSNRにおいて、最近傍法に対して9.4 dB、バイキュービック法に対して8.6 dB、CPSNRが高い。

円形ゾーンプレートを入力画像とした場合の生成画像をFigure 44に示す。Figure 44bに示すように、入力画像 R_{LH} には強い折り返しが含まれる場合であっても、局所的な色画像間の相関に応じて、 G_{HH} に含まれる高周波成分を R_{LH} に重畳することで、Figure 44cに示すように折り返しを低減した画像を生成することができる。

以上の結果から、提案法の局所色相関モデルは、最近傍法やバイキュービック法に比べて、解像度を向上する効果と折り返しを復元する効果が高いといえる。

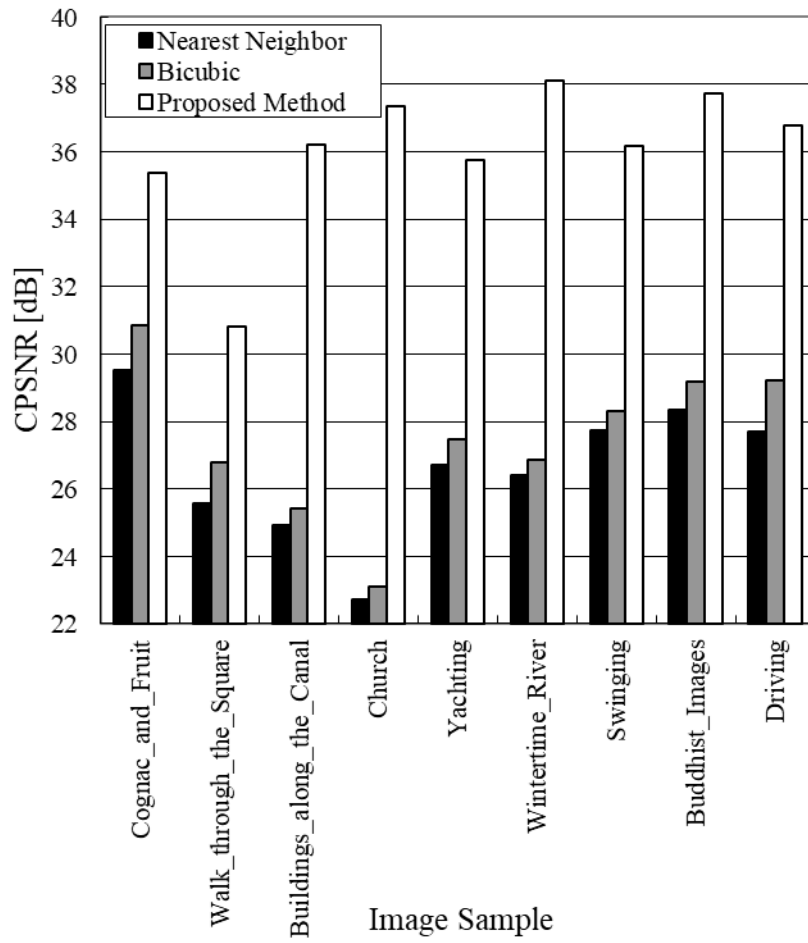
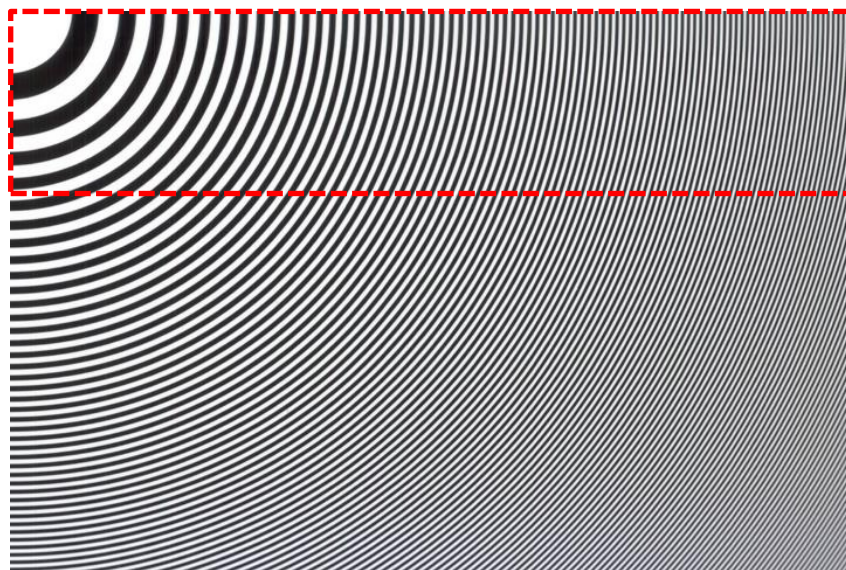
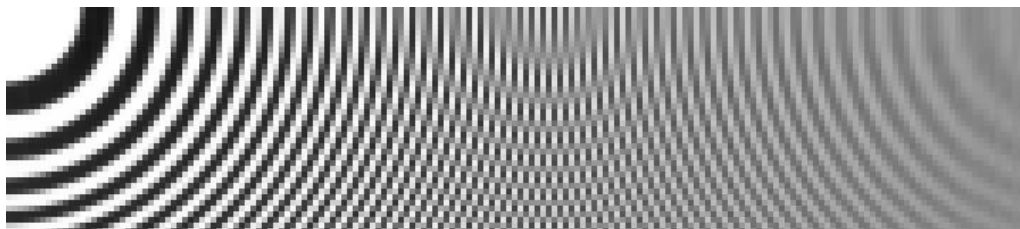


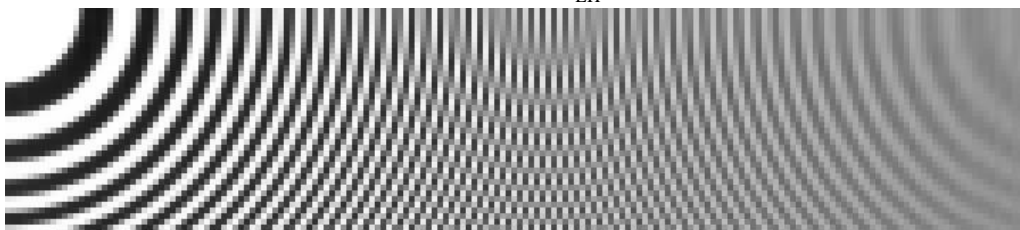
Figure 43 CPSNRs of generated high resolution images by color cross-correlation.



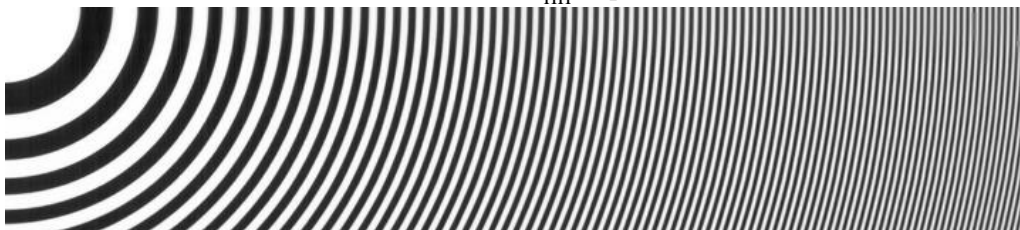
(a) Grand truth image R_{true}



(b) Input image R_{LH} ($N_p=16$)



(c) Bi-cubic R_{HH} ($N_p=16$)



(d) Proposed method R_{HH} ($N_p=16$)

Figure 44 Up-sampled images using color cross-correlation.

(a) Ground truth, (b) Input image, (c) Bi-cubic and (d) Proposed

4.3.5. 考察

提案法における動きと生成画像の同時推定の効果は、4.3.2 節の動きの誤差低減と、4.3.3 節の生成画像の誤差低減及び生成画像の主観的な解像感の向上として確認できた。同時推定による生成画像の誤差を低減する効果は、拡大率が高いほど顕著であり、拡大率 16 倍の動画生成において、CPSNR を 0.9dB 向上する効果を確認した。一方、拡大率が低い場合や、高域成分が少ない動画、及び、色間の相関が低い動画に対しては、同時推定の効果が低く、MEB 法より生成画像の CPSNR が低くなる場合がある。このことから、色間の相関が低い、又は、高域成分が少ない動画に対しては、生成画像 R_{HH} が R_{LH} の Bicubic 補間に近づくように、局所色相関モデルの式(4-1)を変更することによって、提案法の CPSNR の向上が期待できる。

本論文では、生成画像についての評価尺度として、信頼性が高く、広く使われている指標である CPSNR を用いた。一方、CPSNR は、見た目の自然さなどの主観的な基準とは必ずしも一致しない。高解像度高フレームレートと見た目の自然さの両立を検証するためには、主観評価を加えた検証が今後必要である。

4.4. 結言

本章では、仮想視点映像中の局所的な低解像度領域の解像度を向上させることを目的として、車載カメラ映像を高解像度化する方法について提案した。高解像度の動画を、データ伝送量を増やすことなく生成するために、RGB 3 色のカラー画像のうち、G を高解像度低フレームレートで撮影し、高解像度のカラー動画を生成する手法を提案した。従来法の課題であった、低解像度の入力動画から推定した動きの誤差を低減するために、提案法は、高解像度の生成画像においてフレーム間の画素値の差を最小化する位置変化としてフレーム間の動きを定義し、これを解くために動きと生成画像の同時推定手法を適用した。シミュレーション実験により、同時推定手法による動きの推定精度の向上と生成画像の誤差の低減の 2 つの効果を検証し、提案手法の有効性を確認した。また、提案法によって入力動画の総画素数を 4~16 倍に向上した高解像度高フレーム動画の CPSNR は、フレームレート変換手法、再構成型の超解像手法、文献[81]の 3 手法で生成した動画の CPSNR より高いことを確認した。

仮想視点映像中の局所的な低解像度領域の解像度向上の効果は、次章に示す。

5. 提案手法を用いた車両周辺監視用車載カメラシステムの評価実験

本章では，2～4章で提案した3つの手法からなる仮想視点映像生成手法を用いた車両周辺監視用車載カメラシステムが成立することを確認するため，下記の3項目の評価実験を行った．

- ・仮想視点映像中の低解像度領域の高解像度化の効果の評価
- ・データ伝送遅延の評価
- ・実際の駐車シーンの車載カメラ映像から生成した仮想視点映像の評価

5.1. 評価方法

仮想視点映像中の低解像度領域の高解像度化の評価

提案手法によって生成した仮想視点映像中の低解像度領域の高解像度化の効果を確認するため，生成画像の解像度について定量評価した．評価の手法を以下に示す．

生成された仮想視点画像の解像度を評価するために，解像度チャートを含む評価画像を用いたシミュレーション実験を行い，定量的に評価する．まず，解像度チャートを含む高解像度の映像（正解画像） R_{true} ， G_{true} ， B_{true} を用意する．次に，これらの正解画像を用いて， N_t フレームごとに G_{true} からフレームを抽出して G_{HL} を， R_{true} と B_{true} から $N_p (= N_x \times N_y)$ ピクセルの平均値を取るダウンサンプリングして R_{LH} と B_{LH} を，それぞれ生成する．次に，これら3つの画像を入力画像として，4章で提案した高解像度化手法により高解像度化した画像を生成する．最後に，これらの高解像度化した画像を用いて，2章と3章で提案した仮想視点映像生成手法により仮想視点映像を生成し，得られた仮想視点映像の解像度をベースラインの仮想視点映像と比較する（Figure 45）．

実験に用いる，前，左，右および後ろのカメラの4つの高精細映像の正解画像（Figure 46）は，レンダリングソフトウェアであるLightwave 3Dによって合成した．カメラパラメータをTable 7に示す．各拡大率（ $N_t = N_p = 1, 4, 6, 9, 16$ ）の入力画像は，正解画像から生成する．ベースラインビデオシーケンス（640×480ピクセル，30fps）も正解画像から生成する．各拡大率の入力画像シーケンスのピクセルの合計量は，ベースラインビデオシーケンスと同じである（Table 8）．

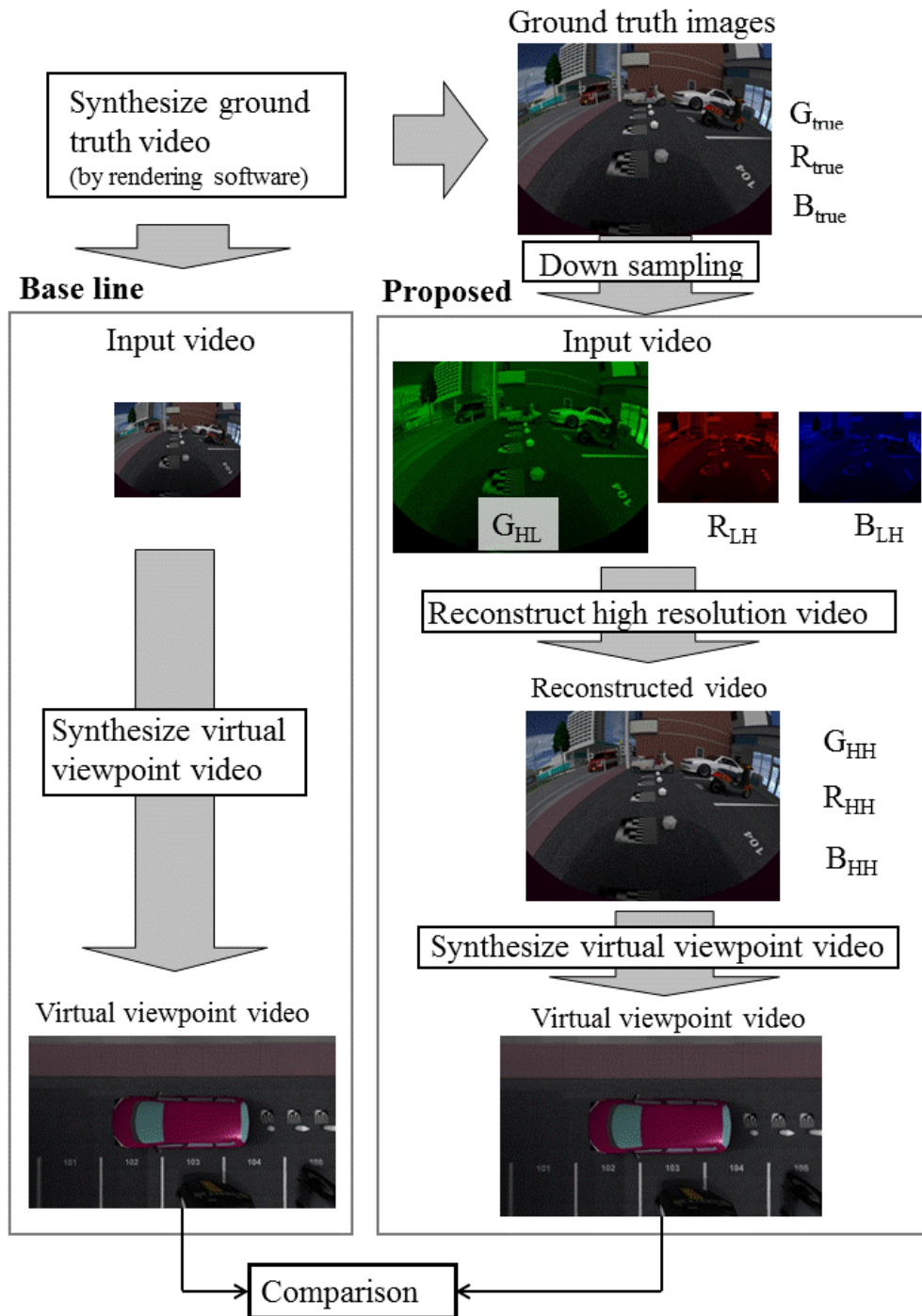
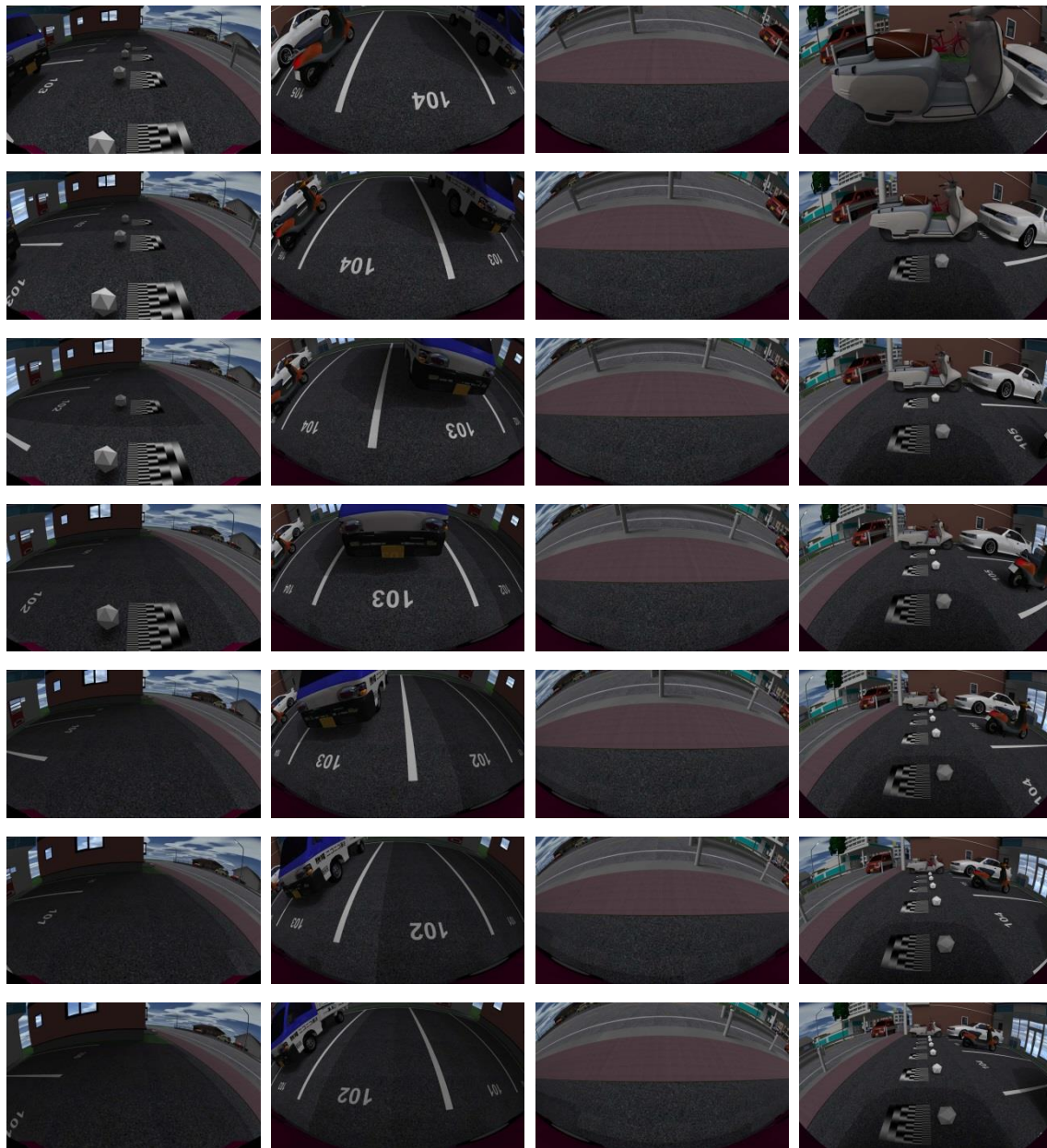


Figure 45 Flowchart of evaluation procedure.



(a) Front

(b) Left

(c) Right

(d) Rear

Figure 46 Test video sequences for evaluation.

Table 6 Camera parameters of vehicle mounted cameras and virtual viewpoint.

Camera	Position [mm]			Rotation [deg.]			FOV [deg.]
	x	y	z	Heading	Pitch	Bank	
Front	0	700	-30	0	30	0	150
Left	-960	1000	-1720	-90	50	0	150
Right	960	1000	-1720	90	50	0	150
Rear	0	860	-4780	180	30	0	150
Viewpoint A	0	10000	-7000	180	180	0	60
Viewpoint B	0	10000	-2500	180	180	0	60
Viewpoint C	0	10000	-6000	0	180	0	60

Table 7 Number of pixels and frame rate of input and reconstructed videos.

Enlarge- ment ratio	Input video			Reconstructed video
	G_{HL} [pix/f] x[f/s]	R_{LH}, B_{LH} [pix/f]x[f/s]	Total [pix/s]	G_{HH}, R_{HH}, B_{HH} [pix/f]x[f/s]
1 (base line)	640 x 480 x 30	640x480 x30	27,648k	640x480 x30
4	1280 x 960 x 7.5	640x480 x30	27,648k	1280x960 x30
6	1280 x 1440 x 5	640x480 x30	27,648k	1280x1440 x30
9	1920 x 1440 x 3.333	640x480 x30	27,648k	1920x1440 x30
16	2560 x 1920 x 1.875	640x480 x30	27,648k	2560x1920 x30

5.2. 評価実験結果

5.2.1. 仮想視点映像の低解像度領域の高解像度化

本節では、2章と3章で提案した手法により仮想視点映像を生成する際に、4章で提案した手法により高解像度化したカメラ映像を入力とすることによる、仮想視点

映像中の局所的な低解像度領域の改善効果を示す。

車載カメラ映像の高解像度化の効果

4章の提案手法による車載カメラ映像の高解像度化の効果を示す。この実験では、解像度チャートを含む正解画像を用いて、ベースライン映像および提案法の拡大率 ($Nt = Np=1,4,6,9,16$) によって生成した高解像度映像の2つの映像の解像度を比較する。提案法の各拡大率の入力画像の総画素数は、ベースライン映像の総画素数と同じである。

ベースライン映像 (640×480 [画素], 30 [フレーム/秒]) の画像例を Figure 47a に、提案法の拡大率16($Nt=Np=16$)により生成した高解像度映像の例を Figure 47b に示す。比較を容易にするため、Figure 47 と同じフレームの一部を拡大した画像を Figure 48 に示す。Figure 47 および Figure 48 から、提案法によって生成した映像は解像度が大幅に向上していることがわかる。

各拡大率で高解像度化した映像中の解像度チャートの解像度と振幅の関係を Figure 49 に示す。さらに、高解像度化した映像の上限解像度として、振幅が半減 (modulation=0.5) する解像度を抽出した結果を Table 8 に示す。Table 8 から、提案法は、拡大率4~16の場合に、解像度を1.9~3.5倍に向上できる。

以上の結果から、本節では4章で提案した高解像度化手法は、ベースライン映像のデータ量を増やすことなく、カメラ映像の解像度を1.9~3.5倍に向上した映像を生成できることを示した。

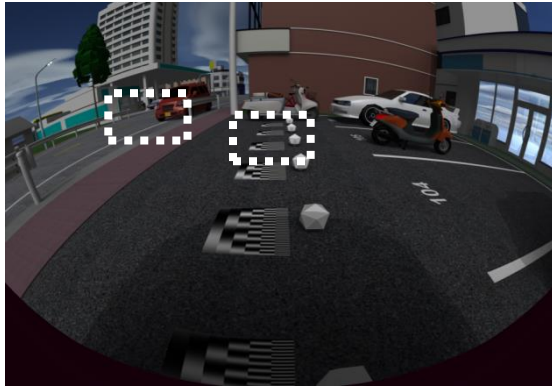


(a) Base line video (Rear, 270)



(b) Reconstructed high-resolution video (Rear, 270)

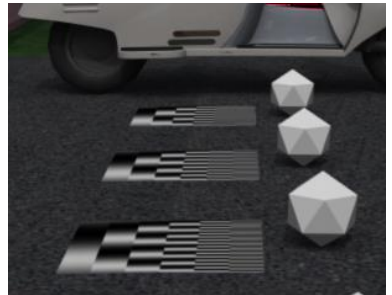
Figure 47 Reconstructed video from same amount of pixel data ($N_p=N_t=16$).



(a) input image (Rear 270 frame)



(b) Base line (enlarged)



(c) Proposed method $N_p=N_t=16$ (enlarged)

Figure 48 Examples of reconstructed image ($N_p=N_t=16$).

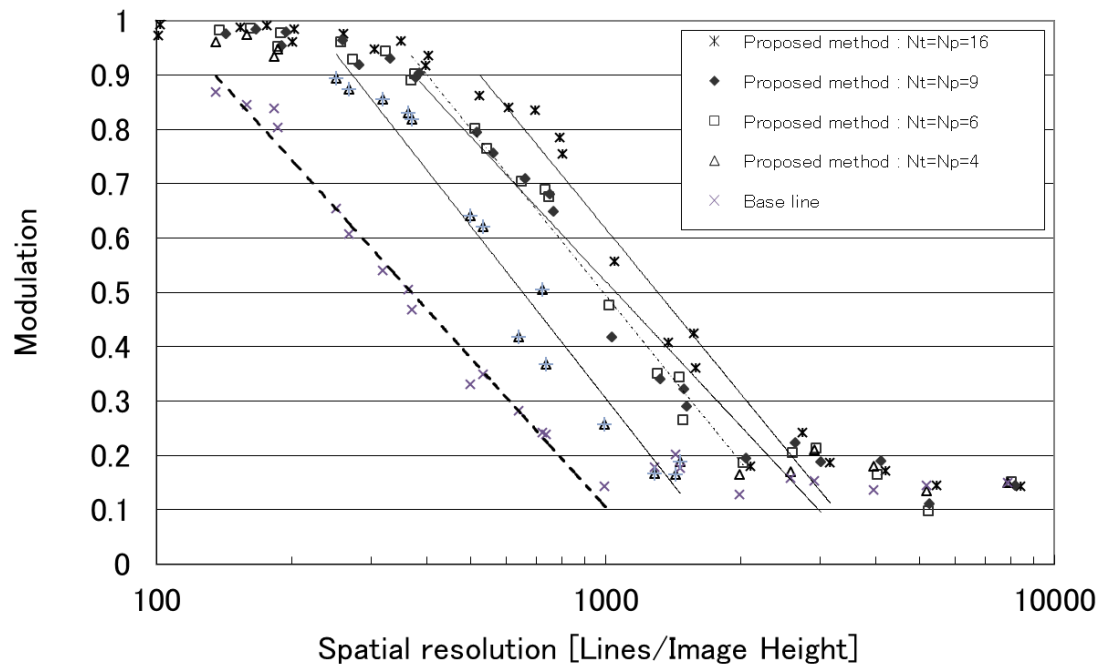


Figure 49 Resolution of reconstructed video

Table 8 Spatial resolution of reconstructed videos.

Enlargement ratio	Spatial resolution [lines/image height]	Resolution ratio to base line
1 (base line)	360	1.0
4	670	1.9
6	980	2.7
9	1030	2.9
16	1260	3.5

仮想視点映像の低解像度領域の高解像度化

2章と3章で提案した手法により仮想視点映像を生成する際に、4章で提案した手法により高解像度化したカメラ映像を入力とすることによる、仮想視点映像中の局所的な低解像度領域の改善効果を示す。

拡大率 1~16 ($Nt = Np=1,4,6,9,16$) で高解像度化した映像を用いて生成した仮想視点映像の例を、Figure 50 に示す。Figure 50 では、車の後方に 1m 間隔で路面上に解像度チャートを配置した。拡大率 1(Base line)の仮想視点映像 Figure 50a は、車から遠くなるほど解像度低下し、車の後方 4~5m 付近では、路面上の白線や文字が判別できないほど視認性が低下していることがわかる。これに対し、拡大率 4~16 で高解像度化した映像を用いて生成した仮想視点映像 Figure 50b では、拡大率が高いほど車の後方 4~5m 付近の解像度が高く、路面上の白線や文字の視認性が向上することがわかる。

Figure 51 は、ベースライン映像および高解像度化した映像($Nt = Np=1, 4, 6, 9, 16$)を使用して生成された仮想視点映像中の解像度チャートにおける解像度を、路面平面上の線密度と振幅の関係で示している。Figure 51a, Figure 51b に、それぞれ車の後方 5m と 4m における、解像度を示す。Figure 51a において、ベースラインの分解能は、約 5[lines/m](50%)であるのに対し、提案法の拡大率 6($Nt=Np=6$)で高解像度化した仮想視点映像の解像度は、約 16[lines/m](50%)であり、ベースラインよりも解像度を約 3 倍向上している。このことは、ベースライン映像から生成した仮想視点映像では、車の後方 5m において、典型的な 0.1m 幅の白線の表示に必要な解像度が無いのに対し、提案法により高解像度化した映像から生成した仮想視点映像では、十分な解像度を有していることを示している。

以上の結果により、4章で提案した高解像度化手法により、仮想視点映像中の局所的な低解像度領域を高解像度化できることと、その結果、低解像度領域における白線などの監視対象物の有無を把握しやすくなることを示した。



(a) Base line ($N_p=N_t=1$)



(b) Proposed method (enlargement ratio: $N_p=N_t=16$)



(c) Proposed method (enlargement ratio: $N_p=N_t=9$)

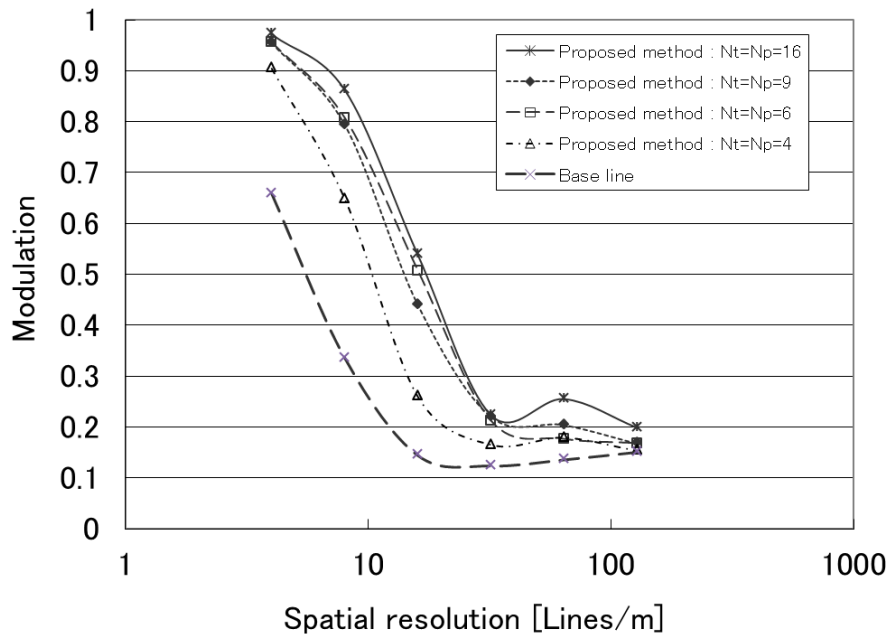


(d) Proposed method (enlargement ratio: $N_p=N_t=6$)

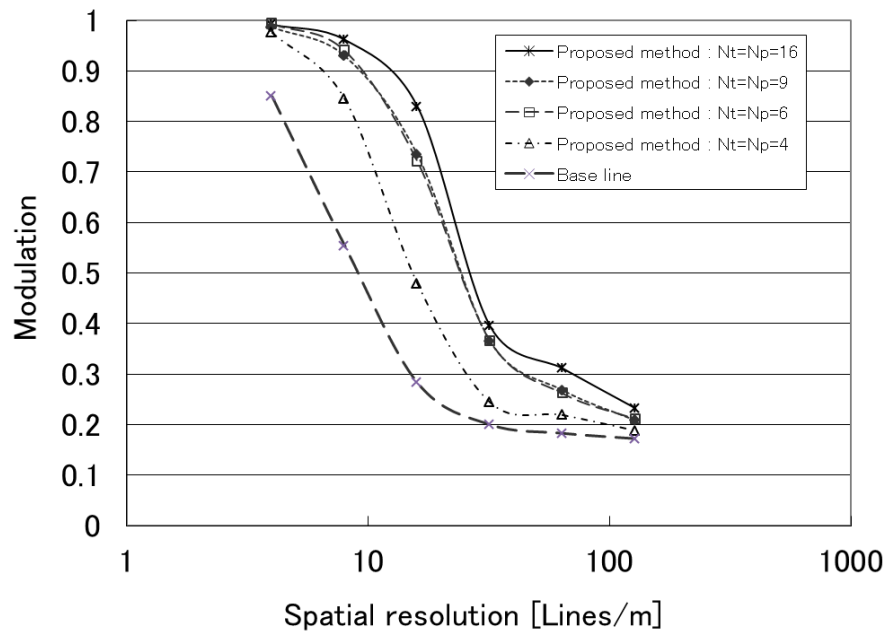


(e) Proposed method (enlargement ratio: $N_p=N_t=4$)

Figure 50 Generated virtual viewpoint video.



(a) Resolution at 5 [m] rear of the vehicle



(b) Resolution at 4 [m] rear of the vehicle

Figure 51 Resolution of virtual viewpoint videos

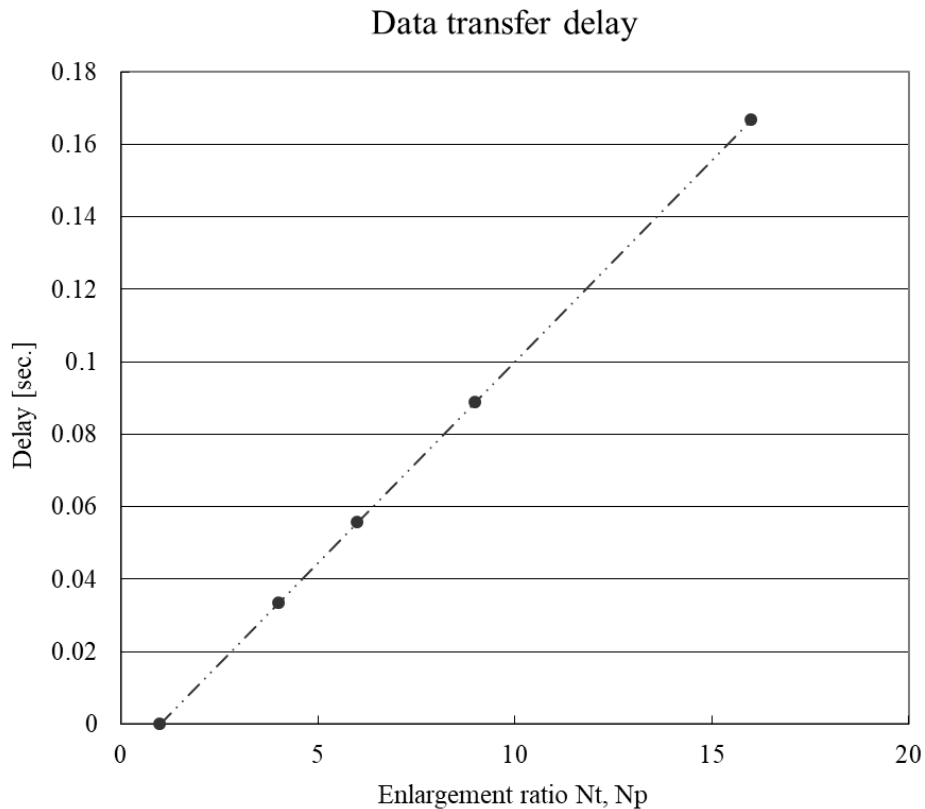


Figure 52 Data Transfer Delay of the Proposed Method

5.2.2. データ伝送遅延

本節では、提案手法を用いた場合のデータ伝送遅延を評価する。4章で提案した高解像度化手法は、1フレームあたり（1/30秒あたり）の画素データ量が異なる。そのため、データ伝送速度が一定のベースライン映像と同じデータ伝送路を利用した場合、撮像した画素データの伝送を開始してから、画像の再構成に必要なデータを伝送し終わるまでに遅延が発生する。1フレーム周期ごとの画素データ量は、拡大率が大きくなるほど差異が大きくなるため、拡大率が大きくなるほど伝送遅延が大きくなる。

そこで本節では、拡大率1～16（ $N_t = N_p = 1, 4, 6, 9, 16$ ）の場合の、ベースライン映像に対する遅延時間をシミュレーション実験により算出した。

拡大率ごとの遅延時間の実験を Figure 52 に示す。実験では、データ伝送速度は一

定（拡大率1のベースライン映像のデータ伝送速度と同じ）と仮定した。拡大率が9($Nt=Np=9$)の場合の伝送遅延は0.088[sec.]、拡大率が6($Nt=Np=6$)の場合の伝送遅延は0.056[sec.]、

車両周辺監視の用途においては、映像表示の遅延が少ないほど望ましい。一方、低速もしくは停止時の用途については、0.1[sec.]以上の遅延は許容されない場合があることから、提案手法は拡大率4, 6, 9の場合に、伝送遅延の点で実用的である。

5.2.3. 3つの手法により生成した仮想視点映像

本節では、2章~4章で提案した3つの手法により生成した仮想視点映像の結果を示す。

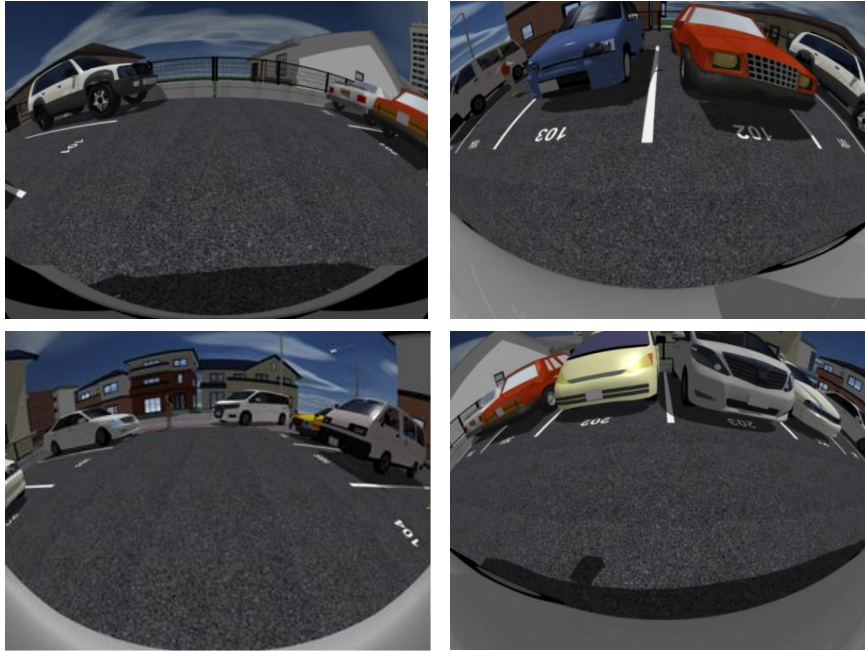
CGにより生成した駐車シーンの映像を用いて、2章~4章の各章で提案した手法により生成した仮想視点映像の例を、Figure 53に示す。Figure 53aは車載カメラで撮影した映像、Figure 53bは2章で提案した近似形状モデルを用いた仮想視点映像生成手法で生成した映像の例である。この手法は、被写体の3次元形状として、予め用意した近似3次元形状モデルを用いる。そのため、ガラスなどの透明物や金属などの鏡面反射物、タイヤなどの低反射率物などの、3次元形状推定が困難な被写体が含まれる映像(Figure 53a)から、車の周りを上空から俯瞰した仮想視点映像(Figure 53b)を生成できる。しかし、2章で提案した手法だけでは、180度以上の広視野で映像を生成すると、生成した像が歪むために、運転者が物の距離や方向の把握が難しくなる。そこで3章では複合射影カメラモデルを用いた仮想視点映像生成手法を提案した。Figure 53cは、3章で提案した複合射影カメラモデルを用いて生成した映像の例である。この手法は、画像生成のカメラモデルとして、透視射影と等距離射影の2つを組み合わせた複合射影モデルを用いる。これにより、車の近くは透視射影を用いることで路面上の直線の歪みが無く、遠くは等距離射影を用いることで視野が広い映像(Figure 53c)を生成できる。しかしながら、近似3次元形状モデル、複合射影モデルを併用しても、仮想視点映像中に、解像度が大きく低下する領域が発生して、低解像度領域中の物の有無を把握しにくくなる。そこで、本研究では、上記のモデルを併用した上で、この問題を解決するカメラ映像の高解像度化法を提案した。Figure 53dは、4章で提案した高解像度化手法を用いて生成した映像の例である。この手法は、カラー映像うちの1色を高解像度かつ低フレームレートで撮像することで、高解像度情報を含む映像を、映像データ量を増やすことなく取得する。そし

て、フレーム間の動きと高解像度映像を同時に推定する動き補償型超解像処理によって、カラー映像を高解像度化する。これにより、映像データ量を増やすことなく、仮想視点映像中の低解像度領域を高解像度化した映像(Figure 53d)を生成できる。

実際の駐車シーンにおいて車載カメラで撮影した映像を用いて、2章~4章の各章で提案した手法で生成した仮想視点映像の例を、Figure 54, Figure 55, Figure 56に示す。Figure 54は、並列駐車、縦列駐車、店舗前駐場での3つの駐車シーンの仮想視点映像の例である。金属光沢のあるポール(Figure 54b)や、網状の柵(Figure 54c)が車の近くにある場合でも、物の有無、車と物との距離やその方向がわかり易い仮想視点映像を生成できることが確認できる。Figure 55a および Figure 56は、2章、3章で提案した手法で、Figure 55b および Figure 56b は、本論文の3つの提案手法で、それぞれ生成した仮想視点映像である。Figure 55a および Figure 56 に対して、Figure 55b および Figure 56b は解像度が向上している。このことから、4章で提案した高解像度化手法は、車載カメラで撮影した映像に対しても有効であることが確認できた。

5.2.4. 考察

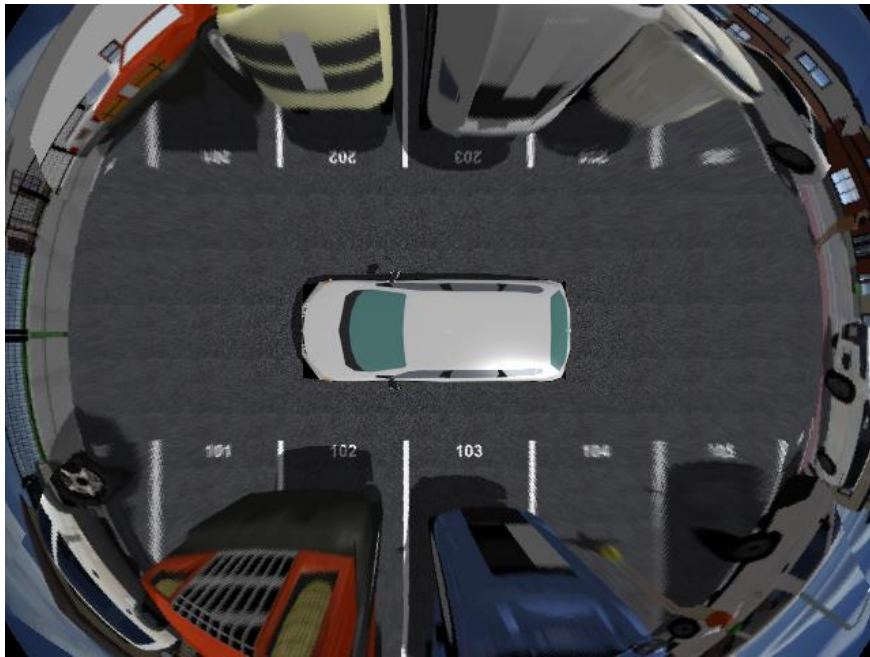
4章で提案した高解像度化手法の効果は、5.2.1節の実験において、2章と3章で提案した方法で生成した仮想視点映像中の低解像度領域の解像度を向上できることと、白線や数字の視認性を改善できることとして、確認できた。4章で提案した高解像度化手法は、データ伝送量一定の条件において伝送が遅延するが、5.2.2節の実験において、データ伝送遅延が実用に耐えうる範囲であることを確認できた。2~4章で提案した3つの手法からなる仮想視点映像生成手法の効果は、5.2.3節の実験において、駐車シーンのCG画像および実際に車載カメラで撮影した映像から、広い視野かつ車の周りの直線が画像上でも直線である全周囲画像を生成できることとして、確認できた。



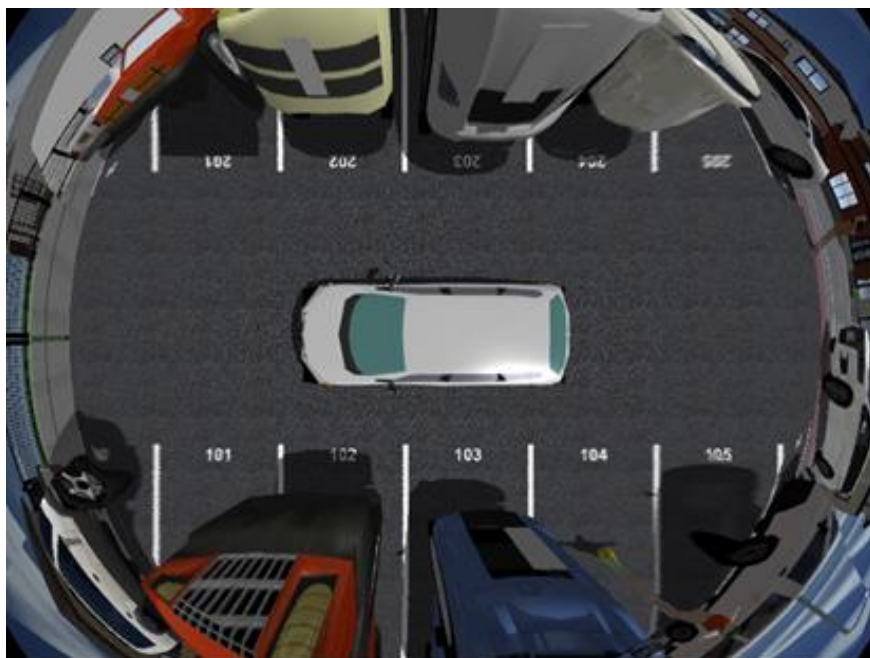
(a) Input images from vehicle mounted cameras



(b) A virtual viewpoint image generated by the method proposed in Chapter 2.
(Perspective projection, a plane shape model)

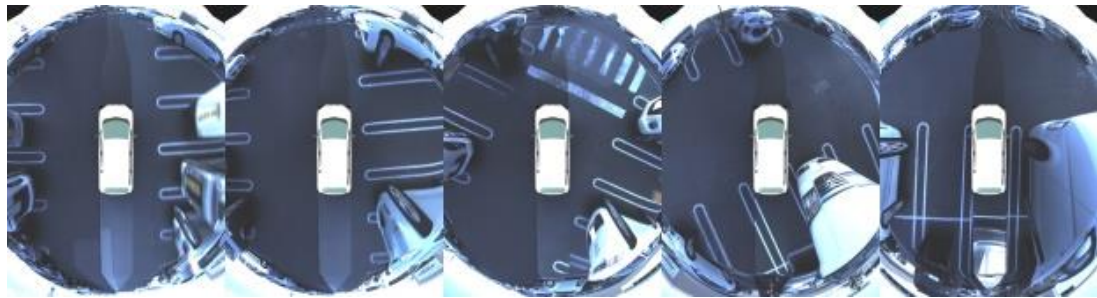


(c) A virtual viewpoint image generated by the method proposed in Chapter 3.
(Composite projection, a half-pipe shape model)



(d) A virtual viewpoint image with improved resolution by the method proposed in Chapter 4 ($N_t=N_p=6$).

Figure 53 Virtual viewpoint video generated by the three proposed methods.



(a) Side by side parking lot

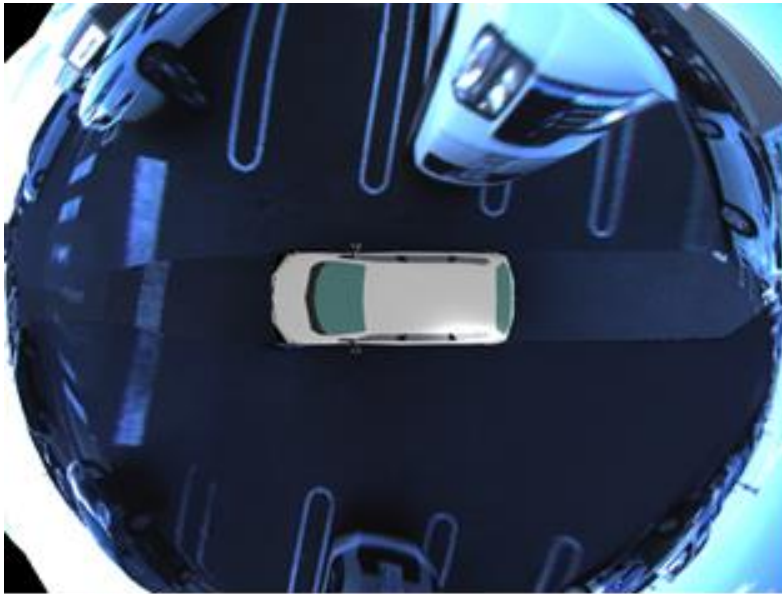


(b) Parallel parking lot



(c) Parking lot adjacent to the store

Figure 54 Virtual viewpoint videos generated by the three proposed methods.



(a) A virtual viewpoint image generated by the method proposed in Chapter 3.
(Composite projection, a half-pipe shape model)



(b) A virtual viewpoint image with improved resolution by the method proposed in Chapter 4 ($N_t=N_p=6$).

Figure 55 Virtual viewpoint images of “Side by side parking lot“.



(a) A virtual viewpoint image generated by the method proposed in Chapter 3.
(Composite projection, a half-pipe shape model)



(b) A virtual viewpoint image with improved resolution by the method proposed in Chapter 4 ($N_t=N_p=6$).

Figure 56 Virtual viewpoint images of “Parallel parking lot“.

5.3. 結言

本章では，本論文で提案した（１）近似３次元形状モデルを用いた仮想視点映像生成手法，（２）複合射影カメラモデルを用いた仮想視点映像生成手法，（３）色ごとに異なる時空間解像度が異なる動画像を用いた高解像度化手法の３つの手法により，車の近くは路面上の直線の歪みが無く，かつ，実空間中の距離と画像上の距離が比例し，車から遠くは，水平線以上を含む視野が広い仮想視点映像を生成できることを，示した．

以上の結果から，2～4章で提案した３つの方法からなる仮想視点映像生成方法は，車に設置した複数のカメラで撮影した映像から，車の周りの物の有無やその距離や方向を把握しやすい映像を生成することができる，車両周辺監視用の車載カメラシステムに適用できることを結論づける．

6. 結論

6.1. 本論文のまとめ

本論文では，車に設置した複数のカメラで撮影した映像から，車の周りにおける障害物や白線などの監視対象物の有無や，監視対象物までの距離や方向の把握が容易な映像を生成できる，仮想視点映像生成方式の開発を目的とした．

確立した方式によって，監視対象物の有無やその距離や方向の把握が容易な全周囲の映像を運転者に提示する車両周辺監視用の車載カメラシステム(Figure 2)の実現を目指した．

上記目的の達成のため，以下の3つの重要な問題に焦点を当ててきた．

Problem 1: モデルベースの自由視点映像生成方法を車載カメラシステムに適用する場合の根本的な課題は，システムが稼働する駐車場や道路には，透明物，鏡面反射物，低反射物などの3次元形状を推定できない被写体が存在することである．そのため，被写体の3次元形状の推定する従来のモデルベース手法では，自由視点映像を生成できないことである．

Problem 2: 自由視点映像生成方法の課題は，180度以上の広い視野角の映像を得るために，魚眼レンズの広視野角カメラモデルを用いて生成した映像において，像が歪み，路面上の直線が画像上では曲線になることである．その結果，画像上での車と物との距離が実際の距離が比例しなくなり，車と物との距離や方向がわかりにくいことである．

Problem 3: 自由視点映像生成方法を，同一領域を撮影するカメラ台数が少なく，かつ，車両の上空から俯瞰する視点など，実カメラから離れた位置に仮想視点がある場合に適用する場合の課題は，仮想視点画像中に局所的な低解像度領域が発生することである．

Problem 1 を解決し，駐車場や道路で，車両上空から俯瞰した仮想視点映像を生成

するために、カメラの映像を3次元形状モデルへ投影し、投影された映像を仮想的な視点から見た映像として生成する方法を提案した。この方法は、車の周りの3次元形状を推定する代わりに、平面や球などの近似的な3次元形状モデルを選択的に用いる。そのため、3次形状を推定できない物が存在する場合でも、仮想視点映像を生成することができる。リアルタイム仮想視点映像生成ハードウェアを搭載した実験車両による路上評価により、提案法が3次形状を推定できない物が存在する場合でも、仮想視点映像を生成できることを示した。

Problem 2 を解決し、車と物との距離や方向が分かりやすい全周囲の仮想視点映像を生成するために、複合射影カメラモデルを用いた映像生成方法を提案した。この方法は、全周囲映像を生成する際のカメラモデルとして、透視射影カメラモデルと等距離射影カメラモデルを組み合わせた複合射影カメラモデルを用いる。これにより、車の近くは透視射影を用いることで歪みが少なく、遠くは等距離射影を用いることで視野が広い、映像を生成することができる。実験では、実際の駐車シーンの映像から、視野が広く、車の近くの歪みが少ない全周囲映像を生成できることを示した。

Problem 3 を解決し、データ伝送量を増やすことなく、カメラ映像を高解像度化する方法を提案した。この方法は、色ごとに時空間解像度の異なる映像を撮影し、時空間の情報を統合することで高解像度高フレームレートの映像を生成するため、データ伝送量を増やすことなく、カメラ映像を高解像度化できる。標準動画像を用いたシミュレーション実験により、カメラ映像を高解像度化できることを示した。

さらに、提案した3つの手法が車両周辺監視システムに適用可能であることを示すため、提案手法で生成される仮想視点映像の解像度の評価と、データ伝送遅延の評価を行った。シミュレーション実験により、提案手法は、車載カメラ映像から、車の近くは路面上の直線の歪みが無かつ実空間中の距離と画像上の距離が比例し、車から遠くは水平線以上を含む視野が広い仮想視点映像を、実用的な解像度とデータ伝送遅延の内で、生成できることを示した。

これらの成果により、車両周辺監視用車載カメラシステムを実現するために、車に設置した複数のカメラで撮影した映像から、車の周りの物の有無やその距離や方向を把握しやすい映像を生成できる仮想視点映像生成方法を確立できたと結論づける。

6.2. 今後の展望

本論文では，車に設置した複数のカメラで撮影した映像から，車の周りにある障害物や白線などの監視対象物の有無や，監視対象物までの距離や方向の把握が容易な映像を生成できる，仮想視点映像生成方式方法について説明した．3つの重要な問題は解決されたが，次に示す事項については考慮する価値がある．

仮想視点映像生成の新たな用途への展開：第2章で説明した仮想視点映像の生成方法は，消費者製品としての車両周辺監視システムに既に適用されている．提案されたロバストな方法は，セキュリティカメラシステムなど，頑強性が必要な他の産業用途にも展開することができる．

3次元計測デバイスとの組み合わせ：Kinect やライトフィールドカメラなどの3次元計測デバイスの性能が，近年，急速に向上している．最新の3次元計測デバイスを用いても，ガラスや鏡面反射物など，未だに3次元計測ができない被写体が存在するものの，3次元計測デバイスと提案手法を相互補完的に併用することにより，被写体形状に依存する仮想視点映像の歪みを改善することが期待できる．

圧縮センシング技術の適用：第4章で説明した高解像度ビデオを再構成する方法は，できるだけ少ない画像シーケンスデータを取得し，目標カラー画像シーケンスが疎であるという仮定の下で，取得した少数のデータから画像シーケンスを再構成する．この概念は，近年，研究開発が進められている圧縮センシング[88]の概念と共通する．カラービデオシーケンスを捕捉するための圧縮センシングの理論を適用することにより，より効果的な方法を達成することが期待できる．

謝辞

本論文をまとめるにあたり，終始温かい激励とご指導，ご鞭撻を頂いた奈良先端科学技術大学院大学情報科学研究科環境知能学研究室 萩田紀博教授に心より感謝申し上げます．学位論文審査において，副査として，貴重なご指導と助言をいただいた奈良先端科学技術大学院大学情報科学研究科光メディアインターフェース研究室 向川康博教授，奈良先端科学技術大学院大学情報科学研究科環境知能学研究室 神原誠之准教授に心より感謝申し上げます．本学博士後期課程在学中から研究指導認定退学も学位取得までの長きにわたり，叱咤激励と的確なご指導，ご鞭撻を頂いた豊田工業大学大学院工学研究科 浮田宗伯教授に心より感謝申し上げます．

本研究を博士論文としてまとめるきっかけを与えていただくとともに，研究全体の方向付けをご指導いただきました木戸出正継様（元奈良先端科学技術大学院大学教授）に心より感謝申し上げます．

本論文の高解像度化の研究において，ベースになった研究当初から長きにわたりご指導をいただきました奈良先端科学技術大学院大学 横矢直和教授（現在学長）に心より感謝申し上げます．

パナソニック株式会社の旧先端技術研究所およびイノベーション推進本部およびの皆様には，業務と学位取得の両立において，ご指導・ご高配いただいたことを感謝いたします．

最後に，長い間温かく見守って下さった両親，社業と学業を影で支えてくれた妻 祥子，息子 竜之進，真志に感謝いたします．

参考文献

- [1] "Federal Motor Vehicle Safety Standard, Rearview Mirrors; Federal Motor Vehicle Safety Standard, Low-Speed Vehicles Phase-in Reporting Requirements", *Department of Transportation, National Highway Traffic Safety Administration*, NHTSA-2010-0162, Nov.2010.
- [2] "The approximation of the laws of the Member States relating to the type-approval of devices for indirect vision and of vehicles equipped with these devices", *The European Parliament and the Council of the European Union*, Amending Directive 70/156/EEC and repealing Directive 71/127/EEC, Nov. 2003.
- [3] "A notification of the partial amendments of the security standard of the road transportation vehicle and the detailed regulations of the security standard of the road transportation vehicle", *Ministry of Land, Infrastructure and Transport*, http://www.mlit.go.jp/kisha/kisha03/09/090707_.html#1, Jul. 2003.
- [4] M. Levoy and P. Hanrahan, "Light Field Rendering", *Proceedings of 23rd annual conference on Computer graphics and interactive techniques (SIGGRAPH'96)*, ACM, pp. 31-42, 1996.
- [5] S.J. Gortler, R. Grzeszczuk, R. Szeliski, and M.F. Cohen, "The lumigraph", *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques (SIGGRAPH '96)*, ACM, vol. 30, pp. 43–54, 1996.
- [6] C. Chang and T. Chen, "A Survey on Image-based Rendering Representation, Sampling and Compression", *Signal Processing: Image Communication*, Vol. 19, Issue 1, pp. 1-28, Jan. 2004.
- [7] S.C. Chan, Heung-Teung Shum, King-To Ng, "Image-based rendering and synthesis", *IEEE Signal Processing Magazine*, vol.24, Issue 6, pp.22-33, Nov. 2007.

- [8] T. Kanade, P. J. Narayanan, and P. W. Rander, "Virtualized Reality: Concepts and Early Results", *Proceedings of the IEEE Workshop on the Representation of Visual Scenes (in conjunction with ICCV'95)*, pp. 69-76, June 1995.
- [9] T. Matsuyama, X. Wu, T. Takai, and S. Nobuhara, "Real-Time 3Dshape Reconstruction, Dynamic 3D Mesh Deformation and High Fidelity Visualization for 3D Video", *International Journal on Computer Vision and Image Understanding*, Vol.96, No.3, pp. 393-434, Dec. 2004.
- [10] 高井勇志, 松山隆司, "Harmonized Texture Mapping", *映像情報メディア学会誌*, 63 卷, 4 号, pp. 488-499, 2009.
- [11] T. Takai, A. Hilton, and T. Matsuyama "Harmonized Texture Mapping", *The Fifth International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT2010)*, Paris, France, May 2010.
- [12] D. Scharstein, "View synthesis using stereo vision", *Springer-Verlag*, 1999.
- [13] Yosuke Ito and Hideo Saito, "Free Viewpoint Image Synthesis using Uncalibrated Multiple Moving Cameras", *Proceedings of Mirage 2005*, INRIA Rocquencourt, France, pp. 173-180, Mar. 2005.
- [14] S. Moezzi, L.C. Tai, and P. Gerard, "Virtual view generation for 3D digital video", *IEEE Multimedia*, vol. 4, no. 1, pp. 18-25, 1997.
- [15] S. Chen and L. Williams, "View Interpolation for Image Synthesis", *Proceedings of ACM SIGGRAPH 93*, pp.279-288, 1993.
- [16] W. Matusik, C. Buehler, R. Raskar, S. J. Gortler, and L. McMillan, "Image-based visual hulls", *Proceedings of the 27th annual conference on Computer graphics and interactive techniques (SIGGRAPH '00)* , pp. 369-374, 2000.
- [17] K. Kutulakos and S. Seitz, "A theory of shape by space carving", *International Journal of Computer Vision*, vol. 38, no. 3, pp. 199-218, 2000.
- [18] G. Slabaugh, W. B. Culbertson, T. Malzbender, M. Stevens, and R. Schafer, "Methods for volumetric reconstruction of visual scenes", *International Journal of Computer Vision*, vol. 57, no. 3, pp. 179-199, 2004.

- [19] Y. Furukawa and J. Ponce, "Accurate, Dense, and Robust Multiview Stereopsis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pp. 1362-1376, Aug. 2010.
- [20] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. "A comparison and evaluation of multi-view stereo reconstruction algorithms", *2006 IEEE Computer society conference on Computer vision and pattern recognition*, pp. 519-528, 2006.
- [21] D.G. Lowe, "Object recognition from local scale-invariant features", *Proceedings on International. Conference on Computer Vision 2.* pp. 1150-1157, 1999.
- [22] H. Bay, T. Tuytelaars, L.V. Gool, "SURF: speeded up robust features", *Proceedings on Ninth European Conference on Computer Vision*, 2006.
- [23] R. Dovgand, R. Basri, "Statistical symmetric shape from shading for 3Dstructure recovery of faces", *ECCV*, 2004.
- [24] A. Tankus, N. Sochen, Y. Yeshurun, "Perspective shape-fromshading by fast marching", *IEEE Computer Vision and Pattern Recognition*, 1, pp.43-49, 2004.
- [25] Y. Furukawa, J. Ponce, "Carved visual hulls for image-based modeling", *International Journal on Computer Vision*, 81 (1), pp.53-67, 2008.
- [26] J.L. Landabaso, M. Pargas, J.R. Casas,"Shape from inconsistent silhouette", *Computer Vision and Image Understandings*, 112, pp.210-224, 2008.
- [27] J. Barron, J. Malik, "Shape, illumination, and reflectance from shading", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 37 (8), 1670-1687, 2015.
- [28] S.M. Seitz, C.R. Dyer, "Photorealistic scene reconstruction by voxel coloring", *IEEE Computer Vision and Pattern Recognition*, pp.1067-1073, 1997.
- [29] A.Yezzi, et al., "A surface evolution approach to probabilistic space carving", *IEEE 3DPVT*, pp.618-621, 2002.
- [30] J.L. Landabaso, M. Pargas, J.R. Casas, "Shape from inconsistent silhouette", *Computer Vison and Image Understandings*, 112, pp.210-224, 2008.

- [31] G. Haro, "Shape from silhouette consensus", *Pattern Recognition*, 45, pp.3231-3244, 2012.
- [32] M. Zollhofer, et al., "Real-time non-rigid reconstruction using and RGB-D camera", *ACM Transaction on Graphics*, vol.33 Issue 4, no.156, July 2014.
- [33] H. Habe, Y. Katsura, and T. Matsuyama, "Skin-off: Representation and Compression Scheme for 3D Video", *Picture Coding Symposium (PCS) 2004*, pp. 301-306, San Francisco, 2004.
- [34] R. Y. Tsai, "An efficient and accurate camera calibration technique for 3D machine vision", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR '86)*, pp. 364-374, 1986.
- [35] S. Okamoto, M. Nakagawa, K. Nobori, A. Morimura, T. Okada, S. Masuda and Y. Fujita. "Development of parking assistance system using virtual viewpoint image synthesis", *Proc. of the 7th World Congress on Intelligent Transport Systems*, 3075S, Nov. 2000.
- [36] 岡本修作, 中川雅通, 登一生, 森村淳, "画像生成装置, 画像生成方法", *特許公報*, 特許第 3286306 号 (US patent, US7307655 B1, European patent EP1115250B1), Nov. 1998.
- [37] 登一生, 中川雅通, 石井浩史, 岡本修作, "画像処理装置および監視システム", *特許公報*, 特許第 3300337 号, 2002.04.19, *United States Patent* US6734869 B2, *United States Patent* US7714887 B2, *European Patents* EP1150252B1.
- [38] 岡本修作, 中川雅通, 登一生, 森村淳, "画像処理装置および監視システム", *特許公報*, 特許第 3300334 号, 2002.04.19, *United States Patent* US7161616 B1, *European Patents* EP1179958B1.
- [39] 登一生, 浮田宗伯, 萩田紀博, "複合射影カメラモデルと半円柱形状モデルを用いた車載カメラシステムのための低ひずみ全周囲画像生成手法", *電子情報通信学会論文誌D*, Vol. J101-D, No.5, pp.780-790, May 2018.
- [40] K. Nobori, N. Ukita and N. Hagita, "A surround view image generation method with low distortion for vehicle camera systems using a composite projection", *Proc. of the 15th IAPR Int. Conf. on Machine Vision Applications*, pp.386-389, May 2017.

- [41] 登一生, 今川太郎, 本村秀人, 吾妻健夫, “色ごとに時空間解像度が異なる動画像に対する動きと画像の同時推定による高解像度・高フレームレート動画像生成”, 画像電子学会論文誌, Vol.42 No.4, pp.527-535, Jul. 2013.
- [42] 登一生, 今川太郎, 本村秀人, 吾妻健夫, “時空間解像度の異なる 3 板撮像画像からの高時空間解像度カラー画像生成:動きと生成画像の同時推定”, 電子情報通信学会技術研究報告, IE2007-321, PRMU2007-305, pp.377-384, Mar. 2007.
- [43] 今川太郎, 吾妻健夫, 登一生, 本村秀人, “時空間解像度の異なる 3 板撮像画像からの高時空間解像度カラー画像生成:再構成型超解像の利用”, 電子情報通信学会技術研究報告, IE2007-322, PRMU2007-306, pp.385-392, Mar. 2007.
- [44] 本村秀人, 吾妻健夫, 登一生, 今川太郎, 上村健二, 津村徳道, 中口俊哉, 三宅洋一, “時空間解像度の異なる 3 板撮像画像からの高時空間解像度カラー画像生成:事例学習型超解像の利用”, 電子情報通信学会技術研究報告, IE2007-323, PRMU2007-307, pp.383-400, Mar. 2007.
- [45] 今川太郎, 吾妻健夫, 登一生, 本村秀人, “時空間解像度の異なる 3 板撮像画像からの高時空間解像度カラー画像生成”, 画像の認識・理解シンポジウム(MIRU2008)講演論文集, pp.221-228, Aug.2008.
- [46] 登一生, 吾妻健夫, 本村秀人, 今川太郎, “画像処理装置, 画像処理方法, コンピュータプログラムおよび当該コンピュータプログラムを記録した記録媒体, フレーム間動き算出方法および画像処理方法”, 特許公報, 特許第 4469018 号, 2010.03.05, *United States Patent*, US7903156B2, *Chinese Patents*, CN101743755B.
- [47] R. Hicks, K. Schofield, P. Tarnow and M. Veiseh, “Panoramic Electronic Rear Vision for Automotive Applications”, *International Congress and Exposition, SAE Technical Paper Series*, 1999-01-0655, Mar. 1999.
- [48] 横矢直和, “全方位画像センサによる情景の映像アーカイビング”, 3次元画像コンファレンス'98 講演論文集, pp.64-69, July 1998.
- [49] M. Wada, K. S. Yoon, H. Hashimoto, N. Shimizu, S. Matsuda, M. Mizutani, H. Masuda, “Development of Advanced Parking Assistance System in the

- Human Interactive Support System Framework”, *Proceedings of the 6th World Congress of ITS*, Nov. 1999.
- [50] P. J. Narayanan, P. W. Rander, T. Kanade “Constructing a Virtual World Using Dense Stereo”, *Proceedings on the 6th IEEE Conference on Computer Vision*, pp.3-10, Jan. 1998.
- [51] K. Fintzel, R. Bendahan, C. Vestri, S. Bougnoux and T. Kakinami, "3D parking assistant system", *in 11th World Congress on Intelligent Transport Systems*, Nagoya, Japan, 2004.
- [52] Y.C. Liu, K.Y. Lin and Y.S. Chen, "Bird’s-eye view vision system for vehicle surrounding monitoring”, *Proceedings of RobVis 2008*, pp.207-218, Auckland, New Zealand, Feb. 2008.
- [53] C.C. Lin and M.S. Wang, "A vision based top-view transformation model for a vehicle parking assistant”, *SENSORS*, Vol.12, no.4, pp.4431-4446, Apr. 2012.
- [54] R. Hicks, K. Schofield, P. Tarnow and M. Veiseh, "Panoramic electronic rear vision for automotive applications”, *International Congress and Exposition, SAE Technical Paper Series*, 1999-01-0655, Mar. 1999.
- [55] 市原栄太郎, 高尾広行, 大田友一, "NaviView:仮想車載カメラ画像による運転者の視覚支援", *電子情報通信学会論文誌 D*, Vol.J82-D-II, No.10, pp.1816-1825, Oct. 1999.
- [56] T. Ehlgen, T. Pajdla and D. Ammon, "Eliminating blind spots for assisted driving", *IEEE Trans. Intelligent Transportation Systems*, Vo.9, No.4, pp.657-665, Dec. 2008.
- [57] 金岡晃廣, 高野照久, 菅原大輔, 大谷荘太郎, 鈴木政康, 知野見聡, 大泉謙, "アラウンドビューモニターの開発", *Nissan technical review*, no.63, pp.37-41, Sep. 2008.
- [58] S. Shimizu, J. Kawai and H. Yamada, "Wraparound view system for motor vehicle", *Fujitsu Scientific & Technical Journal*, Vol.46, No.1, pp.95-102, Jan. 2010.

- [59] J. Pan, V. Appia and A.C. Bovik, "Virtual top-view camera calibration for accurate object representation", *2016 IEEE Southwest Symposium on Image Analysis and Interpretation*, pp.21-24, 2016.
- [60] S.M. Santhanam, V. Balisavira, S.H. Roh and V.K. Pandey, "Lens distortion correction and geometrical alignment for around view monitoring system", *The 18th IEEE International Symposium on Consumer Electronics*, pp.1-2, 2014.
- [61] M. Yu and G. Ma, "A visual parking guidance for surround view monitoring system", *2015 IEEE Intelligent Vehicles Symposium*, pp.53-58, 2015.
- [62] J. Xu, G. Chen and M. Xie, "Vision-guided automatic parking for smart car", *Proceedings of the IEEE Intelligent Vehicles Symposium 2000*, pp.725-730, Dearbon, MI, USA, Oct. 2000.
- [63] 小越咲子, 小越康宏, 木村春彦, 広瀬貞樹, "CCD カメラ画像に基づいた自動車の車庫入れの自動化", *電子情報通信学会論文誌 A*, Vol.J87-A(2), pp.253-264, Feb. 2004.
- [64] T. Soneoka, "Digital cinema today and the future business development", *NTT Technical Review*, vol. 4, No. 7, pp. 12-17, 2006.
- [65] T. Yamashita, K. Mitani, M. Sugawara, H. Shimamoto, and F. Okano, "Ultrahigh-definition 4-CCD Video Camera System with 4,000 Scanning Lines", *Journal of the Institute of Image Information and Television Engineers*, Vol. 58, No. 3, pp. 383-391, 2004.
- [66] J. Arai, M. Okui, T. Yamashita, and F. Okuno, "Integral three-dimensional television using a 2000-scanning video system", *Applied optics*, vol. 45, no. 8, pp. 1704-1712, 2006.
- [67] R. Y. Tsai and T. S. Huang, "Multiframe image restoration and registration", *Advances in computer vision and image processing*, 1(2), pp. 317-339, 1984.
- [68] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview", *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21-36, 2003.

- [69] M. Elad and A. Feuer, “Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images”, *IEEE Transactions on Image processing*, 6(12), pp. 1646-1658, 1997.
- [70] 矢口陽一, 田中正行, 奥富正敏, “超解像処理のための複数モーションに対応したロバストかつ高精度な位置合わせ手法”, *電子情報通信学会論文誌 D*, Vol.J92D, pp.2033-2043, Nov. 2009
- [71] Z. Lin and H. Y. Shum, “Fundamental Limits of Reconstruction Based Super resolution Algorithms under Local Translation”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26, 1, pp.83-97, 2004.
- [72] M. Tanaka and M. Okutomi, “Theoretical analysis on reconstruction-based super-resolution for an arbitrary psf”, *IEEE Computer society conference on Computer Vision and Pattern Recognition*, vol. 2, pp.947-954 , 2005.
- [73] W. T. Freeman, E. C. Pasztor and O. T. Carmichael, “Learning Low-Level Vision”, *International Journal of Computer Vision*, 40, 1, pp.25-47, 2000.
- [74] S.Baker and T. Kanade, “Hallucinating Faces”, *Proceedings on IEEE International Conference on Automatic Face and Gesture Recognition*, pp83-88, 2000.
- [75] D. Glasner, S. Bagon, and M. Irani, “Super-Resolution from a Single Image”, *Proceedings of the 12th IEEE International Conference on Computer Vision*, pp349-356, 2009.
- [76] 松延徹, 長原一, 岩井儀雄, 谷内田正彦, 鈴木俊哉, “モーフィングによる高解像度高フレームレート動画の生成”, *電子情報通信学会論文誌 D*, Vol.90-D, No.4, pp.1073-1084, 2007.
- [77] 渡邊清高, 岩井儀雄, 長原一, 谷内田正彦, 鈴木俊哉, “ウェーブレット領域での動き補償と画像統合による高解像度高フレームレート動画の生成”, *情報処理学会論文誌*, Vol. 47, No. SIG 10. pp. 195-207, July 2006.
- [78] 重本倫宏, 星川章, 長原一, 岩井儀雄, 谷内田正彦, 鈴木俊哉, ” 時間的・空間的分解能の異なる複合センサカメラシステム”, *情報処理学会論文誌*, 47, SIG_5(CVIM_13), pp.35-44, Mar. 2006.

- [79] 今川太郎, 吾妻健夫, 佐藤智和, 横矢直和:”時空間解像度の異なる動画像を用いた高解像度高フレームレート撮影と露光量確保の両立”, *映像情報メディア学会誌*, 63(8), pp.1139-1148, 2009.
- [80] E. Shechtman, Y. Caspi, and M. Irani : "Space-time super resolution", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.27, no.4, pp.531-545, 2005.
- [81] 今川太郎, 吾妻健夫, 登一生, 本村秀人 : “高解像度高フレームレート化と露光量確保のための時空間解像度が異なるカラー画像からの動画像生成”, *電子情報通信学会論文誌 D*, Vol.J92-D, No.8, pp1229-1239, Aug. 2009.
- [82] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, “Joint MAP registration and high-resolution image estimation using a sequence of undersampled images”, *IEEE Transactions on Image Processing* Vol. 6, no. 12, pp. 1621-1633, 1997.
- [83] 後藤知将, 奥富正敏, “画像復元とレジストレーションの同時最適化の実験的検証” *電子情報通信学会論文誌 D* Vol.J-90-D, No.6, pp.1632-1635, Jun. 2007.
- [84] J. F. Hamilton Jr. and J. E. Adams, “Adaptive color plane interpolation in single sensor color electronic camera”, *U.S. Patent 5 629 734*, 1997.
- [85] B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, ”Color plane interpolation using alternating projections”, *IEEE Transaction on Image Processing*, pp.997-1013, 2002.
- [86] W. Lu and Y. Tan, “Color Filter Array Demosaicking, New Method and Performance Measures”, *IEEE Transactions on Image Processing*, Vol.12, No.10, pp.1194-1210, 2003.
- [87] 映像情報メディア学会監修:像標準テレビジョン方式・システム評価用標準動画像” , http://www.ite.or.jp/shuppan/testchart_index.html
- [88] Donoho, D. L., “Compressed Sensing”, *IEEE Transactions on Information Theory*, V. 52(4), 1289–1306, 2006.

ACM : Association for Computing Machinery.

IEEE : Institute of Electrical and Electronic Engineers.

IP SJ : Information Processing Society of Japan.

SAE : Society of Automotive Engineers.

研究業績

論文誌

1. 登一生, 浮田宗伯, 萩田紀博, “複合射影カメラモデルと半円柱形状モデルを用いた車載カメラシステムのための低ひずみ全周囲画像生成手法,” 電子情報通信学会論文誌D, Vol. J101-D, No.5, pp.780-790, May 2018, (related to chapter 3).
2. 登一生, 今川太郎, 本村秀人, 吾妻健夫, “色ごとに時空間解像度が異なる動画像に対する動きと画像の同時推定による高解像度・高フレームレート動画像生成”, 画像電子学会論文誌, Vol.42 No.4, pp.527-535, Jul. 2013, (related to chapter 4).

国際学会

1. Kunio Nobori, Norimichi Ukita and Norihiro Hagita, "A surround view image generation method with low distortion for vehicle camera systems using a composite projection," *Proceedings of the 15th IAPR Int. Conf. on Machine Vision Applications*, pp.386-389, May 2017, (related to chapter 3).
2. Shusaku Okamoto, Masamichi Nakagawa, Kunio Nobori, Atsuhiko Morimura, Takeshi Okada, Satoru Masuda and Yukio Fujita. "Development of Parking Assistance System Using Virtual Viewpoint Image Synthesis", *Proceedings of the 7th World Congress on Intelligent Transport Systems*, 3075S, Nov. 2000, (related to chapter 2).

国内学会 (査読無し)

1. 登一生, 今川太郎, 本村秀人, 吾妻健夫, “時空間解像度の異なる3板撮像画像からの高時空間解像度カラー画像生成: 動きと生成画像の同時推定”, 電子情報通信学会技術研究報告, IE2007-321, PRMU2007-305, pp.377-384, Mar. 2007, (related to chapter 4).
2. 今川太郎, 吾妻健夫, 登一生, 本村秀人, “時空間解像度の異なる3板撮像画像からの高時空間解像度カラー画像生成: 再構成型超解像の利用”,

電子情報通信学会技術研究報告, IE2007-322, PRMU2007-306, pp.385-392, Mar. 2007, (related to chapter 4).

3. 本村秀人, 吾妻健夫, 登一生, 今川太郎, 上村健二, 津村徳道, 中口俊哉, 三宅洋一, “時空間解像度の異なる3板撮像画像からの高時空間解像度カラー画像生成:事例学習型超解像の利用”, *電子情報通信学会技術研究報告*, IE2007-323, PRMU2007-307, pp.383-400, Mar. 2007, (related to chapter 4).
4. 今川太郎, 吾妻健夫, 登一生, 本村秀人, “時空間解像度の異なる3板撮像画像からの高時空間解像度カラー画像生成”, *画像の認識・理解シンポジウム (MIRU2008) 講演論文集*, pp.221-228, Aug.2008, (related to chapter 4).

登録特許

1. 登一生, 中川雅通, 石井浩史, 岡本修作, “画像処理装置および監視システム”, *特許公報*, 特許第 3300337 号, 2002.04.19, *United States Patent* US6734869 B2, *United States Patent* US7714887 B2, *European Patents* EP1150252B1, (related to chapter 2).
2. 岡本修作, 中川雅通, 登一生, 森村淳, “画像生成装置, 画像生成方法”, *特許公報*, 特許第 3286306 号, 2002.03.08, *United States Patent*, US7307655 B1, *European Patents*, EP1115250B1, (related to chapter 2).
3. 岡本修作, 中川雅通, 登一生, 森村淳, “画像処理装置および監視システム”, *特許公報*, 特許第 3300334 号, 2002.04.19, *United States Patent* US7161616 B1, *European Patents* EP1179958B1, (related to chapter 2).
4. 登一生, 吾妻健夫, 本村秀人, 今川太郎, “画像処理装置, 画像処理方法, コンピュータプログラムおよび当該コンピュータプログラムを記録した記録媒体, フレーム間動き算出方法および画像処理方法”, *特許公報*, 特許第 4469018 号, 2010.03.05, *United States Patent*, US7903156B2, *Chinese Patents*, CN101743755B, (related to chapter 4).

受賞 Awards

1. 平成二十三年度全国発明表彰「発明賞」: 登一生, 岡本修作, 中川雅通, 森村淳, “車両周辺の仮想視点画像を合成する基本技術の発明”, *社団*

法人発明協会, June. 2011, (related to chapter 2).

2. 平成二十二年度近畿地方発明表彰「特許庁長官奨励賞」：登一生，岡本修作，中川雅通，森村淳，“車載用の仮想視点画像合成技術の基本発明”，*社団法人発明協会*, Nov. 2010, (related to chapter 2).
3. 電気科学技術奨励賞：登一生，岡本修作，中川雅通，“仮想視点画像合成技術の開発と全周囲車載カメラシステムの実用化”，*財団法人電気科学技術奨励会*, Nov. 2011, (related to chapter 2).