ステレオ動画像を用いた車両の前方環境認識

関 晃仁[†]　奥富 正敏

東京工業大学 大学院理工学研究科

概要　本論文では，車載ステレオカメラから得られた動画像から，車両の前方環境の多様な情報を取得する手法について述べる．提案手法では，道路環境を認識するにあたり，自動車にとって道路の情報がベースとなる点に着目する．そこで，最初に車載カメラの映像から空間中の道路領域を推定し，その情報に基づいて，後の各種処理を行う．まず，道路領域の時系列の変化具合を判断することで危険な領域を検出し，その空間位置を特定する．加えて，道路面を基準とすることで道路上の立体物の検出と位置推定を行う手法も提案する．さらに，道路の情報と車両の運動拘束を利用して，自車両の運動を推定する．また，シーン中の各点の空間位置とモーションを同時に推定し，それを利用して障害物の運動を推定する手法も提案する．各章の最後に，各提案手法による実験結果を示し，その有効性を確認する．

Road Scene Understanding Using Sequential Stereo Images

Akihito Seki[†] and Masatoshi Okutomi

Graduate School of Science and Engineering,
Tokyo Institute of Technology

Abstract In this paper, we propose methods of road scene understanding of vehicle front view using sequential vehicle-mounted stereo images. To understand road scene, we notice the issue that road surface information is fundamental because vehicles run on the road. Therefore, road surface is detected using sequential stereo images at beginning. First, we propose dangerous region detection method which is derived from checking edges of road region in real space. Second, obstacle detection method is proposed. Obstacles which are the 3D objects on the road surface are detected and measured using this road information. Third proposed method is ego-motion estimation. This is done by using road information and vehicle motion constraints on road surface. Final proposed method is simultaneous depth and 3D motion estimation. We present experimental results to demonstrate the effectiveness of our methods at each chapter.

1　はじめに

近年，自動車のインテリジェント化を実現するために，さまざまな技術が世界中で研究開発されている．それらの技術によって，ドライバーの運転の負担を軽減することを目的として，車両を道路白線内へキープさせる車線維持支援機能や，ステアリングやブレーキを自動的に行なうことで駐車を支援する機能などの運転支援技術，交通事故の発生を防ぐことを目的として，夜間などの運転者から見えにくい歩行者を検出するナイトビジョン機能 [26] や，わき見や居眠りなどの運転者の状態を検知する機能などの予防安全技術を搭載した車両の市販が始まっている．

これらのシステムは，センシングデバイスとしてカメラを利用している．カメラは，安価で小型なセンサにもかかわらず，大量の情報を取得することができる．加えて，他のセンサでは得られない標識のパターン認識などにも利用可能であることから，このような運転支援や予防安全技術に有効なセンシングデバイスであり，今後も自車両の前方環境認識において，中心的な役割を果たしていくものと考えられる．しかし，撮影された映像は，そのままでは制御などに利用することができない．そのため，例えば，車線の位置等の所望な情報を得るための画像処理が必要である．加えて，自動車に搭載する先進装置の研究により，車載画像処理用プロセッサである Visconti®[32] や IMAPCAR® などの高速な演算装置が開発されたことで，車載コンピュータの性能も向上し，より広範囲な環境で，安定かつ正確に画像から情報を抽出するための高度な画像

[†]現在，(株) 東芝 研究開発センターに所属．

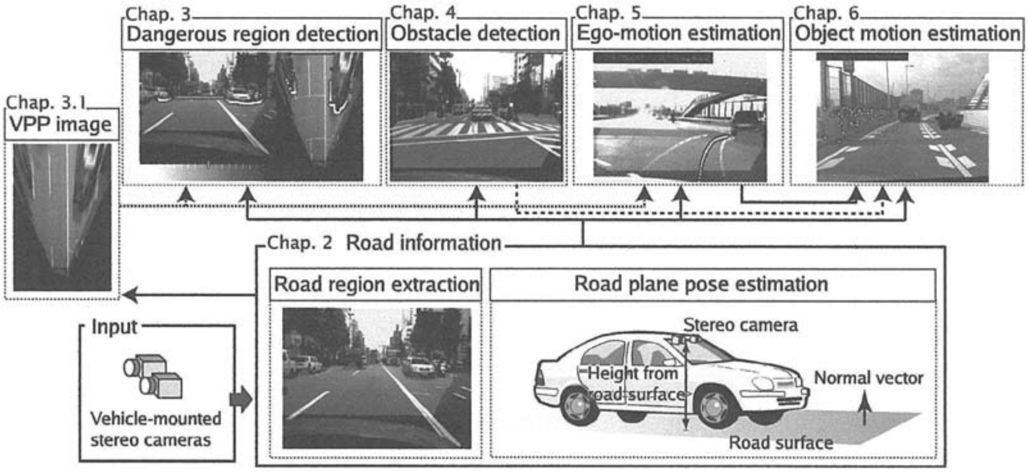

図 1: 本研究の全体図

処理技術の実現が求められている．

運転支援・予防安全技術のうち，前走車との車間距離を一定に保つクルーズコントロール機能や前走車の急停車に対処して，自動的にブレーキをかけるブレーキアシスト機能では，車両前方の障害物を検出し，その 3 次元的な位置とそのモーションを取得することが制御に必要である．市販されているシステムでは，障害物検出やモーション推定にミリ波レーダーを利用したものが大部分を占めている．レーダーセンサは，比較的複雑なシステムではないため，導入はしやすいが，障害物までの距離測定を行える間隔がまばらなため，空間的な解像度が低くなってしまう．また，電波の反射によって物体を検出する仕組みであるため，検出対象の材質によって反射性能が異なる．そのため，例えば，車両などの金属製物体は検出しやすく，人や植木など非金属製物体は検出しにくいといった短所がある．これらの性質のため，利用できる道路が高速道路環境下などに限定されてしまう．しかし，レーダーセンサでは霧のように視界が不良な状況下でもカメラと違い，利用することができる．一方，カメラの映像を利用した画像処理による方法では，検出対象の材質に依存せず，センサの空間解像度が高いことや，画角が広いことから，より広範囲かつ詳細に多様な物体の情報を得られる可能性がある．また，それぞれのセンサを組み合わせて，短所を補い合うセンサフュージョンによって，これまでは高速道路などに限定されていた安全支援技術の想定環境が，市街地のような複雑な環境下へと適用範囲が広がるだけではなく，自動運転のような，より高度な技術への礎になっていく可能性もある．

本論文では図 1 に示すように，カメラのみを利用して，1). 危険領域の検出，2). 障害物の検出，3). 自車両の運動推定，4). 対象のモーション推定を実現する手法をそれぞれ提案する．これらの手法は，道路環境が整備されていない一般道路環境下においても機能することが重要である．また，将来，センサフュージョンを想定した場合にも，カメラ単体で利用した場合に実現できる機能や，適用可能な範囲を見極めておく必要がある．

さて，本研究では，道路環境を認識するために，自動車は道路上を走行することから，道路に関する情報が要であると考える．そのため，道路領域と道路面のカメラに対する姿勢などの道路に関する情報を先に推定し，それを利用して 1)〜4) の機能を実現する．以降では，それぞれの処理について，順に説明していく．

なお，各章の内容と関連の深い研究発表として，危険領域検出は文献 [18]，障害物の検出は文献 [20, 23]，自車両の運動推定は文献 [19, 22]，対象のモーション推定は文献 [21] である．

2 道路情報の取得

車両の前方環境を認識するにあたって，道路に関する情報を先に求めておくことにより，例えば，これから述べるような危険な領域を簡単に検出できるようになったり，障害物検出処理などの高度な環境認識処理の高速化や，頑健性向上につながる．

さて，道路に関する情報である，道路領域やカメラと道路面間の姿勢は，車両の周囲の様子や走行中の車両振動や出現する道路面によって常に変化する．図 2 は，走行中に撮影された画像の一部を取り出して，道路領域を明るく，水平線を点線で書き入れたものである．このときのカメラの車両への設置姿勢はどちらも

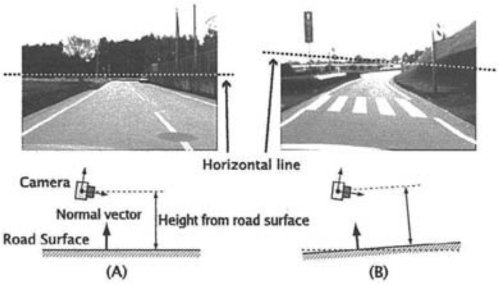

図 2: 走行中の道路領域と道路面とカメラ間の姿勢

同じである．同図より，(B) では (A) に比べて，道路領域が異なり，ピッチ角とロール角が変化していることが分かる．

そこで本章では，道路環境認識において基礎となる道路領域抽出処理および道路面とカメラ間の相対的な姿勢を推定する処理について述べる．

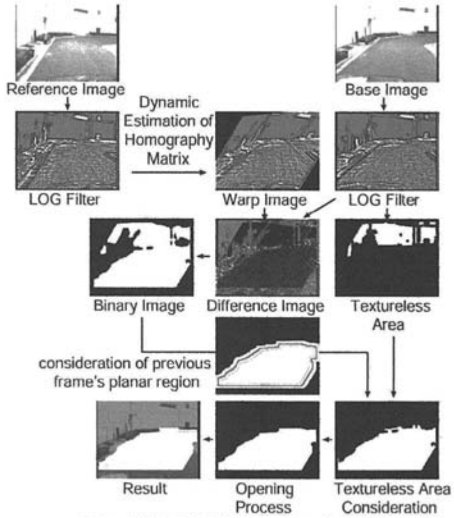

図 3: 道路領域抽出処理の全体図

2.1 道路領域の抽出

まず，道路領域の抽出方法について説明する．

空間中の平面が 2 枚の画像に投影されたとき，画像間の平面内における関係は広く知られているように，射影変換行列 $\mathbf{H}$ を使って以下のように表現できる．

$$\tilde{\mathbf{m}}_r \sim \mathbf{H}\tilde{\mathbf{m}}_b = k\mathbf{A}_r\left(\mathbf{R}+\frac{\mathbf{t}\mathbf{n}^T}{d}\right)\mathbf{A}_b^{-1}\tilde{\mathbf{m}}_b \quad (k \neq 0) \tag{1}$$

ここで，$\tilde{\mathbf{m}}_*$ は同次座標系で表した画像座標，$\mathbf{A}_*$ はカメラの内部パラメータを表し，添字 $*$ が b は基準カメラ，r は参照カメラをそれぞれ表す．$\mathbf{R}$ と $\mathbf{t}$ はそれぞれカメラ間の回転行列と並進ベクトルを表す．$\mathbf{n}$ は平面の法線ベクトル，d はカメラと平面間の距離である．

本論文では，奥富ら [14, 15] の手法を利用して道路領域を求める．この手法では，先に述べた道路平面に対して左右の画像を一致させる射影変換行列 $\mathbf{H}$ を動的に推定し，画像間の一致具合，テクスチャや時系列変化を加味して道路領域を抽出する．図 3 は，その道路領域抽出処理の全体図である．まず基準画像と参照画像を入力する．次いで画像間の輝度を一致させるため，それぞれの画像に対して LoG フィルタを施した後，ヒストグラム平坦化処理によりコントラストを上げる．続いて，道路平面に対する射影変換行列を動的に推定する．道路平面領域を求めるには，まず推定された射影変換行列を用いて参照画像を変形し，変形後の参照画像と基準画像との輝度差を計算する．平面部分は両画像が完全に重なって輝度差が小さくなるため，輝度差に対する閾値を用いて道路領域候補を抽出する．その際，テクスチャの無い領域では平面部分以外も輝度差が小さくなる可能性があるため，別途処理を加える．次に Opening 処理を行い，小領域を除去して最終的な道路領域を得る．

2.2 道路面の姿勢推定

画像を利用して道路面とカメラ間の姿勢を推定するために，いくつかの方法が提案されており，例えば，道路白線の平行性などのモデルを利用した方法が提案されている [16, 5, 13, 12]．しかし，一般道路環境では必ずしも平行な白線があるわけではない．また，ステレオ画像間でマッチングを行ない，ステレオ計測をして道路面を推定する方法も提案されている [2, 10]．しかし，Williamson ら [30] が指摘しているように，道路面のようにフロントパラレルと大きく異なる面に対して，従来の変形を許さないウィンドウベースのマッチングを適用するのでは，左右画像のウィンドウ内での変形が大きく，対応点を正しく求めることが難しい．また，小ウィンドウを使うマッチングのため，道路面に十分なテクスチャがあることや，道路面以外の計測データを除去するために，画像に道路面が十分に写っている必要があると考えられる．

式 (1) の射影変換行列 $\mathbf{H}$ には，道路面の法線ベクトル $\mathbf{n}$ とカメラから道路面までの距離 d が含まれている．そこで，本研究では，道路領域を抽出する際に推定する $\mathbf{H}$ から，道路面の姿勢を推定する．それには，大きくわけて 2 つの方法がある．1 つめは，カメラ間の外部パラメータは未知で，ステレオカメラ間のベースライン長 (並進ベクトルの大きさ) だけが既知な場合である．2 つめは，カメラ間の外部パラメータである，回

転行列と並進ベクトルが既知な場合である [25]. また，どちらの場合もカメラの内部パラメータは既知であることを前提としている.

道路面に対する射影変換行列が動的に推定されるため，走行中に車両が振動して道路面とカメラの位置関係が変わったとしても，その変化に応じた姿勢パラメータを推定することができる．さらに，道路面に対する大きなウィンドウを変形して射影変換行列を計算しているため，テクスチャの少ない路面でも安定に推定することができる.

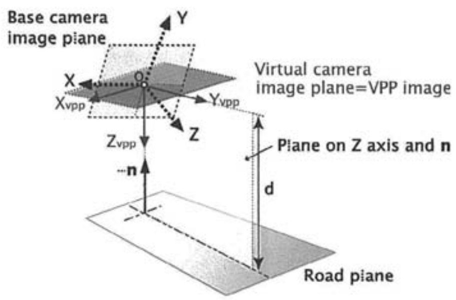

図 4: VPP 画像の幾何関係

3　危険領域検出

本章では，車載ステレオカメラを使った危険領域検出について述べる．危険領域とは，自車両にとって危険となる領域のことであり，例えば，対向車などのように自車両に近付いてくるものは危険性が高くなる．それらを検出することで，視覚誘導や安全走行のための運転支援などに利用できると考えられる.

提案手法では，自車両にとって危険であるのか否かを 2 章で抽出された道路領域と非道路領域の境界の変化具合を時系列的に計測することで求める．例えば境界が近づいている領域は，自車両が障害物などの領域に相対的に近付いたということである．境界の変化を測定する際に，画像中で測定した場合には，空間的に同じ変化をしたとしても，境界が遠くにあった場合には，その変化が小さく観測され，近くでは大きく観測される．また，その境界が変化している空間的な位置や変化の大きさが容易には求まらない．そこで，提案手法では仮想的に道路面を垂直上方から見た画像 (仮想投影面画像*) を生成することで，実空間における道路領域を求め，VPP 画像上での道路領域の境界の変化から，危険領域の検出を行なう．通常，危険領域までの距離を測定するためには，ステレオ画像間の対応を取る必要があるが，VPP 画像を生成したことで，ステレオ対応点探索を行うことなく，危険領域までの距離の測定が可能である．さらに，提案手法ではカメラの光学中心から道路領域の方向毎の境界変化を計測するため，トラッキングなしに領域の危険性を評価することができる.

本章では，まず VPP 画像生成手法について説明し，続いて危険領域検出手法について述べる.

3.1　仮想投影面画像生成

VPP 画像は，原画像を仮想的に道路面と平行に上方から観測した画像に相当している．つまり，図 4 に示す幾何関係にあり，以下に述べるような二つの性質を満たすものとする.

- VPP 画像の座標系は，実空間の座標系に一致する．そのため，VPP 画像上の一点は，3 次元実空間中の一点に対応する.
- VPP 画像の縦軸は，基準カメラの光軸を道路面に投影した方向と一致する.

以上の VPP 画像の性質は，以下のように言い替えることもできる.

1. 仮想カメラの光軸 Z_{vpp} は，平面の法線ベクトル $\mathbf{n}$ に一致する.
2. 仮想カメラの Y_{vpp} 軸は，法線ベクトル $\mathbf{n}$ と基準カメラの光軸である Z 軸で張る平面上に存在する.

この 2 つの性質に基づいて，基準カメラから仮想カメラへの座標変換を考える．その際，2.2 節によって求められる道路面の姿勢パラメータが必要である．まず，この座標変換では，それらの光学中心が変化しないことから，回転行列 $\mathbf{R}_{vpp\to r}$ を使って表現する．$\mathbf{R}_{vpp\to r}$ を一般的な回転行列の表現として，

$$\mathbf{R}_{vpp\to r}=\begin{bmatrix} c\theta c\phi c\psi - s\phi s\psi & c\theta s\phi c\psi + c\phi s\psi & -s\theta c\psi \\ -c\theta c\phi s\psi - s\theta c\psi & -c\theta s\phi s\psi + c\phi c\psi & s\theta s\psi \\ s\theta c\psi & s\theta s\phi & c\theta \end{bmatrix} \tag{2}$$

ここで，$c\phi$ や $s\phi$ はそれぞれ $\cos\phi$ と $\sin\phi$ を示すものとする．上式で，未知数は (θ, ϕ, ψ) であり，上記の性質に基づいて計算すると，

$$\begin{cases} \theta = \arccos(n_z) \\ \psi = \arctan\left(\frac{-n_y}{n_x}\right) \end{cases} \tag{3}$$

が得られる．また，ϕ は，$\pm\pi/2$ と求まり，その符合は，Y_{vpp} 軸と Z 軸は同じ方向であるため，それらの内積が正であるように決めればよい.

以上より，基準画像から VPP 画像への変換は射影変換行列 $\mathbf{H}_{vpp}$ を使って，以下のように表現すること

*この画像を本論文では，仮想投影面画像 (Virtual Projection Plane Image 略して VPP 画像) と呼ぶこととする.

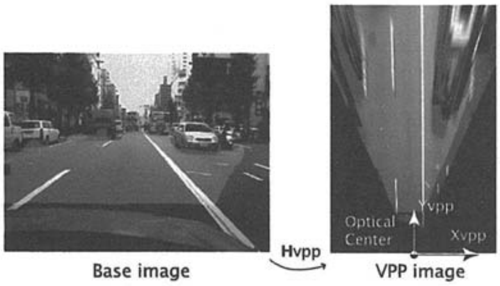

図 5: 道路領域抽出結果および VPP 画像生成例. (色の明るい領域が抽出された道路領域)

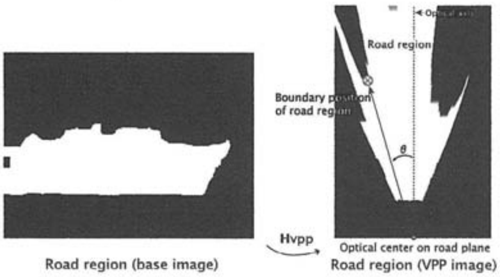

図 6: 方向別距離の測定

ができる.

$$\tilde{\mathbf{m}}_{vpp} \sim \mathbf{H}_{vpp}\tilde{\mathbf{m}}_r = \mathbf{A}_{vpp}\mathbf{R}^{-1}_{vpp\to r}\mathbf{A}_r^{-1}\tilde{\mathbf{m}}_r, \quad (4)$$

ここで, $\tilde{\mathbf{m}}_{vpp}$ は, VPP 画像の同次座標であり, $\tilde{\mathbf{m}}_r$ は, 基準画像の同次座標である. また, $\mathbf{A}_{vpp}$ は仮想カメラの内部パラメータである.

図 5 は, 基準画像とその基準画像に対する VPP 画像を生成した結果である. 画像中の色の明るい領域は, 抽出された道路領域である. 右側の VPP 画像では, 道路白線が平行になっており, 正しく VPP 画像が生成されていることが分かる.

3.2 危険領域検出処理

危険領域検出は, VPP 画像上における, カメラの光学中心を道路面に投影した位置から非道路領域までの距離の変化に基づいて行う.

まず, 道路領域と非道路領域の端点までの距離を方向ごとに求めるため, 図 6 のように道路領域を VPP 画像に変換した画像を用意する.

次に, 基準カメラの光学中心を VPP 画像に投影した点を VPP 画像の光学中心とするとき, この光学中心から VPP 画像縦軸 (基準カメラの光軸を道路面へ投影した方向) に対して, θ だけ傾いた直線を伸ばし, 道路領域の端を計算する. この端が方向別の道路面の拡がりの端点である. VPP 画像上での長さは, 仮想カメラの内部パラメータと基準カメラの光学中心から道路面までの距離を使うことで実際の距離に変換できる. これによって, 光学中心から方向別の道路と非道路の境界までの距離である, 方向別距離が算出できる.

次に道路領域の方向別相対速度を求める. 画像は 1 フレーム 1/30[sec] の等間隔で撮影される. そこで, 1 フレームごとの方向別距離を時系列にわたってフィルタリングすることで, その速度を算出する. この速度は, 基準カメラの光軸を道路面へ投影した方向からの各方向ごとの道路領域の拡大収縮速度であり, 方向別相対速度と呼ぶことにする.

方向別相対速度がマイナスの場合には, 道路領域が削減されていることに相当するため, その方向には対向車などの危険性の高い障害物が存在している可能性がある. 何も変化がない領域では, 方向別相対速度が 0 である.

3.3 危険領域検出結果

図 7 は, 危険領域の検出結果を表している. 同図の左側は, 推定された各種情報を基準画像中に表示している. A は方向別の道路領域の境界を表しており, その色は方向別速度に応じて, C に示す色に従う. 速度が 0 の場合には, 灰色であり, 危険性が高い (近付く方向) 場合には, 白くなり, 危険性の低い (遠ざかる方向) 場合には, 黒く示される. 同図上段と下段の実験結果では, 共に対向車によって道路領域が減少する領域で危険性が高く表示されている. B は, 一定角度ごとの方向別距離をメートル単位で表示している. また, 車載カメラの解像度の関係で, VPP 画像は前方 30[m] まで生成したため, 表示される範囲もその範囲に限られる. 次に同図右側は, VPP 画像に推定された情報を表示している. D は, 基準カメラの光学中心を道路面に投影した位置 E からの距離を 5[m] ごとに描いたものである. 同図上段の結果では, 対向車が 16,7[m] 先に位置していることが分かる.

4 障害物検出

3 章で述べた方法は, 障害物を陽に検出するのではなく, 自車両に近付いてくる領域を検出することで危険領域検出を行なった. 簡易な方法である一方, その検出方法の性質上, 障害物が道路面に接していないと, 障害物までの距離が正しく計測されない可能性があった. そこで, 本章の提案手法では, 障害物を陽に検出することで, この問題を解決する.

さて, ステレオカメラを利用して, 障害物を検出する従来の手法では, ステレオ画像間の対応点探索を行うこ

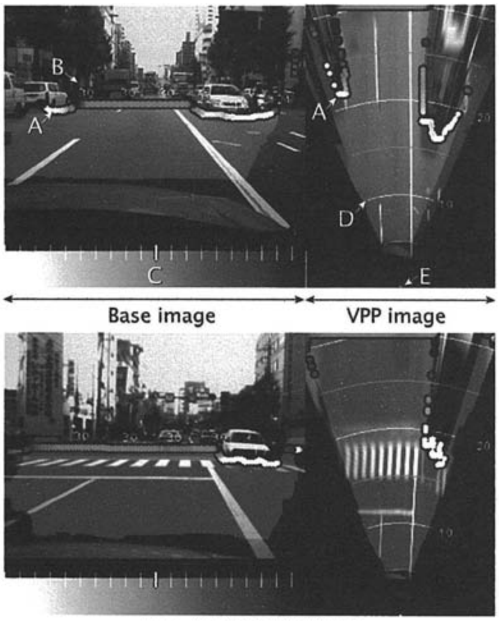

図 7: 危険領域検出結果

とで3次元再構成をし，障害物を求める手法 [10, 31, 17, 11, 8, 9, 28, 29] が提案されている．例えば Labayrade ら [10] は，ステレオ計測された点群の3次元情報から道路面を推定し，車両の後面などをカメラに対して平行な平面と想定することで，障害物を検出する手法を提案している．さらに Hu ら [8] は，Labayrade らの枠組を拡張して道路構造物なども検出している．しかし，どちらも障害物を平面としてモデル化しているため，曲線物体 (例えば，カーブしているガードレール) を正確に検出することができない．また，窪田ら [9] は道路に垂直な方向に短冊状のウィンドウを設けて，ステレオ画像間の対応点探索を行ない，Dynamic Programing を使って物体を検出するため，曲線物体もそのままの形で検出可能である．しかし，短冊状のウィンドウでのマッチングは，建物などの垂直に長い物体には有効である [3] が，車高の低い自動車や踏切のバーなどのようにウィンドウ内における対象物体の占有率が低いと，誤った位置で対応づけされてしまい，正しく検出されない可能性がある．

さて，本論文では環境中に存在する数々の物体のうち，図8に示すように自車両にとって障害となる可能性のある立体物を障害物とする．例えば，標識や立体横断歩道のように道路上方の物体は自車両より高い位置に存在しており，自車両と衝突しない．また，路面上に描かれている道路標識も同様に自車両にとって安全である．道路面は自車両が走行する2次元平面に相当するため，道路面からの物体高さを元にすれば，障

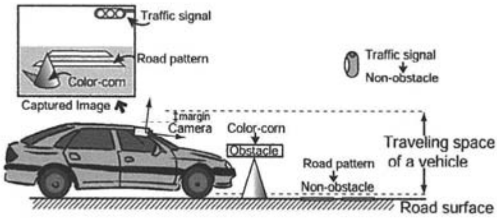

図 8: 障害物: 車両の通過する可能性のある空間に存在する物体

害物か否かを判定することができる．そのため，カメラと道路面の相対的な姿勢を動的に取得できていることが重要である．

4.1 障害物検出手法の概要

図9は，提案手法の流れを表している．まず，道路領域抽出処理と道路面の姿勢推定によって，道路領域と道路面の姿勢を推定する．次に，障害物の空間位置を求める (図9 ③)．障害物は画像中の道路領域以外に存在するため，道路領域以外のエッジ点に関して，ステレオ計測を行なう．抽出された道路領域を活用することでエッジ点の数を減らしたり，対応点探索時の視差探索範囲を限定したりすることで計算コストを抑える．

さて，ステレオ計測をして得られる点の3次元位置は，カメラを基準とする座標系である．カメラと路面の関係が分からないと，その点が障害物に属するものなのか，それ以外に属するものなのかを区別することが困難である．また，先にも述べたように走行中，道路面とカメラの位置関係は変化するため，その関係を固定することはできない．そこで，先ほど求めた姿勢パラメータを用いることで，ステレオ計測結果を道路面を基準とした座標系 (道路座標系) に変換する．道路座標系では道路面に属する点の高さは0となるため，容易に路面パターンであると判断することができる．また，自車両が通過する可能性のある，道路面から一定の高さを持つ空間も容易に特定することができる．

次に，その道路座標系に変換された点の位置を利用して障害物を抽出する (図9 ④)．道路環境中に存在する自動車，歩行者やガードレールなどの障害物は，道路面に垂直であったり，水平方向に幅を持っている．そこで，自車両の通過する空間に含まれる点をそれぞれ1票として，道路面に対して高さ方向に垂直に投影し，道路面を投票グリッドとして投票処理を行なう．さらに投票グリッドを用いて，セグメンテーション処理とノイズ除去をし，障害物を検出する．結果として検出される障害物は，自車両が通過する可能性のある空間に含まれる物体に限られる．

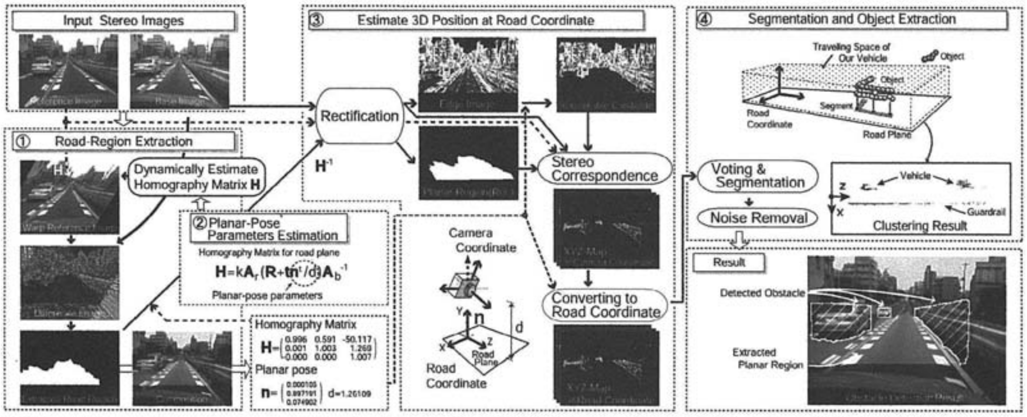

図 9: 障害物検出手法の全体図

4.2 ステレオ計測と障害物の検出

本節では，抽出された道路領域と射影変換行列から推定された道路面の姿勢パラメータを利用して，障害物を検出する方法を図 9 中の③の流れに従って述べる．

4.2.1 道路領域と姿勢を利用した対応点探索の限定

物体の空間位置を計測するには，基準画像の注目点に対して，実空間の奥行き方向に設定された範囲内で参照画像のエピポーラ線上を探索し，注目点と最も類似性の高い位置を対応位置として求める．そこで対応点探索をする注目点の候補 (以下，注目点候補) の決定方法についてまず述べる．自動車のボディのようにテクスチャレスな部分は，対応点を直接求めることができないため，注目点候補はエッジ部分のみにする．さらに前処理によって求められた道路領域には障害物が存在しないことから，道路領域に含まれるエッジ点を注目点候補から除く．また，道路面の姿勢パラメータを使うことで，空間中の検出範囲が画像に射影される範囲を求められるため，範囲外は注目点候補から除く．最終的に残った注目点候補を図 9 の③中の'Remarkable Candidate' に示す．

次に，ステレオ画像間の対応を求める処理について説明する．本研究では，図 10 に示すように，道路面上の点は射影変換行列で対応がついているため，その性質を考慮した対応点探索を行なう．同図で，基準カメラの光学中心 O から AB 方向への光線を考える．A は奥行 $Z_{\min}$ の位置，B は $Z_{\max}$ の位置とする．A が参照画像に投影される点は m_{l2} であり，道路領域に含まれるため，基準画像の m_{r2} と対応がついている．すると，参照画像のエピポーラ線上で道路領域の端点である m_{l1}，空間中では C より遠方が対応点の探索候補である．そのため，通常の探索範囲が m_{lb} から m_{l2} であるとき，道路領域とその射影変換行列を利用することで m_{lb} から m_{l1} に限定することができる．また，参照画像の道路領域は基準画像の道路領域を射影変換行列 $\mathbf{H}^{-1}$ で変換して求める[†]．非道路領域のみ探索を行なうことで，計算コストを抑えることができる．さらに，Cross checking[4] を利用することで，マッチングミスを抑えている．

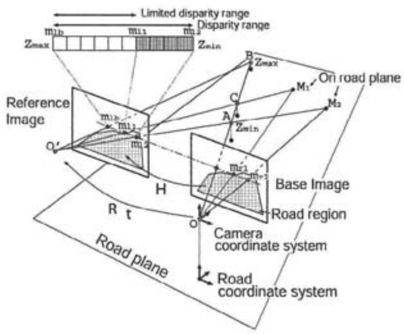

図 10: 対応点探索と道路領域の関係

4.2.2 カメラ座標系から道路座標系への変換

ステレオ画像間の対応を求めることで，三角測量の原理からカメラ座標系における 3 次元位置を求めることができる [6]．カメラ座標系では，路面パターンや道路上方の立体歩道などの非障害物を区別することが困難であるため，道路面を基準とした座標系である道路座標系に変換する．それは，道路面に対する射影変換行列 $\mathbf{H}$ によって推定された道路面とカメラの距離 d と基準カメラに対する道路面の法線ベクトル $\mathbf{n}$ を利用することでカメラ座標系から道路座標系へ座標変換する．この座標変換は，カメラの光軸を道路面に

[†]抽出された道路領域は基準画像に対するものであるため．

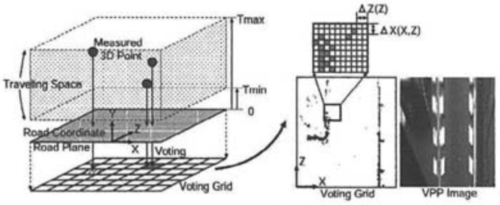

図 11: 車両の通過空間と投票処理

投影したときの方向が座標変換後も変わらず，さらに道路面上は $Y=0$ となるように決める．カメラ座標系上の点 $\mathbf{X}_c=(X_c,Y_c,Z_c)^T$ から道路座標系上の点 $\mathbf{X}_r=(X_r,Y_r,Z_r)^T$ に変換するには，以下の式を利用する．

$$\mathbf{X}_r=\mathbf{R}_{rc}^{-1}(\mathbf{n})\mathbf{X}_c+\mathbf{t}_{rc}(d) \tag{5}$$

回転行列 $\mathbf{R}_{rc}(\mathbf{n})$ と並進ベクトル $\mathbf{t}_{rc}(d)$ は，道路面の姿勢より算出することができる [23]．この座標変換によって，カメラと道路面の位置関係が変化しても，カメラ座標系の 3 次元位置を道路面を基準とした道路座標系に変換することができる．

4.2.3 道路面への投票を利用した障害物の抽出

道路座標系に変換した 3 次元計測結果を利用して障害物を抽出する．障害物は自車両が通過する可能性のある空間に存在する物体であるため，その空間に相当する，道路面からの高さが $T_{\max}$ から $T_{\min}$ [‡] の範囲に存在する 3 次元計測結果をそれぞれ 1 票として道路面に対して垂直に射影し，投票グリッドに投票する．同じ奥行きと幅に計測結果が多く存在するほど，そのグリッドに多くの投票数を得る．その様子を図 11 に示す．同図では，右側のガードレールと左車線の前走車や対応点探索ミスによるノイズに投票値があることがわかる．道路環境に存在する構造物は高さ方向に大きさを持つことが多いため，高さ方向に射影することで点群が少なくても，後述するクラスタリング処理によって障害物を検出可能である．

次に，投票後の投票グリッドを利用してクラスタリングを行なう．クラスタリングは，空間中の幅と奥行きを特徴量とした最短距離法を用いる．ここで，ステレオ計測では遠方になるにつれて計測誤差が大きくなること，注目点候補はピクセル単位で設定するため，遠方の点の密度は近くの点の密度に比べて疎になることを考えると，クラスタの結合範囲を距離にかかわらず一定とすると遠方の物体が検出されなかったり，近接する物体が不必要に結合したりしてしまう．そこで，結合範囲を平行ステレオの誤差に従って，

$$\Delta Z(Z) = \frac{2\delta_u bfZ^2}{b^2f^2-\delta_u^2Z^2}$$

$$\Delta X(X,Z) = \begin{cases} -\frac{\delta_u Z(2X-b)}{bf} & (X<0) \\ \frac{\delta_u Z}{2f} & (0\le X\le b) \\ \frac{\delta_u Z(2X-b)}{bf} & (X>b) \end{cases} \tag{6}$$

とする．ただし，δ_u は有効画素間隔，f は焦点距離，b はベースライン距離を表す．さらに生成されたクラスタからノイズを除去し，障害物を抽出する．障害物は，実空間中の位置とステレオ画像中での位置を求めることができるため，自動車の制御や他の画像処理に利用し易いと考えられる．

4.3 障害物検出結果

提案手法を実画像に適用して，障害物を検出した結果を示す．障害物の検出範囲は，カメラの光軸を道路面に投影した直線を基準として，左右方向に 10[m] 以内かつ，奥行き方向に 50[m] 以内に存在する立体物とした．

図 12 は，一般道路環境下で撮影された画像に対して，提案手法を適用した結果である．色の明るい領域は，抽出された道路領域，斜線の領域は検出された障害物を表している．同図 (A) は，市街地を走行するシーンであり，前走車や側壁，ガードレールなどがそのままの形状に近い形で検出されている．同図 (B) は，カーブのため道路が傾斜しているシーンである．提案手法では，道路面とカメラ間の姿勢を動的に推定しているため，路面パターンを誤検出することなく，前走車や対向車，道路脇の草が正しく検出できている．同図 (C) は，雪道を走行するシーンである．道路上の看板などが検出されることなく，雪壁や前走車，曲がってくる対向車が正しく検出できている．

5 自車両の運動推定

本章では，車載ステレオカメラを利用した自車両の運動推定手法について述べる．自車両の運動推定を行うことにより，その推定結果から，車両の走行状況をモニタリングしたり，ナビゲーションや安全走行のための運転支援などに利用できる．通常，車両の運動を推定するには，車速センサー，舵角センサーやヨーレートセンサーからの情報を使う場合がほとんどである．しかし，車輪の回転を利用した車速センサーなどからの情報では，路面の動きを直接計っていないため，車輪に滑りが発生した場合などに正確な速度を求めることができない．一方，カメラを利用することで直接路面

[‡]本論文では，実験車両の大きさを考えて $T_{\max}=2.0$[m]，$T_{\min}=0.2$[m] とした．

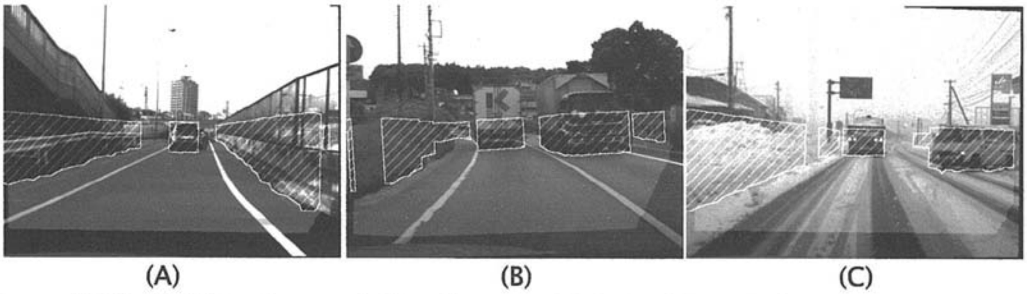

図 12: 障害物の検出結果．色の明るい領域は，抽出された道路領域，斜線の領域は検出された障害物を表している．

の動きを計測することもできる．さて，自車両の運動推定は，その車両に搭載されたカメラのエゴモーション推定を行なうことと同等であるため，以降ではカメラのエゴモーション推定を自車両の運動推定と同じものとして扱う．

時系列画像から，カメラの 3 次元モーションを推定するには，シーンに含まれるモーションからカメラのモーションを取り出す必要がある．さらに，シーンのモーションは，その奥行きと相互に結びついており，どちらかが既知の場合には，残りの未知パラメータを比較的容易に推定することができる [27]．そこで，本章で述べる手法では，道路面を利用することで，その奥行きを既知と考える．加えて，市街地のように，複数の移動物体が含まれるようなシーンにおいて，道路領域は空間的に静止しているため，カメラのモーションのみが含まれる領域として利用するには最適であると考えられる．本章の提案手法では，先に求められた道路領域と道路面の姿勢を有効に利用し，自車両の運動を推定する．

本章では，まず提案手法による自車両の運動推定手法の概要を述べ，続いてそれぞれの手法について詳しく説明し，章の最後で実験結果を示す．

5.1 自車両の運動推定手法の概要

自車両の運動推定手法の処理の流れを簡単に説明する．まず，道路領域を抽出し，さらに道路面の姿勢を推定する．次いで，その姿勢を使って，VPP 画像を生成する．

さらに，隣接したフレームの VPP 画像を使い，道路領域内に対してマッチングをすることで，自車両の運動を推定する．その際，まず車両の運動モデルを用いて計算量を軽減しながら安定に運動パラメータを求めた後，その運動パラメータを勾配法により更新することで，より高精度に求める．以降では，各方法について説明する．

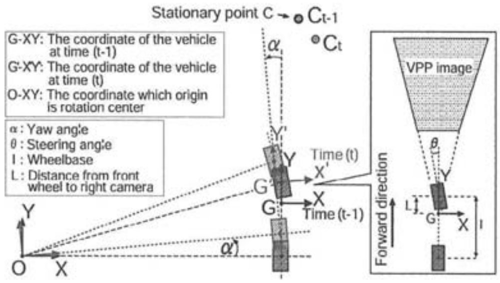

図 13: モデル化した車両の運動

5.2 車両の運動モデルを利用した運動パラメータの推定

まず，アッカーマンステアリングジオメトリ [1] を想定し，自動車を 2 車輪モデルに当てはめる．4 車輪モデルを 2 車輪モデルに当てはめるために，カメラの設置位置を通る車両の中心軸を考える．その様子を図 13 に示す．同図中でカメラは車両上の G に設置されている．ここで，l はホイールベース，L は前輪からカメラの設置位置までの距離，θ は操舵角，α は車両の回転角であるヨー角を示している．

$\Delta x = 0$ すなわち直進している場合 $(\alpha = \theta = 0)$ はモデルに当てはめる必要がないため，車両の並進ベクトルは，

$$(t_x, t_y) = (\Delta x = 0, \Delta y) \tag{7}$$

である．以下では，$\Delta x \neq 0$ の場合について述べる．

時刻 $t-1$ の車両上の点 G を原点とした $G-XY$ 座標系において，静止点 C を $\mathbf{C}_{t-1,G} = (x_G, y_G)$ とする．車両の回転中心 O は $G-XY$ 座標系において，

$$\mathbf{O}_G = \left(-\frac{l}{\tan\theta}\ ,\ -l+L\right) \tag{8}$$

と表される．そこで，静止点 C を $O-XY$ 座標系に変換すると，

$$\mathbf{C}_{t-1,O} = \left(\frac{l}{\tan\theta} + x_G\ ,\ l - L + y_G\right) \qquad (9)$$

となる．

次に時刻 t では，車両が点 O を中心として α だけ回転したとする．静止点 C は $O-XY$ 座標系上では移動しないが，各時刻の車両上の座標系から観測すると，その位置が変化する．そこで，時刻 t の車両上の $G'-X'Y'$ 座標系で観測される静止点 C の位置を $O-XY$ 座標系に変換する．変換された点を図 13 で点 C_t とし，その位置は以下のように表される．

$$\begin{aligned}\mathbf{C}_{t,O} &= \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \frac{l}{\tan\theta} + x_G \\ l - L + y_G \end{bmatrix} \\ &= \begin{bmatrix} \left(\frac{l}{\tan\theta} + x_G\right)\cos\alpha + (l - L + y_G)\sin\alpha \\ -\left(\frac{l}{\tan\theta} + x_G\right)\sin\alpha + (l - L + y_G)\cos\alpha \end{bmatrix}.\end{aligned} \qquad (10)$$

車両の運動によって引き起こされる静止点 C の移動ベクトル $(\Delta x, \Delta y)$ は $\mathbf{C}_{t,O} - \mathbf{C}_{t-1,O}$ である．ここで，α が十分に小さいものと仮定すると，$\cos\alpha \fallingdotseq 1$，$\sin\alpha \fallingdotseq \alpha$ が成り立つので，

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \mathbf{C}_{t,O} - \mathbf{C}_{t-1,O} \fallingdotseq \begin{bmatrix} (l - L + y_G)\alpha \\ -\left(\frac{l}{\tan\theta} + x_G\right)\alpha \end{bmatrix} \qquad (11)$$

である．これにより，注目点 (x, y) と移動ベクトル $(\Delta x, \Delta y)$ から操舵角 θ とヨー角 α が以下のように求められる．

$$\alpha = \frac{\Delta x}{l - L + y_G},\ \theta = \arctan\left(\frac{l}{-\frac{\Delta y}{\Delta x}(l - L + y_G) - x_G}\right) \qquad (12)$$

さらに車両の並進ベクトル (t_x, t_y) は，以下の式により求められる．

$$\begin{cases} t_x = \frac{l}{\tan\theta}(\cos\alpha - 1) - (l - L)\sin\alpha \\ t_y = \frac{l}{\tan\theta}\sin\alpha + (l - L)(\cos\alpha - 1) \end{cases}. \qquad (13)$$

5.3 VPP 画像を利用した移動ベクトル推定と車両の運動推定

自車両の運動を求めるには，静止領域を基準としてカメラの相対的な運動を求める必要がある．2 章で述べた道路領域抽出手法により，抽出される領域は道路面に相当するため，静止領域である．また，VPP 画像は道路面を垂直上方から観察した画像に相当するため，VPP 画像上での車両の運動は 2 次元平面内の運動として扱うことができる．そこで，隣接するフレームの VPP 画像同士を道路領域内に対してのみ領域ベースマッチングをし，移動ベクトルを推定する．そして，その移動ベクトルから車両の運動パラメータを求めることができる．

実際の車両では，回転半径に応じた限界の旋回速度が存在する [7]．その限界旋回速度 $V_{\max}$[m/s] は経験則によって，

$$V_{\max} = \sqrt{\mu R_{\mathrm{ad}} g} \qquad (14)$$

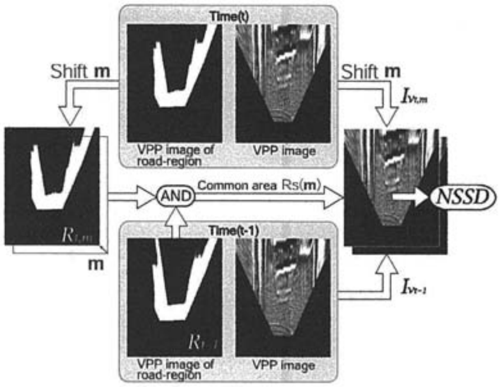

図 14: 移動ベクトルの推定手順

と表される．ただし，μ は路面とタイヤの横滑り摩擦係数，g は重力加速度，R_{ad} は旋回半径である．限界旋回速度から算出される，33[ms](30[fps] の動画像の間隔) 間のヨー角 (ヨーレート)$\alpha_{\max}$[rad/frame] は，

$$\alpha_{\max} = \frac{V_{\max}}{30 R_{\mathrm{ad}}} \qquad (15)$$

である．例えば，カメラのフレームレートを 30[fps] とし，$\mu = 0.7$ としたとき，旋回半径が 10[m] での最大のヨーレート $\alpha_{\max}$ は 0.028[rad]($\fallingdotseq$1.6[degree]) となり，ヨーレート $\alpha_{\max}$ は，小さいことがわかる．そこで，VPP 画像に映る自車両の運動によって生じるフローを一様な平行移動と近似して，移動ベクトル $\mathbf{m} = (\Delta x, \Delta y)$ を求める．

次に，移動ベクトルの推定方法を説明する．この処理の全体図を図 14 に示す．まず，時刻 $t-1$ の道路領域を VPP 画像に変換して R_{t-1} とする．同様に時刻 t の道路領域も変換し，その画像を $\mathbf{m}$ だけ平行移動させて $R_{t,\mathbf{m}}$ とする．次いで，$R_{t,\mathbf{m}}$ と R_{t-1} との共通領域 $R_s(\mathbf{m})$ を式 (16) に従って算出する．この共通領域は両時刻における道路領域に相当する．

$$R_s(\mathbf{m}) = R_{t,\mathbf{m}} \cap R_{t-1}. \qquad (16)$$

次に，時刻 $t-1$ の基準画像に LOG フィルタを施し，ヒストグラム平坦化を行なった後，VPP 画像に変換する．この VPP 画像を $I_{V_{t-1}}$ とする．時刻 t の画像に対しても同様の処理をして VPP 画像に変換した後，$\mathbf{m}$ だけ平行移動させて $I_{V_{t,\mathbf{m}}}$ とする．この 2 枚の VPP 画像の $\mathbf{m}$ に対する相違度を式 (17) に従って，共通領域 $R_s(\mathbf{m})$ 内について計算する．

$$NSSD(\mathbf{m}) = \frac{1}{N}\sum_{\mathbf{x}\in R_s(\mathbf{m})}[I_{V_{t,\mathbf{m}}}(\mathbf{x}) - I_{V_{t-1}}(\mathbf{x})]^2. \qquad (17)$$

ただし，N は共通領域 $R_s(\mathbf{m})$ の面積を示す．この相違度が最も小さくなる (類似度が最も高くなる) 移動ベ

クトル $\mathbf{m}$ を以下のように探索する．

$$\hat{\mathbf{m}} = \arg\min_{\mathbf{m}\in\mathbf{M}_0}\{NSSD(\mathbf{m})\}. \tag{18}$$

ただし，$\mathbf{M}_0$ はフレーム間の移動ベクトルの探索範囲である．

最後に，求められた移動ベクトルから車両の運動を算出する．まず道路面とカメラ間距離および仮想カメラの内部パラメータから VPP 画像の 1 画素に相当する実距離を算出し，移動ベクトルの成分である $(\Delta x, \Delta y)$ の単位系を実座標系に一致させる．次に $\Delta x = 0$ なら車両は直進運動であるので式 (7) に，それ以外は式 (12) に代入する．ここで注目点 (x_G, y_G) は，共通領域の重心位置を用いる[§]．

フレーム間の回転角度が十分に小さいことから，隣接するフレーム間では VPP 画像中の道路領域全体の運動が純粋な平行移動であると仮定してマッチングを行ない，移動ベクトルを推定する．そして車両の運動モデルを利用し，その移動ベクトルと重心座標から回転運動を含めた自車両の運動を推定できる．そのため画像中から特定の注目点を選んだり，そのトラッキングを行なったりする必要がない．

5.4 運動パラメータの更新

本節では，前節で求められた運動パラメータを初期値として，より高精度に運動を推定する方法を述べる．

車両の運動は VPP 画像上では回転運動と並進運動であるため，運動推定は 2 枚の VPP 画像間のユークリッド変換行列を求めることに相当する．ここでは文献 [24] に述べられている勾配法の考え方を利用して，新たに定式化を行う．勾配法では適切な初期値が必要であるため，前節で求められた運動パラメータを初期値として用いる．

まず 2 枚の画像間の関係は，ユークリッド変換行列 $\mathbf{M}$ を用いて，

$$\mathbf{x}' \simeq \mathbf{M}\mathbf{x} \tag{19}$$

と表される．ただし，

$$\mathbf{M} = \begin{bmatrix} \cos\alpha & \sin\alpha & t_x \\ -\sin\alpha & \cos\alpha & t_y \\ 0 & 0 & 1 \end{bmatrix} \tag{20}$$

である．このとき，

$$\mathbf{M} \leftarrow \mathbf{M}(\mathbf{I}+\mathbf{D}) \tag{21}$$

として，ユークリッド変換行列 $\mathbf{M}$ を更新することで最適なユークリッド変換行列 $\mathbf{M}$ を求めるとする．ただし，$\mathbf{D}$ は更新量を表す行列である (後に式 (24) でその内容が示される)．最適なユークリッド変換行列 $\mathbf{M}$ は，時刻 t における画像 I_{V_t} をユークリッド変換行列 $\mathbf{M}$ で変換した画像 $\tilde{I}_{V_t}$ と時刻 $t-1$ における画像 $I_{V_{t-1}}$ の相違度 E が最小になるように求められる．

まず更新量 $(\mathbf{I}+\mathbf{D})$ の求め方を簡単に説明する．$\mathbf{x}'' = (\mathbf{I}+\mathbf{D})\mathbf{x}$ と置き，相違度 E を

$$E(\mathbf{d}) = \frac{1}{N}\sum_{i\in R_s}\left[\tilde{I}_{V_t}(\mathbf{x}''_i) - I_{V_{t-1}}(\mathbf{x}_i)\right]^2 \tag{22}$$

と定義する．ただし，$\mathbf{d}$ は $\mathbf{D}$ の各要素を並べた 9 次元ベクトルである．式 (22) をテーラー展開することで，

$$E(\mathbf{d}) \approx \frac{1}{N}\sum_{i\in R_s}\left[\mathbf{g}_i^T\mathbf{J}_i^T\mathbf{d} + e_i\right]^2 \tag{23}$$

を得る．ただし，$\mathbf{g}_i^T = \nabla\tilde{I}_{V_t}(\mathbf{x}_i)$，$e_i = \tilde{I}_{V_t}(\mathbf{x}_i) - I_{V_{t-1}}(\mathbf{x}_i)$，$\mathbf{J}_i = \frac{\partial\mathbf{x}''}{\partial\mathbf{d}}$ である．ユークリッド変換では，回転成分 ω と並進移動成分 (dt_x, dt_y) の更新量を考える必要がある．それによって，以下の式を得ることができる．

$$\mathbf{I}+\mathbf{D} = \begin{bmatrix} \mathbf{I}+\mathbf{X}(\omega) & \mathbf{dt} \\ \mathbf{0} & 1 \end{bmatrix} \tag{24}$$

となる．ただし，

$$\mathbf{X}(\omega) = \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix},\quad \mathbf{dt} = [dt_x, dt_y]^T \tag{25}$$

である．次いで，式 (23) の $\mathbf{d}$ を $\mathbf{\Theta}$ に置き換えると，

$$E(\mathbf{\Theta}) \approx \frac{1}{N}\sum_{i\in R_s}\left[\mathbf{g}_i^T\mathbf{J}_{\mathbf{\Theta}_i}^T\mathbf{\Theta} + e_i\right]^2 \tag{26}$$

となる．ただし，

$$\begin{aligned} \mathbf{\Theta} &= [\omega, dt_x, dt_y]^T \\ \mathbf{J}_{\mathbf{\Theta}i} &= \frac{\partial\mathbf{x}''}{\partial\mathbf{\Theta}} = \frac{\partial\mathbf{x}''}{\partial\mathbf{d}}\frac{\partial\mathbf{d}}{\partial\mathbf{\Theta}} \\ &= \begin{bmatrix} -y & 1 & 0 \\ x & 0 & 1 \end{bmatrix}^T \end{aligned} \tag{27}$$

である．相違度 E が最小になるには，変数 $\mathbf{\Theta}$ に関して式 (26) が最小になればよい．つまり式 (26) を $\mathbf{\Theta}$ について偏微分したものが 0 と等しい．その条件から以下の式を得ることが出来る．

$$\mathbf{A}\mathbf{\Theta} = -\mathbf{b}. \tag{28}$$

ただし，

$$\mathbf{A} = \sum_{i\in R_s}\mathbf{J}_{\mathbf{\Theta}i}\mathbf{g}_i\mathbf{g}_i^T\mathbf{J}_{\mathbf{\Theta}_i}^T,\quad \mathbf{b} = \sum_{i\in R_s}e_i\mathbf{J}_{\mathbf{\Theta}i}\mathbf{g}_i \tag{29}$$

である．式 (28) から $\mathbf{\Theta}$ を求め，ユークリッド変換行列 $\mathbf{M}$ を繰り返し計算で更新することで，最適なユークリッド変換行列 $\mathbf{M}$ を精度よく求める．求められたユークリッド変換行列 $\mathbf{M}$ から自車両の運動パラメータ (並進ベクトルと回転角) を算出する．

[§]ある画像同士のマッチングをとる場合に、マッチしたときのその領域を代表する運動は，領域の重心であると考えるのが自然である．さらに??節の合成画像実験でその妥当性を検証している．

5.5 運動推定結果

図 15 は，推定された自車両の運動推定結果である．各画像は，その画像が撮影された時刻より後の各時刻の画像を利用して推定した運動推定結果を元にして，軌跡を描いている．画像中にある 3 本の点群のうち，左側の点群は実験車両の左端，右側は右端，中の点はカメラの位置を示している．また画像上部には，推定された現在の時速とヨーレートを示している．本研究では，オフライン処理で計算を行なっているため，このような表示が可能である．

同図 (A) は，交差点を右折しているシーンである．推定された軌跡が交差点の形状と合っており，正しく推定されていることがわかる．続いて，同図 (B) は，降雪中の高速道路を走行している．雪粒や前走車の水飛沫が映り込んでいても，安定に推定されている．同図 (C) は，トンネル内を走行しているシーンである．暗いシーンにおいても，提案手法が有効であることが確かめられた．

6 空間内の位置とモーションの同時推定

移動カメラの映像から，シーン中の 3 次元位置とその 3 次元モーションを取得することで，空間位置とモーションの情報を利用したセグメンテーション，エゴモーション推定，自己位置推定や視覚誘導，3 次元地図の生成，さらに障害物の回避などに利用できる．

通常，ステレオ動画像を利用して 3 次元位置とその 3 次元モーションを推定するには，ステレオ対応点探索とトラッキング作業を別々に行なう必要があった．そのため，モーションによる物体の見え方の変化を考慮して処理することは困難であった．本章では，時系列ステレオ画像間のステレオ対応点探索とトラッキングに相当する作業を同時に行なうことで，画像中の各点ごとの 3 次元位置とその 3 次元モーションを同時に最適化して推定する手法を提案する．また，提案手法では，画像内で複数の異なるモーションが存在するようなシーンにおいても前処理なしに，各点ごとにモーションと奥行きを推定することが可能である．

本章では，まず時系列画像間の奥行き Z とモーション $\mathbf{T}_m = [T_{mx}, T_{my}, T_{mz}]^T$ による座標関係について述べ，続いて奥行きとモーションを変数としたコスト関数を定義し，そのコスト関数を最小化することによって，奥行きとモーションを同時に推定する方法を説明する．

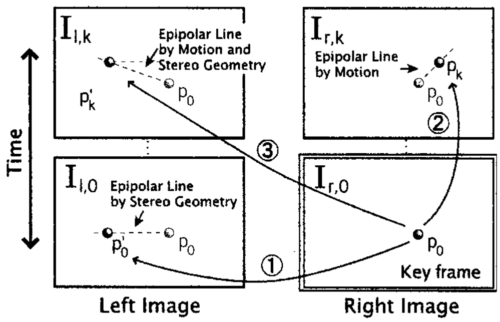

図 16: 時系列ステレオ画像の対応点の関係

6.1 時系列ステレオ画像間の対応位置

本節では，図 16 に示すような時系列ステレオ画像間の対応を考える．k 番目の左画像を $\mathbf{I}_{l,k}$ で示し，それと同時刻の右画像を $\mathbf{I}_{r,k}$ で示す．本章では，キーフレームを右画像の 0 番目とする (左画像をキーフレームとしても同様の議論ができる)．以降ではキーフレームの点 $\mathbf{p}_0 = [p_{x0}, p_{y0}]^T$ を基準として，各画像との関係について述べる．

6.1.1 同一時刻のステレオ画像間の関係

図 16 の①の関係である．奥行きが Z であるときの，点 $\mathbf{p}_0$ が左画像に写る位置を点 $\mathbf{p}'_0$ とする．また，$\mathbf{T}_s$ はステレオカメラ間の並進ベクトルである．この 2 点間には，ステレオカメラ間のエピポーラ拘束が成り立っており，

$$\mathbf{p}'_0 = \mathbf{p}_0 + \frac{\mathbf{A}_x \mathbf{T}_s}{Z} \tag{30}$$

と表される．ここで，

$$\mathbf{A}_x = \begin{bmatrix} -f & 0 & \hat{x} \\ 0 & -af & \hat{y} \end{bmatrix},\ \hat{x} = x - u_0,\ \hat{y} = y - v_0 \tag{31}$$

である．

6.1.2 同一カメラの画像間の関係

図 16 の②の関係である．k 番目の右画像中の点 $\mathbf{p}_k$ とすると，

$$\mathbf{p}_k = \mathbf{p}_0 + \frac{k\mathbf{A}_x \mathbf{T}_m}{Z - kT_{mz}} \tag{32}$$

で表すことができる．右画像には，モーションに依存するエピポーラ拘束が存在する．このエピポーラ拘束は，モーションが未知であるので一意に決まっていないが，モーションが決まると 1〜k 番目のすべての右画像のエピポーラ拘束が一意に決定する．

(A) (B) (C)

図 15: 運動推定結果 (推定された軌跡を併せて表示)

6.1.3 異なる時刻のステレオ画像間の関係

図 16 の③の関係である．k 番目の左画像上の点 $\mathbf{p}'_k$ とキーフレーム上の点 $\mathbf{p}_0$ との関係は，

$$\mathbf{p}'_k = \mathbf{p}_0 + \frac{\mathbf{A}_x(k\mathbf{T}_m + \mathbf{T}_s)}{Z - kT_{mz}} \tag{33}$$

である．左画像にも，モーションとステレオカメラ間の関係で決まるエピポーラ拘束が存在し，モーションが決まると 1〜k 番目のすべての左画像のエピポーラ拘束が一意に決定する．

また，それぞれの時刻のステレオ画像間はエピポーラ拘束が成り立っているため，各時刻のステレオ画像間には，

$$\mathbf{p}'_k = \mathbf{p}_k + \frac{\mathbf{A}_x\mathbf{T}_s}{Z - kT_{mz}} \tag{34}$$

が常に成り立っている．

以上で説明したように，モーション $\mathbf{T}_m$ と奥行き Z が決まるとすべての画像の位置を一意に決定できる．

6.2 推定アルゴリズム

本節では，画像内の点ごとのモーションと奥行きを同時に求める方法について述べる．各画像におけるキーフレーム中の点 $\mathbf{p}_0$ の対応位置は，点 $\mathbf{p}_0$ の回りにウィンドウ領域 W を考え，全画像でその領域と最も合う位置とする．そこで，コスト関数 E を以下のように設計し，最小になるモーションと奥行きのパラメータ $\mathbf{m}$ を求める．

$$\begin{aligned} E(\mathbf{m}) = &\sum_{k=0}^{n}\sum_{\mathbf{x}\in W}[\mathbf{I}_{l,k}(\mathbf{x}+\mathbf{u}_k) - \mathbf{I}_{r,0}(\mathbf{x})]^2 \\ &+ \sum_{k=1}^{n}\sum_{\mathbf{x}\in W}[\mathbf{I}_{r,k}(\mathbf{x}+\mathbf{u}'_k) - \mathbf{I}_{r,0}(\mathbf{x})]^2 \end{aligned} \tag{35}$$

ここで，n は時系列ステレオ画像のペア数，$\mathbf{m} = [T_{mx}, T_{my}, T_{mz}, Z]^T$ であり，$\mathbf{I}(\mathbf{x})$ は $\mathbf{x} = [x, y]^T$ における画像の輝度を表す．また，$\mathbf{u}_k$ と $\mathbf{u}'_k$ は式 (32), 式 (33) より，

$$\mathbf{u}_k = \frac{\mathbf{A}_x(k\mathbf{T}_m + \mathbf{T}_s)}{Z - kT_{mz}} \quad , \quad \mathbf{u}'_k = \frac{k\mathbf{A}_x\mathbf{T}_m}{Z - kT_{mz}} \tag{36}$$

である．式 (35) の前半は時系列の左右画像間の SSSD (sum of SSDs) を表し，後半は時系列の右画像間の SSSD を意味している．

まず，式 (36) の $\mathbf{u}_k$ について，$\Delta\mathbf{m}$ の変化分を考え，$\mathbf{u}^{\Delta}_k$ とする．続いて，その変化量は真値に対して小さい，つまり $(Z - kT_{mz}) >> (\Delta Z - k\Delta T_z)$ と仮定すると，

$$\mathbf{u}^{\Delta}_k = \left(\frac{1}{\dot{Z}_k} + \xi'\right)\mathbf{A}_x\{k(\mathbf{T}_m + \Delta\mathbf{T}) + \mathbf{T}_s\} \tag{37}$$

ここで，

$$\begin{aligned} \dot{Z}_k &= Z - kT_{mz} \\ \xi' &= \tfrac{-\Delta Z + k\Delta T_z}{\dot{Z}_k^2} \end{aligned} \tag{38}$$

である．これを利用して，以下のように整理する．

$$\mathbf{x} + \mathbf{u}^{\Delta}_k = \begin{bmatrix} \gamma & 0 & \frac{\alpha'}{\dot{Z}_k} \\ 0 & \gamma & \frac{\beta'}{\dot{Z}_k} \\ 0 & 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} \frac{\kappa}{\gamma} & 0 & \frac{\zeta}{\gamma} \\ 0 & \frac{\kappa}{\gamma} & \frac{\eta}{\gamma} \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{I}+\mathbf{D}} \tilde{\mathbf{x}} \tag{39}$$

ここで，

$$\begin{aligned} \tilde{\mathbf{x}} &= (x, y, 1)^T \\ \alpha' &= -kfT_{mx} - ku_0T_{mz} - fT_{sx} \\ \beta' &= -kafT_{my} - kv_0T_{mz} \\ \gamma &= 1 + \tfrac{kT_{mz}}{\dot{Z}_k} \\ \kappa &= 1 + \tfrac{k}{\dot{Z}_k}(T_{mz} + \Delta T_z) + \xi' kT_{mz} \\ \zeta &= -\tfrac{k}{\dot{Z}_k}(f\Delta T_x + u_0\Delta T_z) + \xi'\alpha' \\ \eta &= -\tfrac{k}{\dot{Z}_k}(af\Delta T_y + v_0\Delta T_z) + \xi'\beta' \end{aligned} \tag{40}$$

で表される．次に微小な要素を持つ行列 $\mathbf{D}$ を用い，座標変換 $\tilde{\mathbf{x}}'' \sim (\mathbf{I}+\mathbf{D})\tilde{\mathbf{x}}$ を考える．ここで，$\tilde{\mathbf{x}}$ は $\mathbf{x}$ の同次座標表現とし，$\mathbf{I}$ は 3 次の単位ベクトルである．$\tilde{\mathbf{I}}_{l,k}(\mathbf{x}'')$¶について座標 $\mathbf{x}$ の周りで 1 次のテイラー展開をすると，

$$\begin{aligned} \mathbf{I}_{l,k}(\mathbf{x} + \mathbf{u}^{\Delta}_k) - \mathbf{I}_{r,0}(\mathbf{x}) &= \\ \tilde{\mathbf{I}}_{l,k}(\mathbf{x}'') - \mathbf{I}_{r,0}(\mathbf{x}) &\approx \mathbf{g}^T_{l,k}\mathbf{J}^T_{l,k}\Delta\mathbf{m} + e_{lr,k} \end{aligned} \tag{41}$$

¶$\tilde{\mathbf{I}}_{l,k}(\mathbf{x})$ は $\mathbf{I}_{l,k}$ を座標 $(\mathbf{x} + \mathbf{u}_k)$ でサブサンプリングして生成した画像である．

が得られる．ここで，

$$\begin{aligned} \mathbf{g}_{l,k}^T &= \nabla\tilde{\mathbf{I}}_{l,k}(\mathbf{x}) \\ \mathbf{J}_{l,k}^T &= \frac{\partial \mathbf{x}''}{\partial \Delta\mathbf{m}} \\ e_{lr,k} &= \tilde{\mathbf{I}}_{l,k}(\mathbf{x}) - \mathbf{I}_{r,0}(\mathbf{x}) \end{aligned} \tag{42}$$

である．さて，Jacobian を構成する行列は

$$\frac{\partial \mathbf{x}''}{\partial \Delta\mathbf{m}} = \frac{\partial \mathbf{d}}{\partial \Delta\mathbf{m}}\frac{\partial \mathbf{x}''}{\partial \mathbf{d}} \tag{43}$$

と変形できる．$\mathbf{d}$ は，行列 $\mathbf{D}$ を行ごとに並べたベクトルである．右辺の後半は，

$$\frac{\partial \mathbf{x}''}{\partial \mathbf{d}} = \begin{bmatrix} x & y & 1 & 0 & 0 & 0 & -x^2 & -xy & -x \\ 0 & 0 & 0 & x & y & 1 & -xy & -y^2 & -y \end{bmatrix}^T \tag{44}$$

である [24]．右辺の前半は，式 (39) の $\mathbf{D}$ に相当する行列を $\Delta\mathbf{m}$ で偏微分することで得られる．これらにより Jacobian は，

$$\frac{\partial \mathbf{x}''}{\partial \Delta\mathbf{m}} = \frac{1}{\gamma}\begin{bmatrix} -\frac{kf}{Z_k} & 0 \\ 0 & -\frac{kaf}{Z_k} \\ \mu & \nu \\ -\frac{\alpha'+kxT_{mz}}{Z_k^2} & -\frac{\beta'+kyT_{mz}}{Z_k^2} \end{bmatrix} \tag{45}$$

となる．ここで，

$$\begin{aligned} \mu &= \frac{k}{Z_k}(x-u_0) + \frac{k}{Z_k^2}(\alpha' + kT_{mz}x) \\ \nu &= \frac{k}{Z_k}(y-v_0) + \frac{k}{Z_k^2}(\beta' + kT_{mz}y) \end{aligned} \tag{46}$$

である．

$\mathbf{u}'_k$ についても，$\mathbf{u}_k$ と同様に式展開をすることができる．こちらはステレオ画像間の $\mathbf{T}_s$ を $\mathbf{0}$ と考えればよい．よって，コスト関数 E は以下のようになる．

$$\begin{aligned} E(\mathbf{m}+\Delta\mathbf{m}) = &\sum_{k=0}^{n}\sum_{\mathbf{x}\in W}\left[\mathbf{g}_{l,k}^T\mathbf{J}_{l,k}^T\Delta\mathbf{m} + e_{lr,k}\right]^2 \\ &+ \sum_{k=1}^{n}\sum_{\mathbf{x}\in W}\left[\mathbf{g}_{r,k}^T\mathbf{J}_{r,k}^T\Delta\mathbf{m} + e_{rr,k}\right]^2 \end{aligned} \tag{47}$$

コスト関数の最小値を求めるために，$\Delta\mathbf{m}$ で偏微分して 0 となるようにし，整理すると以下が得られる．

$$\mathbf{A}_s\Delta\mathbf{m} = -\mathbf{b}_s \tag{48}$$

ここで，

$$\begin{aligned} \mathbf{A}_s &= \sum_{k=0}^{n}\sum_{\mathbf{x}\in W}\mathbf{J}_{l,k}\mathbf{g}_{l,k}\mathbf{g}_{l,k}^T\mathbf{J}_{l,k}^T + \sum_{k=1}^{n}\sum_{\mathbf{x}\in W}\mathbf{J}_{r,k}\mathbf{g}_{r,k}\mathbf{g}_{r,k}^T\mathbf{J}_{r,k}^T \\ \mathbf{b}_s &= \sum_{k=0}^{n}\sum_{\mathbf{x}\in W}e_{lr,k}\mathbf{J}_{l,k}\mathbf{g}_{l,k} + \sum_{k=1}^{n}\sum_{\mathbf{x}\in W}e_{rr,k}\mathbf{J}_{r,k}\mathbf{g}_{r,k} \end{aligned} \tag{49}$$

と表される．コスト関数を最小にするためは，式 (48) から $\Delta\mathbf{m}$ を繰り返し求め，

$$\mathbf{m} \leftarrow \mathbf{m} + \Delta\mathbf{m} \tag{50}$$

として，$\mathbf{m}$ を更新する．これによって，奥行きとモーションを同時に最適化しており，複数枚の時系列ステレオ画像間の対応を同時に推定していることに相当する．

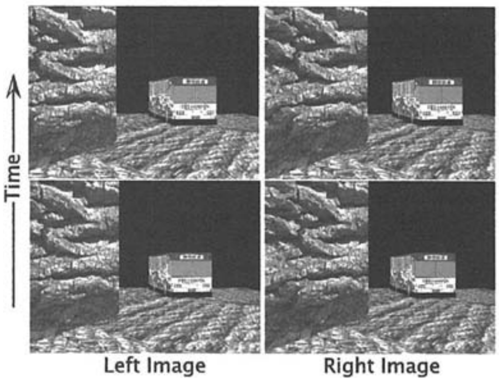

図 17: 入力画像

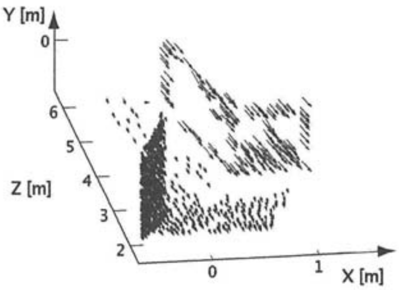

図 18: モーション推定結果．矢印の方向と長さは推定されたモーションの方向と大きさ，矢印の位置は，推定された 3 次元位置を示す．

6.3 位置とモーションの推定結果

図 17 は，入力画像の一部である．このシーン中では，床面と側壁が静止しており，バスが移動している．ステレオカメラは，ベースライン長を 0.15[m] にして，水平に配置し，床面に対して垂直にカメラの Y(縦) 軸を配置した．このシーン中でのモーションは 2 つあり，1 つはカメラが床面上を平行に奥行き方向に 0.1[m/frame] で移動している．もう 1 つはバスがカメラと正対から 10[degree] 傾いて 0.2[m/frame] で移動している．つまり画像中に存在するモーションは，カメラのモーションが加わったバスの相対モーションとカメラモーションによる側壁と床面のモーションである．

図 18 は，推定された 3 次元位置に推定された 3 次元モーションを併せて表示した結果である．矢印の方向は，モーションの方向を，長さはモーションの大きさをそれぞれ表している．同図では，斜め上から観察した様子である．この結果から，バス上の点と壁や床のモーションが矢印の大きさや方向が異なっていることが分かり，3 次元位置とモーションが正確に推定されていることが分かる．

6.3.1 障害物の絶対運動

本実験では，障害物上の特徴点それぞれの絶対速度を推定する．絶対速度に変換することによって，例えば障害物が移動しているのか否かが分かり，障害物の危険性を判断する情報となる．

まず，4 章で述べた障害物検出処理によって，検出された障害物領域内における特徴点に関して，本章で提案した手法により，3 次元モーションを推定し，それを道路座標系へ変換する．それによって，障害物の道路座標系における 3 次元モーション (相対モーション) を推定することができる．続いて，5 章で述べた自車両の運動推定結果を利用することで，その相対モーションを，道路などの静止座標系を基準とした絶対モーションに変換する．

図 19 は，左から X 軸 (横) 方向，Y 軸 (高さ) 方向，Z 軸 (奥行き) 方向の絶対モーション推定結果を示している．推定された絶対モーションは，特徴点ごとに円形で表示し，色で絶対速度を，円の直径は特徴点の推定された距離が近いほど大きく表している．また，左上に表示されている数字は，そのフレームの推定された自車両の速度と回転速度である．同図 (A) では，前走車 2 台と側壁の点について，それぞれ絶対モーションを推定した結果を示している．まず，X 軸方向の速度を見ると前走車上の点は，側壁の点に比べて若干黒みがかっており，横方向に運動していることが分かる．Y 軸方向は，どの点もほとんど運動が無いことが分かる．Z 軸方向については，前走車の点が大きく移動していることが分かる．続いて，同図 (B) は，対向車も右側に写っているシーンである．対向車の Z 軸方向の速度が，前走車とは異なり，明るい色となっており，逆向きの運動であることが分かる．

7 まとめ

本論文では，車載ステレオカメラを利用した車両の前方環境認識に関する手法を提案した．自動車は道路面に沿って走行することから，本研究では，道路に関する情報が要であると考えた．そこで，最初に車載カメラの映像から空間中の道路領域と道路面とカメラ間の姿勢を動的に推定した．次に，3 章では，それらを活用して，仮想的に道路面を垂直上方から観測した画像である VPP 画像を生成した．VPP 画像上での道路領域と非道路領域との境界の時間変化を計測することによって，危険領域を検出し，その空間位置を推定した．4 章では，ステレオ計測を利用して，画像中の各点の 3 次元位置を測定し，それらの点の集合から，道路上に存在する立体物を，道路面の姿勢を利用して分離した後，セグメンテーション処理を行なうことで，障害物を検出した．次に 5 章では，道路領域と道路面の姿勢を用い，自動車の運動を道路面上で拘束のある運動と考えて，自車両の運動を推定する手法を提案した．加えて，6 章では，時系列ステレオ画像間の対応を同時に取ることによって，シーン中の各点の空間位置とモーションを一度に推定した．また，それを利用した障害物の絶対速度推定結果を示した．

謝辞　本研究を進めるにあたり，道路領域抽出処理プログラムを提供して下さった杉本茂樹氏に感謝いたします．また，実験用画像撮影において御協力いただいた，(株) 本田技術研究所栃木研究所の飯星明氏，藤巻朋氏，植田恭史氏，奥富研究室の田中健一氏に感謝致します．

参考文献

[1] 安部正人. 自動車の運動と制御 (第 2 版). 山海堂, 2003.

[2] T. Aizawa, A. Tanaka, Higashikage H, Y. Asokawa, M. Kimachi, and S. Ogata. Road surface estimation robust against vehicle's existence for stereo-based vehicle detection. In *Proc. Intelligent Transportation Systems*, pp. 43–48, 2002.

[3] N. Cornelis, K. Cornelis, and L. Van Gool. Fast compact city modeling for navigation pre-visualization. In *Proc. IEEE CVPR*, Vol. 2, pp. 1339–1344, 2006.

[4] P. Fua. A parallel stereo algorithm that produces dense depth maps and preserves image features. *Machine Vision and Applications*, Vol. 6, pp. 35–49, 1993.

[5] J. Goldbeck and B. Huertgen. Lane detection and tracking by video sensors. In *Proc. Intelligent Transportation Systems*, pp. 77–79, 1999.

[6] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision Second Edition*. Cambridge University Press, 2003.

[7] 林洋. 実用 自動車事故鑑定工学. 技術書院, 1992.

[8] Z. Hu and K. Uchimura. U-v-disparity: An efficient algorithm for stereovision based scene analysis. In *Proc. IEEE Intelligent Vehicles Symposium*, Vol. 1, pp. 48–54, 2005.

[9] 窪田進, 仲野剛, 岡本恭一. 車載ステレオ障害物検出システムのための実時間処理可能な全体最適化アルゴリズム. 画像の認識・理解シンポジウム (MIRU), pp. 106–113, 2006.

[10] R. Labayrade, D. Aubert, and J.P. Rarel. Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation. In *Proc. IEEE Intelligent Vehicles Symposium*, Vol. 2, pp. 646–651, 2002.

[11] 森屋崇, 岩佐まもる, 石川直人, 中島直人. ステレオ画像を用いた市街地における車両衝突防止システム. 第 8 回画像センシングシンポジウム, pp. 77–82, 2002.

[12] 中森卓馬, 石川直人, 中島真人. 動画像処理による車両前方監視システム. 電子情報通信学会技術研究報告, pp. 1–6, 2002.

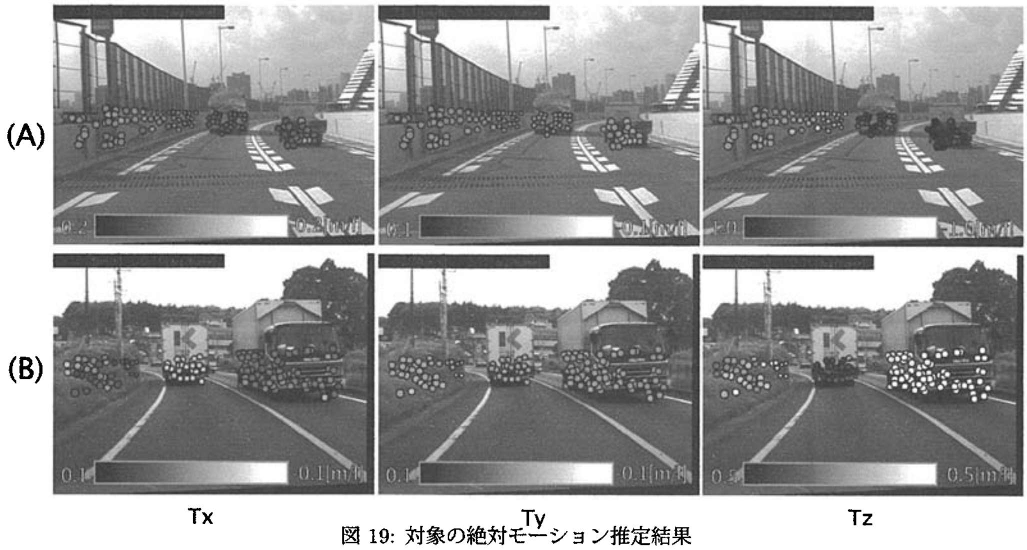

図 19: 対象の絶対モーション推定結果

[13] 中山啓満, 窪田進, 谷口恭弘, 小野口一則. 逆投影画像上での候補追跡処理による白線検出. 電子情報通信学会技術研究報告, Vol. 101, pp. 15–22, 2001.

[14] M. Okutomi, K. Nakano, J. Maruyama, and T. Hara. Robust estimation of planar regions for visual navigation using sequential stereo images. In *Proc. IEEE ICRA*, pp. 3321–3327, 2002.

[15] 奥富正敏, 中野勝之, 丸山純一, 原智章. ステレオ動画像を用いた視覚誘導のための平坦部の連続推定. 情報処理学会論文誌, Vol. 43, No. 4, pp. 1061–1069, 2002.

[16] F. Paetzold and U. Franke. Road recognition in urban environment. *Int. Journal on Image and Vision Computing*, Vol. 18, No. 5, pp. 377–387, 2000.

[17] 実吉敬二, 塙圭二, 十川能之, 荒井一真. ステレオ画像を用いた運転者支援のための前方状況認識システム. 電子情報通信学会技術研究報告, Vol. 30, pp. 39–46, 1997.

[18] 関晃仁, 奥富正敏. ステレオ動画像を利用した平面抽出による障害物検出. 情報処理学会論文誌, Vol. 45, No. SIG 13(CVIM10), pp. 1–10, 2004.

[19] A. Seki and M. Okutomi. Ego-motion estimation by matching dewarped road regions using stereo images. In *Proc.ICRA*, Vol. Vol.1, pp. 901–907, 2006.

[20] A. Seki and M. Okutomi. Robust obstacle detection in general road environment based on road extraction and pose estimation. In *Proc. IEEE Intelligent Vehicles Symposium*, Vol. Vol.1, pp. 437–444, 2006.

[21] 関晃仁, 奥富正敏. ステレオ動画像を用いた動的シーンのモーションと奥行きの同時推定. 画像の認識・理解シンポジウム (MIRU), pp. 352–357, 2006.

[22] 関晃仁, 奥富正敏. ステレオ動画像を利用した道路領域の抽出と追跡による自車両の運動推定. 情報処理学会論文誌, Vol. 47, No. SIG 5(CVIM10), pp. 90–99, 2006.

[23] 関晃仁, 奥富正敏. 道路面の抽出・姿勢推定を元にした一般道路環境下におけるロバストな障害物検出. 電子情報通信学会論文誌, Vol. J89-D, No. 8, pp. 1859–1868, 2006.

[24] H.Y. Shum and R. Szeliski. Systems and experiment paper: Construction of panoramic image mosaics with global and local alignment. *International Journal of Computer Vision*, Vol. 36, No. 2, pp. 101–130, 2000.

[25] 杉本茂樹, 奥富正敏. ステレオ画像を利用した平面姿勢推定手法と多眼カメラへの拡張. 情処学研報 (2005-CVIM-151), Vol. 2005, No. 112, pp. 131–138, 2005.

[26] T. Tsuji, H. Hattori, M. Watanabe, and N. Nagaoka. Development of night-vision system. *IEEE Transactions on Intelligent Transportation System*, Vol. 3, No. 3, pp. 203–209, 2002.

[27] S. Vedula, S. Baker, P. Rander, R. Collins, and T. Kanade. Three-dimensional scene flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 3, pp. 475–480, 2005.

[28] T. Williamson. A high-performance stereo vision system for obstacle detection. Robotics Institute, Carnegie Mellon University CMU-RI-TR-98-24, September 1998.

[29] T. Williamson and C. Thorpe. Detection of small obstacles at long range using multibaseline stereo. In *Proc. IEEE Intelligent Vehicles Symposium*, pp. 311–316, 1998.

[30] T. Williamson and C. Thorpe. A specialized multibaseline stereo technique for obstacle detection. In *Proc. IEEE CVPR*, pp. 238–244, 1998.

[31] Q. Yu, H. Araujo, and H. Wang. Stereo-vision based real time obstacle detection for urban environments. In *Proc. International Conference on Advanced Robotics*, pp. 1671–1676, 2003.

[32] 古川賢司, 岡田隆三, 谷口恭弘, 小野口一則. 車載用画像処理 LSI を用いた車両周辺監視システム. 第 9 回画像センシングシンポジウム, pp. 227–232, 2003.