

第3章 ステレオカメラの基礎

3.1 車載カメラによる画像の取得

本題に入る前に、処理すべき対象である、車載カメラで撮像される画像がどのようなものであるかを知っておくと都合がよい。撮像された画像の一例を図3-1に示す。

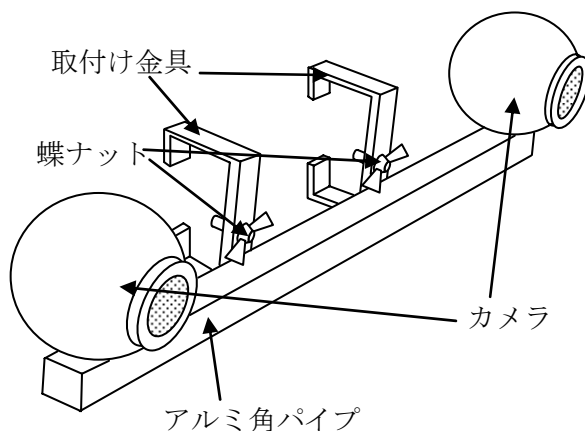
車載カメラで得られる画像は路面、前方車両、歩行者やその他の立体物、白線、道路境界構造物など独特な世界が映し出され、しかもその時間的な変化も独特である。机上であれこれ想像するより、読者の方々が自らカメラを車載して画像を取得し、それを処理することによって、理解を深めていただければと思う。



図3-1 車載カメラで得られる画像の例

画像を取得するためには、まず自動車にカメラを取り付けなければならない。そのとき、どこにカメラを取り付けるべきかが初めに問題となる。1970年代から1980年代に掛けての初期の開発時代にはカメラも大きかったため、ボンネットの上や屋根の上など屋外に取り付けられていた。その後カメラも小型化し、室内に取り付けられるようになった。カメラの位置は、できるだけ高いほうがよい。自動車の周囲にある物体は通常地面に接しているため、高い位置から撮像した方が個々の物体の重なりが少なくなって物体が分離しやすくなり、カメラ位置からの距離の分解能も上がるためである。室内での位置は、視野の妨げにならないようルームミラーの前方によく設置される。この際、カメラの視野

がワイパーの払拭範囲内に入るようにする。簡単な実験用のカメラとしては、市販されている安価な USB 接続のカメラで充分である。図 3-2 に、筆者が手軽な撮影として用いているカメラと自動車に取り付けた様子を示す。



(1) カメラ本体と取り付け具



(2) 車載した様子

図 3-2 実験用車載カメラ

図 3-2 に示すように簡単な取り付け金具を作って、ルームミラーの裏面に直接取り付けているが、取り付け方にはコツがある。カメラを向ける方向は、空をできるだけ含めないように下を向ける。大体、水平線が上から 5~10%になるようにする。そうするとカメラの視野角にもよるが、下の方はボンネットが

写ってしまう。そこで筆者の場合、上から 50%の領域を処理に用いて下の部分は捨てている。無駄な処理をしない分、処理速度が向上する。この類いのカメラはシャッタ速度や絞りを手動で設定（固定）できるので、都合がよい。

近頃のカメラは、カラーカメラがほとんどである。標識や信号を識別するためには、カラーカメラが必須であろう。しかし、ここで対象としているのは道路の白線や障害物であり、これらを認識する場合には、筆者の経験からカラー情報を使うことによる能力の向上はあまりない。しかも安価な単板式 CCD カメラでは、空間分解能が低下するので、総合的にはカラーフィルタのない白黒カメラのほうがよい。但し、カラーカメラの方が入手しやすく安価なため、カラーカメラを使わざるを得ない状況が増えている。その場合は、RGB のままでは処理に時間が掛かる割に性能が良くならないので、

$$B = 0.299R + 0.587G + 0.114B \quad (3.1)$$

の式を用いて白黒の輝度情報に変換する。ここで R、G、B は赤、青、緑それぞれの輝度である。

このカメラシステムは、2 台のカメラを並列させてステレオ撮影できる形になっているが、USB 接続のカメラは今のところ外部同期を取って、2 台同時に撮影することができない。そのため、リアルタイムでステレオ撮影するためにはかなりの改造が必要である。簡単に撮像するには、2 台のパソコンを用いてそれぞれのカメラを接続して動画撮影し、タイミングの合っている場面を取り出すようにする。シャッタ速度や絞りは固定にしておく。最大 1/30 秒のずれが生じるが、うまくタイミングの合った画像を選べば、そこそこのステレオ原画像が得られる。

とは言え、ステレオカメラは同期が命である。理想的にはカメラは必ず外部同期方式の物を使い、外部から 2 つのカメラに同期信号を送って、シャッタタイミングの同期を取る。筆者らは専用のボード（画像メモリ付き）を作って使っているが、近頃は PCI バスの画像取得ボードでステレオカメラに対応した物も発売されている。しかし車載は面倒で、上で述べたように USB カメラを改造して、ノートパソコンでも安くて手軽に撮影できる USB のステレオカメラが手軽に入手できればよいのだが、と思っている。

この他のステレオ撮影上の注意点を述べておく。2 つのカメラの間はしっかりと固定し、少々のことではずれないようにしておく。もちろん、ずれてもそのたびに校正（カメラの相対位置を求め、平行等位になるようにすること）をすればよいのだが、たびたびでは面倒である。また、感度をどちらも同じになるようにするために、できれば片方の電子シャッタやオートアイリスの信号を

両方のカメラで共用する。それができなければ、手動にしていつも同じ設定になるように調整する。ステレオカメラの場合横幅が長くなるので、ワイパーの払拭面積に入りづらくなるが、できるだけ入るように設置する。ルームミラーの位置に取り付けると、ダッシュボードからの反射光がフロントガラスに、更に反射してカメラに入り、写り込みが生じやすい。左右同じ写り込みということはないので、写り込みが起らないようにする。簡単な方法としては、写真撮影などで使う黒い反射防止布をダッシュボードに貼る。これは、通常の運転時でも視界がクリアになり、お勧めである。

3.2 ステレオ画像処理の流れ

通常、画像処理と言えば1台のカメラ（単眼）で得られた画像を処理することを指す。また、多くの車載カメラシステムでは、1台の車載カメラから得られた画像を処理して白線などを検出している。1台のカメラで得られた画像から白線や障害物を検出するということは、3次元の世界を2次元で捉えて、そこから3次元の情報を復元しようというものである。従って、さまざまな困難があり、いろいろな工夫を強いられる。ステレオカメラによる方法は、3次元の世界をそのまま3次元で捉えるのだから、まさに次元が違う。もちろん、ステレオカメラにも固有の問題点はあるが、物体の抽出や路面の決定といった重要で基本的な問題に圧倒的な強みを見せてくれる。本項ではここから、そのステレオカメラについて、基本的な事柄から実用化に当たったの問題点とその解決法に至るまで、順を追って説明していく。

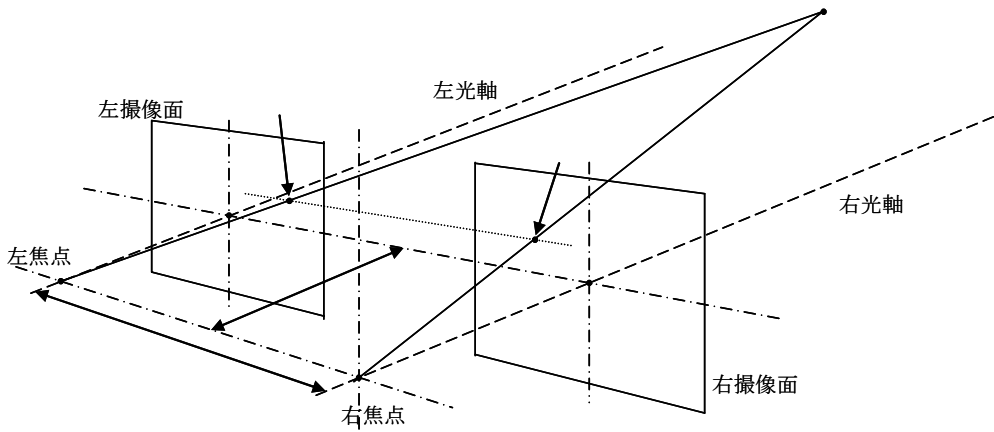


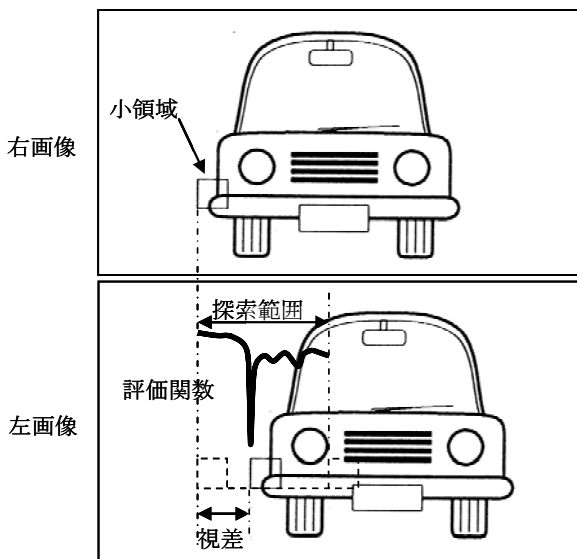
図 3-3 平行等位に配置された左右カメラ

ステレオ画像認識は複数台のカメラで物体を同時に撮影し、各カメラで得られたその物体の画像上での映る位置の違い（視差）から、その物体の位置や立体的な形状を認識する技術である。世の中にはカメラを3台以上使ったシステムも多くあるが、ここでは筆者らが自動車用として開発してきた2台のカメラを使った方法について説明する。まず、2台の同一仕様のカメラを図3-3のように平行等位に設置する。平行等位とは2つのカメラの光軸を平行にして、撮像面を一致させ、更に撮像面の横軸（ x 軸）も一致させた（互いに回転していない）カメラ配置である。図3-3のように、 (X, Y, Z) を右カメラの焦点を原点とした実際の空間の座標系、 (x_l, y_l) 、 (x_r, y_r) をそれぞれ左カメラ、右カメラの撮像面上で、それぞれの光軸との交点を原点とした座標系とする。また、 X, x_l, x_r 軸はすべて、左カメラの焦点から右カメラの焦点に向かう方向に一致させる。

すると $y_l = y_r$ となり、また実際の空間の点 $P(X, Y, Z)$ とそれぞれのカメラでの投影点 $m_l(x_l, y_l)$ 、 $m_r(x_r, y_r)$ は、

$$Z = \frac{Bf}{x_l - x_r}, \quad X = \frac{Z}{f} x_r, \quad Y = \frac{Z}{f} y_r \quad (3.2)$$

という簡単な関係で結ばれる。ここで B はカメラ間隔、 f は焦点距離で共に定数である。そこで、 $x_l - x_r$ （視差）を計測すれば Z が求まり、右カメラでの像の位置 (x_r, y_r) より点 P の座標を求めることができる。



小領域は 4×4 画素。
左画像において、小領域の位置を1画素ずつずらしながら評価関数を計算して、評価が最も良かった（この場合は最小値）ずれ量を視差とする。実際には、これほど明快に視差が求まることは多くない。

図 3-4 視差の求め方

第7章 より高度な手法

第1節 3次元自己位置認識

7-1.1 はじめに

自動車にはまだ応用していないが、ステレオカメラを用いると、自車の姿勢と運動がわかり、更にそれを累積することで出発点からの自己位置を認識することができる。ここでは、筆者らが無人ヘリコプタ用に開発したステレオ画像方式を解説する。

自己位置を認識する方法は大別して2つある。1つはランドマークや地磁気、GPSのように外部空間に固定された情報を収集して、位置を直接的に認識する方法、そしてもう1つは、車輪の回転や舵角、加速度センサで得られる加速度などからロボット自身の動きを検出し、その動きを累積することにより、絶対位置のわかった点からの位置を相対的に認識する方法である。

絶対位置を、常にどんな条件下でも直接求めることができれば理想的であり、相対的に認識するセンサは不用になる。しかし、実際はなかなか理想通りになってくれない。例えば、ランドマークでは常にマークをロボットの見えるところに置いておかねばならず、狭い領域ならまだしも、屋外の不特定領域を移動させる場合には、多数のマークを至る所に設置しなければならず、現実的ではない。また、GPSにしても、精度についてはD-GPSの一般化で、目的によっては問題のないレベルに達しているが、ビル街や森林、山間、地下、トンネルなど衛星からの電波の届きにくいところが結構あり、GPSだけに頼ってシステムを構築することは難しい。

そこで直接計測する方法を補完するための、絶対位置のわかっている点からの動きを累積して求める、相対的な位置計測が重要になる。この方法は外部情報を用いずに、専ら自分の動きを計測するのみで軌跡や現在位置を求めるので、高精度でドリフトのない動き検出センサを用いれば、出発点を与えるのみでその後の位置がわかってしまう。そうなると直接計測の補完ではなく、単独で使用できる位置計測手段となるが、実際には直接計測と同様、理想的にはいかな

い。例えば、車輪の回転検出や舵角検出では車輪の滑りの問題があり、また、船や飛行機では当然使えない。加速度センサも、位置を出すための2回積分に耐える精度を持ったものはなかなか入手できない。また、極低加速では加速度の検出すら困難である。

このように、理想的な位置検出センサがないことから、多くのロボットではそれらを組み合わせて必要な位置精度を確保している。ところで、我々の開発しているリアルタイムステレオ画像認識システムは、現在では混雑した道路でも自律走行が可能なほどに認識レベルが向上している。今回はこの技術を元に、3次元動きベクトルを検出する部分を追加して、このステレオ方式3次元自己位置認識システムを開発した。ステレオ画像方式による位置検出は、動きを累積して相対的な位置計測をする方式に分類されるが、動き検出に対する精度が非常に高く、またドリフトも少なく、これまでのセンサと比べて、より理想的なセンサに近づいている。次項から、この方法の原理、ハードウェアによるリアルタイム処理手法、実験結果などを述べていく。

7-1-2 原理

図7-1-1に大まかな処理の流れを示す。2組（カメラ数は4台）のステレオカメラを用い、1組は前方に向けて遠方を含んだ前方風景を、もう1組は下方に向けて地表面を、同期を取って連続的に（本装置では0.1秒ごとに）撮像する。次に、前方カメラで撮影した画像をステレオ処理で距離分布画像にし、その中で遠方に写っている対象のみを選び出す。遠方の風景は並進運動に対して、あまり影響を受けないことから、選び出した対象が次の画面でどこに移動したかを追跡することにより、その間のロボットの回転運動のみを検出することができる。遠方風景が並進運動の影響をあまり受けないことは、電車に乗って車窓から月を見たとき、電車が直進運動している限り、月の見える位置が全く動かないことからもおわかりいただけると思う。カーブすれば、月はずっと窓から消えてしまう。

一方、下方カメラで撮影した地表面の画像は、やはりステレオ処理して距離分布画像に変換する。ここで、元画像においてコントラストの強い領域をいくつか（本装置では30個）選んで、それらの領域が次の画面でどこに移動したかを3次的に追う。得られた個々の領域の動きは、並進と回転の計6成分を含んでいるので、前方カメラから得られている回転成分を除去して、並進成分のみにする。原理的には、地表面の動きのみから6成分すべてが求められるはずであるが、実際には特にロール成分とピッチ成分に対して、十分な精度が得られず、実用できない。

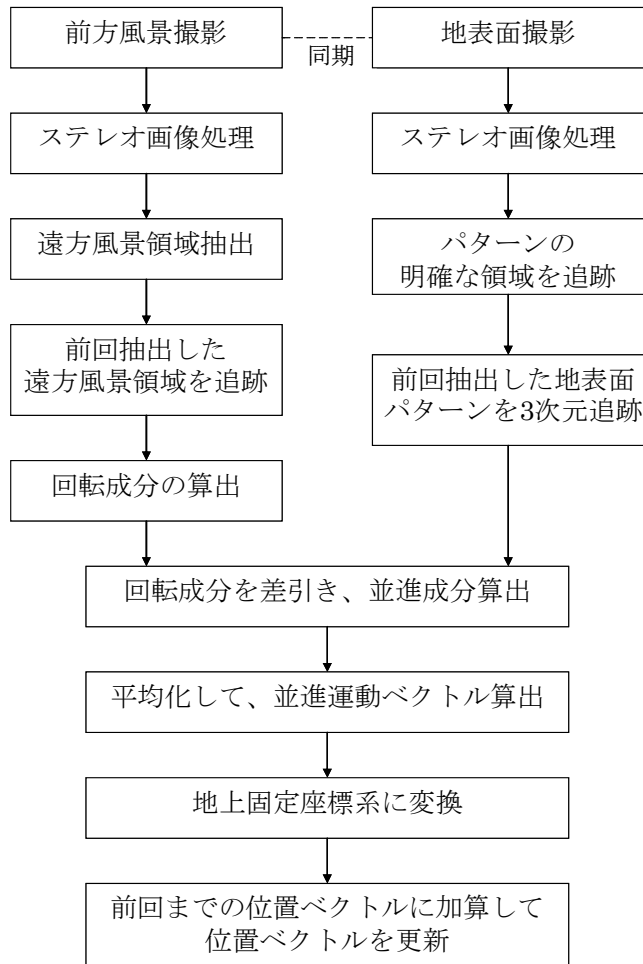


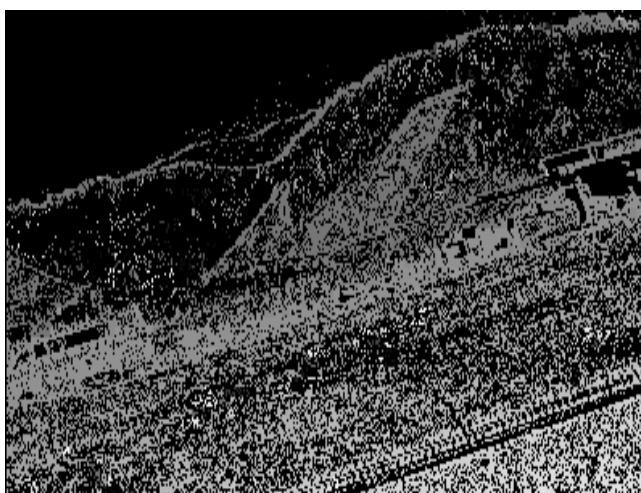
図 7-1-1 処理の流れ

こうして得られた各領域での並進成分を平均化して、3次元並進ベクトルを求めると、これと逆向きのベクトルがロボットの並進運動ベクトルになる。このベクトルは、連続した2つの画像間での動きを、前画面におけるロボットに固定した座標系で見ているので、このベクトルを地上に固定した座標系に変換する。最後に、これまでの位置ベクトルに、変換した新しいベクトルを加算して新しい位置ベクトルとする。これが、原点から移動したロボットの位置を示すベクトルである。

7-1.3 ステレオカメラ



(1) 元画像（右カメラで撮影）



(2) 距離分布画像

図 7-1-2 元画像と距離分布画像

本装置で用いているステレオカメラは、全画面を横 100×縦 50 領域に分割して、それぞれの領域に対して距離を求めて、全体として距離分布画像を生成する。画面の画素数が 400×200 画素、従って 1 領域当たりの画素数は 4×4 画素である。探索範囲は、右画像内の 1 つの領域に対して、左画像内の横方向に 0