

ステレオカメラによる自動車運転支援システム

実吉敬二^{†1}

人間の弱点を補って衝突を防止する運転者支援システムを説明する。ステレオカメラを用いることによって、複数の立体物の位置や大きさ、相対速度を短時間で求めることができる。さらに白線や黄線も頑強に検出できるので、一般的な交通環境において危険度の判断に必要な情報を得ることができる。本システムの概要と実用化のためのポイントなどを述べる。

Drive Assist System of a Car by means of Stereo Vision

KEIJI SANEYOSHI^{†1}

A drive assist system using a stereo vision to avoid collision caused by human error is presented. Using stereo camera positions, sizes and relative speeds of more than one object can be detected simultaneously with short time. The outline and advantages of the system will be mentioned as well as some important points to be realized.

1. はじめに

自動車が生産してから長い間、自動車を動かす主体は運転者であり、事故の責任は運転者にある、というのが常識であった。ところがここまで自動車が普及してくると、人間の生来持っている自然な現象、つまり疲れ、散漫、思い違いなどが原因となる防ぎようがない事故が絶対数として多くなる。もちろんこのような状態になって事故を起こす確率は一人一人にとってはそれこそ万が一である。しかし1億人を超える集団になると、1年間に5000人近くの死亡事故になってしまう。原因が人間の自然な性質なのだから、事故の責任は運転者にある、と言っているだけでは済まないことは明白であろう。

そこで自動車自身が衝突を回避するシステムの登場となる。筆者も1980年代の終わりにそのころ所属していた富士重工業(株)でステレオカメラを利用した衝突回避システムを考案し、1991年の東京モーターショーで公表した¹⁾。それ以来、20年以上に亘って研究を続けてきたが^{2) 3)}、2年ほど前に派手な宣伝効果もあってか、ようやく世の中に認知されるようになってきた。今日の講演では、開発当時の話も交えながら、ステレオカメラの優位性や交通環境認識への応用、そして実用化するためのいくつかのポイントを述べたい。

2. ステレオカメラの優位性

車線や周囲立体物を検出し、衝突の危険のある物体を認識する周囲環境認識システムで実用化されているセンサとしては、レーザーレンジファインダー、ミリ波レーダー、単眼カメラ、そしてステレオカメラが挙げられる。これらのセンサを比較して画像認識、中でもステレオカメラの優位

性を示す。

画像に含まれるデータの数は画素数である。一昔前まではVGAに代表される30万画素程度が普通であったが、今では百万画素以上が当たり前となり、一千万画素を超えるカメラも容易に手に入るようになってきた。画素は一つ一つが異なる情報を持ち、画素の数だけ情報が得られることは考えてみれば驚異的である。これに対してレーザーレンジファインダーは例えば自律走行車によく使われるVELOCITYNE社のものが現在のところ最も情報量の多いセンサであろう。図1の中央にこのレーザーレンジファインダーで得られた38mおよび64m先の前方車と、同じ場面をステレオカメラで同時に撮像した視差画像を同図右に示す。視差画像ではトンネル入り口や周囲の木々、道路面の形状まで精細に計測されている。データ量の違いは明らかであろう。ミリ波レーダーで複数の素子を切り替えて視野を確保するが、その空間分解能はさらに悪く、例えば混雑している道路で車一台一台を見分けることができない。

レーザーレンジファインダーやミリ波レーダーでは小さい物体を見落としてしまうことがある。レーザーレンジファインダーやミリ波レーダーは面を検出するからである。それに対して画像は輝度や色の変化、すなわちエッジを検出する。したがって小さな物体でも画像に写れば検出できる。ポールや柵程度の大きさであれば画像で捉えられるので、検出することは容易である。

情報量が少なくても精度や信頼性の高い情報が得られれば十分ではないかという議論もある。精度の高さという点では確かにレーザーレンジファインダーやミリ波レーダーは距離に関係なく数センチメートルの精度で距離を計測できる。ステレオカメラでは近い方でしか精度良く計測できない。しかし遠方において数センチメートル単位で測る必要はなく、遠方物体は衝突までに時間が掛かることも利用して、衝突回避が出来ることが示されている。人間の両

^{†1} 東京工業大学
Tokyo Institute of Technology

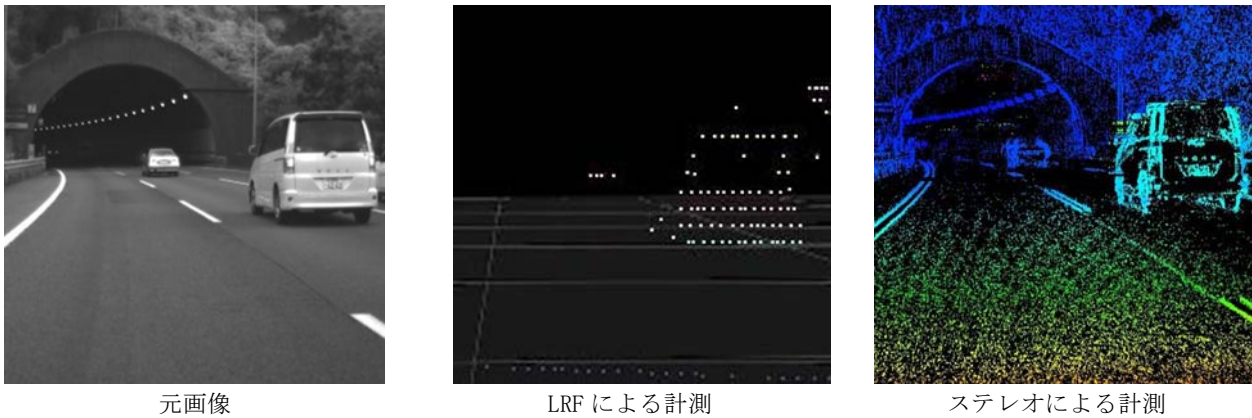


図1 レーザレンジファインダー (LRF)とステレオカメラによる計測における情報量の違い
 Figure 1 The difference of the amount of information between Laser Range Finder (LRF) and stereo vision

眼視も同じ欠点を持っているが、覚醒状態であれば時速100km/hで渋滞の後尾に迫っても衝突することはない。

らに画像による認識の特長は、白線や標識など平面上に描かれたマークを認識できることである。人間はこれらのマークを認識して運転をしている。例えば3車線の幅のある道路で白線が引かれていなかったら、ぶつからないように運転するためには大変な神経を使うだろう。センサも同じで立体物の検出・認識が出来るだけでは衝突防止に車載するセンサとしては片手落ちである。マークの検出は画像の独壇場であろう。白線の反射率が高いことを利用してレーザーで検出する研究もあるが、すべてのマークの反射率が高いわけではなく実用的ではない。

その他に、ステレオカメラではメカニカルな機構を使わずにレンズを交換するだけで視野を広くできることやリアルタイム処理の指標となる動作速度が大きいことも特長である。一昔前まではステレオカメラは処理時間の掛かり、そのために大掛かりなハードウェアが必要なことが大きな欠点と言われていたが、ハードウェアの飛躍的な進歩によって、60fps程度であっても安価なFPGAで実現されている。

なお、ステレオカメラと単眼カメラを比較すると、同じ画像認識でもステレオカメラの場合には単眼カメラと同様の2次元情報に奥行きというもう一つの情報が加わる。まさに次元の違う情報が得られるのである。自律走行車や移動ロボットは奥行き方向に進んでいくのであるから、奥行き情報は本質的に重要であり、単眼カメラのみで実際の交通環境で自律走行させることはよほど環境が整っていない限り不可能であろう。

このようにステレオカメラは、高い空間分解能と小物体やマークにも及ぶ検出率の高さ、高速処理、視野の広さという特長を持つため、他のセンサに比べ、複雑な交通環境の中で衝突の危険性のある物体をする能力に特に優れていると言えるだろう。

3. ステレオカメラ

採用しているステレオマッチング法は小領域マッチング法で以下のような原理によっている。図2に原理の説明図を示す。2台のカメラは光軸が互いに平行で撮像面が同一平面上に乗り、しかも撮像面の横軸（軸）が一致する、いわゆる平行等位に配置してある。その結果、図中の点までの距離は図中の記号を用いて

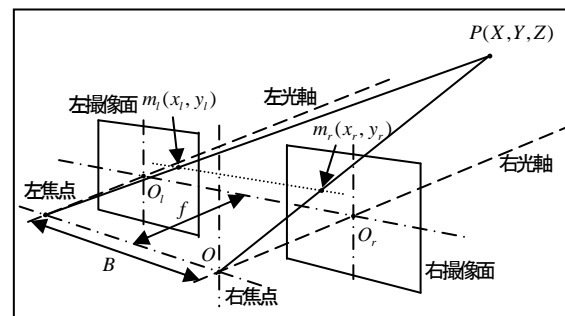


図2 ステレオカメラの配置
 Figure 2 Arrangement of two cameras

$$Z = \frac{Bf}{x_l - x_r}$$

という簡単な関係で表せる。このとき m_r , m_l には同じパターンが写っており、この左右カメラに写った同じパターンを見つけることがステレオマッチングである。我々はパターン的一致度の評価に SAD (Sum of Absolute Difference) を使い、小領域の大きさは初期には 4×2 画素、その後 4×4 画素などに変更した。探索深さはやはり初期には 100 画素、現在では 200 画素前後としている。

図3に得られた視差画像の一例を示す。これは 400 万画素のカメラで撮像したもので、後で述べる FPGA を用いて 8.5fps の処理速度で視差画像が生成されている。この図からもわかるように全面を視差で埋めるわけではない。輝度の変化のないところ、具体的には左右で隣り合う画素の輝度差が小さいところは原理的に視差が求められないところ

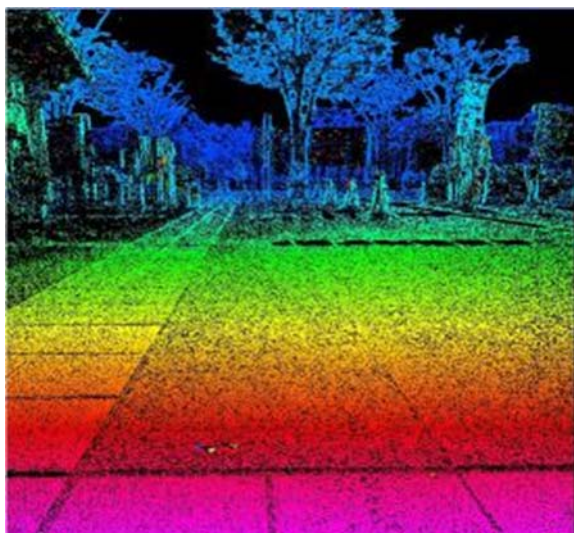


図3 視差画像(400万画素)色で視差を示す。
Figure 3 Disparity image (color indicates disparity)

なので視差データを入れていない。その部分が黒で表示されている。

4. 交通環境認識

(1) 白線の検出

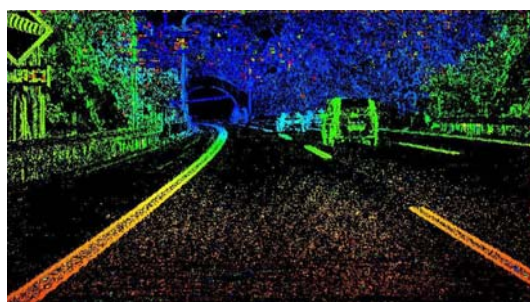
白線は車線の境界、すなわちそれに沿って走行してよい範囲を示す重要な情報である。白線検出そのものは単眼カメラによる方法が多く研究されているが、視差画像の情報を加えることでより強固な検出が可能となる。

具体的には次のとおりである。まず視差画像から道路面を特定する。これは自車の手前は立体物がなければ道路面なので、その部分の視差画像に平面を当てはめることで行なう。実空間での平面は視差空間（奥行きを距離の代わりに視差にした空間）でも平面になるので、視差空間のまま行なう。一方で元画像からソーベルフィルタを使って方向を区別してエッジを検出しておく（図4 b）。そして道路面近傍にある視差データでエッジのある部分だけ残して、道路面にあるエッジの視差画像を生成する（図4 d）。この時点ではまだ白線以外の道路の模様や汚れ、影の境界などが存在する。視差画像を用いると遠近に関係なく白線や車線の幅を求められるので、白線や車線の規格に合ったエッジの組のみを選択する。具体的には、白線 r となっているエッジペアで、その間隔が白線の幅である 15~20 cm のペアを選ぶ。さらにそのペアが車線幅である 3.5m ないし 3.75m の間隔で存在するペアの組を残す。最後にそれらのエッジを実空間に変換してからやはり規格の曲率内になるように考慮しながら接続していく。連続するフレームでは前のフレームで得られた白線の位置を利用し、片側が実線の場合はその位置から反対側の破線の白線の位置精度を高めるなどの工夫をする。図4 e に検出した結果を示す。視差画像

を利用することでガードレールや車両に描かれている白線風の模様など道路平面以外のパターンを除外することができる。さらに絶対的なサイズがわかることも規格情報が使



a. 元画像



b. 視差画像



c. エッジの抽出



d. エッジ部分の視差



e. 検出した結果

図4 白線の検出

Figure 4 White line detection

えるなど有利である。

(2) 立体物の検出

まず、視差画像を図5のように小領域の幅で縦の領域に分割する。次に各領域内のデータで、白線検出で求めた道路面より上にあつて、自車の高さより下にあるものを選択して抽出する。抽出したデータについて、領域ごとに視差を横軸にしたヒストグラムを作ると、例えば図6のようになる。立体物の多くは地面から垂直に立っており、一方カメラ光軸はほぼ水平方向を向いているので立体物はほぼ等しい視差になり、ヒストグラムに明確なピークができる。そこでピーク付近に立体物があるとしてその前後の視差も含めて平均化し、その立体物までの距離を求める。

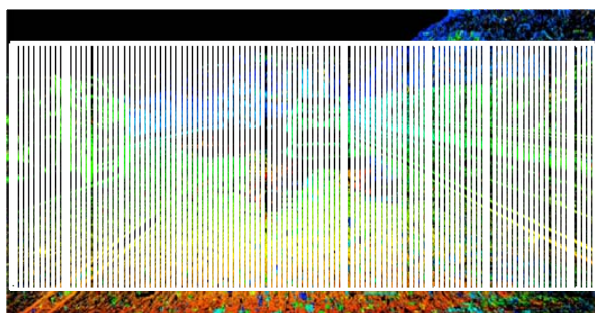


図5 マッチング領域の幅で短冊形に分割
 Figure 5 Divide to rectangles of the width of the same as that of matching area

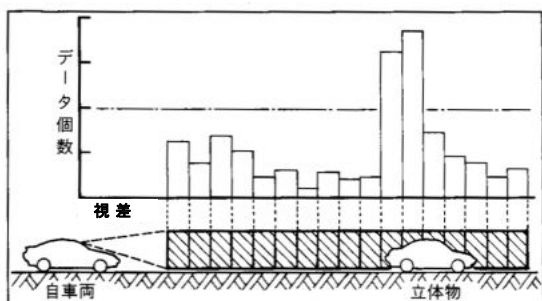


図6 短冊内の視差ヒストグラム
 Figure 6 Histogram within the rectangle



図7 短冊の視差の位置に棒として投影した結果
 Figure 7 Projected image of the rectangles according to the obtained disparities

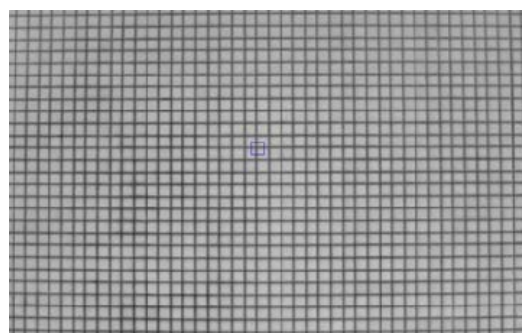
この操作をすべての短冊領域に対して行なう。すると立体物に短冊が含まれる場合、その短冊で得られた視差を距離に直してみると、あたかも立体物の位置に立った棒のようになる。その仮想的な棒を画像上に投影すると図7の色で示した部分になる。この図をみればわかるように、特に立体物として抽出しなくてもこの棒群を避けるように自車を制御すればよいので、前方車との衝突を防止するためのセンサになる。

しかし実際には立体物の左右端の位置や相対速度が重要なので、隣接した短冊が同じような距離になった領域をひとまとめにして一つの立体物とみなし、左右端の位置や全体の動きから相対速度を求める。そして、検出した複数の立体物の距離や、物体追跡から得られる相対速度によって衝突危険性を判断する。また自車の走行する車線内にある立体物のうち最も近いものを先行車両として ACC (Adaptive Cruise Control) などに利用する。

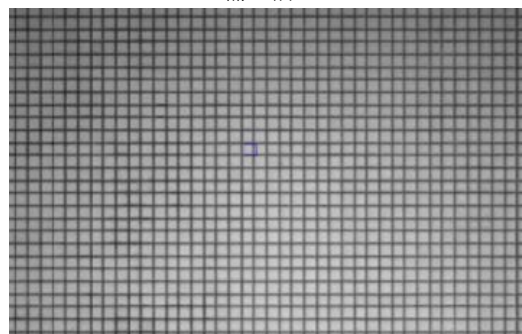
5. 実用化するためのポイント

(1) 補正と校正

ステレオカメラは、二つのカメラの内部パラメータや配置が既知でなければ正しい視差が得られない。またミスマッチングを押しえつつ短時間で探索を行なうためには画像の歪みをなくし、平行等位に配置しなければならない。これらの補正・校正によって左右画像の位置ずれを、経験上0.1画素程度に抑えると正確な視差画像が得られる。これは大



補正前



補正後

図8 補正に用いた格子
 Figure 8 A lattice pattern used for distortion modification

変厳しい制限であり、よほど注意深く補正・校正を行わなければならない。また車載カメラのように急激な温度変化や振動のある環境の中でこの精度を保っていくことは難しく、機械的なカメラの結合に頼っては実現できない。そこでステレオカメラ組み立て後にソフト的に補正・校正を行ない、さらに実際の動作中に常に自動的に校正することで長期間、性能を保てるようにした。これが実用化できた大きなポイントのひとつである。

これらの処理に特に画期的な技術を使ったわけではなく、補正には格子、校正にはいくつかの異なる距離に配したマークを用いた通常の方法で補正・校正を行った。格子は 1.5 画素程度の線幅で格子の間隔は 10 画素程度とした。格子点の計測は、まず格子点を形成する縦横の格子線の位置を点列で求め、縦横の直線をそれぞれ当てはめ、それらの直線の交点として求めた。こうすることで格子点の位置は 0.05 画素程度の精度で求めた。図 8 に補正前と補正後の格子画像を示す。ステレオカメラの校正には 7 つのパラメータを決めなければならぬことが一般に知られている。我々は、1 枚の格子板で左右カメラの補正を同時に行うことで、このパラメータを 7 個から 2 個まで減らすことが出来ることを見出し、校正を単純化した。2 個のパラメータとは光軸回りの回転と上下方向の並進で、10 個ほどの距離の異なるマークの画像上の位置からこの 2 つのパラメータを計算した。7 個のパラメータを求める方法に比べ簡単で、平行等位の配置の近傍で起こる解の不安定さもない。



図 9 校正に用いたマーク
 Figure 9 Marks used for rectification

補正・校正をした後、ユーザーが使用している間に経年変化でカメラの位置関係がずれてくることは防ぎようがない。そこでカメラが動作中にバックグラウンドジョブとして、常に画像の適当な点を選んで左右ばかりでなく上下方向にも探索して、探索範囲が水平方向からずれているかどうかを検出する。そしてずれが無視できなくなると、校正パラメータを変更する。これは経験上片側のカメラの並進・回転の調整のみで対応でき、7 つのパラメータすべてを調整する必要はない。また補正テーブルを変更することもないので、それほど難しいことではない。しかしこの仕組みがないと商品にならないほど重要な機能である。

(2) リアルタイム処理

衝突回避を目的とした障害物検出センサでのリアルタイム

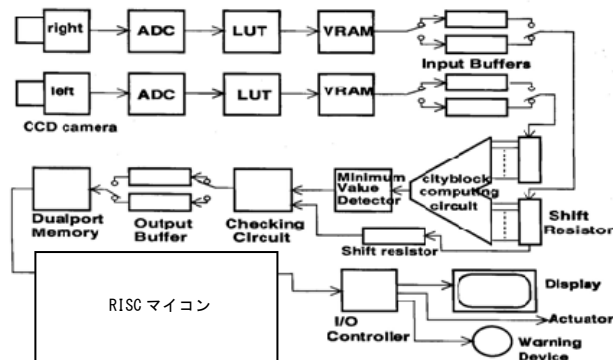


図 10 ステレオ処理回路の概略図

Figure 10 Schematic diagram of stereo processing

処理は本質的に重要である。ステレオ処理は処理が重く時間が掛かると言われてきたが、ハードウェアの飛躍的な進歩のおかげで過去の話となった。筆者が手掛け始めた 20 年前もマイコンの性能が 0.5MIPS という時代で、とてもステレオカメラを車載できるとは思われなかった。しかし処理自体は局所的で単純な繰り返し計算なので、ハードウェアでパラレル・パイプライン処理を行えばリアルタイムで実現できるのではないかと考えた。そして 50 個ばかりの ALU チップを組み合わせて、VGA サイズで探索幅 100 画素、10fps のリアルタイムステレオ処理装置を組み上げた。

図 10 にステレオ処理回路の概略図を示す。この回路構成は、1991 年に特許を申請したときの図をわかりやすく書き直したものであるが、現在でも基本的な構成は同じである。追加した回路として、VRAM の部分に補正・校正回路が加わり、最小値検出回路と並列にサブピクセル計算回路と特異点除去フィルターが加わっている。

我々のシステムでは 4×4 画素単位で処理するので、4 ライン分ごとに処理を独立させることができる。そこで、ステレオマッチング回路の前後に 4 ライン分のバッファを持たせ、前処理回路で 4 ライン分のデータがバッファに溜まるごとに処理を始める。その結果、全画面分がバッファに書き込まれるのを待たなくてよいので、高速化される。前処理もそのことを考慮して、できるだけ少ないライン数だけで処理ができるようにする。

ステレオ処理の中心となるステレオマッチング回路のブロック図を図 11 に示す。図のように初段に減算器、その次に絶対値を演算するユニットを並べ、その後に加算器をピラミッド状に接続する。これは横方向が並列処理、縦方向がパイプライン接続になっていて、各演算ユニットは 1 クロックで演算する。その結果、1 クロックごとに 4×4 画素に対する SAD 評価関数の結果が出力される。

SAD 評価関数の結果が出るごとに、最小値判定回路に送られ、そこでこれまでの最小値と比較され、より小さければ最小値として置き換えられ、そのときの視差が保存される。探索を最後まで行なった時点で保存されている視差がその領域の視差である。これを判定回路 (Checking circuit)

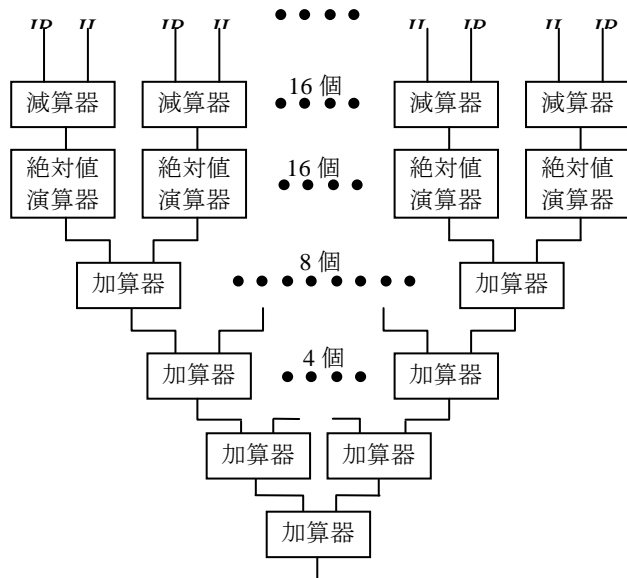


図 11 ステレオマッチング回路
 Figure 10 Stereo Matching Circuit

に送る。ここでは3章で述べた輝度の変化の少ないところのデータを除いたり、得られた視差の前後の SAD 評価値を用いて1次補間することによりサブピクセル精度で視差を求めたり、さらに mismatching 領域を削除するための特異点除去フィルタが含まれる。

得られた視差画像は再び4ラインバッファを介してデュアルポートメモリに書き込まれ、RISC マイコンによって白線検出や立体物検出が行なわれる。

リアルタイム処理を実現するうえで FPGA によるハードウェア化は欠かせない技術である。近年、FPGA (Field Programmable Gate Array) といわれる、プログラムを書き込むことによって自由に論理回路を組むことの出来る IC が急速に発達し、簡単に使えるようになった。安価なものでも容量は数十万ゲートあり、また 100MHz のクロックで動作する。上に述べた回路を FPGA に構築したとすると、演算器一つあたり 100 ゲート程度なので、全部でも 5000 ゲート程度しか使わない。またクロックも 100MHz とすると全画面を 10msec で処理できる。我々はアルテラ社の FPGA を使っているが、VGA の場合、上に述べた処理をすべて入れてもロジックエレメント (LE) 数は 5000 程度で、同社の FPGA ではグレードの低い CycloneIV の最小版でも収まってしまう程度である。

プログラム言語は現在、VHDL などのハードウェア記述言語によって記述されている。1 例としてカウンタのプログラムを図 3-19 に示す。C 言語でも記述できるコンパイラも性能が上がってきて、近い将来 C 言語でも開発出来るようになるであろう。そうなるとうハードウェア、ソフトウェアの区別もあいまいになってくる。

まだまだ実用化のために必要な手法はあるが、細かい話

になるのでここでは割愛させていただく。

6. おわりに

自動車の運動は、奥行き方向に向かっていくのだから、周囲を立体として認識することは当たり前のことである。2 つの眼を使った立体画像認識は猿や人間が行なっているやり方で、進化の最後によく獲得した高度な認識法である。我々が開発してきたステレオカメラもそこからヒントを得ている。今後、自動車に限らず、移動ロボットや人間の助けをする様々なロボットの活躍する時代がくるであろう。そういったロボットには人間と同様の物体認識能力や周囲環境認識能力が求められる。そのとき、人間と同じ手法で認識を行なうステレオカメラは最も重要なセンサとして活躍していると確信している。

謝辞 「ぶつからないクルマ」という夢の自動車の実現に向けて、80 年代の終わりからの創成期に共に努力してくださった富士重工業 (株) スバル研究所の塙氏と喜瀬氏、その後 99 年の第 1 号車発売に至るまで、ステレオカメラの性能向上に心血を注いで下さった同社画像グループの方々、そしてステレオカメラを信じて社内外の逆風に耐え、2009 年の本格的発売に結び付けて下さった同画像グループの若い方々⁴⁾にこの場をお借りして心から御礼を申し上げたい。

参考文献

- 1) K. Saneyoshi, K. Hanawa, , "Image Recognition System for Active Drive Assist", Proc. INT. Symp. AVEC, 1992, Sep. Yokohama, Japan, p280 (1992)
- 2) K. Saneyoshi "Drive Assist System Using Stereo Image Recognition", Proc. of IEEE Intelligent Vehicles '96, Tokyo, p230-p235 (1996)
- 3) 十川能之, 塙圭二, "ステレオ画像認識による車両前方監視システム", 自動車技術, vol.56, no.4, pp.34-39, Apr. 20028
- 4) T. Saito, N. Morimitsu, H. Sekiguchi, Y. Takahashi, S. Katahira "Next Generation Driving Assist System Using New Stereo Camera", FISITA (September 2008), 2008
- 5) (ステレオカメラの動向) Myron Z. Brown, Darius Burschka, Gregory D. Hager, "Advances in Computational Stereo", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 25, N0. 8, pp.993-100