

車載向けステレオ画像処理技術

Stereo Vision Technology for Automotive Applications

服部 寛

■ HATTORI Hiroshi

車や歩行者などの道路上の物体検出に有効なステレオ画像処理技術を開発した。この技術は、検出対象の空間的な位置の制約を再帰相関演算による対応探索処理に組み込むことで、高速かつ高精度なステレオ計算を実現した。

この技術を東芝が開発した画像認識用LSI “Visconti_{TM}” に実装し、2.8 GHzのプロセッサに匹敵するステレオ演算能力があることを確認した。更に、大型トラック左折時の巻き込み検知システムに適用し、実画像を用いた実験でこの技術の有効性を示した。

Toshiba has developed a domain-specific, area-based stereo vision method for safe navigation, which is applicable to the detection of cars, pedestrians, and other objects on the road. High-speed and highly precise stereo calculation is achieved by incorporating a spatial restriction in the height direction of road scenes into the recursive computation of correlation measures.

We mounted this system on our Visconti_{TM} image-recognition LSI and realized stereo calculation operation equivalent to that of a processor with a processing speed of 2.8 GHz. We also applied this technology to a side-collision warning system for a heavy-duty truck when making left turns, and verified the validity of the system in an experiment using actual images.

1 まえがき

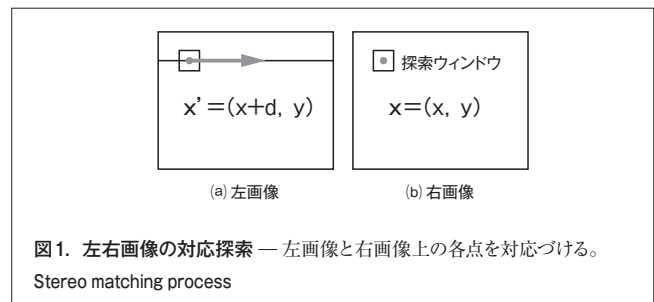
自動車の安全運転を支援するため、車に搭載したカメラで周囲の道路状況を認識する車載画像処理の研究は、自動車メーカーを中心に世界中で盛んに行われている。2台のカメラを用いるステレオ視は、対象物体までの距離計測が可能であり、1台のカメラをベースにしたシステムに比べ、高性能なシステムの構築が可能となる。

東芝は、道路環境の幾何学的な拘束を利用した、高速かつ高精度なステレオ画像処理技術を開発した。ここでは、この技術のアルゴリズムの詳細、当社が開発した画像認識用LSI “Visconti_{TM}^{(1), (2)}” 上への実装方法と性能評価、及び自転車に接近する立体物を識別することで大型トラック左折時の巻き込みを検知するシステムへ適用した事例について述べる。

2 ステレオ視差計算の概要

ステレオ視は、空間中の同一点の左右画像への投影点を対応づけ、三角測量によって対象までの3次元位置情報を得る技術である。この技術を用いれば、対象物までの距離が計測できるだけでなく、復元された視差(奥行き)をベースに対象物の切出しが可能であり、パターン認識や動きの計算など、その後の処理が比較的容易になる。また、道路領域を処理対象から除外できるので、処理コストの削減にもつながる。

ステレオ視においてもっとも困難かつ重要な課題は、左右



画像の対応づけ、すなわち、対応探索問題である。例えば、**図1**に示すように、右画像上の点 $\mathbf{x} = (x, y)$ の対応点を求める際には、周囲に探索ウィンドウを設定し、左画像上の同一走査線を探索し、輝度パターンがもっとも類似した点を求めることになる。左画像上の対応点 $\mathbf{x}' = (x+d, y)$ と点 \mathbf{x} の水平方向の位置のずれ d を視差と呼ぶ。以下では右画像を基準とし、右画像上の各点の左画像上の対応点を求めるものとする。

対応探索処理を**図2**に示す視差の探索空間 V で考える。 V の点 (x, y, d) は右画像上の点 (x, y) と、その左画像上の対応候補点 $(x+d, y)$ から成る1組の対応候補を示す。空間 V を走査し、各点が示す対応候補の類似度を計算して、各点 (x, y) についてピーク検出処理を行えば、基準画像上のすべての点の視差、すなわち視差マップを得る。

探索空間 V を走査し、視差マップを得る手順として次の二つが考えられる。

- (1) 点 (x, y) を固定して視差 d を探索範囲内で動かす、と

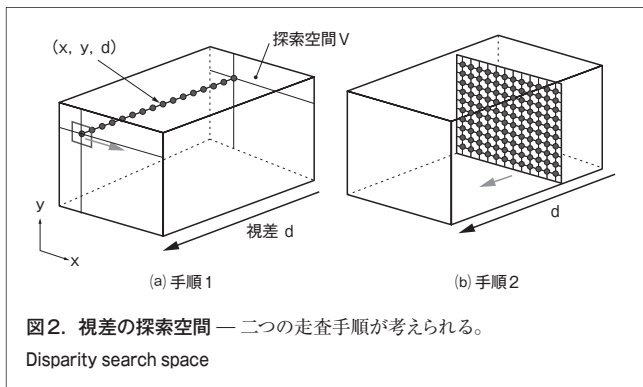


図2. 視差の探索空間 — 二つの走査手順が考えられる。
Disparity search space

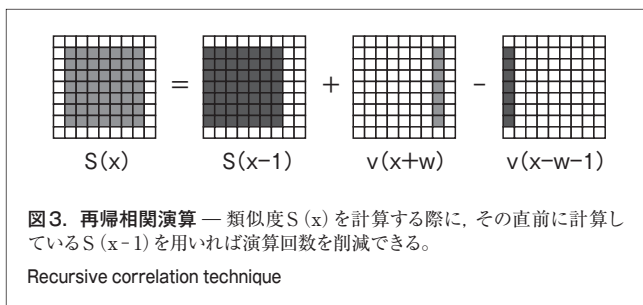


図3. 再帰相関演算 — 類似度 $S(x)$ を計算する際に、その直前に計算している $S(x-1)$ を用いれば演算回数を削減できる。
Recursive correlation technique

いう処理を画像全点に対して行う。

(2) 視差 d を固定して点 (x, y) を動かすという処理を探索範囲内の全視差に対して行う。

探索ウィンドウのサイズが小さい場合には両者の計算コストはほぼ同じとなるが、小さいウィンドウでは正しい対応点を見つけるための十分な輝度変動がないことが多いため、誤対応を招きやすい。したがって、ある程度のウィンドウサイズは必要となるが、こうした実用的なサイズのウィンドウでは後者の方が類似度の計算を再帰的に行うことができるため、少ない演算量で視差マップを得ることができる。これを再帰相関演算法と呼び、概要を図3に示す。

類似度として輝度の差の絶対値のウィンドウ内総和 (SAD) を想定し、各画素は輝度の差の絶対値を示すものとする。点 x の相関値 $S(x)$ は、図3の $S(x)$ の網掛け部 (■) で示した画素 (ウィンドウ内の画素) の総和となる。 $S(x)$ は、1画素左の計算結果 $S(x-1)$ に右端の1列分の総和 $v(x+w)$ を加え、左端1列分の総和 $v(x-w-1)$ を引くことによって、より少ない演算回数で計算することができる。更に、縦1列分の総和自身も同様に、縦方向の重なりを利用して再帰的に計算すればより効率的である。

更に、物体の存在範囲を考慮することで視差の探索範囲を制限し、計算コストを削減する。道路シーンでは、車や人間といった検出対象物は道路面上に存在し、かつ、高さの最大値を持つ。したがって、空間上に高さの上限と下限を表す平面を定義することができる (図4)。

視差の探索空間においても、図5に示すように、同様に対

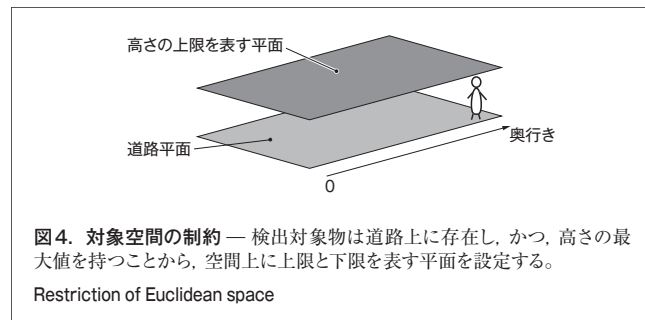


図4. 対象空間の制約 — 検出対象物は道路面上に存在し、かつ、高さの最大値を持つことから、空間上に上限と下限を表す平面を設定する。
Restriction of Euclidean space

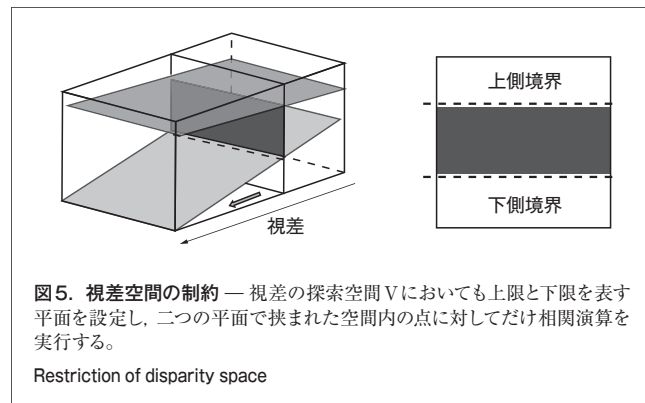


図5. 視差空間の制約 — 視差の探索空間 V においても上限と下限を表す平面を設定し、二つの平面で挟まれた空間内の点に対してだけ相関演算を実行する。
Restriction of disparity space

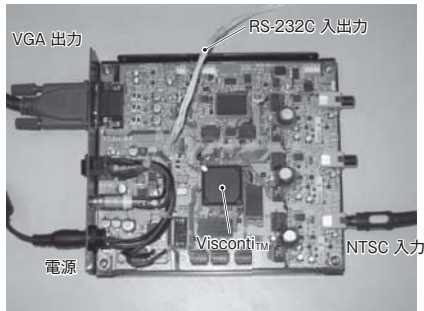
象空間を二つの平面で制限することができる。二つの平面で定義された空間内の点に対してだけ相関演算を行うことで、計算コストが削減されるだけでなく、探索空間が小さくなるので対応精度も向上する。

3 画像認識 LSI への実装方法と性能評価

当社が開発した画像認識用 LSI, Visconti™ の評価ボードを図6に示す。

Visconti™ は、動作温度や消費電力の面で厳しい車載条件をクリアしながら、高度な画像処理が可能な構造となっている。Visconti™ の性能を最大限に利用するため、第一に、対応探索において左右画素の類似度を評価する関数として SAD を用いた。この際、二つのカメラの特性の違いによる左右画像の輝度のバイアスを補正するため、左右画像とも各画素の輝度からウィンドウ内の平均値を引くことによって正規化しておく。元画像の各点の輝度が8ビットであるのに対し、正規化画像は8ビットを超えるが、以後の処理の並列度を維持するため、飽和処理により8ビット内に収める。この Zero-mean SAD は、輝度の差の自乗和 (SSD) や正規化相互相関 (NCC) などほかの関数に比べて計算コストが低く、かつ、ほぼ同等の識別性能を持つ。

第二に、Visconti™ が持つ並列動作可能な四つのモジュールを活用して処理を高速化する。これら四つのモジュールは、一つの座標変換モジュールと三つの処理モジュールから成る。ステレオ画像は通常、探索ライン (エピポーラライン) が走査線と一致するのが理想であるが、実際に二つのカメラをそのよ



VGA : Video Graphic Array (640×480画素)
 NTSC : National Television System Committee (現行テレビ方式)
 RS-232C : コンピュータと周辺装置を接続してデータを交換するための
 インタフェース規格

図6. Visconti_{TM} 評価ボード — 動作温度や消費電力の面で厳しい車載条件をクリアしながら、高度な画像処理ができる。

Prototype processing hardware

うに配置するのは事実上不可能である。したがって、エピソードラインが走査線と一致するように入力画像を補正する。この補正処理はステレオ画像の平行化と呼ばれ、入力画像に対して幾何学変換（2次元の射影変換）を施すことによって実現される。変換パラメータは、カメラキャリブレーションによってあらかじめ求めておくことができ、処理自体は容易であるが、不連続なメモリアクセスが必要となるため、SIMD (Single Instruction Multiple Data)^(注1) 演算による高速化が困難である。

Visconti_{TM}には画像の幾何学変換を行う座標変換モジュールが用意されており、このモジュールにステレオ画像の平行化処理を割り当てる。また、視差の探索空間を $y=const.$ の二つの平面で三つに分割し、それぞれを三つの処理モジュールに割り当てる。この際、処理コストを均等にするため、三つの空間上の点の数ができるだけ均等になるように（図5で示す探索空間の体積ができるだけ同じになるように）分割する。

上記の方式により得られた視差マップの例を図7に示す。画像解像度 320×240 画素、視差範囲 18 画素、探索ウィンドウ 7×7 画素で、処理速度は1フレーム当たり 15.47 ms となった。この処理速度を達成するためには、2.8 GHz 以上のプロセッサ (SIMD 演算による高速化も含む) が必要である。

表1は、視差の探索空間の制約によって計算コストがどの程度軽減されているかを示している。この場合、1347.84 k個の視差空間の点のうち、828.8 k個の点において相関演算を実行している。有効点数Nの削減により、約25%高速化されている。Nの削減の割合に比べて処理時間削減の比率が小さいのは、ステレオ画像の平行化と輝度正規化処理が画像解像度だけに依存し、ステレオ処理全体の計算コストは、Nに対して線形に変化しないためである。視差計算だけの計算時間は表1に示すようにNに比例する。

(注1) マイクロプロセッサにおいて、一つの命令で複数データを扱う処理方式。

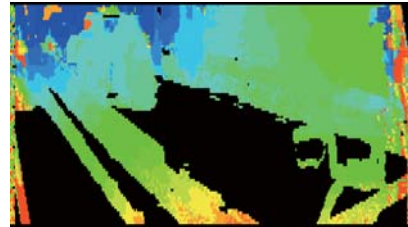


図7. 入力画像と視差マップの例 — 赤いほど近く、青いほど遠いことを示す。黒は、輝度変動が小さいため視差が不定な領域を示す。

Input view and example of disparity map

表1. 処理コストの比較

Comparison of computational costs

項目	全探索	制約付き探索
有効点数N (×10 ³ 個)	1347.84	828.80
視差計算 (ms)	15.12	9.90
全体 (ms)	20.83	15.47

4 巻き込み検知システムへの適用

開発したステレオ法を Visconti_{TM} に実装し、大型トラック左折時の巻き込み検知システムに適用した事例を述べる。図8のように、トラックの左サイドミラーのステイ（支柱）に、1/3インチのCCD（電荷結合素子）カメラを左後方に向けて二つ固定し、昼間の市街地を走行した。カメラ画角は約60°、カメラ

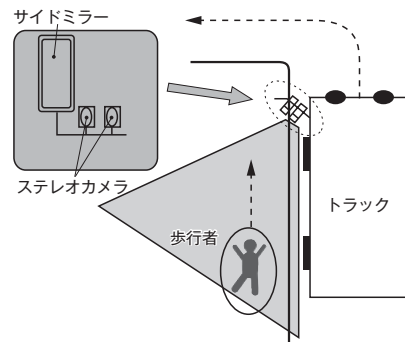


図8. 巻き込み検知システム — サイドミラーに取り付けたステレオカメラを用いて、トラック左折時の巻き込みを検知する。

Collision warning system

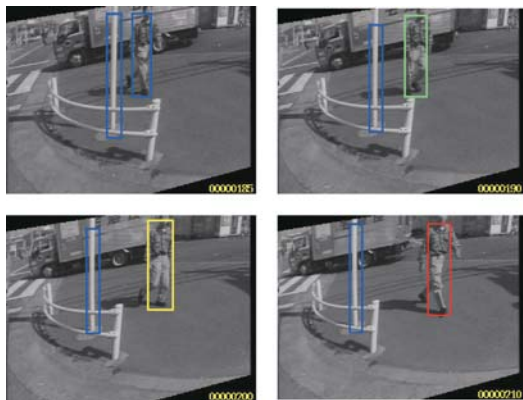


図9.トラックに接近する物体の検出 — 各フレームの矩形は道路平面上の立体物を表し、その色が危険度を表す。歩行者はトラックに接近する動き成分を持つため、時間経過に伴って危険度が増加している(矩形の色が青→赤)。これに対し標識(ボール)はトラックから遠ざかるので、色が青いまま、つまり危険度0と判定されている。

Detection of approaching object

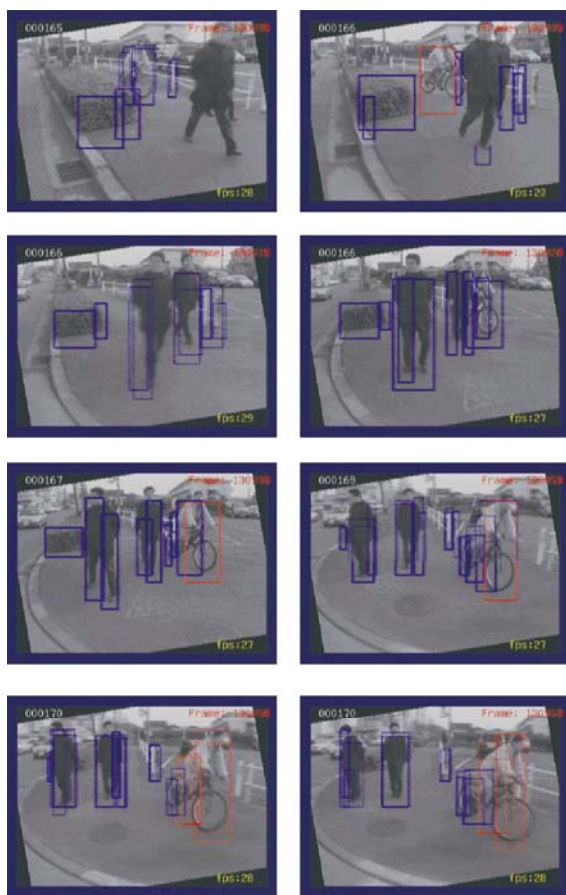


図10. 処理結果の例 — トラックに接近する物体(自転車)を赤い矩形で、遠ざかる物体(ふたりの歩行者)を青い矩形で示している。

Example of output image sequence

間隔は約15 cm, カメラの路面からの高さは約140 cmである。これらの条件下で、日時の異なる映像(天候は晴れ又は曇り)を合計10時間収集して評価対象とした。

このシステムの処理の概要を図9に示す。各フレームで得られる視差マップを用いて、路面に対して高さを持つ立体物を抽出する。更に、動きを計算することで、各物体がトラックに衝突する可能性がどの程度かを予測する。各画像の矩形(くけい)が検出された立体物であり、矩形の色が危険度を表している。歩行者はカメラに接近してくるため、青→緑→黄→赤と危険度が増加しているのに対し、手前の標識はカメラから遠ざかる動き成分を持つため、青く表示されたまま、つまり危険度0となっている。このように、路面に対して高さを持つ物体をステレオ視によって抽出し、その動きによって危険度を評価する。図10の例では、自転車とふたりの歩行者が交差しており、比較的難しいシーンであるが、ステレオ視により検出物体の前後関係がわかり、更に、物体ごとの切出しも比較的容易なため、トラックに接近する自転車だけを抽出することができる。

10時間の評価映像中にトラックに接近してくる物体は153個含まれ、そのうち143個が正しく検出され、接近していない物体を7回誤検出した。すなわち、検出率は94.1%、誤検出率は4.7%であった。

5 あとがき

道路環境における物体検出に有効なステレオ画像処理技術について述べた。今後は、パターン認識の技術を併用することにより、いっそうの性能向上を目指す。

文 献

- (1) Tanabe, J., et al. "Visconti: Multi-VLIW Image Recognition Processor based on Configurable Processor". CICC. San Jose, California, USA, 2003-09, IEEE. 2003, p.185-188.
- (2) Hattori, H., et al. "Dense Stereo Matching in Restricted Disparity Space". IEEE Intelligent Vehicles Symposium. Las Vegas, Nevada, USA, 2005-06, IEEE. 2005, p.118-123.



服部 寛 HATTORI Hiroshi

研究開発センター マルチメディアラボラトリー研究主務。
画像認識の研究・開発に従事。情報処理学会会員。
Multimedia Lab.