

ビデオレートステレオマシンの開発

金 出 武 雄*¹ 蚊 野 浩*² 木 村 茂*³
川 村 英 二*⁴ 吉 田 收 志*⁵ 織 田 和 夫*⁶

Development of a Video-Rate Stereo Machine

Takeo Kanade*¹, Hiroshi Kano*², Shigeru Kimura*³,
Eiji Kawamura*⁴, Atsushi Yoshida*⁵ and Kazuo Oda*⁶

A video-rate stereo machine has been developed at Carnegie Mellon University with the capability of generating a dense range map, aligned with an intensity image, at the video rate. The target performance of the CMU video-rate stereo machine is: 1) multi image input of 6 cameras; 2) high throughput of 30 million point×disparity measurement per second; 3) high frame rate of 30 [frame/sec]; 4) a dense depth map of 240×256; 5) disparity search range of up to 60 [pixel]; 6) high precision of up to 8 [bits] (with interpolation); 7) uncertainty estimation available for each pixel. The stereo machine is based on a multi-baseline stereo method. The basic theory requires some extensions to allow for parallel, low-cost, high-speed machine implementation. The three major ones are: 1) the use of small integers for image data representation; 2) the use of absolute values instead of squares in the SSD computation (SAD instead of SSD); and 3) camera geometry compensation capability for precise camera alignment. The stereo machine successfully generates disparity map of 200×200 with 23 steps of disparity search range in the video rate.

Key Words: Stereo Vision, Real Time, Special Purpose Hardware, Parallel Processing, Depth Image

1. はじめに

ステレオ視による距離計測はさまざまな利点をもつ。第一にステレオ視は受動的計測法であり、光などのエネルギーを発生しない。また適切な撮像系や光学系を用いることで、距離がはなれたシーンに対しても正確で密な距離画像を計測することが可能である。ステレオ視はセンサーフュージョンの一つであり、可視画像と距離画像を同一のカメラ座標上で得ることができる。画像を順次走査するメカニズムがないため、原理的には撮像とほぼ同時に距離画像を得ることができる。走査方式による

レーザレンジファインダのように、測定対象物の動きによる形状歪を生じない。

これまでに開発されたステレオシステムは、これらステレオ視がもつ能力を十分に実現しているとはいえない。とくに、距離画像の計算速度(フレームレート×画像サイズ)と視差の探索範囲(距離計測のダイナミックレンジ)における性能が不十分である[1]~[3]。Teleos社が開発したPRISM 3システム[6]、JPLが開発したDataCubeを用いたシステム[4]、CMUのiWarpでのマルチベースラインステレオ[9]、およびINRIAのシステム[12]などが最も高速なステレオシステムである。しかし、これらのシステムにおいてさえ、入力画像と同程度に密な距離画像をビデオレートで生成することに成功していない。

また、現在のステレオシステムによって得られる距離画像は精度、信頼性ともに不十分である。その理由の一つは、ステレオ画像対の対応点探索問題において、局所演算を用いたアルゴリズムがあいまい性を生ずる点である。これまでに、画像ピラミッドを疎密探索する方法や表面のなめらかさを仮定した大域最適化法などが提案されている[13][14]。しかし、これらの方法は発見的手法になる傾向があり、非常に多くの計算量が必要である。

原稿受付 1996年1月19日

- *¹カーネギメロン大学
- *²三洋電機株式会社
- *³防衛庁
- *⁴株式会社サイヴァース
- *⁵富士電機株式会社
- *⁶株式会社アジア航測
- *¹Carnegie Mellon University
- *²Sanyo Electric Co., Ltd.
- *³Japan Defence Agency
- *⁴CyVERSE Co., Ltd.
- *⁵Fuji Electric Co., Ltd.
- *⁶Asia Air Surrey Co., Ltd.

我々が開発したビデオレイトステレオマシンはカーネギメロン大学で開発し、実験を重ねてきたマルチベースラインステレオ法に基づいている。この方法は基線長の長さや方向が異なるように配置した複数のカメラを用いることに特徴があり、複数のステレオ対から得られる距離情報を一つの評価関数に統合し、信頼性の高い距離計測を可能にした。本マルチベースラインステレオ法は複数のステレオ対に含まれる冗長性を利用して、ハードウェア化に適した単純な構成の並列アルゴリズムである。

本論文では、2章においてマルチベースラインステレオ法を紹介し、3章において実現されたハードウェアを説明する。とくにハードウェア化における工夫がみられるLOGサブシステム、SSADサブシステムについて詳しく説明する。4章でステレオマシンの性能を紹介し、5章を結論にあてる。

2. マルチベースラインステレオ法による距離計測

2.1 マルチベースラインステレオ法

一組のステレオ画像対において、対応点の位置のずれ量を視差と呼ぶ。ステレオ画像対が平行等位化 (rectification) されていると仮定すると、視差 d とシーン中の点までの距離 z の関係は

$$d = \frac{B \cdot F}{\lambda \cdot z} \quad (1)$$

で与えられる (Fig. 1 参照)。ここで λ は撮影された画像の画素サイズ、 B と F はそれぞれ基線長と焦点距離である。ステレオ画像対から直接計測されるのは視差 d であり、この式から視差の推定値の分散が基線長の二乗に反比例することが証明できる [7]。それゆえ基線長を長くすれば計測される距離の精度がよくなる。一方、基線長を長くするとより広い視差の範囲を探索しなければならず、偽対応の可能性が高まる。すなわち基線長の選択に関して、計測の精度とマッチングの正しさの間にトレードオフが存在する。

式(1)から式(2)を得る。

$$\frac{\lambda \cdot d}{B \cdot F} = \frac{1}{z} = \xi \quad (2)$$

この式は、真の視差を基線長で割った値はステレオ対によらず一定であることを示している (F と λ はすべての画像に共通の定数とみなす)。それゆえ、対応点探索において、マッチングの評価関数を距離の逆数に比例した ξ の関数として表現すると、基線長によらず正しい距離に対応した ξ において大きな

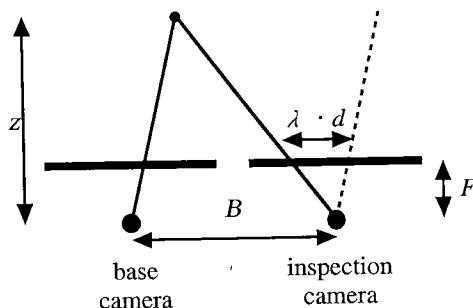


Fig. 1 Rectified stereo camera pair

評価値を示すはずである。複数の基線長に対応した複数の評価関数をたし合わせるにより、計測の精度と信頼性の向上が期待できる。

対応するウィンドウ間での SSD (Sum of Squared Difference) 値はイメージマッチングのための単純かつ有効な評価値である。一組のステレオ画像対において、一方の画像 (基準画像) 内のある一点の周りにウィンドウを定義する。これと同様のウィンドウを他方の画像 (検査画像) にも定義し、エピポーラ線に沿って移動させる。画像 f_0 を共通の基準画像とする $n+1$ 個のステレオ画像 f_k ($k=0 \sim n$) を得たとき、 f_0 のカメラ座標 (i, j) における ξ を変数とする SSD 関数は次のように書くことができる。

$$\begin{aligned} SSD_k(i, j, \xi) &= \sum_{(ii, jj) \in W} SD_k(i+ii, j+jj, \xi) \\ SD_k(i, j, \xi) &= \left(f_k \left(i - t_1 \cdot \frac{B_k \cdot F}{\lambda} \cdot \xi, j - t_2 \cdot \frac{B_k \cdot F}{\lambda} \cdot \xi \right) - f_0(i, j) \right)^2 \end{aligned} \quad (3)$$

ここで、 $t=(t_1, t_2)$ は f_0 のカメラ座標 (i, j) によって決まる f_k 上のエピポーラ線に平行な単位ベクトルであり、 B_k はステレオ画像対 (f_0, f_k) における基線長、 W はウィンドウ領域である。このようにして得られる k 個の SSD 関数をたし合わせることで次のような SSSD 関数を得る。

$$\begin{aligned} SSSD(i, j, \xi) &= \sum_{k=1}^n SSD_k(i, j, \xi) \\ &= \sum_{k=1}^n \left(\sum_{(ii, jj) \in W} SD_k(i+ii, j+jj, \xi) \right) \end{aligned} \quad (4)$$

Fig. 2 における SSD1 から SSD3 は個々のステレオ対における SSD 関数の一例である。これらの SSD 関数はすべて、真の距離に対応する ξ の位置で極小値をとっている。SSD 関数をたし合わせ SSSD 関数を作ったものが Fig. 2 の SSSD である。SSSD 関数はよりはっきりしたあいまい性の少ない最小値をとる。また、最小値付近での曲率が大きくなっており、このことは距離計測の精度が良くなっていることを示すものである。

一方、ステレオ画像対において物体間の隠蔽により死角領域が生じる場合、そのステレオ画像対から得られる SSD 関数は明確な最小値をとらないと予想される。SSSD 関数として統合される複数の SSD 関数にこのような性質がある場合、SSSD 関数の最小値は十分に小さくならず、また最小値近傍における曲率が小さくなる。このような現象は死角領域のみならず、物体の輪郭部であるジャンプエッジで生じる可能性がある

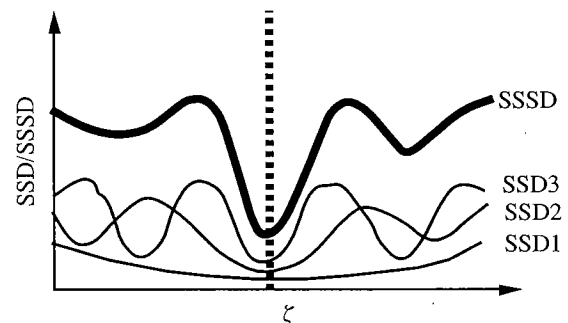


Fig. 2 SSD and SSSD functions

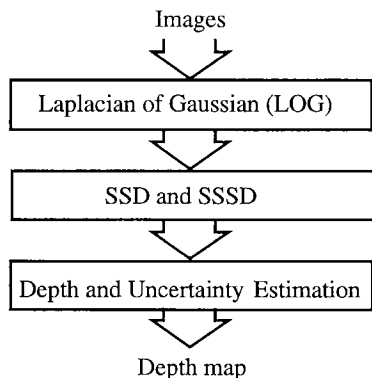


Fig. 3 Outline of stereo method

り、偽対応や計測精度の劣化という問題を生じる。

理論的には、繰り返しパターンのように本質的にあいまいさを含む場合でも、異なった基線長のステレオ対を組み合わせることであいまいさを除去することができる。またより多くのステレオ対を用いることで計測精度を向上させることが可能である[7]。この考え方はどのような基線長の組み合わせに対しても成り立つ。計算過程は完全に局所的であり、探索や最適化、平滑化などは必要ない。このアルゴリズムを行うことは距離の逆数に対するSSD関数を計算し、それをたし合わせてSSSD関数を求め、各画素ごとにSSSD関数の最小値を決定することだけである。

本アルゴリズムをさまざまな条件における室内、室外で実験を重ねた。これらは照明条件などが校正された実験室において15 [mm] から 38 [mm] の最大基線長で1 [m] から 2 [m] の距離に置いた模型を計測したものや、9 [cm] から 15 [cm] の最大基線長で15 [m] から 34 [m] の距離にある建造物などを計測したもの、30 [cm] の最大基線長で60 [m] 程度の距離の屋外を計測したものなどである。良好な対応点探索結果が得られた位置における計測誤差は、実験室内での計測で約0.8%、屋外の計測で数%であった[8]。

2.2 距離計測アルゴリズム

マルチベースラインステレオ法に基づく距離計測アルゴリズムの全体は Fig. 3 のようになる。まず入力画像にLOG (Laplacian of Gaussian) フィルタ処理を施す。このフィルタは入力画像のノイズを平滑化するとともに特徴を強調するだけでなく、個々のカメラの特性やカメラ位置の違いによる画像の輝度変化を補正する働きがある。第二段階は全ステレオ対のSSD関数を計算したのち、それらをたし合わせることでSSSD関数を得る。この段階で、距離の逆数に対するSSD関数を求めるために画像の補間処理も実行される。最終段階ではSSSD関数の最小値を見つけることにより ζ を決定する。最小値付近におけるSSSD関数の曲率を解析することで計測値の信頼性を求めたり、サブピクセル精度での ζ を求めることも可能である。

3. ビデオレートステレオマシンの設計

3.1 ステレオマシンのアーキテクチャ

我々はマルチベースラインステレオ法、およびさまざまな条

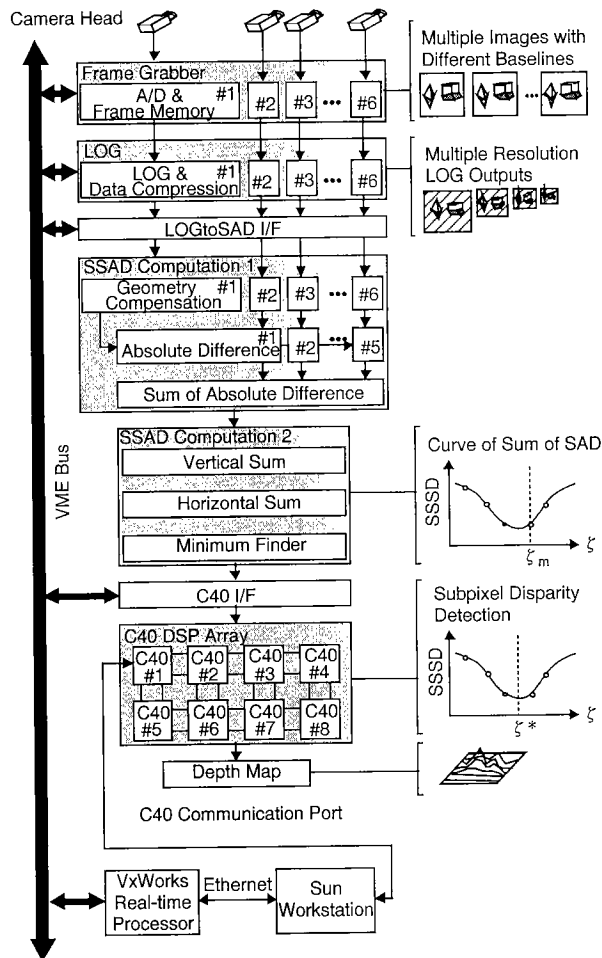


Fig. 4 Architecture of stereo machine

件での実験結果を基にしてビデオレートステレオマシンの設計を行った。本方式の特徴の一つは完全に局所的な計算アルゴリズムにある。距離の逆数に対するSSSD関数の計算に必要なのは、各ピクセル位置における莫大な局所ウィンドウ演算のみであり、大域的な最適化や比較はまったく必要ない。この特徴がステレオマシンを実現するために最も重要な点である。

しかしながらこのアルゴリズムを直接ハードウェア化することは、回路規模、複数カメラの正確な校正の必要性などの点でむだが多い。そこで、高速かつ安価な並列ハードウェアを実現するために、基本アルゴリズムに以下の3点の拡張を行った。

1) 画像データ表現にビット数の少ない整数を用いた。2) SSD計算において二乗の代わりに絶対値を用いた (SSDではなくSAD (Sum of Absolute Difference))。3) 入力画像の幾何補正回路を導入した。

Fig. 4 に試作システムのブロック図を示す。システムは以下の5つの重要なサブシステムで構成される。1) 複数のカメラで構成されるステレオカメラヘッド部。2) 複数画像を同時に取り込むことができる画像取り込み部。3) LOG フィルタ部。4) 並列SSAD計算部。5) SSAD関数のサブピクセル計算と計測距離の信頼度計算を行う並列DSP部。ビデオレートステレオマシンは、これらのサブシステムが入力画像のデータ列に

対してパイプライン的に処理を進め、並列 DSP 部に 30 [msec] ごとに 1 枚の距離画像 (視差画像) を生成する。

いくつかのサブシステムは VME バスに接続され、VxWorks 実時間処理プロセッサで制御される。このプロセッサはステレオマシンのさまざまなレジスタやメモリにアクセスする。画像読み取り部の画像メモリを読み書きすることや LOG フィルタの係数の設定、LOG 処理された画像の読み書き、SSAD 計算におけるウィンドウの大きさの設定、幾何補正回路の変換特性の設定などを行う。このように、ステレオマシンを直接制御するプログラムを VxWorks に実現するとともに、ステレオマシンのユーザインタフェースを Sun ワークステーションに実現した。利用者はグラフィカルなインタフェースを通じてステレオマシンの機能を活用することができる。並列 DSP 部に生成される視差画像やカメラから入力された濃淡画像をワークステーションに転送することも可能である。

3.2 LOG サブシステム

本サブシステムは並列に入力される 6 系統の入力画像に対応して 6 系統の LOG フィルタ回路を実現した。画像読み取り部から濃淡画像を入力し、LOG 画像を SSAD 計算部に出力する。Fig. 5 に 1 系統分の LOG 回路のブロック図を示す。1 系統あたり 4 個の空間フィルタ LSI を用いた。各 LSI は 49 個の係数をプログラムすることにより、任意の 7×7 の空間フィルタを実現できる。たとえば、最初の 3 個のフィルタをガウシアン、最後のフィルタをラプラシアンに設定し、各出力を次段の入力とすることで非常に大きなウィンドウサイズの LOG フィルタを実現できる。このようにフィルタ処理を直列化することにより LOG フィルタの最大サイズは 25×25 になる。

LOG サブシステムは多重解像度の LOG 画像を生成する能

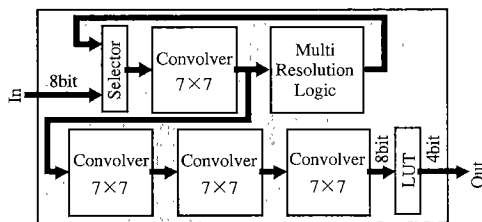


Fig. 5 Block diagram of LOG subsystem

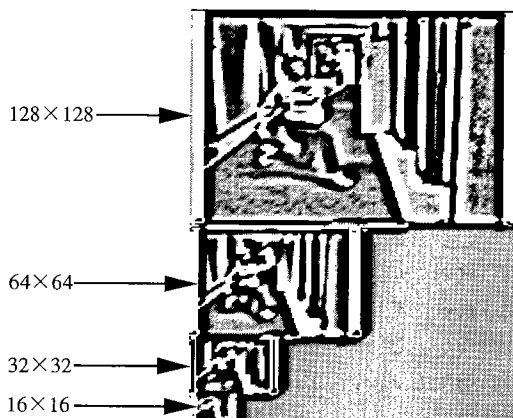


Fig. 6 Multi resolution output of LOG

力も備えている。これは Burt の方法[11]で画像をぼかせながら間引くことにより、複数の解像度の画像ピラミッドを作るものである。Fig. 6 に多重解像度の LOG 画像の例を示す。この例では最も大きい画像のサイズが 128×128、その次が 64×64、最も小さいものが 16×16 となっている。

LOG フィルタ処理されたデータは適当なシフト演算により 8 ビットを取り出す。さらに LOG サブシステムに続く SSAD 計算部の回路規模を小さくすることを目的として、8 ビットから 4 ビットにデータ圧縮される。ソフトウェアで行った予備実験により、8 ビット画像データから得られる視差画像と、8 ビットデータをヒストグラム平均化処理したのち 4 ビットに圧縮したデータから得られる視差画像にほとんど差がないことを確認した。

実際のハードウェアによる実現では、ヒストグラム平均化処理を近似するために画素値の分布を非線形関数により変換、圧縮する方法を採用した。Fig. 7 左上に入力画像の一例を示す。この画像を LOG フィルタ処理した画像を Fig. 7 右上に示す。LOG 処理した画像の画素値のヒストグラムを Fig. 8 (a) に示す。LOG 画像のヒストグラムはこのようにゼロ付近に鋭いピ

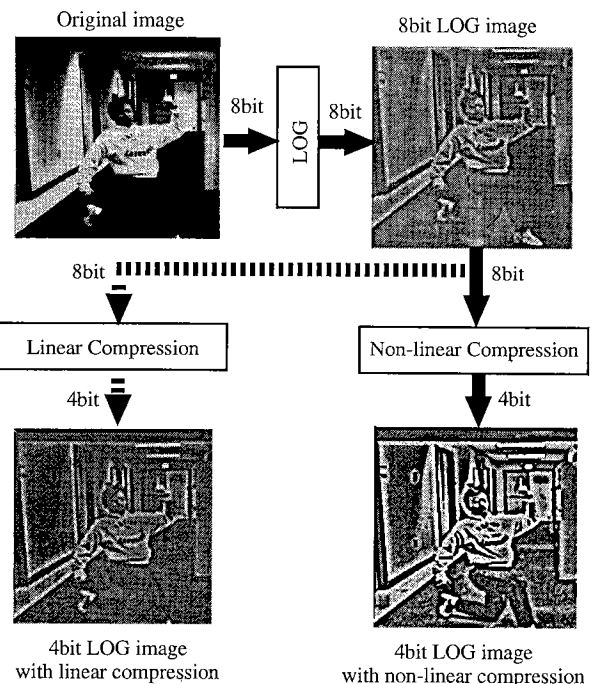
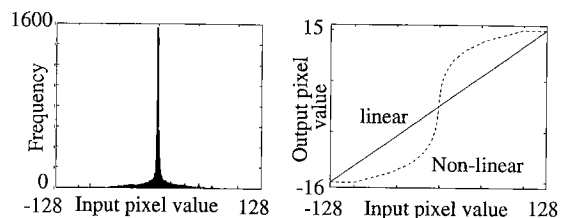


Fig. 7 Data compression of LOG image



(a) Histogram of 8bit LOG image (b) linear and non-linear data compression

Fig. 8 Histogram of LOG image and data compression

ークを持つ単峰性の分布になる。8ビット LOG 画像から4ビット LOG 画像への線形関数および非線形関数の特性の一例を Fig. 8(b) に示す。このような非線形圧縮を用いることでゼロ近傍の画素値は細かく分割され、ゼロから離れるにしたがい画素値は粗く分割される。Fig. 7 左下に線形圧縮した LOG 画像、Fig. 7 右下に非線形圧縮した LOG 画像を示す。線形圧縮した LOG 画像は平均的な濃淡は保存されるが細かい特徴が失われるのに対し、非線形圧縮した LOG 画像では多くの特徴が保存され、信号レベルが相対的に拡大される。

3.3 SSAD サブシステム

SSAD サブシステムは次の3つの機能をもっている。すなわち、ステレオ対ごとに正しい対応点候補を抽出する幾何補正部。抽出された対応点候補の差の絶対値を全ステレオ対、およびウィンドウ内についてたし合わせる SSAD 演算部。SSAD 関数から最小値を抜き出す最小値検出部である。これらの設計において重要なポイントは入力画像の幾何補正回路の導入と、SSAD 計算に含まれる2つの加算処理の入れ替え、およびパイプライン処理に適したウィンドウ演算部の実現である。

3.3.1 幾何補正部

入力画像の幾何補正は SSAD 演算の前処理として実行され、あらかじめ求めておいたステレオカメラヘッドの幾何学的特性を、ステレオマシンの立ち上げ時に幾何補正回路部の Lookup テーブルにダウンロードすることで機能する。このモジュールを導入することにより、カメラ特性のばらつきやカメラヘッドの位置決め精度を補正することが可能になっただけでなく、輻輳をつけたステレオなど、カメラを任意に配置することが可能になった。

f_0 を基準画像とする $n+1$ 個のステレオ画像 $f_0 \sim f_n$ が得られたとき、幾何補正後の対応点候補の差の絶対値 AD_k ($k=1 \sim n$) は以下ようになる。

$$AD_k(i, j, \xi) = |f_k(I_k(i, j, \xi), J_k(i, j, \xi)) - f_0(I_0(i, j), J_0(i, j))| \quad (5)$$

ここで (i, j) は幾何補正前の画像の座標値であり、基準画像からは $(I_0(i, j), J_0(i, j))$ なる座標が対応点として抽出され、検査画像からは $(I_k(i, j, \xi), J_k(i, j, \xi))$ なる座標が対応点として抽出されることを示している。基準画像に対する補正は主にレンズ歪のためであり、検査画像に対する補正はそれに加えてエピポーラ線に沿った対応点の抽出を含むものである。これらの座標値は一般に、画像取り込み時にサンプルされる格子点上にないため、近傍4画素を用いた線形補間を行う。 I_k, J_k などの関数は従来から知られている強カメラ校正[10][15]や、新しい手法である弱カメラ校正[16]を用いて得ることができる。

3.3.2 SSAD 演算部

SSAD 関数の計算を定式化すると式(6)のように表現できる。

$$SSAD(i, j, \xi) = \sum_{k=1}^n \left(\sum_{(ii, jj) \in W} AD_k(i+ii, j+jj, \xi) \right) \quad (6)$$

この計算式には画素値の差の絶対値に関して、二種類の加算処理(ウィンドウ内の加算処理 $\sum_{(ii, jj) \in W}$ と複数ステレオ対での加算処理 $\sum_{k=1}^n$) が含まれる。前者は二次元的、後者は一次元的で

あることから、式(6)の加算処理の順序を入れ換え、

$$SSAD(i, j, \xi) = \sum_{(ii, jj) \in W} \left(\sum_{k=1}^n AD_k(i+ii, j+jj, \xi) \right) \quad (7)$$

とすると、ウィンドウ内加算の演算器を一個で済ませることができる。

ステレオ対加算された SAD 関数をウィンドウ内加算することで SSAD 関数を求める。このウィンドウ内加算の計算は簡単であるが、単純な方法を用いると計算量が膨大になる。本ステレオマシンは近傍画素の SSAD 関数の計算においてかなりの重複計算を含むことを利用することで、計算量の削減に成功している。

$D(i, j, \xi)$ を次のように定義する。

$$D(i, j, \xi) = \sum_{k=1}^n AD_k(i, j, \xi) \quad (8)$$

すると、SSAD を次のように書くことができる。

$$SSAD(i, j, \xi) = \sum_{(ii, jj) \in W} D(i+ii, j+jj, \xi) \quad (9)$$

$-m \leq ii \leq m, -m \leq jj \leq m$ ($m \geq 0$) と仮定すると、式(9)は次のように書くことができる。

$$SSAD(i, j, \xi) = \sum_{ii=-m}^m \left(\sum_{jj=-m}^m D(i+ii, j+jj, \xi) \right) \quad (10)$$

$VSUM(i, j, \xi)$ を次のように定義すると、

$$VSUM(i, j, \xi) = \sum_{jj=-m}^m (D(i, j+jj, \xi)) \quad (11)$$

式(10)は最終的に次のように再帰的な形で書ける。

$$SSAD(i, j, \xi) = SSAD(i-1, j, \xi) - VSUM(i-m-1, \xi) + VSUM(i+m, j, \xi) \quad (12)$$

同様に $VSUM(i, j, \xi)$ も次のように再帰的に書ける。

$$VSUM(i, j, \xi) = VSUM(i, j-1, \xi) - D(i, j-m-1, \xi) + D(i, j+m, \xi) \quad (13)$$

すなわち、一つのウィンドウ加算を垂直加算と水平加算に分け、垂直加算の結果得られる部分和などをローカルメモリに蓄え、再帰的に利用することで SSAD 関数を得る。これらの演算をパイプライン的に実行することで、ウィンドウ単位でのマッチング処理をビデオレートで実行することを可能にした。

3.3.3 最小値検出部

SSAD サブシステムの最終段に位置する最小値検出モジュールは SSAD 関数の最小値を検出する。また、このモジュールは SSAD 関数から最小値近傍の点数を抽出することも可能である。これらの情報は後段の並列 DSP 部に転送され、視差のサブピクセル補間処理や計算される視差情報の信頼度計算に使われる。

4. ビデオレートステレオマシンの性能

4.1 ハードウェアおよび仕様

ビデオレートステレオマシンは市販のデバイスのみを用いて開発した。PLD, Complex-PLD, ROM, RAM などのプログラマブルデバイスおよび、空間フィルタ用 LSI, ALU, パイプラインレジスタなどが主要な部品である。ビデオカメラ、TI 社の TMS 320 C 40 を用いた並列 DSP 部、実時間プロセッサボードを除いてすべて CMU で設計、開発を行った。現在、200 × 200 の画像サイズで 23 段階の視差探索範囲に対して、毎秒

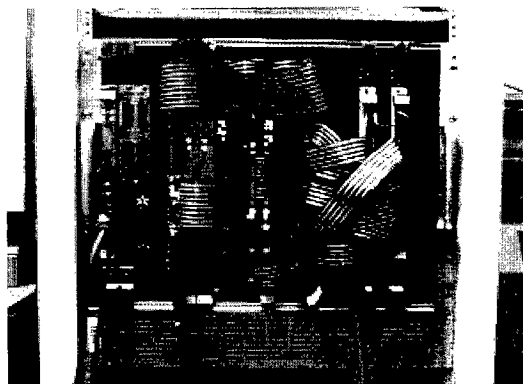
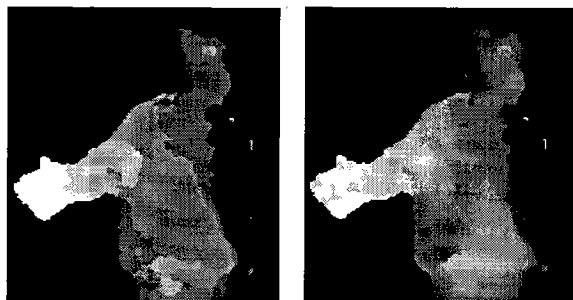


Fig. 9 CMU video-rate stereo machine



(a) Disparity map with 30 disparity range

(b) Disparity map with disparity interpolation

Fig. 10 Example scenes demonstrating disparity interpolation

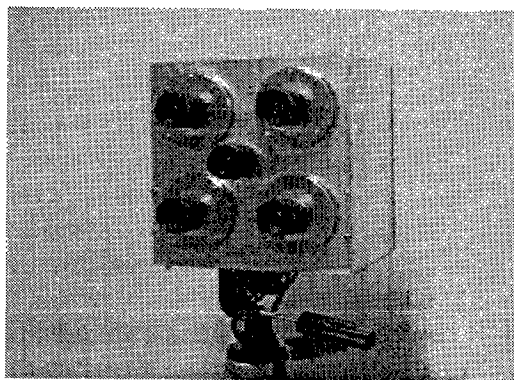


Fig. 11 Camera head of stereo machine

30 フレームの速度で動作する。ハードウェアの外観を Fig. 9 に、主な仕様を Table 1 に示す。

4.2 サブピクセル精度での視差検出

DSP 並列演算部にサブピクセル精度での視差検出を行うソフトウェアを実現した。SSAD サブシステムで行われる視差の探索は最大 60 段階であり、離散的な視差に対応した離散的な SSAD 関数が得られる。この SSAD 関数の最小値を含む近傍 3 画素を並列 DSP 部に転送する。SSAD 関数の最小値近傍を 2 次関数で近似し、その極小値を求めることで 8 ビット精度 (256 段階) の視差を得ることが可能になる。

Fig. 10 (a) は 30 段階の視差で表現した視差画像、(b) は視差の補間計算を行った視差画像である。(b) の画像は (a) に比

Table 1 Performance of stereo machine

Number of Camera	2 to 6
Processing time/pixel	33 [ns] × (disparity range + 2)
Frame rate	up to 30 [frame/sec]
Depth image size	up to 256 × 240
Disparity search range	up to 60 [pixels]

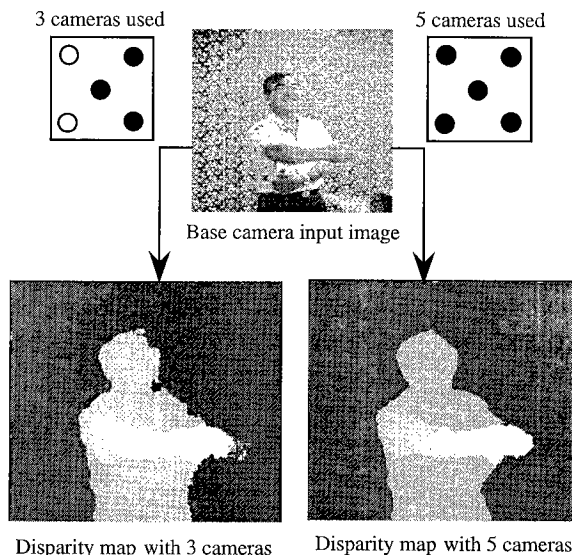


Fig. 12 Example scenes of disparity map with 3 and 5 cameras

較して、スムーズな濃淡変化が得られている。視差の補間計算を含んだステレオマシンの動作速度は 200 × 200 の画像サイズ、23 段階の視差探索範囲に対して毎秒 15 フレームである。

4.3 カメラヘッド

Fig. 11 に示すように中央に基準カメラを配置し、その周りの 4 台のカメラとステレオ対を構成する対称型配置のカメラヘッドを開発した。このような対称型のカメラ配置により、遮蔽領域の影響による偽対応を解決できる場合がある。Fig. 12 は 5 台のカメラ全てを用いる場合と、右半分の 3 台のカメラを用いる場合で、得られる視差画像を比較したものである。5 台を用いた場合、ほぼすべての領域で視差が正しく求められているのに対し、3 台の場合には遮蔽領域が発生するエッジ部でノイズ的な視差が計測されていることが確認できる。

5. 結 論

本論文ではマルチベースラインステレオ理論に基づく CMU ビデオレートステレオマシンを紹介した。特に、ステレオマシンの最も重要なサブシステムである LOG サブシステムおよび SSAD サブシステムについて、高速かつ安価なハードウェアを実現するための工夫について詳しく述べた。また、ステレオマシンの性能、視差のサブピクセル補間、対称型カメラ配置の効果について実例を示した。

今後の課題として死角による計測誤差、および物体の輪郭部における計測誤差が問題である。前者は SAD 関数をたし込んで SSAD 関数を得る場合に、十分に最小値が小さい SAD 関数のみから SSAD 関数をつくるなどで対応可能である。後者

はイメージマッチングのためのウィンドウサイズを適応的に設定する処理, すなわちジャンプエッジなど明確な特徴が存在する領域ではウィンドウサイズを小さく設定し, 輝度変化が少ない領域ではウィンドウサイズを大きく設定するなどの処理で対応可能である. ステレオマシンの精度を向上させるためには, これらを含めてステレオ計測そのもののアルゴリズムを改良する必要がある.

ステレオマシンは自律移動ロボットのための視覚センサとして利用することが可能だけでなく, これまでにない新しい応用分野に利用することが可能である. CMU では, 動きをとまなう実シーンのリアルタイム 3次元モデリングや, ステレオマシンから得られる距離情報を使った 3次元画像合成などの応用分野の研究開発を進めている.

謝 辞

本ステレオマシンの開発は CMU に在籍する多くの人の協力で見事なものであり, 特に画像取り込み部などの開発に協力していただいた Larry Lyle に感謝する.

参 考 文 献

- [1] Nicholas Ayache and Francis Lustman: "Trinocular stereo vision for robotics," Technical Report 1086, INRIA, 1989.
- [2] Pascal Fua: "A parallel stereo algorithm that produces dense depth maps and preserve image features," Technical Report 1369, INRIA, 1991.
- [3] Ali E. Kayaalp and James L. Eckman: "A pipeline architecture for near real-time stereo range detection," Technical Report GDLS-AI-TR-88-1, General Dynamics AI Lab., 1988.
- [4] L. H. Matthies: "Stereo vision for planetary rovers: stochastic modeling to near real time implementation," International Journal of Computer Vision, 8(1), pp. 71-91, 1992.
- [5] T. Nakahara and T. Kanade: "Experiments in multiple-baseline stereo," Technical Report, Carnegie Mellon University, Computer Science Department, 1992.
- [6] H. K. Nishihara: "Real-time implementation of a sign-correlation algorithm for image-matching," TR-90-02, Teleos Research, 1990.
- [7] 奥富正敏, 金出武雄: "複数の基線長を利用したステレオマッチング," 電子情報通信学会誌, vol. J75-D-II, no. 8, pp. 1317-1327, 1992.
- [8] Masatoshi Okutomi, Takeo Kanade and T. Nakahara: "A multiple-baseline stereo method," In Proc. of DARPA Image Understanding Workshop, pp. 406-426, DARPA, 1992.
- [9] J. Webb: "Implementation and performance of fast parallel multibaseline stereo vision," In Proc. of DARPA Image Understanding Workshop, pp. 1005-1012, DARPA, 1993.
- [10] Roger Y. Tsai: "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV camera and lenses," IEEE Journal of Robotics and Automation, vol. RA-3, no. 4, 1987.
- [11] P. J. Burt and E. H. Adelson: "The laplacian pyramid as a compact image code," IEEE Trans. on Communication, vol. COM-31, no. 4, pp. 532-540.
- [12] Oliver Faugeras, et al.: "Real time correlation based stereo: algorithm and application," Research Report 2013, INRIA, 1993.
- [13] W. E. Grimson: "Computational experiments with a feature based stereo algorithm," IEEE Trans. on Pattern Anal. and Mach. Intell., 7, 1 pp. 17-34, 1985.
- [14] D. Marr and T. Poggio: "A theory of human stereo vision," In Proc. of Roy. Soc. London, pp. 301-328, 1979.
- [15] S. Kimura, T. Kanade, H. Kano, A. Yoshida, E. Kawamura and K. Oda: "CMU video-rate stereo machine," In Proc. of Mobile Mapping Symposium, 1995.
- [16] Olivier Faugeras: "What can be seen in three dimensions with an uncalibrate stereo rig?," In Computer Vision-ECCV '92, LNCS-Series vol. 588, Springer-Verlag, pp. 563-578, 1992.



金出武雄 (Takeo Kanade)

1945年10月24日生。1973年京都大学電子工学科博士課程修了。同年同情報工学科助手。1976年同助教授。1980年CMUに移る。1985年同教授。1992年ロボティクス研究所所長。1994年U. A. and Helen Whitaker教授。1995年Joseph Engleberger賞受賞, 米国国家研究評議会航空宇宙エンジニアリング委員, NASA先端技術諮問委員, カナダ先端研究所諮問委員。IEEE, AAAIの各Fellow。(日本ロボット学会正会員)



木村 茂 (Shigeru Kimura)

1955年11月4日生。1978年名古屋工業大学工学部電気学科電気卒業。1980年名古屋大学修士課程修了。同年防衛庁技術研究本部入庁(防衛庁技官)。1991年工学博士(慶大)。1992年~1994年CMUロボティクス研究所客員研究員。戦闘機搭載用コンピュータおよび実時間動画像処理システムの研究開発に従事。現在, 2研1部情報第1研究室主任研究員。



吉田收志 (Atsushi Yoshida)

1963年1月28日生。1985年名古屋大学工学部電気学科卒業。同年富士電機株式会社入社。その後, 株式会社富士電機総合研究所(現在, 富士ファコム制御株式会社)に出向。1993年~1995年CMUロボティクス研究所客員研究員。入社以来, 文字認識, 文字読取装置の開発に従事し, 現在はマシンビジョンの研究に従事。電子情報通信学会会員。



蚊野 浩 (Hiroshi Kano)

1960年3月5日生。1984年京都大学大学院情報工学専攻修士課程修了。同年三洋電機(株)中央研究所入社。カラー記録装置, カラー画像処理技術の研究開発に従事。1993年~1995年CMUロボティクス研究所客員研究員。電子情報通信学会, 情報処理学会の各会員。

(日本ロボット学会正会員)



川村英二 (Eiji Kawamura)

1962年5月10日生。1986年東京工業大学理学部数学科卒業。同年日本電気マイクロコンピュータ事業部入社。LSIの設計, CAD開発環境の構築に従事。1992年~1994年CMUロボティクス研究所客員研究員。1995年(株)サイヴァース設立, 代表取締役就任。



織田和夫 (Kazuo Oda)

1965年1月13日生。1988年京都大学理学部地球物理学科卒業。同年アジア航測株式会社入社, 現在に至る。1994年9月よりCMUロボティクス研究所客員研究員。画像による距離・形状計測, 写真測量, 地理情報システムの研究開発に従事。

American Society of Photogrammetry and Remote Sensingの会員。