

航空宇宙技術研究所報告

TECHNICAL REPORT OF NATIONAL AEROSPACE LABORATORY

TR-1358

テンプレート・マッチング法による実時間動き検出

五味 広美

1998年7月

航空宇宙技術研究所
NATIONAL AEROSPACE LABORATORY

目次

ABSTRACT	1
1. はじめに	1
1.1. 構成	2
2. テンプレート・マッチング法	2
2.1. 動き検出モジュールの位置づけ	2
2.2. 動き検出アルゴリズムの前提条件	3
2.3. テンプレート・マッチング法	3
(1) テンプレート・マッチング法の使い方	3
(2) 画像特性	3
(3) 一致度の評価基準	4
(4) テンプレートの大きさや形状	4
(5) 探索範囲と画像更新レイト	4
(6) 探索パラメータ	4
(7) 処理時間	5
2.4. 実時間処理の重要性	5
3. 実験装置	6
3.1. 実験装置	6
(1) 模型と移動方法	6
(2) ビデオカメラと照明	6
(3) 画像処理部	6
3.2. テンプレート・マッチング法の処理速度	7
(1) TRV の処理速度	7
(2) CTRV の処理速度	7
4. 動きの信頼性	7
4.1. 高信頼性の動きを獲得する方法の比較検討	7
4.2. 選んだ方法：時間的変化制約による高信頼性動きの抽出アルゴリズム	9
(1) プログラム開発	9
4.3. 実験	9
(1) 実験の制限	9
(2) フィルターとしての特性	9
(3) 加減速実験	10
(4) 追跡実験	11
(5) 空間的変化制約による実験	11
5. 動きの精度	12
5.1. 補間の理論	12
(1) 評価値変化のモデル	12
5.2. 予備実験	13
(1) 予備実験方法	13
(2) 予備実験結果	13
5.3. 実験	14
(1) テンプレート画素数の影響	15

(2) 移動速度の影響	15
6. 考察	16
6.1. 動き検出モジュールのハードとしてのTRVの評価	16
(1) 処理能力	16
(2) 使い勝手	17
6.2. 今後の課題：動き情報の利用アプリケーション	18
7. まとめ	18
参考文献	19
付録A. テンプレート・マッチング法の専用ボード	20
付録B. パルスモータ回転速度の安定性	21

テンプレート・マッチング法による実時間動き検出*

五味 広美*¹

Optical flow computation at video rates by template matching

Hiromi GOMI

ABSTRACT

Template matching method was investigated and evaluated by experiments. Three hundred optical flows were calculated at video rates with motion estimation processors and the randomly fluctuating optical flows were excluded as unreliable ones on the assumption that optical flows vary smoothly with time due to mass. The algorithm used bit data handling and succeeded in the segmentation and tracking of a miniature model on a uniform background. The accuracy of motion estimation was improved by 3-point-linear interpolation. The motion estimation error was about 1/10 picture element (p.e.) with a template size of 16 × 16 p.e. when the velocities of miniatures were less than 5 p.e. / frame.

Keywords : visual tracking, optical flow, velocity detection, template matching, reliability, sub-pixel accuracy, real-time vision system, computer vision

概 要

画像上の動きを実時間で求める方法として、テンプレート・マッチング法を模型実験により得られた画像を使い評価した。急激に変化する動きは信頼性が低いという力学的なモデルに基づく動きの信頼性評価アルゴリズムを提案し、ビデオレートで300点の動きの信頼性を判定した。このアルゴリズムにより抽出した高い信頼性の動きを使い、移動対象を背景から切り出すことが出来た。また、3点線形補間を併用することにより動きの精度を2ビット向上した。5画素/フレームの動きまでは、動きによる補間精度の低下は見られなかった。

1. はじめに

工学的に作られたシステムは想定された条件の下では機能するが、それ以外の条件での動作が保証されていないので、航空機の位置と姿勢に関する情報を得る系を多重にし、高い信頼性を達成するのが基本である。現在では、GPS、ILS等、電波を使った航法を組み合わせることにより、着陸の最終段階を除き、大型の旅客機は自動化飛行が可能になっている。

パイロットは、想定外の緊急事態に於いても、各種の計器からの情報と視覚や体感等によって直接得た情報を総合的に判断する、最終的な飛行管理者として存在する。また、パイロットは、これらの電子機器によらないで、窓

から外を眼で直接見て飛行機の位置や飛行方向を確認し、体感により航空機の運動を経験的に推定する1つの独立した系を構成することが出来る。

GPSやILSは数個の信頼性の高いシステム/機器を基準として高い精度を達成しているが、外的な基準の動作と電波の伝播が正常であるものとして(確認が難しい)、位置と姿勢を計算により間接的に求めている。それに対して、パイロットは外界の多数の特徴点/モノの全体的な見え方(視覚情報)から、位置と姿勢を経験的、直接的に求めることにより、信頼性を得ている。したがって、電波航法の結果を視覚により確認する組み合わせは多重系として適している。

パイロットのこれらの機能を機械化・自動化で置き換えることは現在の工学のレベルでは不可能であるが、中長期的には、パイロットの視覚による確認を代行/支援することには工学的実現性がある。ヒトの視覚メカニズ

* 1998年2月23日受付(received 23 February, 1998)

*¹ 数理解析部(Computational Sciences Division)

ムはまだ解明されていない部分が多く、計算機による画像処理に基づく工学的な視覚機能はヒトの視覚機能に及ばないが、計算機視覚の優れた点は画像(視野)の全面探査を行ない見落としが無いことである。安全にとり見落としが無いことは重要である。また、汎用の視覚機能が工学的に実現されていない現在、用途が限定的なパイロットの視覚機能を中間目標として設定することは、計算機視覚にとっても必要なことである。

視覚の最も重要な機能は移動である。動物だけが視覚機能を持っていて、移動出来ない植物が視覚機能を持っていないことから、視覚は移動する為に必要な感覚と考えられる。動物は移動することにより、敵から逃げる、獲物を追いかける、食べ物を探す、快適な/安全な場所を獲得する等、より良い生存条件を得ることが出来、植物より優位な立場にいる。視覚は移動の為に有効な遠隔センサーであっただけではなく、視覚情報処理を行なう為に大脳皮質が発達し、ヒトは知的にも進化し、今日の文明を築いた。したがって、視覚機能を工学的に実現することを目指す計算機視覚としても、移動の為に計算機視覚の研究を行なうべきと考え、さらに、現在の計算機視覚の限界を乗り越えることを期待して、動き情報検出アルゴリズムの研究を進めている。

空港を移動する航空機は背景と動きが異なるので、動き情報を使うことにより航空機を背景から分離することが可能になり、その移動を追跡することが出来る。さらに、画面上の動き情報からビデオカメラと対象間の相対的な3次元運動と外界形状を抽出することが原理的に可能である。例えば、航空機に搭載したビデオカメラで外界を撮影した画像を処理することにより、航空機の運動パラメータ、位置の確認、障害物の検出等が出来る。

この報告では、動き検出アルゴリズムの中で計算機視覚に適すると考えられるテンプレート・マッチング法に関して検討を進めた。それにより求めた動き情報の信頼性を判定する方法を提案し、精度を向上する方法を実験により調べた。

1. 1. 構成

第2章では、動き検出アルゴリズムとして、テンプレート・マッチング法の特性と実時間処理の重要性について検討した。

第3章では、使用した実験装置について述べ、画像処理部のテンプレート・マッチング法の処理速度について測定した。

第4章では、テンプレート・マッチング法で得られた動きの信頼性を確認する新しい方法を提案し、第5章では、テンプレート・マッチング法で得られた動きの精度を補間の併用により向上する方法を実験で確認した。

第6章では、以上の研究成果を含め、実験に使用した処理装置を計算機視覚システムの動き検出モジュールとして使う立場から評価し、動き検出モジュールの出力を利用する今後の研究課題を考察した。

第7章は全体のまとめである。

2. テンプレート・マッチング法

2. 1. 動き検出モジュールの位置づけ

ヒトの視覚では、網膜像を入力として、明度、色、動き、両眼視差等を抽出するモジュールが並列に配置され、それらの結果を使う外界の奥行きや対象の運動等を推定する高次のモジュールが階層的に配置されていると考えられている。計算機視覚システムでも、ヒトの視覚と同様に、ビデオ画像の処理モジュールを並列的かつ階層的に結合して視覚システムを構成するのが合理的である(脚注)。

現在、研究を進めている視覚システムでは、対象を撮像するビデオカメラモジュール、ビデオ画像から画像上の動き(2次元)を検出するモジュール、画像上の動きから運動パラメータ(3次元)を抽出するモジュールの直列的な3階層を考えている。ヒトの視覚系でも動きと運動とは2つの独立した問題として処理されていることが物理心理学実験により示されている(文献1)。

画像処理過程を入力と出力とが明確に定義された、交換可能なモジュールに分けて考えることによりモジュールの目的と評価が明確となり、また、研究・開発も進め易い。モジュールはアルゴリズム(ソフトウェア)とハードウェアとから構成され、私の研究の中心はアルゴリズムである。

ビデオカメラモジュールはレンズ系、ビデオカメラ、ビデオ信号のAD変換からなり、レンズによりCCD受光面上に結像された光の強度分布を電気信号に変換するモジュールである。基本的なパラメータは焦点距離等のレンズ特性、光(強度と波長(色))を電気信号に変換するCCDの感度特性である。視覚センサーとして一般に使われているビデオカメラに関しては文献2と3で評価を行った。

動き検出モジュールはデジタル化されたビデオ画像を入力として、出力は動き情報である。入力されるビデオ

脚注) 動物の視覚は脳の機能である。複雑系である脳の機能を要素(モジュール)に還元するのは不可能という主張がある。各モジュールを寄せ集めた計算機視覚システムはヒトの視覚の様な汎用性と頑強性とを現在のところ達成していない。しかし、要素の分解と統合という近代科学技術の標準的アプローチをここでは採用する。

画像は、NTSC を利用した場合、1 画面 512 画素 × 480 画素、毎秒 30 画面、画素明度 8 ビット、画素明度誤差の標準偏差は 1 ~ 2 ビット程度である。出力すべき動き情報の評価項目は後段の運動パラメータ抽出モジュールの研究を進めないといけない部分があるが、一般に動きの精度、信頼性、密度、実時間処理が考えられる。動きの密度は動き画像と呼べる位に画面全体で均一に多数求まることが望ましいので、動き検出モジュールの役割は明度画像を動き画像に変換することである。動き情報を利用する上では実時間処理が重要である。

運動パラメータ抽出モジュールは次期研究課題であり、第 6 章で議論する。

2.2. 動き検出アルゴリズムの前提条件

画像から動きを求めるには、時系列画像間の対応を取らなければならない。対応を取ることが出来る為には、次の 3 つの前提条件が必要である。

照明条件は急激に変化しない。

対象表面の光反射係数は変化しない。

対象の形状と姿勢は大きく変化しない。

これらはすべての動き検出アルゴリズムに必要な共通の前提条件で、画像中の変化が動きにより生じたものと仮定して、動き検出アルゴリズムは動きを求めている。これらの前提条件をどの程度緩和することが出来るかが動き検出アルゴリズムの頑強性の評価になる。

動き検出アルゴリズムの代表的なものは勾配法とテンプレート・マッチング法であるが、それらの比較と勾配法の実験は文献 4 で行なった。ここでは、計算機視覚に適すると考えられるテンプレート・マッチング法に関する研究について報告する。

2.3. テンプレート・マッチング法

テンプレート・マッチング法とは基準画像(これをテンプレート(鋳型)と呼ぶ)を探索画像上を移動させながら、基準画像と探索画像との一致度を評価して最大の一一致度の位置を対応点とする方法で、相関法、シティブロック法と呼ばれることもある(Fig.1.)。この節では、テ

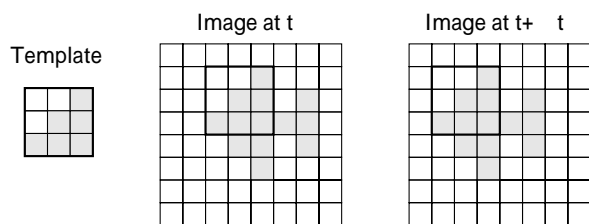


Fig.1. Template matching method.

ンプレート・マッチング法に関して、テンプレート・マッチング法の使い方、画像特性、一致度の評価基準、テンプレートの大きさと形状、探索範囲と画像更新レイト、探索パラメータ、実時間処理の視点から検討をする。

動きの信頼性と精度に関しては、それぞれ第 4 章と第 5 章とで報告する。画像から得られる動きの信頼性と精度とは画像データに大きく依存するので、動き情報検出アルゴリズム単体の信頼性と精度を定義し求めることは困難である。この報告では、得られた動きの信頼性を評価し、高い信頼性の動きだけを選び出し、さらに補間により動きの精度を向上することを提案している。

(1) テンプレート・マッチング法の使い方

テンプレート・マッチング法は対応点を求めるアルゴリズムなので、時系列画像間でテンプレート・マッチング法を行なえば動きや移動体の追跡が出来、視点の異なる複数の画像間でテンプレート・マッチング法を行なえば三角測量の原理で距離を求めることが出来る。ここでは動き検出を中心に検討を進める。画像上の動き(速度場)はオプティカルフローまたはオプティックフローと呼ばれる。

動きや距離を求める場合、動きや距離を求めたい場所を切り出すことにより簡単にテンプレート作ることが出来る。この場合、テンプレートは入力画像に応じて更新し続けられるので、2.2. の前提条件が満足される可能性が大きい。

それに対して、移動体を追跡する場合、移動体のテンプレートを前もって作る必要があり、移動体の移動により移動体の見え方が変化するので、見え方の変化に応じてテンプレートを何らかの方法で更新する必要がある。テンプレートを作る方法としては、適当な画像の中から手作業で切り出したり、移動体のモデルを考え、CAD を使って作成する方法などが考えられる。単に、移動体を発見し追跡するだけならば、動き情報を使って移動体を自動的に切り出しテンプレートを作成し、動き情報に応じて適当にテンプレートを更新する方法が考えられる。

この報告ではテンプレート・マッチング法により動きを求める実験を行なう。

(2) 画像特性

テンプレート・マッチング法はデジタル画像に対して実行される。カメラのレンズにより対象の明るさ(輝度)に比例した像が CCD 受光面上に結像し、その光強度が電気信号に変換され、カメラ内部で信号処理され、AD 変換されたものがデジタル画像の画素明度である。CCD 内のホトダイオードの光 / 電気変換特性は直線で、ビデオカメラ画像をモニターに表示する場合、カメラ内部でモニ

ターの特性に合わせる為の補正(脚注1)を行なうことがある。視覚センサーとしての光/電気の変換特性としては、動物の眼の特性に倣って対数変換とすることが良さそうである。

さらに、画像のノイズを取り除いたり、特徴を強調したりする空間的/時間的フィルターを掛けることもあるが、フィルターにより特徴を強調することは画像の特徴を良く表わす場合もあるが、画像データに含まれるノイズを強調し誤対応に陥ることもある。

この報告の実験では、ビデオレイトで対数変換を実行することが出来なかったため、輝度/明度の変換特性は直線とした。

(3) 一致度の評価基準

明度を特徴とするテンプレート・マッチング法の一致度の評価基準には、

- ・統計学上の相関係数：

$$\{I_r(i, j) - \bar{I}_r\} \{I_s(i+x, j+y) - \bar{I}_s\} / \sigma_r \sigma_s$$

- ・明度差の2乗の総和：

$$\{I_r(i, j) - I_s(i+x, j+y)\}^2$$

- ・明度差の絶対値の総和：

$$|I_r(i, j) - I_s(i+x, j+y)|$$

等が使われている。ここで、 I は画素明度、 σ は画素明度の標準偏差、下付き添字の r, s はそれぞれテンプレート(基準画像)、探索画像を表わす。座標 (x, y) は一致度を評価する位置を表わし、テンプレート全体の画素 (i, j) に渡り総和を取る。

相関係数は数学的に厳密である。その物理的な意味は、画素の明度を領域の標準偏差で割ったものは対象表面の光の反射率に比例するので、相関係数による評価基準は、照明が空間的に均一であれば、照明の明るさの影響を受けない。しかし、一致度を評価する位置ごとにテンプレートと重なる探索領域の明度平均と標準偏差を計算しなければならないので、計算量が多くなる。簡略化した相関として明度積の総和を利用して平均値と標準偏差の計算を省略する方法も考えられる。この場合でも、テンプレートと重なる探索領域の明度の最大値で規格化する

必要がある。輝度/明度の変換特性を対数とした場合、明度最大値による規格化は、

$$\ln(I \times a) = \ln(I) + \ln(a)$$

となるので、一定値を加算することにより行なうことが出来る。最大値による規格化をしないで、明度積の総和を評価基準に使ったものもあるが(文献5)、探索範囲が広いと対応点で評価値が最大値を取るとは限らない。

明度差の2乗の総和と明度差の絶対値の総和はテンプレート・マッチング法の一致度の評価基準としては解かり易い(脚注2)。明度差の2乗の総和はテンプレートが大きくなると4バイトの整数表現の上限を越えることがあり、また、乗算より絶対値の方が処理時間が短いので、明度差の絶対値の総和(SAD)がテンプレート・マッチング法の評価基準として、多くの場合、使われる。この報告でも、実験に使用した画像処理装置の関係から、明度差の絶対値の総和を評価値として使った。

(4) テンプレートの大きさと形状

テンプレートは大きい方がノイズに強く、追跡と動き検出が正確になるが、テンプレートの大きさ(面積)に比例して、計算機負荷が増加する。また、1枚のデジタル画像の大きさは512 × 512画素程度なので、大きなテンプレートを使うと、動きの空間分解能が悪くなり、追跡の範囲が狭くなる。ビデオカメラのレンズの焦点距離を含めて利用対象に応じたテンプレートの大きさを決めなければならない。

デジタル画像は直交2次元配列なので、テンプレートの形状は一般的に長方形である。マスキングにより任意の形状に対応することが出来るが、計算効率は良くない。

(5) 探索範囲と画像更新レイト

画像上での対象の動きが適当な大きさになるように、処理する画像間の時間間隔を調整する必要がある。画像上の動きが大きくなるように画像間の時間間隔を大きくすると、3つの前提条件が満足されなくなり、また、テンプレートの探索範囲が広くなり処理時間が増加する。逆に、画像間の時間間隔が小さ過ぎると、得られる動きが小さくなりSN比が悪くなり、動きの検出精度が低下する。

(6) 探索パラメータ

動きを求める場合、探索画面上でテンプレートを上下、左右の2方向に平行移動させながら対応点を求めることが多い。テンプレートをアフィン変換してマッチングを取ることにより、さらにテンプレートの拡大縮小、回転、傾斜歪み、縦横比の合計6個のパラメータを決めることが出来る。

脚注1) モニターに入力される映像信号強度とそれに対するモニターの画像の明るさは、蛍光材料の特性の為に比例しない。撮像対象の輝度とモニターの明るさとが比例する様に、対象輝度と映像信号強度との特性を補正することを補正という。対象の輝度、あるいは、それに比例するビデオカメラ結像面の照度を I 、ガンマ特性値を γ 、映像信号強度を E とすると、 $E = I^\gamma$ である。

脚注2) これらの評価値は対応した時に最小値を取るのではなく、非対応の評価値と言うのが正確である。

パラメータの求め易い順番は平行移動、回転、拡大縮小である。ビデオ画像のぶれを取り除く為に、家庭用ビデオカメラでは並行移動だけを考慮して、画像を安定化しているが(文献6) さらに画像全体をテンプレートとしてアフィン変換のパラメータを高速に求め、画像を安定化する方法も実用化されている(文献7、8)。

ビデオカメラによる画像が正射影近似出来る場合はアフィン変換で充分である。遠近効果が無視出来ない場合は、テンプレートに透視変換を適応することにより原理的にはより正確なマッチングを取ることが可能であるが、対象面の傾きと対象面までの距離が必要になる。

探索空間は探索パラメータの数の次元を持つので、処理時間は指数的に増加する。また、画像データにはノイズが含まれていて、パラメータ同士が干渉し合っているため、多数のパラメータを求めることは難しい。動きを求めるこの報告ではテンプレートは更新され続けるので3つの前提条件がほぼ満足されるので、平行移動の2つのパラメータの探索空間で対応点探索を行なった。

(7) 処理時間

画像から動きが得られるまでの時間は画像データを計算機に取り込むのにかかる時間と画像データを処理して動きを求めるのにかかる時間の和である。外界が光学系によりCCD受光面上に結像する時間は光の速度が速いので一般に無視出来る。CCD上の光学像が電気信号に変換される時間は、フレーム蓄積モードで、インターレースのNTSCビデオ信号の場合、電荷蓄積時間に1フレーム(1/30秒)電荷の読み出し時間に1/30秒かかる。動きを計算する為には時間的にずれた2枚の画像が必要なので、さらに1/30秒かかる。読み出しとAD変換は同時に行なわれるので、必要な画像データを計算機に取り込む時間は1/10秒となる。ビデオカメラをフィールド蓄積モードで使うことにより画像データを取り込む時間は半分になるが解像度も悪くなる。

テンプレート・マッチング法で画像データから動きを求める時間は計算機の処理速度の向上により急激に短くなっているが、その時間は1フレーム以下であることが望ましい。

2.4. 実時間処理の重要性

実時間処理による動き検出の研究を進めるのは次の3つの理由からである。

実時間処理の必要性

画像処理により得られる情報を利用しようとするとき多くの場合、画像の実時間処理が必要になる。例えば、移動する場合、現在の周囲の状態を認識し、目標を見出し、障害物を避け、目標に向かう必要がある。周囲の状況が移動により変化し、周囲の状況と画像処理により得られ

るそれらに関する情報との間に時間遅れがある場合、移動が不安定になり、障害物を避けることが難しくなったりする。画像処理が間に合わない場合、画像処理の速度に合わせて移動速度を遅くする方法や、安全が確認された所まで移動して、画像処理が終了するまでその場所に一時停止する方法がある。しかし、画像処理が間に合わない多くの場合、画像処理は採用されなくなる可能性が大きい。

移動するビデオカメラで得られた画像には、対象が静止していてもカメラの移動や移動に伴う振動による動きが見られる場合がある。これらの動きをビデオレイトで取り除かなければ、移動対象の動きを検出することが出来ない。

シミュレーションの限界

現在使える計算機の処理速度に制限がある以上、すべてを実時間処理することは出来ないため、シミュレーション実験は必要である。対象が移動する場合、対象の移動によって画像が流れぼけることがある。飛び越し走査をしている場合、対象の移動による画像のぼけの影響は大きくなる。コマ撮りした画像データに画像処理アルゴリズムを適応して評価する場合、これらの影響を考慮して実験を行なうが、しばしば思わぬ見落としがあるため、最終の使用状態での評価実験が必要である。

誤差が蓄積して測定値が発散する様なことはシミュレーションするのが難しい。シミュレーションの評価は難しく、実環境で実時間処理を実行することによる障害の発見は新たな研究課題にもなる。

新しいアプローチ

従来のアプローチでは、計算コストを意識して、観測時間を長くすることにより移動体の運動状態の測定精度を向上したり、より丁寧な画像処理を行なうことにより画像処理の精度を向上していた。計算機の処理速度は10年間に約100倍向上している。この傾向が今後とも続くものとすれが、現在は処理量が多くて使えそうにないと思われるアルゴリズムでも実時間処理が可能になると考えられる。単純なアルゴリズムを高速処理して、精度の低い動きを多数求めて、全体として信頼性を上げるアプローチも考えられる。

テンプレート・マッチング法はアルゴリズムが単純で研究的なおもしろさが無かった為と計算量が多く実用に供しないと考えられ、今までの画像処理の教科書にはあまり書かれていない。しかし、テンプレート・マッチング法は計算機視覚での応用が拡大している。その理由はノイズに強いからだけではなく、広範囲のパターンを扱うことの出来るアルゴリズムでもあるからである。

3. 実験装置

3.1. 実験装置

実験は、ヒトの眼と同程度の解像度を持つビデオカメラを使い、2km離れた航空機の動きを求める実験を1/500のスケールで行なった(Fig.2.)

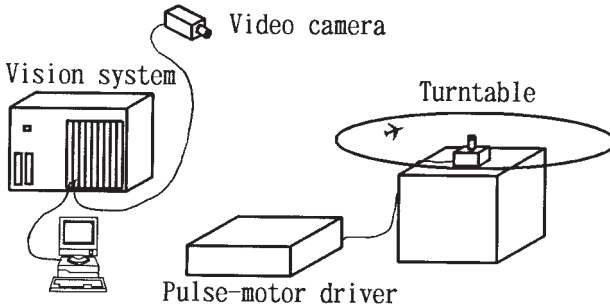


Fig.2. Experimental arrangement.

(1) 模型と移動方法

計算機視覚の適用分野として航空機着陸後の地上走行と管制塔からの空港面監視を想定しているため、2kmの距離の航空機を選んだ。動きを検出する対象は客室窓の無い貨物飛行機(Boeing747-200F)の1/500の模型(文献4)を使用した。画像的な特徴の少ない貨物機で動きを求めることが出来れば、画像的特徴の多い空港内を走行する車両等の動きの検出は容易である。ヒトの眼で動きの監視が可能な距離は数キロメートル程度と考えられる。

模型を回転テーブルに載せ、テーブルを回転させることにより、模型を移動させた。この方法では模型の背景になるテーブル面も移動するが、テーブル面を均一な布で覆うことにより画像上では動きが検出されない様にした。また、この方法により図と地の分離という画像処理上の難しい問題を避けることが出来る。模型を糸で吊り空中を回転させる方法も検討したが、模型の揺れを押さえることが出来なかった。

模型の移動速度精度

回転テーブルの回転面は厚さ5mmの合板から切り出した円盤で、円盤の中心をパルスモータ駆動の回転台(中央精機(株)PS-125)に固定し、回転させた。パルスモータ駆動用電源の周波数は500パルス/秒間隔で最大5000パルス/秒である。最大周波数の時、テーブルは25度/秒の回転速度で回転する。回転速度の精度は電源周波数の精度により決まり、付録B.に示す様にその精度は高い。模型の移動速度精度に影響するものは、i)回転半径が異なることにより模型の部分により移動速度が異なること、ii)模型の周速が一定でもビデオカメラから模型を見る視角が異なることにより模型の速度が異なることである。

i) 回転半径の影響

模型胴体を半径400mmの円周上の接線方向に置いた場合、胴体中心線から主翼先端と胴体外面とはそれぞれ30mmと6.5mm離れているので、胴体中心の回転半径に較べて主翼先端と胴体外面との回転半径はそれぞれ7.5%と1.6%異なる。実験では機首から主翼付け根までの部分の画像データを使うことにより、回転半径の影響が少なくなる様にした。

ii) 視角の影響

充分遠くのカメラから見た場合、周速一定で移動する模型の見かけの速度は \cos で変化する。ただし、はカメラ・回転中心・模型の成す角度である。模型の全長は140mmなので、i)の様に置かれた場合、回転中心に対して張る角度は約20度(± 10 度)である。 $\cos 10^\circ$ は0.985であるので、模型がカメラと回転中心とを結ぶ線上付近のデータだけを使うことにより、視角の影響を少なくした。

(2) ビデオカメラと照明

焦点距離50mmのレンズを着けた1/3インチCCDのビデオカメラ(ソニー(株)のモノクロカメラXC-73)を使用した。CCD画素の大きさは縦 $7.4\mu\text{m}$ ×横 $6.35\mu\text{m}$ なので、CCD画素当たりの視角は約0.5角度分である。ヒトの眼と同程度の解像度のカメラを使う理由は、計算機とヒトの脳に入力される情報を同程度にし、計算機処理とヒトの視覚処理の結果を比較する為である。カメラの特性は直線($=1$)に設定した。

照明は天井の蛍光灯を主として使ったが、必要に応じて、蛍光灯スタンドや直流点灯の白熱灯を併用した。蛍光灯の明るさの変動の影響を少なくする為にビデオカメラの1/100秒の電子シャッターを使った。

移動対象の模型とビデオカメラ間との距離は4mとした。この条件で、回転テーブルの中心から400mmの位置に模型を置き、5000パルス/秒で回転させた場合、模型の周速は180mm/秒(実機サイズに換算して90m/秒)、画像上での移動速度は10画素/フレームになる。

(3) 画像処理部

ビデオカメラからのNTSCのビデオ信号を8ビットの精度でAD変換した。テンプレート・マッチング法を高速に実行する画像処理ボードとして、富士通(株)のトラッキングビジョン(以下、TRV。第4章の実験に使用)とカラートラッキングビジョン(以下、CTRV。第5章の実験に使用)を使用した(両者を区別する必要がない場合は、トラッキングビジョンまたはTRVと書く)。使用した画像処理ボードの詳細は、他のテンプレート・マッチング法専用の画像処理ボードとの比較を含めて、付録A.にまとめた。TRVではデジタル画像の縦横比が1になるようにビデオ信号のAD変換のサンプリングが固定しており、今までの様にCCD画素とデジタル画像の画素と

を1対1に対応させることが出来なかった。

3.2. テンプレート・マッチング法の処理速度

8 × 8画素のテンプレート、16 × 16画素の探索範囲のテンプレート・マッチングの処理速度は、TRVで180回/フレーム、CTRVで500回/フレームとカタログに書かれている。以下では、テンプレート・マッチングが長い単語なので、相関という単語をしばしば使う。

(1) TRVの処理速度

TRVの制御はMC68040搭載のVMEボード計算機(株)アパルデータ、25MHz)と実時間オペレーティングシステムのOS-(マイクロウェア・システムズ(株)を使って行なった。TRVはビデオ信号をAD変換するビデオモジュール(以下、VMDL)1枚と相関演算を行なうトラッキングモジュール(以下、TMDL)最大5枚から構成される。TMDLを増やすことにより処理速度は上がるが、頭打ちの傾向が見られる(Fig.3.)。各TMDLの相関演算は並行して行なわれるが、各TMDLの制御は1台の制御用計算機で行なわれるので、処理能力が直線的に伸びない。制御用計算機によるTMDLの制御は、相関演算条件の設定と計算結果の読み出しである。

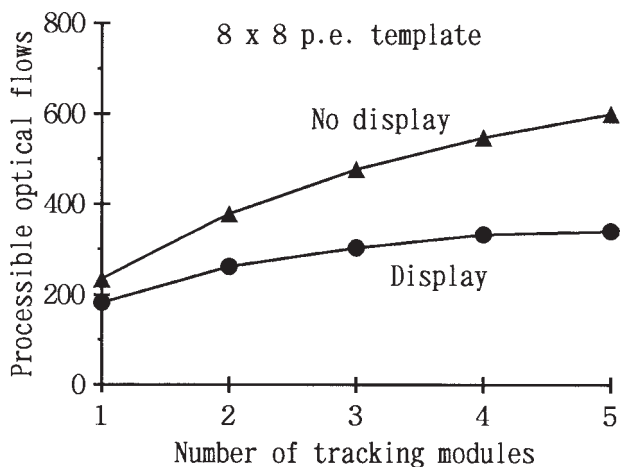


Fig.3. Maximum of processible optical flows with TRV.

相関演算1回当たりの時間の内、並列化出来る部分を t_1 、並列化出来ない部分を t_2 とし、TMDL1枚が1フレーム当たりに計算可能な相関演算を n 回、TMDLの枚数を m 枚とすると次式が成立する。

$$n \times t_1 + m \times n \times t_2 = 1 \text{ フレーム (33m秒)}$$

TMDLの枚数 m を1枚から5枚に変えた時の実験結果(Fig.3.)を代入して、最小2乗法で相関演算1回当たりの時間を求めると、

$$t_1 = 108 \mu \text{ 秒}$$

$$t_2 = 33 \mu \text{ 秒}$$

となった。 t_1 はTRV取り扱い説明書の記載110 μ 秒とほぼ

一致する。

この結果から、TMDLの枚数をいくら増やすことが出来ても1フレーム当たり1000回以上の相関演算を実行することは出来ないことが解かる。

実際には、相関演算の結果を表示したり利用する処理が必要になるので、1フレーム当たりの相関演算はTMDLを最大の5枚使った場合で300回程度である。

(2) CTRVの処理速度

CTRVは上記のVMEボード計算機では制御出来なかったため、Sbus-VMEbus変換ボードを通してUNIXワークステーション(Sun4/2、40MHz)で制御した。実験では1フレーム当たりの相関演算が120回であれば比較的安定した処理が可能であった。カタログ性能の1/4程度であったが、その原因としては非実時間オペレーティングシステムのUNIXを使ったことと、Sbus-VMEbus変換ボードを通してCTRVとのデータのやり取りをしたことが考えられる。非実時間オペレーティングシステムなので、処理速度の測定値がばらつき、正確な測定は意味がないと考え、実験はこれまでとした。

4. 動きの信頼性

動きを計算する過程で微分(差分)操作を行なう勾配法に較べれば、テンプレート・マッチング法により得られる動きの信頼性は高いが、画像データ(対象、撮像条件、ノイズ)によっては誤った動きを出力することがある。この章では、動きの信頼性を確認し、信頼性の低い動き情報を取り除くアルゴリズムを検討する。力学的なモデルに基づく新しいアルゴリズムを提案し、それを実時間で実行するプログラムを作成し、実験によりその有効性を確認する。

4.1. 高信頼性の動きを獲得する方法の比較検討

高い信頼性の動きを獲得する方法は、動きを求める前に画像データを評価する方法と動きを求めてから動きの信頼性を評価する方法に大きく分けられる(Table1.)。後

Table1. Methods of obtaining reliable motions.

- Methods prior to motion calculation
 - Estimation of motion reliability
 - Image data evaluation ①
- Methods posterior to motion calculation
 - Smoothing ②
 - Evaluation of motion reliability
 - Algorithm dependent ③
 - Algorithm independent
 - Comparison ④
 - Physical model ⑤

者は動きの平均を取る方法と信頼性の高い動きだけを抽出する方法があり、抽出する方法は、さらに、動き検出アルゴリズムに依存する方法と依存しない方法に分類される。動き検出アルゴリズムに依存する方法以外の方法は、原理的には動き検出アルゴリズム全般に渡って適応が可能である。

動きを求める前に画像データを評価する方法

画像処理により求める動きの信頼性は、画像データを検討することにより推定することが出来る。動きを求めようとする位置で、あらゆる方向に画像的特徴があれば正確な動きが求まり、画素明度の空間分布が均一な場合は動きの信頼性は低く、空間分布が1次元的な場合は窓問題に陥る危険性が高い(例えば、文献4)。インタレストオペレータ(文献9)の様な事前の評価方法が提案されているが、厳密な事前評価にはパターン認識の難しさがあり、処理時間がかかる。

動きの平均を取る方法

動きを求める位置で得られた動きをそのまま使うのではなく、その周囲の動きを含めた平均を取り、その位置の動きとする方法である。信頼性が低い動きの信頼性を上げるが、信頼性の低い動きの誤差が伝搬して、信頼性の高い動きの信頼性が低下する。平均は時間平均と空間平均が考えられるが、平均の取り方に応じて動きの時間的、空間的な分解能が悪くなる。

動き検出アルゴリズムに依存する方法

この報告で採用したテンプレート・マッチング法について適応可能な方法について議論する。テンプレート・マッチング法では、時系列画像間で非対応の程度を表わす画素間の明度差の絶対値の総和(SAD)が最小になる位置を対応点として、動きを求めている。したがって、対応点におけるSADの値がその周囲の全てのSADの値より充分小さければ対応点(動き)の信頼性は高く、SADの空間分布が均一な場合や1次元である場合は、得られる動きの信頼性が低くなる。厳密な信頼性の判定には、動きを求める前に画像を評価する方法と同様に、パターン認識の問題に直面する。

テンプレート・マッチング法では、SADの空間分布から対応点の信頼性を評価する方法は標準的な方法である(文献10、11、27)。8近傍で評価すると、真の対応点が画素と画素との境界付近にある場合、信頼性が低いと判定する可能性があるので、16近傍で評価しなければならない(脚注)また、対応点の信頼性が高いと判断するSADの増加分(閾値)は画像データ(対象、撮像条件等)により変化するので、閾値の決定は試行錯誤で決めなければならない。二宮らは計算量を少なくする為に、対応点におけるSADの最小変化率(SADの最小値とその次に小さい値との差)により信頼性を判定している(文献12、13)。

CTRVではこの方法が可能であるが、TRVでは読み出し可能なSADの数が制限され、かつ、TMDL当たり1つの動きしか処理出来なくなるので、この方法は適さない。この方法は第5章のCTRVを使った補間の実験で試用した。

複数の方法で求めた動きを比較する方法

動きを正しい方法で測定すれば、どのような方法で測定しても得られる動きは等しいという考え方に基づいている。テンプレート・マッチング法だけを使う方法としては、拡大率(付録A3.参照)や大きさが異なるテンプレートを使って動きを求め、得られた動きが等しい場合、動きの信頼性が高いと判断する方法である。この方法では画像データに含まれるノイズに対しては有効であるが、窓問題に直面しているかどうか解かりにくい。また、動き検出アルゴリズムを複数回実行しなければならないので、処理速度が半分以下に低下する。

動きの変化から評価する方法

移動体の構造(質量分布)とそれに作用する力が解かれれば移動体の運動が厳密に計算出来、画像上の動きを予測出来る。最も単純な運動モデルの1つは移動体には質量があるので急激な速度変化が出来ないと言うモデルで、動きの時間的変化あるいは空間的変化が大きいものは信頼性が低いとして棄てる方法である。

この方法の特徴は、移動体の3次元世界での動きを考えることにより、動きの変化の閾値の物理的意味が明確になり、閾値の設定が照明条件等の撮像条件の影響を受けにくいことである。例えば、普通の移動体では衝突事故でもない限り1g以上の加速度を受けることはない。しかし、この方法は、ブラウン運動が支配的なマイクロな世界では適用することが出来ない可能性がある。また、この方法では移動体と背景の境界付近の動きを信頼性が低いと判断するので、その対策が必要である。

動きの時間的変化と空間的変化は移動速度 v により次式で結びつけられているので、時間的変化と空間的変化とを制約する2つの方法は物理的に等価である。

$$\frac{dv}{dt} = v \times \frac{dv}{dx}$$

動きの空間的変化を制約する方法では、8近傍を考える場合、動きの信頼性はその位置の動きがその8近傍のすべての動きと等しい時、最も信頼性が高く、その8近

脚注) 8近傍は考えている画素を取り囲む8画素である。

16近傍は8近傍を取り囲む16画素である。4近傍(考えている画素の上下・左右の4画素)と8近傍は画像処理で一般的な用語だが、16近傍は筆者の造語である。

傍のすべての動きと等しくない時、最も信頼性が低いと評価するのが自然である。信頼性を判定する動きの位置が移動体と背景との境界に位置する場合は、その位置の動きはその8近傍の動きの約半分と等しいことが期待される。信頼性を判定する動きの位置が移動体の角にあれば、8近傍の約1/4程度と動きが等しいことが期待される。したがって、近傍の動きとの関係は空間的な配置も考慮しなければならない。8近傍の動きに対して等しい、あるは、等しくないの2つの状態が考えられるので、 $256 (= 2^8)$ 通りの近傍関係の配置に対して評価をしなければならない。

時間的な変化を制約する方法は時間的前後(1近傍)を考えれば良いので、空間的な変化を制約する方法に較べて処理プログラムが簡単になり、その分、高速処理が可能になる。次節でこの方法に付いて検討を進める。

4.2. 選んだ方法：時間的な変化制約による高信頼性動きの抽出アルゴリズム

このアルゴリズムでは、動きはなめらかに変化するという物理的根拠に基づき、ある時刻に速度 v であるものが t 時間後には $\pm v$ の変化範囲にあれば、動きの信頼性が高いと判断する (Fig.4.)。動きは画像上に固定した位置で求めているので、動きの位置は移動体上を移動する。さらに、3次元的な動きが画像上に2次元的な動きとして投影される。したがって、画像上の動きの時間的な変化は単純な加速度にはならない。

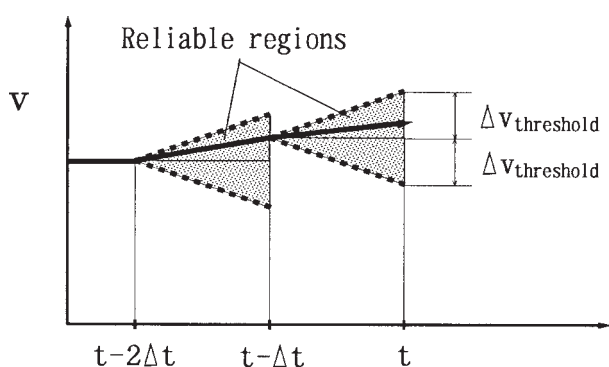


Fig.4. Concept of the temporal smoothness constraint.

この判断により高い信頼性の動きを抽出する操作は、時間的な変化が大きい動きは充分減衰し、時間的に安定な動きはほとんど減衰しない一種のフィルターリングである。このフィルターのパラメータは、 t 時間の動きの許容変化範囲 v と、許容変化範囲に収まっている期間 $n \cdot t$ である。さらに、動きの許容変化範囲に収まる確率 p を考えることにより、許容変化範囲に収まる条件を緩くすることが考えられる。実際に、このアルゴリズムが上手く機能する適当な t 、 v 、 $n \cdot t$ 、 p の組み合わせが存

在するかどうかは実験により確かめなければならない。

動きの大きさがテンプレートの大きさに較べて小さい場合、同じ大きさの動きが連続して出力される可能性が大きいので、 v を小さくし、 $n \cdot t$ を大きくすることにより信頼性を高めることが出来る。 $v \times n \cdot t$ をテンプレートより大きくすることにより、対象と背景との境界で複雑に変化する動きを取り除くことが出来る。

(1) プログラム開発

このアルゴリズムのプログラムはC言語で書き、TRV制御用計算機上で実行する。プログラム作成上の注意点は高速処理である。

動きの変化が許容範囲 v の中に収まっているかどうかの判定は、動きの x と y 成分それぞれが条件を満足する場合、収まっているとした。動きの x と y 成分の2乗和の平方根から動きの大きさを求め、許容範囲に収まっているかどうかを判定する方法は、大きさを求める段階で実数演算があるので時間がかかると予想し、採用しなかった。

ある位置での動きが信頼性が高い場合は1、低い場合は0として、ビット情報として記録し、期間 $n \cdot t$ の間、動きが許容変化範囲に収まっているかどうかの判定処理にビット演算を使用した。その結果、MC68040(25MHz)上で信頼性判定を動き1点当たり約15 μ 秒で実行出来る様になった。

4.3. 実験

(1) 実験の制限

実験にはTRVを使用したので、

- t はNTSC方式の1フレーム(1/30秒)である。
- 飛び越し走査の影響を避ける為に、拡大率を2としたので、 v は2画素/フレーム²間隔で0、2、4、・・・、32画素/フレーム²である。
- 信頼性の高い変化範囲に連続的に収まる条件は厳しく、確率 p により緩い条件にしないと条件を満足出来る動きが存在しないのではと考えたが、その必要が無かったので、実験では $p=1$ とした。

(2) フィルターとしての特性

動きの変動幅は対象、照明、撮像条件等に依存し、変動幅が任意の動きデータを作ることは難しいので、ランダムな動きと安定した動きの両極端に近い場合に対してフィルターの減衰特性を測定した。この両極端の変動幅に対する減衰の差が充分大きくなければ、両極端の動きの変動幅の差に較べて実際に使う場面での信頼性の高低を分ける変動幅の差は小さいので、減衰の差により高い信頼性の動きを抽出することは出来ない。

ランダムな動きに対する減衰特性

テンプレート・マッチング法の探索範囲を考えること

により、時間的に完全にランダムな動きを作り、減衰されるかどうかを判定し、減衰される動きの割合を計算出来る。しかし、ランダムに変動する時系列の動きは 256^n の場合が考えられ、その数が大きいので数学的に減衰を求めないで、ランダムな動きを以下に示す方法で作り実験により減衰を求めた。

時間的にランダムな動きをTRVから出力する為に、比較的均一な画像(実験室の模様のない壁)を大きく焦点をずらして撮像したビデオ信号をVMDLに入力した。デジタル画像の明度が完全に均一になった場合(AD変換器の下限または上限(脚注)を越える暗いまたは明るい画像を撮像すれば得られる) TRVは一定(安定した)の動きを出力するので、デジタル画像の明度の最低値は0ADU以上、明度の最大値は255ADU以下になるように照明を調節する必要がある。このようにした場合、CCD受光面上の像の光強度分布は均一になるが、各CCD画素に入射する光子の数は統計的に変動するので、この画像データをTRVに入力した場合、得られる動きはランダムに変動する。CCDの画素感度がほぼ均一でない場合など、得られる動きがランダムでなくなる可能性は考えられるが、得られた動きのランダムさの程度の厳密な評価は行っていない。

拡大率2のテンプレート 8×8 画素を縦横16画素ピッチで縦6 \times 横20の動きを求め、それらをフィルターを通し、時間的にランダムに変動する動きに対する減衰を求めた(Fig.5.) 減衰は n に対して指数的に小さくなっているので、TRVにより得られた動きは充分ランダムであると考えられる。

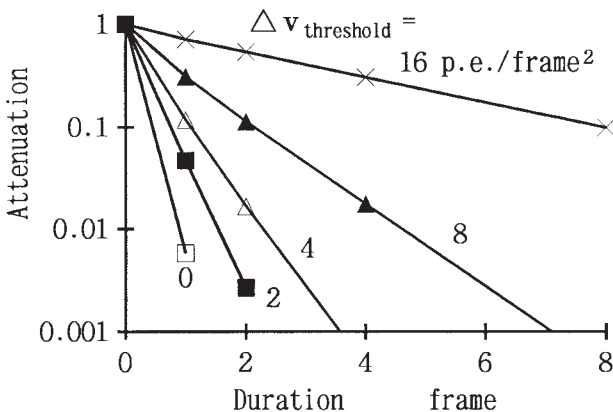


Fig.5. Attenuation of randomly fluctuating optical flows.

安定した動きに対する減衰特性

完全に安定した動きに対しては、提案したアルゴリズム

脚注) 8ビットのAD変換器を使用した場合、下限は0ADU、上限は255ADUになる。ここでADUはAD変換単位である。

ムの減衰は無い。実際に移動している対象をビデオカメラで撮像し、画像処理により得られる安定した動きは変動する。その原因は移動体の速度の変動、ビデオ画像の写り方と安定性、AD変換タイミングのドリフト、テンプレート・マッチング法の動き検出が画素単位であることによる桁落ちが考えられる。安定した動きのデータにはこの動きの変動成分が含まれていることが必要である。

時間的に安定した動きを得る為に、パルスモーター駆動の回転テーブルの上に模型を置き、模型を一定速度で移動させて撮像したビデオ信号をVMDLに入力した。回転テーブルの上(模型の背景)には黒い布を敷き、VMDLでデジタル化した画像で背景明度が0ADUとなる様にした。原理的には模型の背景の明るさをAD変換器の飽和明度とする方法も考えられるが、模型の表面の光反射率が大きいので、背景を全て飽和させる実験は非常に難しい。TRVでは明度0ADUの均一な画像に対しては0画素/フレームの動きを出力するので、大きさが0画素/フレーム以外の動きは移動対象の動きである。

拡大率2のテンプレート 8×8 画素を縦横16画素ピッチで縦6 \times 横20の動きを求め、安定した動きに対する減衰を求めた(Fig.6.)

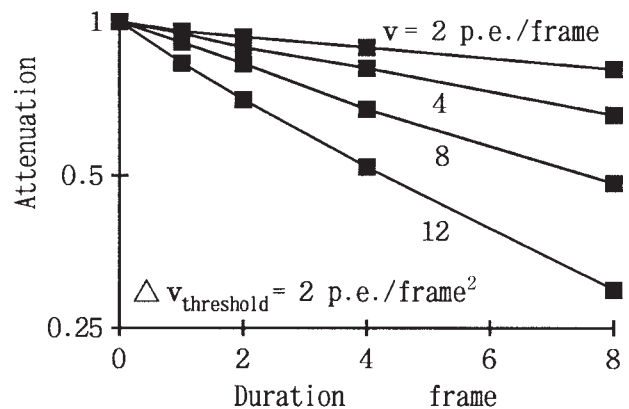


Fig.6. Attenuation of stable optical flows.

例えば、 $v=2$ 画素/フレーム²、 $n_t=2$ フレームの場合、ランダムな動きに対して減衰0.003、安定した動きに対して減衰0.93なので、時間的変化制約による方法が上手く機能することが期待出来る。実際の信頼性の低い動きは、ここで作り出したランダムな動きよりは安定しているため、安定した動きの減衰との差はかなり縮まることが予想されるが、どの程度になるのかは不明である。最終的には、このアルゴリズムを現場で使ってみて評価しなければならない。

(3) 加減速実験

模型の移動速度が一定の場合、テンプレート・マッチング法の動き検出が画素単位であるので桁落ちし動きの

変動成分が取り除かれ、安定した動きがほとんど減衰しない可能性も考えられるので、加減速実験を行なった。

設定可能な実験条件が実験装置の為に大きく制約された。パルスモータ起動時 1000 パルス/秒 ($v=2$ 画素/フレーム) と最大速度時 5000 パルス/秒 ($v=10$ 画素/フレーム) の間を 0.5、1、2 画素/フレーム² の一定加速度で加減速する実験を行なった。静止した状態から急加速するので、回転テーブルが振動して、加速度を正確に一定にすることが出来なかったが、一定加速時の動きの減衰は定速時の安定した動きの減衰よりも少し小さいだけで、ほぼ等しかった。

(4) 追跡実験

回転テーブル上に明るい布を敷き、その上に模型を載せ、テーブルを回転させ、模型を 6 画素/フレームの速さで移動させた。模型をビデオカメラで撮像し、16 画素ピッチで縦 10 × 横 30 の動きを求め (Fig.7.)、上記アルゴリズムにより信頼性の高い動きをビデオレイトで抽出した (Fig.8.)、得られた信頼性の高い動きの重心は、動きを求めた空間的なピッチ (16 画素) 程度の誤差で、移動対象に追跡した (Fig.9.)、この実験では背景に対してもピントが合っているので、背景の不均一さが原因で安定した動きが求まる傾向にあるが、その様な動きは少ないので、重心の位置にはほとんど影響を与えなかった。

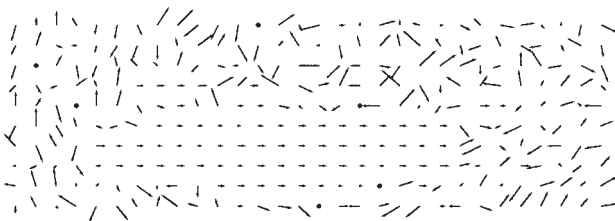


Fig.7. Optical flows output from TRV.

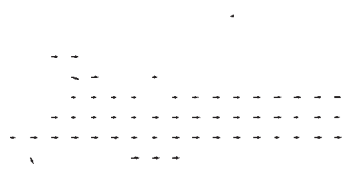


Fig.8. Reliable optical flows by the temporal smoothness constraint.

背景に模型の影が出来、影の動きが出力される。影の動きはこのアルゴリズムでは取り除くことが出来ないが、影は模型と一緒に動くので、追跡には影響しなかった。

(5) 空間的变化制約による実験

4.1. で述べた様に、8近傍の動きの関係は空間的な配置を考えた場合は 256 通りになる。空間的な配置を考

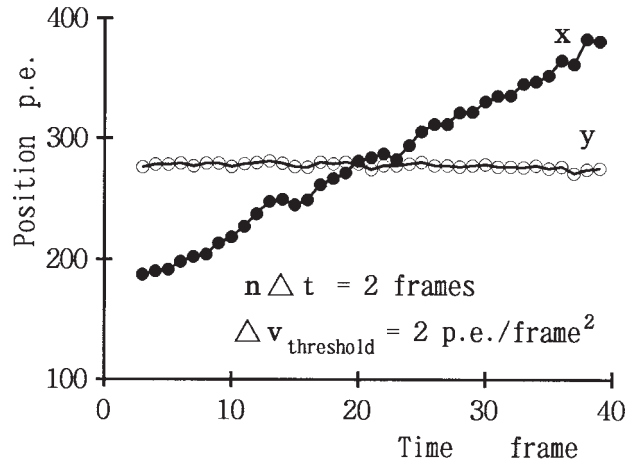


Fig.9. Central locus of reliable optical flows.

ない場合は 8 通りになり、動きの信頼性は 8 近傍の動きの中の等しい動きの数と正の相関があるとするのが自然である。空間的に動きが等しいということは、空間的な動きの差が v 画素/(フレーム・ピッチ)以内であることである (脚注)。

実験データとしては 4.3.(2) の動きを使い、2 画素/(フレーム・ピッチ)として、この簡単化した空間的な動き変化制約のアルゴリズムを手作業で実行した (Fig.10.)、図の縦軸は 8 近傍の動きが 0、1、2、3 個以上等しい条件を満足する動きの数である。その中でヒトが見て信頼性が高い動きと低い動きを分けて示してあるが、8 近傍の動きが 3 個以上等しい条件を満足する動きは全て信頼性の高い動きになった。8 近傍の動きが 4 個以上等しい条件を満足する動きの数は、さらに減少するので、今回使用した実験データに対しては 3 個以上とするのが適当と考えられる。

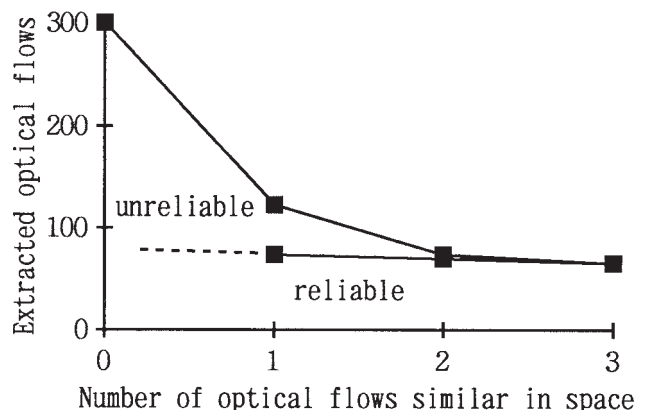


Fig.10. A result of a simplified space smoothness constraint.

脚注) ピッチは動きを求める画像上での空間的な間隔で、この実験では縦、横ともに 16 画素である。

時間的変化制約による実験結果との比較は難しいが、空間的変化制約による方法の方が信頼性の高い動きが多く残った。周囲と動きが異なる点(孤立点)の信頼性は低いと考えられるが、空間的な孤立点の除去には空間的変化制約による方法が、時間的な孤立点の除去には時間的変化制約による方法が適しているので、両者を併用するのが良さそうである。

動きの精度

テンプレート・マッチング法により得られる動きの精度は4ピット(± 3 ピット)程度である。実験に使用したTRVとCTRVでは動きの精度(探索範囲)はx方向、y方向とも-8から+7画素である。動き情報を利用する目的により必要な精度は決まるが、例えば、画像上の動きの空間分布から対象の形状と対象とカメラ間の相対的運動を推定する「動きからの構造復元」の場合、動きの空間分布(動きの相対的な差)が重要なので、4ピットの動き精度では不足と考えられる(文献14)。

動きの精度を4ピットから6ピットに2ピット向上しようとするすると探索範囲が 4×4 倍になり、それに比例して処理時間が増加する。テンプレートと探索範囲は同程度の大きさが望ましいので、処理時間は (4×4) の2乗程度になる。画像の大きさは普通 512×512 画素程度に限られているので、テンプレートと探索範囲を大きくすると動きの空間分解能が低くなり、その結果、画像の視野が狭くなる。

動きの精度に対して関心を持つ様になった個人的な動機は、第4章で提案した動きの信頼性を確認するアルゴリズムで v を1画素以下どの程度まで小さく出来るかということを知りたいからである。動きを4ピットの粗い精度で求めた為に誤差が桁落ちにより取り除かれ、提案したアルゴリズムが助けられた可能性があるからである。

テンプレート・マッチング法に補間を併用することにより $1/\text{数} \sim 1/10$ 画素程度の精度が可能と言われているが、補間アルゴリズムと精度の根拠となる実験結果について公表しているものは少なく、対象の移動速度の影響については調べられていない(文献6、15~17)。この章では、補間アルゴリズムについて調査・検討し、CTRVでテンプレート・マッチング法を、ワークステーションで補間をビデオライト処理して、対象が移動する場合の補間精度を実験により求めた結果について報告する。

5.1. 補間の理論

補間により位置決め精度が向上するのは2つの理由が考えられる。画像として(画素単位)の位置決め精度が1画素程度になる様にビデオシステムは設計されているが、

テンプレート全体としての位置決め誤差(標準偏差)は、テンプレートが n 個の画素から構成されているとすると、 $1/n$ となるからである。もう1つの理由は、最近ではビデオカメラの内部での信号処理がデジタル化され、アナログのビデオ信号の質が向上していることが考えられる。近い将来、デジタルのビデオ信号が一般的になった場合、ビデオカメラのCCD画素とデジタル画像の画素とは厳密に1対1に対応することになり、さらに補間の精度が向上することが期待出来る。

補間アルゴリズムとしては、テンプレートまたは探索画像を補間して解像度の高い画像を作りテンプレート・マッチング法を実行する方法も考えられるが、計算量が非常に多くなる。また、この方法は2値化画像に対して適用することが出来ない。

しばしば使われる補間アルゴリズムでは、一致度の評価値が画素間でも連続的に変化するものと仮定して、画素間で評価値が最大値を取る位置を補間により推定する。したがって、評価の種類と評価値変化のモデルが重要である。

テンプレート・マッチング法の一致度の評価基準には2.3.(3)で述べた様にいろいろあるが、実験に使用したCTRVでは評価にSADが採用されているので、評価の種類を比較することは出来なかった。

(1) 評価値変化のモデル

画素間の評価値がどのように変化するのは、撮像する対象の画像的特徴とビデオカメラの解像度特性に依存する。

ビデオカメラを水平に左右2台並べたステレオカメラの配置で、ビデオカメラの光軸の向きを制御しながら、左右画像の中心部分の対応評価にSADを使った実験結果(文献18)をFig.11.に示す。この実験では、パルスモータ駆動の回転台に載せたビデオカメラを、同じ対象(位置)を見る様にカメラ光軸を上下・左右に調整する。次に、左

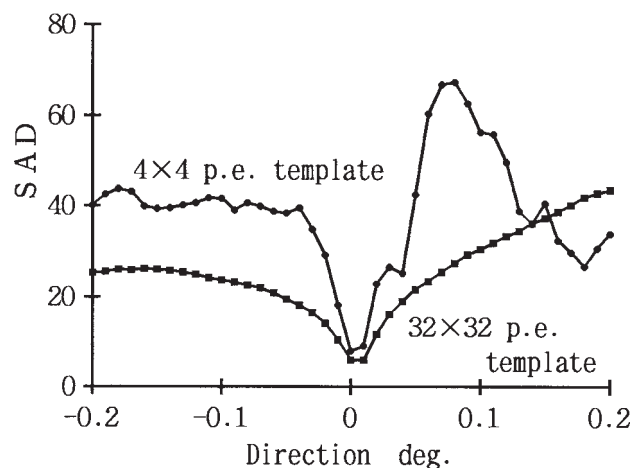


Fig.11. SAD variation near correspondence.

右方向だけを1パルス当たり0.005度(デジタル画像上で0.6画素相当)の精度で移動させながら、SADを計算した。

ほぼ対応点を中心として対応を取る部分(テンプレートの半分程度の範囲で、SADは対応点に向けて左右非対称で単調に減少し、対応点に近づくにつれて減少が大きくなった。この実験結果より、広い範囲のSADの変化は多項式近似に向かないことが解かる。また、テンプレートの大きさはSADの最小値(1画素当たりの平均値)に影響を与えないが、小さなテンプレートの方がSADの変化が大きいのので正確な位置決めが出来ることが予想される。

回転台を使ったこの実験方法では機械的ながたが1パルス程度(回転台バックラッシュのカタログ値)あるので、1画素以下の動きに対するSADの変化を知ることは出来ない。しかし、この実験結果からSADの変化が大きくなる対応点付近のデータだけを使って補間する方が補間精度が良くなり、補間の計算量も少なくなると予想されるので、3点と4点の線形補間と2次函数(放物線)の3点補間について検討を進めた。3点補間では2画素間の3個のSAD、4点補間では3画素間の4個のSADが必要になる。SADが最小値を取る付近でSADが折れ線的に変化するのか放物線的になめらかに変化するのかにより、直線補間が良いのか放物線補間が良いのかが決まる。

3点線形補間

3点で線形補間するにはSADが左右対称に変化すると仮定することが必要である(Fig.12.)。この時、SADが最小値を取る位置xは次式で与えられる。

$$x = \frac{SAD(-1) - SAD(+1)}{2 \times \{ \max\{SAD(-1), SAD(+1)\} - SAD(0) \}}$$

ただし、SADの引数0は画素単位で最小値を取る位置を示し、-1と+1はその左右の位置を示す。この3点線形補間はよく使われている。

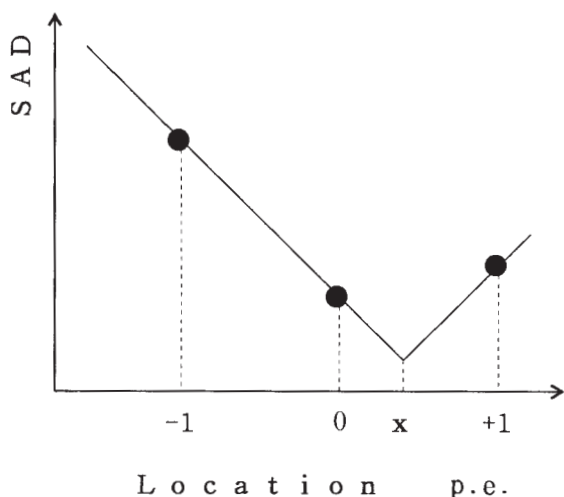


Fig.12. 3-point linear interpolation.

4点線形補間

4点のSADを左右2点づつに分け、それぞれの2点を通る直線の交点をSADが最小値を取る位置とする。SADが最小値を取る位置xは次式で与えられる。

$$x = \frac{SAD(0) - 2 \times SAD(+1) + SAD(+2)}{SAD(-1) - SAD(0) - SAD(+1) + SAD(+2)}$$

ただし、SADの引数0または1が画素単位で最小値を取る位置を示し、-1は4点のSADの左端、+2は右端の位置を示す。xは引数0の位置からの距離である。

2次函数補間

3点のSADの値から2次函数 $ax^2 + bx + c$ の3つの係数を決める。この時、SADが最小値を取る位置xは次式で与えられる。

$$x = \frac{SAD(-1) - SAD(+1)}{2 \times \{ SAD(-1) - 2 \times SAD(0) + SAD(+1) \}}$$

ただし、SADの引数0は画素単位で最小値を取る位置を示し、-1と+1はその左右の位置を示す。

5.2. 予備実験

(1) 予備実験方法

航空機模型を水平移動台(移動量読み取り誤差1/100mm)に載せ、焦点距離50mmのレンズを使い、カメラと模型との距離を約2.6mとした。この時、0.8mmの模型の移動が画像上で約2画素に相当する。最初に取り込んだ画像の模型上から基準画像を切り出す。この基準画像と、0.1mmずつ模型を移動して撮像した画像との間でSADを計算した。

(2) 予備実験結果

基準画像に対する模型移動によるSADの変化をFig.13.に示す。参照画像の縦の大きさは32画素、拡大率2で、

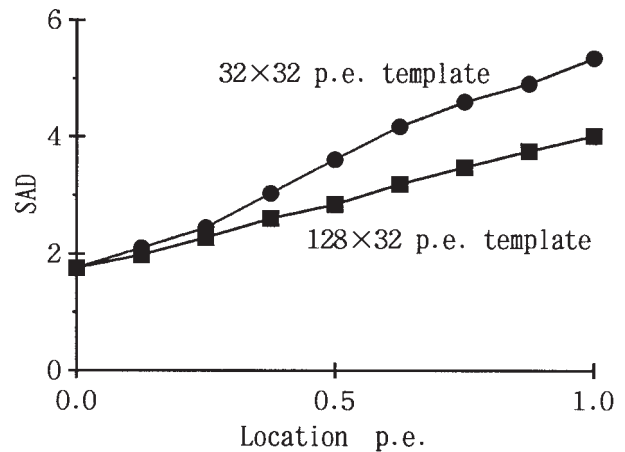


Fig.13. SAD variation versus one picture element displacement.

模型の胴体直径よりやや大きい。横128画素、拡大率2は機体の長さの70%程度になる。大きなテンプレートの方がSADの値は直線的に変化した。

次に、 32×32 画素、拡大率2のテンプレートのデータを使って、3種類の補間法により求めた模型の位置(SADが最小値になる補間位置)をFig.14.に示す。3点補間では移動量0、0.5、1画素付近で補間の精度が良い。4点線形補間では移動量0.5画素付近だけの精度が良く、移動量0、1画素付近の精度が悪い。このようになる理由は対応点付近でSADがほぼ左右対称に変化することとSADの最小値の求め方から推定出来る。

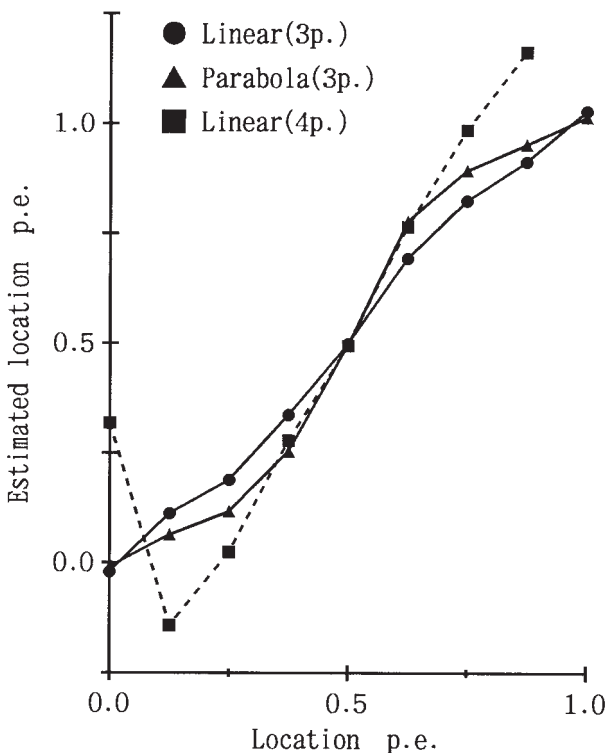


Fig.14. Comparison among three interpolations.

予備実験では意図的に大きめのテンプレートを使い、正確な位置決めが出来る様にし、補間法の優劣がはっきり出る様にした。この実験結果から、精度の優れた3点線形補間を採用した。画素単位の対応点を中心に画像上で水平と垂直方向に独立に補間を行ない、画像上のサブ画素単位の対応点の座標を求めた。

5.3. 実験

補間の精度に影響するものとして次のものが考えられる。

一致度の評価方法

評価値変化のモデル

対象の輝度分布(画像的特徴)

ビデオカメラの特性

ビデオカメラとAD変換部の同期の正確さ

テンプレートの大きさ

対象の移動速度

と については前節で既に議論した。 に関しては、補間の位置決め精度は対象の画像的な特徴(明度の空間分布)の影響を受けるので、ここで得られた実験結果は、厳密には、ある特定の対象に対する補間精度である。したがって、補間を利用する場合、補間を適応する条件ごとに精度を評価することが必要になるが、多くのものには画像的特徴があるので大体の補間精度は決めることが出来ると考えられる。ここで使った画像は貨物機の胴体前半(機首から主翼付け根まで)が写っていて、半分程度は背景を含む横128画素×縦32画素の領域である。主翼の先端は回転テーブルの中心からの距離が異なり、その移動速度が変わるので避けた。航空機模型は画像的特徴が少ないので、比較の為に、補間精度の上限が得られると期待出来る画像的特徴の大きい布製袋を丸めたものも使った(Fig.15.) のビデオカメラの特性は直線とした。 は補間の精度に重要だが、同期の正確さを確認するのは困難なので、普通のモノクロカメラを使い、その

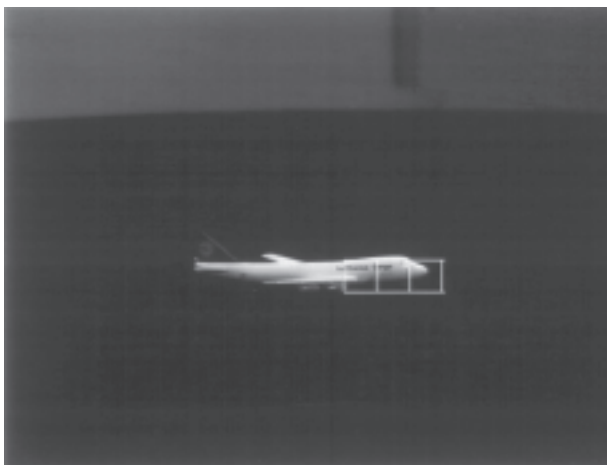


Fig.15. Moving objects

映像信号の同期信号に同期を取りAD変換した場合の補間精度を調べた。とがこの章の実験課題である。

補間の精度として、画像上に固定した位置での動きの標準偏差を使った。その理由は、テンプレートを指定する位置決めより動きの方が多数の結果を簡単に得ることが出来ることと、正規分布の再生性により動きの標準偏差は位置決めの標準偏差の2倍大きいので正確に測定出来るからである。

(1) テンプレート画素数の影響

32 × 32画素(拡大率1)、16 × 16画素(拡大率2)、8 × 8画素(拡大率4)のテンプレートを使い、1画素/フレームの速度で対象を移動させながら得られる動きの標準偏差を求めた(Fig.16.)。入力画像の影響を少なくして、画素数の影響だけを調べる為に、画素数によりテンプレートの大きさが変わらない様に拡大率を調整した。しかし、より厳密に画素数の影響だけを調べる為には、元の画像をガウス空間フィルターでぼかす必要があるが、ビデオレイト処理が出来ないので、拡大率を変える方法を採用した(付録A3.参照)。その影響を見る為に、レンズの焦点をずらすことにより光学的に明度を平均化することを試みた。こうすることにより、拡大率(1画素置き、3画素置きの明度)を使うことの影響が少なくなり、位置決め精度が向上することが期待出来るが、その効果は少なかった。

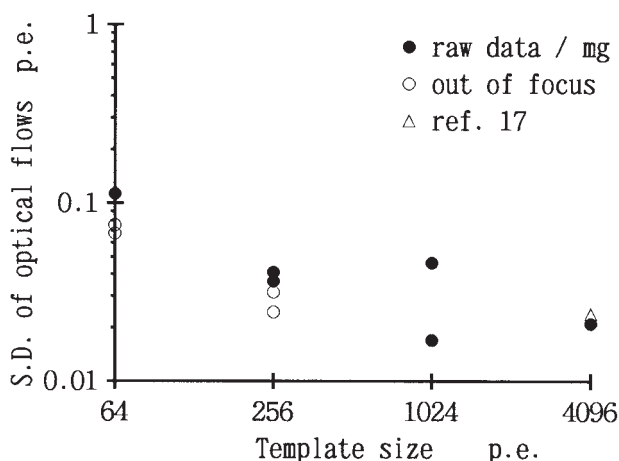


Fig.16. Template size effect on standard deviation of optical flows.

4096画素のテンプレートは、画像に合わせて、上記の32 × 32画素のテンプレートを横に4個並べて作った。IBMのデータが示してあるが、このデータは相関係数を高速に求めるIBMパソコンの拡張用ボードのカタログ(文献17)から取った。

補間の精度を表わすものとして、得られた動きの標準偏差をそのまま使う方法と、拡大率で割った標準偏差を

使う方法の2つが考えられる。前者の方法では拡大率が小さい方が有利になり、後者の方法では拡大率が大きい方が有利になる。第5章(補注)では後者の考え方でデータ整理を行なった。

予備実験で小さなテンプレートの場合、補間精度が対象の位置(画像的特徴)により異なったので、対象を移動させることにより対象の異なる位置での補間精度を求め、実験結果を統計的に処理した。この実験では対象の移動による補間精度への影響が少なくなる様に低速で実験を行なった。

今回の実験では、256画素以上のテンプレートを使っても精度はほとんど変わらない。256画素以下のテンプレートでは精度はかなり悪くなるので、今回の対象に対しては256画素のテンプレートを使うのが効率的ということになる。

(2) 移動速度の影響

移動速度が速くなりビデオ画像が流れる(ぼける)ことによりどの程度補間精度が低下するのかを調べるのがこの実験の主目的である。飛び越し走査の影響を避ける為に、以下の実験では、16 × 16画素(拡大率2)の大きさのテンプレートを使った。航空機模型を使った最初の実験では移動速度に比例して標準偏差が増加したが、移動速度を上げると回転テーブルの振動が発生した。テーブルの振動に注意した実験結果をFig.17.に示す。さらに、対象を光の反射の少ない布製袋にすると、移動速度の増加による標準偏差(誤差)の増加は小さくなった。模型は最大10画素/フレームの速度で移動させたが、図の横軸の移動速度は、画像上の移動速度を拡大率2で割ったものである。

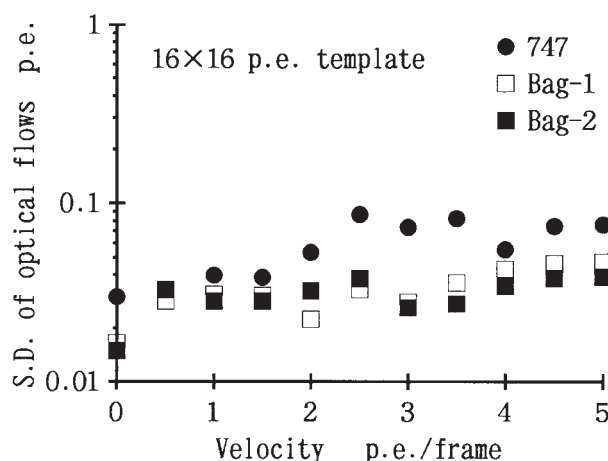


Fig.17. Velocity effect on interpolation error.

移動することによりビデオ画像が流れ(ぼけ)補間精度が低下したのであれば、CCDの露光時間を電子シャッターを使い短くすることにより、補間精度の低下が少な

くなることが期待出来る。直流点灯光源を使い照明の変動の影響を取り除き、対象として光の反射の少ない布製袋を使用して、実験を行なった。移動速度が4画素/フレーム(拡大率で割った値)の時、電子シャッター速度により補間の誤差はほとんど変わっていない(Fig.18.)

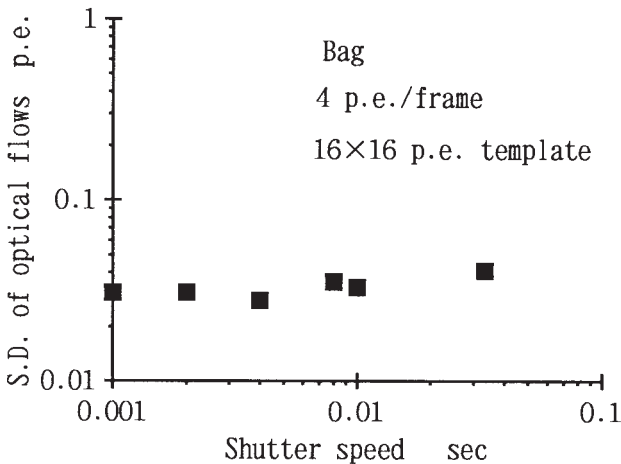


Fig.18. Shutter speed effect on interpolation error.

したがって、Fig.17.で移動速度の増加に伴って補間の精度が低下したのは、テーブルが振動し、模型の移動速度の精度自体が低下したことが原因と考えられる。画像上で1画素の大きさは回転テーブル上で約0.6mm、1/10画素は0.06mm、0.03画素は0.02mmになる。合板を切り出して手作りした直径1mの回転テーブルを、15秒で1回転させた時、1/10mm ~ 1/100mmの振動が発生する可能性は大きい。

テンプレート・マッチング法に補間を併用することにより動き精度は2ビット程度良くなったが、UNIXの下での補間を含めた処理時間は2倍以内であった。これをテンプレート・マッチング法だけで達成しようとする、この章の始めて述べた様に、処理時間は16倍あるいは16²倍になる。

動きにより補間の精度が低下するものと予想して実験を進めたが、動きの計算ではテンプレートの画像も、探索画像も同じ様に流れてぼけるので、動きの大きさがテンプレートの大きさより小さければ、動きによる補間精度の低下は無視出来たと考えられる。初めにテンプレートを登録する追跡の場合は、移動速度に合わせてぼかしたテンプレート画像を使いテンプレート・マッチング法を実行することにより、動きによるぼけの影響を少なくすることが出来ると考えられる。

6考察

トラッキングビジョン(TRV)を計算機視覚システムの動き検出モジュールのハードとして使う立場から、TRV

を評価し、今後どの様に計算機視覚システムの研究を進めるかを検討する。

6.1. 動き検出モジュールのハードとしてのTRVの評価

理想的な(完成した)動き検出モジュールはブラックボックス化され、ビデオ画像信号を動き検出モジュールに入力すれば、モジュールのパラメータは内部で自動調整されて、処理結果の動きを出力する。その出力は複数のモジュールからランダムにアクセスが可能である。現在のTRVは以下の点でこの理想からかなり遠い。

(1) 処理能力

動き検出モジュールとしての処理能力は処理速度、精度、測定範囲(ダイナミックレンジ)、処理時間で評価される。第4章と第5章で報告した技術を組み合わせることにより、ビデオ画像から信頼性と精度の高い動きをビデオレイトで数百点求める見通しを得た。

処理速度

現在の処理速度は数百点/フレームであり、増設ボードと実時間オペレーティングシステムを整備すれば、1000点/フレームを実現することが可能である。

NTSCのビデオカメラの画素数はヒトの眼の視細胞の数に較べて数百分の一なので、ヒトの眼と同程度の機能を実現する為には、ビデオカメラを複数台使うか、高品位のビデオカメラを使うことが必要になる。視覚システムの動き検出モジュールの処理点数として、現在の10倍程度の第1次視覚野のハイパーコラムの数(5000程度)を目標にしている。現在の技術レベルの向上が今後も続くものとすれば、テンプレート・マッチング法でこの程度の処理が可能になるのは5年以内と予想される。計算機視覚がヒトの視覚以上の性能を出そうとすれば、それ以上の処理速度が要求される。

精度

補間を併用することにより6ビット程度の精度を達成した。動きの精度をこれ以上向上しようとすると、ビデオカメラの光学系による画像の歪みが問題になる。カメラ較正によりある程度は取り除くことが出来るが、広い視野の場合、画像の歪みを補正することは難しくなる。

ヒトの眼の網膜像は視野の中心以外は非常に歪んでいて、網膜像と1対1に投影されている視覚野はさらに歪んでいるので、ヒトの視覚ではカメラ較正的な方法により精度を向上しているとは考えられない。したがって、計算機視覚でもカメラ較正により画像の歪みを補正する方法ではなく、精度を向上する別の方法を探すべきである。

一般に、精度を上げると計算機負荷が増えるので、処理可能な動きの点数が減少する。また、動きの精度を上げるには多くのデータ(大きなテンプレート)が必要になり、動きの空間分解能が低下するので、高密度で動きを

求めても有効な動きの数は増えない。したがって、処理点数と精度との間には最適化が必要である。

1つの動きに対して精度と信頼性を考えるのが一般的である。しかし、この報告では、得られた動きの信頼性を初めに判定して、信頼性の高い動きだけを抽出することを考えたので、信頼性と精度とを独立に考えた。

測定範囲(ダイナミックレンジ)

動きの大きさに合わせて測定可能な最大値が変化し、最大値に対して測定精度が6ビットの精度で保証されることが必要である。標準的なテンプレート・マッチング法では8画素/フレーム(=240画素/秒)程度の動きが検出可能である。フィールドレイト(1/2フレーム)を採用したり、拡大率を大きくしたり、解像度の低い画像を作ることにより、さらに動きの測定範囲の最大値を大きくすることが可能である。今回の実験範囲(5画素/フレーム、拡大率で割った値)では画像が流れぼけることによる精度の低下は検出されなかったが、より高速になれば精度が低下すると考えられる。

ゆっくりした動きを検出するには、画像間の時間間隔を長くすることにより、精度良く検出することが可能になる。CTRVでは時間間隔を1/2~255フレームの範囲で変えることが出来るので、測定範囲は8ビット取れる。測定範囲はこの程度で充分である。

処理時間

NTSCのビデオカメラを使った場合、動きを計算するのに必要な画像の取り込みに3フレーム(0.1秒)かかり、動き情報により制御する上での時間遅れになる。NTSCの場合、フィールド画像を使うことにより時間遅れを半分(3フィールド)にすることが可能であるが、解像度も半分に落ちる。プログレッシブカメラで4倍速の読み出しをすることにより視野は狭くなるが解像度を低下することなく、時間遅れを1/4にすることが出来る。

画像間隔を短くするとCCDの露光時間が短くなるので画像が暗くなり、また、検出される動きの大きさも比例して小さくなる。したがって、計算機の処理速度がいくら速くなくても時間遅れ(動きを求める処理時間)を零に近づけることは出来ない。評価項目として実時間処理を重要と考えているが、必要な露光時間により決まる画像取り込み時間が処理時間のかなりの割合を占める。

(2) 使い勝手

コスト

動きの計算コストは研究を進める上でも、実用化する上でも重要である。1994年にTRVが発売された時、ミニマム構成で約300万円、約180点/フレームの処理能力であった。テンプレート・マッチング法を実行する他の画像処理ボードに比較してTRVはコスト・パフォーマンスが大幅に優れていたが、国立研究所の研究室でも容易に

購入出来る金額ではなかった。1996年に発売されたCTRVはパソコンで使えるPCIバス版も発売され、76万円で約500点/フレームの処理能力がある。2年間でコストは1/10程度に低下し、さらに、大学ではアカデミックディスカウントを受け50万円以下でテンプレート・マッチング法のビデオレイト処理が可能になった。この程度で、テンプレート・マッチング法の利用研究が促進されると考えられるが、実用される為にはさらに1/10程度のコストの低下が必要である。

ソフトの開発環境

現在はまだTRVを使う為には、TRVの制御用プログラムを実験条件に合わせて利用者/研究者自身が作成しなければならないので、ソフトの開発環境は重要である。TRVの開発製造元と同じ開発環境を使うことにより、移植作業とその時に発生するトラブルを避けることが出来る。

動き情報を制御に使うには実時間処理が必要である。現在、TRVの開発製造元が使っていて、ユーザサポートを行なっている実時間オペレーティングシステムはVxWorks/Tornadoである。アカデミックディスカウントが受けられる大学は安く購入出来るが、そうでないと、TRVの5倍程度の導入費用がかかる。VxWorks/Tornadoを導入しても、開発製造元が使っているTRVの制御用計算機はMotorolaの68系を使っているため、処理速度の速いPowerPC系を使おうとすると、ライブラリーの修正と再コンパイルが必要になる。

計算機の技術進歩が速過ぎて、TRVの開発製造元と利用者が共通の開発環境を一致させるのが難しい。

ユーティリティ

TRVを使うのに必要な基本的なドライバーソフトは開発製造元が提供している。現在のTRVは処理能力が充分でないため、より高レベルの機能をライブラリー/アプリケーションとして提供することは出来ない。TRVが動き検出モジュールとして機能する為に次の機能が必要である。

- ・ 画像全体を粗く動きの監視(周辺視)
 - ・ 動きを広いダイナミックレンジで
 - ・ 動き検出領域の切り出し
- ・ 動きが検出された付近の注視(中心視)
 - ・ 動き検出領域に合わせて領域の拡大縮小
 - ・ 注視領域内の動きを密に

画像入力

TRVへの画像データはアナログとデジタルの入力が可能である。ビデオカメラ等からの画像入力はNTSC規格のアナログ入力を使う。デジタルはTRV間の画像データのやり取りに使われ、富士通独自の規格である。現在、ビデオ信号のデジタル化が進行中であるが、次世代のTRVはビデオカメラや他の画像処理機器から非圧縮のデジタ

ル画像入力ができることが望ましい。アナログに較べてデジタルのビデオカメラの方が、

- ・ CCD とデジタル画像との画素間の 1 対 1 対応が正確に取れているので、動き検出に適する。
- ・ ビデオカメラ内部では最近デジタル処理されている。それを DA 変換してカメラからアナログ出力したものを、TRV 側で AD 変換するのはコストと画質の点で不利である。

動きの出力

TRV では制御と後処理とを同じ計算機で行なうので、プログラムが複雑になり、計算機が後処理に専念出来ない。理想的なシステムでは、動き検出モジュールの制御はモジュール内部で行ない、モジュール間のデータのやり取りとは独立していることが必要である。動き検出モジュールの出力は、他のモジュールからランダムにアクセス可能で、モジュール間のデータのやり取りは規格化され、異種画像処理装置間の接続が容易に出来ることが望ましい。

6.2. 今後の課題：動き情報の利用アプリケーション

動き検出モジュールの開発は、動き情報を利用するアプリケーションの要求に合う様に進める必要がある。動き情報の利用アプリケーションの開発は今後の課題であるが、視覚システム自体の汎用性を意識し、共通基盤を形成するように開発を進める必要がある。その代表的なものとしては、次に示す前処理的な と、主要な とのアプリケーションが考えられる。これらのアルゴリズムが原理的には可能であることは既に示されている。解決しなければならない問題は現在 / 近未来の技術を使ってどの程度の精度で情報が得られるかと言うエンジニアリング上の問題である。

画像の安定化

ビデオカメラが移動する時に振動したり、支持台が動揺する場合、画像データがぶれる。この様なカメラの振動や動揺による画像上の動き成分を画像処理により取り除き、後処理をし易くする技術である(文献6、7、8)。

ヒトの視覚では網膜像は絶えず揺れ動いていて、動いている方がより鮮明な外界が見えている(文献19)。画像の安定化はある仮定 / モデルに基づいて行なうので、仕方によっては実際とは異なる結果を得る場合もある。処理過程をヒトがモニターする場合は画像の安定化は必要だが、画像処理によりカメラの動きの影響を取り除くことが出来るのであれば、画像の安定化を前処理として独立に行なう必要は無いと考えることも可能である。

カメラ視線(注視点)制御

カメラの解像度に制約があるので、全ての部分を高解像度で見ることは出来ない。必要な部分にカメラを向け、

より詳細な情報を得る為にカメラを視線制御する。また、外界のある点を連続して注視することにより、解くべき問題が簡単になり(文献20、21)、カメラとその点との間の相対的な運動や距離を、また、注視点の近くの注視点に対する相対的な動きを正確に求めることが可能になる。カメラの視線制御も画像の安定化と同じ問題がある。

運動検出、消失点検出

画像処理によりカメラと対象の相対的な運動が原理的に求まる。応用としては、静止した外界に対するカメラ(を搭載した物)の運動検出、または、固定したカメラに対する撮像対象の運動検出である(文献22)。航空機に搭載したカメラにより撮影された画像から動きを求め、カメラ自身の回転による動き成分を取り除いた動きの交点(消失点)は航空機が現在向かっている位置を示す。

外界形状、障害物検出

1 台のカメラが移動することにより視点の異なる複数の画像を得て、対応点問題を解くことにより対象までの方位と距離が求まり、外界の形状や障害物を検出することが可能になる(文献23)。視点が異なる画像を得る時刻に差があるので、対象が移動する場合は対応点問題を解くのが難しくなる。しかし、視点間距離を大きく取ったり、視点の位置を自由に選ぶことが可能なので、2 台のカメラを使う立体視よりも精度を向上することが期待出来る。

7. まとめ

ビデオ画像上の動きを求める方法として、テンプレート・マッチング法を調査・検討し、回転テーブルに載せた航空機模型を使った実験により評価した。富士通(株)のトラッキングビジョンでテンプレート・マッチング法を実行し数百点の動きを求め、動きの信頼性評価と精度向上をビデオレイトで実行するアルゴリズムを研究・開発した。

テンプレート・マッチング法の特徴は、数学的に簡単で、物理的 / イメージ的な理解が容易である。

パターンの種類だけテンプレートの数を増やすことにより、どの様なパターンでも直接扱うことが出来る。積分的なアルゴリズムなのでノイズに強い。計算機負荷が大きい。

動きの信頼性を確認する方法として、質量があるので急激な速度変化は出来ないという力学的なモデルに基づく、時間的変化制約による高い信頼性の動きを抽出するアルゴリズムを提案した。このアルゴリズムの特徴は、画像データの特性の影響を受けにくい。信頼性のレベルを力学的に設定することが出来るので、

閾値の設定が容易である。

信頼性の判定が1次元なので、高速処理が可能である。単純な力学的モデルに基づくので、ほとんどの対象の動きに対して適応可能である。

テンプレート・マッチング法の動き検出精度を向上する方法として、補間を検討し、実験によりテンプレートサイズと対象に対する移動速度の影響を調べた。

3点、4点の線形補間と3点の放物補間を実験で評価し、精度の良い3点線形補間を採用した。

単純なテンプレート・マッチング法の精度が4ビットなのに対して、補間を併用することにより動きの精度は2ビット程度良くなり、処理時間の増加は2倍以内であった。

5画素/フレームの移動速度の範囲では、移動により補間精度の低下は検出されなかった。

参考文献

- 1) Marr,D.; Vision, Freeman, New York(1982)
乾敏郎、安藤広志訳；ビジョン、産業図書、東京(1987)。
- 2) 五味広美；視覚センサーとしてのビデオカメラ、航空宇宙技術研究所報告 TR - 1132 (1991)。
- 3) 五味広美；視覚センサーとしてのビデオカメラ(その2)、航空宇宙技術研究所報告 TR - 1231 (1994)。
- 4) 五味広美；Local 勾配法による動き測定精度、航空宇宙技術研究所報告 TR - 1301 (1996)。
- 5) Burt,P.J., Anandan,P., Hanna,K., van der Wal,G.and Bassman,R.; A front end vision processor for vehicle navigation, Proc. Int. Conf. Intelligent Autonomous Systems (1993)
- 6) 森村淳、魚森謙也、石井浩史、赤堀裕志、北村好徳、坂口隆；純電子式画像揺れ補正技術、National Technical Report 37, 308-314 (1991)。
- 7) 根岸秀樹、武本憲幸；ピラミッドビジョン、映像情報インダストリアル 1995/1、18-24。
- 8) Morimoto,C. and Chellappa,R.; Fast electronic digital image stabilization for off-road navigation, Real-Time Imaging 2, 285-296 (1996)
- 9) Barnard,S.T. and Thompson,W.B.; Disparity analysis of images, IEEE Trans. Pattern Anal. Mach. Intell. 2, 333-340 (1980)
- 10) Inoue,H., Tachikawa,T. and Inaba,M.; Robot vision system with a correlation chip for real-time tracking, optical flow and depth map generation, Proc. Int. Conf. Robotics and Automation, 1621-1626 (1992)
- 11) 井上博允、稲葉雅幸、森武俊、立川哲也；局所相関演算に基づく実時間ビジョンシステムの開発、日本ロボット学会誌 13、134-140 (1995)
- 12) Ninomiya,Y., Matsuda,S., Ohta,M., Harata,Y. and Suzuki,T.; A real-time vision for intelligent vehicles, Proc. Intelligent Vehicles Symposium, 101-106(1995)
- 13) 安藤道則、丹下幸一、太田充彦、二宮芳樹；高速画像処理装置の開発、画像ラボ 7/7、33-36 (1996)。
- 14) Barron,J.L., Jepson,A.D. and Tsotsos,J.K.; The feasibility of motion and structure from noisy time-varying image velocity information, Int.J. Comput. Vis. 5, 239-269 (1990)
- 15) 清水雅夫、桑島茂純；画像間相関を利用したリアルタイム計測、画像ラボ 5/1、74-77 (1994)。
- 16) 太田充彦、松田幸子、二宮芳樹；高速画像処理装置の開発、画像ラボ 7/9、44-47 (1996)。
- 17) IBM PM55 シリーズ 説明資料 (1997)。
- 18) 五味広美；ヒトの視覚に学んだビデオカメラの視線制御による対応点探索、航空宇宙技術研究所報告 TR-1185 (1992)。
- 19) Steinman,R.M., Levinson,J.Z., Collewijn,H. and van der Steen,J.; Vision in the presence of known natural retinal image motion, J. Opt. Soc. Am. A2, 226-233(1985)
- 20) 大井祥照、吉川澄人、小沢慎治；注視点を原点とする極座標によるカメラ運動パラメータ推定法、電子情報通信学会論文誌 D- 80、745-753 (1997)。
- 21) Daniilidis,K.; Fixation simplifies 3D motion estimation, Comp.vision image understanding 68, 158-169(1997)
- 22) Horn,B.K.P.; Robot vision, MIT Press, Cambridge, Massachusetts (1986)
NTTヒューマンインタフェース研究所プロジェクト RVT 訳；ロボットビジョン、朝倉書店、東京(1993)。
- 23) Zhang,Z. Weiss,R. and Hanson,A.R.; Obstacle detection based on qualitative and quantitative 3D reconstruction, IEEE Trans. Pattern Anal. Machine Intell. 19, 15-26 (1997)
- 24) 内山隆、沢崎直之、青木孝、森田俊彦、稲本康、佐藤雅彦、稲葉雅幸、井上博允；ビデオレイトラッキングビジョンの実用化、第12回日本ロボット学会学術講演会予稿集 345-346 (1994)。
- 25) 森田俊彦、沢崎直之、内山隆、佐藤雅彦；カラートラッキングビジョン、第14回日本ロボット学会学術講演会予稿集 279-280 (1996)。
- 26) 実吉敬二；ステレオ式画像認識技術、計測と制御 36、190-192 (1997)。
- 27) Anandan,P.; A computational framework and an algorithm for the measurement of visual motion, Int.J. Comput. Vis. 2, 283-310 (1989)

付録 A. テンプレート・マッチング法の専用ボード

A1. 歴史

テンプレート・マッチング法は、計算機負荷が大きく、実時間処理に不向きと考えられていたが、シストリックアレー構造の計算機により並列パイプライン処理が可能で、LSI 化に適したアルゴリズムであることが解かった。動き予測により時系列画像圧縮用に開発された集積回路 MEP (Motion Estimation Processor) を使って、1991 年東京大学で局所相関演算 (テンプレート・マッチング法) ボードが試作され (文献 11)、1994 年末に富士通がトラッキングビジョン (TRV) として製品化した。自動車メーカーも社内的に画像処理ボードを開発している。処理条件の設定もハード化されているので、処理速度はさらに速い。

富士通は集積回路 ACP (Advanced Correlation Processor) を開発し、1996 年末にはカラートラッキングビジョン (CTRV) を製品化し、テンプレート・マッチング法の処理速度を向上し、動き検出アルゴリズムの研究と、その応用を加速している。

A2. テンプレート・マッチング法の専用ボード

TRV と CTRV は VME 規格ボードで (脚注)、TRV と CTRV の制御とテンプレート・マッチング法を実行して得られた結果の後処理も同じ計算機上で行なう。

TRV

NTSC のビデオ信号をビデオモジュール (VMDL) に入力し、AD 変換する。VMDL からのデジタル画像データは、各トラッキングモジュール (TMDL) の画像入力用メモリーに入力される。TMDL は 3 枚の画像メモリーを参照画像用、探索画像用、画像入力用に切り替えて使用することにより、VMDL からの画像データの入力と並行して、参照用画像と探索用画像のメモリーを使って相関演算を行なうことが可能である (文献 24)。

CTRV

TRV の最小構成は VMDL、TMDL、制御用計算機の合計 3 枚であるが、CTRV は基本ボードと制御用計算機の合計 2 枚である (文献 25)。CTRV の基本ボードには ACP が 1 つ搭載されているが、2 つの ACP が搭載されている増設ボードを増設することにより、処理速度を上げることが出来る。

自動車企業で開発されたボード

エレクトロニクスや情報通信技術の活用によって、より高度な道路交通システム (Intelligent Transport Systems) を構築し、自動車交通の諸問題を抜本的に解決

しようという取り組みが世界各国で進められている。自動車の予防安全・運転支援の基盤技術として走行環境認識の 1 つとして画像処理技術が注目され、画像認識処理アルゴリズムを開発する為の実験用ツールとして、(株)豊田中央研究所、(株)スバル研究所などで、車載可能な高速画像処理装置が開発されている (文献 13、文献 26)。(株)スバル研究所では、ステレオ視による画像認識技術を中心に研究・開発を進めているので、比較の対象とはしない。

(株)豊田中央研究所では、画像処理アルゴリズムごとにテンプレート・マッチングボード、オプティカルフローボード、ステレオ視ボード、線分抽出ボードを開発している。テンプレート・マッチングボードは MEP を 1 個使い TRV とほぼ同じ処理能力を持っている。オプティカルフローボードは MEP を 2 個使い、テンプレート・マッチング法で動き検出する条件を固定し、処理速度を上げている。ステレオ視ボードもアルゴリズムはテンプレート・マッチング法であるが、探索範囲が異なるので、専用の LSI を開発して処理能力を上げている。

時系列画像の時間間隔の変更、サブピクセル補間、SAD による動きの信頼性確認が可能である点で TRV よりも優れている (CTRV では同様なことが可能である)。ビデオ入力に ITI 社の MVC150/40 を使っているため、種々のビデオ信号入力が可能である。しかし、このビデオ入力ボードが高価なのが、社内での利用を妨げているとの話である。処理結果を DSP バスに出力することにより、バスのボトルネックを解消している。

自動車企業で開発された画像処理技術は組み込まれて商品として売ることが目的なので、その途中で、研究・開発の成果である画像処理ボードを社外に出すことは難しい。研究発表という成果の公表も、社内的に評価されていないので、難しい。

A3. ピラミッド画像と間引き画像

テンプレート・マッチング法を実行するハードの探索範囲は 16 画素 × 16 画素程度が一般的で、それ以上に探索範囲を拡大する場合、探索範囲は変えないで、空間的な解像度が低い画像を使うことにより実質的な探索範囲を拡大することが行なわれている。空間的に解像度が低い画像を作る方法としては、元の画像にガウスフィルターを掛けて画像をぼかし、1 画素置きにサブサンプリングする方法や、単純に、元の画像の 2 画素 × 2 画素の明度の平均を明度とする方法がある。

ガウスフィルターを掛けるには専用のハードが必要になるので、TRV と CTRV では元の画像をぼかさずにサブサンプリングした間引き画像が使われている。2 画素から 1 画素をサブサンプリングするのを間引き率 1 また

脚注) CTRV には PCI 規格ボードもある。

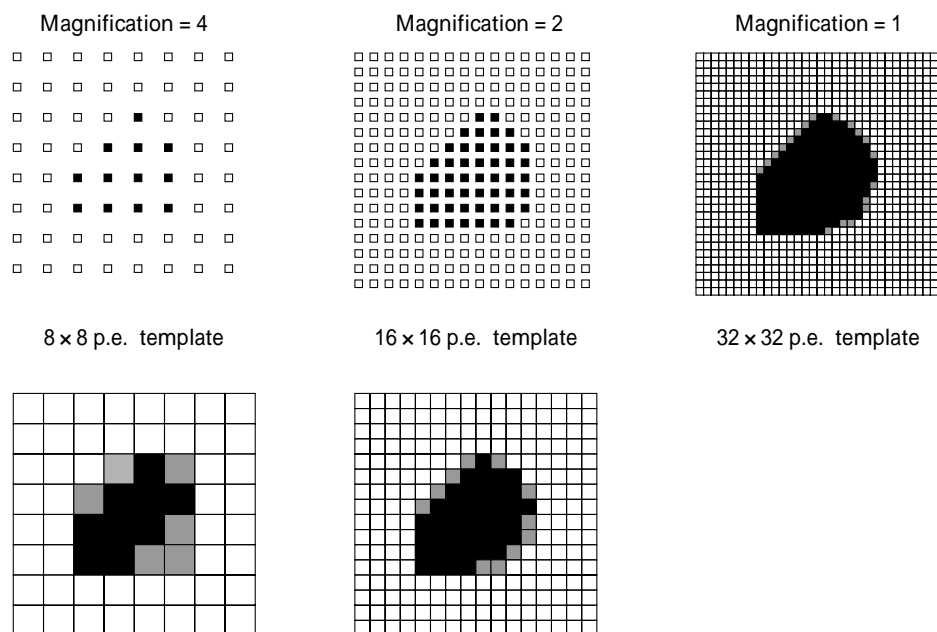


Fig.A. Magnification.

は拡大率2、8画素から1画素をサブサンプリングするのを間引き率7または拡大率8と呼ぶことにすれば、CTRVでは間引き率は0から7または拡大率1から8まで画素単位で変えることができる (Fig.A.)

元の画像をぼかさないうち、画像にサンプリング周波数よりも高い周波数が残っている可能性があるため、間引き画像を使ったテンプレート・マッチング法の信頼性は低下する。

A4. 今後の課題

画像処理市場はテレビジョンと計算機の市場に較べてはるかに小さい。ビデオカメラとテレビ放送のデジタル化、放送と通信のボーダレス化、テレビ受像機とパソコンの統合などの技術革新の流れに画像処理技術は乗らなければ効率的な発展はない。

画像処理アルゴリズムを高速に実行する専用のLSIを開発して処理能力の高い画像処理装置(モジュール)を作る必要があるが、それと同時に画像処理装置(モジュール)間で大量の画像データと少量の処理結果をやり取りする方法の標準化を進めなければ、研究成果の蓄積が出来ない。

付録 B. パルスモータ回転速度の安定性

パルスモータの回転速度は脱調(駆動用電源のパルスにモータの回転が追いついて行けなくなる)がなければ、駆動用電源のパルスによって決まる。実験に使用した駆動用電源は中央精機(株)のパルスモータコント

ローラ CPC-3D である。最大周波数が5000パルス/秒の特別仕様で、メーカー側にも資料が残っていないので、手元の資料からその駆動用パルスの安定性に付いて検討する。

駆動用パルスの発生には株アンペールのLSI:PPMC-101Cが使われている。基本周波数は100kHzでICを使って発振している。この周波数の安定性は不明だが、実験を行なう数秒間の短期的な安定性は高いものと考えられる。PPMC内部のカウンターで基本周波数を数え、設定された数(パルスレート)ごとに1つのパルスモータ駆動用のパルスを発生する。基本周波数とカウンターとの同期が取られていないことにより、オフセット誤差とランダム誤差が発生する。

オフセット誤差は駆動用パルスを出す為の処理に要する時間で、駆動用パルスを出すたびに約50μ秒の固定値が駆動用パルス間隔に加えられる。これはパルスモータの回転速度の大きさに影響するが、安定性には関係しない。

ランダム誤差は、例えば、パルスレートが20の場合、基本周波数とカウンターとの同期が取られていないことにより、駆動用パルスの間隔は基本周波数で数えて20以上から21未満の範囲で変動する可能性が考えられる。すなわち、駆動用パルスの個々のパルス間隔の変動幅は約5%になる。オフセット誤差を考慮しない時、200μ秒(=1/5000秒)間の駆動用パルスの測定値は約4762Hzから5000Hzの範囲で変動する。

今回の実験で動きを求めたビデオレイト(1/30秒間)での駆動パルスの変動は統計学的にそれ程大きくはなら

ない。1/30秒間に駆動用パルスは約 167 パルス発生する。統計学によれば、「正規母集団(その標準偏差を σ とする)から取り出した n 個からなる標本の標準偏差は σ / \sqrt{n} 」になる。駆動用のパルス変動は標準偏差に比例するので、1/30秒間の駆動パルスの変動は、約 0.4% ($=5\% / \sqrt{167}$)になる。この程度の変動であれば、今回の実験には影響しない。

今回の実験で問題になるのは短期間の周波数安定性で、それを計測により確認することは困難である。また、最近開発された、PPMC-111シリーズのLSIは基本周波数とカウンターとが同期しているので、現在ではオフセッ

ト誤差もランダム誤差も考慮する必要が無くなったので、理論的考察に留めた。

補注(15頁)

動きの大きさを表わす方法として、画像上での動きをそのまま使う方法と、拡大率で割った動きを使う方法とが考えられ、第4章では前者の方法で、第5章では後者の方法で整理した。動きの情報を使う立場からは前者の方法が良く、補間精度に対する動きの影響を知る上では後者の方法が良いと考えたからである。

航空宇宙技術研究所報告1358号

平成10年7月発行

発行所 科学技術庁航空宇宙技術研究所
東京都調布市深大寺東町7-44-1
電話(0422)47-5911 ㊦182-8522
印刷所 株式会社実業公報社
東京都千代田区九段北1-7-8

©禁無断複写転載

本書(誌)からの複写、転載を希望される場合は、管理部
研究支援課資料係にご連絡ください。

