

動画像処理による3次元奥行き分布計測

三池 秀敏*・坂井 美則**・田辺 敦***

野村 厚志****・古賀 和利*****

Measuring 3-D Depth Map by Sequential Scene Processing

Hedetoshi MIIKE, Yoshinori SAKAI, Atsushi TANABE,

Atsushi NOMURA and Kazutoshi KOGA

Abstract

Here, we propose realistic methods measuring three-dimensional (3-D) depth map by digital sequential image processing. Three different approaches as representative techniques for optical flow analysis are developed from 1) the gradient-based method, 2) the matching method and 3) our spatio-temporal mutual-correlation method. A simple idea of the motion stereo with translational displacement of a camera is introduced in the analysis. Utilizing artificial image sequences and real dynamic scenes, an accuracy and a resolution of these methods are tested. We clarified that the spatio-temporal correlation method is superior in accuracy of the optical flow velocity. The occlusion problem is discussed briefly.

1. 緒 言

三次元情報計測は、大別して能動的な手法と受動的な手法に分けられる。能動的な手法としては、レーザ光や超音波などを発して対象物からの反射波の到着時間を計測し奥行き情報を解析する方法や、縞模様や格子模様などの特殊なパターン光源を用いて、その物体表面上での幾何学的変形等の画像情報より、物体の形状などの三次元情報を得る方法¹⁾がある。こうした手法は、工業的あるいは実用的な見地から測定対象を限定した場合には、精度の良い計測手段として多用されているが、ロボットの目などの人工知能的機能を求める場合には柔軟性などの点で不向きであろう。こうした目的には、TVカメラからの画像情報を処理することにより三次元情報を得る受動的な手法が適していると

思われるが、一般には困難な問題である。受動的な方法としては、最近予備的な知識を必要としないボトムアップ的解析法が注目されるようになり²⁾、1)左右のカメラからの入力画像間の特徴点の対応処理を行ない視差を解析する両眼立体視と、2)既知方向に並進運動するカメラが捉えた動画像から変位ベクトル場(オプティカルフロー：光流動)を解析し奥行き情報を得る運動立体視法が知られている³⁾。カメラが右目の位置から左目の位置に対応するところまで並進移動したと考えれば、両眼立体視も運動立体視も解析すべき問題としては等価である。運動立体視では、両眼視での右目の画像から左目の画像に至る間に、さらに複数枚の画像が存在することになり、この情報を十分利用すればいわゆる対応点探索という困難な問題を回避できる可能性もあり、最近注目を集めている。

光流動の解析法としては従来、A)一方の画像内の特徴点を含む小領域(テンプレート)と、対応するもう一枚の画像での小領域を空間相互相関などにより評価し対応をとるマッチング法^{3,4)}と、B)動画像の任意の画素における濃淡値の時間変化と、空間変化の間にある一定の関係があることを用いるグラディエント法⁵⁻⁷⁾が知られている。マッチング法の場合得られる光流動の

*電気電子工学科

**学部電気工学科

***学部電気工学科(現在：松下電器)

****大学院電気工学専攻

*****知能情報システム工学科

密度が小さいこと、グラディエント法の場合時間空間的に滑らかな画像を必要とするなどの問題点があり、得られる光流動の精度を定量的に示したものは少ない⁷⁾。我々は、数年前より動画像の時空間相関法を用いた高精度決定法を提案して来たが⁸⁻¹⁰⁾、解析に必要な画像枚数、計算速度などの改良すべき点が多い。本研究では、問題をカメラの水平方向の並進運動から三次元奥行き情報を計測する場合(運動視)に限定し、マッチング法、グラディエント法、及び我々が提案して来た時空間相関法をそれぞれ発展させた手法を提案し、各々の手法を比較検討することにより実用的な三次元奥行き分布計測法の確立を目指している。

2. 理 論

2.1 マッチング法

ここでは、両眼立体視の考え方にそって¹¹⁾、光軸が互いに平行な二台のTVカメラが捉えた2枚の画像間のマッチングにより画像中の各画素における視差を解析する手法を提案する。考え方自体は特に目新しいものではなく、一方の画像中の小領域をテンプレートとして、もう一方の画像中の対応する小領域との空間相互相関を解析することで得ている。ただし、計算コストを押さえ、得られる光流動場の空間分解能をできるだけ高くする為に

a) テンプレートとして 1×5 画素の水平方向に長い小領域を採用する。

b) マッチングのための相関計算は画像の水平方向のみを変化させて計算する。

という、非常に簡略化した計算法を採用した。これは、2台のカメラが同一の水平線上に置かれていると仮定することや、画像中のある特徴(水平方向)を表現するには一画素の濃淡値のみでは不十分で、少なくとも2次式以上の多項式が必要ではないかとの考えが背後にある。また、マッチングに用いるテンプレートを大きくすると解析精度は上がるものの、得られる光流動場の空間分解能は悪くなることが考えられ、一つの妥協点として 1×5 画素の小領域を設定している。すなわち、空間分解能を上げるため 1×3 画素の領域を選べばミスマッチする割合が大きくなり、 1×7 画素では得られた光流動の場より物体の奥行きだけでなく形状もあわせて評価しようとする、3次元奥行き空間分布を計測する目的には空間分解能が悪くなると考えている。もちろん、このテンプレートに相当する画像領域のサイズは絶対的なものではなく、画像の性質や画像を構成する画素サイズ等により相対的に変化するも

のと考えられる。より一般的な取扱いでは、画素サイズを変更可しながら階層的なマッチング処理を行なっていく必要があると考えられる。なお、2台のカメラの垂直方向の誤差はあらかじめ補正しておけば良いという立場で、a)、b)の解析を行なっている。実際の計算では、2台のカメラを用いずに、一台のカメラが水平方向に一定の速度で並進運動する場合に、静止したシーンを捉えた動画像より適当な時間離れた2枚の画像フレームをピックアップし、それぞれ左カメラ及び右カメラが捉えた画像シーンとして計算に用いている。この意味では、カメラの直垂方向の誤差はほとんど考えなくてもよい。

2.2 グラディエント法

グラディエント法の基本的考え方は、Rigidな濃淡分布を持つ運動物体が、短い時間 δt 内では一定速度で運動し、A点: $p(x, y, t)$ からB点: $p(x + \delta x, y + \delta y, t + \delta t)$ に移動し、その間では物体上の濃淡値 g は変化しないという仮定より、

$$g(x, y, t) = g(x + \delta x, y + \delta y, t + \delta t) \quad (1)$$

が成立し、Taylor展開により一次の項までの近似で(速度一定)、

$$g_x u + g_y v + g_t = 0 \quad (2)$$

が成立する。ここに、 $g_x = \partial g / \partial x$ 、 $g_y = \partial g / \partial y$ 、 $g_t = \partial g / \partial t$ 、 $u = dx/dt = V_x$ 、 $v = dy/dt = V_y$ ここで、未知数は光流動の速度 V の x 成分 u と、 y 成分 v の2つであり、方程式が(2)式のみでは解けず、独立な拘束条件(Constraint Equation)を付加する必要がある。この手法としては従来、速度場が滑らかとして、画像全体で評価誤差の二乗和が最小となるよう最適化する手法(Global Optimization)と、近傍の領域内では速度場が一定と見なして、近傍各画素において(2)式に対応する n 個の式を立て($n \geq 2$)、Over-Deterministicな複数個の方程式から最小二乗法的な考え方で最適解を見出す手法(Local Optimization)が知られていた。

我々は、最近の研究でLocal Optimizationの考え方を発展させ、同一画素で時間方向での複数個の拘束方程式を立てることによる解法(Temporal Optimization)や⁷⁾、カメラと物体の相対的な運動方向を画像面に平行な方向(カメラの光軸に垂直な方向)に限定した場合、流体力学で用いられる保存の式を基礎とした(2)式の積分表現による解法が可能であることを示して来た¹²⁾。ここでは、提案している理論を簡単に紹介するとともに、今回対象とした光流動の方向が既知の場合の解法について述べる。

(2) 式は, $\text{div}(\mathbf{V}) = 0$ の条件のもとで3次元に戻して考えると, 次式のように変形できる.

$$g_t + \text{div}(g\mathbf{V}) = 0 \quad (3)$$

ここに,

$$\begin{aligned} \text{div}(g\mathbf{V}) &= \text{grad}(g) \cdot \mathbf{V} + g \cdot \text{div}(\mathbf{V}) \\ &= \text{grad}(g) \cdot \mathbf{V} \end{aligned} \quad (4)$$

である. 速度場 \mathbf{V} の発散が零であることは, 画面内の物体の運動がカメラの光軸に垂直な方向に限定されていることを意味することが, 指摘されている¹³⁻¹⁵⁾. (3) 式を基に, 2次元表現で微分表示から積分表示に改めると,

$$\int_{\delta C} g \mathbf{V} \cdot \mathbf{n} ds = - \frac{\partial}{\partial t} \int_{\delta S} g ds \quad (5)$$

となり, ここで \mathbf{n} は閉路 δC の外向きの法線単位ベクトルを示す. 流体力学では (5) 式は連続の式として知られている. また, 電磁気学では, 電流 $i = \rho \mathbf{V}$ (ρ は電荷密度) の連続の式あるいは電荷の保存測

$$\text{div}(i) = - \partial \rho / \partial t \quad (6)$$

が対応する. ここでは, 画像関数 g が電荷密度 ρ に, $g\mathbf{V}$ が電流 i に対応する. 以下, 画面の奥行き方向の運動が無いことを前提として (5) 式を用いて, カメラの水平方向への一定速度の運動による光流動を解析する手法を提案する.

実際の計算法は, 速度ベクトル場が定常 (時間的に一定) と仮定し, (5) 式の拘束方程式を異なる複数個 (n) の時間に対して立て, Over Deterministic な方程式群を最小二乗法を用いて解く方法を用いた (積分形の拘束条件式に対する Temporal Optimization¹²⁾). 図1は閉領域 δS として矩形領域をとり具体的な計算法を示している. 光流動の速度ベクトルの方向を $+x$ とすれば, $\mathbf{V} = (u, 0)$ であり (5) 式は次式のように変形できる.

$$u(A-C) = - \frac{\partial}{\partial t} G(t) \quad (7)$$

ここに,

$$A = \int_b^a g dx \quad C = \int_d^c g dx \quad (8)$$

$$G = \int_{\delta S} g dx dy \quad (9)$$

ここで, G は閉路 δC で囲まれた領域 δS 全体の輝度 g の総和である. 今回の解析では, Fig. 1 に示した矩形領域のサイズをマッチング法で取り扱った 1×5 画素と実質的に同等にするため 1×3 画素に選んでいる. これは(8)式の A 及び C を求める演算において, 矩形の右端及び左端の境界線上の輝度を求めるのに, 相隣る二つの画素の輝度情報を用いていることによる.

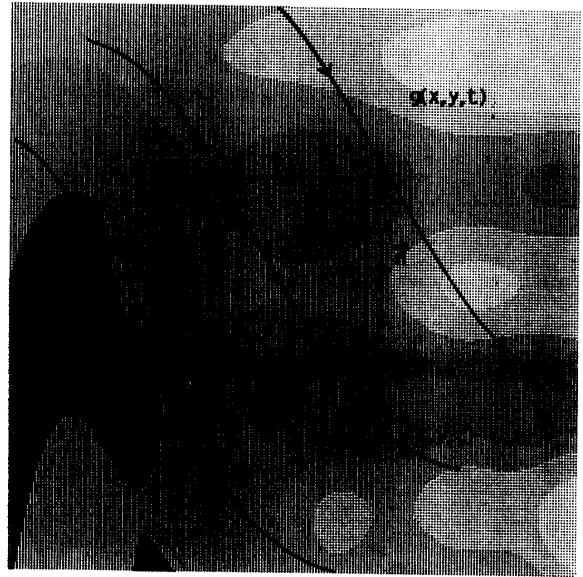


Fig. 1 Application of integration formula of temporal local-optimization (IF-TLO) method to a small rectangular area.

2.3 時空間相関法

この解析法は, 定常な二次元の速度場を計測するオリジナルな手法として提案して来たものであり⁸⁻¹⁰⁾, 速度を求めたい任意の画素での輝度の時間変化の系列 (以下, 画素時系列と呼ぶ) と近傍の画素時系列との相互相関解析を基本とする.

すなわち, 画素時系列を $E(t)$ としたとき, 中心画素 C と近傍画素 L との相互相関関数 $M_C^L(\tau)$ を次のように定義する.

$$\begin{aligned} M_C^L(\tau) &= (1/T)(1/S) \cdot \int_{-T/2}^{T/2} \{E_C(t) - \bar{E}_C\} \\ &\quad \times \{E_L(t+\tau) - \bar{E}_L\} dt \end{aligned} \quad (10)$$

ここで, S は規格化定数で, ここでは画素 C, L の時系列の分散のうち大きい方を採用している. また, \bar{E}_C, \bar{E}_L はおのおの $E_C(t), E_L(t)$ の時間平均値を示す. 一方, 同一運動物体が二つの画素 C, L を通過していることを確認し速度解析の信頼性を上げるため, 画素 C に対し L と対象な位置にある画素を R とした時, 次式のような相関関数の積 $M_C^{L,R}$ を新たな評価関数とした.

$$M_C^{L,R}(\tau) = M_C^L(\tau) \cdot M_C^R(-\tau) \quad (11)$$

近傍の画素のうち積の相関が最大となる方向より, 速度ベクトル方向に関する情報が得られ, その方向での相関のラグタイム (相互相関関数がピークとなる遅れ時間) より速度の絶対値の情報が得られる. この際間

題となるのは、計算機で取り扱う動画像が、連続なシーンを時間・空間的にデジタル化したものであり、そのままでは方向及び速さの解析精度が低いことである。こうしたデジタル誤差を解消し、高精度の解析手法とすることを目的に、1) 画像関数モデルによる理論的アプローチ⁹⁾や、2) 真の最大相関方向と真のラグタイムを、与えられた離散データより二次曲線補間により推定する近似法的アプローチ¹⁷⁾を提案して来た。ここでは、基本的には2) の方向に従い、運動方向が水平方向と推定している場合に、できる限り速度の解析精度を上げダイナミックレンジを大きくする解析法の提案を行なっている。具体的には、以下の手順に従う。

- 1) 最隣接の左右の画素時系列 (距離 l) との相互相関を計算し、速度の第一候補を計算する (予備計算)。
- 2) 1) の解析で得た速度の情報に基づき、相関を計算する相手の画素との最適距離 D を決定する。
- 3) 2) と同様 1) で得た速度の情報に基づき、相関計算に必要な時系列長 T を選定する。
- 4) 相関の具体的な計算法は、従来の積和の形 (10式参照) ではなく、次式で与える差の絶対値和を採用する。

$$M_{c^L}(\tau) = \int_{-T/2}^{T/2} |E_c(t) - \{E_L(t+\tau)\}| dt \quad (12)$$

ここで、1) から 3) の手順で相関をとる際に、近傍画素との距離や必要な時系列長を最適化することで、速度の解析精度とダイナミックレンジを高く保ちながら、得られる速度場の空間分解能と時間分解能をあげる工夫が行なわれていることになる。具体的な解析法の詳細については 3 章で述べる。また、(10) 式の代わりに (12) 式を新たに採用しているのは、経験的に (12) 式による評価の方が、比較する画素時系列間が互いに類似しているか否かの分離度が高いことになる。

2.4 運動立体視の理論と校正曲線

得られた光流動の速度ベクトル場より、3次元情報解析に関する手法にはいろんな提案があるが²⁾、ここでは最も単純な、1つのカメラの水平方向への並進運動により得られる動画像を基にした、運動立体視 (Motion Stereo) の手法により奥行き情報を計算する方法を簡単に紹介する。Fig. 2 (a) に示すように、移動するカメラが捉えたシーンにおいて点 P の奥行き D は、

$$D = h \cdot (\Delta X / \Delta U) = h \cdot (V_0 / V') \quad (13)$$

で与えられる^{16,17)}。ここで、 ΔX は一定時間 Δt 中に TV

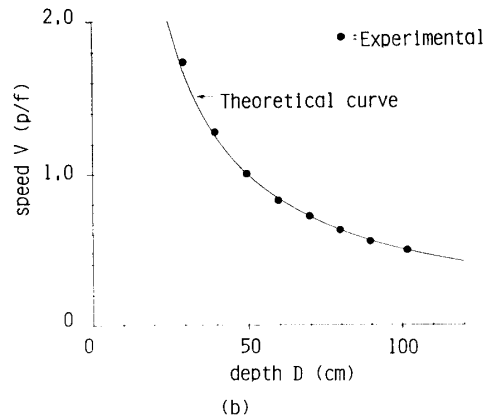
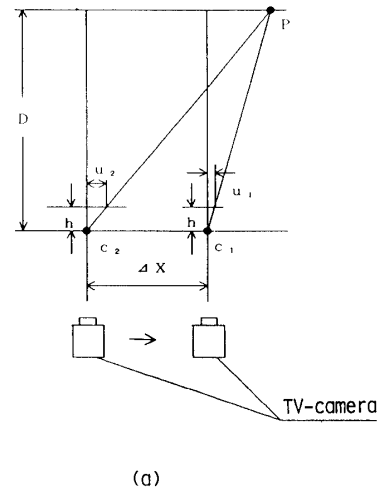


Fig. 2 (a) Principle of the "motion stereo". A relationship among displacement of a camera ΔX , disparity of the object P on the image plane $\Delta U (=u_2 - u_1)$ and depth D is illustrated. (b) A relation between depth D and optical flow velocity V .

カメラの平行移動する距離、 $\Delta U (=u_2 - u_1)$ は TV カメラの移動による画面内の点 P の位置の変化量を示す。また、 h は点 P が十分離れていれば TV カメラのレンズの焦点距離に相当する。一方、 V_0 はカメラの移動速度、 V' は物体 P の光電変換素子上の光流動の速度である。実際に得られる、画像上の光流動の値 $V(p/f)$ を用いれば (13) 式は、

$$D = h \cdot (V_0 / V) (I/F) \quad (14)$$

となる。ここに、 $F(f/s)$ は動画像のサンプリング周波数 (一秒当たりのフレーム数)、 $I(p/m)$ は装置定数で光電素子上の 1 画素当たりのサイズの逆数に相当する。このように、光流動 V を解析することにより、奥行き

Dの情報が得られることになる。

Fig. 2 (b) は、(14)式の理論式と、実際に現実のシーン ($V_0=12.5$ (mm/s) で並進するカメラが捉えた(コップの移動シーン), を解析して得られた光流動と奥行きの関係を示す実験値を比較したものである。詳細は、文献(17)に譲るが、両者は良く一致しており、解析法の妥当性及び理論の正当性を保証している。ここで用いた光流動の解析法は、2. 3節で提案した時空間相関法である。Fig. 2 (b) の光流動と奥行きの関係を示す曲線を校正曲線として用いることで、現実のシーンを解析した光流動場より奥行き分布に変換することができる。

3. シミュレーションによる検討

3.1 方法

本章では計算機シミュレーションによって作成した人工的な動画画像を用いて、前章で示した3種類のアルゴリズムにより解析し、比較検討することで、各手法の精度や信頼性を定量的に調べる。ここで利用する動画画像は、ガウス関数的な輝度分布を持ち、直径が15画素 (以下 p と略す) である円形の6つの粒子が、それぞれ0.2, 0.5, 0.8, 1.1, 1.4及び1.7画素/フレーム (以下 p/f と略す) の一定速度で水平方向に並進運動するものである。Fig. 3に、疑似カラー (16色) 表示を用いて、シミュレーション動画画像の1シーンを示している(モノクロ)。動画画像の枚数は71枚で、一枚の画像は $128 \times 128(p)$ で構成されている。また、結果の表示方

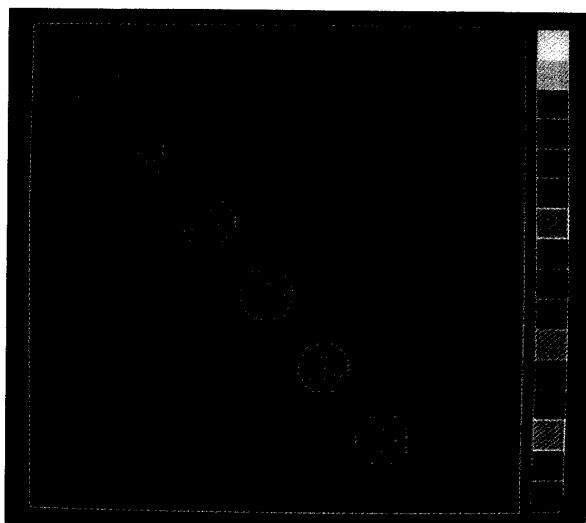


Fig. 3 A snapshot from the simulation image sequence

法として、速度の大きさの頻度分布を表すヒストグラムと、速度ベクトルの空間分布を示すベクトル分布図の2種類を用いた。ここで解析に用いたコンピュータは、PC9801 XL (80286 CPU 8MHz, 80287コプロセッサ20MHz, MS-FORTRAN 及びMS-C) である。

3.2 結果

a) マッチング法

この方法で計算に用いた画像枚数は35枚目 (両眼視の右画像に相当) と45枚目 (両眼視の左画像に相当) の2枚で、計算時間は2分程度である。具体的な計算アルゴリズムは、原理 (§2. 1) でも述べていたように、1) 右の画像に於いて、マッチングを取りたい画素を中心に、縦1 (p) 横5 (p) の小領域 (1×5 画素) をテンプレートとしてとり、その小領域が左の画像の同一水平線上の同一サイズの小領域が左の画像の同一水平線上の同一サイズの小領域と、どの程度類似しているかを計算する。

2) 類似度が最大となる位置 (左の画像でのテンプレートの中心画素) を見出すことで視差を決定する。

3) 2) で求めた右と左の視差の情報より光流動の速度を求める。

の手順にしたがう。シミュレーション画像に対する計算結果を Fig. 4 に示す。この Fig. 4 (b) の速度の頻度分布のヒストグラムより、与えた6つの異なる速度を持つ物体の各速度の位置に分布の中心があり、各分布のピークは十分鋭く、比較的精度のよい結果が得られている。また Fig. 4 (a) の速度ベクトル分布図に於いては、マッチングに用いたテンプレートのサイズが 1×5 画素と小さく、一定であるため、多少運動方向に引き伸ばされているものの、特別な形状抽出処理を施すことなく、運動物体の位置と形状を反映した光流動の分布が自動的に得られている。

b) グラディエント法

速度解析に用いた画像枚数は35枚目を中心とする30枚目から39枚目までの10枚で、計算時間は3分程度である。計算原理は2. 2節に述べているが、具体的には、

1) まず3枚の連続した画像を取り出し、求めたい画素を中心に縦1 (p) 横3 (p) の小領域 (1×3 画素) を考えて、それぞれの画像について各小領域に於ける輝度の和 ($G = \sum g$) を算出し、まん中の画像について輝度の和の時間変化の微係数 ($\delta G / \delta t$) を求める (式(9) 参照)。

2) 1) の3枚の動画画像の中央のフレームの画像に於て、2. 2節のように小領域を囲む閉路に沿う線積分

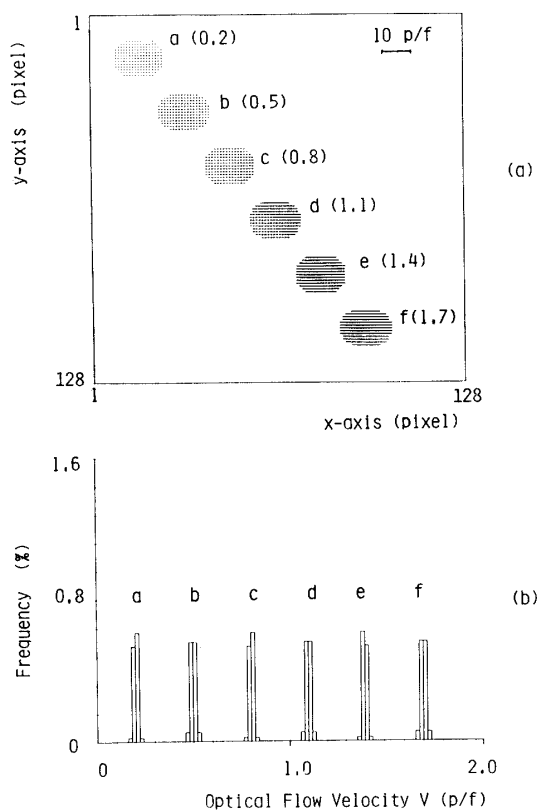


Fig. 4 Analyzed results obtained by the "matching method" for simulation image. (a) Velocity-field map, (b) Frequency histogram of the magnitude of the velocity.

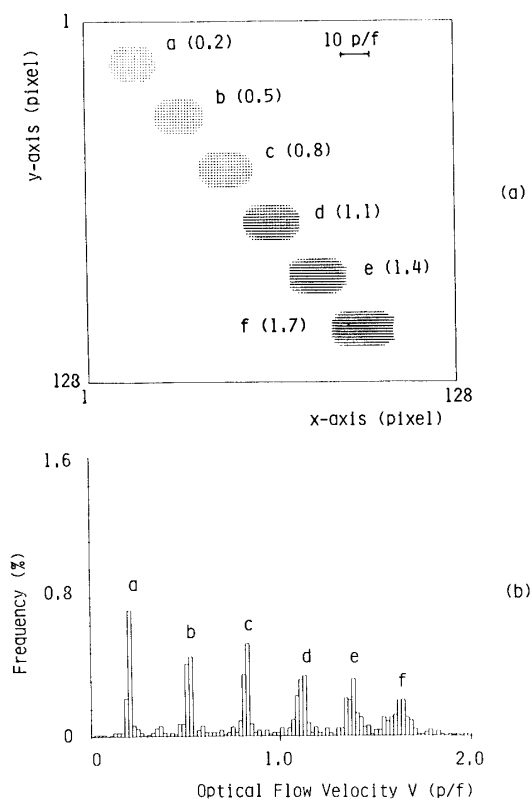


Fig. 5 Analyzed results obtained by the "gradient method" for simulation image. (a) Velocity-field map, (b) Frequency histogram of the magnitude of the velocity.

に相当する左右両端境筒での輝度の差を求め (式(8)参照), (7)式に相当するグラディエントの拘束方程式をたてる。

3) 1), 2) の操作を, 異なる 8 個の時点を中心とした動画像について求め, 拘束方程式の 8 つの組をつくる。この 8 つの方程式を基に, 最小二乗法により速度の最確値を推定する。

と言う手順に従う。ここで小領域を 1×3 画素とした理由は, 実際に左右両端の境界における輝度の違いを計算するとき, 1×5 画素の小領域を考慮することと等価であることや, 小領域 1×5 画素として計算した解析結果に大差が無く計算コストも低いことが挙げられる。計算結果を Fig. 5 (a), (b) に示した。これにより, グラディエント法ではマッチング法に比べかなり解析精度が悪く, 特に速度の大きい領域でその方向が強いことが分かる。これは, グラディエント法が基本的に画像の時間・空間的な連続性を前提としていることに拠ると考えられる。

c) 時空間相関法

この方法では, 全体として用いた画像枚数は 71 枚で, 計算時間は 10 分程度である。具体的な計算方法は, まず画像平面中のある解析対象画素 C の画素時系列と, 画素 C から水平方向へ左右にそれぞれ同一距離 D ($1 \sim 9$ (p)) 離れた画素 L, R での画素時系列間とのそれぞれの時空間相互相関関数を求める。その相互相関関数の計算は, 画素 C の時系列データに於いて 35 フレーム目を中心とした時間窓枠 F (サイズは可変で $5 \sim 15$ (f) とした) の画素時系列と, 点 R, L のそれぞれの同一時間窓枠を持つ画素時系列との間で行われる。その距離 D と時間窓枠下の具体的な関係は Table 1 に示している。また, 解析結果を Fig. 6 に示した。Fig. 6 (b) の頻度分布ヒストグラムより, 速度の絶対値の解析精度は十分高く, 本論文で提案した 3 つの手法の内で最も良いことがわかる。ただ, 速度ベクトル分布図 (Fig. 6 (a)) に於いては, 速度が大きくなるにつれて進行方向に物体が引き伸ばされて表示されてしまう。これ

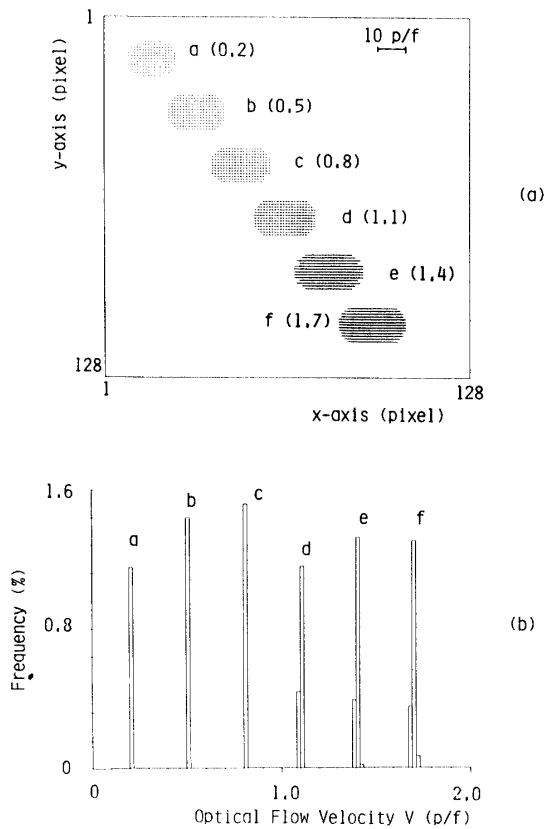


Fig. 6 Analyzed results obtained by the "spatio-temporal correlation method" for simulation image. (a) Velocity-field map, (b) Frequency histogram of the magnitude of the velocity

Table 1 Selection of the distance D between correlation pairs and the length of temporal window F at each pixels. D and F are determined according to the velocity of optical flow

Velocity $V(p/f)$	Distance $D(p)$	Temporal Window $F(f)$ size
~ 0.25	1	15
0.25 ~ 0.50	2	13
0.50 ~ 0.75	3	9
0.75 ~ 1.00	4	7
1.00 ~ 1.25	5	7
1.25 ~ 1.50	6	5
1.50 ~ 1.75	7	5
1.75 ~ 2.00	8	3
2.00 ~	9	3

は、高速で運動する物体ほど同一時間に描く軌跡が大きいことによる。Table 1の時間窓枠Fの設定においては、できるだけ軌跡長が一定となるよう、速度との関係を考慮しているが、高速側では相関解析に用いる時系列長が短くなり過ぎても解析精度上問題であることから制限される。このことは、3次元の奥行き分布を自動的に求めると言う観点からは、得られた光流動の場の空間分解能が、特に近距離側で悪くなることを意味する。

4. 実画像における検討

4.1 方法

本章では、実画像を用いて前章と同様の実験を行った。用いた動画画像は、背景が黒で画像中に静止し、それぞれ異なった奥行きにある白っぽい3つの物体(ディスクの箱、コップ、修正液の瓶)を、カメラを水

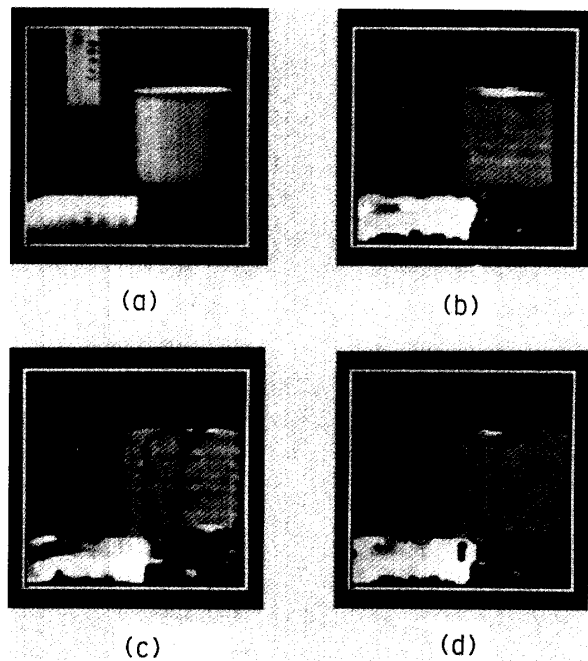


Fig. 7 (a) A snapshot of the analyzed real scene, (b) Analyzed depth map of the real scene obtained by the "matching method", (c) Depth map of the real scene analyzed by the "gradient method", (d) Depth map of the real scene analyzed by the "spatio-temporal correlation method". All depth maps are illustrated by 256 gray-values representation. White region corresponds to shallow depth.

平方方向に移動して撮影したものである。この動画画は1枚が128×128(p)の画像サイズを持ち、そのスナップをFig. 7 (a)に示している。また、ここで用いる画像は実画像であり、通常種々の原因よりなる雑音を多く含んでいる。雑音を軽減し解析精度を上げるために、以下に述べる前処理、中処理、後処理の3つの処理を行っている。

a) 前処理：ここでは、雑音を軽減するために3×3(p)からなる平滑化フィルターを掛ける。他にガウスフィルター、メディアンフィルター等¹⁾も知られている。しかし、ガウスフィルターはノイズの軽減率が低く、メディアンフィルターは背景と物体の境界が明確に現れ、マッチング法と時空間相関法に於いては問題がないがグラディエント法に於いては、“画像平面上の輝度は滑らかに変化する”という仮定に反する恐れがあるので、ここでは採用していない。

b) 中処理：時空間相関法では、空間的に1(p)、時間方向に5～15(f)からなる時空の窓、マッチング法では空間的に1×5(p)の窓、グラディエント法は空間的に1×3(p)の窓、というふうに3つの手法とも窓領域を設けて計算している。そこで、各窓領域に於て輝度の一番大きいものと一番小さいものの差を振幅(AMP)としたとき、1) AMPがある一定値を越えている場合は、計算した光流動の値の信頼性が高いとして、そのまま中心画素の光流動の値とする。2) AMPが一定値を越えない場合は、その近傍8画素の領域に於て計算した光流動の値の分散を計算し直し、さらにその分散がある一定値を越えない場合のみ、中心画素の光流動の値を採用する、という処理を導入した。また2)のケースで、分散が一定値を越えた場合は、その光流動の値は信用できない、あるいは不定であるとして0とおいた。

c) 後処理：この後処理は、計算した光流動の異常値を取り除く効果と共に、B)の中処理によっても決定できなかった、濃淡勾配の小さな物体上の光流動の値の未定の部分を、推定し補間する効果がある。そのアルゴリズムは次の3つのステップよりなる。

1) 3×3の小領域を考え、その小領域の中心画素を缺んだ対角線上の画素に於いて、両方に光流動の値を持つもの(速度が未定でない画素)の存在を確認する。存在しなければ次の小領域を考察し、存在すれば次のステップ2)に進む。

2) ステップ1)の小領域に於いて各画素の持つ光流動の値を、大きいものから順に並べ、中央値付近の値を平均(速度が未定の場合を除く)し、その値を中心画素の値として置き換える(一種のメディアンフィル

ター)。

3) ステップ2)の中心画素の値と元の値との差が一定値を越えたとき、その中心画素を含む全ての小領域(9つ)に於てもう一度ステップ1)から計算を繰り返す。以上の操作により、解析した光流動の場を滑らかにすることができる。

4.2 結果

解析結果の表示方法は各方法とも速度の頻度分布を示すヒストグラムと、速度が大きくなるにつれて高輝度を持つように、256段階で濃淡画像表示された速度ベクトル分布図を2つの表現法で示している。

a) マッチング法

使用した画像枚数は3章と同じく動画中の2枚を用い、計算時間は10分程度である。解析結果を、光流動の速度ベクトル分布図としてFig. 7 (b)に、また頻

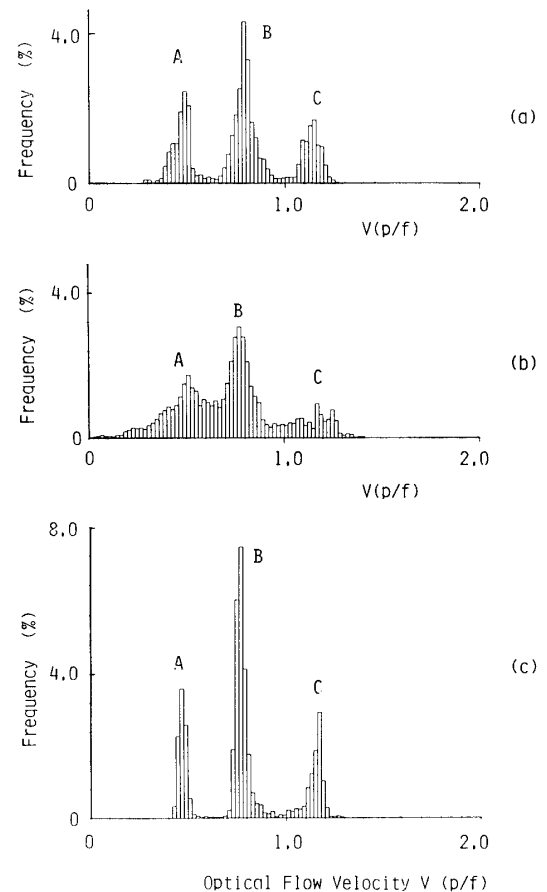


Fig. 8 Frequency histograms of the optical flow velocity obtained by (a) "matching method", (b) "gradient method", and (c) "spatio-temporal correlation method"

度分布ヒストグラムとして Fig. 8 (a) に示した, Fig. 8 (a) の頻度分布ヒストグラムでは, 3つの物体の光流動がほぼ分離した分布を示しており, 3つの異なる奥行きを持つ対象の存在が推定される. また, Fig. 7 (b) の速度ベクトル分布図を見ると, シミュレーション動画の例と同様に, 物体の形状をほぼ反映したベクトル分布が示されており, この方法の3次元奥行き分布解析に於ける空間分解能が良いことが認識できる.

b) グラディエント法

使用した画像枚数は, シミュレーションの場合と同様に10枚で, 計算時間は5分程度である. 結果を Fig. 7 (c) と Fig. 8 (b) に示した. Fig 8 (b) の頻度分布ヒストグラムより, 3つの物体の速度の分離が悪く, 解析精度が相対的に低い. Fig. 7 (c) の速度分布図の中で物体形状 (特に物体の内部) が欠けて見える部分は, 4. 1 節で述べた中処理, 後処理によっても値が特定できない部分であるが, 画像の性質から画像輝度の時間的, 空間的グラディエントがこの領域で小さいことに起因していると考えられる.

c) 時空間相関法

使用した全画像枚数は31枚で, 計算時間は17分程度である. 解析結果を Fig. 7 (d) と Fig. 8 (c) に示した. Fig. 8 (c) の頻度分布ヒストグラムでは, 3つの物体の光流動の大きさの違い, すなわち奥行きの違いが, 明確に示されている. Fig. 7 (d) の速度ベクトル分布図に於いても, 同一物体上にはほぼ同一濃淡値が付けられており, 解析精度の高さが確認できる. しかし, 解析法の性質上やむを得ないが, 物体は水平方向 (カメラの運動方向) に引きのばされた形を示しており, マッチング法に比べて奥行き分布図としての空間分解能は劣っていると言えよう.

5. 議 論

まず, 時空間相関法について述べると, 空間分解能が特に高速運動する対象に対して悪くなるのは, マッチング法のように空間的に相関をとるのではなく, 時間方向で相関をとることが原因となっている. 得られる光流動の場が物体形状をより正確に反映するためには, 時間方向の相関を出来るだけ短い時間に対して計算することが要求される. このことは精度の面からはマイナスであり動画のサンプリング速度が変えられない場合には, 適当な空間分解能と時間分解能及び解析精度の妥協点を見い出すべきであろう. この時空間相関法は, 我々の研究室で5年前より提案してきた方法であり^{8-10,17)}, これからまだ改良の余地があると考え

られる. 例えば, 相関を取る相手の画素との距離 D と時間窓枠 F の関係を最適化し, 精度や空間分解能を改善できる可能性があると思われる.

また, マッチング法に於いても, 最終目標を3次元の奥行きの空間分布を計測するという観点からすれば, まだ十分手法として確立しているとは言えない. マッチング法では, 相関の取り方, 相関関数の評価法を見直すことや, 階層処理的なマッチングを実行することで, 解析精度の向上や解析法のロバストネスを改良できる可能性はある.

グラディエント法に関しては, 滑らかな模様を持った球がゆっくり回転しているような動画や, カメラの回転により相対的に画像全体が回転しているような動画の解析には, 比較的高い精度を示すが¹²⁾, 今回ここで取り扱ったような輝度変化が激しく, 物体の運動速度が速い動画の解析では, 精度のよい結果を得るのは困難のようである. 今後, 計算に用いる閉鎖域 δC の大きさ, 形状などを工夫することにより, 解析精度の向上を計れる余地はあると考えられる.

また, ここでは取り上げなかったがオクルージョン (物体が別の物体を後に隠す状態) の問題については, より自然な実験室のシーンを用いて予備的な解析を行った結果, 時空間相関法がマッチング法より空間分解能に於いても優れていることを確認している. この理由としては, まずマッチング法では, 2枚の画像のみで空間的に相関 (対応) をとるため特に境界部分の左右の対応が困難となってくることから結果の信頼性が悪くなると考えられる. 一方, 時空間相関法では時間的に相関をとり, 対応を連続的な変化の中で捉えているため, マッチング法より奥行き境界部分の解析精度及び信頼性が高いものと考えられる.

6. ま と め

本研究で明らかとなったことは, 奥行き空間分布の計測を目指し提案した3つの実用的なオプティカルフロー解析法について, 解析精度的には時空間相関法が最適であり, また空間分解能の点から言えばマッチング法が最適であることである. グラディエント法については, 先に述べたようにここで用いた動画のように輝度の時間, 空間変化の比較的大きな画像に対しては不利のようである.

これかいらの課題としては,

- 1) グラディエント法の解析精度の改良
- 2) 時空間相関法の距離 D と時間窓枠 F の関係の最適化と, 精度を保ちながらの解析に必要な画像枚数の削

減

3) マッチング法の解析精度の向上

などが考えられる。特に、ここで提案したグラディエント法が、他の手法と同等の解析精度を持つようになれば、アルゴリズムの単純さ、計算コストの面から有力な手法となることが期待される。

今後、それぞれの手法の特徴を活かしながら、実用的な解析法を確立していく必要がある。

謝 辞

本研究に関し、助力頂いた河村好子氏及び当研究室の諸氏に謝意を表す。また、本研究の一部は、大川情報通信基金の援助を頂いた。

参考文献

- 1) 江尻正員：工業用画像処理（昭晃堂，1988）
- 2) 金谷健一：画像理解—三次元認識の数理—（森北出版，1990）
- 3) S. T. Barnard and W. B. Thompson, IEEE Trans. Pattern Anal. Mach. Intell., **PAMI-2**, (1980) 333.
- 4) F. Glazer, G. Reynolds and P. Anandan, Proc. IEEE (1983) 432
- 5) B. K. P. Horn and b. Schunck, Artif. Intell., **17**, (1981) 205
- 6) H. -H. Nagel and W. Enkelmann, IEEE Trans. Pattern Anal. Mach. Intell., **PAMI-8**, (1986) 565
- 7) J. K. Kerney, W. B. Thompson and D. L. Boley, IEEE Trans. Pattern Ana. Mach. Intell., **PAMI-9**, (1987) 229
- 8) H. Miike, K. Kurihara, H. Hashimoto and K. Koga, Trans. IEICE Japan **E69**, (1986) 877
- 9) 古賀和利, 三池秀敏, 電子情報通信学会論文誌, **J72D**, (1989) 507
- 10) K. Koga and H. Miike, Proc. ISSSE' 89, (1989) 517
- 11) デビッド・マー：ビジョン（産業図書：乾，安藤訳，1987）
- 12) A, Nomura, H. Miike and K. Koga, Pattern Recogn. Letter: in press.
- 13) B. G. Schunck, Proc. Workshop Motion : Representation and Analysis, Charleston S. C., (1986) 88
- 14) J. M. Fitzpartrick, Proc. IEEE Conf. Comput. Vision Pattern Recogn., (1985) 78
- 15) H. -H. Nagel, IEEE Trans. Pattern Anal. Mach. Intell., **11**, (1989) 12
- 16) 服部哲郎, 中田 誠, 久保克巳：第19回画像工学コンファレンス論文集, (1988) 157
- 17) 野村厚志, 古賀和利, 三池秀敏：電子情報通信学会論文誌, **J73D-2**, (1990) 728

(平成2年10月15日受理)