

ステレオ視差パターンの統計学習に基づく移動ロボット視覚

北川 景介[†] 福井 和広[†][†] 筑波大学大学院システム情報工学研究科コンピュータサイエンス専攻

〒 305-8573 茨城県つくば市天王台 1-1-1

E-mail: †kitagawa@viplab.is.tsukuba.ac.jp, ††kfukui@cs.tsukuba.ac.jp

あらまし 本論文では、移動ロボットの視覚を想定したパターン認識によるステレオ視覚を提案する。従来の“計測的な”ステレオ視覚を移動ロボット視覚に用いる場合、正確な3次元形状を復元した後、その情報から「その空間に移動することが可能か」等のロボットに必要な判断を行う。この処理過程における3次元形状復元は点の対応付けの難しさ等により現時点では十分な安定性や精度が得られているとは言えない。そこで本論文では、ステレオ画像の入力から3次元計測を介さず直接上記の判断結果を与える、つまり、「注目する範囲内の最も近い点がどれだけ遠いか」で決まるクラスを出力する“パターン認識的な”ステレオ視覚の実現を目指す。

キーワード 統計学習, ステレオ視覚, ロボット視覚, 視差パターン, カーネル判別分析, 3次元シーン生成

Robot Vision Based on Statistical Learning of Stereo Parallax Patterns

Keisuke KITAGAWA[†] and Kazuhiro FUKUI[†][†] Graduate School of Systems and Information Engineering, University of Tsukuba

Tennodai 1-1-1, Tsukuba-shi, Ibaraki, 305-8573 Japan

E-mail: †kitagawa@viplab.is.tsukuba.ac.jp, ††kfukui@cs.tsukuba.ac.jp

Abstract In this paper, we introduce a framework of a new stereo vision system based on statistical learning of stereo patterns for mobile robot. Robot vision systems with the conventional stereo visions firstly gather 3D information of their environments, and then judge whether to go or not, from the 3D information. However, the gathering 3D information is generally unstable due to the difficulty of stereo matching between left and right images. In order to overcome this problem, we propose a method of obtaining the above judgment by using the pattern recognition of stereo image patterns without the stereo matching process. This means that the proposed method can output a distance class of the nearest point from a baseline even if the 3D information is not reconstructed.

Key words Statistical learning, Stereo vision, Robot vision, Parallax pattern, Kernel discriminant analysis, 3D scene generation

1. はじめに

現在、ロボットの研究が盛んであり、自分で判断し行動するロボットを世界中の研究者が創ろうとしている。その実現に欠かせない最重要課題の一つに空間に関する情報の取得がある。

「空間の情報を得る」、つまりこの場合は、「自分の周囲の形状を知る」手法の代表的なものとして、“計測的な”ステレオ視覚がある [1]。“計測的な”ステレオ視覚では異なる画像間での対応付けが重要であり、そのための様々な手法が提案されている。

ここでステレオ視覚の使用目的をロボットの移動に限ると、正確な距離情報が測れれば移動ができるが、移動を行うために正確な距離計測が必要なのかという疑問が生まれる。移動ロ

ボットの視覚は「進むことが可能か」や「動作に必要な空間が確保できているのか」などの問いに答えられれば十分であると考えるならば、必ずしも“計測的な”ステレオ視覚は必要ない。

本論文ではこのように一般的なステレオ視覚とは異なる視点に立ち、ステレオ画像の入力から直接上記の判断結果を与える“パターン認識的な”ステレオ視覚を提案し、動作・移動が行える空間が存在しているのかというロボットに有用な情報の取得を目指し、有効な特徴抽出・識別法を提案する。

以下、第2章で“パターン認識的な”ステレオ視覚である提案手法のフレームワークとその実現への課題を、第3章で提案手法の詳細を述べ、第4章で実験を行い、その考察を第5章で、結論と今後の課題を第6章で述べる。

2. “パターン認識的な”ステレオ視覚

2.1 提案手法のフレームワーク

「進むことが可能か」を考えると、これはステレオ画像から得られるパターンを、「進むことが可能」か「進むことが不可能」に分類するという2クラスの識別問題と捉えることができる。このパターン識別に基づく提案手法のフレームワークを図1に示す。

まず、ステレオカメラから得られる左右の輝度画像(図2)と画像内で最も近い点が属する距離クラス(図4)の対応関係を、大量の左右画像と距離情報(図3)のセットから統計学習する。識別時には、入力された左右の輝度画像に対して、先に学習した識別器を用いて画像中で自分自身に最も近い点がどの距離クラスにあるかを識別する。

2.2 提案手法実現への課題

提案手法を実現するためには以下の課題を解決する必要がある。

(1) 視差情報を残したパターン圧縮

識別に有効な情報は、被写体までの距離の情報、つまり、ステレオ画像では視差の情報である。しかし、ステレオ画像には視差以外の情報が大量に含まれているため、識別に貢献しない情報を含む大きなデータを扱うことになり、パターン認識の負担を増している。そこで、データのサイズを小さくしつつ、視差情報を保存するようなパターン圧縮を前処理として考える。

(2) 複雑なクラス識別境界

左右画像に含まれる対象物は多種多様であり、各クラスの境界は極めて複雑であると予想される。したがって、線形的手法[2]では識別が困難であり、非線形識別の導入が必要と考えられる。

(3) 学習データ取得の困難さ

統計的に学習するためには、正解ラベルが付いているサンプルが大量に必要となる。しかし、実環境においては、距離情報のラベルが付いた左右画像の大量取得は極めて困難であり、何らかの方法に基づいて学習サンプルを生成する必要がある。

3. 提案手法の詳細

以下、前述の課題に対する解決策を挙げることで提案手法の詳細を述べる(解決策の番号は課題の番号と対応している)。

(1) 前処理(パターン圧縮)

基本的には、不必要な情報を含むデータであっても識別器で視差情報の抽出が行われるが、今回はパターン認識の負担を軽減するために、パターン圧縮を行う。さらに、画像全体を複数の小画像に分割し、各小画像に対し上記の統計学習を行う。

この前処理によって作られたデータを視差パターンと呼ぶ。入力された画像内に複数の小画像を設定し、そのそれぞれに対し視差パターンの算出を行う。本手法では、この視差情報の濃度を高くしたパターンを識別することにより、その領域内の最も近い点によって定まるクラスを得ることを考える。この前処理をどのようなものとするかは重要であるが、今回は暫定的に後述の処理とし、実験を行った。

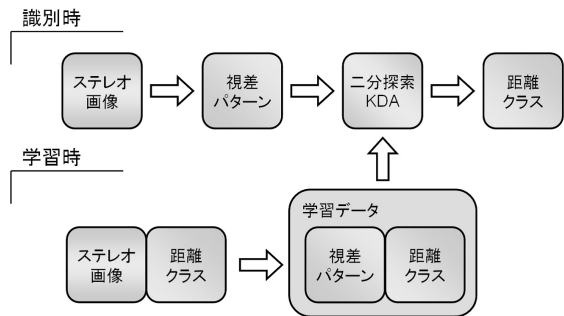


図1 提案手法のフレームワーク



図2 入力左右輝度画像例



図3 入力距離情報例

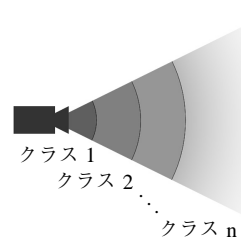


図4 クラス定義

以下、設定されている小画像に対する前処理を3つの部分処理に分けて述べた後、小画像の設定について述べる。以降、その時に注目している小画像を基準小画像、入力左右画像のうち基準小画像を含まないものを比較画像と呼ぶ。

i) 処理A: 平滑化

まず、入力された左右画像を基準小画像より小さい局所領域の窓で走査し、局所領域の輝度総和をとることで平滑化を行う。以降、左右画像として平滑化されたものを用いる。

ii) 処理B: SAD算出

比較画像内に基準小画像に対応する小画像(基準小画像の領域を視差方向に g 画素移動したもの)を設定する。

この2つの小画像間で、相違度としてSADを計算する。 $X(n)$, $Y(n)$ をそれぞれ基準小画像と比較画像内の小画像の n 番目の画素の値、 M を小画像と同じ大きさの領域の画素番号とすると、次式ようになる。

$$SAD = \sum_{n \in M} |X(n) - Y(n)| \quad (1)$$

上記の処理を十分に大きい範囲の g 全てに対し計算することで、それらを要素とするベクトルを得る。

iii) 処理C: 正規化

処理Bで得られたベクトルの大きさを1に正規化する。

iv) 小画像の設定

小画像の高さは処理Aの局所領域の高さに、小画像の幅は平滑化後の画像の幅の半分とした。よって、小画像は図5のよう

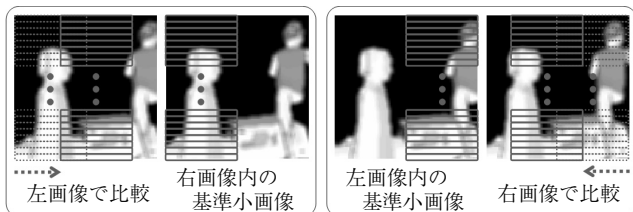


図5 小画像の設定と処理 B

な横に長い長方形である。小画像の高さを処理 A の局所領域の高さとしたのは、その高さが処理 A 後の平滑化画像にとって 1 行に相当する最小の高さであり、クラス内分散を抑えられると考えたためである。小画像の高さを小さくすることが可能なのは、ステレオカメラが平行等位に配置されている場合はエピポーラ線が水平となり [3]、画像を行列としてみたときの行ベクトルは全て独立したサンプルとなるためである。小画像の幅を平滑化後の画像の幅の半分としたのは、その大きさの小画像を図 5 のように、平滑化後の左画像の右側と右画像の左側に 1 行ずつ設定することによって、全ての領域が処理 B の g を考慮した上で視差方向の比較領域を確保できるためである。また、これにより「左右どちらに移動すべきか」というロボットに有用な情報も得られると考えられる。

右画像の左側を距離画像の左側に、左画像の右側を距離画像の右側に対応させ、それぞれの小画像領域内の最も近い点までの距離を距離画像より求め、その小画像の正解ラベル（距離クラス）を定めた。

(2) 二分探索 KDA

複雑な識別境界に対応できるよう、非線形的手法であるカーネル判別分析 (KDA) [4] [5] を用いる。

必要な空間の基準がある場合は全ての問題を注目境界の前後に対する 2 クラス問題と見ることでもできるが、あらかじめ注目境界が定まっていない場合を考慮し、対象としている問題を多クラスの識別とする。多クラス KDA は多くのクラスの中から 1 つのクラスを選択するのに有効な判別空間を得ることができるが、今回の問題では必ずしもその必要はないと考える。本手法では多クラス KDA を用いるのではなく、二分探索の要領で 2 クラス KDA を繰り返すことにより多クラスの識別を実現しており、今回の問題に対して有効となる、必要な情報が得られた時点で残りの処理を放棄できるという特徴があり、精度を選択することによってコストを調節できる。

この方法では学習・識別ともに、1 回目の識別からサンプルが変わらず減るのみであり、2 回目以降の識別のカーネル行列は 1 つ前の識別で使用したカーネル行列の小行列となるため、カーネル非線形手法で問題となるカーネル行列の計算時間が二分探索内の識別回数に比例して増加するということはない。

図 6 は 8 クラスの問題 (図 4 における $n = 8$) の例であり、学習用のサンプルで学習した最初の 2 クラス問題の判別軸に識別用の 8 クラスのサンプルを射影したときのヒストグラムである。この例では距離の近い 4 つのクラスと距離の遠い 4 つのクラスが判別できるような軸が選ばれており、パターンの中の視

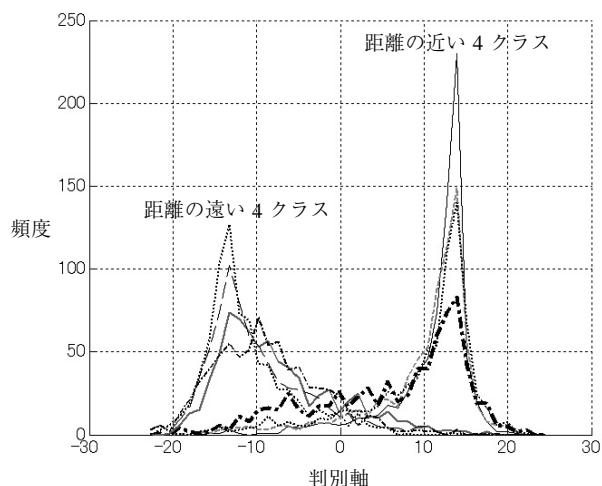


図6 二分探索 KDA の例

差情報に注目した軸であると考えられる。

(3) CG による 3 次元シーン生成

大量のサンプルを得るために、本手法では OpenGL を用いて 3 次元シーンを生成し、ステレオカメラの画像とそこに映るシーンの距離情報を取得した。

3D モデルをいくつか集めたモデルセットを用意し、その中からランダムにモデルを選択・配置することでシーンの生成を行う。

まず、ランダムに複数のモデルを選択し、そのそれぞれに対し、カメラに映る範囲内でランダムにモデルの位置を設定する。(ステレオカメラのカメラ間距離を 0.25 m、画角を 60 度とし、図 7 のような左右カメラの midpoint からの奥行き方向を 0.5 m から 20 m の範囲、上下左右方向は左右カメラの midpoint から広がる画角 60 度の四角錐の範囲をカメラに映る範囲として用いた。) 次に、それぞれのモデルの水平方向の向きをランダムに設定し、配置が完了する。ここで、透視投影であることから、モデルの位置がカメラから遠くなるにつれて画像上での大きさが小さくなるという現象が生じる。モデルの大きさの初期値を正規化しているため、このままではこの情報から距離がわかってしまい、CG による実験が手法の評価に不適切になってしまう。それを防ぐためにここで、距離による画像上の大きさの変化を打ち消すようにモデルの大きさを変更し、画像上の大きさを距離によらず一定にしている (図 8)。

このときの左右のカメラの画像と左右カメラの midpoint から測った距離画像を出力する。

4. 実験

提案手法の性能、特に正しく視差情報が使われているかを調べるために、4 つの実験を行った。

まず、全ての実験で共通する、サンプル、実験方法・評価方法について述べた後、各実験の詳細を述べる。

4.1 サンプル生成

今回はモデルセットを 2 つ用意した。モデルセット M_α は車と人の 3 次元モデル、計 42 個から構成され、モデルセット M_β

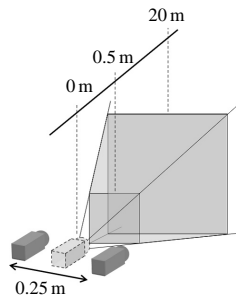


図7 モデルの配置範囲

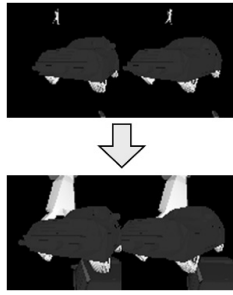


図8 モデルの大きさ

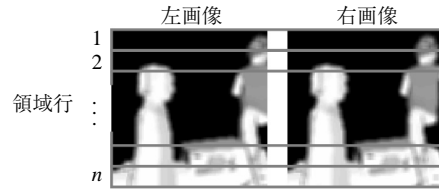


図9 領域行

は道路周辺にあるものと海の生物の3次元モデル、計54個から構成されている。

1つのモデルセットに対し、同時にシーン内に入るモデル数を3から7まで変化させ、そのそれぞれにおいて1500シーンを生成する。これにより、そのモデルセットを用いた7500シーンに対する左右画像及び距離画像のサンプルセット1つを得る。この処理を、モデルセット M_α において乱数を変えて2回行い、サンプルセット $S_{\alpha 1}$, $S_{\alpha 2}$ を得た。また、モデルセット M_β において1回行い、サンプルセット S_β を得た。サンプルセット $S_{\alpha 1}$, $S_{\alpha 2}$ を合わせたものは、同時にシーン内に入るモデル数を3から7まで変化させ、そのそれぞれにおいて3000シーンを生成したものに相当する。よって、サンプルセット $S_{\alpha 1}$, $S_{\alpha 2}$ を合わせたものは、サンプルセット S_β の倍の大きさのサンプルセット S_α として扱う。

なお、左画像、右画像、距離画像はどれも高さ・幅ともに100画素である。

4.2 実験方法・評価方法

まず、クラスは手前から0.5mごとに3.5mまで区切り、クラス1, クラス2, ..., クラス7とし、それより遠いものをクラス8とした。

KDAのカーネル関数にはガウスクーネルを用い、そのパラメータ $2\sigma^2$ と正則化パラメータ λ は、予備実験により経験的に求めた $2\sigma^2 = 0.03$, $\lambda = 0.1$ という値を全実験共通で使うが、その値が最適なものとは限らない。

図9のように入力画像を n 個の領域に分割し、上下方向の画素単位的位置を行と呼ぶのに対し、この領域単位的位置を領域行と呼ぶ。これは、画像の行方向に特性の異なるサンプルで実験・評価を行うための設定であり、提案手法に由来するものではない。小画像1つに対しベクトルが1本なので小画像の集まりに対し実験を行うことになるが、その小画像の集まりを、全入力画像の同一領域行内の小画像とし、その集まりごとに学習・識別を行う。 $n = 1$ とすれば全サンプルをまとめて実験していることになる。全ての領域行の大きさが等しくなることを優先し、画像の行数が領域行数で割り切れない場合は余りとなる行を実験に使用しない。今回の実験ではこの n の値は全て共通で $n = 8$ とした。

サンプルにはサンプルセットのいずれかを用いるが、クラスによってサンプル数に大きな偏りがあるため、1クラスあたりのサンプル数を定め、サンプルセット内からランダムに抽出する。抽出されたものは2等分され、半分が学習用、半分が識別

実験用となる。この1クラスあたりのサンプル数は実験によって異なる。

視差パターン算出時の処理A：平滑化で使用する局所領域の大きさは 4×4 画素とする。従って、小画像の高さも4画素となる。

また、処理B：SAD算出で使用する g は、全実験で共通して0から45画素の範囲とするが、 $g = 0$ での値は識別を行う前に一度使用する。なぜなら、この値が0に近いということは視差が現れないほど映っているものまでの距離が大きいということを示しており、KDA等を行うまでもなくクラスがわかるためである。この実験では背景が単色のCG画像を用いるため、 $g = 0$ での値が0となるものだけをあらかじめクラス8として扱う。これを識別に成功したサンプルとして扱うと識別率が上がったように見えてしまうので、識別率はそのようなサンプルが初めから無かったものとして算出する。

評価には通常、

$$\frac{\text{正しいクラスに識別されたサンプル数}}{\text{全サンプル数}} \times 100 \quad [\%] \quad (2)$$

という識別率を用いるが、ここではこの識別率を純識別率と呼び、本実験ではもう一つ指標を設ける。

$$\frac{\text{正しいクラスに識別されたサンプル数} + \text{正しいクラスの隣のクラスに識別されたサンプル数}}{\text{全サンプル数}} \times 100 \quad [\%] \quad (3)$$

式(3)の識別率を緩和識別率と呼ぶ。これは画像上での視差がクラスの境界で変化するとは限らず、隣のクラスとは重なりを持ってしまうためである。

4.3 実験1：提案手法

提案手法の性能を測る実験を行う。サンプルセットには S_α を用い、1クラスあたりのサンプル数は3000とした。

この実験では左右・領域行による特性の違いがないため、左右・領域行毎の結果を統合した識別率を表1に示す。

4.4 実験2：モデルの大きさの変化に対する実験

シーンの生成時に、モデルの位置がカメラから遠くなるにつれて画像上での大きさが小さくなるという現象が生じ、それを打ち消すようにモデルの大きさを変更するという処理を行っていた。これは、背景が単色のCGであることとモデルの大きさの初期値を正規化していることから、被写体の大きさの情報が距離情報を与えてしまい、学習時に視差ではなく被写体の大きさに注目してしまうことを避けるためであった。

モデルの大きさに対する処理がパターンに与える影響を調べるために、モデルセット M_α を用いて、同時にシーン内に入るモデル数を3から7まで変化させ、そのそれぞれにおいて300

シーンを上記の処理を行わずに生成することで、モデルが遠くにあるほど小さく映っているサンプルセット $S_{\alpha N}$ を得た。

処理を行わなかったサンプル $S_{\alpha N}$ 、処理を行った通常のサンプル $S_{\alpha 1}$ を用いて、1クラスあたりのサンプル数を 300 として実験を行った。

$S_{\alpha N}$ のサンプルを用いた学習成果で、他の $S_{\alpha N}$ のサンプル、および、 $S_{\alpha 1}$ のサンプルを識別した場合の識別率、並びに、 $S_{\alpha 1}$ のサンプルを用いた学習成果で、他の $S_{\alpha 1}$ のサンプル、および、 $S_{\alpha N}$ のサンプルを識別した場合の識別率を表 1 に示す（括弧内の数字は識別サンプルが学習サンプルと同じサンプルセットからとられた場合を 1 とした比である）。

なお、上記 4 つの実験で用いられた $S_{\alpha N}$ の識別サンプル、および、 $S_{\alpha 1}$ の識別サンプルは、それぞれ $S_{\alpha N}$ 、 $S_{\alpha 1}$ からランダムに選ばれているが、実験ごとに選びなおすことはしていない。

4.5 実験 3：未学習のモデルを用いた画像に対する識別

学習時と識別時で異なるモデルのサンプルセットを使い、未学習のモデルに対する識別を行う。

サンプルセットには $S_{\alpha 1}$ と S_{β} を用い、1クラスあたりのサンプル数は 1500 とした。

$S_{\alpha 1}$ のサンプルを用いた学習成果で、他の $S_{\alpha 1}$ のサンプル、および、 S_{β} のサンプルを識別した場合の識別率、並びに、 S_{β} のサンプルを用いた学習成果で、他の S_{β} のサンプル、および、 $S_{\alpha 1}$ のサンプルを識別した場合の識別率を表 1 に示す（括弧内の数字は識別サンプルが学習サンプルと同じサンプルセットからとられた場合を 1 とした比である）。

なお、上記 4 つの実験で用いられた $S_{\alpha 1}$ の識別サンプル、および、 S_{β} の識別サンプルは、それぞれ $S_{\alpha 1}$ 、 S_{β} からランダムに選ばれているが、実験ごとに選びなおすことはしていない。

4.6 実験 4：未学習の歪みを持った画像に対する識別

学習時と識別時で異なる歪みのサンプルセットを使い、画像上での視差の大きさの変化が出力に与える影響を調べる。

レンズの歪みのモデルとしてよく知られる次式 (x , y は歪みの無い画像での座標、 x_d , y_d は歪んだ画像での座標、 k は歪み係数) の r の 2 次の項までを使用して歪み (歪み A と呼ぶ) を生成した (座標原点は画像中央とし、補間には共一次内挿法を用いた)。

$$x_d = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots) \quad (4)$$

$$y_d = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots) \quad (5)$$

$$r = \sqrt{x^2 + y^2} \quad (6)$$

ただし、距離画像は歪みの無いものを共通で使用するため、できる限り同じ範囲が映るよう x_d , y_d の最大値が x , y の最大値と一致するように拡大縮小を行った。また、上記変換の逆変換を用いて別の歪み (歪み B と呼ぶ) も生成した。こちらにも同様の拡大縮小を適用した。

サンプルセット $S_{\alpha 1}$ に対し、歪み A と歪み B の変換を施し、そのそれぞれにおいて歪み係数 k が 0.1 と 0.2 の場合の、計 4 種類の歪みサンプルセットを用意し、それらに $S_{\alpha 1}$ を加えた計 5 セットを用い、1クラスあたりのサンプル数を 1500 として実験を行った。

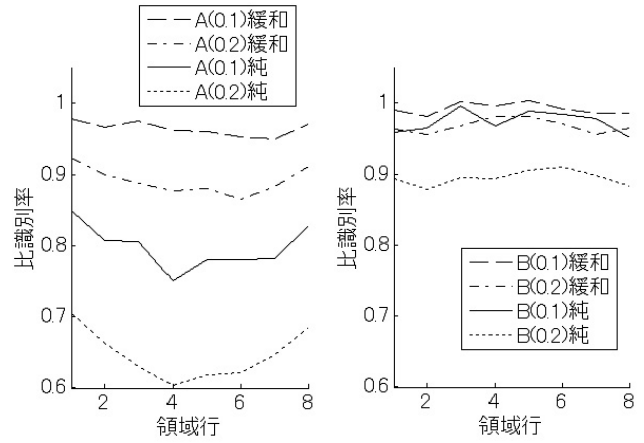


図 10 歪み A に対する領域行ごとの比識別率

図 11 歪み B に対する領域行ごとの比識別率

表 1 実験結果

実験	比較 学習 : 識別	単位 : [%]		
		純識別率	緩和識別率	
1		61.49	86.53	
2	$S_{\alpha N}$	$S_{\alpha N}$	77.75 (100.0)	93.10 (100.0)
		$S_{\alpha 1}$	44.41 (57.1)	75.15 (80.7)
	$S_{\alpha 1}$	$S_{\alpha 1}$	51.48 (100.0)	81.40 (100.0)
		$S_{\alpha N}$	54.66 (106.2)	84.86 (104.3)
3	$S_{\alpha 1}$	$S_{\alpha 1}$	61.93 (100.0)	86.11 (100.0)
		S_{β}	45.42 (73.4)	76.39 (88.7)
	S_{β}	S_{β}	56.74 (100.0)	79.57 (100.0)
		$S_{\alpha 1}$	48.88 (86.2)	79.73 (100.2)
4	$S_{\alpha 1}$	$S_{\alpha 1}$	61.93 (100.0)	86.11 (100.0)
		$A_{0.1}$	60.24 (97.3)	85.39 (99.2)
		$A_{0.2}$	55.37 (89.4)	83.30 (96.7)
		$B_{0.1}$	49.46 (79.9)	83.05 (96.5)
		$B_{0.2}$	40.13 (64.8)	76.73 (89.1)

$S_{\alpha 1}$ のサンプルで学習し、他の $S_{\alpha 1}$ のサンプルを識別した場合の識別率 (実験 3 の実験結果) を表 1 に示す。同じ学習成果を用いて、歪み係数 $k = 0.1$ の歪み A サンプル、歪み係数 $k = 0.2$ の歪み A サンプル、歪み係数 $k = 0.1$ の歪み B サンプル、歪み係数 $k = 0.2$ の歪み B サンプルを識別した場合の識別率を表 1 に示す (括弧内の数字は $S_{\alpha 1}$ 学習による $S_{\alpha 1}$ に対する識別率を 1 とした比である)。

また、 $S_{\alpha 1}$ 学習による $S_{\alpha 1}$ に対する識別率を 1 とした場合の、各歪みサンプルに対する領域行ごとの比識別率を図 10, 図 11 に示す。

なお、上記 4 つの実験で用いられた $S_{\alpha 1}$ による学習・識別結果は同一のものであるが、識別で用いた歪みサンプルが同じ $S_{\alpha 1}$ サンプルから得られたものとは限らない。

5. 考 察

5.1 実験 1 に対して

サンプル純識別率が約 6 割とあまり良い識別率が得られていない。これには大きく 3 つの原因が考えられる。1 つ目は、視差情報の抽出が不十分であり、被写体に依存する情報が識別を困難にしたということである。2 つ目は、輝度画像における小

画像領域の設定、および、その領域の距離画像との対応が適当でなかったということである。3つ目は、二分探索 KDA においてパターン分布が異なるにもかかわらず全ての識別に同じパラメータを用いていることである。

5.2 実験 2 に対して

$S_{\alpha N}$ による学習成果で $S_{\alpha 1}$ を識別すると識別率が大きく下がることから、画像上のモデルの大きさに距離情報を持たせたまま学習サンプルとして使用することが不適切であると確認できた。しかし、これは CG を用いて生成した今回のサンプルの問題であり、視差情報への注目が不十分なために識別器が視差以外の情報を使ってしまったとも言える。今回の視差以外の情報の排除は、前処理だけでなく学習サンプルによっても行われていたと考えられる。

また、 $S_{\alpha 1}$ による学習成果で $S_{\alpha N}$ を識別しても識別率が下がらないことから、 $S_{\alpha 1}$ による学習成果は画像上のモデルの大きさによらない視差情報を用いて識別していると考えられる。さらに、モデルの大きさの調整が主に、遠くなったことで小さく映ってしまわないようモデルを拡大する処理であったことから、 $S_{\alpha 1}$ 学習による $S_{\alpha N}$ に対する識別率が $S_{\alpha 1}$ 学習による $S_{\alpha 1}$ に対する識別率より高い識別率が得られていることに関しては、モデルの拡大処理を行わないことで画像上でモデルが占める面積が小さくなり、SAD の反応が正しい視差に集中したためと考えられる。このことにより、 $S_{\alpha N}$ サンプルの識別は $S_{\alpha 1}$ サンプルの識別より易しい問題と考えられる。

5.3 実験 3 に対して

未学習のモデルセットを用いたサンプルに対しては識別率が低下するという結果が得られた。これは前処理が不十分であり、視差パターンの中に被写体に依存する情報が多く残っていたためと考えられる。しかし、未学習のモデルに対しても学習モデルの 7 割以上の識別率で識別できたということから、被写体に依存しない視差情報の抽出が行われたことも確認できる。

また、 $S_{\alpha 1}$ と S_{β} で識別率の低下に差が出たことは、それぞれのサンプルセットが持つパターン分布の違いによるものと考えられる。

5.4 実験 4 に対して

まず、歪み A に対しては画像中央部より画像上部・下部の純識別率低下が大きい。これは、歪み A の画像は歪みの無い画像の図 12 に示す点の輝度値を用いて作成されており、画像中央部は歪みの無い画像に近く、上部・下部は歪みの無い画像との違いが大きいためと考えられる。 k の値が大きくなれば歪みが大きくなり識別率は低下している。

歪み B も同様に、歪みの無い画像の図 13 に示す点の輝度値を用いて作成されていることから、画像中央部と上部・下部では純識別率の低下に差が出たと考えられる。歪み A では変化の大きい場所が画像の四隅であったのに対し、歪み B では中央行・中央列の変化が大きかったため、全ての領域行で歪みの影響を受け、歪み A より全体的に識別率が低くなっている。

純識別率は領域行によって識別率の低下に差が見られるのに対し、緩和識別率は領域行による差があまり見られない。このことから、正しいクラスの隣のクラスへの誤識別が生じた、つ

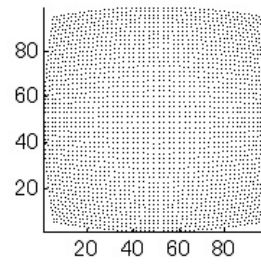


図 12 歪み A の参照点
($k = 0.2$)

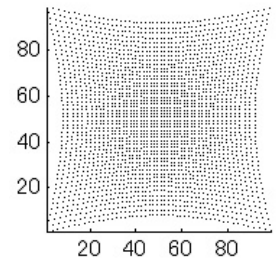


図 13 歪み B の参照点
($k = 0.2$)

まり、入力画像から読み取った視差の大きさにわずかな変化が生じた、ということがわかり、これらの識別率の低下は視差情報を用いて識別を行った結果であると考えられる。

6. 結論と今後の課題

本論文では、得られる情報の使用目的を限定することでステレオ視覚をパターン認識の問題として捉え、特徴抽出・識別法を提案した。いくつかの実験により、視差パターンに視差情報が含まれること、並びに、その情報を識別器により読み取れることを確かめた。

今後の課題としては、まず、前処理を改善し視差パターンをより識別に適したものにすることが挙げられる。そのために、サンプル生成も今回用いた実際の状況を再現するようなものではなく、より直接的に生成し、効率よく正確なサンプルを大量に得ることが考えられる。次に、KDA を行うにあたってカーネル行列に使われるサンプル数を削減する必要がある。さらには KDA 以外の識別法に変更することも考えられる。また、2 眼のステレオ視では原理的に視差を検出できない入力が存在してしまうため、カメラを増やすことも検討しなければならない。

文 献

- [1] 奥富正敏, 金出武雄, “統計的モデルに基づく適応型ウィンドウによるステレオマッチング”, 電子情報通信学会論文誌 (D-II), Vol. J74-D-II, No. 6, pp. 669–677, Jun. 1991
- [2] 石井健一郎, 上田修功, 前田英作, 村瀬洋, わかりやすいパターン認識, オーム社, 東京, 1998.
- [3] 佐藤淳, コンピュータビジョン—視覚の幾何学—, コロナ社, 東京, 1999.
- [4] John Shawe-Taylor, Nello Cristianini, Kernel Methods for Pattern Analysis, Cambridge University Press, Cambridge, 2004.
- [5] 麻生英樹, 津田宏治, 村田昇, パターン認識と学習の統計学, 甘利俊一 (編) 岩波書店, 東京, 2003.