

サルの視覚認識における特徴解明に関する研究

徳田 裕季 勝間 大輔 清水 大輔 渡邊 睦

鹿児島大学 大学院 理工学研究科

〒890-0065 鹿児島市郡元 1-21-40 Tel: 099-285-8453

E-mail: {sc103038,sc097012,mutty}@ics.kagoshima-u.ac.jp

あらまし 脳における視覚認識の研究を行う上で、重要な問題の1つは、物体像の変化に対して物体を識別するメカニズムの解明である。我々は、視覚認識における特徴の解明、認識メカニズムの解明の2段階に分けることによって研究を進めている。視覚特徴は、明度情報と形状情報に分類されると考えられる。たとえば、明度ヒストグラムでの局所的な違いや、局所的な形状の大きさ、ウェーブレット変換係数に関する特徴の変化は、物体像の変化の間でのサルの認識成功率によって判明し、有意のある特徴は馴染みの薄い物体の画像によって、選定される。本論文では、サルの視覚認識に用いた有意のある特徴の選定するために、第1段階であるサルの視覚認識に関する解析結果について報告する。

キーワード 視覚認識 特徴 サル

A Study about feature elucidation in sight recognition of a monkey

Yuki TOKUDA, Daisuke KATSUMA, Daisuke SIMIZU and Mutsumi WATANABE

Graduate School of Science and Engineering, Kagoshima University

1-21-40 Korimoto Kagoshima, 890-0065, Japan

E-mail: {sc103038, sc097012, mutty}@ics.kagoshima-u.ac.jp

Abstract One of the most important problems to investigate visual recognition in a brain is the mechanism to identify an object against changes of appearance. We are attacking the problem by dividing it into two stages: that is, the visual feature searching stage and the recognition mechanism investigation stage. Visual features are considered to be classified into two classes; the brightness information class and the shape information class. Changes of several features involved into the both classes, such as, the intensity histogram, local difference of brightness, local shape size and wavelet transform coefficients, between different views, are examined with the success recognition rate of monkeys and possible features are selected by using unfamiliar object images. This paper reports the analysis of the first stage as for selecting possible features used for visual recognition of monkeys.

Keyword Visual Recognition Features Monkeys

1. はじめに

脳における視覚認識の研究に関して、重要な問題の1つは、物体像の変化に対して物体を識別するメカニズムの解明である。なかでも、物体像の変化が複雑な「観察角度の変化を伴う物体像の変化」への対処は、困難な問題であると考えられる。

本研究では、サルの視覚認識に関する知見を得るために、サルの視覚認識に反映された物体認識システムの構築を目的としており、サルの視覚認識において、用いられる画像特徴候補の解明(網膜レベル)、特徴に基づく認識メカニズムの解明(脳レベル)の2段階で研究を進めている。従来は、「脳内で物体像はまず、観察角度ごとに別々に表現され、物体が回転するときに異なる観察角度の物体像を時系列での経験によって、同一物体の異なる観察角度の物体像の脳内表現が脳内で結び付けられる」と考えられてきたが、2章で述べるサルの視覚認識実験において、「回転像を時系列での経験がない場合でも、それぞれの観察角度を見慣れるだけで、角度によらずに物体を認識する能力が成立する」という結果が判明した。

2. サルの視覚認識に関する代表的な研究

サルの視覚認識に関する代表的な研究として、サルを対象とした観察角度に依存しない物体認識の能力を測定する手法[1]が、挙げられる。

対象物体は、剛体である、かつ、馴染みの薄い3次元物体とする。よく似ている物体が4個、ある回転軸を元に30度ずつ回転している4×4の物体像をモデルとし、これを刺激セットとする。それぞれの物体像を学習することで、物体認識が可能であることが判明した。

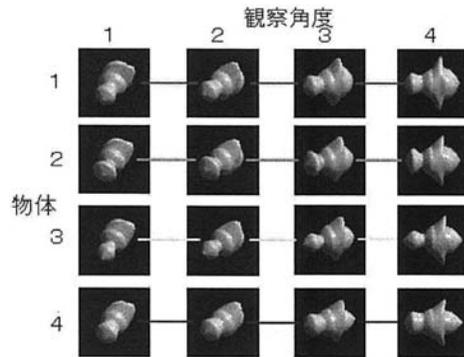


図 1. 刺激セット

サルの視覚認識実験では、2頭の訓練されたサル(ニホンザル)を見た場合の行動課題遂行の成績について実験を行っている。

サルはコンピュータのモニタに映った刺激セットの物体像を見て「何であるか」を判断し、体の前に置かれた押しレバースイッチを操作して応答する。刺激像を掲示し応答するという一連の課題遂行の最小単位(試行)の中では、刺激として物体像を1ないし5個経時的に提示する。各刺激像は1秒に1個、それぞれ0.5秒間に提示した。

サルがレバースイッチを押してモニタの中心に目を向けると刺激像の提示が始まる。この実験で主に用いたテスト課題(図1)では、最初の物体(物体1)の異なる投影(観察角度)像を1~4個提示し、次に異なる物体(物体2)の像を見せる。その際、物体2の像が提示された1秒以内にサルがレバースイッチを離したら正解とし、報酬としてジュースを少量与えた。物体1の提示回数はランダムとしたので、サルは2回目の物体像以降は像を見るたびに前と同じ物体像であるか、それとも新しい物体像であるかを判断しなければならない。このテストを1日あたり400回程度行った。

このテストの目的は、前の刺激から今の刺激へ変化する中で、あるひとつの物体が回転した像であるか、または、物体も変わり、かつ、回転している物体像であるかを判断させることによって、

観察角度に依存しない物体認識の能力を測定することである。

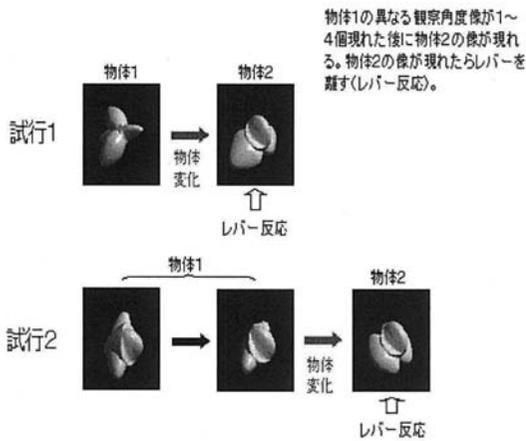


図2 観察角度をまたいだ物体認識を調べるテスト課題

新しい物体像を見慣れるときには、脳の中でその物体像の効率的な表現が出来上がると、考えられる。「新しい物体像の表現は、観察角度の変化によって変化しにくい図形特徴を使って構成される」と、仮定すると、視覚認識実験の結果を説明できるが、現在、どのような特徴を用いて物体認識しているか不明である。そこで、我々は第1段階として、サルの視覚認識における特徴を解明するために、サルの物体認識に用いられている特徴候補として、主に、明度情報、形状情報による特徴抽出して、サルの認識正解率と特徴量の変化間の相関解析を行った。

3. 解析手法の流れ

解析手法の流れについて以下に示す。

まず、刺激セットから画像をランダムで2枚選出し、画像特徴を検出する。次に、画像特徴を定量的に評価するためにパラメータ化を行う。その後、サルの視覚認識実験における生データと比較を行う。

画像の比較については、2つの物体が同じ物体

である場合と、異なる物体である場合に分けて比較を行う。観察角度が同じであると仮定すると、同じ物体である場合は、物体の差(違い)がないので、認識率が上がるほど物体の特徴量の差が小さくなると考えられる。また、異なる物体の場合は、物体の差(違い)があるので、認識率が上がるほど物体の特徴量の差が大きくなると、考えられる。

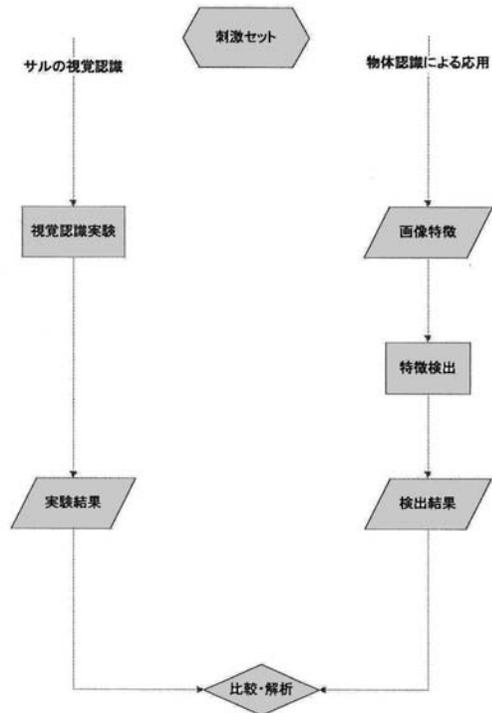


図3. 解析手法の流れ

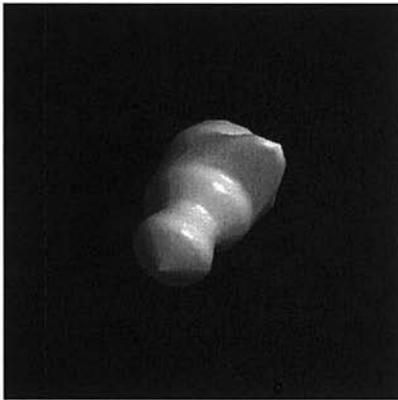
4. 画像特徴候補の抽出[2]

剛体かつ、普段見慣れない3次元物体が持つ情報は、明度情報と形状情報に大きく分類される。これらを大局的、および局所的に区別すると、明度情報に基づいた画像特徴として、大局的手法は、明度ヒストグラムによる差分、2次元離散ウェーブレット、局所的手法は、局所的な画像差分に分類する。また、形状情報に基づいた画像特徴として、大局的手法は、Canny Edge Detectorによる差分、局所的手法は、形状特徴を用いた線分検出に分類する。

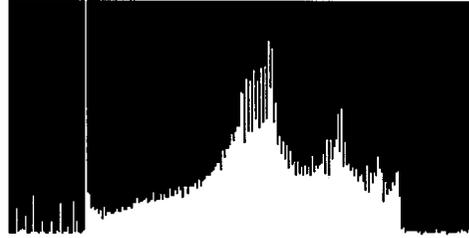
4.1 明度ヒストグラムによる解析

図4の(a)は刺激セットからの原画像、(b)はそのヒストグラムである。

ここで、用いているヒストグラムでは、縦軸を画素数、横軸を明度(輝度値)としてヒストグラムを取得している。原画像の背景の黒い部分、つまり、明度の低い部分は、ヒストグラムでは左側の表され、また、原画像の鏡面反射部分、つまり、明度の高い部分はヒストグラムでは右側に表される。



(a)



(b)

図4.(a)原画像 (b)ヒストグラム

4.2 2次元離散ウェーブレットによる解析[3]

信号を時間と周波数の両面から同時に捉える時間周波数解析において、ウェーブレットは信号の部分を切り出す単位となるものである。ウェーブレット変換は、この単位で切り出した信号各部の大きさを表す。信号は最小の単位があるため、ウェーブレット変換を離散化して効率のよい時間周波数解析が得られる。

4.2.1 分解と再構成

ほとんどの信号(イメージ)は、低周波成分がより重要で、これによりほとんどの信号が認識できると考えられる。それに対して、高周波成分は微小な部分を伝えるものと考えられる。

このため、ウェーブレットによる解析では、低周波成分を Approximation 成分、高周波成分を Detail 成分と分類され、信号をこれら2つに分解する。

2次元のウェーブレット変換の場合、行方向と列方向にイメージを分解するため、1回の分解で、Approximation 成分と3方向(水平、垂直、対角)の Detail 成分からイメージが再構成される。

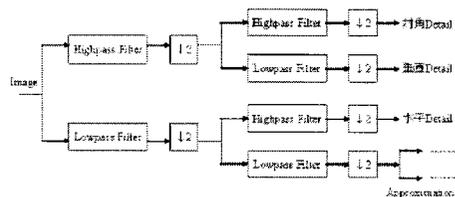


図5.2次元離散ウェーブレットにおける分解

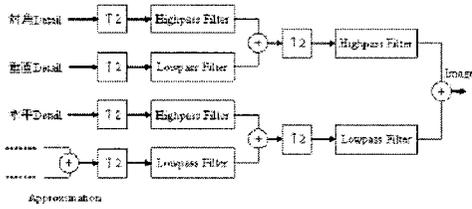


図 6. 2次元離散ウェーブレットにおける再構成

4.2.2 イメージの圧縮

2次元の信号は2変数関数 $f(x, y)$ で表される。2変数関数 $f(x, y)$ は次のようにスケーリング関数の線型結合で表される。

$$f_j(x, y) = \sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} c_{LL,k,l}^{(j)} \phi(2^j x - k) \phi(2^j y - l)$$

この関数の展開関数 $\{c_{LL,k,l}^{(j)}\}$ は次のアルゴリズムによって分解される。

$$c_{LL,k,l}^{(j-1)} = \sum_n \sum_m \frac{1}{2} g_{2k-n} \frac{1}{2} g_{2l-m} c_{LL,m,n}^{(j)}$$

$$c_{HH,k,l}^{(j-1)} = \sum_n \sum_m \frac{1}{2} h_{2k-n} \frac{1}{2} h_{2l-m} c_{LL,m,n}^{(j)}$$

$$c_{LH,k,l}^{(j-1)} = \sum_n \sum_m \frac{1}{2} g_{2k-n} \frac{1}{2} h_{2l-m} c_{LL,m,n}^{(j)}$$

$$c_{HL,k,l}^{(j-1)} = \sum_n \sum_m \frac{1}{2} h_{2k-n} \frac{1}{2} g_{2l-m} c_{LL,m,n}^{(j)}$$

ここで、添字 LL などの L は低周波成分、 H は高周波成分を表す。ウェーブレットは、Haar ウェーブレットを適用し、分解レベルは5とし、画像圧縮のスレッシユホールド値は、Birge-Massart 手法による2次元ウェーブレット用閾値である準中央値(Scarce medium)を適用した。

4.3 局所領域による画像差分による解析

あらかじめ、刺激セットの中の全ての画像に存在する物体の局所領域のうち、領域上部を領域A、領域下部を領域Bと定めた。刺激セットからの原画像からの領域指定例を図7に示す。

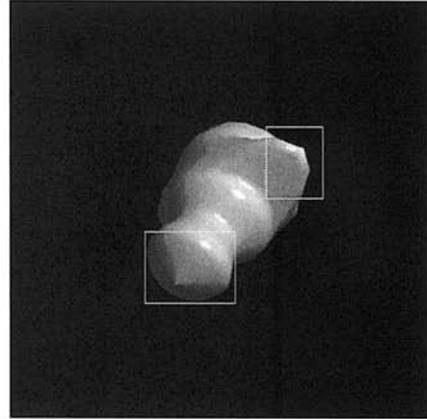


図 7. 領域指定

4.4 形状特徴に基づいた解析

まず始めに、Canny Edge Detector によってエッジ処理を行い、画像の輪郭線を抽出した。抽出された画像を基に刺激セットにおいて、全ての画像の中で存在するという条件の下で、特徴点を定め、特徴点で線を結び、それを特徴線分とした。検出結果例を図8に示す。

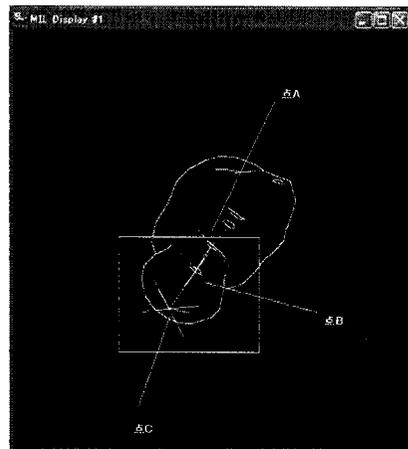


図 8. 検出結果

5. 実験

5.1 実験条件

サル の 視 覚 認 識 実 験 と 同 じ 条 件 で 行 う た め に、刺 激 セ ッ ト か ら ラ ン ダ ム で 選 出 さ れ た 2 枚 の 画 像 か ら、同 一 物 体 で あ る か ど う か を 判 断 す る と い う こ と を 実 験 条 件 と し た。

実 験 に 使 用 し た P C の C P U は、Pentium IV 3.2GHz、OS は、Windows XP Professional、仕 様 言 語 は、Microsoft 社 製 Visual Studio 2005 C++ で あ る。2 次 元 離 散 ウ ェ ヴ レ ッ ト に よ る 解 析 に 関 し て は、CYBERNET 社 製 MATLAB7.1 を 使 用 し た。

5.2 実験結果

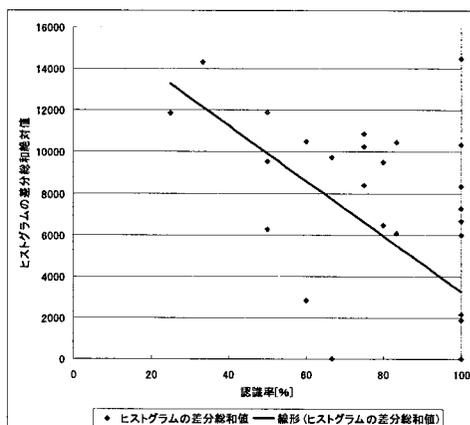
前 章 で 述 べ た 各 手 法 に よ っ て、特 徴 を 抽 出 す る 実 験 を 行 っ た。各 手 法 は、明 度 情 報 を 伴 う 手 法、形 状 情 報 を 伴 う 手 法 に 大 き く 分 類 さ れ、さ ら に、大 局 的、お よ び 局 所 的 手 法 に 分 類 さ れ る。

明 度 情 報 を 伴 う 大 局 的 手 法 と し て、明 度 ヒ ス ト グ ラ ム を 用 い た 手 法 で は、刺 激 セ ッ ト か ら ラ ン ダ ム で 選 出 さ れ た 2 枚 の 画 像 か ら ヒ ス ト グ ラ ム を そ れ ぞ れ 取 得 し て、そ れ ぞ れ の ヒ ス ト グ ラ ム の 差 分 を 行 う こ と に よ り、差 分 ヒ ス ト グ ラ ム を 取 得 し て、そ の 差 分 ヒ ス ト グ ラ ム の 総 和 を 求 め た。ま た、2 次 元 離 散 ウ ェ ヴ レ ッ ト に よ る 手 法 で は、画 像 圧 縮 を 行 い、画 像 そ れ ぞ れ の 圧 縮 率 で あ る 圧 縮 後 の 残 さ れ た イ メ ー ジ エ ネ ル ギ ー (Compressed Image Energy) を 取 得 し た。ま た、局 所 的 手 法 と し て、画 像 差 分 を 用 い た 手 法 で は、予 め 領 域 指 定 し た 領 域 内 で 差 分 さ れ た 画 像 数 を 取 得 し た。

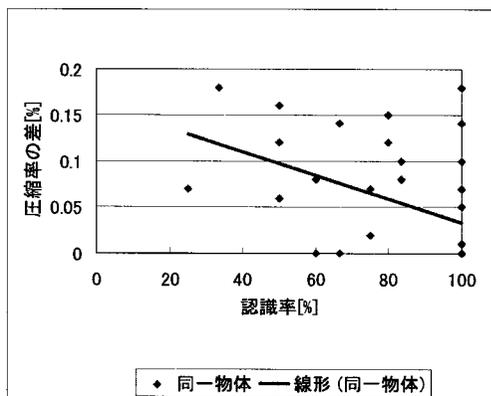
形 状 情 報 を 伴 う 大 局 的 手 法 と し て、エ ッ ジ 差 分 を 行 い、形 状 を 形 成 す る 輪 郭 部 分、陰 影 部 分、鏡 面 反 射 部 分 の 画 素 数 を 取 得 し た。ま た、局 所 的 手 法 と し て、エ ッ ジ 処 理 後 の 画 像 を 基 に、図 7 の 青 領 域 内 で、特 徴 線 分 を 取 得 し た。

取 得 し た 特 徴 量 に 基 づ き、サ ル が 物 体 を 認 識 し た 正 解 率 を 2 枚 の 画 像 の 物 体 が 同 一 物 体 で あ る 場

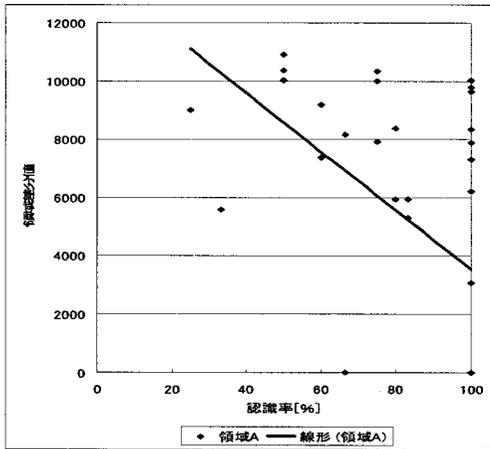
合 と、異 なる 物 体 で あ る 場 合 と で 分 類 し て 比 較 を 行 っ た。比 較 を 行 っ た 実 験 結 果 例 を 図 9、図 10 に 示 す。



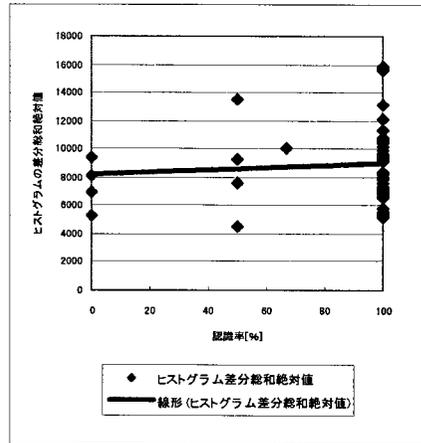
明度ヒストグラム



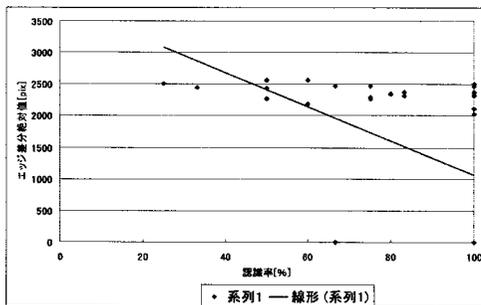
ウェーブレット変換による圧縮



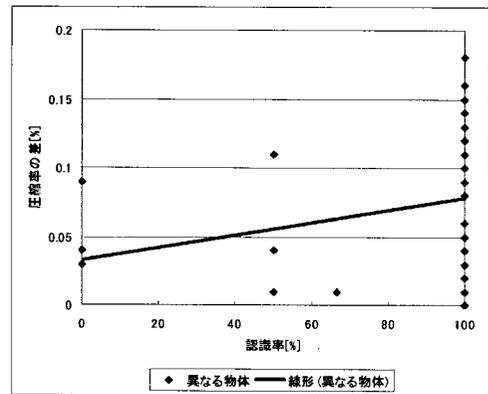
局所領域差分



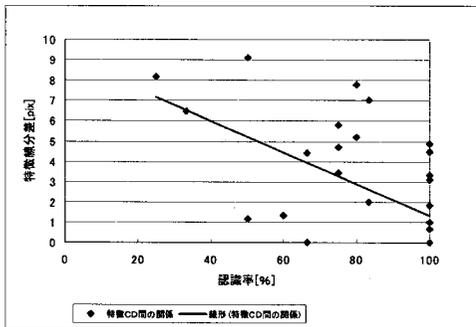
明度ヒストグラム



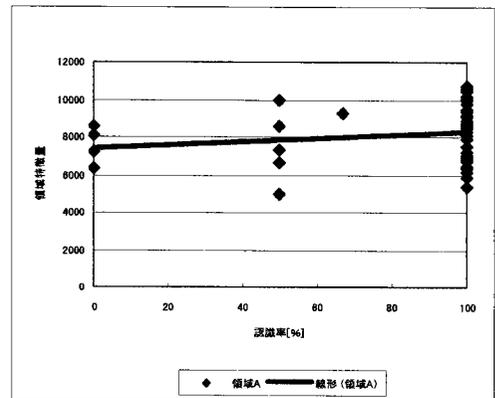
エッジ差分



ウェーブレット変換による圧縮

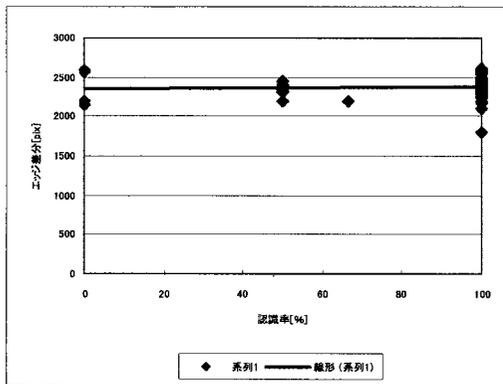


特徴線分

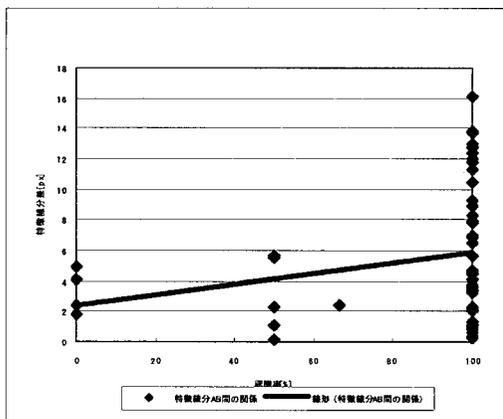


局所領域差分

図 9. 同一物体での特徴量とサル正解率との比較



エッジ差分



特徴線分

図 10. 同一物体での特徴量とサルの見解率との比較

6. 考察

画像特徴に基づく手法について、考察を行う。

各特徴量とサルが物体を認識した正解率の関係を比較・検証を行った。同一物体に関して、サルが物体を認識した正解率が上がるにつれて、特徴量が下がるという負の相関関係が得られたことを確認した。また、異なる物体に関して、サルが物体を認識した正解率が上がるにつれて、特徴量も上がるという正の相関関係が得られたことを確認

した。しかし、正の相関関係が得られない画像特徴も確認した。それらの画像特徴は、有意のある画像特徴として、有効性は見られないと考えられる。また、本研究で用いたサルの生データ量が少ないこともあったため、データ量を増加、それに伴う画像特徴候補の検討など考えていく必要があると考えられる。

7. まとめ

本研究では、サルの視覚認識に用いた有意のある特徴を選定するために、第1段階であるサルの視覚認識に関する研究を行った。

各解析手法において、特に明度情報を伴う解析手法においては、大局的手法であるウェーブレット変換による手法、形状情報を伴う局所的な手法である形状特徴を用いた特徴線分による解析手法に対して、同一物体と異なる物体、双方の場合において、より相関関係を示した。従って、相関傾向から、有効性が見られ、サルが用いる画像特徴として、特にウェーブレット変換による手法、形状特徴を用いた特徴線分による手法について候補となることが判明した。

参考文献

- [1] 「Prior experience of rotation is not required for recognizing objects seen from different angles」
鹿兒島大学 大学院 理工学研究科
Gang Wang, Shinji Obama,
Wakayo Yamashita, Tadashi Sugihara, Keiji Tanaka
ネイチャー・ニューロサイエンス VOL8,
No12,2005
- [2] 榊原 進：“ウェーブレット ビギナーズガイド”，
電機大出版局，1995
- [3] Dana H. Ballard, Christopher M. Brown
“コンピュータ・ビジョン”，1987