研究論文

把持運動において対象物の画像情報の変更が 手形状計算へ与える影響^{*1}

福田 浩 士, 福 村 直 博, 宇 野 洋 二

豊橋技術科学大学情報工学系システム応用研究室^{*2}

Influence of Processed Visual Information of Objects on Computation of Hand Shapes in Human Grasping Movements

Hiroshi Fukuda, Naohiro Fukumura and Yoji Uno

Department of Information and Computer Sciences, Toyohashi University of Technology

From the viewpoint of computational theory, it has been postulated herein that certain properties of objects are extracted from their visual images for visuo-motor transformation in human grasping movements. To examine the computational feasibility of this hypothesis, a five-layer neural network model has been constructed for transforming the visual images of objects into the prehensile hand shape. After sufficient learning, the property information corresponding to the size and shape of objects was extracted in the third layer. The values of the third-layer and output-layer units changed widely for shadingless images and only slightly for blurred images, in comparison with the original images. Essentially the same results were obtained from a psychophysical experiment conducted on human grasping movements with the same images (original, shadingless and blurred) as those used in the neural network study, thus supporting the plausibility of the present computational hypothesis.

1. は じ め に

本研究は、ヒトの把持運動における視覚-運動変換と対象物認知の関係について、計算論的に解明することを目的としている。

ヒトが熟練度の高い把持を正確かつ滑らかに行うことができるのは、把持対象となる物体の視覚情報に基づいて手形状や把持力などの把持運動情報を事前に計算しているからであると考えられる。我々は、そのようなフィードフォワードの視覚-運動変換は、脳内に獲得されている把持運動に関する対象物の内部表現に基づいているという仮説を立てた。それでは、その把持運動に関する対象物の内部表現とはどういうものであろうか。Marr²⁾は視覚の目的は網膜における外界の2次元的な画像からその3次元構造を推測することであると述べている。ある物体の2次元的な画像からその3次元形状を推定するためには、それが不良

設定問題であるために、脳内に物体の3次元構造の内部表現が必要となる。しかしながら、このMarrの見解は視覚のみの計算で閉じてしまっている。把持運動における視覚-運動変換の目的は、網膜における対象物の2次元的な画像からそれを把持するための把持点、手形状などの把持運動情報を計算することである。我々は把持運動の計算理論を展開し、把持運動を実行するためには少なくとも合力決定、力分配、把持点決定、手形状形成、制御の5つの計算問題が解かれなければならないと主張しているが、それらの計算問題は不良設定問題であるので、脳内にそれらに関する内部表現が必要となる。その内部表現は、過去の把持の経験によって視覚情報や体性感覚情報などが統合されることによって脳内に獲得されると我々は考えている。従って、本論文で言うところの把持運動の内部表現とは、対象物の視覚的特徴とそれを把持するための把持点（対象物上の把持点と、それと接触する手の部位）や手形状の対応関係をコーディングしているものである。

内部表現に基づいた視覚-運動変換時に、視覚情報

^{*1} 1999年10月29日受付, 2000年6月20日受理

^{*2} 〒441-8580 豊橋市天伯町雲雀ヶ丘 1-1

を細部まで使うとなると膨大な情報処理量とメモリ量が必要となる。そこで、我々は、「対象物の視覚情報から把持運動にとって必要不可欠な特徴情報が抽出され、その情報から内部表現に基づいて把持運動情報が計算される」という仮説を立てた。その仮説に基づいて、我々は、心理物理実験と5層の神経回路モデルを用いた計算機シミュレーションを通して、ヒトが対象物の視覚情報のみから事前にある程度把持運動情報を計算していること、その際に対象物の特徴情報が抽出され得ることを示した。

本研究では、視覚-運動変換において、入力される対象物の視覚情報、処理過程で抽出される対象物の特徴情報、出力される手形状の関係を調べた。まず、神経回路モデルを用いて、視覚-運動変換の際に対象物の特徴が抽出され得ること、そして、抽出される特徴情報はその対象物を把持するための手形状情報とも密接に関係していることを示す(2章前半)。次に、同じ神経回路モデルを用いて、対象物の視覚情報が全体にわたって乱されたときの、対象物の特徴抽出と手形状の計算への影響を調べる(2章後半)。次に、心理物理実験において、神経回路モデルの入力と同様の乱された視覚情報を被験者に示し、それを把持するための手形状をつくらせ、その結果から視覚情報と手形状の関係を考察する(3章)。最後に、シミュレーション結果と心理物理実験結果を比較検討し、ヒトの把持運動における視覚-運動変換について考察する(4章)。

2. 神経回路モデルによる検討

2.1 特徴抽出を行う神経回路モデル

把持運動の視覚-運動変換の処理過程において抽出される対象物の特徴は、ヒトが対象物を識別する(対象物を視覚的に同定する)ために用いるような対象物固有の特徴ではなく、手形状などの把持運動情報を計算するのに適した表現における特徴であると考えられる。そこで、本節では、5層の神経回路モデルを構築し、対象物の2次元画像を入力し、対応する手形状を出力するように学習を行った。そのとき、第3層に抽出される特徴情報と手形状との関係を解析した。

2.1.1 神経回路モデルのアーキテクチャ

5層の神経回路モデルの入力に対象物の2次元画像、出力を手形状データとしたとき、第3層に対象物の特徴が情報圧縮されて抽出されることを期待し、神経回路モデルのユニット数を第1層から3600, 20, 2, 27, 18とした(Fig. 1)。第1, 3層, 第5層のユニットは線形ユニットで、第2, 第4層はシグモイド関数を用いた非線形ユニットである。対象物として、

一辺がそれぞれ20, 30, 40, 50, 60 [mm]の正方形を上底面に持ち、高さが均一(120 [mm])の四角柱5種類、直径がそれぞれ20, 30, 40, 50, 60 [mm]で高さが均一(120 [mm])の円柱5種類、直径が20, 30, 40, 50, 60 [mm]の球5種類を用意し、それらの60画素×60画素の2次元画像を学習の入力データとして用いた。入力データは各画素の輝度情報[0, 255]を[-1, 1]の範囲に正規化している。そして、各対象物の画像に対してある被験者がつくった手形状をCyber Glove (Virtual Technologies社)で計測し、教師データとして用いた。そのデータはCyber Gloveの出力値を[0, 2]の範囲に正規化している。

Fig. 2に示すCyber Gloveの22センサのうち、手首のpitch, yaw, 手の平の湾曲のためのセンサと故障のため信頼性のない小指のDIP関節のためのセンサ(それぞれセンサ番号19, 20, 21, 17)を除いた18センサを使用した。その18センサは、拇指のCM関節のrotationとabduction, MP関節とIP関節のflexionに対する4センサ、他の4本の指のMP関節, PIP関節, DIP関節(小指を除く)のflexionおよびそれぞれの間のabductionに対する14センサである。

2.1.2 学習結果

学習後の神経回路モデルに、学習に用いた15種類

Fig. 1 A five-layered neural network model which computes a prehensile hand shape for the image of an object.

Fig. 2 Sensor numbers for finger joints.

の対象物の画像を入力したときの第3層の2つのユニットの出力値を Fig. 3 に示す。横軸は第3層のユニット1の出力値、縦軸は第3層のユニット2の出力値を示す。また、●印は角柱、▲印は円柱、■印は球を表し、印の側の数値は対象物のサイズを表している。この図より、ユニット1の出力値で{角柱, 円柱}, {球}という分類ができ、ユニット2の出力値は対象物のサイズが大きくなるにつれて大きくなっていることがわかる。つまり第3層のユニット1が対象物の形に関する情報を、ユニット2が対象物のサイズに関する情報を表していることになる。

Fig. 4 は、学習後の神経回路モデルの第1層と第3層のユニットの関係をイメージ化したものである。そのイメージの各画素値は、その画素に対応する神経回

路モデルの第1層のユニットの値を最大にし、それ以外のユニットの値を最小にしたときの第3層の各ユニット値を $[0, 255]$ で量子化した値である。イメージの白い部分は第3層のユニット値の高い値、黒い部分は低い値を示している。Fig. 4 のユニット1では、角柱と円柱のような長い対象物と球とで明暗が分かれている。これは形に関する情報を表していると考えることができる。また、ユニット2では、各対象物の共通部分に重み付けがされ、グラデーションのようにになっている。今回の神経回路モデルでは、対象物の位置が固定であるので、この重み付けはサイズを表していると考えることができる。

Fig. 5 は、第3層と第5層の関係を示している。第4層はシグモイド型の非線形ユニットであるが、15種類全ての対象物において第4層が取り得る範囲ではほぼ線形であると考えてよい構造となっていた。そこで、第3層の各ユニットの値と第5層の各ユニットの値の線形回帰を調べることで、第3層と第5層の関係を見ることにした。Fig. 5(a) は、学習後の神経回路モデルの第3層のユニット1 (形に関するユニット) の値を -2 から 2 まで 0.2 ステップで変化させたときの第5層の各ユニットの値との回帰直線の傾きを示している。同図(b) は第3層のユニット2 (サイズに関するユニット) の値を 5 から 7 まで 0.2 ステップで変化させたときのものである。第3層のユニット値を変化させる範囲は、ユニット1では Fig. 3 において角柱・円柱と球の境界部分、ユニット2では 40 [mm] の対象物の値の近辺をとっている。各図の横軸は第5層のユニット番号で、これは Fig. 2 に示してある Cyber Glove のセンサ番号に相当する。第3層のユニット1の値と第5層の各ユニットの値の回帰における決定係数はユニット4を除いて 0.9 以上であった。そのときの回帰直線の傾きは、Fig. 5(a) より、第5層の abduction に関するユニット (No. 3, 10, 14, 18) で正の値となっている。これは、第3層のユニット1の値が球に近づく、abduction が広がり、円柱や角柱に近づく、abduction が閉じるような手形状がつくられることを示している。また、中指と薬指の MP 関節に関するユニット (No. 7, 11) で正の値となり、ほとんど全ての指の IP 関節に関するユニット (No. 2, 5, 8, 9, 12, 13)、拇指の rotation 関節と MP 関節、小指の MP 関節に関するユニット (No. 0, 1, 15) で負の値となっている。これは、球だけに注目するとユニット2だけでなくユニット1の値とサイズの間にも高い相関があり、球のサイズがこのユニットの振舞いによっても影響されることを示している。また

Fig. 3 Values of the 3rd layer units in the neural network.

Fig. 4 Relation between the 1st layer and the 3rd layer in the neural network.

Fig. 5 Relation between the 3rd layer and the 5th layer.

同図(b)より, 第3層のユニット2の値と第5層の各ユニットの値の回帰直線の傾きは, 拇指のrotationと他の指のMP関節で負の方向に大きくなっており, 各指のIP関節で比較的小さな正の値となっている。また, 各ユニットに対する回帰における決定係数は全て0.9以上であった。このことは, 第3層のユニット2における表現がサイズの大きい方向にシフトすると各指のMP関節が伸び, PIP, DIP関節がそれを補正するように若干曲がるような手形状をつくるようになることを示している。

2.1.3 考 察

対象物の2次元画像情報を入力とし, その対象物を把持するための手形状を出力とし, 5層の神経回路モデルを学習させた結果, 第3層の各ユニットに対象物の形とサイズに関する特徴情報が抽出された。以前の研究⁹⁾では, 第3層の各ユニットに特定の特徴情報が現れるような学習結果にはならず, 学習後に15種類の対象物に対する第3層の値に主成分分析したところ, 対象物のサイズに関する主成分と対象物の形に関する主成分が抽出された。それに対して, 今回のシミュレーションでは, 第3層の各ユニットがそれぞれ対象物のサイズと形に関する特徴情報を表すような学習結果となった。このように, 神経回路モデルの結合荷重の総数72,580に対してデータ数が15と少ないために, 初期状態が異なると学習後の結合荷重の値は変化してしまう。しかしながら, 結合荷重の値が変化して

も, 学習後の第3層の値からは常に対象物サイズと形に関する主成分が抽出された。

第1層と第3層の関係を調べた結果, 第1層の情報から第3層のユニット1の結合で主に対象物の形に関する特徴情報を, 第1層の情報から第3層のユニット2への結合で主に対象物のサイズに関する特徴情報を計算するような神経回路モデルが構築されていたことがわかった。この結果は, 視覚-運動変換の際に, 対象物の視覚情報から対象物の特徴情報が抽出され得ることを示している。

第3層と第5層の関係を調べたところ, 第3層の特徴情報に基づいて, 入力された対象物に適した手形状を計算する処理を行っていることが示された。

抽出された対象物の特徴は形とサイズに関する情報であったが, これは把持運動情報を計算するのに適した表現の特徴情報であると考えられる。なぜなら, Fig. 3の特徴情報の分布が角柱と円柱とで極めて近いからである。対象物を識別する(対象物を視覚的に同定する)ための特徴がそのまま抽出されるのであれば, 角柱, 円柱, 球に対して特徴情報の分布は明確に分かれるべきである。このことは, 今回のように, 把持運動情報の中で手形状のみを計算するタスクであれば, 角柱と円柱がはっきりと分かれて抽出される必要はないことを意味している。従って, 我々は, 把持運動における対象物の特徴抽出では, 視覚情報から対象物の固有の情報が特徴としてそのまま抽出されるので

はなく、把持運動情報を計算するのに適した表現の特徴情報が抽出されることを主張する。

しかしながら、手形状の他に把持点や把持力などの把持運動情報を考慮してモデル化すれば、角柱と円柱の特徴情報の分布はおそらく分かれるであろう。その場合でも対象物の固有の情報に基づいて分かれるのではなく、その対象物に対する把持運動情報に基づいて分かれるはずである。そのため、対象物の視覚的同定の結果が異なっても、把持運動のための特徴情報は同じになる、またはその逆となる対象物が存在する可能性も否定はできない（5章でより詳細な考察を行う）。

2.2 視覚情報の変更，対象物の特徴抽出，手形状計算の関係

本節では、前節で構築した神経回路モデルを用いて、対象物の視覚情報が全体にわたって乱されたとき

に、対象物の特徴情報の抽出と手形状の計算がどのような影響を受けるのかを調べた。

2.2.1 シミュレーション方法

15種類全ての対象物に対して Fig. 6 のような処理を施した画像を用意する。Fig. 6(a)は無処理（unprocessed）画像，(b)は陰影除去（shadingless）画像，(c)はぼかし（blurred）画像である。陰影除去画像は、無処理画像の対象物部分の画素値を最大値に、背景部分を最小値とした2値化画像である。ぼかし画像は、無処理画像に 3×3 の平滑化フィルタをかけたものである。陰影除去画像は実世界において対象物に強い光が照射された状況を、ぼかし画像は観測者の目の疲労などから対象物がぼやけて見える状況を想定している。前節で学習した神経回路モデルに、陰影除去画像とぼかし画像を入力したときの第3層と第5

Fig. 6 Examples of processed images of objects.

層の出力値を無処理画像を入力したときの出力値と比較する。

2.2.2 シミュレーション結果

Fig. 7は神経回路モデルに無処理画像を入力したときと処理された画像を入力したときの第3層の出力値を示している。Fig. 7(a)は無処理画像と陰影除去画像との比較、同図(b)は無処理画像とぼかし画像との比較である。図の横軸は第3層のユニット1の出力値、縦軸はユニット2の出力値であり、●印、▲印、■印はそれぞれ角柱、円柱、球の無処理画像に対する出力値を、○印、△印、□印はそれぞれ角柱、円柱、球の処理された画像に対する出力値を表している。また印の側の数値は対象物のサイズを表している。無処理画像に対する出力値は前節 Fig. 3と同じである。Fig. 7(a)より、陰影除去画像では無処理画像に比べて全体的にユニット2の値が大きくなる方向にシフトしており、対象物によってはユニット1の値のシフトも見られる。Fig. 7(b)より、ぼかし画像では60 [mm]の角柱と全ての円柱でユニット2の値が小さくなる方向にシフトしている。60 [mm]を除いた角柱でユニット2の値が大きくなる方向にシフトしている。球では、30 [mm]でユニット1と2の両方向への若干のシフトが見られるもののほとんど差はない。

Fig. 8(a)に角柱30 [mm]、(b)に円柱30 [mm]、(c)に球30 [mm]の処理画像に対して

ネットワークが出力した手形状を示す。これらの図では、無処理画像と比較しやすいように、縦軸に無処理画像における出力値との差分を示している。flexionに関するセンサはその関節が曲がるにつれて出力値が大きくなり、abductionに関するセンサは対応する指の間が閉じているほど大きくなる。拇指のrotationに関するセンサは拇指が手の平の方に入り込むほど、そして拇指と示指の間が閉じているほど大きくなる。今回のシミュレーションでは、処理画像における手形状が無処理画像の手形状に対してどのように変化するかを調べることができればよいので、比較には神経回路の出力値をそのまま用いた。この神経回路モデルの出力値はCyber Gloveの生の出力値を正規化したものである。縦軸に単位はないが、実際の関節角度とほぼ線形の関係にある。横軸は第5層のユニット番号で、これは手形状計測装置Cyber Gloveのセンサ番号に相当する。図の◇印は基準となる無処理画像に対する値(差分であるので全て零)、△印は陰影除去画像に対する値、□印はぼかし画像に対する値を表している。

角柱30 [mm]では、陰影除去画像の手形状は、無処理画像に比べて拇指のrotationと他の指のMPが伸び、PIP、DIPがその開きを補うように若干曲がっている。また、各指の間のabductionが狭くなっている。ぼかし画像の手形状は、拇指のrotationと他

Fig. 7 Values of the two units in the 3rd layer: (a) when shadingless images and unprocessed images of objects are fed to the neural network, and (b) when blurred images and unprocessed images of objects are fed to the neural network.

Fig. 8 Output values of the neural network when processed images of the objects are fed to the neural network. Figures (a), (b) and (c) show the results for 30 [mm] quadrangular prism, 30 [mm] cylinder and (c) 30 [mm] sphere, respectively.

の指のMPが若干伸びているが、無処理画像とほぼ同じ出力になっている。また、他のサイズの角柱でも同じような出力の傾向が見られた。円柱30 [mm]では、陰影除去画像の手形状は、無処理画像に比べて拇指のrotationと他の指のMPが伸びている。ぼかし画像の手形状は、拇指のrotationと他の指のMPが若干曲がっているが、無処理画像とほぼ同じ出力になっている。また、他のサイズの円柱でも同じような出力の傾向が見られた。球30 [mm]では、陰影除去画像の手形状は、無処理画像に比べて拇指のrotationと他の指のMP、IPが伸び、各指の間のabductionが広がっている。ぼかし画像では、無処理画像とほぼ同じ出力になっている。また、他のサイズの球でも同様の出力の傾向が見られた。

2.2.3 考 察

陰影除去画像では、無処理画像とは異なる特徴が第3層に抽出された。これは、画素情報の変更が特徴抽出に大きく影響したことを示している。従って、陰影除去画像では無処理画像と同じ特徴を抽出するために必要な視覚情報が欠けていたと考えられる。それに対して、ぼかし画像において第3層に抽出された特徴は、無処理画像のものとほぼ同じであった。従って、ぼかし画像では無処理画像と同じ特徴を抽出するために必要な視覚情報がほとんど欠けていないと考えられる。

しかしながら、これらの結果から“把持運動のための特徴抽出には対象物の陰影情報が重要である”と結論づけるのは難しい。神経回路モデルでは、第3層のユニット2の出力が画像の輝度に場所ごとの重み付け和をとっているために、輝度が一様に高い入力画像ではユニット2の出力が大きくなってしまふ。このことが、陰影除去画像で第3層の値が大きくなり、輝度情報が比較的保たれているぼかし画像で第3層の値がそれ

ほど変化しなかった理由として考えられるからである。

角柱30 [mm]に対する手形状で、拇指のrotationと他の指のMPが伸びているのは、第3層のユニット2の値が大きくなる方向にシフトした結果によるものである。また、各指の間のabductionが狭くなっているのは、第3層のユニット1の値が大きくなる方向にシフトした結果によるものである。このことは、第3層に抽出された特徴情報に応じた手形状が計算されていることを示している。また、円柱30 [mm]と球30 [mm]においても、第3層に抽出された特徴情報に応じた手形状が出力されていることがわかる。これらの結果は、処理された画像でも、無処理画像と同じような特徴を抽出できるだけの情報が含まれていれば、無処理画像と同じような手形状を計算することができることを示している。従って、我々は、手形状が適切に計算されるためには、適切な特徴情報が視覚情報から抽出されなければならないことを主張する。

3. 心理物理実験による検討

本章では、心理物理実験により、前章の神経回路モデルのシミュレーションと同様の処理画像をヒトに与えたときの手形状の振舞いを調べていく。

3.1 実験装置と実験方法

被験者 3人の成人男子(22~25歳)が被験者として実験に参加した。被験者A、Bは実験に関して完全にナイーブであるが、被験者Cは著者の1人である。また、その3人の被験者はシミュレーションで用いたデータの被験者とは異なる。

実験環境 被験者は、Fig. 9のようにIndigo 2 (Silicon Graphics 社)のディスプレイが置かれているテーブルの前に座り、顎台で頭を固定される。顎台とディスプレイの間には覆いが設けてあり、自分の手に関する視覚的なフィードバックを被験者に与えない

ようにしてある。手形状は被験者の右手に装着されている Cyber Glove で計測する。今回の実験では Cyber Glove の 22 センサのうち、手首の pitch, yaw と手の平の湾曲に関するセンサを除いた、19 センサを使用した。その 19 センサは 2.1.1 節で説明してあるものと同じである（小指の DIP に関するセンサは修復済）。対象物の 2 次元画像はシミュレーションと同様に、5 種類のサイズの角柱、円柱、球、計 15 種類の対象物の陰影除去画像、ぼかし画像、無処理画像の計 45 種類を用いた。全ての画像のサイズは 640 画素×480 画素である。この画像を高解像度のディスプレイモニタに表示することによって、無処理画像に対して被験者が実物の対象物とほぼ同じように認識できるようになっている。陰影除去画像は、無処理画像の対象物部分の画素値を最大値に、背景部分を最小値とした 2 値化画像である。ぼかし画像は、無処理画像に 11×11 の平滑化フィルタをかけたものである。

実験手順 被験者は全ての指関節を伸ばした状態で手の平を下にしてテーブルの上に手を伏せる。これを初期姿勢とする。ディスプレイに処理された対象物の 2 次元画像が表示される。実験者の合図で、被験者は表示された対象物を 5 本の指を全て使って掴んで持ち上げようことを意識して把持形状をつくる。把持形状をつくったら被験者は実験者に「はい」という合図を返し、実験者は手形状を計測する。計測が終了したら被験者は手を初期姿勢に戻す。ここまでするを 1 トライアルとする。

まず、陰影除去画像 15 種類を 1 回ずつランダムに表示し、それぞれに対して 1 トライアルずつ計測を行う。これを 1 セットとし、15 セット行う。つまり処理画像 1 つに対して 15 点のデータを採取することになる。陰影除去画像が終了後、ぼかし画像で、そして最後に無処理画像で同様の計測を行う。最後に無処理画像を行ったのは、被験者が陰影除去画像やぼかし画像を見たときに処理前の画像を予測しないようにする

Fig. 9 Experimental setup.

ためである。

3.2 実験結果

Fig. 10 (a) は角柱 30 [mm], (b) は円柱 30 [mm], (c) は球 30 [mm] に対して被験者 A がつくった手形状である。図の△印、□印、◇印はそれぞれ陰影除去画像、ぼかし画像、無処理画像に対する手形状を示している。図の縦軸は、無処理画像と処理画像の結果の比較のために、無処理画像に対する出力値との差分を示している。flexion に関するセンサはその関節が曲がるにつれ出力値が大きくなり、abduction に関するセンサは対応する指の間が閉じているほど大きくなる。拇指の rotation に関するセンサは拇指が手の平の方に入り込むほど、また拇指と示指の間が閉じているほど大きくなる。今回の実験では、処理画像に対する手形状の変化がどのような傾向にあるかを調べることを目的としているので、比較には Cyber Glove の生の出力値をそのまま用いた。よって、縦軸に単位はないが、Cyber Glove の出力値は実際の関節角度とほぼ線形の関係にある。横軸は Cyber Glove のセンサ番号である。

角柱 30 [mm] において、陰影除去画像に対する手形状は無処理画像に比べて、拇指の rotation とその他の指の MP が伸び、それを補うように PIP, DIP が曲がっている。円柱 30 [mm] において、陰影除去画像に対する手形状は無処理画像に比べて、拇指の rotation が内側に入り、MP と IP が伸び、abduction が閉じている。示指では MP が伸び、PIP, DIP が曲がっている。中指では MP が伸び、PIP が曲がっている。薬指と小指では MP に変化はなく、PIP が曲がっている。球 30 [mm] において、陰影除去画像に対する手形状は無処理画像に比べて、拇指の MP, IP が伸び、abduction が閉じている。他の指では、MP, PIP, DIP の変化は比較的小さい。また、abduction が閉じているというのが大きな違いとなっている。このように、陰影除去画像に対する手形状は無処理画像と大きく異なっているが、ぼかし画像に対する手形状は全ての対象物で無処理画像と比較的近いものとなっている。

Fig. 11, Fig. 12 はそれぞれ被験者 B, C の結果であるが、陰影除去画像は無処理画像と大きく異なり、ぼかし画像は無処理画像に近い手形状になっている。また、他の対象物に対しても同様な結果が得られている。

3.3 考察

心理物理実験の結果は、全ての被験者で、陰影除去画像では手形状が無処理画像と大きく異なり、ぼかし画像では手形状が無処理画像と近くなっていた。

Fig. 10 Output values of the Cyber Glove sensors for hand shapes of subject A.

Fig. 11 Output values of the Cyber Glove sensors for hand shapes of subject B.

Fig. 12 Output values of the Cyber Glove sensors for hand shapes of subject C.

それでは何故、陰影除去画像では無処理画像と大きく異なったのだろうか。被験者 A は、角柱 30 [mm] の無処理画像では、前面の左の側面に拇指の把持点を、その対面に他の指の把持点を置くような手形状をつくっていた。それに対し、陰影除去画像では、左側の縦方向のエッジ部分に拇指の把持点を、右側の縦方向のエッジ部分にその他の指の把持点を置くような薄い平面的な物体に対する手形状をつくったと考えられる。各指の MP が伸び、PIP、DIP 関節が曲がっているという結果となったのは、その薄いエッジ部分を指先で把持するためである。円柱 30 [mm] でも、陰影除去画像では、おそらく薄い平面的な物体に対する手形状をつくったと考えられる。これは、この

被験者の無処理画像に対する手形状は拇指の rotation が内側に入り込み、MP と IP が曲がる傾向にあったのだが、陰影除去画像では、rotation が比較的外側となり、abduction が閉じ、MP と IP が伸びていることから推測している。球の 30 [mm] でも、陰影除去画像では薄い平面的な物体に対する手形状をつくったために大きく異なったと考えられる。これは、無処理画像は球を指先で掴むために、拇指の MP、IP、他の指の PIP、DIP を曲げ、abduction が広がるような手形状をつくっているのに対し、陰影除去画像では、拇指の MP、IP、他の指の PIP、DIP が伸び、abduction が閉じていることからわかる。被験者 B の角柱における手形状は被験者 A の角柱における手形状と同じ

ような振舞いをしている。円柱において、無処理画像では、拇指のIP、その他の指のPIPとDIPを曲げるような円柱に対する手形状をつくったのに対して、陰影除去画像では、拇指のIP、他の指のPIP、DIPを伸ばし、MP関節を曲げるような薄い平面的な物体に対する手形状をつくったと考えられる。被験者Bの球からは被験者がどのような対象物に対する手形状をつくったかを推測するのは難しいが、陰影除去画像で大きく手形状が異なっている。但し、被験者Bは「5本の指先を全て使って」と指示されているにもかかわらず、サイズの小さい対象物では薬指と小指には意識がいかず、拇指、示指、中指のみで手形状をつくる傾向にあったため、主に拇指、示指、中指の振舞いを比較している。被験者Cの角柱と円柱における手形状は被験者Aの角柱における手形状と同じような振舞いをしている。被験者Cの球は被験者Aの球と同じような振舞いをしている。これらのことから、全ての被験者はほとんどの対象物の陰影除去画像では、薄い平面的な対象物に対する手形状をつくったために、無処理画像の手形状と大きく異なったことがわかる。それに対し、ぼかし画像では無処理画像とほぼ同じような対象物を掴むための手形状をつくったため、比較的近くなったと考えられる。

4. 神経回路モデルによるシミュレーション結果と心理物理実験結果の比較検討

2.1節の神経回路モデルによるシミュレーションの結果は、視覚情報から手形状への視覚-運動変換の際に、手形状を計算するのに適した大きさと形に対応するパラメータが第3層に抽出されること（把持運動のための特徴抽出）を示していた。2.2節のシミュレーション結果は、無処理画像と同じ特徴が抽出できなくなるほど変更された陰影除去画像からは元の画像に対する手形状が計算できなくなり、無処理画像とほぼ同じ特徴を抽出できたぼかし画像からは無処理画像に対するものに近い手形状を計算できることを示していた。3章の心理物理実験の結果は、ヒトにおいても陰影除去画像からは無処理画像に対する把持とは大きく異なる手形状をつくり、ぼかし画像では近い手形状をつくることを示していた。すなわち、神経回路が出力した手形状とヒトが計画・計算した手形状が同じ傾向を示したことになる。

本論文の神経回路モデルは、非常に単純な構造であり、その学習アルゴリズムも人工的であるので、ヒトの脳が実際に行っている情報処理とは程遠いものである。しかしながら、神経回路モデルによるシミュレ-

ションと心理物理実験の結果を比較することで、以下で述べるように脳が解くべき計算論的問題の可能性を論じることができる。2.2.3節で述べたように、神経回路モデルにおいて、陰影除去画像で手形状が大きく異なった原因は、輝度情報が一様に高くなったために、抽出された特徴が大きくシフトしたからである。それに対して、3.3節で述べたように、ヒトにおいて、陰影除去画像で手形状が大きく異なった原因は、ヒトが薄い平面的な物体に対する手形状をつくったからである。このように神経回路モデルとヒトの脳とは、入出力関係は同じ傾向を示しても手形状を求めるアルゴリズムに違いがある。しかし、神経回路モデルで対象物の画像情報から第3層のわずかなユニットで特徴が抽出され、その特徴から手形状が計算できたことから、ヒトにおいても把持運動のための特徴情報が抽出された後、手形状が求められる可能性がある（計算論的可能性）。更に、ヒトで陰影除去画像が与えられたときに手形状が大きく異なるのは、神経回路モデルと同様に、把持運動のための適切な特徴情報が抽出できなかった可能性がある。

問題は、脳が画像からどのような情報を抽出しているかであるが、シミュレーション結果と心理物理実験結果の比較により、これを明確にすることはできない。なぜなら、神経回路モデルは角柱、円柱、球しか学習しておらず、それらの特徴の空間内で入力画像に対する特徴を計算してしまうからである。おそらく、神経回路モデルが薄い平面的な物体を学習していれば、陰影除去画像に対する特徴はその物体の特徴情報に近い値となり、手形状もそれに基づいて計算されるであろう。そのときの、神経回路モデルの出力とヒトがつくった手形状が同じ傾向となれば、ヒトにおいても、陰影除去画像に対して薄い平面的な物体に対応する特徴が得られている可能性を示すことになると考えられる。今後、心理物理実験において無処理画像と同じように対象物を認識できるが把持のための手形状は異なる、あるいは、その逆となるような処理画像を見出すことができ、更に、その実験結果を再現できる神経回路モデルを構築できれば、把持運動に必要不可欠な情報について明快に述べることができるだろう。

5. おわりに

本論文では、ヒトの把持運動の視覚-運動変換において、入力される対象物の視覚情報、処理過程で抽出される対象物の特徴情報、出力される手形状の関係を調べ、考察を行った。まず、視覚情報から手形状を計算する神経回路モデルのシミュレーションより、視

覚-運動変換の際に、対象物の特徴が抽出され得ることを示す結果が得られ、そしてその特徴は把持運動情報の計算に適した特徴情報（把持運動のための対象物の特徴情報）であることを主張した。次に、その神経回路モデルを用いて、視覚情報の変更が特徴情報の抽出と手形状の計算にどう影響するかを調べた。その結果、陰影除去画像では特徴情報も手形状も無処理画像のものとは大きく異なり、ぼかし画像では比較的無処理画像のものに近くなった。このことから、適切な手形状が計算されるためには、適切な特徴情報が視覚情報から抽出されなければならないことを主張した。次に、心理物理実験により、ヒトにおいて視覚情報の変更が手形状にどう影響するかを調べたところ、シミュレーション結果と同様に、陰影除去画像では特徴情報も手形状も無処理画像のものとは大きく異なり、ぼかし画像では無処理画像のものに比較的近くになっていた。最後にシミュレーションと心理物理実験の結果を比較検討した結果から、把持運動における視覚-運動変換の際に、ヒトの脳内でも把持運動のための特徴抽出が行われ、その抽出結果が手形状計算に大きく影響する計算論的可能性を主張した。

我々が、「把持運動のための対象物の特徴情報」という言葉を使い、(どのような幾何学的形状をしているか、サイズはどのくらいかというような) 視覚的同定のための対象物の特徴情報とは異なるものと考えているのは次の理由からである。Sakata ら³⁾ は、サルが様々な種類とサイズの対象物を固視するか把持したときに、その大脳皮質の前部頭頂間領野 (AIP 野) にそれらの対象物に対して選択性のあるニューロンが存在することを発見し、手や腕の運動に関する対象物の3次元構造的な場所に関する情報は最終的に AIP 野で処理されることを主張している。また、過去の様々な研究より、第一次視覚野から始まり、視覚前野の腹側部を通して下側頭領野へ投射する流れ (腹側経路) で、対象物がどのようなものであるかを認識すると考えられている。これらの知見から、把持運動では、AIP 野もしくはその関連領野と、腹側経路中の領野で対象物の特徴抽出が行われると考えられる。しかしながら、それらの領野で全く同じ特徴抽出を行っているとは考え難く、腹側経路では対象物を認識するための処理に必要不可欠な特徴のみが抽出され、AIP 関連領野では把持運動情報を計算する処理に必要不可欠な特徴のみが抽出されると考えられる。前者が我々のいう「視覚的同定のための対象物の特徴」で、後者が「把持運動のための対象物の特徴」に対応する。上述した

「2つの経路」と「把持運動のための特徴」を仮定するならば、心理物理実験において、陰影除去画像に対してヒトが薄い平面的な物体と認識したために手形状が大きく異なったとは必ずしも解釈できない。そうではなく、把持運動情報を計算するための特徴として薄い平面的な物体に対応するものが抽出され、それが手形状の計算に使われた可能性がある。

更に我々は、把持運動のための視覚情報処理の最終目標は「把持運動に適した表現の特徴情報をつくり出すこと」であることを提案する。それは単なる対象物の3次元立体形状ではなく、より把持に密接な、例えば把持点や手形状の情報を含んだような特徴情報である。そして、その特徴情報から、把持運動に関する対象物の内部表現に基づいて、把持運動情報が計算されるのである。我々の主張する把持運動に適した特徴情報は AIP 野もしくはその関連領野で獲得され、そして内部表現は AIP 野から運動前野の F5 野にかけて存在していると期待している。また、現段階ではまだ把持運動にとって essential な特徴情報が定性的、定量的に明らかになっていない。そこで今後は、手形状だけではなく、把持点も考慮したシミュレーションや心理物理実験を通して、そのことを明らかにすることを目的として研究を進めていく。

謝 辞 有益な御討論を賜った豊橋技術科学大学情報工学系 片山正純講師に深く感謝いたします。また本研究の一部は、平成 11 年度科学研究費補助金基盤研究(c)(2)09650467、科学技術振興調整費「ヒトを含む霊長類のコミュニケーションの研究」、および栢森情報科学振興財団研究助成金によって行われた。

参 考 文 献

- 1) 福田浩士, 福村直博, 片山正純, 宇野洋二 (1999): 対象物の認知と手の把持形状の計算との関係—ヒトの把持運動への計算論的アプローチ—, 信学論 D-II, Vol. J 82-D-II, No. 8, pp. 1315-1326
- 2) Marr, D. (1982): Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. New York: W. H. Freeman and Company. 乾 敏郎, 安藤広志訳 (1987): ビジョン—視覚の計算理論と脳内表現—. 東京: 産業図書
- 3) Jeannerod, M., Arbib, M. A., Rizzolatti, G. and Sakata, H. (1995): Grasping objects: The cortical mechanisms of visuomotor transformation, Trends in Neurosciences, Vol. 18, pp. 314-320