

標準正則化理論に基づく面再構成過程のシミュレーション

白岩 史*¹ 奥畑 亮馬*² 佐賀野正行*³ 林 武文*⁴

要 旨

初期視覚における不良設定の情報復元問題は、標準正則化理論の枠組みにおいて定式化され統一的な扱いが可能であることが示されている。その中で面の再構成に用いられる滑らかさの拘束条件には、薄膜あるいは薄板の曲げポテンシャルエネルギーに対応するエネルギー汎関数が提案されているが、両者の比較と妥当性の検証は必ずしも十分になされていない。本研究では、数値実験により、薄膜と薄板のポテンシャルエネルギーに基づく面の再構成計算を比較し、両者の間の差違を定量的に評価した。また、ランダムドットステレオグラムを用いた面の知覚に関する心理物理実験により、視覚系の行っている計算が薄膜ポテンシャルエネルギーに基づくものに近いことを示した。

Numerical Simulation of Surface Reconstruction Process Based on Standard Regularization Theory

Aya SHIRAIWA*¹ Ryouma OKUBATA*²
Masayuki SAGANO*³ Takefumi HAYASHI*⁴

Abstract

Standard Regularization Theory has been applied to solve the ill posed problems appearing in the early stages of human visual processing. In the case of the surface reconstruction process from retinal images, the perceived surface can be determined to minimize a certain type of energy functional. As the energy functional, two types of formulas based on the, "membrane potential" and "thin plate potential", have been applied. In the present research, the surface reconstruction of human early vision was simulated using these two energy functionals and results were compared quantitatively. Furthermore, a psychophysical experiment was performed to study the properties of surface perception from binocular disparity using a random dot stereogram. The experimental results indicate that the energy functional based on the membrane potential energy can predict human performance more accurately.

*¹ 関西大学大学院総合情報学研究科

*² 任天堂株式会社

*³ 奈良先端科学技術大学院大学

*⁴ 関西大学総合情報学部

1. まえがき

我々は3次元世界の情報を、両眼の2次元網膜で受容している。従って、我々の脳内では2次元網膜像から3次元世界の構造を再構成していることになる。Marrは、彼の著書『ビジョン』の中で、視覚の初期過程（初期視覚）の目的を「網膜に投影された二次元画像データから三次元世界の可視表面の幾何学的構造を推測復元すること」と定義付けている^[1]。しかし、元の3次元世界を復元するためには、何らかの先験的知識に基づく制約条件など、事前知識を用いなければ、初期視覚の情報復元問題は、一般に不良設定問題であるため解くことが出来ない。3次元世界が2次元網膜上に投影されたときに次元が一つ落ちてしまうため、解が一意に求まらないのである。この2次元網膜像から3次元世界を復元する問題を逆問題と呼び、Tikhonovらの標準正則化理論^[2]では、逆問題に対して、2次形式で表せるような制約条件を解に付加して解いている。また、Poggioら^[3]によって、視覚計算理論で提案されてきた幾つかのアルゴリズムが、標準正則化理論の枠組みを用いて統一的に理解できると提案された。

標準正則化による面再構成は、本質的にはスプライン平滑化と同じであり、離散的に与えられたデータにノイズが含まれている場合に、ノイズを取り除きつつ離散データを補間するというものである^[4]。近年、この正則化法で、エネルギー汎関数の最小化問題を解くことによる画像から3次元情報の復元が、コンピュータビジョンのさまざまな問題に適用され、等高線画像とわずかな標高データから詳細な地形形状を復元することが試みられている^{[5], [6]}。

一方、Poggioらの提案した、面の再構成についての正則化汎関数における拘束条件は、一般に滑らかさの拘束の微分演算子が用いられる。この微分演算子は、1次微分と2次微分の場合が考えられ、それぞれ薄膜ポテンシャルエネルギーと薄板ポテンシャルエネルギーに対応している^[7]。

Grimsonは、検出された奥行きを通り、画像放射照度に最もよく適合する面は、薄板ポテンシャルエネルギーを最小にすると証明した^[8]。ただし、用いられている2次微分の微分演算子は正確な物理的な薄板ポテンシャルエネルギーでない^{[7], [9]-[12]}。また、方向の不連続を有する面を薄板ポテンシャルエネルギーで再構成を行うと、角の部分が盛り上がってしまう。Terzopoulosは、面の方向不連続部分を表現するために薄膜ポテンシャルエネルギーを導入し、薄板ポテンシャルエネルギーと区別した^[13]。これら2つのポテンシャルエネルギーの特性のうち、心理物理学的なデータによく合うのは薄膜ポテンシャルエネルギーであることが指摘されている。

本研究は、人間の視覚系で行われていると考えられる面の再構成をコンピュータ上で実行させることにより、人間の視覚機能を解明することを目的としている。そのために、数値実験により拘束条件である薄膜と薄板のポテンシャルエネルギーについての違いを比較し、両者の差異を定量的に評価した。また、両眼視差による面再構成について心理物理実験により人間の特性の評価を行い、両ポテンシャルエネルギーの生体モデルとしての妥当性を検討した。

2. 数値計算

2次元網膜像から3次元世界を復元する逆問題を解くためには、まず、この逆問題を定式化し、標準正則化問題を導かなければならない。本章では、標準正則化問題を解くための一般的な枠組みである変分原理を説明し、オイラー・ラグランジュ方程式を導く。

2.1 変分原理

ある関数 $f(x)$ が x_0 で極値になる必要条件是、 $f(x)$ の微係数が0、つまり $df(x)/dx|_{x=x_0} = 0$ が成り立つことである。このとき $d^2f(x)/dx^2|_{x=x_0} = 0$ が正、負、0の場合に応じて、 x_0 はそれぞれ極小、極大、変曲点となり、これらをまとめて停留点といい、そこでの関数の値を停留値という。

変分原理は極値問題という点では上と同じであるが、上で述べたものが変数の関数についての極値問題であるのに対して、変分原理は関数の関数についての極値問題である。関数の関数を汎関数というので、変分原理とは、汎関数の極値（停留値）を与えるような関数の形と、そのときの汎関数の値を決める問題である。

x, y を独立変数、 $z(x, y)$ を x と y の関数とする。 x, y, z, z_x, z_y の関数 $F(x, y, z, z_x, z_y)$ を x と y について積分したもの、

$$I = \iint F(x, y, z, z_x, z_y) dx dy \quad (1)$$

を極大、または極小にする $z(x, y)$ を求めるのが変分原理である。境界条件として積分域の端 $((x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2))$ での z の値は、

$$z(x_1, y_1) = z_1, \quad z(x_1, y_2) = z_2, \quad z(x_2, y_1) = z_3, \quad z(x_2, y_2) = z_4 \quad (2)$$

と決まっていると仮定する。 F は、 x, y, z, z_x, z_y に関する二階微分までが連続であるとする。積分 I は x と y に関する定積分なので、関数 $z(x, y)$ を決めることによって決まる汎関数である。

2.2 オイラー・ラグランジュ方程式

2.1と同様にして、微分可能な任意の関数 $\zeta(x, y)$ を考える。ただし $\zeta(x, y)$ は、積分区間の端の境界での値は0とする。つまり、

$$\zeta(x_1, y_1) = \zeta(x_1, y_2) = \zeta(x_2, y_1) = \zeta(x_2, y_2) = 0. \quad (3)$$

そして小さなパラメータ α をもつ新しい関数、

$$Z(x, y) = z(x, y) + \alpha \zeta(x, y)$$

を考える。 $Z(x, y)$ は、積分 I に極値を与える関数を $z(x, y)$ としたとき、この $z(x, y)$ と比較するための関数である。 $\zeta(x, y)$ が任意の関数であることから $Z(x, y)$ も任意の関数である。また $Z(x, y)$ は式(3)から、 $z(x, y)$ と同じ境界条件 (式(2)) を満たしている。境界条件 (式(2)) で決まる4辺を端とすると、 $Z(x, y)$ は α が小さいとき、 $z(x, y)$ と端の境界を共有している

境界近くの曲面群を表す。

この $Z(x, y)$ を式(1)に示した汎関数の $z(x, y)$ と置き換えると I はパラメータ α の関数となる。

$$I(\alpha) = \iint F(x, y, z + \alpha\zeta, z_x + \alpha\zeta_x, z_y + \alpha\zeta_y) dx dy. \quad (4)$$

$I(\alpha)$ は α の関数で、 $\alpha = 0$ のとき Z は z に一致する。したがって $\alpha = 0$ のとき $I(\alpha)$ は極値をとる。つまり z が汎関数 $I(\alpha)$ に極値を与える関数となる必要条件是、

$$\left. \frac{dI(\alpha)}{d\alpha} \right|_{\alpha=0} = 0 \quad (5)$$

である。汎関数 I を差分を用いて表すと、

$$I(\alpha) - I(0) = \iint \{F(x, y, z + \alpha\zeta, z_x + \alpha\zeta_x, z_y + \alpha\zeta_y) - F(x, y, z, z_x, z_y)\} dx dy. \quad (6)$$

$F(x, y, z + \alpha\zeta, z_x + \alpha\zeta_x, z_y + \alpha\zeta_y)$ を変数 α でテーラー展開すると、

$$\begin{aligned} F(x, y, z + \alpha\zeta, z_x + \alpha\zeta_x, z_y + \alpha\zeta_y) &= F(x, y, z, z_x, z_y) + \frac{\partial F}{\partial x}(x-x) + \frac{\partial F}{\partial y}(y-y) \\ &\quad + \frac{\partial F}{\partial z}(z + \alpha\zeta - z) + \frac{\partial F}{\partial z_x}(z_x + \alpha\zeta_x - z_x) \\ &\quad + \frac{\partial F}{\partial z_y}(z_y + \alpha\zeta_y - z_y) + \frac{\alpha^2}{2}\{\dots\} \\ &= F(x, y, z, z_x, z_y) + \alpha \left(\frac{\partial F}{\partial z}\zeta + \frac{\partial F}{\partial z_x}\zeta_x + \frac{\partial F}{\partial z_y}\zeta_y \right) \\ &\quad + \frac{\alpha^2}{2}\{\dots\}. \end{aligned}$$

これを式(6)に代入すると、

$$I(\alpha) - I(0) = \alpha \iint \left(\frac{\partial F}{\partial z}\zeta + \frac{\partial F}{\partial z_x}\zeta_x + \frac{\partial F}{\partial z_y}\zeta_y \right) dx dy + \frac{\alpha^2}{2} \iint \{\dots\} dx dy.$$

右辺の第2項以降は非常に小さい値となり無視できるので、第1項を δI とおき、

$$\delta I = \alpha \iint \left(\frac{\partial F}{\partial z}\zeta + \frac{\partial F}{\partial z_x}\zeta_x + \frac{\partial F}{\partial z_y}\zeta_y \right) dx dy. \quad (7)$$

これを I の第1変分という。また式(4)を α で微分すると、

$$\begin{aligned} \frac{dI(\alpha)}{d\alpha} &= \iint \left(\frac{\partial F}{\partial Z} \frac{\partial Z}{\partial \alpha} + \frac{\partial F}{\partial Z_x} \frac{\partial Z_x}{\partial \alpha} + \frac{\partial F}{\partial Z_y} \frac{\partial Z_y}{\partial \alpha} \right) dx dy \\ &= \iint \left(\frac{\partial F}{\partial Z}\zeta + \frac{\partial F}{\partial Z_x}\zeta_x + \frac{\partial F}{\partial Z_y}\zeta_y \right) dx dy. \end{aligned}$$

しかも $\alpha = 0$ で $Z = z(x, y)$, $Z_x = z_x(x, y)$, $Z_y = z_y(x, y)$ となるから、式(5)は、

$$\left. \frac{dI(\alpha)}{d\alpha} \right|_{\alpha=0} = \iint \left(\frac{\partial F}{\partial z}\zeta + \frac{\partial F}{\partial z_x}\zeta_x + \frac{\partial F}{\partial z_y}\zeta_y \right) dx dy = 0. \quad (8)$$

式(7)と式(8)を比べると、 $I'(0) = 0$ の条件は $\delta I = 0$ と同じであることがわかる。この式を $(\partial F/\partial z_x)\zeta_x$ について部分積分すると、

$$\frac{dI(\alpha)}{d\alpha}\Big|_{\alpha=0} = \int \left\{ \left[\frac{\partial F}{\partial z_x} \zeta \right]_{x_1}^{x_2} + \int \left(\frac{\partial F}{\partial z} \zeta - \frac{\partial}{\partial x} \frac{\partial F}{\partial z_x} \zeta + \frac{\partial F}{\partial z_y} \zeta_y \right) dx \right\} dy.$$

右辺の第1項は式(3)より0となるので、

$$\frac{dI(\alpha)}{d\alpha}\Big|_{\alpha=0} = \iint \left\{ \zeta \left(\frac{\partial F}{\partial z} - \frac{\partial}{\partial x} \frac{\partial F}{\partial z_x} \right) + \frac{\partial F}{\partial z_y} \zeta_y \right\} dx dy.$$

次に $(\partial F/\partial z_y)\zeta_y$ について部分積分すると、

$$\frac{dI(\alpha)}{d\alpha}\Big|_{\alpha=0} = \iint \left\{ \left[\frac{\partial F}{\partial z_y} \zeta \right]_{y_1}^{y_2} + \int \left(\zeta \left(\frac{\partial F}{\partial z} - \frac{\partial}{\partial x} \frac{\partial F}{\partial z_x} \right) - \frac{\partial}{\partial y} \frac{\partial F}{\partial z_y} \zeta \right) dy \right\} dx.$$

右辺の第1項は、式(3)より0となるので、

$$\frac{dI(\alpha)}{d\alpha}\Big|_{\alpha=0} = \iint \zeta \left(\frac{\partial F}{\partial z} - \frac{\partial}{\partial x} \frac{\partial F}{\partial z_x} - \frac{\partial}{\partial y} \frac{\partial F}{\partial z_y} \right) dx dy = 0.$$

この式は、 ζ の任意の値において成立しなければならないので、必要条件(式(6))は、

$$\frac{\partial F}{\partial z} - \frac{\partial}{\partial x} \left(\frac{\partial F}{\partial z_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial z_y} \right) = 0 \quad (9)$$

となる。式(9)はオイラー・ラグランジュの方程式に相当する。

2.3 滑らかさの拘束

既に述べたように視覚計算論のうちの面の再構成など、情報復元問題は、一般的に不良設定問題であり、求める解に対して何らかの制約条件を付与する必要がある。しかもこの制約条件は、我々の先験的知識を反映したものでなければならない。面再構成の場合の制約条件とは、ほとんどの場合、「可視表面は滑らかで連続である(滑らかさの拘束)」という制約条件が用いられる。ここでは滑らかさの拘束として用いられる汎関数の物理的意味について述べる。

(1) 薄膜ポテンシャルエネルギー

膜が変形するときのポテンシャルエネルギーは面積の変化に比例し、この比例定数を張力という。膜を $z = f(x, y)$ としたとき、その表面積 S は、

$$S = \iint (1 + z_x^2 + z_y^2)^{\frac{1}{2}} dx dy$$

で与えられ、 $(1+x)^\alpha \cong 1 + \alpha x$ を用いると、

$$S \cong \iint \left(1 + \frac{z_x^2 + z_y^2}{2} \right) dx dy$$

となる。したがって定数項を除いた膜のポテンシャルエネルギーは、

$$\frac{1}{2} \iint (z_x^2 + z_y^2) dx dy$$

で表される。

(2) 薄板ポテンシャルエネルギー

xy 平面と yz 平面とが板のたわみ曲面の主面になるように、等分布曲げモーメントによって薄板が曲げられる場合には、薄板に蓄えられるポテンシャルエネルギーは次式のように表される^[14]。

$$\frac{1}{2} D \iint \left\{ \left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} \right)^2 - 2(1-\nu) \left[\frac{\partial^2 z}{\partial x^2} \frac{\partial^2 z}{\partial y^2} - \left(\frac{\partial^2 z}{\partial x \partial y} \right)^2 \right] \right\} dx dy.$$

D は板の曲げに対する抵抗値、 ν は横歪みと縦歪みの比を表すポアソン比である。実際の金属材料では、 ν は 0.2~0.3 程度の値をとるが、ここでは $\nu = 0$ として簡略化した式を薄板の曲げのポテンシャルエネルギーとして用いる。

$$\frac{1}{2} \iint (z_{xx}^2 + 2z_{xy}^2 + z_{yy}^2) dx dy.$$

2.4 標準正則化理論に基づく問題の定式化

3次元空間 z に対して、 A という変換を施した結果、その像である2次元画像 y が得られたとする。

$$y = A \cdot z. \quad (10)$$

ここでは、 y は網膜に投影された2次元画像の輝度情報を、また z は面の奥行き情報を表すベクトルであり、 A は線形変換であるとしている。視覚システムが解くべき問題は、 y から z を逆計算することに相当する。すなわち、

$$z = A^{-1} \cdot y$$

である。しかし、一般に行列 A は逆行列を持たないため式(10)は解くことができない。標準正則化理論では、与えられたデータ y をよく近似し、ある種の拘束条件を満足する解 z を求めるために、新たに次のような評価関数を定義し、これを最小化するような z を見いだす。

$$\|A \cdot z - y\|^2 + \lambda \|P \cdot z\|^2 \quad (11)$$

式(11)の第1項は、式(10)の両辺の差のノルムの2乗であり、観測データと解の適合度を表す拘束条件である。また第2項は、 z に関する先験的な知識を表す作用素 P との積のノルムの2乗であり、先験的な知識に関する拘束条件を表す。一般に観測データにはノイズや欠落が含まれるため、第2項を大きくするような解は避けるべきである。 λ は、これら2つの拘束条件の重み付けを行う正則化パラメータであり、観測データと先験的な知識のどちらに重きを置いて解を求めるか、問題によって適切な値を定める必要がある。

標準正則化理論で最も重要なことは、3次元世界のモデルをうまく P の選択に反映させる

ことといえる。滑らかさの拘束では、前節に述べた通り P は z の 1 階微分もしくは 2 階微分を求めるオペレータとなり、式(11)は次のように表される。

$$\iint [\lambda(z_x^2 + z_y^2) + (Sz - d)^2] dx dy. \quad (12)$$

$$\iint [\lambda(z_{xx}^2 + 2z_{xy}^2 + z_{yy}^2) + (Sz - d)^2] dx dy. \quad (13)$$

積分範囲は、網膜上の投影像の全ての領域である。

式(12)が薄膜ポテンシャルエネルギー汎関数に基づく汎関数で、式(13)が薄板ポテンシャルエネルギー汎関数に基づく汎関数である。ここで、 S はデータが観測された点のみを抽出するサンプリング行列、 d は観測データである。

滑らかさの拘束条件が式(12)と式(13)の第 1 項にあたり、観測データの解の適合度を表す拘束条件が式(12)と式(13)の第 2 項にあたる。両評価関数が最適な解においてその値が最小となるという条件を満たせば、式(12)や式(13)を最小化することによって最適な 3 次元形状モデルが得られる。

(1) 薄膜ポテンシャルエネルギーに対応するオイラー・ラグランジュ方程式

$F = \lambda(z_x^2 + z_y^2) + (Sz - d)^2$ とおいて、式(9)に式(12)を代入すると、

$$\frac{\partial F}{\partial z} - \frac{\partial}{\partial x} \left(\frac{\partial F}{\partial z_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial z_y} \right) = 0$$

となる。よって薄膜ポテンシャルエネルギーに対応するオイラー・ラグランジュ方程式は、

$$-\lambda(z_{xx} + z_{yy}) + (Sz - d) = 0$$

となり、これが解くべき方程式である。なお、 S は前述のサンプリング行列であるため、左辺第 2 項は観測された点以外は考慮しないこととしている。

間隔 h を持つ正方格子にデータがあるものとして中心差分を用いれば、2 次微分は、

$$z_{xx} = \frac{(z_{i-1,j} - 2z_{i,j} + z_{i+1,j})}{h^2},$$

$$z_{yy} = \frac{(z_{i,j-1} - 2z_{i,j} + z_{i,j+1})}{h^2}$$

と表現できるので、

$$-\lambda \frac{(z_{i-1,j} + z_{i+1,j} + z_{i,j-1} + z_{i,j+1} - 4z_{i,j})}{h^2} + (Sz - d) = 0$$

が得られる。 $h = 1$ としてヤコビ法による弛緩アルゴリズムで表現すれば、

$$z_{i,j}^{(n+1)} = \lambda \frac{z_{i-1,j}^{(n)} + z_{i+1,j}^{(n)} + z_{i,j-1}^{(n)} + z_{i,j+1}^{(n)}}{4} - \frac{(Sz - d)}{4}$$

となり、これを繰り返して解けばいい。

(2) 薄板ポテンシャルエネルギーに対応するオイラー・ラグランジュ方程式

薄板ポテンシャルエネルギーの場合には、2.3(2)に示した通り、エネルギー関数は $x, y, z, z_{xx}, z_{xy}, z_{yy}$ の関数であり2.2と同様にオイラー・ラグランジュ方程式を導出すると次のようになる。

$$\frac{\partial F}{\partial z} + \frac{\partial^2}{\partial x^2} \left(\frac{\partial F}{\partial z_{xx}} \right) + 2 \frac{\partial^2}{\partial x \partial y} \left(\frac{\partial F}{\partial z_{xy}} \right) + \frac{\partial^2}{\partial y^2} \left(\frac{\partial F}{\partial z_{yy}} \right) = 0.$$

$F = \lambda(z_{xx}^2 + 2z_{xy}^2 + z_{yy}^2) + (Sz - d)^2$ とおいて、上式に代入する。その結果、

$$\lambda(z_{xxxx} + 2z_{xyyy} + z_{yyyy}) + (Sz - d) = 0$$

となり、これが解くべき方程式である。

間隔 h を持つ正方格子にデータがあるものとして中心差分を用いれば、4次微分はそれぞれ、

$$\begin{aligned} z_{xxxx} &= \frac{\partial^2}{\partial x^2} \frac{(z_{i-1,j} - 2z_{i,j} + z_{i+1,j})}{h^2} \\ &= \frac{(z_{i-2,j} - 2z_{i-1,j} + z_{i,j}) - 2(z_{i-1,j} - 2z_{i,j} + z_{i+1,j}) + (z_{i,j} - 2z_{i+1,j} + z_{i+2,j})}{h^4} \\ &= \frac{z_{i-2,j} - 4z_{i-1,j} + 6z_{i,j} - 4z_{i+1,j} + z_{i+2,j}}{h^4}, \\ z_{xyyy} &= \frac{\partial^2}{\partial y^2} \frac{(z_{i-1,j} - 2z_{i,j} + z_{i+1,j})}{h^2} \\ &= \frac{(z_{i-1,j-1} - 2z_{i-1,j} + z_{i-1,j+1}) - 2(z_{i,j-1} - 2z_{i,j} + z_{i,j+1}) + (z_{i+1,j-1} - 2z_{i+1,j} + z_{i+1,j+1})}{h^4} \\ &= \frac{z_{i-1,j-1} - 2z_{i-1,j} + z_{i-1,j+1} - 2z_{i,j-1} + 4z_{i,j} - 2z_{i,j+1} + z_{i+1,j-1} - 2z_{i+1,j} + z_{i+1,j+1}}{h^4}, \\ z_{yyyy} &= \frac{\partial^2}{\partial x^2} \frac{(z_{i,j-1} - 2z_{i,j} + z_{i,j+1})}{h^2} \\ &= \frac{(z_{i,j-2} - 2z_{i,j-1} + z_{i,j}) - 2(z_{i,j-1} - 2z_{i,j} + z_{i,j+1}) + (z_{i,j} - 2z_{i,j+1} + z_{i,j+2})}{h^4} \\ &= \frac{z_{i,j-2} - 4z_{i,j-1} + 6z_{i,j} - 4z_{i,j+1} + z_{i,j+2}}{h^4} \end{aligned}$$

と表現できるので、

$$\begin{aligned} \frac{\lambda}{h^2} (z_{i,j-2} + z_{i,j+2} + z_{i-2,j} + z_{i+2,j} + 2z_{i-1,j-1} + 2z_{i-1,j+1} + 2z_{i+1,j-1} + 2z_{i+1,j+1} \\ - 8z_{i-1,j} - 8z_{i+1,j} - 8z_{i,j-1} - 8z_{i,j+1} + 20z_{i,j}) + (Sz - d) = 0 \end{aligned}$$

が得られる。 $h = 1$ としてヤコビ法による弛緩アルゴリズムで表現すれば、

$$z_{i,j}^{(n+1)} = -\frac{\lambda}{20} (z_{i,j-2}^{(n)} + z_{i,j+2}^{(n)} + z_{i-2,j}^{(n)} + z_{i+2,j}^{(n)} + 2z_{i-1,j-1}^{(n)} + 2z_{i-1,j+1}^{(n)} + 2z_{i+1,j-1}^{(n)} + 2z_{i+1,j+1}^{(n)} - 8z_{i-1,j}^{(n)} - 8z_{i+1,j}^{(n)} - 8z_{i,j-1}^{(n)} - 8z_{i,j+1}^{(n)}) - \frac{(Sz-d)}{20}$$

となり、これを繰り返して解けばいい。

2.5 ヤコビ法

上で述べたように、標準正則化理論に基づく汎関数に対応するオイラー・ラグランジュ方程式を解くことによって面を再構成しているのだが、このオイラー・ラグランジュ方程式はその計算する座標（格子点）の4近傍もしくは12近傍の座標の値が必要になる。さらに全座標ごとに離散化することによって得られる連立方程式を解かなければならないので、この連立方程式の係数行列は大規模なものとなることが多い。

大規模の連立方程式の数値解析法としての弛緩法（あるいは緩和法）には、ガウスザイデル法、ヤコビ法、加速緩和（SOR）などがあり、あらゆる分野で広く使われている^{[15], [16]}。そこで本研究ではその中のヤコビ法による弛緩アルゴリズムによってこの問題を解いた。そのヤコビ法について簡単に説明する。

まず次のような n 次元の1次連立方程式を考える。

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n. \end{aligned}$$

ヤコビ法では、まず各行の非対角項を右辺に移動させて、対角項にかかる係数で割る。

$$\begin{aligned} x_1 &= \frac{1}{a_{11}} \{b_1 - (a_{12}x_2 + \cdots + a_{1n}x_n)\}, \\ x_2 &= \frac{1}{a_{22}} \{b_2 - (a_{21}x_1 + \cdots + a_{2n}x_n)\}, \\ &\vdots \\ x_n &= \frac{1}{a_{nn}} \{b_n - (a_{n1}x_1 + \cdots + a_{nn}x_{n-1})\}. \end{aligned}$$

これを一般式で書くと、

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j \right) \quad (i = 1, 2, \dots, n) \quad (15)$$

となる。ヤコビ法では、上式を反復公式とし次のアルゴリズムによって解を求める。

- ① 解の初期値 $x_i^{(0)}$ を仮定する.
- ② $k = 1$ とする.
- ③ 式(15)より導いた次の反復公式により解を更新していく.

$$x_i^{(k)} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ i \neq j}}^n a_{ij} x_j^{(k-1)} \right) \quad (i = 1, 2, \dots, n).$$

- ④ 任意の正のパラメータ ε に対して, 解が次の判定式を満たすとき, その解を正解として反復計算を終える.

$$|x_i^{(k)} - x_i^{(k-1)}| \leq \varepsilon.$$

- ⑤ 判定式を満たさない場合は $k = k+1$ としてステップ③に進む.

3. 数値実験

3.1 薄膜ポテンシャルエネルギーと薄板ポテンシャルエネルギーの比較

2次元の格子点上にランダムに離散的に与えられた奥行きデータから面を再構成する場合のシミュレーションを行った. これは, 後述するランダムドットステレオグラムから面が知覚される現象に対応しており, 対応点の視差からその点の奥行きが検出された後の過程をシミュレートしていることになる.

図1のように50×50の座標を作成し, その格子点上に図2のようにランダムドットを配置して, 前章で述べた方法により, 面の再構成を行った. 図3に薄膜ポテンシャルを利用した場合の面が再構成されるシミュレーションの様子を示す. (a)の初期段階では, 離散的に与えられた奥行き値(観測データ)とその他の点の初期値(奥行き0とした)が示されており, 繰り返し計算により, (b), (c)のように面が復元されていくことがわかる. この場合では, (d)に示す10回程度の反復計算で解が収束している.

すべての格子点における正しい面上の奥行き値と再構成計算による値との差の総和(エラー)を求めることによって, 薄膜ポテンシャルエネルギーと薄板ポテンシャルエネルギーの2つの汎関数の比較を行った. 再構成するモデルはsinカーブ曲面を用い, ノイズは付与していない.

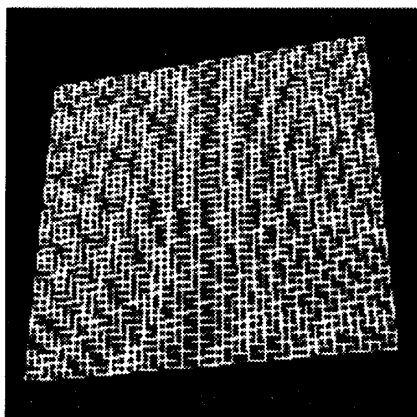


図1 50×50の座標図



図2 ランダムドット

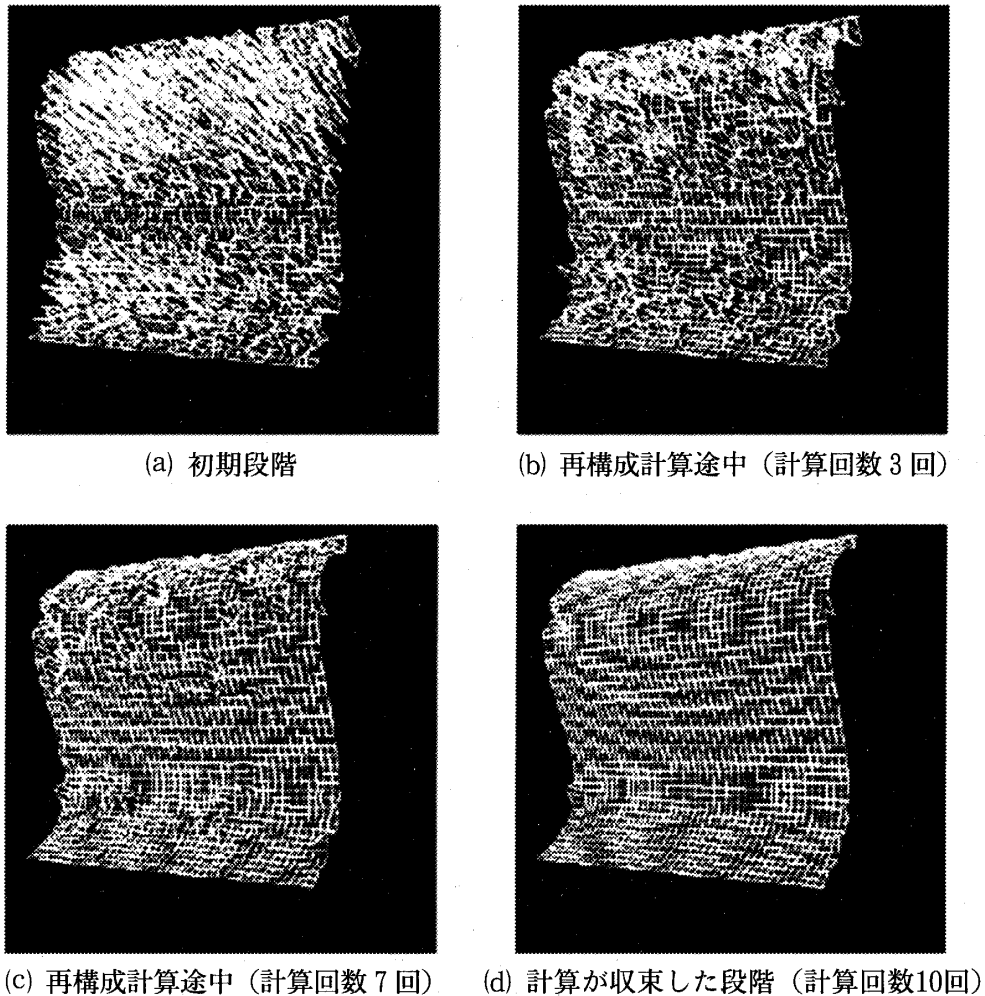


図 3 面再構成のシミュレーションの様子

なお、数値計算のプラットフォームは Windows PC を用い、開発環境は MS-Visual C/C++ とグラフィックス API (OpenGL) を利用した。

3.2 実験および結果

[実験 1] ドット数とエラーの比較

横軸にランダムドットの密度 (ドット数) を、縦軸にエラーの値をとったグラフを図 4 に示し、ドットの密度とエラーの関係を比較した。エラーの値は 10^{-6} の桁で変化がなくなったところで値が収束したとみなし、そのときの値をとった。

[実験 2] 計算の繰り返し回数とエラーの比較

次に横軸に計算の繰り返し回数を、縦軸にエラーの値をとったグラフを図 5 に示し、計算の繰り返し回数とエラーの関係を比較した。このときドット密度は 30% (ドット数 750) に固定した。

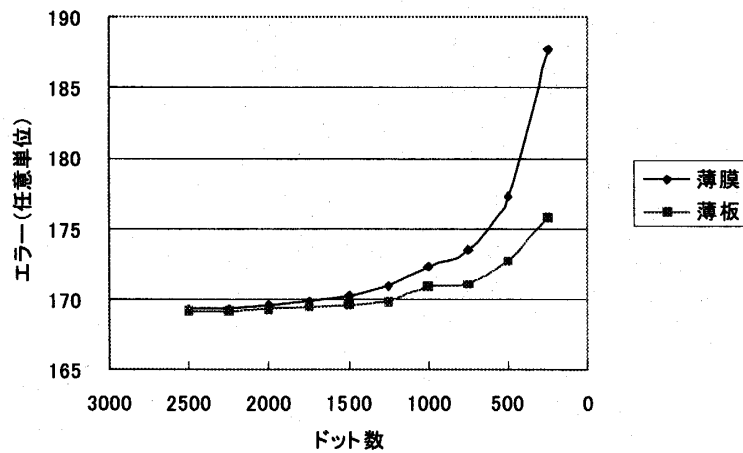


図4 ドット数とエラー

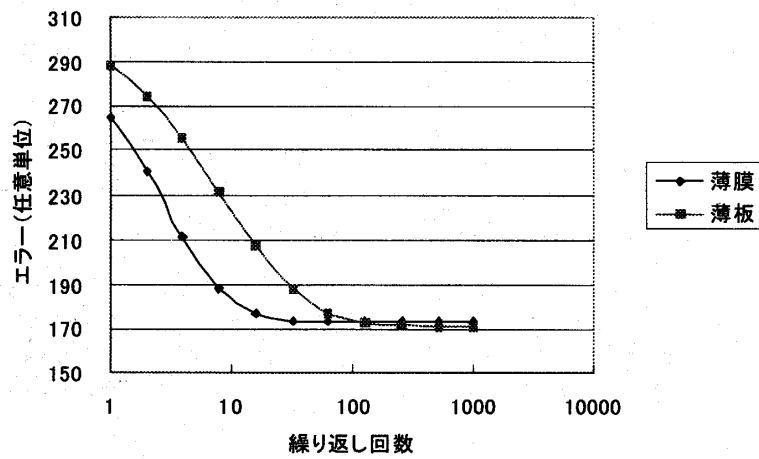


図5 繰り返し回数とエラー

[実験3] sin 曲面の空間周波数とドット密度の比較

横軸に空間周波数を、縦軸にエラーの値をとったグラフを図6から図9に示し、空間周波数

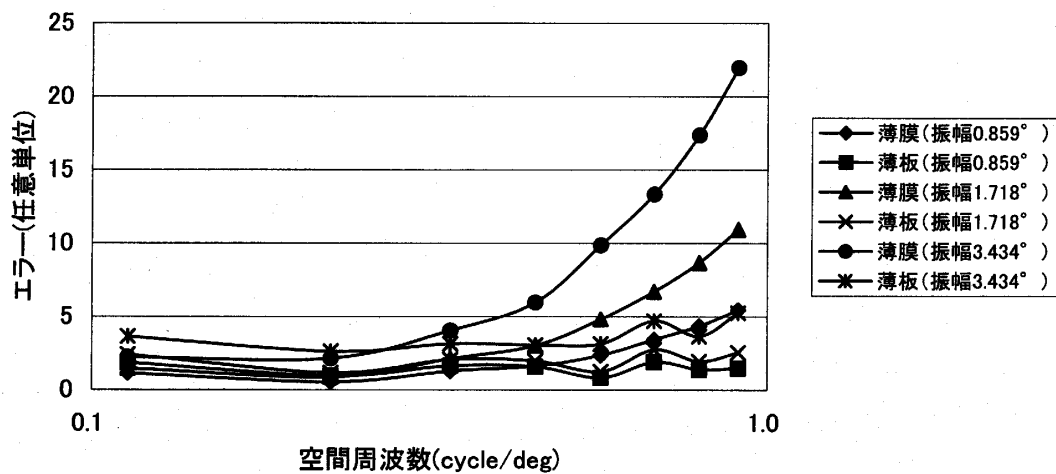


図6 空間周波数とエラー (ドット数2500)

とエラーの関係をもット密度と正弦波の振幅ごとに比較した。

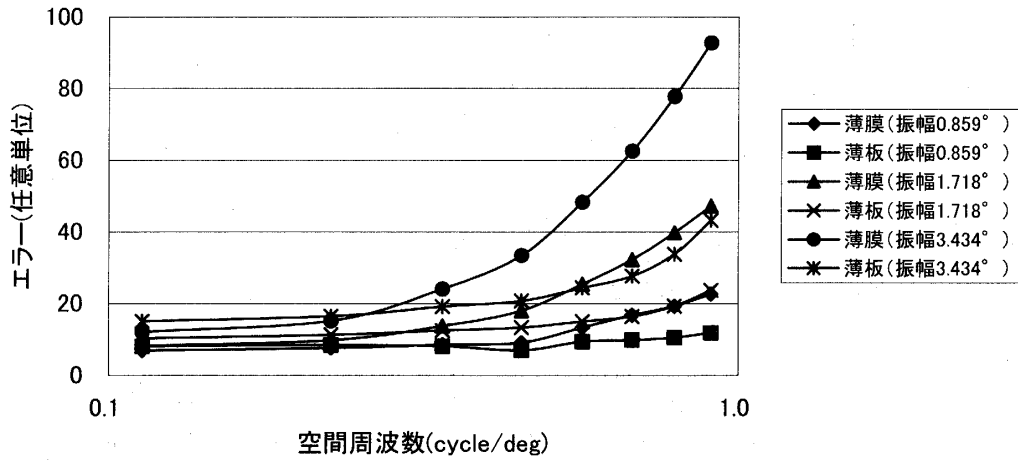


図7 空間周波数とエラー (ドット数1250)

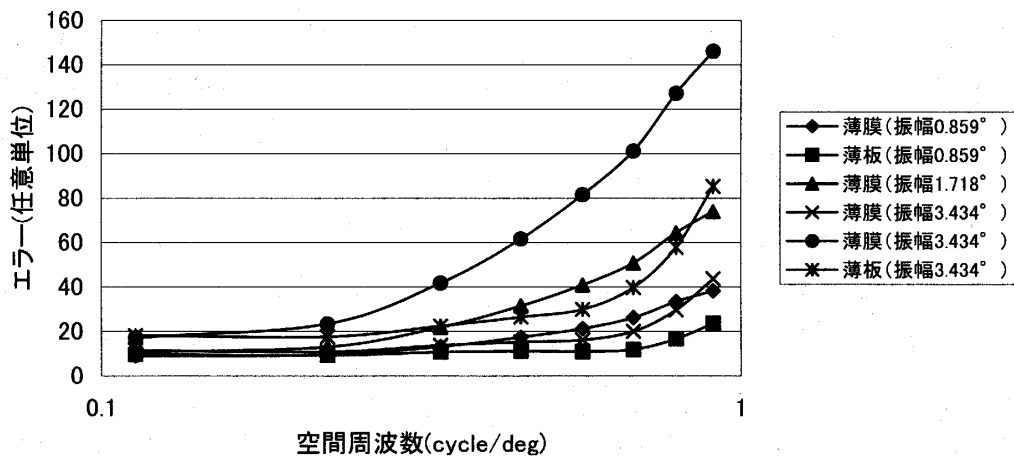


図8 空間周波数とエラー (ドット数750)

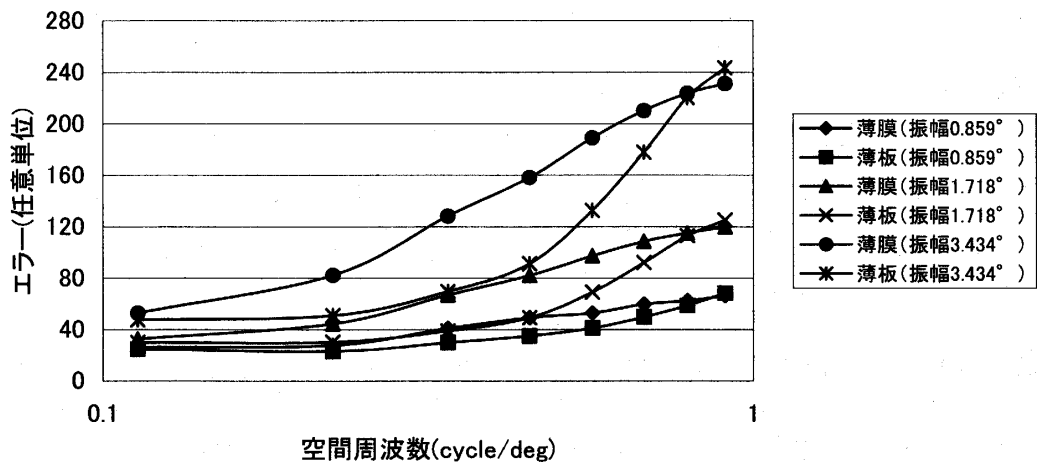


図9 空間周波数とエラー (ドット数250)

3.2 考察

図4からわかるように、ドット密度が約30%（ドット数750）以上のときは薄板ポテンシャルエネルギーも薄膜ポテンシャルエネルギーもエラーの値は、ほぼ変わらない。しかし、それ以下の密度になると薄膜ポテンシャルエネルギーで面を再構成した領域は、薄板ポテンシャルエネルギーで再構成した領域に比べて滑らかに再構成されず、エラーの値も大きい。これは、補間能力が薄板ポテンシャルエネルギーの方が高いことを示している。すなわち、薄板ポテンシャルエネルギーでは2次微分まで（12近傍まで）計算して再構成しているのに対して、薄膜ポテンシャルエネルギーでは1次微分まで（4近傍まで）しか計算して再構成をしていないからであると考えられる。

しかし、図5からわかるように、薄膜ポテンシャルエネルギーの方が薄板ポテンシャルエネルギーよりも圧倒的に計算の収束するスピードが速く、さらに、前述の通り、ドット密度が全体の約30%（ドット数750）以上あれば、収束後のエラーの値は薄板ポテンシャルエネルギーとほとんど変わらなかった。

図6から図9に、空間周波数とエラーの比較をドット密度と振幅ごとに行った実験の結果を示す。この比較を行った理由は、上記2つの比較では、空間周波数は約0.1cycle/deg、振幅は視角にして1.718°に固定されているためである。

図6から図9よりわかるように、ここでも薄膜ポテンシャルエネルギーの方が補間の限界が訪れるのが早い。図4と図5では、ドット密度が全体の約30%（ドット数750）以上あれば、収束後のエラーの値は薄板ポテンシャルエネルギーとほとんど変わらなかったが、空間周波数と振幅を変えることにより、この両方の値が大きくなると、薄膜ポテンシャルエネルギーと薄板ポテンシャルエネルギーの収束後のエラー値の差が大きくなることが示された。また、ドット密度10%（ドット数250）で、空間周波数約0.8cycle/deg以上では、薄膜ポテンシャルエネルギーと薄板ポテンシャルエネルギーのエラーの値は、ほぼ等しくなっている。これは、ドット密度が小さくなることで、一周期内のドット数が少なくなり、より直線的に補間してしまうことによりエラーの値が大きくなるためと考えられる。また、空間周波数が高くなる時、および振幅が大きくなる時も同様のことが考えられる。

実験1、および実験2の結果、人間の視覚系で瞬時に面が再構成されていることから、視覚系が行っている計算理論モデルは薄膜ポテンシャルエネルギーに近いのではないかと考えられる。また、実験3の結果は、人間の視覚系でもドット密度が低い場合は面が認識できないことに相当すると考えられる。

4. 面再構成に関する心理物理学的検討

ここでは、Juleszのランダムドットステレオグラム^[17]を用いて両眼視差からの面の知覚特性を調べ、薄膜と薄板のポテンシャルエネルギーの妥当性を検証した。先行研究では、人間は低密度（2～3%）のランダムドットステレオグラムでさえも、融合可能であるとされてい

る^[18]。また、MTF（空間周波数伝達特性：Modulation Transfer Function）は帯域通過型の特性を示し、空間周波数が0.3~0.5cycle/degの時、感度がピークになり、低周波数領域と高周波数領域の双方で感度が低下することが明らかにされている^[19]。

本実験は、薄膜と薄板のポテンシャルエネルギーの各々の特性のどちらが心理物理学的データによく合うかについて検討を行った。ここでは、正弦波状の面の頂点部分にドットの無い領域を作り、その領域の振幅をどの程度知覚できるかを調べるためにマッチング実験を行った。

人間の視覚系の再構成過程が薄膜ポテンシャルエネルギーに近いのであれば振幅を浅く見積もり、薄板ポテンシャルエネルギーに近いのであれば振幅を深く見積もるであろうと予想される。

4.1 実験方法

[装置]

解像度1024×768の17インチCRTを用いて刺激を提示した。被験者から画面までの距離は約100cmで実験を行った。画面の大きさは32cm×24cm、ドットの表示領域は、12cm×12cmであり、視角6.8°×6.8°に相当する。なお、この実験は暗室にて観測を行った。

[刺激]

図10に刺激として用いたランダムドットパターンを示す。

両眼視差により正弦波状の面が知覚されるランダムドットパターンを用い、ディスプレイ上に、1種類の空間周波数を持つパターンを提示した。両眼視差の刺激は、奥行き方向が x 方向に一定で、 y 方向に正弦波状に変化する（ $x = \sin y$ ）。なお、正弦波の周期は、図11に示されるように、正弦波の頂上部分が丁度中央にくるように配置した、半周期である。

本実験では、ドット数を変えた2種類の刺激、振幅を変えた3種類の刺激、およびドットの無い領域を変えた4種類の刺激を組み合わせ提示した。以下にその値を示す。

空間周波数：0.113 (cycle/deg)

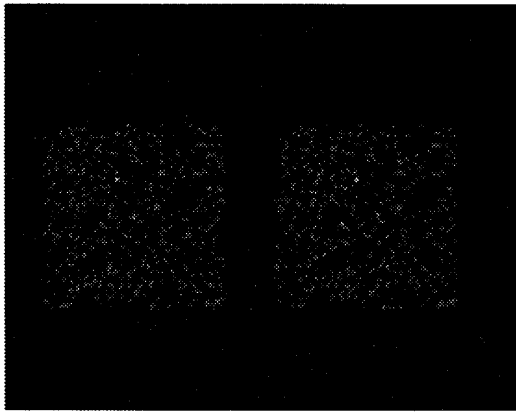
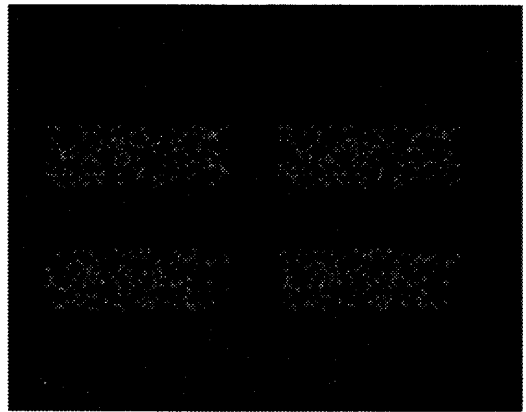
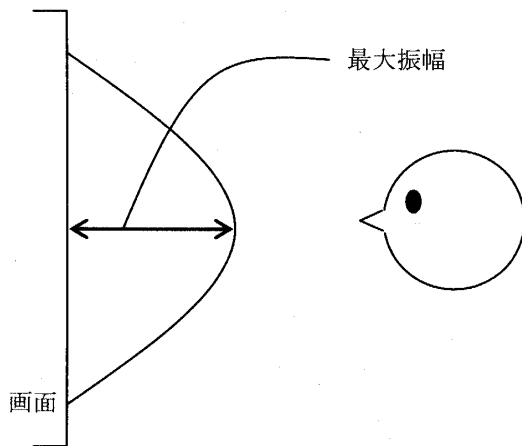
ドット数 : 250, 750

振幅 : 0.20, 0.40, 0.80 (deg)

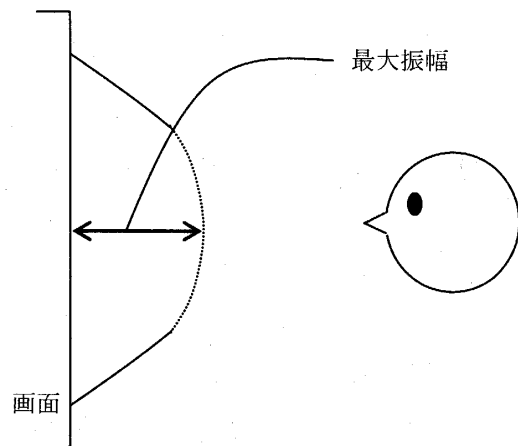
ドットの無い領域：0, 0.82, 2.18, 3.37 (deg)

正弦波の振幅はキーボードの矢印キーで調節することが出来る。

図11(a)は、正弦波状の面の頂点部分にもドットがあるパターンの例であり、図11(b)は、正弦波状の面の頂点部分にドットが無いパターンの例である。

(a) ドットの無い領域 $6.8^\circ \times 0^\circ$ (b) ドットの無い領域 $6.8^\circ \times 2.58^\circ$ 図10 実験に用いたランダムドットパターン (ドット数750, 振幅 0.40° の例)

(a) 面の頂上部分にドットがある場合



(b) 面の頂上部分にドットが無い場合

図11 知覚されると考えられる面のイメージ

[実験の手続き]

あらかじめ奥行き情報およびドットの無い領域を持たせた基準の刺激と、初めは両眼視差を持たない刺激が、キーボード操作により交互に提示される。視差を持たない刺激の両眼視差を0から順に増加させてゆき、基準面で知覚される奥行きと両眼視差による奥行きが一致したと感じた時の振幅を記録した。なお、振幅の値は、 50×50 の座標を使用しているため、 25×25 の位置の振幅の値を最大振幅として記録した。

被験者は、ディスプレイから約100cm離れた場所に座り、条件のコントロールはキーボードの左右の矢印キーで行った。なお、被験者は立体視鏡等は用いず、裸眼で交差法にて観測を行った。

試行回数は、ドット数の1つのパターンにつき、振幅が 0.20° の場合、 0.40° の場合、 0.80° の場合の3セットを各10セット繰り返し、合計60セット行った。

4.2 結果および考察

実験結果を図12, 図13に示す。図中の薄膜と薄板の値は、数値シミュレーションの結果であり、提示した最大振幅の値は、座標における 25×25 の位置の振幅の値である。また、評定値は、基準面で知覚される奥行きと両眼視差による奥行きが一致したと感じた時の振幅の値である。

図からもわかるように、薄膜のポテンシャルエネルギーを用いた数値計算では、正弦波の頂点部分にドットの無い領域が狭い段階から面の再構成の精度が低いことがわかる。また、薄板のポテンシャルエネルギーでの数値計算では、ドットの無い領域が視角にして約 1.5° から 3° のとき、元の正弦波より振幅を大きく再構成した。これに対し、被験者の内観報告により、ドットの無い領域を連続的な面として知覚していることを確認したが、図12および図13より、人間の視覚系の面再構成過程の精度は薄膜ポテンシャルエネルギーより劣り、ドットの無い領域の振幅をより浅く見積もっていることがわかった。また、主観的ではあるが、被験者は、ドットの無い領域に知覚された面は、2つの面を滑らかに正弦波の延長として結合させ補間を行っているのではなく、直線ではないが、直線に近い凸状の滑らかな面により2つの面を結合させているとした。

連続的な面の知覚に関しては、先行研究で、Grimson^[20]は、このような正弦波状の面の頂点部分にドットの無い領域を持つランダムドットステレオグラムの補間の候補は3つあるとし

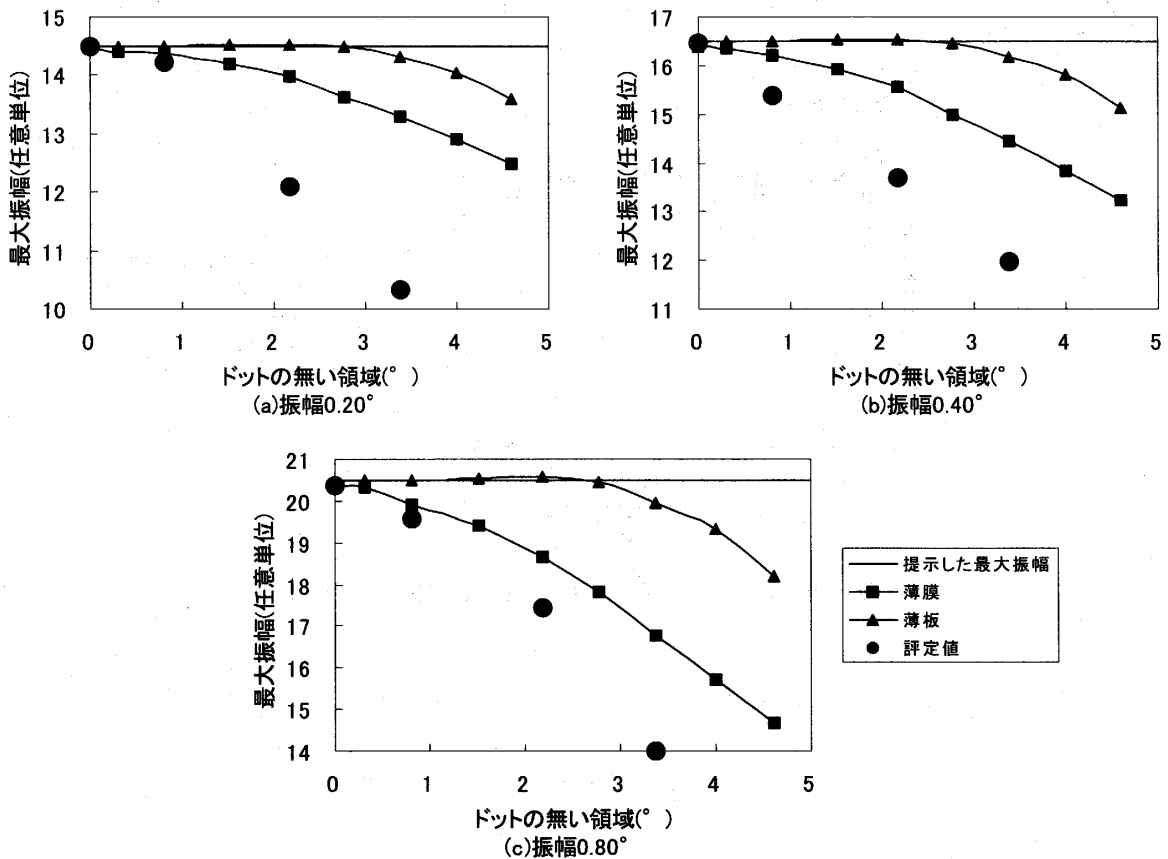


図12 数値計算と心理実験との比較 (ドット数250)

ている。すなわち、(1)一定視差で直線的に補間する、(2)表面方向が不連続にならないように、2つの面を滑らかに結合させる補間を行う、(3)2つの面が交差するまでそれらを直線的に延長する、の3つである。しかし、鮮明な見えは得られず、被験者は(1)と(3)の可能性を排除し、(2)に関してもはっきりしないが、表面方向の不連続は存在しないとしている。(1)および(3)の例を、図14に示す。

一方、ドット密度が大きくなると、ドットの無い領域に連続的な面を知覚することが困難となった。これは、ドットが描画されている領域とドットが無い領域との境目を不連続部分として認識しているからと考えられる。

また、図12、図13から、人間の視覚系の面再構成過程の精度は薄膜ポテンシャルエネルギーより劣り、ドットの無い領域の振幅をより浅く見積もっているのではないかと考えられる結果

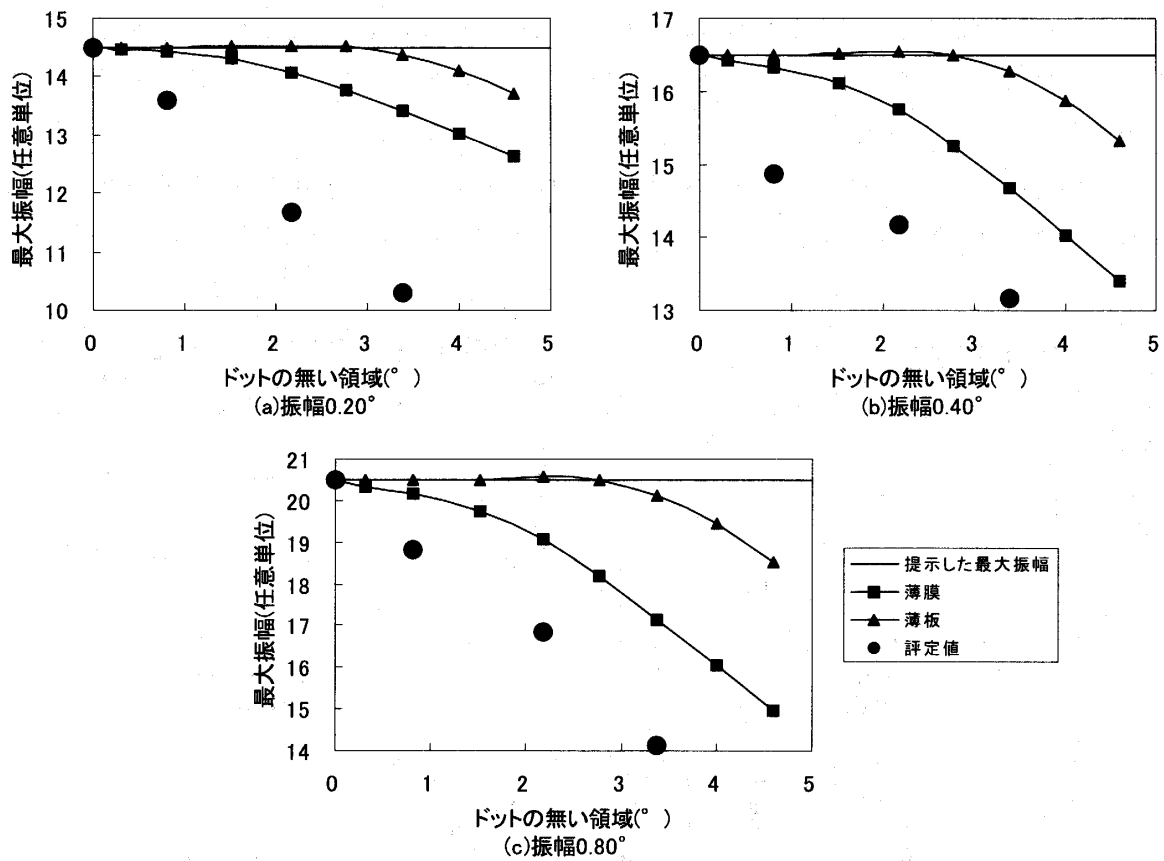


図13 数値計算と心理実験との比較 (ドット数750)

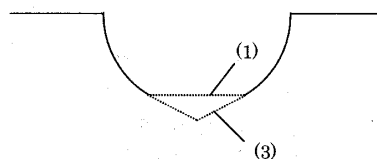


図14 補間の例 [20]

となったが、これは、基準面の振幅をマッチングする方法にも左右されると考えられる。本実験でのマッチング方法は、描画されたドットの集合は違うが、基準となる刺激と同じ正弦波の式を使い、マッチング面を描画した。それ故に、マッチングの際に、図15に示したように知覚される面とマッチング面の形状を合わせてしまい、正しい最大振幅の評定が行われない場合があり、比較が困難となったのではないかと考えられる。この問題の改善点として、図16に示すように、マッチング面に正弦波状の面を使わずに、例えば、中心部分のみが浮き上がって見えるようなランダムドットパターンを用い、この中心部分の飛び出し量と基準面の最大振幅とを比較することなどが考えられる。

以上を総合した結果より、人間の視覚系が行っている計算は、薄膜ポテンシャルエネルギーに近いのではないかと考えられる。

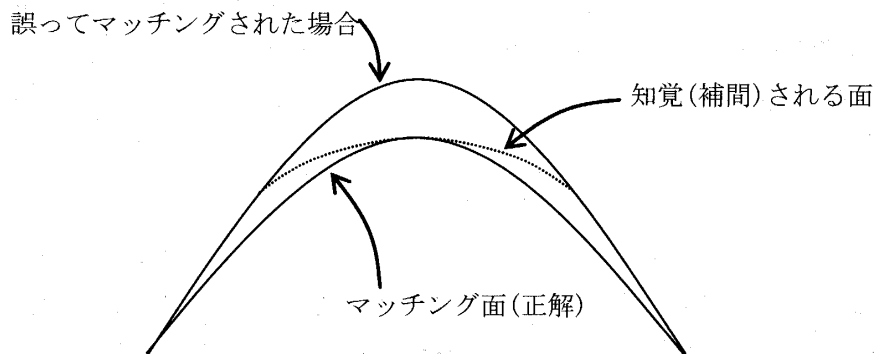


図15 マッチングの例

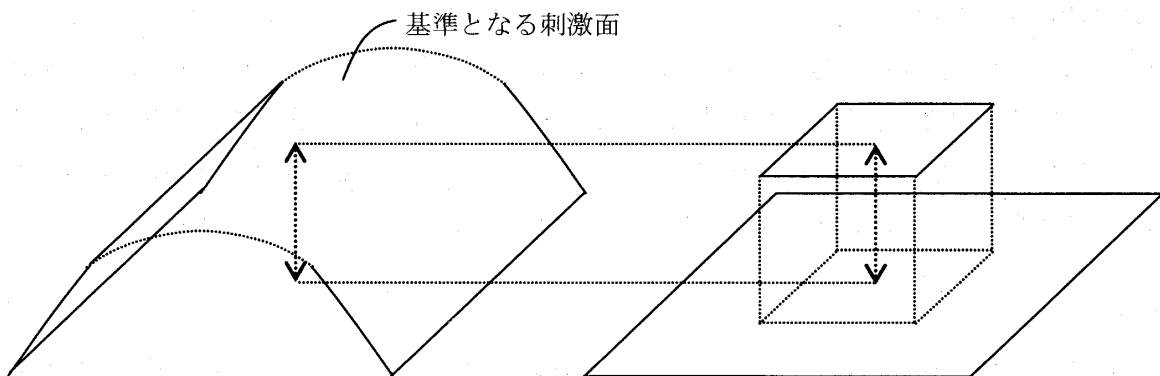


図16 マッチング方法の改善の例

4.3 不連続に関する考察

本実験において、ドット密度が大きくなると、ドットの無い領域に連続的な面を知覚することが困難となった。この時、ドットの無い領域を一定視差で直線的に補間していたのではなく、その領域に全く面を知覚できなかったのである。これは、ドットの在る領域と無い領域との境目を不連続部分として認識している、すなわち、この境界部分にエッジを検出しているという

ことになる。

この不連続を検出する拘束条件を線過程 (line process) といい、エッジの存在を表す確率変数を表す。Geman & Geman^[21] は、線過程を、濃淡レベルの確率変数とは別に準備することにより、不連続の問題を解決している。これは、マルコフ確率場 (MRF) モデルに基づいて、画像中のある画素の状態が、その近傍の画素の状態に依存して決まるという性質に基づいて、最大の確率で起こりうる原画像を後ろ向きに推定している。

つまり、近傍の画素の状態に依存するので、ディスプレイ上に白く表示されたドットとその背景の黒のコントラストが高かったために不連続を検出したのではないかと考えられる。また、ドット数が多くなったため、ドット密度も高くなり、よりエッジを際立たせることになったのではないかと考えられる。

以上より、実験室の環境、実験装置の影響を念頭に置き、正弦波を用いた面再構成実験を行う場合、ドットの輝度値やドットと背景面のコントラストの差異を考え、新たな実験方法を考案していく必要がある。

5. あとがき

本研究では、薄膜と薄板のポテンシャルエネルギーに基づく面の再構成計算を数値シミュレーションにより比較し、両者の生体モデルとしての妥当性を検討した。特に、両眼視差による面の知覚に注目し、心理実験によってこの理論を検証した。その結果、視覚系が行っている計算は、薄膜ポテンシャルエネルギーに近いことを示した。

標準正則化理論は、初期視覚の様々な処理過程を統一的に扱うことが可能であり、その神経生理学的な妥当性やニューラルネットワークのモデルに関しても脳科学の分野で研究が進められている。今後は、単眼の運動視差や陰影などの両眼視差以外の手がかりによる面の再構成に関しても標準正則化理論を適用した数値シミュレーションと心理物理実験の比較を行い、エネルギー関数の妥当性について検証を重ねていきたい。また、不連続を含む面の再構成の計算を可能とするために、線過程を考慮に入れた MRF モデルへの拡張も検討する予定である。

謝辞

本研究の一部は科学研究費 (基盤研究(B), 課題番号16300085) によって行われたことを付記する。

参考文献

- [1] Marr, D. : Vision, Freeman (1982), 乾敏郎, 安藤広志訳: ビジョン, 産業図書 (1987).
- [2] Tikhonov, A. N. and Arsenin, V. Y. : Solution of Ill-Posed Problems, Washington, D. C., Winston and Wiley Publishers (1977).
- [3] Poggio, T., Torre, V. and Koch, C. : Computational vision and regularization theory, Nature, 317, pp.314-319 (1985).
- [4] 水谷博之: 標準正則化における滑らかさの尺度の一般化と面復元への応用, 電子情報通信学会論文誌, J77-A, 9, pp.1284-1291 (1994).
- [5] 村木茂, 横矢直和, 山本和彦: 正則化法による等高線画像からの曲面再構成, 電子情報通信学会論文誌, J73-D-II, 11, pp.1854-1862 (1990).
- [6] 小田悟朗, 今西茂: 曲面の大まかな滑らかさを考慮した正則化法による曲面再構成, 信学技報, PRU92-134, pp.1-9 (1993).
- [7] 朝倉暢彦, 乾敏郎: 両眼視差からの表面復元における視点不変性, 信学技報, PRU95-83 (1995).
- [8] Grimson, W. E. L. : On the reconstruction of visible surface, In Ullman, S. and Richards, W. (Eds.), Image Understanding 1984. Norwood, NJ: Ablex Publishing Corporation (1984).
- [9] 朝倉暢彦, 乾敏郎: 正則化理論による両眼立体視系の空間周波数特性の解析, 信学技報, HIP96-6 (1996).
- [10] Bolle, R. M. and Vemuri, B. C. : On three-dimensional surface reconstruction methods, IEEE Transactions on Pattern Analysis and Machine Intelligence, 13, pp.1-13 (1991).
- [11] Stevenson, R. L. and Delp, E. J. : Viewpoint invariant recovery of visual surface from sparse data, IEEE Transactions on Pattern Analysis and Machine Intelligence, 14, pp.897-909 (1992).
- [12] Yuille, A., Geiger, D. and Bülthoff, H. H. : Stereo integration, mean field theory and psychophysics, Network, 2, pp.423-442 (1991).
- [13] Terzopoulos, D. : Regularization of inverse visual problems involving discontinuities, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8, pp.413-423 (1986).
- [14] Timoshenko, S. and Woinowsky-Krieger, S. : Theory of plates and shells, McGRAW-HILL BOOK COMPANY, INC. (1959).
- [15] 坂上勝彦, 横矢直和: 弛緩法と正則化, 情報処理, 30, 9, pp.1047-1057 (1989).
- [16] 坂上勝彦: エネルギー最小化原理に基づく画像処理 - 弛緩法と正則化 -, 電子技術総合研究所, IE89-89, pp.1-8, (1989).

- [17] Julesz, B. : Binocular depth perception of computer-generated patterns, Bell System Technical Journal, 39, 1125-1162 (1963).
- [18] Rogers, B. and Graham, M. : Similarities between motion parallax and stereopsis in human depth perception, Vision Research, 22, 261-270 (1982).
- [19] White, B. W. : Stimulus-conditions affecting a recently discovered stereoscopic effect, Am. J. Psychol., 411-420 (1962).
- [20] Grimson, W. E. L. : Differential geometry, surface patches and convergence methods, MIT A. I. Lab. Memo 510 (1979).
- [21] Geman, S. and Geman, D. : Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 6, pp.721-741 (1984).
- [22] 石黒真木夫, 松本隆, 乾敏郎, 田邊國士: 階層ベイズモデルとその周辺 時系列・画像・認知への応用, 岩波書店 (2004).
- [23] 川人光男: 脳の計算理論, 産業図書 (1996).
- [24] 白井支朗 (編), 日本生物物理学会シリーズニューバイオフィジックス刊行委員会 (編): 脳・神経システムの数理モデル, 共立出版社 (1997).
- [25] 乾敏郎: Q&A でわかる脳と視覚 ー人間からロボットまでー, サイエンス社 (1993).
- [26] 平井有三: 視覚と記憶の情報処理, 培風館 (1995).
- [27] 岡崎誠: べんりな変分原理, 共立出版 (1993).