

視覚背側経路細胞の反応特性およびヒトの速度知覚特性を
統一的に説明する数理モデルの構築

中村 大樹

電気通信大学 大学院 情報システム学研究科
博士（工学）学位申請論文

2018年3月

視覚背側経路細胞の反応特性およびヒトの速度知覚特性を
統一的に説明する数理モデルの構築

主査	佐藤	俊治	准教授
委員	阪口	豊	教授
委員	末廣	尚士	教授
委員	檜森	与志喜	教授
委員	南	泰浩	教授

著作権所有者

中村 大樹

2018 年

Abstract

Middle Temporal area (MT) neurons which are related to motion perception have been regarded as velocity filters because the response curve of MT neurons illustrates a unimodal function with respect to stimulus speed. This paper proposes a simple computational model of MT neurons that is not based on velocity filters. The proposed model reproduces not only MT responses of speed selectivity but also illusory motion perception. Moreover, to evaluate the plausibility of the MT model, I generated an enormous set of possible visual patterns as inputs to the MT model: $8^8 = 16,777,216$. Numerical quantities of model outputs by computer simulation for 8^8 inputs were used to estimate human illusory perception. Psychophysical experiments show that the model prediction is consistent with human perception.

概要

細胞の入出力特性を記述する既存の数値モデルはどれも、電気生理学実験で得られた「選択性」という考えを基に、特定の特徴量を抽出する「フィルタ」として構築されている。もしこの「特徴量抽出フィルタ」という考えに依らない新しい数値モデルが構築できれば、細胞や視覚機能に対し新しい解釈が得られ、視覚研究を推し進めることができる。

本論文では、視覚機能の1つである動き知覚に焦点を当て、関連する脳細胞の反応特性およびヒトの速度知覚特性を統一的に再現・説明する数値モデルを構築する。対象とする脳細胞は、視覚背側経路に属する Middle Temporal Area (MT 野) 細胞とした。

「特徴量抽出フィルタ」に代わる概念として、「個々の MT 細胞は局所的な速度推定を行っている」という前提のもと数値モデルを構築した。提案モデルは「選択性」に依らない画像処理工学的速度推定手法を計算基盤としているが、MT 細胞同様、提示刺激の速度に対して単峰性のグラフを描いた。この結果から、MT 細胞の計算論的役割は「速度抽出フィルタ」ではなく、「速度推定器」であるという新しい解釈が得られた。また、「速度抽出フィルタ」という既存概念では解釈が難しい現象も提案モデルで再現できた。提案モデルを解析したところ、モデル導出時に設定した仮定の不成立やモデル式においてゼロ除算を避けるために加えた処理の副次的効果に起因した、「演繹的に生じる現象」であることが分かった。

本論文では、物理的な動きと知覚される動きとが乖離している状況である運動錯視にも着目した。もし数値モデルが視覚系細胞や視覚機能を精度良く記述できているのであれば、ヒト同様、数値モデルも錯視を誘発するはずである。また、数値モデルに

において、現象の再現・説明だけでなく、未知の入力に対する予測も大事な要件の1つである。

モデル予測として、ヒトに対して行うには現実的でない膨大な数の入力画像に対し、錯視を生じさせるパターンか否かの選別を行った。具体的には、円環状刺激において取り得る全てのパターン 16,777,216 種に対し、モデルシミュレーションによる選別を行った。得られたモデル予測を評価するために、無作為に選んだ一部の画像に対し心理物理実験を行い、実際にヒトが錯視を誘発するか調べた。モデル予測とヒトの知覚との相関係数を求めたところ、0.96 のとても強い正の相関があった。

キーワード MT 細胞, 運動錯視, Lucas-Kanade 法, 数理モデル

目次

第 1 章	序論	1
1.1	視覚システムを解明するためのアプローチ	1
1.2	視覚領野と選択性発見の歴史	3
1.3	Middle Temporal Area (MT 野) 細胞	6
1.4	既存の MT 細胞モデル: SH モデル	9
1.5	「特徴抽出フィルタ」では解釈の難しい現象	11
1.6	本研究の目的およびアプローチ	13
1.7	本論文の構成	14
第 2 章	計算原理および提案モデル	15
2.1	Lucas-Kanade(LK) 法	16
2.2	ゼロ除算の回避	19
2.3	提案する MT 細胞モデル	21
2.4	微分計算カーネル	24
2.5	むすび	26
第 3 章	MT 細胞の反応特性の再現	29

3.1	様々な刺激速度に対する速度推定結果	30
3.2	微分カーネルサイズと上限速度との関係	32
3.3	方向選択性の再現	35
3.4	Band 幅のコントラスト依存性の再現	38
3.5	Preferred speed のコントラスト依存性の再現	41
3.6	MT 細胞集団からの read-out	43
3.7	むすび	45
第 4 章	速度知覚特性の説明とモデル予測の検証	48
4.1	Drift illusion の背景輝度依存性の再現	49
4.2	Stepping Feet illusion の再現	56
4.3	Drift illusion を生じさせるパターンの網羅的探索	59
4.4	むすび	70
第 5 章	Grating 刺激に対する解析解	73
5.1	MT 細胞の空間周波数依存性	73
5.2	Grating 刺激に対する LK 法による推定速度の解析解	75
5.3	生理データに対するフィッティング	77
5.4	考察	81
第 6 章	結言	83
参考文献		85
謝辞		88

図目次

1.1	脳の領野間の関係図	4
1.2	初期視覚野 (V1 野) の方位選択性	5
1.3	二次視覚野 (V2 野) の曲率選択性	6
1.4	ある MT 細胞の刺激速度に対する反応特性の例	7
1.5	4 個の MT 細胞の反応例	8
1.6	刺激の動く方向に対するある MT 細胞の反応特性の例	8
1.7	Grating 刺激の例	10
1.8	MT 細胞が最大反応を示す速度のコントラスト依存性の例	12
1.9	半値幅のコントラスト依存性	13
2.1	$1/(S_{yy} + \varepsilon^2)$ と $1/S_{yy}$ の違い	20
2.2	(x, y) 座標系と (ξ, η) 座標系.	22
2.3	式 (23) による ξ 方向の推定速度 \hat{v}_ξ を, 対応する視覚領野で表した図	24
2.4	カーネルサイズ $k \in \{5, 9, 17, 33\}$ での x 偏微分を計算する微分カーネル	26
3.1	入力画像の例	30
3.2	刺激速度に対するモデル出力の平均値 (平均推定速度)	31

3.3	様々なカーネルサイズに対するモデル出力	33
3.4	図 3.3 の各推定結果 $\hat{v}_x(k_i)$ を最大値が 1 になるよう正規化してプロットした図	34
3.5	様々な方向に刺激を動かした時の x 軸方向 (0° 方向) の推定速度 \hat{v}_x の平均値	36
3.6	局所座標系の回転角 ϕ を $\phi = 219^\circ$ とした時の式 (23) による ξ 方向 (動径方向) の推定速度 \hat{v}_ξ の平均値	38
3.7	Grating 刺激の例	39
3.8	コントラストを変化させたときの, $\phi = 0^\circ$ でのモデル出力	40
3.9	コントラストを変化させた時の, 刺激速度に対するモデル出力 (推定速度)	42
4.1	Drift illusion の例	49
4.2	Read-out モデルから得られた動きベクトル	52
4.3	背景輝度に対する平均回転量 \bar{R}	53
4.4	極座標系 (r, θ) および局所座標系 (ξ, η) での FW 刺激の表現	54
4.5	Drift illusion の背景輝度依存性の要因	55
4.6	Stepping feet illusion	57
4.7	Stepping feet illusion に対するモデル出力	58
4.8	円環刺激の構成例	60
4.9	全 16,777,216 パターンに対する平均回転量 \bar{R} のヒストグラム	61
4.10	心理物理実験に用いた刺激一覧	62
4.11	心理物理実験で被験者回答を得るために用いた回転デバイス	64

4.12	パラメータ s による変換式の振る舞いの変化	65
4.13	モデル式 (43) において微分カーネル数 $N = 1$, 微分カーネルサイズ $k = 5$ とした場合の, モデル予測とヒトの知覚との散布図	66
4.14	微分カーネルサイズに対する, モデル予測とヒトの知覚との相関係数 r	67
4.15	最も相関係数 r が高くなる条件での, モデル予測に対するヒトの知覚 の散布図	68
4.16	$f = 1/2$, $N = 1$ かつ $k = 17$ での全 16,777,216 パターンに対する平 均回転量 \bar{R} の再シミュレーション結果	68
4.17	時計回りの回転錯視を誘発するパターンで $ \bar{R} $ が大きい 10 種	69
4.18	Rotating snakes	72
5.1	Grating 刺激に対する 3 種類の MT 細胞の反応例	74
5.2	Simoncelli & Heeger モデルにおける速度 2deg/s を検出するための理 想的な時空間周波数フィルタ	76
5.3	式 (60) を用いて生理データをフィッティングした結果	78
5.4	式 (61) を用いて再度生理データをフィッティングした結果	79

表目次

3.1	Read-out モデルの比較	45
3.2	既存の MT 細胞モデル (SH モデル) と提案モデルの比較	46
5.1	式 (60) でモデルフィッティングを行った際の適合度および最適パラ メータ	80
5.2	式 (61) で再度モデルフィッティングを行った際の適合度および最適 パラメータ	80

第1章

序論

本章では，本論文が視覚脳科学研究においてどのような位置づけにあるのか，対象とする脳細胞や視覚特性，構築する数理モデルの粒度などについて記す．また，既存数理モデルの基盤となる考えである「選択性」について紹介する．これまでの数理モデルは「選択性」を基に，「特徴量を抽出するフィルタ」を構築している．本章では，「選択性」から安直に「特徴抽出フィルタ」を構築するというアプローチに対し疑問を呈する．

1.1 視覚システムを解明するためのアプローチ

ヒトの視覚システムを解明するために，神経生理学的アプローチ，認知心理学的アプローチ，計算論的アプローチと，様々なアプローチで研究が行われてきた．

神経生理学的アプローチは，個々の脳細胞または細胞集団を対象とするアプローチである．脳細胞がどのような化学的構造をしているか，どのようにして情報を伝播させているかといったミクロな視点からのアプローチである．ヒトの脳新皮質は数十

億個もの脳細胞で構築されている [1]. 1 つ 1 つの脳細胞について理解を深めることで全体が理解できるという, ボトムアップによるアプローチである. 脳細胞は, 細胞内外のイオン濃度差から生じる活動電位や化学物質の分泌によって, 細胞内, 細胞間の情報伝達を行っている [2]. 視覚刺激 (動画像) や聴覚刺激 (音) など様々な刺激を提示し, 関連する脳細胞のイオン濃度変化や細胞内の電位変化を測定する電気生理学実験が行われている.

一方, 認知心理学的アプローチは, ヒトの視覚機能を対象としたアプローチである. 視覚機能を 1 つのシステムと見なしその入出力特性を調査する, マクロな視点からのアプローチである. 先に全体を把握し, その後で細胞集団や個々の脳細胞の役割を理解するトップダウンによるアプローチである. ヒトや, サルなどの動物に対し視覚刺激や聴覚刺激など様々な刺激を提示し, 提示刺激の物理的性質と知覚される性質との対応関係を調べる心理物理実験が行われている.

計算論的アプローチは, 脳細胞の反応特性や視覚機能の特性を再現・説明するモデルを構築することで, 個々の細胞がどのような役割を担っているか, 視覚機能がどのような計算原理で実現されているかを調べるアプローチである. 数理モデルは, イオンの流出入など分子レベルでのモデル化を行う細粒度モデルと, 細胞や視覚機能を 1 つのシステムと見なす疎粒度モデルとに大別できる.

本論文では, 計算論的アプローチによる視覚システムの解明を目指す. 具体的には, 脳細胞および視覚機能を 1 つのシステムと見なし, その入出力関係を再現・説明する疎粒度モデルを構築する. ヒトの視覚システムは複雑な細胞ネットワークを構築しており, 様々な機能を持つ. 本論文では速度知覚に焦点をあて, 関連する脳細胞である Middle Temporal Area (MT 野) 細胞の反応特性および速度知覚特性を統一的に再現・

説明する数理モデルを構築する。

1.2 視覚領野と選択性発見の歴史

脳細胞はそれぞれが様々な処理を行っており、類似した反応特性を示す細胞群をまとめて「領野」と呼ぶ [3]。図 1.1 に示すように、大脳新皮質には様々な領野があり、各領野は様々な領野と結合し、情報の伝達を行っている。視界に入っている物体を認識するためには、主に $RGC \rightarrow LGN \rightarrow V1 \rightarrow V2 \rightarrow V4 \rightarrow \dots$ といった経路で処理が行われる。一方視界に入っている物体の速度を知覚するためには、主に $RGC \rightarrow LGN \rightarrow V1 \rightarrow MT \rightarrow MST \rightarrow \dots$ といった経路で処理が行われる。領野ごとに行われる処理は異なり、様々な脳領野を上手くつなぎ合わせることでヒトは複雑な視覚機能を実現している。

これまでの単一細胞を対象とした電気生理学実験はどれも、「選択性」の発見に注力されてきた。図 1.2 に示すように、Hubel と Wiesel は線分の傾きを様々に変化させて提示した時に、初期視覚野（V1 野）細胞が特定の方位に傾いた線分に対して強く反応を示すことを発見した [4]。V1 細胞はそれぞれ最大反応を示す方位が異なるため、V1 細胞は「方位選択性」があると言われている。また、最大反応を示す方位は *preferred orientation* と呼ばれている。同様に、図 1.3 に示すように、Ito と Komatsu は 2 つの線分のなす角（曲率）を様々に変化させた時に、二次視覚野（V2 野）細胞が特定の曲率に対して強く反応することを発見した [5]。V2 細胞はそれぞれ最大反応を示す曲率が異なるため、V2 細胞は「曲率選択性」があると言われている。また、最大反応を示す曲率は *preferred curvature* と呼ばれている。同様に、V4 野で形状選択性 [6]、V6 野で頭部運動方向選択性 [7] など、様々な領野で様々な選択性が発見されてきた。これら

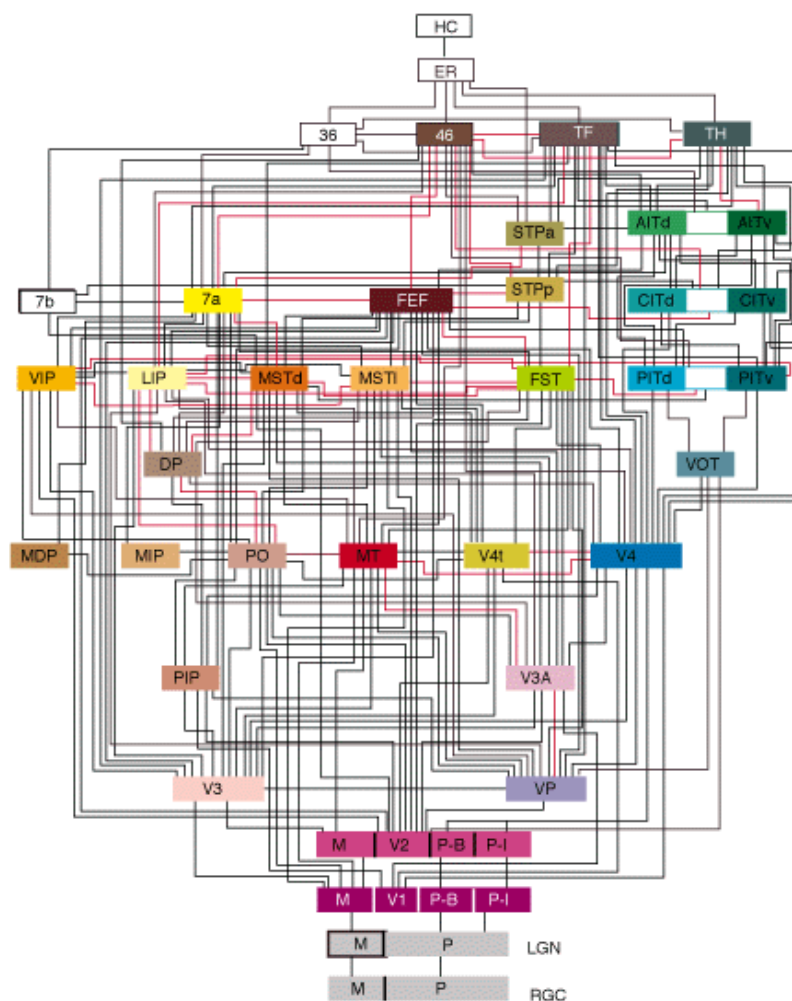


図 1.1 脳の領野間の関係図 [3]. 四角が領野を表し，線は各領野の結合関係を表している．網膜 (Retina) から得られた外界の情報は様々な領野で処理され，より高次の (図の上部に位置する) 領野へと送られる．

は，提示画像の特徴量 X を様々に変化させ，測定対象の脳細胞が $X = X_0$ で最大反応を示す単峰性のグラフを描くことから， X 選択性があると結論づけられてきた．

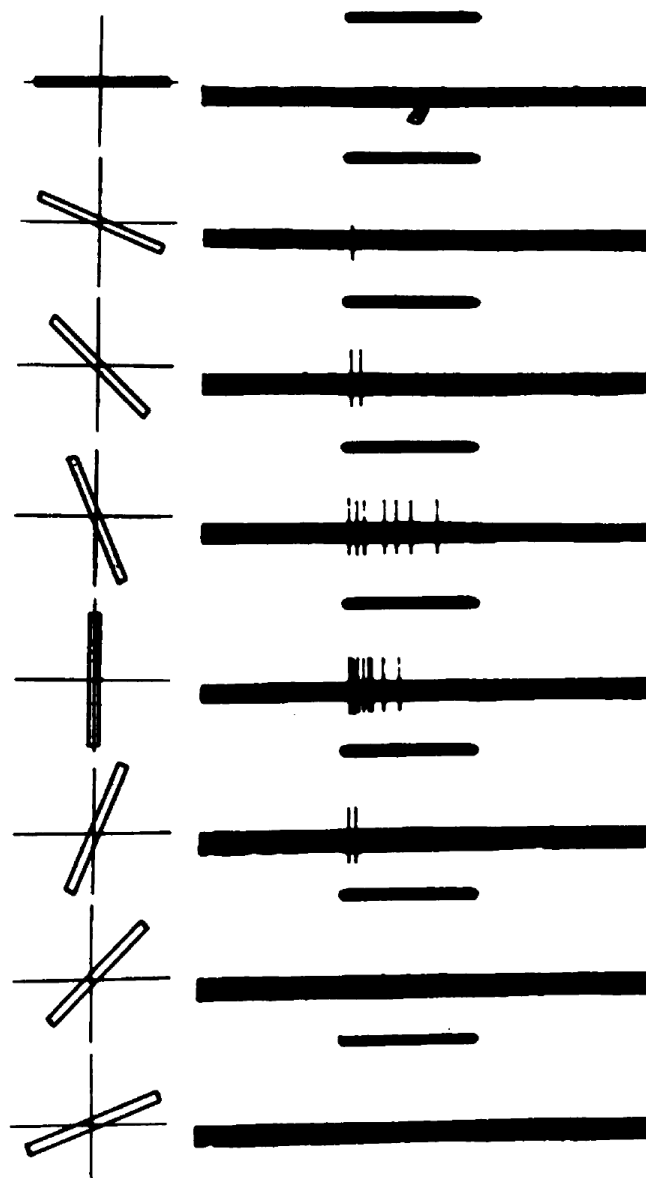


図 1.2 初期視覚野 (V1 野) の方位選択性 [4]. 左列は提示した視覚刺激 (線分) を表している. 右列は, 測定している V1 細胞が発火 (反応) したかどうかを表す図であり, 縦線 1 本が 1 回発火したことを意味している. 縦線の数が多く密度が高いと, 提示刺激に強く反応していることを意味する. また, 視覚刺激は右図の短い黒線が書かれているタイミングで左列の刺激を提示している. この図から, この V1 細胞は 90° に傾いた線分 (縦線) に対して強く反応していることが分かる.

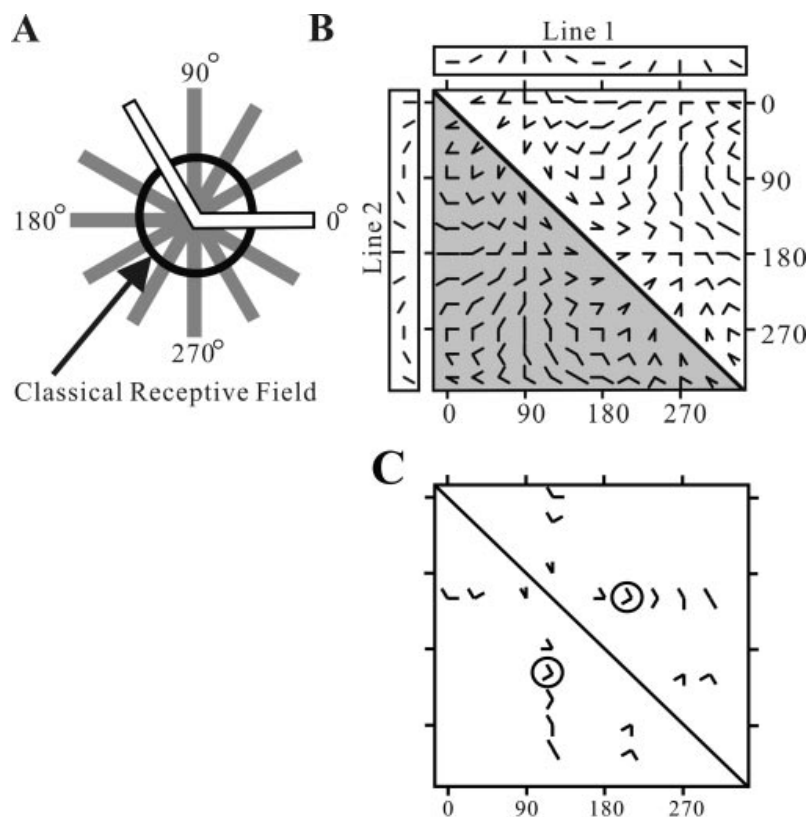


図 1.3 二次視覚野 (V2 野) の曲率選択性 [5]. A:提示刺激. 2本の線分を組み合わせ, 古典的受容野 (classical receptive field; その細胞が担当している視覚領域) の中心に提示する. B:線分の組み合わせによる提示刺激一覧. 対角線で対称な刺激であり, 実験では上半分の刺激を提示している. C:測定対象の V2 細胞が強く反応した提示刺激一覧. 丸で囲まれた刺激が最も強く反応した刺激である.

1.3 Middle Temporal Area (MT 野) 細胞

MT 細胞もまた, 「速度選択性」があると言われている [8]. これは, 図 1.4 に示すように, 横軸を提示刺激の速度, 縦軸をある MT 細胞の発火頻度 (反応度合い) として

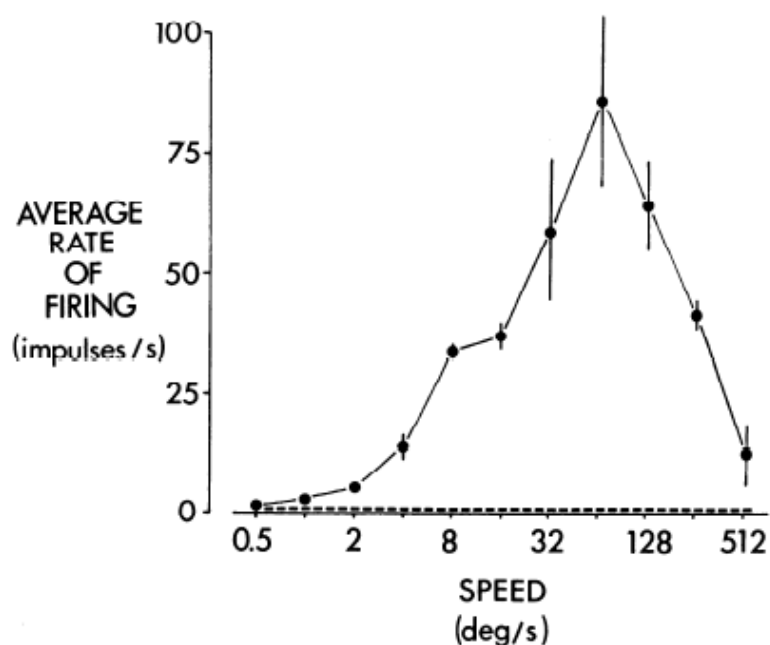


図 1.4 ある MT 細胞の刺激速度に対する反応特性の例 [8]. この MT 細胞は約 64 deg/s の速度で動く刺激に対し最も強く反応している.

プロットした時に、特定の速度（この MT 細胞では約 64 deg/s⁽¹⁾の速度）で最大反応を示すためである。図 1.5 に示すように、MT 細胞ごとに最大反応を示す速度は異なる。また、MT 細胞は刺激の動く方向に対しても選択性がある。図 1.6 に示すように、様々な方向に動く刺激に対する MT 細胞の発火頻度をプロットすると、ある特定の方向に動く刺激に対して強く反応することが分かる [8]. このように、MT 細胞には最大反応を示す速さと最大反応を示す方向がある。これらの速さおよび方向は、preferred speed および preferred direction と呼ばれている。図 1.4 の MT 細胞は preferred speed が 64 deg/s，図 1.6 の MT 細胞は preferred direction が 210° である。

⁽¹⁾ deg/s : 視野角に対する速度。視覚研究では、物体が単位時間辺りに動いた視野角で動きを表す。これは、網膜に写る像を基準にするためである。同じ 1 deg/s で動く物体でも、近くにある物体は時速換算では遅く、遠くにある物体は時速換算では速くなる。

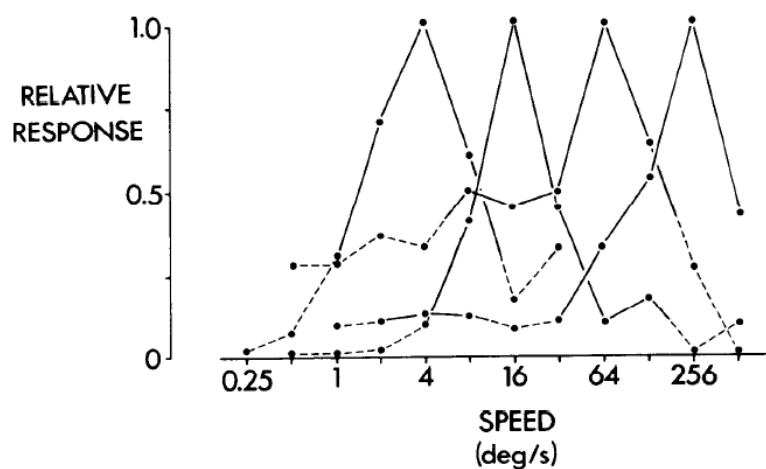


図 1.5 4 個の MT 細胞の反応例 [8]. 4 つの MT 細胞はそれぞれ約 4, 16, 64, 256 deg/s の速度で最も良く反応している。

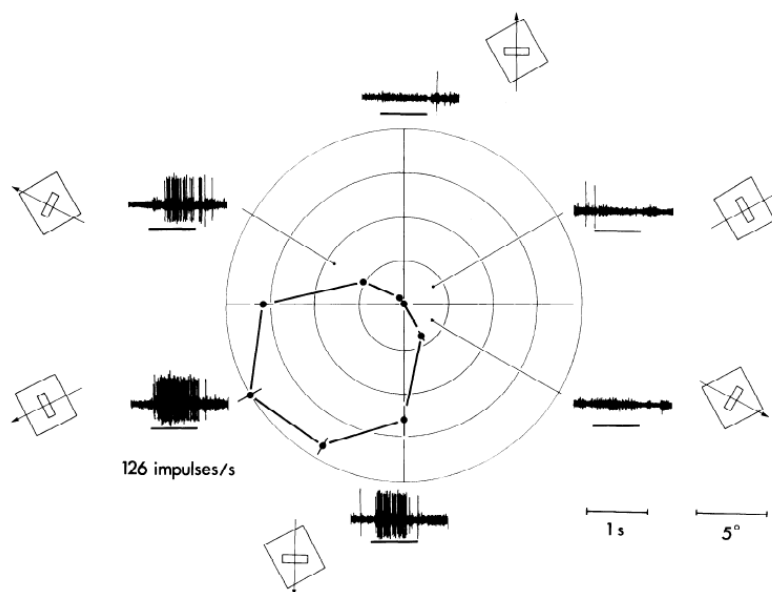


図 1.6 様々な方向に動く刺激に対するある MT 細胞の発火頻度をプロットしたグラフ [8]. 最外周の図は、提示された bar 刺激およびその移動方向を示している。その内側の図は、提示された刺激に対する MT 細胞の発火頻度を表している。中心の図は、刺激の動く方向 ϕ に対して MT 細胞の発火頻度を中心からの距離 r で表したレーダーチャートである。この図から、この MT 細胞は約 210° 方向に動く刺激に対して最も良く反応していることがわかる。

1.4 既存の MT 細胞モデル : SH モデル

Simoncelli & Heeger モデル (SH モデル) は, 前節に記した「速度選択性」という考えを基に構築された MT 細胞モデルである [9]. 工学的観点から見れば, MT 細胞は preferred speed および preferred direction を抽出するバンドパスフィルタであると見なすことができる. MT 細胞は主に V1 細胞から情報を受け取っており, V1 細胞の出力を基に速度を抽出するバンドパスフィルタを構築できれば良い.

V1 細胞には, 時空間周波数に対して選択的に反応する細胞がある [10]. これは, 次式に示す grating 刺激と呼ばれる cos 波状刺激を提示した場合に, 特定の時間周波数 f_t および空間周波数 f_x, f_y に対して最大反応を示すためである.

$$I(x, y, t; f_x, f_y, f_t, \phi) = \cos(2\pi f_x x + 2\pi f_y y + 2\pi f_t t + \phi) \quad (1)$$

但し, 上式において ϕ は位相である. 図 1.7 は, 空間周波数を様々に変化させた場合の grating 刺激の例である. 空間周波数 f_x, f_y は単位空間距離あたりの波数であり, f_x および f_y が小さい場合 (図 1.7 の原点付近) には縞の数が少ないパターンとなる. 一方, f_x および f_y が大きい場合には縞の数が多くなるパターンとなる. f_t は単位時間あたりに動く波数であり, f_t が小さい場合は時間変化の少ないパターン, f_t が大きい場合には時間変化が大きいパターンとなる. フーリエ変換可能な任意の動画像 $I(x, y, t)$ は, cos 波の重ね合わせで表現できる⁽²⁾.

$$I(x, y, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{I}(f_x, f_y, f_t) e^{2\pi i(f_x x + f_y y + f_t t)} df_x df_y df_t \quad (2)$$

⁽²⁾ オイラーの公式より $e^{2\pi i(f_x x + f_y y + f_t t)} = \cos [2\pi(f_x x + f_y y + f_t t)] + i \sin [2\pi(f_x x + f_y y + f_t t)]$ である. また, $\tilde{I}(f_x, f_y, f_t)$ は $I(x, y, t)$ に含まれる時空間周波数成分の強度を意味し, $\tilde{I}(f_x, f_y, f_t) = \iiint I(x, y, t) e^{-2\pi i(f_x x + f_y y + f_t t)} dx dy dt$ で求めることができる.

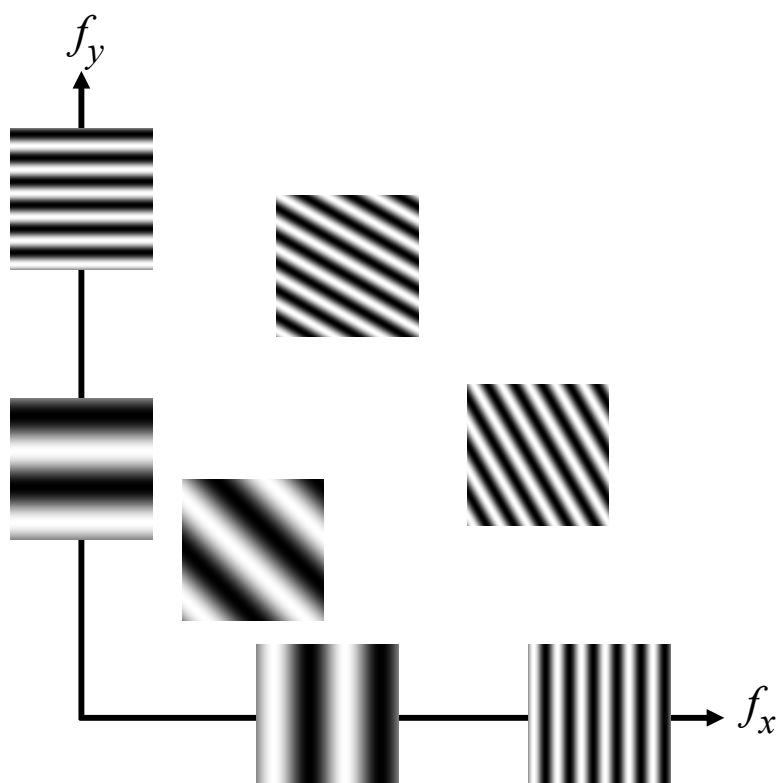


図 1.7 Grating 刺激の例. 空間周波数 f_x, f_y が小さい場合 (原点付近の場合) には縞の数が少ないパターンとなる. 一方, f_x, f_y が大きい場合には縞の数が多くなるパターンとなる. 縞の水平方向からの傾き θ は, $\theta = \tan^{-1} \frac{f_y}{f_x}$ で決まる.

従って, 特定の時空間周波数 $(f_{x_0}, f_{y_0}, f_{t_0})$ に最大反応を示す V1 細胞は $\tilde{I}(f_{x_0}, f_{y_0}, f_{t_0})$ を求めていると見なせ, V1 細胞集団はフーリエ変換による周波数解析を行っていると言える.

一方で, 空間周波数 f_x の \cos 波を速度 v_x, v_y で動かす場合を考えると, grating 刺激は次式で表せる.

$$\begin{aligned} I(x, y, t; f_x, f_y, f_t) &= \cos(2\pi f_x(x - v_x t) + 2\pi f_y(y - v_y t)) \\ &= \cos(2\pi f_x x + 2\pi f_y y + 2\pi(-f_x v_x - f_y v_y)t) \end{aligned} \quad (3)$$

式 (1) と式 (3) を比較すると, 時空間周波数 f_x, f_y, f_t と速度 v_x, v_y との間には $f_t =$

$-f_x v_x - f_y v_y$ すなわち $f_x v_x + f_y v_y + f_t = 0$ の関係がある。従って、速度 $v = (v_{x_0}, v_{y_0})^T$ に選択的に反応する MT 細胞モデルは、最大反応を示す時空間周波数 $(f_{x_0}, f_{y_0}, f_{t_0})$ が $f_{x_0} v_{x_0} + f_{y_0} v_{y_0} + f_{t_0} = 0$ を満たす V1 細胞の出力を集めることで実現できる。

1.5 「特徴抽出フィルタ」では解釈の難しい現象

既存の MT 細胞モデルは、MT 細胞に速度選択性があるため、特定の速度を抽出するバンドパスフィルタと見なして数理モデルが構築されてきた。しかし、速度抽出フィルタという考えでは解釈の難しい電気生理学実験結果がいくつかある。

図 1.8 は、preferred speed が提示刺激のコントラストに依存して変化する例である [11]。この MT 細胞は、高コントラスト時（70%；濃い黒線）には約 20 deg/s の速度で最大反応を示すが、低コントラスト時（5%；薄い灰色線）には約 8 deg/s の速度で最大反応を示す。速度抽出フィルタという考えを基にこの現象を解釈すると、提示刺激のコントラストに依存してフィルタ特性が変化するという解釈になる。

図 1.9 は、刺激方向に関する感度（分解能）が提示刺激のコントラストに依存して変化する例である [12]。高コントラスト時には刺激の動く方向と preferred direction とがある程度異なっても反応するが、低コントラスト時には刺激の動く方向が preferred direction に近くないと反応しなくなる。これは、細胞集団の反応から提示刺激の動く方向を推定することを考えると、信号強度が低下するほど方向に関する分解能が向上することを意味する。しかし、信号強度が低下する（SN 比が低下する）と一般的に分解能は低下するはずである。この現象は、単純なゲインコントロールや iceberg 効果 [13] では説明ができない。速度抽出フィルタの入出力特性を考えると、コントラストを変化させても各方向に対する出力が定数倍されるだけであり、半値幅

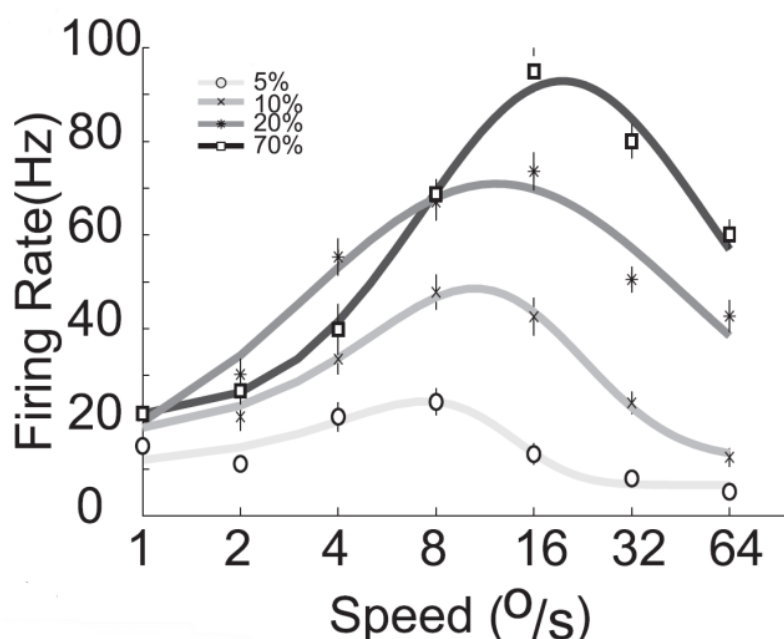


図 1.8 MT 細胞が最大反応を示す速度のコントラスト依存性の例 [11]. 提示刺激のコントラストを変化させた時の、刺激速度に対する MT 細胞の発火頻度をプロットしている. 高コントラスト時 (70% ; 濃い黒線) は約 20 deg/s の速度で最大反応を示すが, 低コントラスト時 (5% ; 薄い灰色線) では約 8 deg/s の速度で最大反応を示す.

に変化はない. また, ゲインコントロールによって信号強度が増強されても, 各方向に対する出力が定数倍されるだけであり半値幅は変化しない. コントラスト変化による半値幅の変化は V1 細胞の方位選択性に対しても見られる (iceberg 効果) [13]. しかし, iceberg 効果ではむしろコントラストを低下させた場合に半値幅が広がってしまい, MT 細胞の反応特性は説明できない.

上述したように, MT 細胞の反応特性は提示刺激のコントラストや模様依存して様々に変化する. だが, 頑健にオプティカルフロー推定を行うのであれば, このような依存性は望ましくない. そもそも, 「MT 細胞が速度抽出フィルタである」という

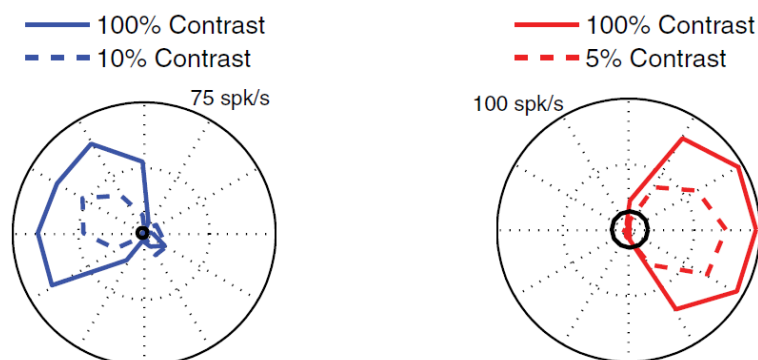


図 1.9 コントラストを変化させたときの、刺激の動く方向に対する MT 細胞の反応の変化 [12]. 各レーダーチャートは、中心からの距離が発火頻度を表し、偏角が刺激の動く方向を表している. 左図と右図はそれぞれ異なる MT 細胞に対する測定結果である. コントラストが低下しても MT 細胞が反応する方向は変化しないが、方向に対する感度は上昇する (preferred direction に近い方向でないと反応しなくなる).

考えに依らない視点からこれらの現象を考察できれば、新しい解釈や気づきが得られるのではないだろうか.

1.6 本研究の目的およびアプローチ

既存の MT 細胞モデルはどれも「速度抽出フィルタ」として構築されてきたが、本論文では速度抽出フィルタではない新しい考えのもと、新しい MT 細胞モデルを構築することを目的とする. 具体的には、「速度抽出フィルタ」に代わる概念として、「個々の MT 細胞はそれぞれが速度推定を行っている」という考えのもと数理モデルを構築する. 異なる概念から構築された MT 細胞モデルは、MT 細胞の反応特性に対して異なる解釈を与えるだけでなく、MT 細胞の出力を利用する MST 野などの細胞モデルや細胞集団の反応強度から知覚される速度を求める read-out モデルの構築方法にも影響

を与える。また、MT 細胞が入力として利用する V1 野の細胞モデルに要求される計算原理も異なる。

提案モデルで MT 細胞の電気生理学実験結果が再現できるかを調べ、その妥当性を評価する。また、1.5 節で紹介した「速度抽出フィルタ」の視点からでは解釈の難しい現象に対し、新しい解釈を与える。更に、MT 細胞集団の反応から提示された刺激の速度を求める read-out モデルを構築し、ヒトの速度知覚特性を説明できるか調べる。

1.7 本論文の構成

第 1 章では、本研究の位置づけ、研究対象、既存モデルの紹介、本研究の目的を記した。第 2 章では、新しい概念をもとにした新しい MT 細胞モデルの導出を行う。第 3 章では、研究対象である MT 細胞の電気生理学実験を提案モデルで再現できるか調べる。また、MT 細胞集団の反応から提示された刺激の速度を求める read-out モデルを導出する。第 4 章では、3 章で導出した read-out モデルを用いてヒトの速度知覚特性が説明できるか調べる。第 5 章では、式 (1) による grating 刺激に対し解析的に提案モデルの出力を求め、MT 細胞の反応特性を定量的に再現できるか調べる。第 6 章では、本論文のまとめおよび今後の展望について記す。

第 2 章

計算原理および提案モデル

本章では，MT 細胞の反応特性を再現する新しい視覚数理モデルの導出を行う．既存の MT 細胞モデルである Simoncelli & Heeger モデルは「特徴抽出フィルタ」という考えを基にしており，特定の時空間周波数に選択性を持つ V1 細胞の出力を集めることで，特定の速度に選択性を持つ MT 細胞モデルを構築している [9]．言い換えれば，時空間周波数抽出フィルタの出力を集め，速度抽出フィルタを構築している．これに対し本論文では，「個々の MT 細胞は Lucas-Kanade (LK) 法 [14] による速度推定を行っている」という前提のもと新しい数理モデルを構築した．LK 法は画像処理工学的速度推定手法であり，数学的関係性から提示刺激の速度を推定しようと試みる手法である．すなわち，速度選択性を再現するために構築された手法ではない．LK 法は提示刺激の速度をできる限り正確に推定しようと試みる手法であり，MT 細胞のような刺激速度 v に対する推定速度 \hat{v} のグラフは，理想的には直線となるはずである．しかし数値シミュレーションを行ったところ，詳細は第 3 章で示すが，その入出力特性 ($\hat{v} - v$ グラフ) は MT 細胞の反応特性と同様の単峰性のグラフを描くことが分かった．

本章ではまず，計算基盤である LK 法の導出方法を記す．次に，モデル式において

ゼロ除算を避ける処理を加える。最後に，MT 細胞の発火頻度および正規化された反応強度を求めるモデルを導出する。本章では数値シミュレーションにおける微分計算の実装方法や，提案モデルが細胞ネットワークで実現できることについても記す。

2.1 Lucas-Kanade(LK) 法

本節では，LK 法による 3 次元動画像から物体の動きを推定する手法を紹介する。まず 3 次元動画像 $I(x, y, t)$ と速度場 $\mathbf{v}(x, y, t) = (v_x(x, y, t), v_y(x, y, t))^T$ との数学的関係性を示す。次にその関係性から，未知の速度場 $\mathbf{v}(x, y, t)$ を 3 次元動画像 $I(x, y, t)$ から推定する方法を示す。

着目点 (x, y) 付近を速度 $\mathbf{v} = (v_x, v_y)^T$ で平面的に運動する物体を，固定点から撮影する状況を考える。「仮定 (a)：撮影画像の輝度の時間変化が物体の運動にのみ依存する」とすると，ある時刻 t で撮影された画像 $I(x, y, t)$ と，微小時間 Δt だけ経過した後

に撮影された画像 $I(x, y, t + \Delta t)$ との間には，局所的に次の関係が成り立つ。

$$I(x, y, t + \Delta t) = I(x - v_x \Delta t, y - v_y \Delta t, t) \quad (4)$$

ずれ量 $v_x \Delta t$ および $v_y \Delta t$ が十分に小さく，速度場 (v_x, v_y) が一定であり，「仮定 (b)：輝度値 I の空間変化が一次の Taylor 展開で近似できる」とすると，式 (4) の右辺は次のように近似できる。

$$I(x - v_x \Delta t, y - v_y \Delta t, t) \simeq I(x, y, t) - v_x \Delta t \frac{\partial I(x, y, t)}{\partial x} - v_y \Delta t \frac{\partial I(x, y, t)}{\partial y} \quad (5)$$

式 (4) と式 (5) から

$$I(x, y, t + \Delta t) \simeq I(x, y, t) - v_x \Delta t \frac{\partial I(x, y, t)}{\partial x} - v_y \Delta t \frac{\partial I(x, y, t)}{\partial y} \quad (6)$$

であり，整理すると

$$v_x \frac{\partial I(x, y, t)}{\partial x} + v_y \frac{\partial I(x, y, t)}{\partial y} + \frac{I(x, y, t + \Delta t) - I(x, y, t)}{\Delta t} \simeq 0 \quad (7)$$

である．微小時間 Δt が十分に小さい $\Delta t \rightarrow 0$ とすると，速度場 (v_x, v_y) と撮影画像 $I(x, y, t)$ との間には次の関係が成り立つ．

$$v_x \frac{\partial I(x, y, t)}{\partial x} + v_y \frac{\partial I(x, y, t)}{\partial y} + \frac{\partial I(x, y, t)}{\partial t} \simeq 0 \quad (8)$$

式 (8) は輝度勾配ベクトル ∇I を用いて表すと

$$\nabla I \cdot \begin{pmatrix} v_x \\ v_y \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\partial I(x, y, t)}{\partial x} & \frac{\partial I(x, y, t)}{\partial y} & \frac{\partial I(x, y, t)}{\partial t} \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ 1 \end{pmatrix} \simeq 0 \quad (9)$$

であり，数学的には輝度勾配ベクトルと動きベクトルとが直交することを意味する．

上記の関係性から，撮影画像 $I(x, y, t)$ から未知の速度場 (v_x, v_y) を推定する場合，式 (8) を満たす (v_x, v_y) が尤もらしい速度場 (\hat{v}_x, \hat{v}_y) である．しかし，式 (8) は2つの未知数が含まれるため一意に解が求まらない．そこで窓関数 $w(x, y)$ を考え，「仮定 (c) : 窓関数 $w(x, y)$ のサポート領域内で速度場は一定である」とする．尤もらしい速度場 (\hat{v}_x, \hat{v}_y) であれば窓関数内のいたるところで式 (8) が成り立つはずである．従って，次式に示すエネルギー関数 E が最小となる速度場 (v_x, v_y) が尤もらしい速度場 (\hat{v}_x, \hat{v}_y) である．

$$E = \iint w(x, y) \left(v_x \frac{\partial I(x, y, t)}{\partial x} + v_y \frac{\partial I(x, y, t)}{\partial y} + \frac{\partial I(x, y, t)}{\partial t} \right)^2 dx dy \quad (10)$$

式 (10) は v_x, v_y に対し下に凸な関数であるため，式 (10) が最小となる (\hat{v}_x, \hat{v}_y) は $\partial E / \partial v_x = 0$ かつ $\partial E / \partial v_y = 0$ となる (v_x, v_y) である．式 (10) の両辺を v_x, v_y でそれ

ぞれ偏微分してゼロとおくと、次の連立方程式が得られる。

$$\begin{cases} v_x S_{xx} + v_y S_{xy} + S_{xt} = 0 \\ v_x S_{xy} + v_y S_{yy} + S_{yt} = 0 \end{cases} \quad (11)$$

但し、 S_{ij} を次式で定義した。

$$S_{ij} \stackrel{\text{def}}{=} \iint w(x, y) \frac{\partial I(x, y, t)}{\partial i} \frac{\partial I(x, y, t)}{\partial j} dx dy \quad (12)$$

$(i, j = x, y, \text{ or } t)$

式 (11) を行列形式で表すと

$$\begin{pmatrix} S_{xx} & S_{xy} \\ S_{xy} & S_{yy} \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix} + \begin{pmatrix} S_{xt} \\ S_{yt} \end{pmatrix} = \mathbf{0} \quad (13)$$

であり、整理すると

$$\begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} S_{xx} & S_{xy} \\ S_{xy} & S_{yy} \end{pmatrix}^{-1} \begin{pmatrix} -S_{xt} \\ -S_{yt} \end{pmatrix} \quad (14)$$

である。従って、着目点 (x, y) における推定された尤もらしい速度場 $\mathbf{v} = (\hat{v}_x \hat{v}_y)^T$ は次式で求まる。

$$\hat{\mathbf{v}} = \begin{pmatrix} \hat{v}_x \\ \hat{v}_y \end{pmatrix} = \begin{pmatrix} S_{xx} & S_{xy} \\ S_{xy} & S_{yy} \end{pmatrix}^{-1} \begin{pmatrix} -S_{xt} \\ -S_{yt} \end{pmatrix} \quad (15)$$

式 (15) は画像中の 1 点に対して速度場の推定を行う式である。複数箇所の速度場を推定するために毎回式 (15) を適用しては、冗長な計算が多く効率が悪い。本論文では、 S_{ij} を関数として再定義し、式 (15) を関数として再表記することで計算時間の短縮を図った。

$$\begin{aligned} S_{ij}(x, y, t) &\stackrel{\text{def}}{=} w(x, y) * \frac{\partial I(x, y, t)}{\partial i} \frac{\partial I(x, y, t)}{\partial j} \\ &= \iint w(x - x', y - y') \frac{\partial I(x', y', t)}{\partial i} \frac{\partial I(x', y', t)}{\partial j} dx' dy' \end{aligned} \quad (16)$$

$$\begin{aligned}\hat{\mathbf{v}}(x, y, t) &= \begin{pmatrix} \hat{v}_x(x, y, t) \\ \hat{v}_y(x, y, t) \end{pmatrix} \\ &= \begin{pmatrix} S_{xx}(x, y, t) & S_{xy}(x, y, t) \\ S_{xy}(x, y, t) & S_{yy}(x, y, t) \end{pmatrix}^{-1} \begin{pmatrix} -S_{xt}(x, y, t) \\ -S_{yt}(x, y, t) \end{pmatrix}\end{aligned}\quad (17)$$

但し、式(16)中の*は畳み込み演算を意味する。

2.2 ゼロ除算の回避

式(17)は行列式 $\det \begin{pmatrix} S_{xx} & S_{xy} \\ S_{xy} & S_{yy} \end{pmatrix} = S_{xx}S_{yy} - S_{xy}^2$ がゼロとなる状況では逆行列が求まらず、速度場の推定が行えない。行列式がゼロとなる状況は、窓関数 $w(x, y)$ 内で特定の方向に対する輝度変化がゼロとなる状況である。例えば横縞や縦縞は $\partial I(x, y, t)/\partial x = 0$ または $\partial I(x, y, t)/\partial y = 0$ であり、行列式がゼロとなる。このような状況でも速度推定が行えるよう、本論文では次式による改良を行った。

$$\begin{aligned}\hat{\mathbf{v}}(x, y, t) &= \begin{pmatrix} \hat{v}_x(x, y, t) \\ \hat{v}_y(x, y, t) \end{pmatrix} \\ &= \left\{ \begin{pmatrix} S_{xx}(x, y, t) & S_{xy}(x, y, t) \\ S_{xy}(x, y, t) & S_{yy}(x, y, t) \end{pmatrix} + \varepsilon^2 E \right\}^{-1} \begin{pmatrix} -S_{xt}(x, y, t) \\ -S_{yt}(x, y, t) \end{pmatrix}\end{aligned}\quad (18)$$

ここで、 ε^2 はゼロ除算を避けるために加えたパラメータであり、 E は単位行列である。 ε^2 が十分に小さい場合には、式(18)は式(17)と等価である。式(18)を展開すると

$$\hat{\mathbf{v}} = \begin{pmatrix} \hat{v}_x \\ \hat{v}_y \end{pmatrix} = \begin{pmatrix} -\frac{(S_{yy} + \varepsilon^2)S_{xt} - S_{xy}S_{yt}}{(S_{xx} + \varepsilon^2)(S_{yy} + \varepsilon^2) - S_{xy}^2} \\ -\frac{(S_{xx} + \varepsilon^2)S_{yt} - S_{xy}S_{xt}}{(S_{xx} + \varepsilon^2)(S_{yy} + \varepsilon^2) - S_{xy}^2} \end{pmatrix}\quad (19)$$

である。横縞のような x 軸方向の輝度変化がゼロ ($\partial I(x, y, t)/\partial x = 0$) の場合には

$$\hat{\mathbf{v}} = \begin{pmatrix} \hat{v}_x \\ \hat{v}_y \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{S_{yt}}{S_{yy} + \varepsilon^2} \end{pmatrix}\quad (20)$$

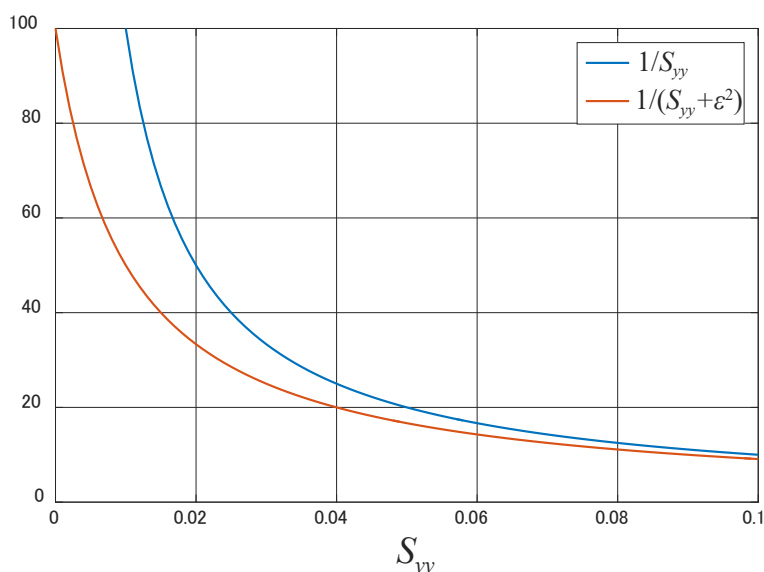


図 2.1 $1/(S_{yy} + \epsilon^2)$ と $1/S_{yy}$ の違い ($\epsilon^2 = 1.0 \times 10^{-2}$ とした場合). S_{yy} が大きい場合には $1/(S_{yy} + \epsilon^2) \approx 1/S_{yy}$ である. S_{yy} がゼロに漸近すると $1/S_{yy}$ は ∞ に発散してしまうが, $1/(S_{yy} + \epsilon^2)$ は $1/\epsilon^2$ (この図では $1/0.01 = 100$) に漸近する.

となり, x 軸方向の推定速度 \hat{v}_x はゼロとなる. 同様に, 縦縞のような y 軸方向の輝度変化がゼロ ($\partial I(x, y, t)/\partial y = 0$) の場合には

$$\hat{\mathbf{v}} = \begin{pmatrix} \hat{v}_x \\ \hat{v}_y \end{pmatrix} = \begin{pmatrix} -\frac{S_{xt}}{S_{xx} + \epsilon^2} \\ 0 \end{pmatrix} \quad (21)$$

となり, y 軸方向の推定速度 \hat{v}_y はゼロとなる. $\epsilon^2 \neq 0$ と $\epsilon^2 = 0$ の場合の比較として, $1/(S_{yy} + \epsilon^2)$ と $1/S_{yy}$ の違いを図 2.1 に示す. S_{yy} が十分に大きい場合には ϵ^2 の影響が無視でき, $1/(S_{yy} + \epsilon^2) \approx 1/S_{yy}$ である. 一方, S_{yy} がゼロに漸近すると $1/S_{yy}$ は ∞ に発散してしまうが, $1/(S_{yy} + \epsilon^2)$ は $1/\epsilon^2$ (図 2.1 では $1/0.01 = 100$) に漸近する.

このような処理を加えることで, 縦縞や横縞のようなゼロ除算が生じる入力パターンに対しても速度推定が行えるようになる.

2.3 提案する MT 細胞モデル

詳細は第3章で記すが、数値シミュレーションによるLK法の入出力特性から、式(18)で求まる \hat{v}_x が、 x 軸正の方向($\phi = 0^\circ$ 方向)に最大反応を示すMT細胞の発火頻度 $\text{MT}_{\phi=0^\circ}$ に比例するとした。また、式(18)で求まる \hat{v}_y が、 y 軸正の方向($\phi = 90^\circ$ 方向)に最大反応を示すMT細胞の発火頻度 $\text{MT}_{\phi=90^\circ}$ に比例するとした。細胞ごとに最大発火頻度は異なるため比例定数 a をパラメータとして、受容野中心が (x, y) のMT細胞の時刻 t における発火頻度を次式でモデル化した。

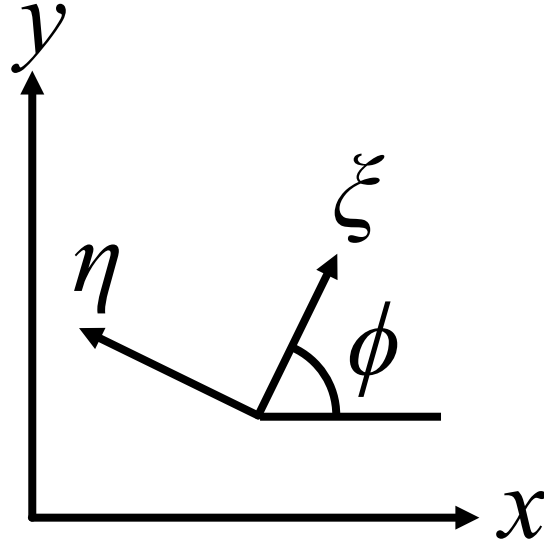
$$\begin{cases} \text{MT}_{\phi=0^\circ}(x, y, t) = a(x, y) \hat{v}_x(x, y, t) \\ \text{MT}_{\phi=90^\circ}(x, y, t) = a(x, y) \hat{v}_y(x, y, t) \end{cases} \quad (22)$$

任意の方向 ϕ に最大反応を示すMT細胞は、式(18)の座標軸を回転させることでモデル化できる。図2.2に示すように、 (x, y) 座標系を反時計回りに ϕ だけ回転させた新しい座標系 (ξ, η) を考えると、 ξ 方向および η 方向の推定速度 $(\hat{v}_\xi, \hat{v}_\eta)$ は次式で求まる。

$$\begin{aligned} \hat{\mathbf{v}}(\xi, \eta, t) &= \begin{pmatrix} \hat{v}_\xi(\xi, \eta, t) \\ \hat{v}_\eta(\xi, \eta, t) \end{pmatrix} \\ &= \left\{ \begin{pmatrix} S_{\xi\xi}(\xi, \eta, t) & S_{\xi\eta}(\xi, \eta, t) \\ S_{\xi\eta}(\xi, \eta, t) & S_{\eta\eta}(\xi, \eta, t) \end{pmatrix} + \varepsilon^2 E \right\}^{-1} \begin{pmatrix} -S_{\xi t}(\xi, \eta, t) \\ -S_{\eta t}(\xi, \eta, t) \end{pmatrix} \end{aligned} \quad (23)$$

$\phi = 0^\circ$ の場合には $\hat{v}_\xi = \hat{v}_x$ かつ $\hat{v}_\eta = \hat{v}_y$ であり、式(18)と等価である。また、 $\phi = 90^\circ$ の場合には $\hat{v}_\xi = \hat{v}_y$ かつ $\hat{v}_\eta = -\hat{v}_x$ である。但し、式(16)の i 偏微分および j 偏微分はそれぞれ、 i 方向の方向微分、 j 方向の方向微分に置き換えた。なお、 ξ 方向および η 方向の方向微分は x 偏微分および y 偏微分を用いて次のように表せる。

$$\begin{pmatrix} \frac{\partial I(\xi, \eta, t)}{\partial \xi} \\ \frac{\partial I(\xi, \eta, t)}{\partial \eta} \end{pmatrix} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} \frac{\partial I(x, y, t)}{\partial x} \\ \frac{\partial I(x, y, t)}{\partial y} \end{pmatrix} \quad (24)$$

図 2.2 (x, y) 座標系と (ξ, η) 座標系.

ゼロ除算を避けるパラメータ ε^2 の効果は回転に対して不変であり，窓関数 $w(x, y)$ がガウス窓の場合には次の関係がある．

$$\begin{pmatrix} \hat{v}_\xi(\xi, \eta, t) \\ \hat{v}_\eta(\xi, \eta, t) \end{pmatrix} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} \hat{v}_x(x, y, t) \\ \hat{v}_y(x, y, t) \end{pmatrix} \quad (25)$$

また，極座標系 (r, θ) は (ξ, η) 系の一例である． $\phi = \theta$ とすれば r 方向および θ 方向は ξ 方向および η 方向と一致する． ξ 方向および η 方向の方向微分は r 偏微分および θ 偏微分を用いて次のように表せる．

$$\begin{pmatrix} \frac{\partial I(\xi, \eta, t)}{\partial \xi} \\ \frac{\partial I(\xi, \eta, t)}{\partial \eta} \end{pmatrix} = \begin{pmatrix} \frac{\partial I(r, \theta, t)}{\partial r} \\ \frac{1}{r} \frac{\partial I(r, \theta, t)}{\partial \theta} \end{pmatrix} \quad (26)$$

MT 細胞の最大発火頻度は細胞ごとに異なるため，比例定数 a をパラメータとして，受容野中心が (x, y) の MT 細胞の時刻 t における発火頻度を次式でモデル化した．

$$\text{MT}_\phi(x, y, t) = a(x, y) \hat{v}_\xi(x, y, t) \quad (27)$$

これまで MT 細胞は特定の速度を抽出する速度抽出フィルタと見なされており、フィルタの入出力特性を比較するためには出力の最大値が 1 となるよう正規化を行って比較の方が都合が良い。MT 細胞の最大発火頻度は細胞ごとに異なり、図 1.5 のように最大発火頻度で正規化された反応特性で議論されることが多い [8, 15]。本論文でも、位置 (ξ, η) に受容野中心を持ち、 ϕ 方向に最大反応を示す MT 細胞の時刻 t における正規化された反応 $\text{MT}_\phi^{\text{norm}}$ を次式で定義した。

$$\begin{aligned} \text{MT}_\phi^{\text{norm}}(\xi, \eta, t) &= \frac{\text{MT}_\phi(\xi, \eta, t)}{\max_{\mathbf{v}} \text{MT}_\phi(\xi, \eta, t)} \\ &= \frac{a(\xi, \eta) \hat{v}_\xi(\xi, \eta, t)}{\max_{\mathbf{v}} \{a(\xi, \eta) \hat{v}_\xi(\xi, \eta, t)\}} \\ &= \frac{\hat{v}_\xi(\xi, \eta, t)}{\max_{\mathbf{v}} \hat{v}_\xi(\xi, \eta, t)} \end{aligned} \quad (28)$$

$\hat{v}_\xi(\xi, \eta, t)$ は提示刺激の ξ 方向の速度を推定した結果であり、 $\hat{v}_\xi - v_\xi$ グラフを描けば、理論的には ξ 方向の刺激速度 v_ξ に対し単調増加関数になるはずである。しかし数値シミュレーションの結果、詳細は 3.1 節に記すが、MT 細胞の入出力特性と類似した特定の速度 v_0 で最大値をとる単峰性のグラフを描くことが分かった。

式 (23) による ξ 方向の推定速度 \hat{v}_ξ を対応する視覚領野で表したのが図 2.3 である。既存モデルである SH モデルは、初期視覚野 (V1 野) 細胞の計算原理をフーリエ変換による時空間周波数解析と見なして MT 細胞モデルを構築している。これは、gabor 関数が V1 野細胞の受容野特性を精度良く記述できるからである [16]。一方提案モデルでは、V1 野細胞の計算原理を時空間微分と見なしている。これは、Gaussian derivative モデルでも V1 野細胞の受容野特性を精度良く記述できるからである [17]。

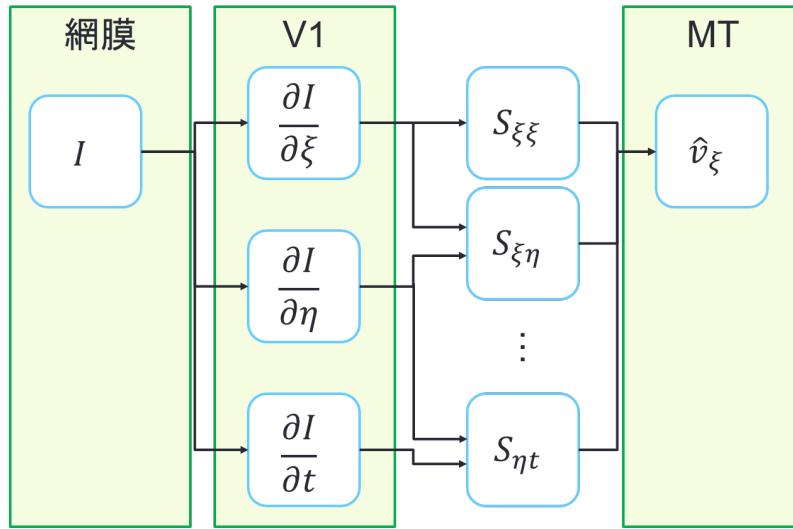


図 2.3 式 (23) による ξ 方向の推定速度 \hat{v}_ξ を，対応する視覚領野で表した図。

2.4 微分計算カーネル

数値計算において，微分計算には Sobel フィルタや Prewitt フィルタなど様々な方法がある [18]．本論文で行った数値シミュレーションでは，微分計算を Gaussian derivative [17] で実装した．

Dirac のデルタ関数 $\delta(x, y, t)$ の定義

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x - a, y - b, t - c) I(x, y, t) dx dy dt = I(a, b, c) \quad (29)$$

から，デルタ関数と任意の動画像との畳み込み $\delta(x, y, t) * I(x, y, t)$ は

$$\delta(x, y, t) * I(x, y, t) = \iiint \delta(x - x', y - y', t - t') I(x', y', t') dx' dy' dt' \quad (30)$$

$$= I(x, y, t) \quad (31)$$

であり， $I(x, y, t)$ の偏微分 $\partial I(x, y, t) / \partial x$, $\partial I(x, y, t) / \partial y$, $\partial I(x, y, t) / \partial t$ はそれぞれ次式で

表せる.

$$\begin{aligned}
\frac{\partial I(x, y, t)}{\partial x} &= \frac{\partial}{\partial x} (\delta(x, y, t) * I(x, y, t)) \\
\frac{\partial I(x, y, t)}{\partial y} &= \frac{\partial}{\partial y} (\delta(x, y, t) * I(x, y, t)) \\
\frac{\partial I(x, y, t)}{\partial t} &= \frac{\partial}{\partial t} (\delta(x, y, t) * I(x, y, t))
\end{aligned} \tag{32}$$

ここで, $\delta(x, y, t) = \delta(x) \delta(y) \delta(t)$ であり, デルタ関数 $\delta(X)$ はガウス関数 $G(X; \sigma_X^2) = \frac{1}{\sigma_X \sqrt{2\pi}} \exp\left(-\frac{X^2}{2\sigma_X^2}\right)$ において $\sigma_X \rightarrow 0$ の状況と等価である. 従って式 (32) はガウス関数 $G(x; \sigma_x^2), G(y; \sigma_y^2), G(t; \sigma_t^2)$ を用いてそれぞれ次式のように近似できる.

$$\begin{aligned}
\frac{\partial I(x, y, t)}{\partial x} &\simeq \frac{\partial}{\partial x} \left\{ \left(G(x; \sigma_x^2) G(y; \sigma_y^2) G(t; \sigma_t^2) \right) * I(x, y, t) \right\} \\
\frac{\partial I(x, y, t)}{\partial y} &\simeq \frac{\partial}{\partial y} \left\{ \left(G(x; \sigma_x^2) G(y; \sigma_y^2) G(t; \sigma_t^2) \right) * I(x, y, t) \right\} \\
\frac{\partial I(x, y, t)}{\partial t} &\simeq \frac{\partial}{\partial t} \left\{ \left(G(x; \sigma_x^2) G(y; \sigma_y^2) G(t; \sigma_t^2) \right) * I(x, y, t) \right\}
\end{aligned} \tag{33}$$

微分および畳み込み演算の線形性より, 上式は次のように表せる.

$$\begin{aligned}
\frac{\partial I(x, y, t)}{\partial x} &\simeq \frac{\partial}{\partial x} \left\{ \left(G(x; \sigma_x^2) G(y; \sigma_y^2) G(t; \sigma_t^2) \right) * I(x, y, t) \right\} \\
&= \left\{ \left(G(y; \sigma_y^2) G(t; \sigma_t^2) \frac{\partial}{\partial x} G(x; \sigma_x^2) \right) * I(x, y, t) \right\} \\
&= \left\{ G(y; \sigma_y^2) G(t; \sigma_t^2) G(x; \sigma_x^2) \right\} * \left\{ \frac{\partial}{\partial x} I(x, y, t) \right\} \\
\frac{\partial I(x, y, t)}{\partial y} &\simeq \frac{\partial}{\partial y} \left\{ \left(G(x; \sigma_x^2) G(y; \sigma_y^2) G(t; \sigma_t^2) \right) * I(x, y, t) \right\} \\
&= \left\{ \left(G(x; \sigma_x^2) G(t; \sigma_t^2) \frac{\partial}{\partial y} G(y; \sigma_y^2) \right) * I(x, y, t) \right\} \\
&= \left\{ G(y; \sigma_y^2) G(t; \sigma_t^2) G(x; \sigma_x^2) \right\} * \left\{ \frac{\partial}{\partial y} I(x, y, t) \right\} \\
\frac{\partial I(x, y, t)}{\partial t} &\simeq \frac{\partial}{\partial t} \left\{ \left(G(x; \sigma_x^2) G(y; \sigma_y^2) G(t; \sigma_t^2) \right) * I(x, y, t) \right\} \\
&= \left\{ \left(G(x; \sigma_x^2) G(y; \sigma_y^2) \frac{\partial}{\partial t} G(t; \sigma_t^2) \right) * I(x, y, t) \right\} \\
&= \left\{ G(x; \sigma_x^2) G(y; \sigma_y^2) G(t; \sigma_t^2) \right\} * \left\{ \frac{\partial}{\partial t} I(x, y, t) \right\}
\end{aligned} \tag{34}$$

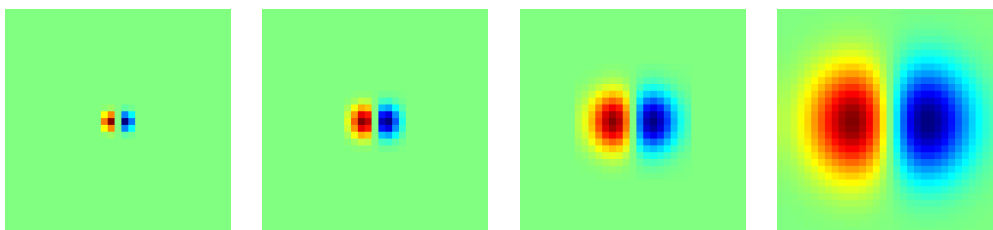


図 2.4 カーネルサイズ $k \in \{5, 9, 17, 33\}$ での x 偏微分を計算する微分カーネル. 左から順に, $k = 5, k = 9, k = 17, k = 33$ の場合である. 但し比較しやすいよう, 全て 33×33 pixel の画像とし, それぞれのカーネルにおいて最大値が赤, 最小値が青となるよう可視化した.

式 (34) は, 動画像 $I(x, y, t)$ の時空間微分 $\partial I(x, y, t)/\partial x, \partial I(x, y, t)/\partial y, \partial I(x, y, t)/\partial t$ は, ガウス関数の偏微分と動画像 $I(x, y, t)$ との畳み込みで近似できることを意味している.

本論文では $G(t; \sigma_t^2) = \delta(t), \sigma_x = \sigma_y = \sigma$ とし, 時空間微分の計算には $3\sigma = k/2$ となる $k \times k$ pixel の微分カーネルを用いた. 例えば 5×5 pixel の空間微分カーネルの場合, $\sigma = 0.83$ pixel のガウス関数を偏微分した微分カーネルを用いた. 図 2.4 に, カーネルサイズ $k \in \{5, 9, 17, 33\}$ での空間微分カーネルを例示する. $3\sigma = k/2$ としているため, カーネルサイズ k が大きくなると大域的な輝度変化を計算するフィルタであることが分かる. 時間微分に関しては, フレーム差分に対し $3\sigma = k/2$ となる $k \times k$ pixel のガウス関数を畳み込んだ結果を用いた.

2.5 むすび

本章では, 式 (27) による MT 細胞の発火頻度を求めるモデル, および式 (28) による MT 細胞の正規化された反応強度を求めるモデルを提案した. 式 (28) はモデルパラメータとして, 受容野中心 (x, y) , 最大反応を示す方向 (速度推定を行う方向) ϕ , 時空間微分カーネルサイズ k , 窓関数の空間的広がり Γ の大きさがある. また, 式 (27)

はこれに加えて推定速度から平均発火頻度へ変換する比例定数 a がある。提案モデルは導出するにあたり、以下の仮定を置いている。

- (a) 輝度値の時間変化は物体の動きによってのみ生じる
- (b) 輝度値の空間的な変化に対し Taylor の一次近似が成り立つ
- (c) 窓関数 $w(x, y)$ 内で速度 \mathbf{v} は一定である
- (d) 位置 (x, y) における ξ 方向の推定速度 \hat{v}_ξ が、受容野中心 (x, y) で ξ 方向に最大反応を示す MT 細胞の発火頻度に比例する

(a)~(c) は Lucas-Kanade 法導出時においた仮定であり、これらの仮定が崩れる状況では正しく速度推定が行えない。具体的には以下の状況では各仮定が崩れ、正しく速度推定が行えない。

- (i) 画像中の物体が突然消失・出現する場合
- (ii) 2 枚の画像間で動き $\mathbf{v}(x, y, t)$ が大きすぎる場合やルーフエッジのような輝度値が 2 次関数的に変化する場合
- (iii) 窓関数 $w(x, y)$ のサポート領域内で速度が一定でない場合

このような状況で得られた推定結果は工学的には意味のない値である。だが興味深いことに、詳細は第 4 章に記すが、工学的には意味のない値がヒトの知覚する運動錯視（動きに関する錯覚）と定性的に一致することが分かった。

MT 細胞の計算原理を「速度抽出フィルタ」と見るか「速度推定器」と見るかによって、V1 野細胞モデルに要求される計算原理も異なる。既存モデルの計算原理である「速度抽出フィルタ」と見なした場合には V1 野細胞は時空間周波数解析を行っている

と見なせ, 提案モデルの計算原理である「速度推定器」と見なした場合には V1 野細胞は時空間微分を行っているから見なせる.

第 3 章

MT 細胞の反応特性の再現

本章では，第 2 章で導出した MT 細胞モデルが，これまでに行われてきた MT 細胞の電気生理学実験結果と定性的に一致することを示す．具体的には，入力画像の速度に対する LK 法による速度推定結果 ($\hat{v} - v$ グラフ) が，図 1.4 に示すような MT 細胞の反応特性同様，単峰性のグラフを描くことを示す．また，微分カーネルサイズを変えることで，図 1.5 に示すような異なる速度で最大反応を示す入出力特性が得られることを示す．更に，刺激の動く方向に対する反応特性や，第 1 章で紹介した，既存概念である「速度抽出フィルタ」では解釈の難しい現象に対しても，提案モデルで再現できるか調べる．

提案モデルは既存の MT 細胞モデルとは異なる概念で構築したため，細胞集団の反応強度から提示刺激の速度を求める read-out モデルも異なるモデルを構築する必要がある．本章では MT 細胞の反応特性に対する新しい解釈を提唱し，得られた新しい解釈に基づく新しい read-out モデルの導出も行う．また，既存の read-out モデルとの比較も行う．

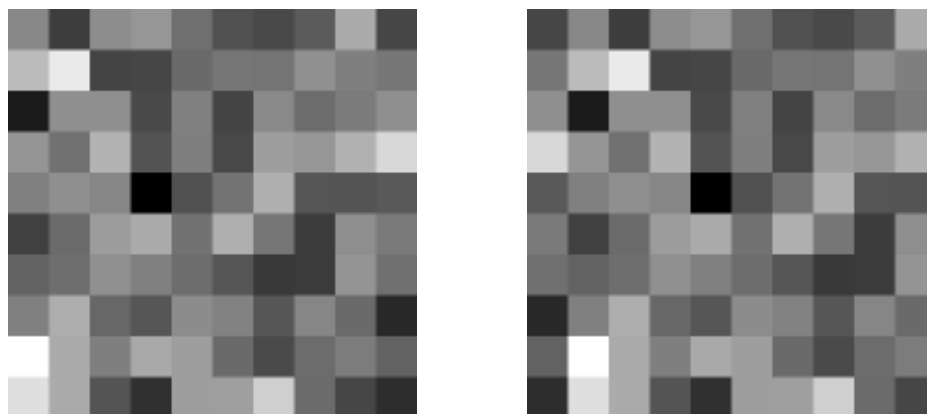


図 3.1 入力画像の例. 各ピクセル値がガウス分布からのサンプル値で決まるガウシアンランダムドット画像. 最小値を黒, 最大値を白として可視化した. 左図は元画像, 右図は元画像を 1 pixel/frame 右方向に動かした画像.

3.1 様々な刺激速度に対する速度推定結果

本節では, 様々な刺激速度に対する LK 法による速度推定結果 ($\hat{v} - v$ グラフ) を求め, MT 細胞同様, 特定の速度で最大となる単峰性のグラフを描くことを示す. 数値シミュレーションに用いた画像は, 図 3.1 に示すようなガウシアンランダムドット画像 (各ピクセル値がガウス分布からのサンプリング値である画像) とした. 画像サイズは 150×150 pixel とし, 20 組用意した. 動き v は x 軸方向 (水平方向) に限定し, ピクセル値を循環シフトすることで実現した. また, サブピクセルの動き (1 ピクセル以下の動き) は, 入力画像を縮小することで実現した. なお, 式 (18) におけるモデルパラメータは $\varepsilon^2 = 1.0 \times 10^{-4}$, $\Gamma = 11$ pixel とし, 微分カーネルサイズ k は $k = 5$ pixel とした. 式 (18) による速度推定結果はピクセルごとに求まるが, 以降特に断りがなければ画像中心に対する推定結果に着目をする.

図 3.2 に式 (18) による LK 法による速度推定結果の平均値 ($\hat{v}_x - v_x$ グラフ) を示す. 横

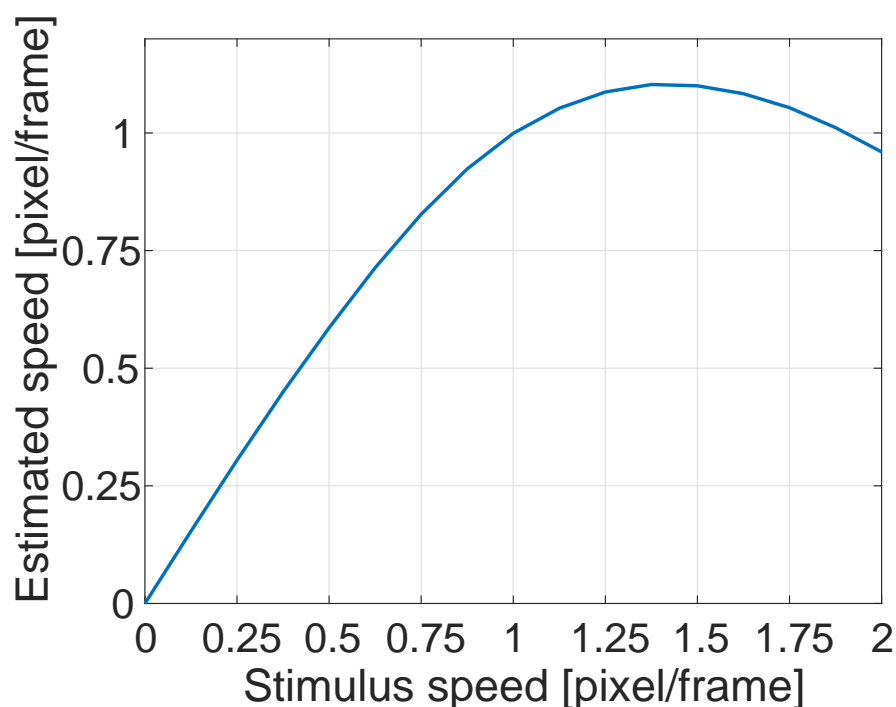


図 3.2 刺激速度に対するモデル出力の平均値（平均推定速度）。理想的には傾き 1 の直線となるはずだが、提示刺激の速度が約 1 pixel/frame を超えると推定速度が減少する傾向にある。

軸は x 軸方向の刺激速度 v_x ，縦軸は x 軸方向の推定速度 \hat{v}_x の平均値である。LK 法は刺激速度を出来るだけ正確に推定しようと試みる手法であり， $\hat{v}_x - v_x$ グラフは理想的には傾き 1 の直線となるはずである。しかし実際には，約 1 pixel/frame を境に正しく速度推定が行えていないことがわかる。1 pixel/frame を超える刺激速度に対し正しく速度推定が行えない理由は，モデル導出時に設定した「仮定 (b)：輝度値の空間変化は一次の Taylor 展開で近似できる」という仮定を満たさなくなるためである。すなわち，LK 法による速度推定には，速度推定可能な上限の速度が存在する。

3.2 微分カーネルサイズと上限速度との関係

前節では、モデル導出時の仮定を満たさなくなるため、特定の速度を超える動画像に対しては正しく速度推定が行えないことを示した。本節では、微分カーネルサイズを大きくすることで、この速度推定可能な上限の速度を高速側にシフトさせられることを示す。また、様々な微分カーネルサイズでの速度推定結果が図 1.5 に示す MT 細胞の反応特性と類似した結果となることを示す。

前節の数値シミュレーション(図 3.2)では、微分計算に $k \times k = 5 \times 5$ pixel の Gaussian derivative カーネルを用いた。本節では octave コンセプトに従い、 $k_i = 2^{i+1} + 1$ の Gaussian derivative カーネルを用いて再シミュレーションを行った。図 3.3 に、 $k_i \in \{5, 9, 17, 33\}$ におけるそれぞれの速度推定結果 $\hat{v}_x(k_i)$ の平均値を示す。この図から、小さい微分カーネル ($k_1 = 5$ pixel) を用いた場合は速度推定可能な上限の速度が約 1 pixel/frame と低速であるが、大きい微分カーネル ($k_3 = 17$ pixel) を用いた場合には上限の速度が約 4 pixel/frame と高速側にシフトしていることが見てとれる。式 (34) から、Gaussian derivative はガウス関数による空間的平滑化(ぼかし)と微分計算とを同時に行っていることが分かる。大きい微分カーネルを用いるということは、小さい微分カーネルよりも広い空間に対して平滑化を行うことと等価である。平滑化を行うことで局所的な輝度変化が弱くなり、相対的に大域的な輝度変化が重視されるため、Taylor の一次近似が成り立ちやすくなる。 $k_4 = 33$ pixel の微分カーネルでの速度推定結果が全体的に遅く見積もられているのは、平滑化をかけすぎたために輝度変化が小さくなってしまったことが原因であると考えられる。空間微分が小さくなると、ゼロ除算を避けるために導入したパラメータ ε^2 の影響が相対的に強くなる。 ε^2 は 1 推定

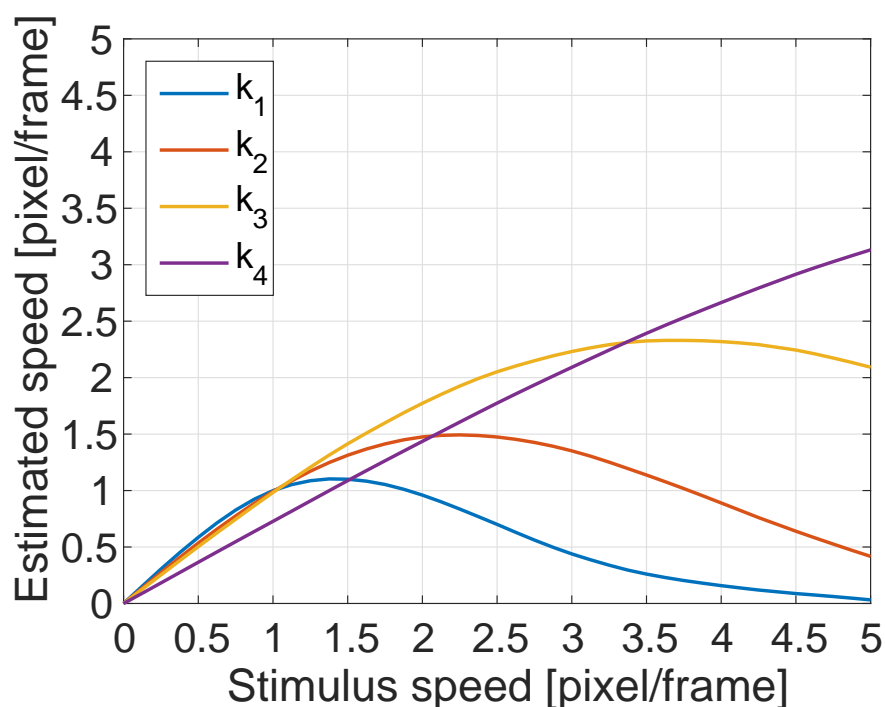


図 3.3 様々なカーネルサイズに対するモデル出力。横軸は入力刺激の速度 v_x ，縦軸はカーネルサイズ k_i で求めた推定速度 $\hat{v}_x(k_i)$ である。小さい微分カーネル (k_1 ；図 3.2 で用いた微分カーネル) では速度推定可能な上限の速度が約 1 pixel/frame だが，大きい微分カーネル (k_3) を用いた場合には約 4 pixel/frame に上昇している。

速度をゼロに落とし込む作用があるため，空間的な輝度変化が小さくなった $k_4 = 33$ pixel の微分カーネルでの速度推定結果は真の刺激速度よりも全体的に低く見積もられた。

図 3.3 から，一見すると大きめ微分カーネルを用いるのが良いように思われる。今回のシミュレーションで用いたような全体が一様に動く画像を入力とした場合には，確かに大きい微分カーネルを用いた方が良い。しかし，小さな物体が動いているような局所的な動きを推定する必要がある場合，大きすぎる微分カーネルでは正しく速度推定が行えない。なぜなら，大きすぎる微分カーネルではモデル導出時の仮定 (c) 「窓

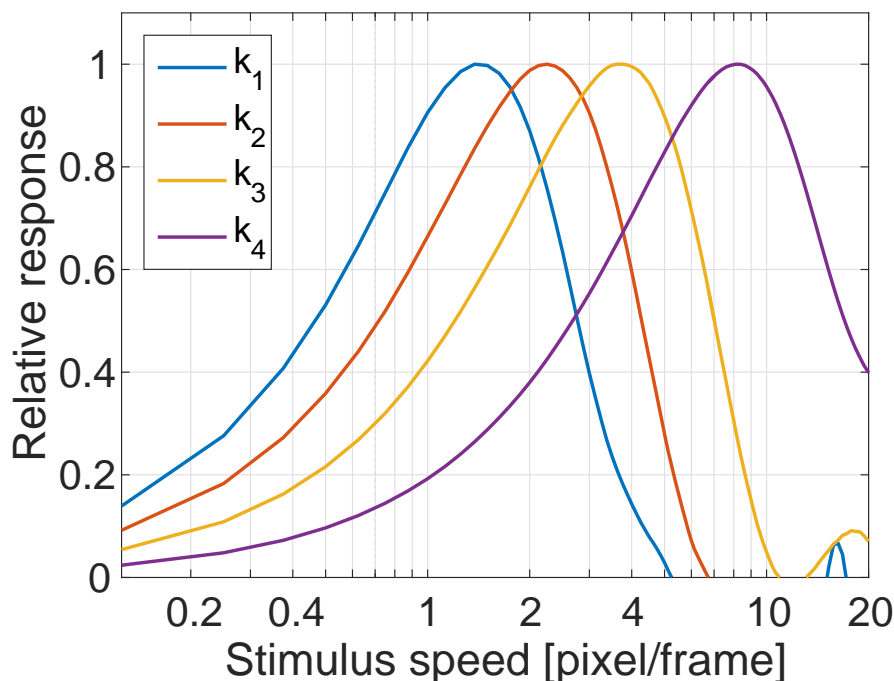


図 3.4 図 3.3 の各推定結果 $\hat{v}_x(k_i)$ を最大値が 1 になるよう正規化してプロットした図。但し、横軸を \log スケールの片対数とした。図 1.5 との類似性から、異なる速度に最大反応を示す MT 細胞は異なる微分カーネルサイズで速度推定を行っていると思なすことができる。

関数 $w(x, y)$ 内で動きは一定である」を満たさなくなる可能性があるからである。

図 3.4 は、図 3.3 の各推定結果 $\hat{v}_x(k_i)$ を最大値が 1 になるよう正規化してプロットした図である。すなわち、 $MT_{\phi=0^\circ}^{\text{norm}}(k_i)$ の入出力特性である。図 1.5 と図 3.4 の類似性から、異なる速度で最大反応を示す MT 細胞は、異なる微分カーネルサイズで速度推定を行っていると思なすことができる。また、preferred speed と呼ばれる最大反応を示す速度は、速度推定可能な上限の速度であると言える。

最大反応を示す速度以外に、半値幅（刺激速度に対する半値全幅）もモデル出力を評価する上で大事な指標である。Maunsell と van Essen は、MT 細胞の刺激速度に対

する半値幅は、平均して 2.9 octaves であると報告している [8]. 図 3.4 の $\hat{v}_x(k)$ の半値幅を調べたところ、 $k = 5, 9, 17, 33$ pixel でそれぞれ 2.6, 2.6, 2.5, 2.7 octaves であった. この点からも、提案モデルが妥当であると言える.

既存の「速度抽出フィルタ」の観点から見れば、MT 細胞が異なる速度に対して最大反応を示すのは、フィルタとして異なる速度を抽出するためであると解釈されてきた. しかし、MT 細胞を「速度推定器」として見れば、MT 細胞が異なる速度に対して最大反応を示すのは、個々の MT 細胞がそれぞれ異なる微分カーネルサイズで速度推定を行っているからである、という新しい解釈が得られた. また、最大反応を示す速度に対しても、「速度抽出フィルタ」と見なした場合「抽出すべき速度」であるのに対し、「速度推定器」と見なした場合「正しく速度推定が行える上限の速度」であるという新しい解釈が得られた.

3.3 方向選択性の再現

前節では、刺激速度に対する LK 法の入出力特性が、MT 細胞同様、単峰性のグラフを描くことを示した. 本節では、1.3 節で述べた MT 細胞の方向選択性について、提案モデルで再現できることを示す. 用いた刺激は、3.1 節同様ガウシアンランダムドット画像で、任意の方向に 1 pixel/frame 動かした画像を 500 組用意した. また、モデルパラメータも空間微分カーネルサイズを $k = 3$ とした以外は 3.1 節と同じパラメータを用いた.

図 3.5 に、様々な方向に動かした時の式 (18) による x 軸方向 (0° 方向) の推定速度 \hat{v}_x の平均値を示す. \hat{v}_x は x 軸方向 (0° 方向) の速度を推定するため、 0° 方向に動く刺激に対し最も強く反応する. 図 3.5 が円形の軌跡を描くのは解析的にも明らかであ

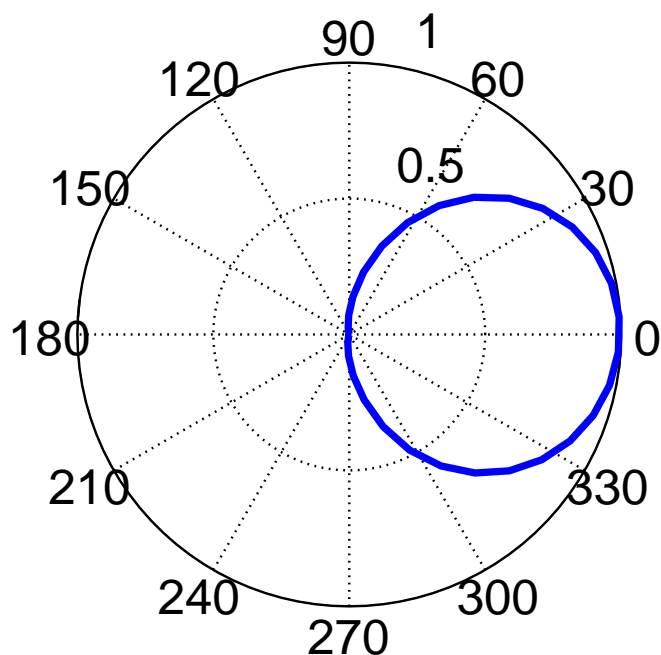


図 3.5 様々な方向に刺激を動かした時の x 軸方向 (0° 方向) の推定速度 \hat{v}_x の平均値。偏角が刺激の動く方向 ϕ_s ，中心からの距離が刺激方向 ϕ_s に対する推定速度 \hat{v}_x の平均値である。刺激速度 v が推定可能域の速度の場合 $v \cos(\phi_s)$ がモデル出力となり，このような正円状のグラフになる。

る。刺激の動く方向 ϕ_s に対する推定速度 \hat{v}_x は，正しく速度推定が行える刺激速度 v であれば， $\hat{v}_x = v \cos(\phi_s)$ となる。この推定結果 \hat{v}_x を図 3.5 のような極座標系 (r, θ) でプロットすると， $r = \hat{v}_x, \theta = \phi_s$ であり，直交座標系 (x', y') とは次の関係にある。

$$\begin{cases} x' = r \cos \theta = v \cos^2(\phi_s) = \frac{v \cos 2\phi_s}{2} + \frac{v}{2} \\ y' = r \sin \theta = v \cos(\phi_s) \sin(\phi_s) = \frac{v \sin 2\phi_s}{2} \end{cases} \quad (35)$$

整理すると

$$\begin{cases} x' - \frac{v}{2} = \frac{v}{2} \cos(2\phi_s) \\ y' = \frac{v}{2} \sin(2\phi_s) \end{cases} \quad (36)$$

である。ここで、 $\cos^2(2\phi_s) + \sin^2(2\phi_s) = 1$ を用いると、

$$\left(x' - \frac{v}{2}\right)^2 + (y')^2 = \left(\frac{v}{2}\right)^2 \quad (37)$$

となり、半径 $\frac{v}{2}$ 、中心座標 $(x', y') = (\frac{v}{2}, 0)$ の円を描く。式 (23) に示す ϕ 方向の推定速度 \hat{v}_ξ に関しても、同様に円形の軌跡を描く。刺激の動く方向 ϕ_s に対する推定結果 \hat{v}_ξ は、正しく速度推定が行える刺激速度 v であれば、 $\hat{v}_\xi(\phi) = v \cos(\phi_s - \phi)$ となる。この推定結果 $\hat{v}_\xi(\phi)$ を図 3.5 のような極座標系 (r, θ) でプロットすると、 $r = \hat{v}_\xi, \theta = \phi_s$ であり、直角座標系 (x', y') とは次の関係にある。

$$\begin{cases} x' = r \cos \theta = v \cos(\phi_s - \phi) \cos(\phi_s) = v \cos \phi \left(\frac{\cos 2\phi_s}{2} - \frac{1}{2}\right) + v \sin \phi \left(\frac{\sin 2\phi_s}{2}\right) \\ y' = r \sin \theta = v \cos(\phi_s - \phi) \sin(\phi_s) = v \cos \phi \left(\frac{\sin 2\phi_s}{2}\right) - v \sin \phi \left(\frac{\cos 2\phi_s}{2} - \frac{1}{2}\right) \end{cases} \quad (38)$$

整理すると

$$\begin{cases} x' - \frac{v}{2} \cos \phi = \frac{v}{2} \cos(2\phi_s - \phi) \\ y' - \frac{v}{2} \sin \phi = \frac{v}{2} \sin(2\phi_s - \phi) \end{cases} \quad (39)$$

である。ここで、 $\cos^2(2\phi_s - \phi) + \sin^2(2\phi_s - \phi) = 1$ を用いると、

$$\left(x' - \frac{v}{2} \cos \phi\right)^2 + \left(y' - \frac{v}{2} \sin \phi\right)^2 = \left(\frac{v}{2}\right)^2 \quad (40)$$

となり、半径 $\frac{v}{2}$ 、中心座標 $(x', y') = (\frac{v}{2} \cos \phi, \frac{v}{2} \sin \phi)$ の円を描く。

次に、刺激を動かす方向 ϕ_s を 30° 間隔にし、局所座標系の回転角 ϕ を $\phi = 219^\circ$ とした時の式 (23) による ξ 方向（動径方向）の推定速度 \hat{v}_ξ の平均値を求めた。その結果を図 3.6 に示す。 ϕ_s の間隔が十分に小さい場合には円状の軌跡を描くが、 30° 間隔では多角形状の軌跡を描く。図 1.6 と図 3.6 との類似性から、MT細胞の方向選択性についても提案モデルで定性的に再現できることが分かる。MT細胞が最大反応を示す方向（preferred direction）は、MT細胞を既存の速度抽出フィルタと見れば抽出すべ

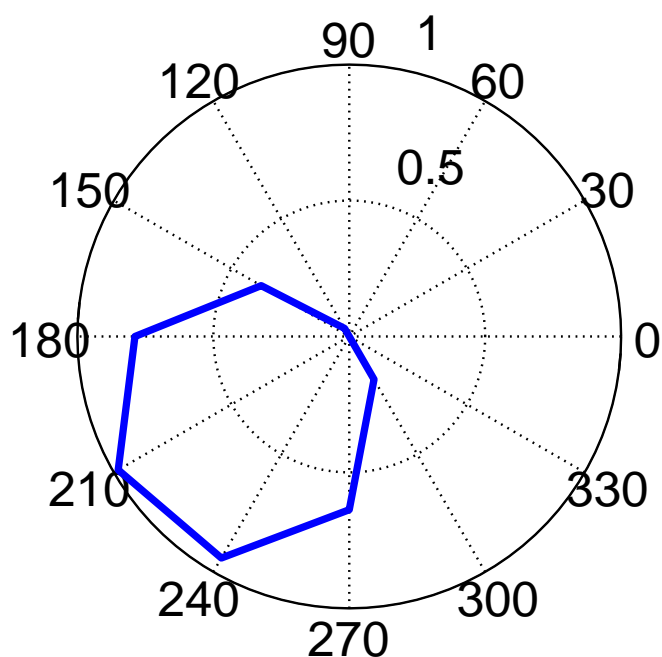


図 3.6 局所座標系の回転角 ϕ を $\phi = 219^\circ$ とした時の式 (23) による ξ 方向（動径方向）の推定速度 \hat{v}_ξ の平均値. 刺激の動く方向 ϕ_s のサンプル間隔を 30° とした場合. MT 細胞の刺激方向に対する反応特性（図 1.6）と類似した入出力特性を示す.

き刺激の移動方向であると解釈でき、新しい計算原理である速度推定器と見ればその MT 細胞が速度推定を試みる方向であると解釈できる.

3.4 Band 幅のコントラスト依存性の再現

本節では、1.5 節で述べた「速度抽出フィルタ」という既存の観点からでは解釈の難しい現象である、band 幅（半値全幅）のコントラスト依存性が提案モデルで再現できることを示す.

図 1.9 は、図 3.7 に示す grating 刺激を様々な方向に動かした時の MT 細胞の発火頻

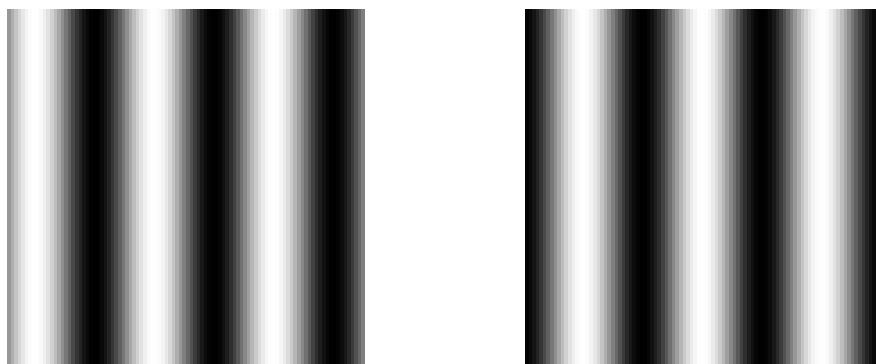


図 3.7 Grating 刺激の例.

度をプロットした図である。高コントラスト時には band 幅が広く, preferred direction ϕ と刺激の動く方向 ϕ_s がある程度異なっても反応を示す。一方で低コントラスト時には band 幅が狭く, $|\phi - \phi_s|$ がある程度小さくないと反応を示さない。

提案モデルでこの現象が再現できるか試みた。入力画像のサイズは 200×200 pixel で, 空間周波数は $f_x = 6$, $f_y = 0$ cycle/image とし, 時間周波数は $f_t = 6$ cycle/s とした。また, フレームレートは 30 frame/s として 50 frame 分用意した。コントラストに関しては, grating 刺激の振幅が 1 である画像をコントラスト 100% とし, 振幅を c 倍することで様々なコントラストの画像を作成した。モデルパラメータは 3.3 節同様, 推定する方向を $\phi = 0^\circ$, 時空間微分カーネルサイズを $k \times k = 3 \times 3$ pixel, 受容野に相当する窓関数 w は $\Gamma = 11$ ($3\sigma^2 = \Gamma/2$) のガウス関数とした。また, 画像中心における ξ 方向の推定結果 \hat{v}_ξ に着目をした。

図 3.8 に, コントラストを変化させたときの, $\phi = 0^\circ$ での速度推定結果 \hat{v}_ξ の平均値を示す。但し, 最大値が 1 となるよう正規化して可視化した。高コントラスト時 (100%; 振幅 1) では刺激の動く方向が $\phi = 0^\circ$ と異なってもある程度大きな出力が得られるが, 低コントラスト時 (10%; 振幅 0.1) では刺激の動く方向が 0° に近くないと出力が大きくならないという結果が得られた。

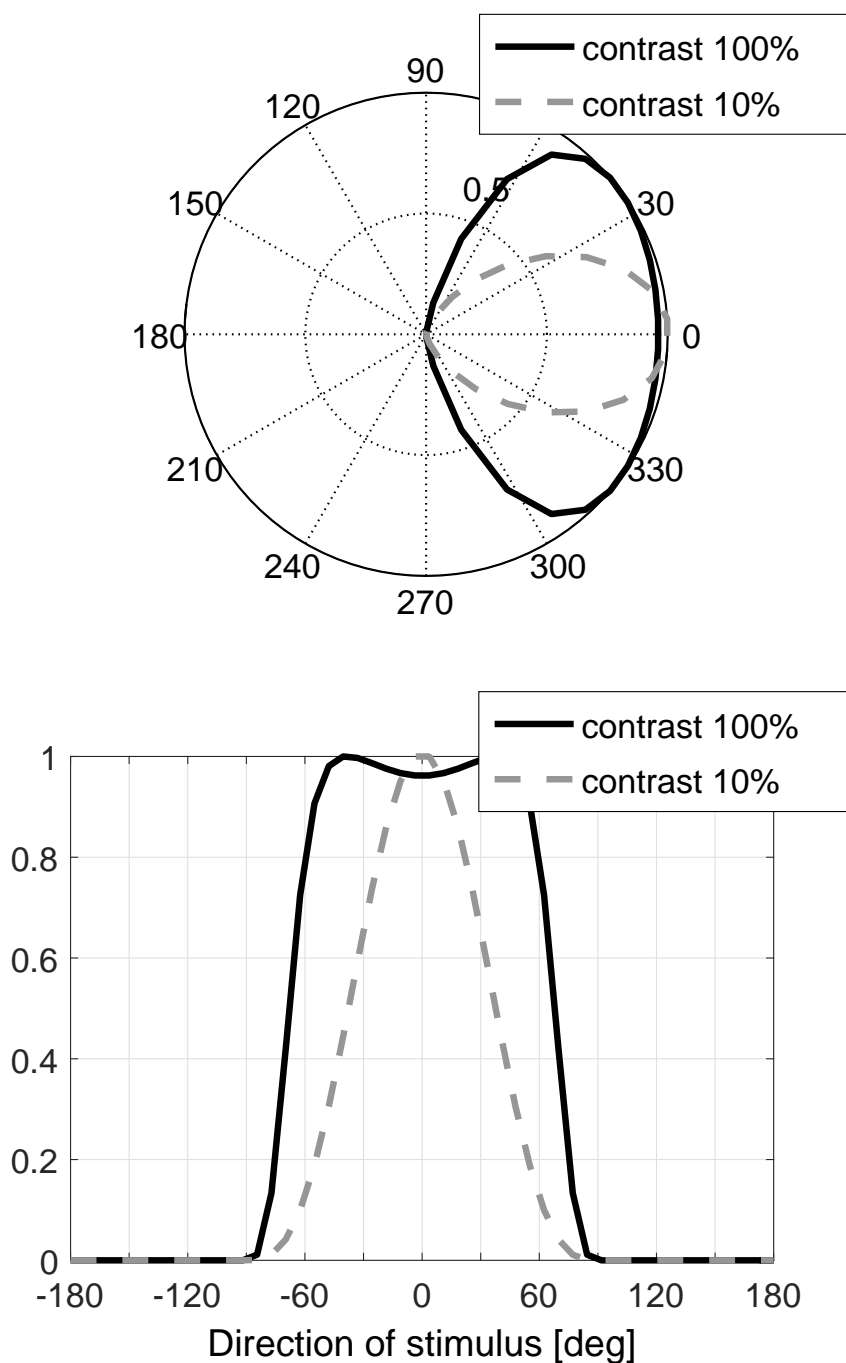


図 3.8 コントラストを変化させたときの、 $\phi = 0^\circ$ でのモデル出力。上図：刺激の動く方向を偏角，その時のモデル出力を中心からの距離で表した極座標プロット。下図：刺激の動く方向に対するモデル出力。高コントラスト（100%）時では刺激の動く方向が 0° と異なっているにもかかわらずある程度大きな出力が得られるが，低コントラスト（10%）時では刺激の動く方向が 0° に近くないと出力が大きくなる。

提案モデルでこのような依存性が生じる原因を調べるため、入力画像 I の振幅を c 倍した cI に置き換え、推定結果 \hat{v}_ξ にどのような変化が生じるか解析的に調べた。 $\phi = 0^\circ$ において、 ξ 方向の推定速度 \hat{v}_ξ は

$$\hat{v}_\xi = \hat{v}_x = -\frac{(S_{yy} + \varepsilon^2)S_{xt} - S_{xy}S_{yt}}{(S_{xx} + \varepsilon^2)(S_{yy} + \varepsilon^2) - S_{xy}^2} \quad (41)$$

であり、 I を cI に置き換えると、

$$\begin{aligned} \hat{v}'_\xi &= -\frac{(c^2S_{yy} + \varepsilon^2)c^2S_{xt} - c^4S_{xy}S_{yt}}{(c^2S_{xx} + \varepsilon^2)(c^2S_{yy} + \varepsilon^2) - c^4S_{xy}^2} \\ &= -\frac{(S_{yy} + \varepsilon^2/c^2)S_{xt} - S_{xy}S_{yt}}{(S_{xx} + \varepsilon^2/c^2)(S_{yy} + \varepsilon^2/c^2) - S_{xy}^2} \end{aligned} \quad (42)$$

となる。従って、入力画像のコントラストを c 倍することは、 ε^2 を $(\varepsilon/c)^2$ に置き換えることと等価である。 c は 1 以下の正の値であり、コントラストを小さくすると相対的に ε^2 の値は大きくなる。

図 1.9 と図 3.8 の類似性から、band 幅のコントラスト依存性を定性的に示せた。提案モデルでコントラスト依存性が生じる原因は、モデル式においてゼロ除算を避けるために導入したパラメータ ε^2 の副次的効果であることが分かった。高コントラスト時には ε^2 の値は無視できるほど十分に小さいが、低コントラスト時には ε^2 の値が無視できないためである。

3.5 Preferred speed のコントラスト依存性の再現

本節では、1.5 節で述べた「速度抽出フィルタ」という既存の観点からでは解釈の難しい現象である、MT 細胞が最大反応を示す速度のコントラスト依存性が提案モデルで再現できるかどうか調べる。

入力画像は 3.1 節でも用いたガウシアンランダムドット画像で、画像サイズは

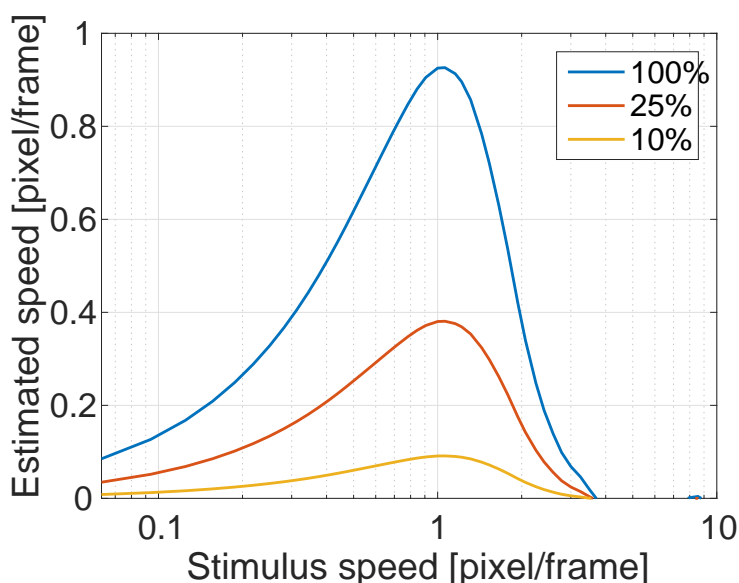


図 3.9 コントラストを変化させた時の、刺激速度に対するモデル出力（推定速度）。コントラストを低下させると推定速度が遅く見積もられるが、最大反応を示す刺激速度はコントラストに依存せず一定である。

128 × 128 pixel とし、50 組用意した。動き \mathbf{v} は x 軸方向（水平方向）に限定し、ピクセル値を循環シフトすることで実現した。また、サブピクセル（1 ピクセル以下）の動きは、入力画像を縮小することで実現した。なお、モデルパラメータは $\phi = 0^\circ$ 、 $\varepsilon^2 = 1.0 \times 10^{-5}$ 、 $k = 3$ 、 $\Gamma = 11$ とした。コントラストは元画像を 100% とし、100%（元画像）、25%（ピクセル値を 0.25 倍した画像）、10%（ピクセル値を 0.1 倍した画像）の 3 つとした。

図 3.9 に、コントラストを変化させた時の刺激速度 v_x に対する推定速度 \hat{v}_x の平均値を示す。この図から、推定速度 \hat{v}_x はコントラストが低下すると小さくなるが、どのコントラスト条件においても、最大出力をとる刺激速度は変化しないことが分かる。

残念ながら、提案モデルで MT 細胞が最大反応を示す速度のコントラスト依存性は再現できなかった。これは、提案モデルに改良の余地があることを示唆している。速度推

定可能な限界の速度を超える刺激を提示した場合、得られた速度推定結果は意味を持たない。従って、提示刺激の速度を正しく推定するためには、得られた速度推定結果が正しく推定できているか、それとも速度推定可能な限界の速度を超えているのかを判断する必要がある。安藤繁は LK 法による速度推定結果を自己評価する方法について提案している [19]。例えば、式 (23) において行列式 $\det \left(\begin{pmatrix} S_{\xi\xi}(\xi, \eta, t) & S_{\xi\eta}(\xi, \eta, t) \\ S_{\xi\eta}(\xi, \eta, t) & S_{\eta\eta}(\xi, \eta, t) \end{pmatrix} + \varepsilon^2 E \right)$ を考えると、刺激の SN 比が大きいと $\det()$ が大きく、SN 比が小さいと $\det()$ が小さくなる。また、窓問題が生じるパターンについても $\det()$ が小さくなる。他の指標としては、得られた推定速度を式 (10) に代入した値で自己評価が行える。式 (10) で定義されるエネルギー関数は、尤もらしい推定速度を代入したときに小さい値をとる。逆に、仮定 (a)~(c) が成立していない状況では大きな値をとる。上述の自己評価量は入力動画像の時空間微分から算出でき、リアルタイムに計算できる。提案モデルの改良については今後の課題とする。

3.6 MT 細胞集団からの read-out

本節では、MT 細胞集団の反応から提示刺激の速度を推定 (read-out) することを考える。既存モデルと提案モデルでは MT 細胞の計算原理や反応強度に対する解釈が異なるため、既存モデルとは異なる read-out モデルを考える必要がある。3.2 節で、個々の MT 細胞はそれぞれ異なる微分カーネルサイズで速度推定を行っているという解釈が得られた。すなわち、同じ空間位置に対して複数の速度推定結果が得られてしまうことを意味する。個々の MT 細胞はそれぞれができる限り刺激速度を正しく推定しようと試みているので、本論文では推定速度の平均値を read-out 結果 (MT 細胞集団の反応から予測される入力刺激の速度) とした。Read-out 結果である推定速度の平均値

\bar{v} は次式で求まる.

$$\bar{v} = \frac{1}{N} \sum_{i=1}^N \hat{v}(k_i) \quad (43)$$

式 (43) を, MT 細胞の正規化された反応強度である式 (28) を用いて書き直すと次のように書ける.

$$\bar{v} = \frac{1}{N} \sum_{i=1}^N L_i \cdot \text{MT}^{\text{norm}}(k_i) \quad (44)$$

但し, $L_i = \max_v \hat{v}(k_i)$ である. 式 (44) は, 個々の MT 細胞を速度推定器と見なした場合に得られる read-out モデルである.

一方で, 既存の read-out モデルとして重心モデルがある. これは, MT 細胞を速度抽出フィルタと見なし, どのフィルタをどの程度通過したかで提示刺激の速度を求めるモデルである. Preferred speed を PS_i , その MT 細胞の正規化された反応強度を $\text{MT}_i^{\text{norm}}$ とした時の重心 (read-out 結果である推定速度) \hat{v}_p は次式で求まる.

$$\hat{v}_p = \frac{\sum_i PS_i \cdot \text{MT}_i^{\text{norm}}}{\sum_i \text{MT}_i^{\text{norm}}} \quad (45)$$

また, Boyraz と Treue によって重心モデルを改良した read-out モデルが提案されている [20]. Boyraz & Treue の read-out モデルは, 重心モデルの分母 $\sum_i \text{MT}_i^{\text{norm}}$ を定数に置き換えることで, 重心モデルでは説明できなかった, 知覚される速度が提示刺激のサイズに依存して変化する現象を説明できる [20].

本論文で提案する read-out モデルと既存の read-out モデルである重心モデルや Boyraz & Treue の read-out モデルは, MT 細胞の反応特性について異なる解釈を基に構築されている. しかし, これらのモデルは同一の形式で書けることが分かった. 各項目を比較した表を表 3.1 に示す. ある MT 細胞の相対的な反応度合いを $\text{MT}_i^{\text{norm}}$, その MT 細胞に対応する重みを L_i とすると read-out 結果 \hat{v}_p は次式で表せる.

$$\hat{v}_p = \frac{\sum_i L_i \cdot \text{MT}_i^{\text{norm}}}{\alpha} \quad (46)$$

$\alpha = N$ かつ $L_i = \max_v \hat{v}(k_i)$ の場合は本論文で提案した read-out モデル (式 (46)), $\alpha = \text{const.}$ かつ $L_i = PS_i$ の場合には Boyraz & Treue のモデル, $\alpha = \sum_i MT_i^{\text{norm}}$ かつ $L_i = PS_i$ の場合には重心モデルである.

Boyraz と Treue は, 重心モデルでは説明できない現象が α を定数に置き換えることで説明できるとしているが, α を定数に置き換える計算論的意義については言及していない. 提案モデルの α は read-out に用いる MT 細胞の個数 N であり, 提示刺激に依存しない定数である. すなわち, 提案モデルの観点から Boyraz と Treue の改良について考察すれば, α を定数とするその計算論的意義は平均値計算に置き換えていると言える.

3.7 むすび

本章では, 第2章で提案した MT 細胞モデルの妥当性を示すため, 様々な電気生理学実験結果の再現を試みた. また, MT 細胞集団の反応から提示刺激の速度を求める read-out モデルを提案し, 既存の read-out モデルとの比較を行った.

これまでの, MT 細胞を「速度抽出フィルタ」と見なしてモデルの構築および現象

表 3.1 Read-out モデルの比較.

	提案モデル (平均値)	重心モデル	Boyraz & Treue
モデル式	提示刺激の推定された速度 $\hat{v} = \frac{\sum_i L_i \times MT_i^{\text{norm}}}{\alpha}$		
MT 細胞の反応強度 MT_i^{norm}	正規化された推定速度 $\frac{\hat{v}(k_i)}{\max_v \hat{v}(k_i)}$	速度抽出フィルタを通過した強度	
重み L_i	推定可能な限界の速度 $\max_v \hat{v}(k_i)$	Preferred speed PS_i	
分母 α	細胞数 N	$\sum_i MT_i^{\text{norm}}$	定数 const.

の考察が行われてきた。しかし本章のシミュレーションで、MT細胞を「速度推定器」と見なす新しい視点が得られた。既存モデルと提案モデルの違いを表3.2に示す。MT細胞の計算原理を速度検出器と見なす既存モデルでは、MT細胞の発火頻度は事前定義されたモデルパラメータ（preferred speed）との相関である。すなわち、発火頻度が高い場合には preferred speed に近い刺激が提示されていることを意味し、発火頻度が低い場合には preferred speed とは異なる刺激が提示されていることを意味する。また、最大反応を示す速度はそのMT細胞が検出すべき速度であり、最大反応を示す刺激方向はそのMT細胞が検出すべき刺激方向である。一方、MT細胞の計算原理を速度推定器と見なす提案モデルでは、MT細胞の発火頻度は推定速度に比例する値であると言え、発火頻度が高い場合には提示刺激の速度が速く発火頻度が低い場合には提示刺激の速度が遅いまたは速すぎることを意味する。また、最大反応を示す速度は正しく速度推定を行える上限の速度であり、最大反応を示す刺激方向はそのMT細胞が速度推定を試みる方向であると言える。

既存モデルと提案モデルではMT細胞の反応強度（平均発火頻度）に対する解釈が異なるため、既存モデルとは異なる read-out モデルを考える必要がある。しかし、推定速度の平均値を計算する read-out モデル（式44）が偶然にも Boyraz & Treue の

表3.2 既存のMT細胞モデル（SHモデル）と提案モデルの比較。

モデル	既存モデル（SHモデル）	提案モデル
計算原理	速度検出器	速度推定器
発火頻度の解釈	検出すべき速度との相関	推定速度に比例した値
最大反応を示す速度	検出すべき速度	速度推定可能な上限の速度
最大反応を示す方向	検出すべき刺激方向	速度推定を試みる方向

read-out モデルと同一の形式となった。従って、Boyraz & Treue の read-out モデルで再現・説明できる現象は、本論文で提案する read-out モデルでも再現・説明できる可能性がある。具体的には、提示刺激の大きさに応じて知覚される速度が異なるという現象を、Boyraz & Treue の read-out モデルでは説明できる。また、重心モデルでも Boyraz & Treue モデルでも再現・説明できる現象についても、提案モデルで再現・説明できる可能性がある。実際に提案モデルでこれらの現象が再現できるかどうかのシミュレーションは今後の課題とする。

既存モデルの計算原理である速度抽出器の場合には、どの速度抽出器がどの程度反応したか、すなわち細胞集団における個々の MT 細胞の発火頻度から、提示刺激の速度をリアルタイムに推定できる。一方提案モデルの計算原理である速度推定器の場合には、発火頻度と提示刺激の速度との関係を経験的に記憶しておく必要がある。従って、TMS や薬品などを用いて記憶や経験に関する情報に影響を与えることで、既存モデルと提案モデルのどちらの計算原理が正しいか検証できる可能性がある。

Band 幅のコントラスト依存性など既存の視点からでは解釈が難しい現象についても提案モデルで再現でき、提案モデルの視点から再考察すれば、演繹的に生じる当然の結果であるという新しい解釈が得られた。しかし、最大反応を示す速度のコントラスト依存性など再現出来ない現象もあることが分かった。これらについては数理的に妥当な処理を加え、提案モデルを改良する必要がある。

第 4 章

速度知覚特性の説明とモデル予測 の検証

前章では、Lucas-Kanade 法を計算基盤とする MT 細胞モデルの出力が多くの電気生理実験結果と定性的に一致することを示し、MT 細胞集団の出力から提示刺激の速度を求める read-out モデルを提案した。

MT 細胞モデルを導出する上で、「MT 細胞は Lucas-Kanade 法を計算基盤としている」という前提を置いている。本章ではこの前提の妥当性を示すために、運動錯視に着目して検証を行った。運動錯視とは物体の動き知覚に関する錯覚であり、物理的な動きと知覚される動きとが乖離している状況と見なすことができる。運動錯視を誘発するパターンを入力として与えた場合、視覚機能を精度良く記述した数理モデルであればヒトと同様に錯視を誘発するはずである。本章では、運動錯視を誘発するパターンに対し式 (43) によるオプティカルフロー推定を行い、ヒトが知覚する運動錯視と類似した推定結果が得られることを示す。

また、未知の入力に対する現象の予測も、数理モデルに求められる大事な要件の 1

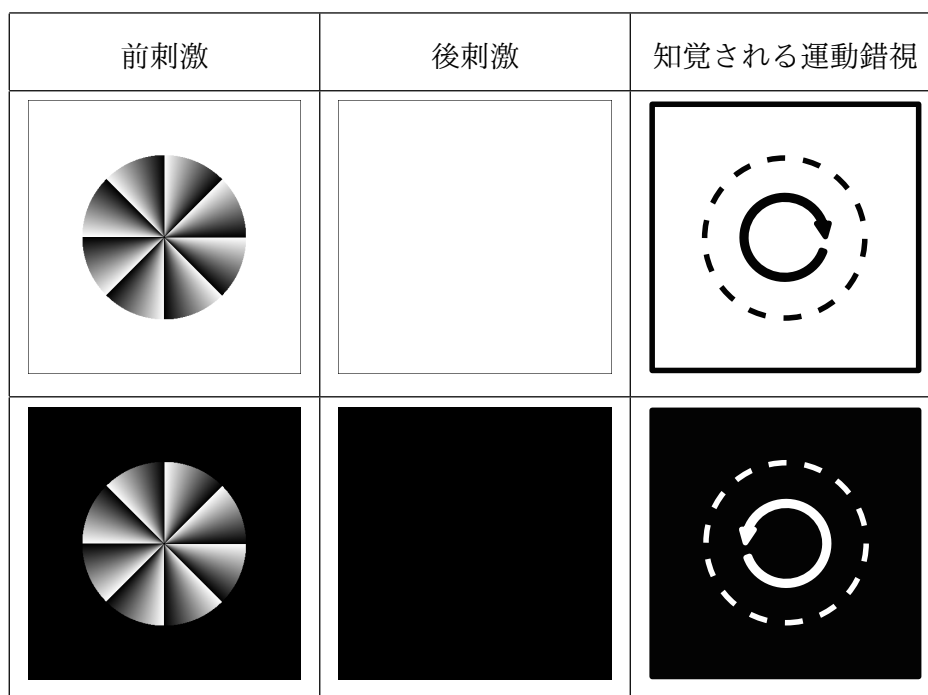


図 4.1 “Drift illusion” の例 [22]. 前刺激から後刺激へ画面を切り替えた際に、背景色が白（上段）の場合は時計回りの回転，背景色が黒（下段）の場合は反時計回りの回転が知覚される。

つである。本章では、運動錯視を生じさせる Fraser-Willcox 刺激（図 4.1 左列）をベースに、考え得る全パターン（ $8^8 = 16,777,216$ パターン）の入力画像に対し、運動錯視を生じさせるか否かのモデル予測を計算した。更にその中の一部の画像に対し実際にヒトが錯視を誘発するか否か、心理物理実験による検証を行った。

4.1 Drift illusion の背景輝度依存性の再現

図 4.1 左列に示す Fraser-Willcox (FW) 刺激を周辺視で観察すると、時折回転して見える [21]。またこの刺激は、単一色で塗られた背景のみが提示されている状態から FW 刺激を提示した場合にも回転錯視を誘発し、同様に FW 刺激を消失させた時にも回転錯視が生じる [23]。簡単のため、FW 刺激の消失時に生じる回転錯視を “drift

illusion”と呼ぶことにする。Drift illusion の回転方向は、図 4.1 に示すように、背景輝度に依存して変化することが報告されている [22]。また、林らは網膜細胞の時間応答特性を記述するモデルで、この錯視が説明できると報告している [22]。

本節では、drift illusion の背景輝度依存性が LK 法でも説明できることを示す。

4.1.1 既存モデル（林らのモデル）による説明

網膜細胞は単純に光刺激を電気刺激に変換して伝達しているだけではなく、様々な処理を行った上で大脳新皮質へ情報を伝達する [24]。網膜から一次視覚野までの伝達経路は、大細胞系（Magnocellular pathway）と小細胞系（Parvocellular pathway）に大別できる [25]。大細胞系に属する細胞は刺激が提示された時および消失した時に一過性の反応を示し、時間解像度は高い。また、反応までの時間遅れも小さい。一方、小細胞系に属する細胞は刺激が提示されている間反応し続け、高い空間解像度がある。しかし、反応までの時間遅れは大きい。

大細胞系に属する細胞は、輝度値が増加したときに反応する ON 型細胞と、輝度値が減少したときに反応する OFF 型細胞とがある [24]。林らは、この ON/OFF 型細胞の反応強度が等しくなる空間位置を追跡することで、drift illusion の回転方向が求まると報告している [22]。しかし、ON/OFF 型細胞の存在は確認されているものの、ON 型細胞と OFF 型細胞の出力差を求める細胞や、出力差がゼロとなる空間的位置を追跡する細胞の存在は今のところ確認されていない。また、そのような細胞を仮定する計算論的意味については述べられていない。

4.1.2 提案モデルによるシミュレーション

本節では, LK 法でも drift illusion の背景輝度依存性が説明できることを示す. 入力画像のサイズは 500×500 pixel とし, FW 刺激の直径は 300 pixel とした. 入力画像はグレースケールで, 相対輝度値 0.0 を最暗 (黒), 相対輝度値 1.0 を最明 (白) とした. モデルパラメータは, 空間微分カーネルが $k_i \in \{5, 9, 17, 33\}$ の 4 種類, $\varepsilon^2 = 1.0 \times 10^{-4}$, $\Gamma = 11$ とした.

図 4.2 に, 式 (43) を用いたオプティカルフロー推定結果を示す. オプティカルフローはピクセル毎に求まるが, 見やすいようダウンサンプリングして可視化した. 図 4.2 上段は背景輝度が 1.0 (最明) の状況であり, 時計回りの回転ベクトルが得られていることが分かる. 一方, 図 4.2 下段は背景輝度が 0.0 (最暗) の状況であり, 反時計回りの回転ベクトルが得られることが分かる. 図 4.1 と図 4.2 を比較すると, 数理モデルによって得られた回転ベクトルはヒトが知覚する運動錯視の回転方向と一致していることが分かる.

得られた回転ベクトルをより定量的に評価するため, ベクトル解析で用いられる 2 次元回転量 rot_{2D} の平均値 \bar{R} を求めた.

$$\begin{aligned} \bar{R} &= \frac{1}{|S|} \iint_S \text{rot}_{2D} \bar{v}(x, y, t) dS \\ &= \frac{1}{|S|} \iint_S \frac{\partial \bar{v}_y(x, y, t)}{\partial x} - \frac{\partial \bar{v}_x(x, y, t)}{\partial y} dS \end{aligned} \quad (47)$$

但し, 上式の S は円環刺激の内側領域であり, $|S|$ はその面積である. $\bar{R} > 0$ は反時計回りを意味し, $\bar{R} < 0$ は時計回りを意味する. 図 4.3 は, 背景輝度に対する平均回転量 \bar{R} をプロットした図である. 最大輝度 $I = 1.0$ で最も小さい \bar{R} (最も強い時計回りの回転) が得られ, 最小輝度 $I = 0.0$ で最も大きい \bar{R} (最も強い反時計回りの回転) が

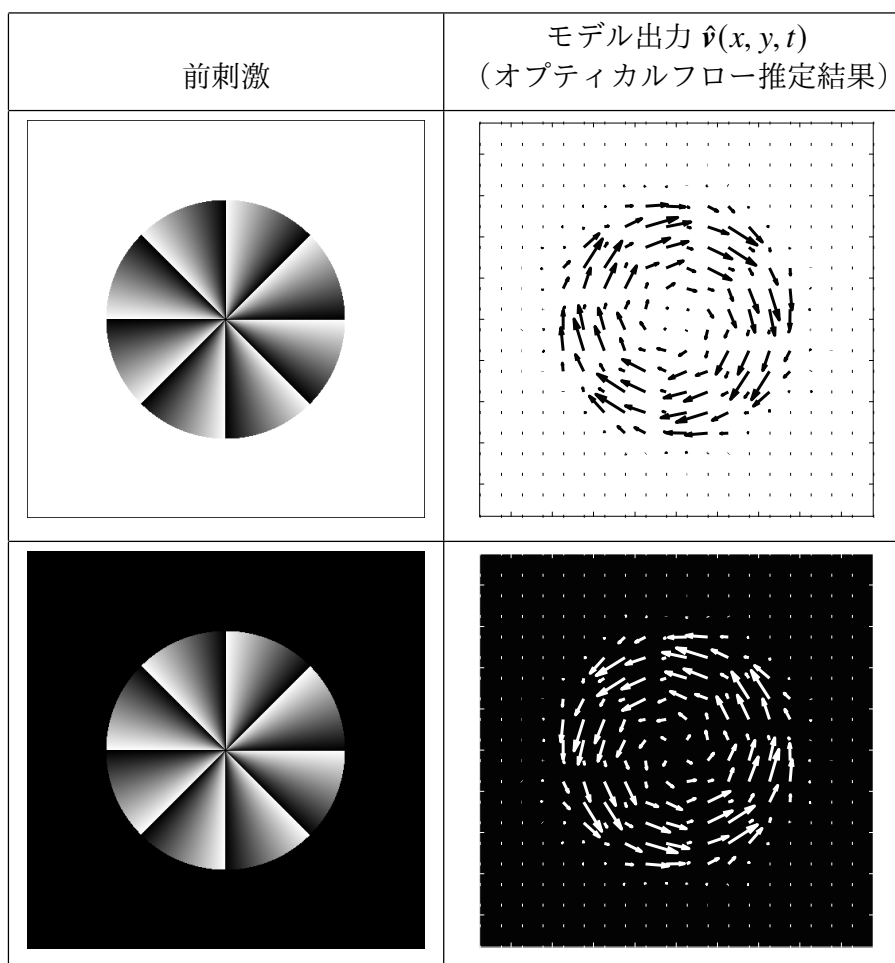


図 4.2 式 (43) による read-out モデルから得られた動きベクトル (オプティカルフロー推定結果, 動き知覚予測). 背景が白 (上段) の場合には時計回りの回転ベクトルが, 背景が黒 (下段) の場合には反時計回りの回転ベクトルが出力された.

得られた. また, 中間輝度 $I = 0.5$ で $\bar{R} = 0$ となった.

Drift illusion は FW 刺激が突然消失するパターンであり, モデル式を導出する上で置いた仮定 (a) 「輝度値の時間変化は物体の動きによってのみ生じる」を満たさない. 従って, 得られた速度推定結果 (オプティカルフロー推定結果) は工学的には無意味な値である. しかし興味深いことに, ヒトの知覚と一致するという結果が得られた.

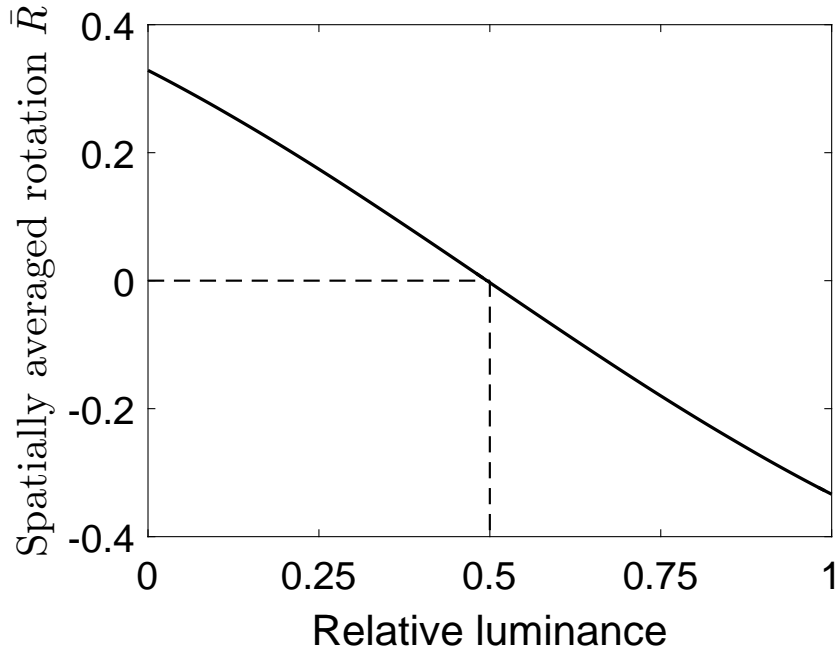


図 4.3 背景輝度に対する平均回転量 \bar{R} . 背景が最暗 ($I = 0.0$) で \bar{R} が最大 (最も強い時計回りの回転), 最明 ($I = 1.0$) で \bar{R} が最小 (最も強い反時計回りの回転) となる. また, 中間輝度 ($I = 0.5$) で回転ベクトルはゼロとなる.

4.1.3 背景輝度依存性の要因

式 (43) を解析した結果, 時間微分 $\partial I / \partial t$ の項が, drift illusion の背景輝度依存性に影響を与えていることが分かった. 簡単のため, 図 4.4 左図に示すように, 円環刺激の中心を原点とする極座標系 (r, θ) を考える. 図 4.4 右図は偏角 θ に対する相対輝度値 $I(\theta)$ である. 但しシミュレーションに用いた FW 刺激は 45° 間隔で周期的なパターンのため, $\theta = 0 \sim 45^\circ$ にのみ着目をした. 図 4.4 のような状況 ($\phi = \theta$) では, $\partial I / \partial \xi = \partial I / \partial r = 0$ であり, 式 (23) は次のように書き表せる.

$$\begin{pmatrix} \hat{v}_r(r, \theta, t) \\ \hat{v}_\theta(r, \theta, t) \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{S_{\theta t}(r, \theta, t)}{S_{\theta\theta}(r, \theta, t) + \varepsilon^2} \end{pmatrix} \quad (48)$$

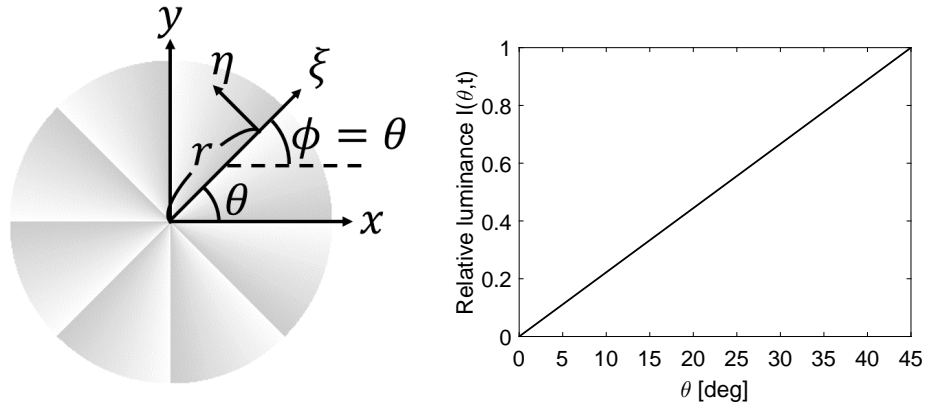


図 4.4 極座標系 (r, θ) および局所座標系 (ξ, η) での FW 刺激の表現. 左図は (x, y) 座標系と (ξ, η) 座標系の対応関係を示す図. 右図は極座標系 (r, θ) で表現した時の, 偏角 θ に対する FW 刺激の相対輝度値 $I(\theta, t)$. FW 刺激は相対輝度値 I が中心からの距離 r に依存しない. 本実験で用いた FW 刺激は $\theta = 45^\circ$ 周期のパターンのため, $\theta = 0 \sim 45^\circ$ に飲み着目をした.

更に, 窓関数をディラックのデルタ関数とし ($w(x, y) = \delta(x, y)$), ゼロ除算を避けるパラメータ ε^2 をゼロとすると, 式 (48) は次のように書き表せる.

$$\begin{pmatrix} \hat{v}_r(r, \theta, t) \\ \hat{v}_\theta(r, \theta, t) \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{\partial I(r, \theta, t)}{\partial t} / \left\{ \frac{1}{r} \frac{\partial I(r, \theta, t)}{\partial \theta} \right\} \end{pmatrix} \quad (49)$$

更に推定角速度 $\hat{\omega}(r, \theta, t)$ を考えると,

$$\begin{aligned} \hat{\omega}(r, \theta, t) &= \frac{1}{r} \hat{v}_\theta(r, \theta, t) \\ &= -\frac{\partial I(r, \theta, t)}{\partial t} / \frac{\partial I(r, \theta, t)}{\partial \theta} \end{aligned} \quad (50)$$

である. 式 (50) から, 推定角速度 $\hat{\omega}(r, \theta, t)$ は原点からの距離 r に依存しないことが分かる. また, 式 (50) は 1 次元の Lucas-Kanade 法からも同様の式が得られる.

図 4.5 は, 背景輝度がそれぞれ $I = 0.0, I = 1.0$ の時の, 偏角 θ に対する推定角速度 $\hat{\omega}(\theta, t)$, 時間微分 $\partial I / \partial t$, および空間微分 $\partial I / \partial \theta$ を示している. また, 推定角速度の図における点線は, 推定角速度の偏角 θ における平均値を表している. 左列と右列を

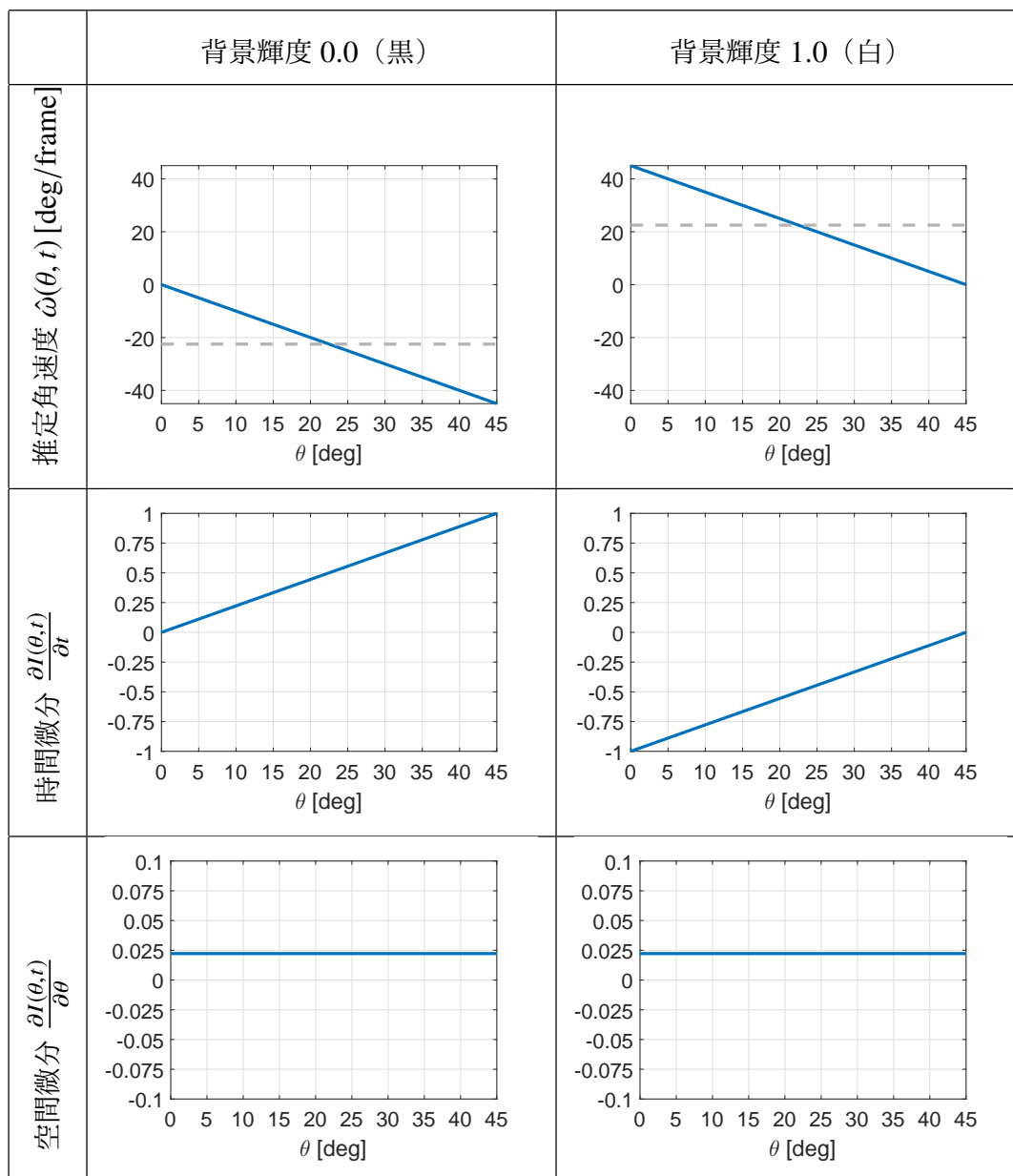


図 4.5 Drift illusion の背景輝度依存性の要因. 左列は背景輝度 $I = 0.0$, 右列は背景輝度 $I = 1.0$ の場合であり, 上段から順に推定された角速度 $\hat{\omega}(\theta, t)$, 時間微分 $\partial I(\theta, t)/\partial t$, 空間微分 $\partial I(\theta, t)/\partial \theta$ である. また, 上段の図中の点線は推定角速度の平均値を表している. この図から, 背景輝度を変えたときに推定角速度に影響を与えているのは時間微分 $\partial I(\theta, t)/\partial t$ の項であることが分かる.

比較すると、空間微分 $\partial I/\partial \theta$ は同じであり、時間微分 $\partial I/\partial t$ の項が推定角速度 $\dot{\omega}(\theta, t)$ に影響を与えていることが分かる。

今回用いた LK 法でも既存モデルである林らのモデルでも、輝度値の時間変化（時間微分）が drift illusion の背景輝度依存性に影響を与えているという結論は同じである。しかし、林らのモデルでは本論文執筆時には発見されていない仮想的な脳細胞を用いており、また、そのような脳細胞が存在する意義については述べられていない。一方、本論文で用いた LK 法は、第3章で示したように、実在する MT 細胞の反応特性を再現でき、2.3 節で示したように他の視覚系細胞を記述する既存の数理モデルを用いて構築できる。従って、提案モデルの方が神経生理学的にも妥当な数理モデルであると言える。

4.2 Stepping Feet illusion の再現

運動錯視の1つとして、図4.6に示す stepping feet illusion がある。提示刺激は、静止している白と黒の縦縞の背景上を、明るい灰色と暗い灰色の四角形が等速で同期して動いている刺激である（図4.6上段）。しかし、ヒトがこの刺激を観察すると、2つの動く四角形が非同期に、かつ早く動いたり遅く動いたりして見える（図4.6下段）[26]。本節では、提案モデルでこの現象が説明できるか数値シミュレーションを行う。

入力画像は、画像全体を 400×200 pixel、背景の縞の幅を 10 pixel、動く四角形の大きさを 40×20 pixel とした。相対輝度値は $0 \sim 1$ の範囲とし、背景の縞をそれぞれ 0.0 と 1.0、動く四角形の相対輝度をそれぞれ 0.8 と 0.2 にした。また、物体の動く速度 \mathbf{v} は、 $v_x = 1$ pixel/frame とした。モデルパラメータは $k = 3$, $\varepsilon^2 = 1.0 \times 10^{-6}$, $\Gamma = 5$ とした。

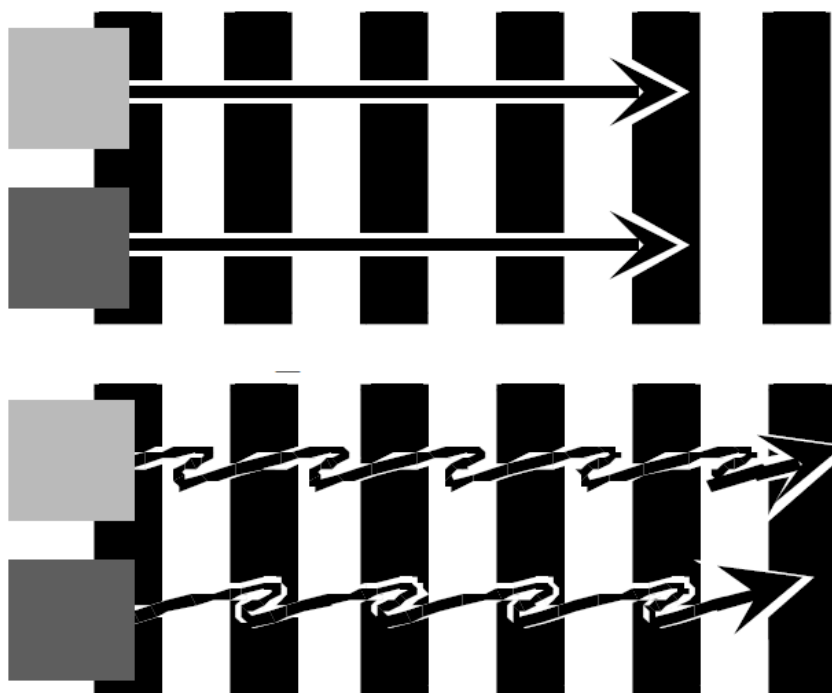


図 4.6 Stepping feet illusion [26]. 上図は物理的な動きを表している．静止している白黒の縞模様上を，明るい灰色の四角形および暗い灰色の四角形が等速度同期して動く．下図は，上図を観察したときに知覚される動きを表している．2つの四角形が等速で同期して動いているにもかかわらず，異なる速度で非同期的に動いて知覚される．

図 4.7 に，モデル式 (18) によるオプティカルフロー推定結果を示す．左図は，2つの動く四角形の右端が暗い縦縞上を通過している時と，明るい縦縞上を通過している時の推定速度を矢印プロットした図である．右図は，2つの動く四角形の右端中央部に着目し，その着目点の背景上の位置に対する x 軸方向の推定速度 \hat{v}_x をプロットした図である．この図から，暗い四角形が暗い背景の部分を通過している時は，明るい四角形が暗い背景の部分を通過している時よりも推定速度が遅く見積もられていることが分かる（図 4.7 左上）．一方で，明るい四角形が明るい背景の部分を通過している時

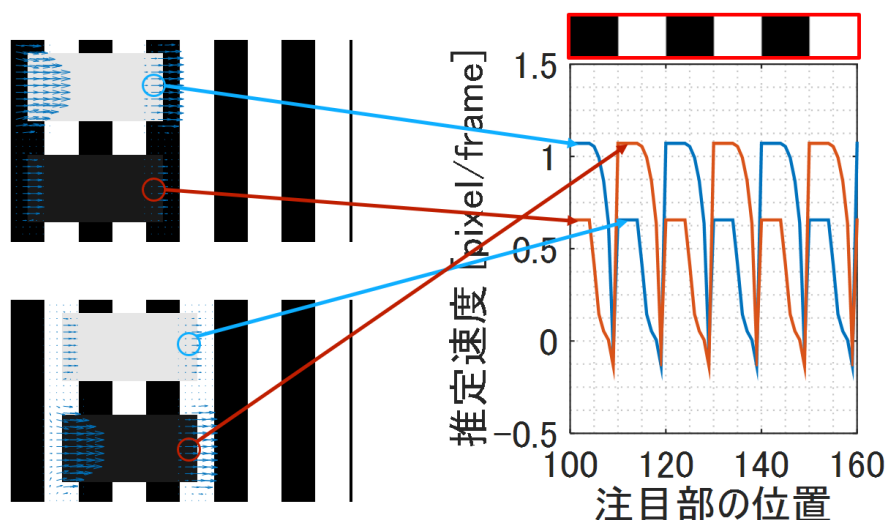


図 4.7 Stepping feet illusion に対するモデル出力。左側の図はどちらもあるフレームに対するオプティカルフロー推定結果であり，上図は移動する物体の右端が黒い背景を通過している時，下図は移動する物体の右端が白い背景を通過している時の図である。右図は，動いている四角形の右端中央部に着目し，着目部の位置に対する速度推定結果をプロットした図である。暗い四角形が暗い背景の部分を通過している時は，明るい四角形が暗い背景の部分を通過している時よりも推定速度が低く見積られる（図左上）。一方，明るい四角形が明るい背景の部分を通過している時は，暗い四角形が明るい背景の部分を通過している時よりも推定速度が低く見積られる（図左下）。

は，暗い四角形が明るい背景の部分を通過している時よりも推定速度が遅く見積もられていることが分かる（図 4.7 左下）。また，図 4.7 右図から，2 つの物体の推定速度が非同期に変化していることが分かる。

提案モデルで上述の結果が得られた原因を解析したところ，モデル式においてゼロ除算を避けるパラメータ ε^2 の副次的効果であることが分かった。簡単のため，窓関数

$w(x, y)$ をデルタ関数 $\delta(x, y)$ に置き換えると、着目点において $\partial I / \partial y = 0$ のため、

$$\hat{v}_x = -\frac{\frac{\partial I}{\partial x} \frac{\partial I}{\partial t}}{\left(\frac{\partial I}{\partial x}\right)^2 + \varepsilon^2} \quad (51)$$

と書ける。 $\partial I / \partial x$ が ε^2 に比べ十分に大きい場合には ε^2 が無視でき、

$$\begin{aligned} \hat{v}_x &= -\frac{\partial I}{\partial t} / \frac{\partial I}{\partial x} \\ &= v_x \end{aligned} \quad (52)$$

となる。一方で $\partial I / \partial x$ が ε^2 に比べてあまり大きく無い場合、

$$\frac{\frac{\partial I}{\partial x}}{\left(\frac{\partial I}{\partial x}\right)^2 + \varepsilon^2} \leq \frac{1}{\frac{\partial I}{\partial x}} \quad (53)$$

のため、 \hat{v}_x は本来の推定速度よりも小さい値をとる。すなわち、2つの動く四角形が暗い背景部分を通過している時、暗い四角形の空間微分 $\partial I / \partial x$ は明るい四角形の空間微分 $\partial I / \partial x$ よりも小さく、暗い四角形の推定速度の方が遅く見積もられる。一方、2つの四角形が明るい背景部分を通過している時は、明るい四角形の空間微分 $\partial I / \partial x$ は暗い四角形の空間微分 $\partial I / \partial x$ よりも小さく、明るい四角形の推定速度の方が遅く見積もられる。

4.3 Drift illusion を生じさせるパターンの網羅的探索

前節までで、様々な運動錯視を提案モデルで説明できることを示した。しかし、現象の再現・説明以外に、現象を予測できることも数理モデルの大事な要件の1つである。本節では、4.1節で紹介した drift illusion を生じさせるパターンの網羅的探索を行う。具体的には、考え得る全パターン ($8^8 = 16,777,216$ パターン) に対し、それぞれ白一色の背景画像に切り替えた時にどの程度錯視を誘発するかのモデル予測を求めた。

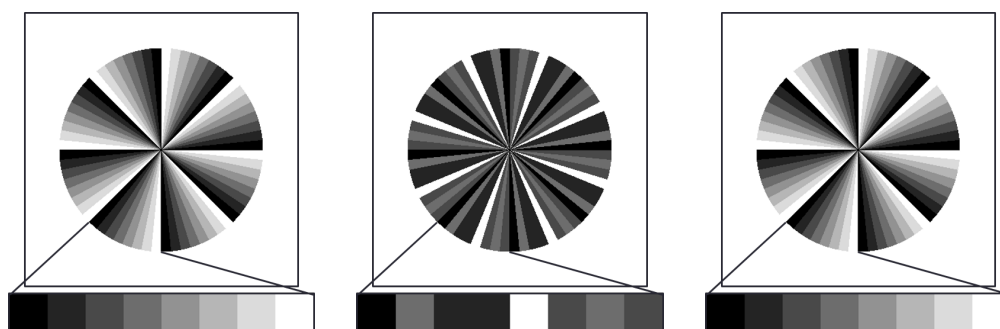


図 4.8 円環刺激の構成例. 45° 周期で 1 セクター辺り 8 色に塗り分けた. 相対輝度値 I は $I \in \{0/7, 1/7, 2/7, \dots, 7/7\}$ の 8 階調. 考え得る全パターン数は $8^8 = 16,777,216$ パターン.

モデル予測の検証として, 一部の画像に対し心理物理実験を行い, ヒトの知覚と一致するかどうかを調べた.

4.3.1 モデル予測の網羅的探索

まず初めに, 考え得る全パターンに対し網羅的にモデル予測を計算した. 簡単のため, 背景輝度は白 ($I = 1.0$) のみとした. 図 4.8 に前刺激の構成例を示す. 前刺激は 500×500 pixel のグレースケールで, 円環状パターンの直径は 300 pixel とした. 円環状パターンは 45° 周期で, 1 セクター辺り 8 色で塗り分けた. また, 相対輝度値 I は $I \in \{0/7, 1/7, 2/7, \dots, 7/7\}$ の 8 階調とした. 従って, 状態として取り得る総パターン数は $8^8 = 16,777,216$ パターンである. ヒトに対し心理物理実験で 8^8 パターンを網羅的に探索するのは現実的に不可能である. なぜなら, 1 パターンあたり 2 秒で回答できたとしても, 不眠不休で 400 日もかかってしまうからである. 一方数理モデルであれば, 1 パターン辺り 20 ミリ秒程度で求まり, 4 日もあれば十分である.

全 8^8 パターンに対する式 (47) による平均回転量 \bar{R} のヒストグラムを図 4.9 に示す. 但しシミュレーション時間の関係上, 微分カーネルサイズは $k = 5$ のみとした. また,

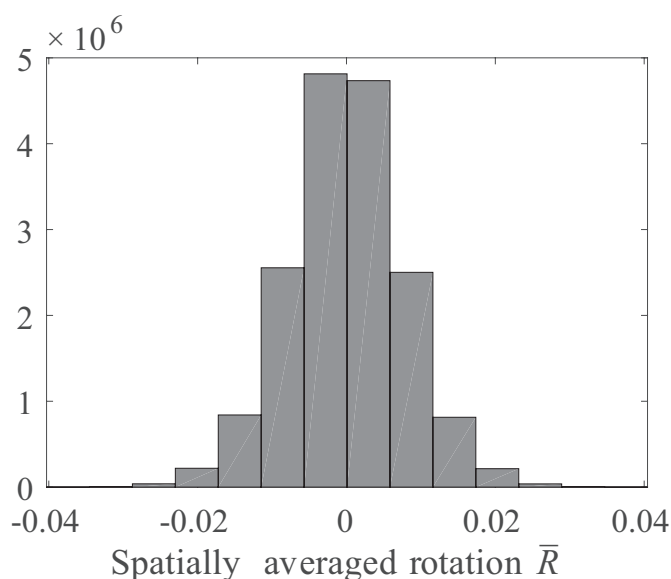


図 4.9 全 16,777,216 パターンに対する平均回転量 \bar{R} のヒストグラム．多くのパターンが $\bar{R} \approx 0$ （錯視を誘発しない）と予測されたが，一部のパターンは $\bar{R} \neq 0$ （時計回りまたは反時計回りの回転錯視を誘発する）と予測された．

他のモデルパラメータは $\varepsilon^2 = 1.0 \times 10^{-4}$, $\Gamma = 11$ とした．他のカーネルサイズや複数のカーネルサイズでの推定結果については 4.3.4 節で考察する．図 4.9 から，多くのパターンは $\bar{R} \approx 0$ であり，錯視を誘発しないと予測された．しかし一部のパターンは $\bar{R} \neq 0$ であり，時計回りまたは反時計回りの回転錯視を誘発すると予測された．

4.3.2 心理物理実験によるヒトの知覚の調査

モデル予測の検証を行うため，実際にヒトが錯視を誘発するか心理物理実験で調査した．実験に用いた前刺激は 8^8 パターンの中から，最明 ($I = 1.0$) および最暗 ($I = 0.0$) を含み，平均回転量 \bar{R} に偏りがないようランダムに選ばれた 33 パターンとした．また，後刺激は白一色 ($I = 1.0$) の画像とした．実際に実験に用いた前刺激を図 4.10 に示す．各枠の数値は上から順に，刺激番号，モデル予測として得られた平均



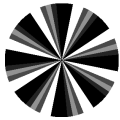
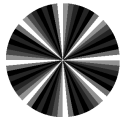
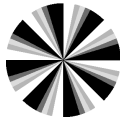


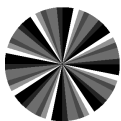









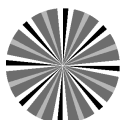


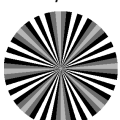
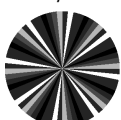









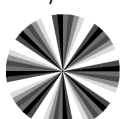
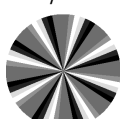

Ex -0.0189 50/50 		#1 -0.0079 49/50 	#2 -0.0036 49/50 	#3 -0.0075 49/50 	#4 -0.0044 48/50 	#5 -0.0081 47/50 
#6 -0.0100 45/50 	#7 -0.0096 44/50 	#8 -0.0035 37/50 	#9 -0.0019 36/50 	#10 -0.0033 35/50 	#11 -0.0015 33/50 	#12 -0.0035 32/50 
#13 -0.0200 31/50 	#14 -0.0015 26/50 	#15 0.0029 23/50 	#16 0.0002 23/50 	#17 0.0006 20/50 	#18 -0.0035 18/50 	#19 0.0006 17/50 
#20 0.0001 16/50 	#21 0.0046 16/50 	#22 0.0050 13/50 	#23 -0.0069 13/50 	#24 0.0070 8/50 	#25 0.0136 6/50 	#26 0.0030 6/50 
#27 0.0038 5/50 	#28 0.0071 5/50 	#29 0.0083 4/50 	#30 0.0113 3/50 	#31 -0.0005 2/50 	#32 -0.0003 2/50 	#33 0.0054 2/50 

図 4.10 心理物理実験に用いた刺激一覧（ヒトが時計回りに回転したと回答した割合でソート）。各セルは上から、刺激番号、平均回転量 \bar{R} 、ヒトが時計回りに回転したと回答した割合、刺激画像である。左上の Ex は、Fraser-Willcox 刺激を 8 階調で描画した場合の結果である。

回転量 \bar{R} , 心理物理実験で得られた被験者が時計回りに回転したと回答した割合を表している。

実験は刺激提示用 CRT モニタ (GDM-F520; Sony Corp.) 以外に光源のない暗室で行い, あご台を用いることで視距離を 1m に統一した。モニタの解像度は 1024×768 pixel で, 視野角にして $22.0 \times 16.6^\circ$ であった。また, リフレッシュレートは 85Hz に設定し, ガンマ補正を行った。提示刺激はモニタの中央部分に描画され, 円環刺激の直径は 13.0° (300 pixel) とした。最大輝度 (白, $I = 1.0$) は 81.3 cd/m^2 であった。被験者は 23 ~ 24 歳の成人男性 5 人で, 刺激消失時に知覚した回転方向 (時計回りまたは反時計回り) を図 4.11 に示す回転デバイス (PowerMate NA16029; Griffin Technology) で回答させた。なお, 本実験は電気通信大学倫理委員会の承認を得て行った。

実験は次の手順で行われた。まず初めに画面中央に注視点が 300ms 提示される。次に, 図 4.10 に示す 33 種の円環状パターンの中からランダムに 1 パターンが選ばれ, モニタ中央に提示される。円環状パターンは 1500 ms 提示された後に消失する。被験者は円環状刺激消失時に知覚した回転方向 (時計回りまたは反時計回り) をどちらか必ず回答する (two-alternative forced choice; 2AFC)。この流れを各パターン 10 回ずつ, 合計 330 回行った。

刺激番号に対する全被験者が時計回りと回答した割合 $P_H(CW)$ は, 図 4.10 の分数で表されている。回答は 2AFC であるため, 回答割合が 100% または 0% に近い場合は時計回り, 反時計回りの錯視を誘発するパターン, 50% に近い場合は錯視を誘発しないパターンであることを意味する。この図から, #1 や #33 といった刺激は回転錯視を誘発することが分かる。



図 4.11 心理物理実験で被験者回答を得るために用いた回転デバイス（Griffin Technology 製 PowerMate NA16029）。回答が「時計回り」「反時計回り」の二択のため、回転するデバイスを用いることで被験者が直感的に扱えるよう配慮した。

4.3.3 モデル予測とヒトの知覚との比較

次に、シミュレーションによって得られたモデル予測と心理物理実験によって得られたヒトの知覚との比較を行った。平均回転量 \bar{R} と被験者が時計回りに回転したと回答した割合 $P_H(CW)$ とを直接比較することはできないため、平均回転量 \bar{R} から時計回りに回転したと回答する割合 $P_M(CW)$ への変換を行った。但し、平均回転量がゼロ ($\bar{R} = 0$) の場合には時計回りに回転したと回答する割合は 50% ($P_M(CW) = 0.5$) で

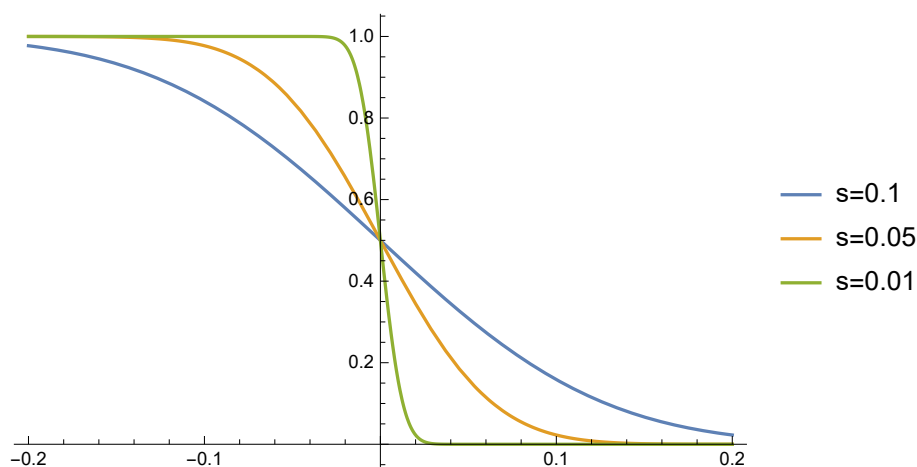


図 4.12 パラメータ s による変換式 (54) の振る舞いの変化. s を大きくするとなだらかな曲線, s を小さくすると急峻な曲線となる.

あるという仮定を置いた. 変換式を次式に示す.

$$P_M(CW) = \frac{1}{2} \left(1 - \operatorname{erf} \left(\frac{\bar{R}}{s\sqrt{2}} \right) \right) \quad (54)$$

$$\operatorname{erf} \left(\frac{\bar{R}}{s\sqrt{2}} \right) = 2 \int_0^{\bar{R}} \frac{1}{s\sqrt{2}} \exp \left(-\frac{t^2}{2s^2} \right) dt \quad (55)$$

ここで, s は正のパラメータである. 図 4.12 に示すように, 式 (54) においてパラメータ s を大きくするとなだらかな曲線に, s を小さくすると急峻な曲線になる. パラメータ s は, 図 4.10 に示す平均回転量 \bar{R} とヒトが時計回りに回転したと回答した割合 33 組から非線形フィットによって求めた. その時の最適パラメータは $s = 0.013$ であった.

各刺激におけるモデル予測 $P_M(CW)$ に対する心理実験結果 $P_H(CW)$ の散布図を図 4.13 に示す. 図中の白丸は 8 階調で描いた FW 刺激 (図 4.10 の Ex.) に対する割合である. 図中の r は Pearson の相関係数, p は無相関であるという帰無仮説に対する検定の p 値である. もし数理モデルが視覚系を完全に記述できていれば, 図 4.13 は全て

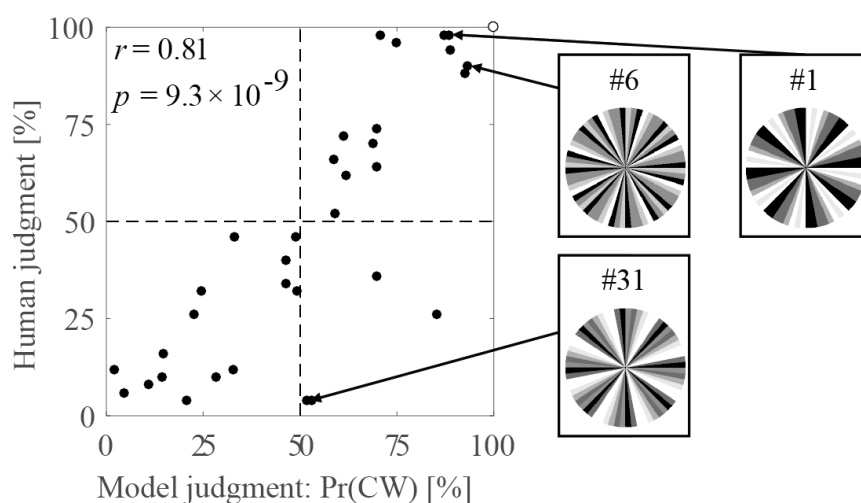


図 4.13 モデル式 (43) において微分カーネル数 $N = 1$, 微分カーネルサイズ $k = 5$ とした場合の, モデル予測とヒトの知覚との散布図. 白丸は Fraser-Willcox 刺激を 8 階調で表現した刺激 (図 4.10 の Ex) に対する結果. 図中の r は Pearson の相関係数, p は無相関を帰無仮説とした検定に対する p 値.

の点が対角線上にプロットされるはずである. 図 4.13 から完全に対角線上にはプロットされないが, 多くの点に対角線に近い位置にプロットされていることが分かる. モデル予測とヒトの知覚との相関係数を求めたところ, $r = 0.81$ という強い正の相関があることが分かった.

4.3.4 他のカーネルサイズでのモデル予測

前節まではシミュレーション時間削減のため, 微分カーネルサイズを $k = 5$ のみでモデル予測を計算した. 本節では他の微分カーネルサイズ $k \in \{9, 17, 33\}$ および, カーネル数 $N = 4$ の場合のモデル予測を計算し, ヒトの知覚との比較を行う. また, 入力画像のスケール (拡大率) f も, $f \in \{1/4, 1/2, 1/1\}$ の 3 種類を試した.

図 4.14 に, 様々な微分カーネルサイズで得たモデル予測とヒトの知覚との相関係数

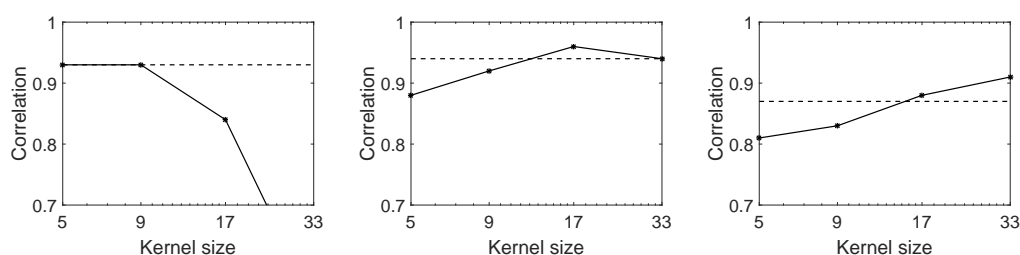


図 4.14 微分カーネルサイズに対する，モデル予測とヒトの知覚との相関係数 r 。実線は単一カーネル ($N = 1$)，点線は複数カーネル ($N = 4$ かつ $k_i \in \{5, 9, 17, 33\}$) での結果。各パネルは異なる入力画像の拡大係数に対する結果 (左: $f = 1/4$, 中央: $f = 1/2$, 右: $f = 1/1$)。

r を示す。図中の実線は単一カーネル ($N = 1$) でカーネルサイズ $k \in \{5, 9, 17, 33\}$ でのモデル予測，点線は複数カーネル ($N = 4$ かつ $k_i \in \{5, 9, 17, 33\}$) でのモデル予測の結果である。各パネルは左から順に，入力画像の拡大係数 $f = 1/4, 1/2, 1/1$ での結果である。この図から，用いる微分カーネルのサイズや入力画像のスケールに依存して相関係数が変化する事が分かる。最も相関係数が高いのは， $f = 1/2$, $N = 1$ かつ $k = 17$ で $r = 0.96$ であった。この時のモデル予測に対するヒトの知覚の散布図を図 4.15 に示す。どの点もほぼ対角線上にプロットされており，高い精度でヒトの知覚を予測できていることが分かる。 $f = 1/2$, $N = 1$ かつ $k = 17$ で 8^8 パターンに対する網羅的探索を再度行った結果が図 4.16 である。図 4.9 同様，多くのパターンが $\bar{R} \approx 0$ (回転錯視を誘発しない) と予測されたが，一部のパターンは $\bar{R} \neq 0$ (時計回りまたは反時計回りの回転錯視を誘発する) と予測された。図 4.17 に，時計回りの回転錯視を誘発するパターンを $|\bar{R}|$ が大きい順に 10 種示す。この図から，強い回転錯視を誘発するパターンは，どのパターンも FW 刺激のような滑らかな輝度変化を持っていることが分かる。しかし，FW 刺激を 8 階調表現したパターン (図 4.10 の Ex.) は 1270 番目であった。従って，FW 刺激のような直線的な輝度変化ではなく，例えば二次関数の

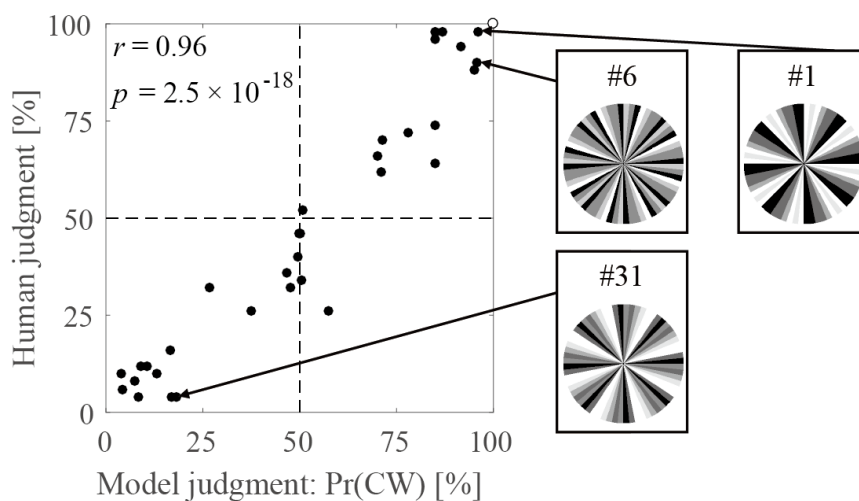


図 4.15 最も相関係数 r が高くなる条件 ($f = 1/2$ かつ $N = 1$ かつ $k = 17$) での、モデル予測に対するヒトの知覚の散布図。

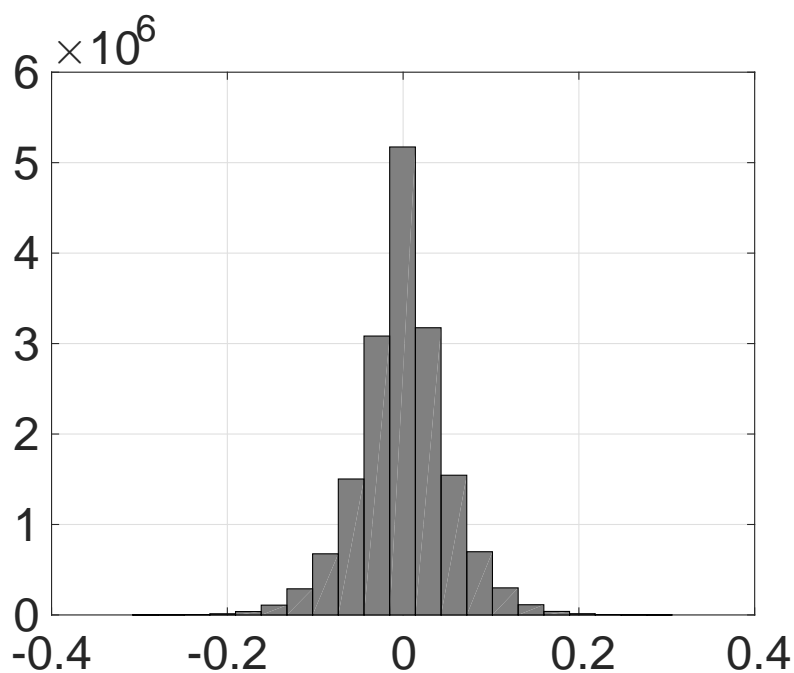


図 4.16 $f = 1/2$, $N = 1$ かつ $k = 17$ での全 16,777,216 パターンに対する平均回転量 \bar{R} の再シミュレーション結果。図 4.9 同様、多くのパターンが $\bar{R} \approx 0$ (錯視を誘発しない) と予測されたが、一部のパターンは $\bar{R} \neq 0$ (時計回りまたは反時計回りの回転錯視を誘発する) と予測された。

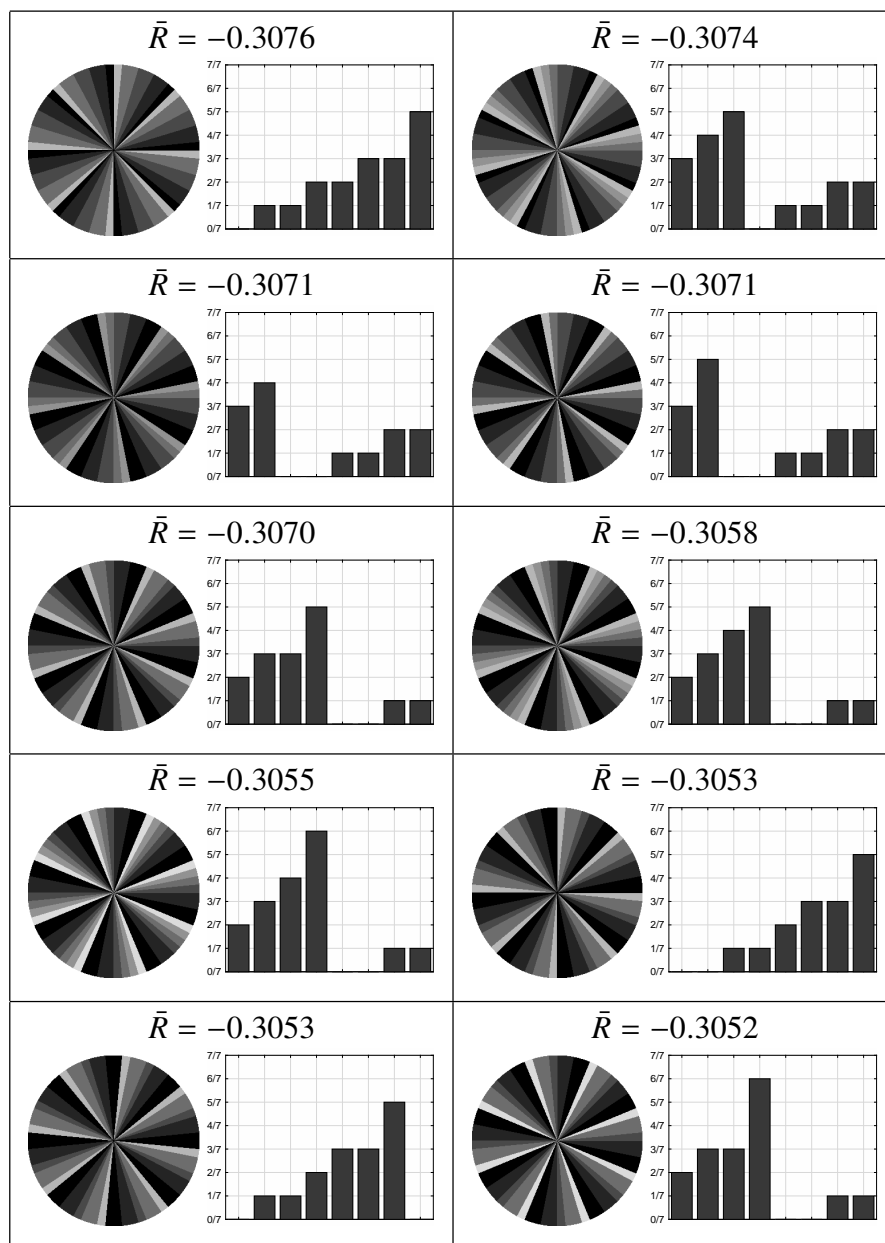


図 4.17 時計回りの回転錯視を誘発するパターンで $|\bar{R}|$ が大きい 10 種. 強い回転錯視を誘発するパターンは, どのパターンも FW 刺激のような明暗が滑らかな輝度変化を持っている.

ような、曲線的な輝度変化をするパターンの方がより強い錯視を引き起こす可能性がある。

今回のシミュレーションにおいて、ヒトの知覚を最も良く予測できる微分カーネルサイズは $k = 17$ であった。しかし、常に $k = 17$ の微分カーネルを用いれば良いわけではない。図 4.14 から分かるように、 $k = 17$ であっても $f = 1$ (右図) の場合には $r = 0.88$, $f = 1/4$ (左図) の場合には $r = 0.84$ となる。すなわち、最適な微分カーネルサイズは刺激サイズに依存することが分かる。一般に、速度を推定したい物体の大きさは未知なため、適切な解像度や微分カーネルサイズは分からない。複数の微分カーネルを用いれば平均的に良い結果は得られるが、単純な平均値を計算するだけでは良い結果は得られない (図 4.14 点線)。3.5 節で紹介した自己評価量など、最適な画像スケールや微分カーネルサイズでの推定結果を重視するようなアルゴリズムを加える必要がある。

4.4 むすび

本章では、read-out モデルの妥当性や「ヒトは Lucas-Kanade 法を計算基盤とする速度推定を行っている」という前提の尤もらしさを示すため、Lucas-Kande 法による運動錯視の説明を行った。また、モデルの汎化性能を評価するため、錯視を誘発するパターンか否かの選別を数理モデルによる数値シミュレーションで網羅的に行った。数値シミュレーションで得られたモデル予測と心理物理実験で得られたヒトの知覚とを比較し、提案モデルが精度良くヒトの知覚を予測できることを示した。

錯視を誘発するパターンの網羅的探索は、新しい錯視パターンを見つける方法としても有効な手法である。実際、図 4.10 の#1 や#33 は新しく発見された回転錯視を誘発

するパターンである。Drift illusion はモデル導出時に置いた仮定 (a) 「輝度値の時間変化は物体の動きによってのみ生じる」を満たさないパターンである。従って、他の仮定を満たさないパターンもまた、運動錯視を誘発する可能性がある。2.5 節にも記したが、2 枚の画像間で動きが大きすぎる場合やルーフエッジのような輝度値が 2 次関数的に変化する場合には「仮定 (b) : 輝度値の空間的な変化に対し Taylor の一次近似が成り立つ」が成り立たなくなる。また、動く物体が小さすぎる場合や局所的な動きが一定でないパターンは「仮定 (c) : 窓関数内で速度は一定である」が成り立たなくなる。このようなパターンに対し数理モデルを用いて網羅的に探索することで、全く新しい錯視を発見できる可能性がある。また、数理モデルから現象を予測することは、複数のモデル同士を比較する上でも大事なアプローチとなる。本論文で行った、錯視を誘発するパターンの網羅的な探索は、モデル同士の比較検討にも使える手法である。既存モデルと提案モデルとの予測性能の比較については今後の課題とする。

提案モデルの計算基盤である Lucas-Kanade 法は輝度勾配モデルとして分類でき、輝度勾配モデルで説明できる運動錯視は提案モデルでも再現できる可能性が高い。しかし、提案モデルで再現できない運動錯視もいくつかある。提案モデルは時間変化を前提としており、静止画にもかかわらず動いて見えるといった現象は今の提案モデルでは説明ができない。例えば、図 4.1 左列に示した Fraser-Willcox 刺激は周辺視で観察すると動いて見えることが報告されている [21]。また、図 4.18 に示す rotating snakes [27] を観察すると、視線移動やまばたきに伴って運動錯視が知覚される [28]。提案モデルにこれらの錯視パターンを入力として与えても、時間微分が常にゼロ ($dI/dt = 0$) のため推定速度は常にゼロ ($\hat{v} = \mathbf{0}$)、即ち静止し続けているという出力が得られてしまう。しかし、視線移動やまばたきといった時間変化を生じさせる眼球モデルを入力画

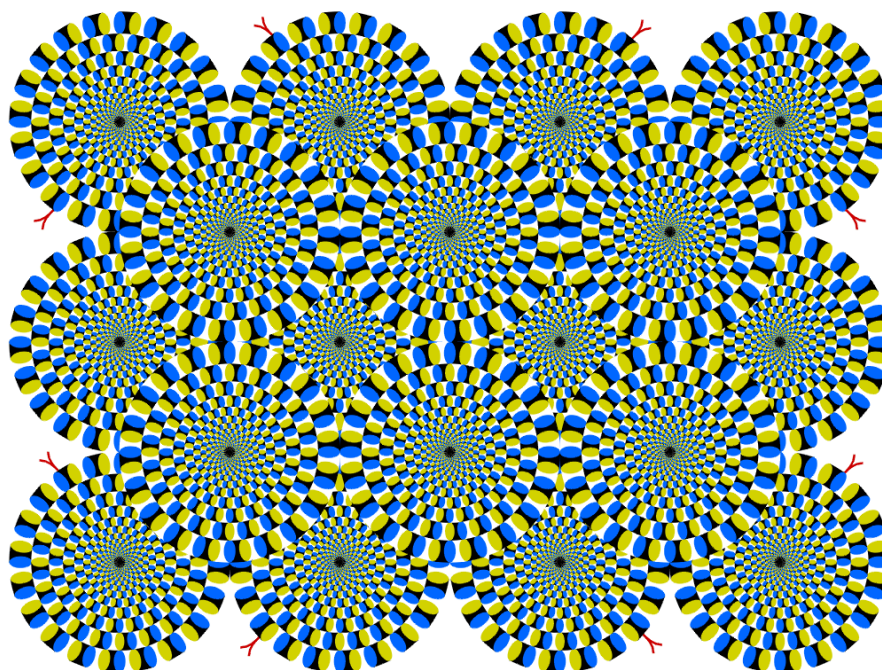


図 4.18 Rotating snakes [27]. この静止画を観察すると、視線移動やまばたき時に回転運動が知覚される [28].

像の前処理として導入することで、これら静止画像に対する運動錯視を説明できる可能性はある。モデルの改良および静止画に対する錯視の説明については今後の課題とする。

第 5 章

Grating 刺激に対する解析解

第 3 章では MT 細胞の様々な反応特性に対し，モデル出力が定性的に一致することをいくつか示した．本章では，式 (1) で定義される grating 刺激に対する 2 次元 LK 法の解析解を求め，MT 細胞の空間周波数依存性を再現できるか定量評価を行う．

5.1 MT 細胞の空間周波数依存性

図 5.1 は，grating 刺激に対する 3 種類の MT 細胞の反応特性を示した図である [29]．左列は最大反応を示す方向の空間周波数を f_s ，時間周波数を f_t とし，各時空間周波数に対する平均発火頻度を円の大きさで表している．時空間周波数 f_s, f_t と最大反応を示す刺激方向の速度 v_s には

$$v_s = -\frac{f_t}{f_s} \quad (56)$$

の関係があるため，横軸を提示刺激の速度として各空間周波数ごとにプロットし直した図が 5.1 右列である．この図から，上段の MT 細胞は空間周波数が増えなくても最大反応を示す刺激速度は変わらないが，中段及び下段の MT 細胞は提示刺激の空間周波

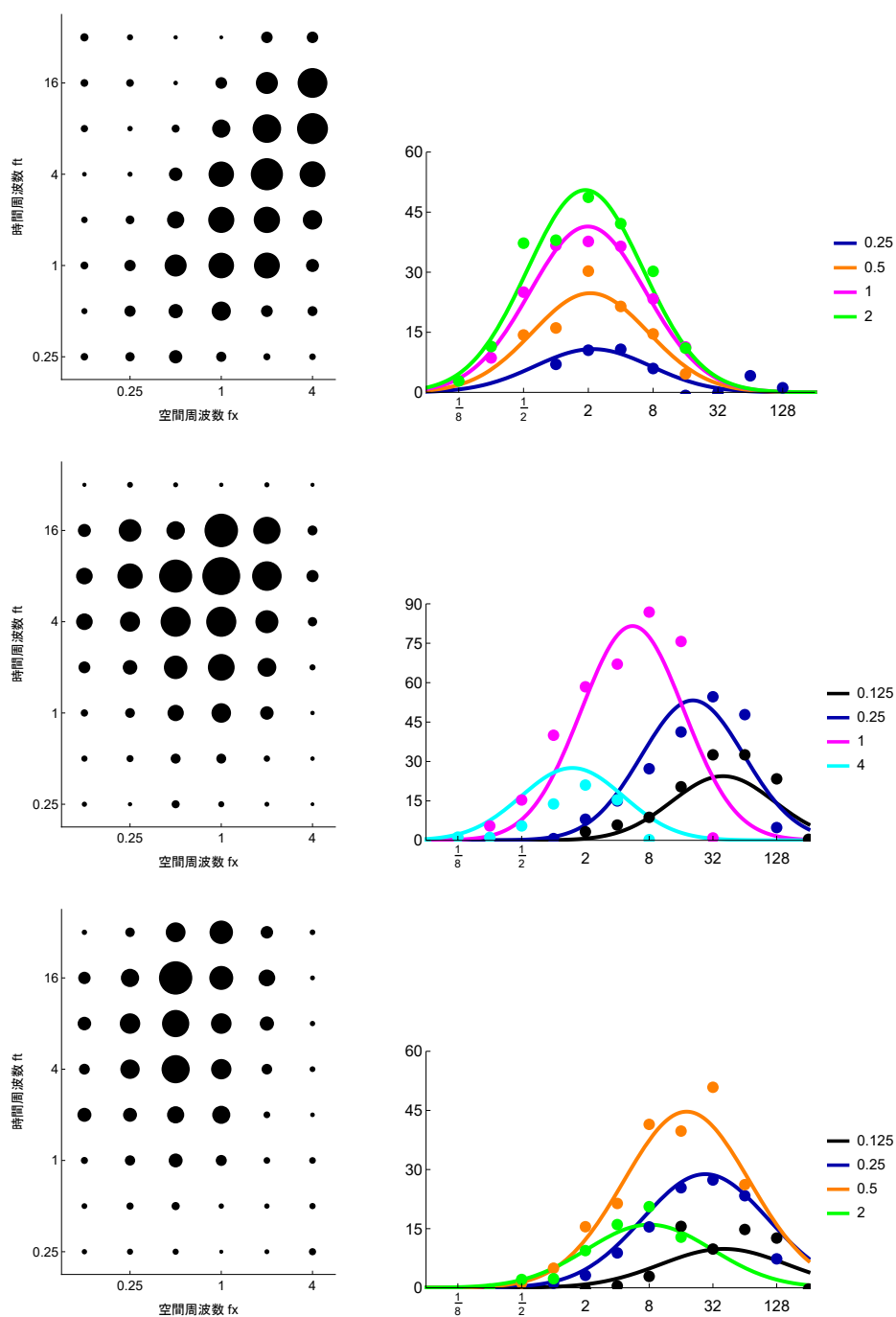


図 5.1 Grating 刺激に対する 3 種類の MT 細胞の反応例 [29]. 左列：最大反応を示す方向の空間周波数を f_s ，時間周波数を f_t とし，各時空間周波数に対する平均発火頻度を円の大きさで表している．右列：左列の図を空間周波数ごとに提示刺激の速度を横軸としてプロットし直した図．

数に依存して最大反応を示す刺激速度が変化していることがわかる。Priebe らによると、測定を行った 104 個の MT 細胞のうち 75 個の MT 細胞が、図 5.1 中段や下段の MT 細胞のように最大反応を示す速度が空間周波数に依存すると報告している [29].

既存モデルの計算原理である「速度検出フィルタ」にとっては、このように提示刺激の空間周波数に依存してフィルタ特性が変化するのは好ましくない。また、第 1 章で紹介した既存モデルである SH モデルの場合、 $v_s f_x + f_t = 0$ を満たす等速度ラインは斜めに傾いた直線となるので、どの MT 細胞も理想的には図 5.2 に示すような傾いた楕円状の時空間周波数受容野を持つはずである。SH モデルは図 5.1 中段や下段の MT 細胞のような受容野となるよう構築することも可能ではあるが、本来の計算原理である「速度検出フィルタ」とはかけ離れたモデルとなり、このような MT 細胞が存在する計算論的意義は分からなくなってしまう。

5.2 Grating 刺激に対する LK 法による推定速度の解析解

式 (19) で定義される 2 次元の Lucas-Kanade 法の時空間微分を式 (34) で定義される Gaussian derivative で近似し、窓関数をディラックのデルタ ($w(x, y) = \delta(x, y)$) とすると、式 (1) で定義される grating 刺激に対する推定速度の解析解は次式となる。

$$\begin{pmatrix} \hat{v}_x(x, y, t) \\ \hat{v}_y(x, y, t) \end{pmatrix} = \begin{pmatrix} -\frac{f_x f_t}{f_x^2 + f_y^2} \frac{1}{1 + \frac{2\pi\epsilon^2}{A^2(f_x^2 + f_y^2) \sin^2(2\pi f_x x + 2\pi f_y y + 2\pi f_t t)} e^{4\pi^2(f_x^2 \sigma_x^2 + f_y^2 \sigma_y^2 + f_t^2 \sigma_t^2)}} \\ -\frac{f_y f_t}{f_x^2 + f_y^2} \frac{1}{1 + \frac{2\pi\epsilon^2}{A^2(f_x^2 + f_y^2) \sin^2(2\pi f_x x + 2\pi f_y y + 2\pi f_t t)} e^{4\pi^2(f_x^2 \sigma_x^2 + f_y^2 \sigma_y^2 + f_t^2 \sigma_t^2)}} \end{pmatrix} \quad (57)$$

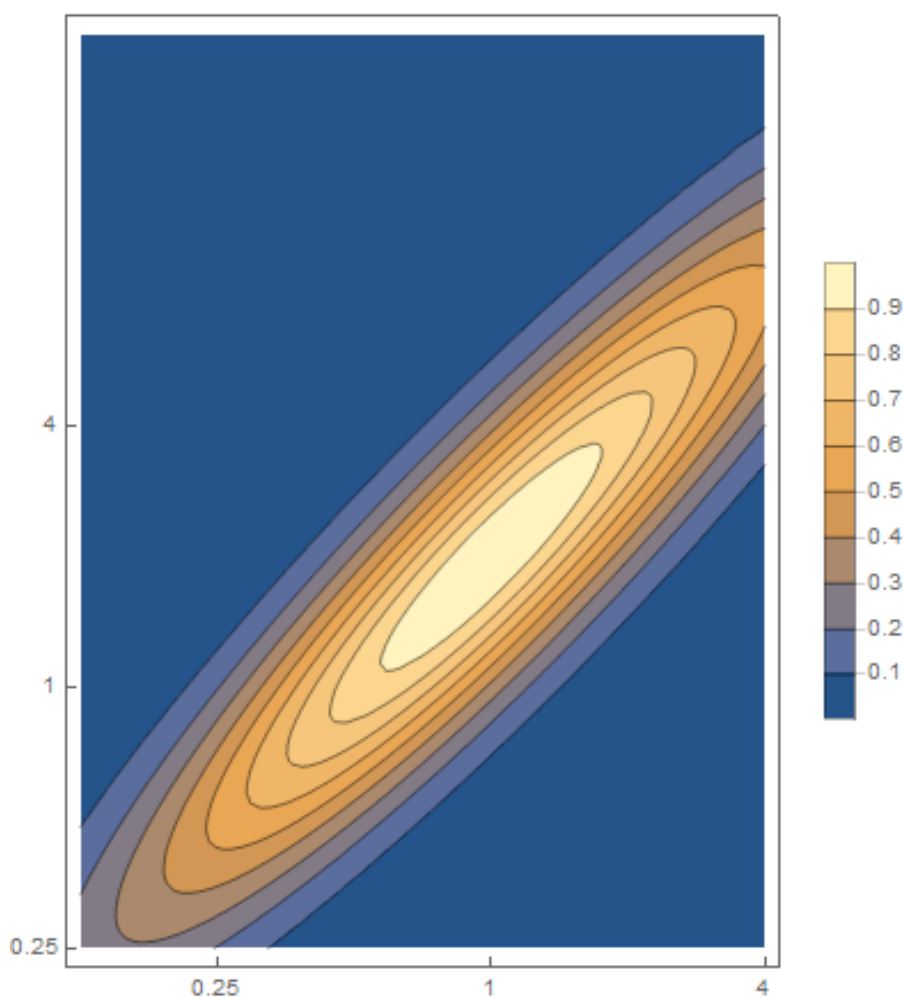


図 5.2 Simoncelli & Heeger モデルにおける速度 2deg/s を検出するための理想的な時空間周波数フィルタ．速度 2deg/s を検出するためには， $2f_s + f_t = 0$ となる受容野を構築すれば良い．

Grating 刺激の原点 $(x, y) = (0, 0)$ に着目し，位置 $(x, y) = (0, 0)$ における x 軸方向の平均推定速度 \bar{v}_x を求めると，

$$\begin{aligned} \bar{v}_x &= \int_{-\frac{T}{2}}^{\frac{T}{2}} \hat{v}_x(0, 0, t) dt \\ &= -\frac{f_x f_t}{f_x^2 + f_y^2} \left(1 - \sqrt{\frac{1}{1 + \frac{A^2(f_x^2 + f_y^2)}{2\pi e^2} e^{-4\pi f_x^2 \sigma_x^2} e^{-4\pi f_y^2 \sigma_y^2} e^{-4\pi f_t^2 \sigma_t^2}}} \right) \end{aligned} \quad (58)$$

である。Grating 刺激は循環シフトするパターンであるため、任意の位置 (x, y) に対して平均推定速度 \bar{v} は上式で求まる。本論文では MT 細胞の平均発火頻度は推定速度に比例すると仮定しており、 0° 方向に最大反応を示す MT 細胞の平均発火頻度 $\overline{\text{MT}}_{0^\circ}$ は比例定数を c として

$$\begin{aligned} \overline{\text{MT}}_{0^\circ}(f_x, f_y, f_t; A, \sigma_x, \sigma_y, \sigma_t, \varepsilon, c) &= c \cdot \bar{v}_x(f_x, f_y, f_t; A, \sigma_x, \sigma_y, \sigma_t, \varepsilon) \\ &= -c \frac{f_x f_t}{(f_x^2 + f_y^2)} \left(1 - \sqrt{\frac{\frac{2\pi\varepsilon^2}{A^2(f_x^2 + f_y^2)} e^{4\pi f_x^2 \sigma_x^2} e^{4\pi f_y^2 \sigma_y^2} e^{4\pi f_t^2 \sigma_t^2}}{1 + \frac{2\pi\varepsilon^2}{A^2(f_x^2 + f_y^2)} e^{4\pi f_x^2 \sigma_x^2} e^{4\pi f_y^2 \sigma_y^2} e^{4\pi f_t^2 \sigma_t^2}}} \right) \end{aligned} \quad (59)$$

と表せる。

5.3 生理データに対するフィッティング

前節で解析的に得られた式 (59) を定量評価するため、図 5.1 左列のデータに対して非線形モデルフィッティングを行った。但し、図中の空間周波数 f_s は最大反応を示す方向に対する空間周波数であるため $f_x = f_s, f_y = 0$ とし、grating 刺激の振幅（コントラスト） A は $A = 1$ とした。 $f_x = f_s, f_y = 0, A = 1$ として式 (59) を整理すると次式となる。

$$\overline{\text{MT}}_{0^\circ}(f_x, f_t; A, \sigma_x, \sigma_t, \varepsilon, c) = -c \frac{f_t}{f_x} \left(1 - \sqrt{\frac{\frac{2\pi\varepsilon^2}{A^2 f_x^2} e^{4\pi f_x^2 \sigma_x^2} e^{4\pi f_t^2 \sigma_t^2}}{1 + \frac{2\pi\varepsilon^2}{A^2 f_x^2} e^{4\pi f_x^2 \sigma_x^2} e^{4\pi f_t^2 \sigma_t^2}}} \right) \quad (60)$$

図 5.1 左列のデータに対して式 (60) で非線形モデルフィッティングを行い、図 5.1 右列に対応する形式でプロットした結果を図 5.3 に示す。また、表 5.1 に適合度および最適パラメータを示す。図中の各点は各空間周波数における MT 細胞の平均発火頻度を表し、各線は最適パラメータにおける各空間周波数でのモデル出力を表している。この図から、図中段や下段のような反応特性が空間周波数に依存する MT 細胞の生理データは再現できることが分かる。また、表 5.1 の R^2 値が 1 に近いことから定量的

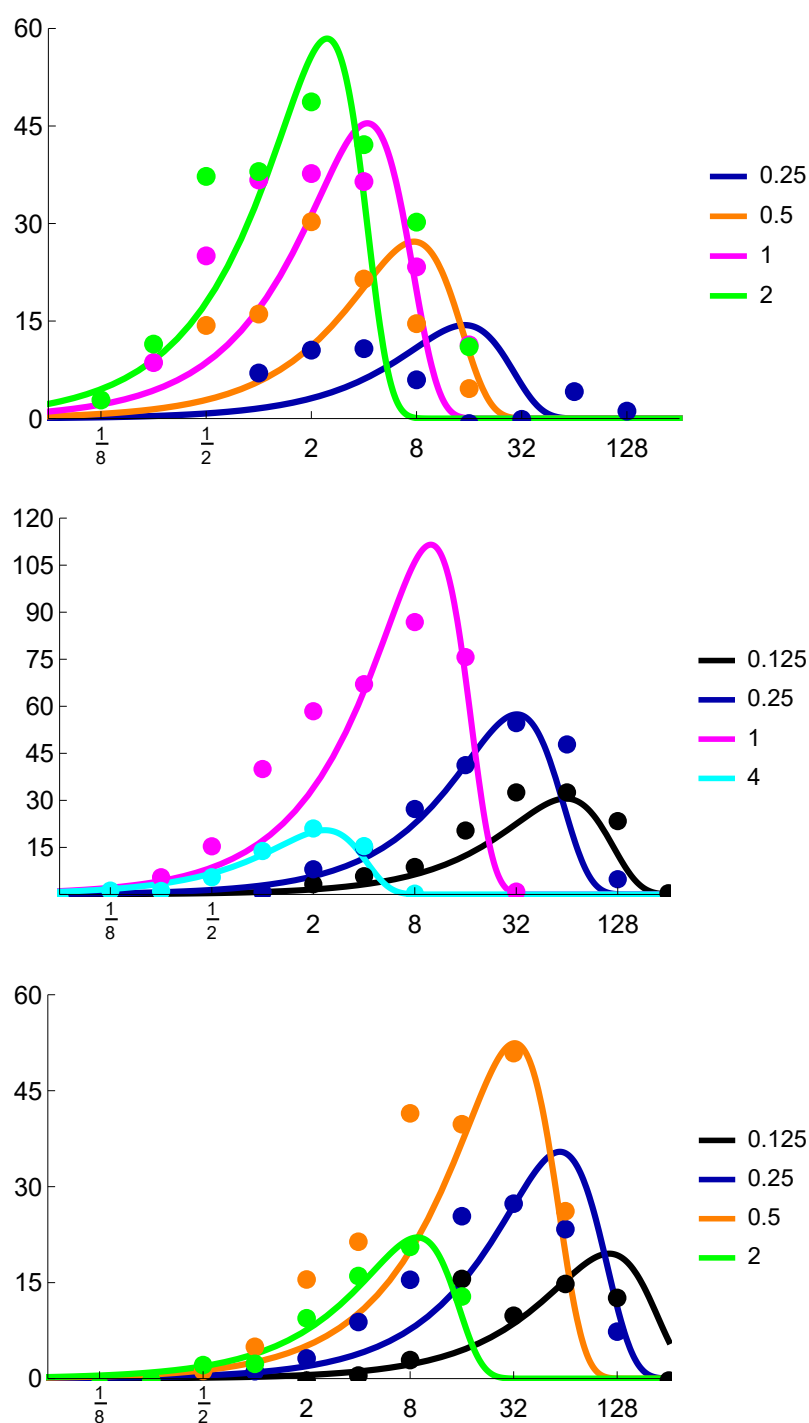


図 5.3 式 (60) を用いて生理データをフィッティングした結果。各点は各空間周波数における生理データの値，各曲線はフィッティング結果。中段，下段のような空間周波数に依存するパターンは良くフィットしているが，上段のような空間周波数に依存しない MT 細胞の反応特性は再現できない。

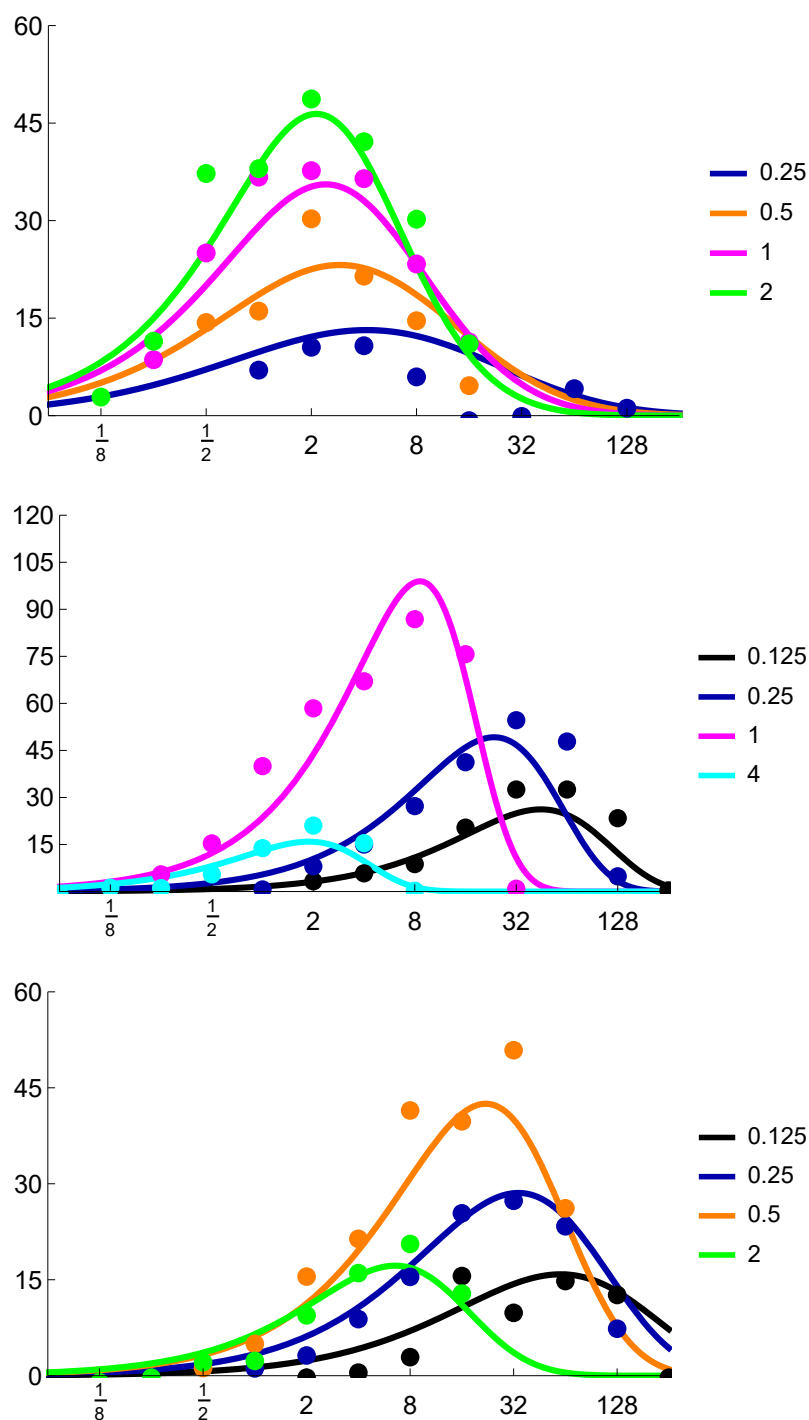


図 5.4 式 (61) を用いて再度生理データをフィッティングした結果. 式 (60) では再現できなかった上段のような空間周波数に依存しない MT 細胞の反応特性が再現できるようになった.

に再現できていることがわかる。しかし，上図のような空間周波数に依存しない MT 細胞の反応特性は再現できていない。上図の MT 細胞の反応特性を再現できない理由としては，モデルパラメータ σ_x, ε を変化させても反応特性が空間周波数 f_x 軸で並行移動するだけであり，モデルパラメータ σ_t を変化させても反応特性が時間周波数 f_t 軸で並行移動するだけだからである。すなわち，解析的に求めた式 (59) または式 (60) による MT 細胞モデルでは，空間周波数に依存しない MT 細胞の反応特性を原理的に再現できない。

天下りの的ではあるが，式 (60) の時間周波数 f_t に関する項を改良した次式で再度モデルフィッティングを行った。

$$\overline{\text{MT}}'_0(f_x, f_t; A, \sigma_x, \sigma_t, \varepsilon, c, n) = -c \frac{f_t}{f_x} \left(1 - \sqrt{\frac{\frac{2\pi\varepsilon^2}{A^2 f_x^2} e^{4\pi f_x^2 \sigma_x^2} e^{4\pi |f_t \sigma_t|^n}}{1 + \frac{2\pi\varepsilon^2}{A^2 f_x^2} e^{4\pi f_x^2 \sigma_x^2} e^{4\pi |f_t \sigma_t|^n}}} \right) \quad (61)$$

表 5.1 式 (60) でモデルフィッティングを行った際の適合度および最適パラメータ。

	R^2	AIC	σ_x	σ_t	ε	c
MT 細胞 A	0.734	378.4	0.0092	0.030	1.20	74.92
MT 細胞 B	0.935	359.7	0.074	0.014	0.61	39.15
MT 細胞 C	0.867	327.7	0.117	0.0079	0.507	9.66

表 5.2 式 (61) で再度モデルフィッティングを行った際の適合度および最適パラメータ。

	R^2	AIC	σ_x	σ_t	ε	c	n
MT 細胞 A	0.949	301.5	0.015	0.00016	0.12	86.19	0.26
MT 細胞 B	0.962	336.3	0.076	0.0039	0.45	46.37	0.94
MT 細胞 C	0.945	286.7	0.120	0.00047	0.234	13.14	0.535

式 (61) は高時間周波数に対する抑制効果 $e^{4\pi f_t^2 \sigma_t^2}$ を $e^{4\pi |f_t \sigma_t|^n}$ に置き換え、新しくモデルパラメータ n を導入したものである。式 (61) での非線形モデルフィッティングの結果を図 5.4 に示す。また、表 5.2 に適合度および最適パラメータを示す。図 5.3 と図 5.4 を比較すると、式 (60) でも再現できていた中段および下段の MT 細胞の反応特性は問題なく再現できており、式 (60) では再現できていなかった上段の MT 細胞の反応特性も式 (61) では再現できていることが分かる。また、表 5.1 と表 5.2 を比較すると、どの MT 細胞の反応特性に関しても R^2 値が向上しており、改良した式 (61) の方が生理データを良く再現していることが分かる。また、AIC に関しても全ての生理データに対して値が小さくなっているため、モデルパラメータ n を追加した式 (61) の方が良い細胞モデルであることが分かる。

5.4 考察

式 (59) における時空間周波数に関する項 $e^{-4\pi f_x^2 \sigma_x^2}$, $e^{-4\pi f_y^2 \sigma_y^2}$, $e^{-4\pi f_t^2 \sigma_t^2}$ は、時空間微分を Gaussian derivative で近似した結果生じた項である。従って、他の関数による微分近似を行えば、異なる解析解が得られるはずである。特に表 5.2 の n の値はどれも 1 以下の値をとっており、MT 細胞が入力情報として利用する V1 細胞モデルに Gaussian derivative よりも良いモデルがある可能性を示唆している。Gauss 関数のように単峰性の曲線となる関数は t 分布やガンマ分布などいくつかある。式 (61) において $n \approx 1$ や $n \approx 0.5$ となる解析解を生じさせる微分近似関数の調査については今後の課題とする。

MT 細胞の生理データに対するフィッティング結果 (σ_x, σ_t) から、対象とする MT 細胞が情報を受け取っている V1 細胞の受容野サイズが推定できる。表 5.2 から σ_x は 0.015 0.12 の範囲をとっており、この値が V1 細胞の生理データを Gaussian derivative

モデルでフィッティングした時のパラメータ分布とかけ離れていないか調べることで、本論文で提案する MT 細胞モデルの妥当性を評価できる。同様に、 σ_i についても評価が行える。V1 細胞の生理データに対する Gaussian derivative モデルでのフィッティングや、モデルパラメータ分布との比較は今後の課題とする。

第 6 章

結言

本論文では、既存の単一細胞を対象とした疎粒度モデルの計算基盤となっている「速度抽出フィルタ」に対して疑問を呈し、「速度抽出フィルタ」に代わる概念として「個々の MT 細胞はそれぞれが速度推定を行っている」という考えのもと、MT 細胞モデルを構築した。MT 細胞細胞の基本的な電気生理学実験結果を再現しただけでなく、既存の概念では解釈の難しい現象に対しても新しい解釈を与えた。また、速度知覚特性についても提案モデルで再現・説明し、構築した数理モデルの妥当性を示した。更に、数値シミュレーションによるモデル予測を行い、提案モデルの汎化性能を評価した。

MT 細胞は「速度抽出フィルタ」ではなく「速度推定器」という新しい視点から現象を考察することで、これまで複雑だと思われていた現象の多くが演繹的に生じる当然の結果であるということが分かった。しかし、提案モデルで再現できない現象もいくつかあり、提案モデルの改良が必要である。モデル改良や精度の向上は今後の課題とする。

本論文で提案した新しい MT 細胞モデルは既存の MT 細胞モデルを否定するもの

ではなく，MT 細胞の速度選択性に対してこれまでとは異なる解釈を提案するものである．提案モデルの良さは，入力画像の時間微分および空間微分から簡単に現象を考察でき，任意の動画像が扱えるという点が挙げられる．既存の MT 細胞モデルである Simoncelli & Heeger モデルでは入力画像の時空間周波数を求め，周波数空間で考察する必要がある．しかし，錯視パターンは *grating* 刺激のような周波数空間で考察しやすいパターンではない場合が多いため，提案モデルでの考察の方が容易であると言える．実際，第4章では提案モデルで運動錯視についての考察を行った．

参考文献

- [1] , . Visual Field Maps in Human Cortex. *Neuron*, pp. 366–383, 2007.
- [2] year. 第3版 カールソン神経科学テキスト－脳と行動－. 丸善株式会社, 2010.
- [3] year. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, pp. 1–47, 1991.
- [4] year. Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, pp. 574–591, 1959.
- [5] year. Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, pp. 3313–24, 2004.
- [6] year. Shape representation in area V4: position-specific tuning for boundary conformation. *Journal of neurophysiology*, pp. 2505–2519, 2001.
- [7] year, year, year. Heading Tuning in Macaque Area V6. *Journal of Neuroscience*, pp. 16303–16314, 2015.
- [8] year. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology* pp. 1127–1147, 1983.

-
- [9] year. A model of neuronal responses in visual area MT. *Vision research*, pp. 743–761, 1998.
- [10] year, year. Spatiotemporal tuning in mouse primary visual cortex. *Neuroscience Letters*, pp. 165–169, 2012.
- [11] year, year. Interactions between speed and contrast tuning in the middle temporal area: implications for the neural code for speed. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, pp. 8988–98, 2006.
- [12] year. The role of V1 surround suppression in MT motion integration. *Journal of Neurophysiology*, 2010.
- [13] year. Melting the iceberg: contrast invariance in visual cortex. *Neuron*, pp. 11–3, 2007.
- [14] year. An Iterative Image Registration Technique with an Application to Stereo Vision, 1981.
- [15] year, year. Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception*, pp. 105–26, 1985.
- [16] year. Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America*, p. 1297, 1980.
- [17] year, year. The Gaussian Derivative model for spatial-temporal vision: I. Cortical model. *Spatial Vision*, pp. 261–319, 2001.
- [18] year. 新編 画像解析ハンドブック. 東京大学出版会, 2004.
- [19] year. 画像の時空間微分算法を用いた速度ベクトル分布計測システム. 計測自動制

- 御学会論文集, pp. 1330–1336, 1986.
- [20] year. Misperceptions of speed are accounted for by the responses of neurons in macaque cortical area MT. *J Neurophysiol* pp. 1199–1211, 2011.
- [21] year. Perception of illusory movement. *Nature*, pp. 565–566, 1979.
- [22] year, year. A Computational Model of Afterimage Rotation in the Peripheral Drift Illusion Based on Retinal ON/OFF Responses. *PLoS ONE* p. e115464, 2014.
- [23] year, year. Illusory motion induced by the offset of stationary luminance-defined gradients. *Vision Research*, pp. 970–978, 2006.
- [24] year. 第3版 カールソン神経科学テキストー脳と行動ー. 丸善株式会社, 2010.
- [25] year. 視覚科学. 勁草書房, 2010.
- [26] year. Moving objects appear to slow down at low contrasts. *Neural Networks*, pp. 933–938, 2003.
- [27] year. Phenomenal Characteristics of the Peripheral Drift Illusion. *Vision*, pp. 261–262, 2003.
- [28] year, year. Microsaccades and blinks trigger illusory rotation in the "rotating snakes" illusion. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, pp. 6043–51, 2012.
- [29] year, year. The neural representation of speed in macaque area MT/V5. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, pp. 5650–61, 2003.

謝辞

本研究を進めるにあたり，研究に対する考え方や様々なアイデア，基礎的なツールである数学や物理学の定義・定理の深い解釈など，多岐にわたるご指導を賜りました佐藤俊治准教授に深く感謝をいたします。また，全体ゼミなど様々な場面での確かな鋭い質問を投げかけて下さった阪口豊教授にも深く感謝をいたします。中嶋豊特任助教には心理物理実験を行う上で多くのご助言を承りましたことを感謝いたします。最後に，人間情報学講座の皆様に感謝いたします。

研究業績一覧

学術雑誌に発表した論文

- (1) 占部一輝, 中村大樹, 佐藤俊治, 韓雪花, “脳の数理モデル構築を目的とした RT ミドルウェアの応用と結果”, 計測自動制御学会論文集, vol. 52, no. 5, pp.264-275, 2016.
- (2) Daiki Nakamura, Shunji Satoh, “Simple speed estimators reproduce MT responses and identify strength of visual illusion”, Neural Computing and Applications, 2017.
DOI:10.1007/s00521-017-3211-5 (印刷中)

国際会議における発表 <査読有>

- (3) Xuehua Han, Shunji Satoh, Daiki Nakamura, Kazuki Urabe, "Unifying computational models for visual attention yields better scores than state-of-the-art models", INCF Japan Node International Workshop: Advances in Neuroinformatics 2014 (AINI 2014), RII-6, Saitama, September 2014.
- (4) Kazuki Urabe, Daiki Nakamura, Shunji Satoh, “HI-brain: a software platform for numerical simulation of the complex vision system by connecting and reusing

existing models” , 4th AINI 2016 and 14th INCF Nodes Workshop, DS-1, RIKEN Wako, May 28, 2016.

- (5) Daiki Nakamura, Shunji Satoh, “A simple visual model accounts for drift illusion and reveals illusory patterns” , The 23rd International Conference on Neural Information Processing (ICONIP 2016), Kyoto, October 2016.

国際会議における発表 <査読無>

- (6) Daiki Nakamura, Shunji Satoh, “A Novel Computational Theory of MT Neurons: Do MT Neurons Actually Prefer Their ‘Preferred Speeds’ ?” , Neuro2013 (36th Annual Meeting of Japanese Neuroscience Society), P2-1-239, Kyoto, June 2013.

国内学会における発表

- (7) 飯野希, 佐藤俊治, 中村大樹, “速度知覚のパターン依存性に関する計算論的考察” , 日本視覚学会 2013 年冬季大会, 2p05, 東京, 2013 年 1 月.
- (8) 中村大樹, 佐藤俊治, “MT 細胞の電気生理実験結果に関する計算論的再考察 – MT 細胞が最大発火する速度は本当に「preferred speed」なのか? –” , 電子情報通信学会技術研究報告, vol. 113, no. 500, NC2013-95, pp. 41-46, 東京, 2014 年 3 月.
- (9) 中村大樹, 佐藤俊治, 韓雪花, 占部一輝, “視覚脳科学研究を目的とした RT ミドルウェアの応用と結果” , 第 15 回 計測自動制御学会 システムインテグレーション部門講演会 (SI2014) , 1A1-4, 東京, 2014 年 12 月.

- (10) 韓雪花, 佐藤俊治, 中村大樹, 占部一輝, “Neuroinformatics 的観点から構築された新規 Saliency map モデル”, 日本視覚学会 2015 年冬季大会, 2o05, 東京, 2015 年 1 月.
- (11) 中村大樹, 佐藤俊治, “視覚数理モデルによる錯視の説明と錯視画像の生成”, 電子情報通信学会技術研究報告, vol. 114, no. 515, NC2014-94, pp. 151-156, 東京, 2015 年 3 月.
- (12) 中村大樹, 佐藤俊治, “簡素な MT 細胞モデルによる複雑な細胞特性の再現”, 第 25 回日本神経回路学会全国大会, P-01, 東京, 2015 年 9 月.
- (13) 菊池勇作, 中嶋豊, 佐藤俊治, 中村大樹, “回転振動錯視に対する心理物理実験と計算論的考察”, 第 25 回日本神経回路学会全国大会, P-06, 東京, 2015 年 9 月.
- (14) 占部一輝, 佐藤俊治, 中村大樹, “視覚研究用シミュレーション基盤: 数理モデルの結合, 追加及び置換を行うための手法”, 第 25 回日本神経回路学会全国大会, P-21, 東京, 2015 年 9 月.
- (15) 中村大樹, 佐藤俊治, “計算論的に最適な速度推定器によって MT 野細胞の複雑な反応特性を説明する”, 第 27 回日本神経回路学会全国大会, P-35, 福岡, 2017 年 9 月.

受賞など

- (1) 第 15 回 計測自動制御学会 システムインテグレーション部門講演会 (SI2014)
RT ミドルウェアコンテスト 2014 日本ロボット工業会賞, 2014 年 12 月 15 日.
- (2) 第 15 回 計測自動制御学会 システムインテグレーション部門講演会 (SI2014)
RT ミドルウェアコンテスト 2014 ベストサポート賞, 2014 年 12 月 15 日.

- (3) 日本神経回路学会 優秀研究賞, 2015 年 9 月 4 日.
- (4) ロボットビジネス推進協議会 RT ミドルウェア普及貢献賞, 2015 年 12 月 2 日.
- (5) 電気通信大学 学生表彰 (研究活動), 2016 年 3 月 25 日.
- (6) ICONIP2016 Student Travel Grants, October 2016.
- (7) 第 27 回日本神経回路学会全国大会 JNNS2017 大会奨励賞, 2017 年 9 月 22 日.