

ランダムドット運動刺激の速度と運動残効の持続時間の関係：計算科学的モデルの構築に向けて

Relationship between Random Dot Motion Stimulus and Motion Aftereffect Duration Toward Computational Scientific Modeling

朱 剣[†], 日高 昇平[†]

Jian Zhu, Shohei Hidaka

[†]北陸先端科学技術大学院大学

Japan Advanced Institute of Science and Technology

Abstract

Motion aftereffect is a phenomenon in which people adapt to a motion stimulus for a relatively long time, and after the motion stops, perceive a motion of a stimulus that should not exist (opposite to the presented direction). Excluding the phenomenon of motion aftereffect from the visual model is the main limitation of mainstream theoretical models. Researchers have proposed models by adding structural units in which motion aftereffects occur to the mainstream theoretical model. There are two attributes of motion aftereffect: the direction of the stimulating motion and the speed of the stimulating motion. However, these theoretical models are limited to the explanation of the direction of motion aftereffect, and the relationship between the duration of motion aftereffect and the speed of motion stimulus is not clarified. The purpose of this study is to construct a computational model that predicts the relationship between the speed of exercise stimulus and the duration of exercise aftereffect, and to explain the findings of psychophysical experiments of exercise aftereffects computationally.

Keywords — Motion Aftereffect, Visual Cortex, Velocity Perception, Computational Model

1. はじめに

運動残効は人々が比較的長い時間に運動刺激に順応し、その運動が停止した後、存在しないはずの刺激の(提示された方向とは逆向きの)動きを知覚する現象である。運動残効は、運動知覚を促進するための指標と見なされる。Mather らの実験では、ストライプカバレッジがない場合と比較し、ストライプカバレッジがランダムドット運動刺激の動きの知覚を促進することを示している[1]。その理由は、後者のほうが運動残効は短いからである。一方、スリット錯視と呼ばれる錯視では、ストライプカバレッジで覆われたランダムドット運動刺激がストライプカバレッジで覆われていないランダムドット運動刺激と比較される。ランダムドット運動刺激の速度はどちらの場合も同じだが、被験者は前者の動きが速いようだと言及する傾向がある。Mather らの実験とスリット錯視によって明らかにされた現象を組み合わせると、速度知覚の逸脱は運動残効と何らかの相関関係があると考えられ

る。

運動残効の現象を視覚モデルから除外することが、主流な理論モデルの主な制限である。この制限を解決するために、van de Grind らは第五次視覚野に相当する脳の領域において運動残効が発生する構造単位を追加し、モデルを提案した[2]。これに対し、Pavan らは第一次視覚野に相当する脳の領域において運動残効が発生する構造単位を追加し、順応時間と運動残効の持続時間の関係を説明した[3]。2つの研究において追加された構造単位は同等であり、広く受け入れられている。一方、運動残効には刺激運動の方向性と刺激運動の速度の2つの属性がある。刺激運動速度と持続時間は密接に関連している[4]。ただし、これらの理論モデルは運動残効の方向性の説明に限定されており、運動残効の持続時間と運動刺激の速度の関係については、明らかにされていない。

2. 目的

本研究の目的は、運動刺激の速度と運動残効の持続時間との関係を予測する計算モデルを構築し、運動残効の心理物理実験の知見を計算論的に説明することである。

3. 心理物理実験

ランダムドット運動刺激の速度と運動残効継続時間の単調に減少する関数という関係を追試で確認する。

■方法

北陸先端科学技術大学院大学の院生 5 名 (年齢 $M=24.00$, $SD=2.26$) は実験に参加した。被験者は全員、正常視力であった。実験ではレーザー視距離計 (Leica 社 DISTO plus 型)、実験専用コンピューターを用いた。図 1 に実験刺激のイメージを示している。中央の十字形は $2\text{ mm} \times 2\text{ mm}$ であり、ランダムドットは $1.5\text{ mm} \times 1.5\text{ mm}$ であった。

刺激呈示の速度は 17 種類があった。各試行の最初に、画面の中心に注視点が呈示された。次に運動刺激が呈示された。被験者は 30 秒続けて注視点を注目してい

た。そして、30秒後に静止刺激が提示された。被験者は注視点を注目しつづけ、元の方向と反対の動きを知覚した。被験者にこの現象が止まったと判断する時点で、キー押しを求められた。これらの手続きを1試行とし、14試行を行った。

■結果

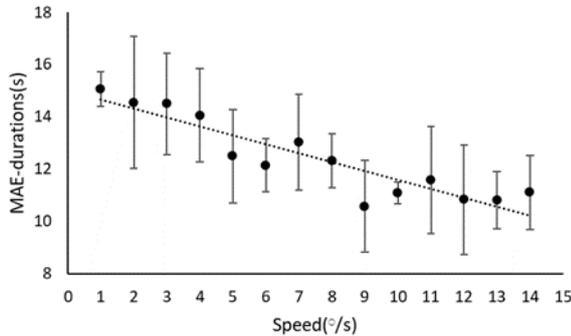


図1 ランダムドット運動刺激の速度(横軸)と運動残効の平均持続時間の関係(縦軸)

図1にランダムドット運動刺激の速度と運動残効の持続時間の関係を示す。提示した運動刺激の速度の増加に従い運動残効の継続時間は減少する傾向が示される。

「運動残効の平均持続時間」に対して1要因参加者内計画の分散分析を行った結果、ランダムドット運動刺激の速度の効果が有意に認められた ($F(13, 52)=12.17, p<.05$)。そこで、有意水準5%でHolm法による多重比較を行ったところ、速度1°/sの場合は、8, 9, 10, 11, 12, 13, 14°/sの場合より、運動残効の平均持続時間は有意に長かった。

速度2°/sの場合は、6, 8, 9, 10, 11, 12, 13, 14°/sの場合より、運動残効の平均持続時間は有意に長かった。速度3°/sの場合は、8, 9, 10, 11, 12, 13, 14°/sの場合より、運動残効の平均持続時間は有意に長かった。速度4°/sの場合は、9, 10, 11, 12, 13, 14°/sの場合より、運動残効の平均持続時間は有意に長かった。速度7°/sの場合は、9, 12, 13°/sの場合より、運動残効の平均持続時間は有意に長かった。他のペアは有意差が認められなかった。

4. モデル化に向けて：先行研究のモデルの整理

長年にわたり、コンピュータビジョンや神経科学などのさまざまな分野で多くの運動処理モデルが提案された。提案されている生物学に解釈可能なモデルのほとんどは、視覚運動処理に関与する異なる脳領域間の役割と相互作用を理解する試みである。既存のモデルの変換と再構築により、運動後の効果と運動刺激の速度を表す計算モデルを確立する。

Adelson らは、位相に依存しない動き検出器を提案した(図2)。Adelson-Bergen の運動エネルギーモデルでは2つの線形時空間要素を組み合わせ、それらの2乗応答を合計する。数学的便宜上、Gabor関数の理想的な場合を考える。1つのケースには偶数フェーズ(コサイン)があり、もう1つのケースには奇数フェーズ(サイン)がある。典型的な一次元Gabor時空間単位は以下の1次元空間と時間の関数で表される：

$$F_{odd}(x, t) = \exp\left(-\frac{x^2}{2\sigma^2}\right) \sin(\omega_x x + \omega_t t)$$

$$F_{even}(x, t) = \exp\left(-\frac{x^2}{2\sigma^2}\right) \cos(\omega_x x + \omega_t t)$$

これらの式においてσはガウス分布の標準偏差を表す。ω_xとω_tはそれぞれ空間周波数と時間周波数を表すである。

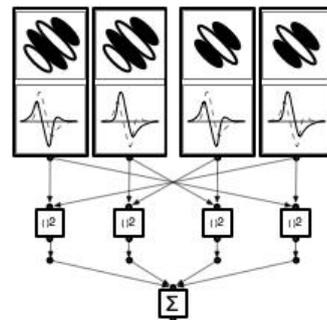


図2 Adelson and Bergen (1985)によって提案された運動エネルギーモデル

Nowlan and Sejnowski(1994, 1995)は、オクルージョンされた透明な移動オブジェクトの2次元速度を計算するためのモーション処理モデルを提案した。彼らの目標は、正確な速度表現を持つことではなく、画像をコヒーレントな動きの領域にセグメント化し、各領域の速度の推定値を提供し、最も信頼できる推定値を積極的に選択することである。モデルは、処理の最初の段階でモーションエネルギーフィルターを使用し、2つの異なる単位応答を並列で計算する。1つの単位のセットは局所速度を推定し、2番目のセットはこれらの局所推定からグローバル速度をサポートするものを選択する(図3)。

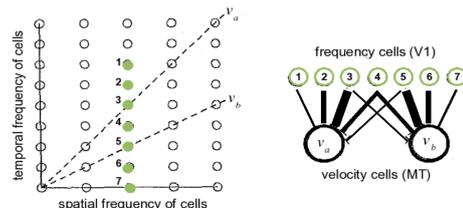


図3 Grzywacz and Yuille(1990)によって提案された速度計算の戦略

ネットワーク構造と脳の操作アルゴリズムを参照して開発されたディープニューラルネットワーク (DNN) は、コンピューターを含む幅広い分野で顕著な成功を収めている。人間の脳、特に大脳皮質は、少なくとも部分的に教師なしの方法で世界を学習する。教師なし学習に基づくこれらの理論は、脳が絶えずリアルタイムで取得される視覚情報からオブジェクトの時空間構造を抽出し、移動するオブジェクトの将来の位置と図形を予測することを示唆している。知覚の予測能力が不可避の神経遅延を補間し、人間の行動がリアルタイムで進行している世界に適応していると想定されている。自動エンコーダネットワークまたは反復性メモリセル (長期短期メモリ, LSTM) を組み込んだ生成的敵対的ネットワークを使用すると、動画内のオブジェクトの将来の状態を予測することが可能になる。PredNet と呼ばれる DNN は、予測コーディング理論に基づいて本質的に設計されている[9][10]。大脳皮質の視覚系の情報処理メカニズムを包括的に説明する。PredNet は、ビデオシーケンスの将来のフレームを予測することを学習する。ネットワークの各レイヤーは、上位レイヤーからの後方情報を使用してローカル予測を行い、それらの予測からの差分値のみを後続の上位ネットワークレイヤーに転送する (図4)。

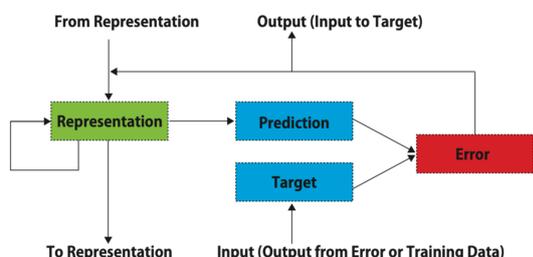


図4 Lotter(2016)によって提案された PredNet モデル

5. Prednet モデルでの探測

人間の視覚予測は非常に正確ですが、錯覚の場合など、「間違い」が時々行われます。予測コーディング理論は、錯覚を生成する理論的なメカニズムであることが示唆されると考えている。予測コーディング理論がこの錯覚を説明するとしたら、PredNet が人間の視覚的錯覚に関連するものと同様の「間違い」も生じると推測する。

■実験

自然画像シーケンスの学習のためのランダムハイパーパラメーター検索の以前の結果によると、すべての畳み込みに対して3×3のフィルターサイズを使用す

る4層モデルと、層ごとのスタックサイズが3, 48, 96, および192の4層モデルが採用される[9]。モデルの重みは、Adam アルゴリズムとデフォルトのパラメーターを使用して最適化される[11]。モデルは、Charades データセットのビデオを使用して、平均二乗誤差でトレーニングされた。ランダムに選択されたビデオがダウンサンプリングされ、MP4 (幅160×高さ120ピクセル, 30 fps) にフォーマットされた。シリアル静止 JPEG 画像 (160×120ピクセル) に再フォーマットされた。全体として、トレーニングセットは約490Kの画像で構成されていた。各画像セットは、ネットワークに順次入力された。モデルは、バッチの最後の予測エラーを使用したバックプロパゲーションにより、連続する32の画像 (バッチ) ごとに更新された。トレーニングは、Crey-XC40 で約28時間続いた。

■結果

PredNet 予測では、ランダムドット運動刺激が用いられた。刺激は15フレームシーケンスとして作成された。

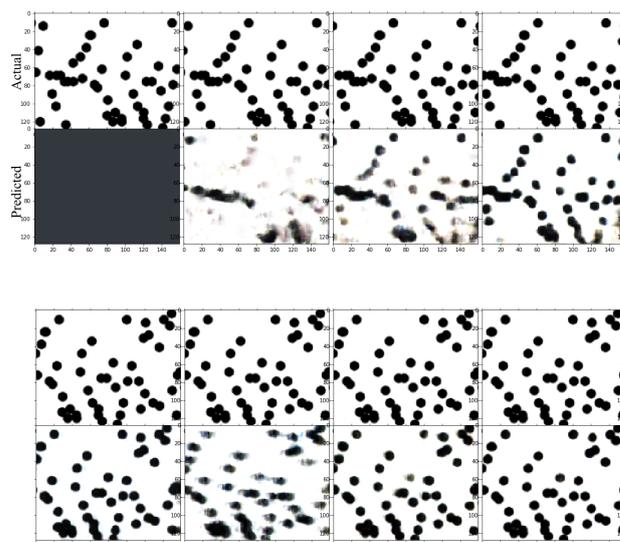


図5 PredNet によるフレームごとの予測結果

6. RC 積分

本研究では、モデルの新しい拡張機能は、「漏れやすい積分器」(図6)回路を使用してモーションセンサーに分裂的なフィードフォワードゲイン制御を実装する。一般的な特徴は、任意の時点での出力信号が、過去の入力大きさに比例して、入力の一部であることである。漏洩積分回路の最も単純な形式は、RC 積分器として知られている。出力は、入力値が増加するにつれて増加する割合で、漸近的にゼロに向かう傾向が

ある。このような単純な RC 積分器では、出力にゼロの制限値がある。これは、特定の神経回路が十分に長い時間刺激を受けた場合にゼロ効率に達することを意味する。その結果、一定の方向性刺激にさらされると、対応するセンサーが完全に無音になる。この行動が生理学的または精神物理学的に観察されることは非常に多い。

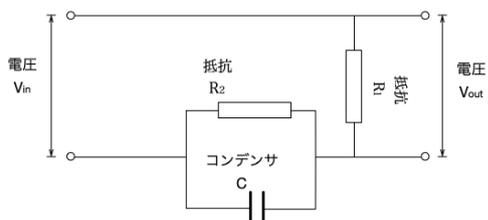


図6 モデルの新しい拡張機能の漏洩積分回路

この「漏洩積分回路」の効果は次のとおりである。一定強度の入力信号の場合、出力信号の強度は入力信号の $1/\alpha$ ($1 < \alpha < \infty$) である。入力信号強度が 0 の場合、出力端は一定期間信号を出力し続け、出力信号の強度は徐々に減少してゼロに戻る。

図6に示す回路の入力電圧と出力電圧（電圧は時間の関数です）の関係を次の関数で表される：

$$V_{out} = V_{in} - c_1 e^{-\frac{t}{R_1 C} - \frac{t}{R_2 C}} + \frac{1}{R_1 C} e^{-\frac{t}{R_1 C} - \frac{t}{R_2 C}} \int_1^t e^{\frac{\xi}{R_1 C} + \frac{\xi}{R_2 C}} \cdot V_{in}(\xi) d\xi$$

7. 今後の研究

以上の先行研究を踏まえて、今後の課題として漏洩積分回路を PredNet モデルに埋め込み、新たなモデルを構築する。大会発表ではモデルシミュレーションの結果とモデルの有効性の分析について報告する予定である。

参考文献

- [1] Mather G., Pavan A., Bellacosa R. M., & Casco C., (2012) “Psychophysical evidence for interactions between visual motion and form processing at the level of motion integrating receptive fields”, *Neuropsychologia*, Vol. 50, No. 1, pp. 153-159.
- [2] van De Grind W. A., Lankheet M. J. M., & Tao R., (2003) “A gain-control model relating nulling results to the duration of dynamic motion aftereffects”, *Vision research*, Vol. 43, No. 2, pp. 117-133.
- [3] Pavan A., Contillo A., & Mather G., (2013) “Modelling adaptation to directional motion using the Adelson-Bergen energy sensor”, *PloS ONE*, Vol. 8, No. 3.
- [4] Pantle A., (1974) “Motion aftereffect magnitude as a measure of the spatio-temporal response properties of direction-sensitive analyzers”, *Vision Research*, Vol. 14, No. 11, pp. 1229-1236.

- [5] Adelson E. H., & Bergen J. R., (1985) “Spatiotemporal energy models for the perception of motion”, *Josaa*, Vol. 2, No. 2, pp. 284-299.
- [6] Nowlan S. J., & Sejnowski T. J., (1995) “A selection model for motion processing in area MT of primates”, *Journal of Neuroscience*, Vol. 15, No. 2, pp. 1195-1214.
- [7] Nowlan S. J., & Sejnowski T. J., (1994) “Filter selection model for motion segmentation and velocity integration”, *JOSAA*, Vol. 11, No. 12, pp. 3177-3200.
- [8] Wilson H. R., Ferrera V. P., & Yo C., (1992) “A psychophysically motivated model for two-dimensional motion perception”, *Visual neuroscience*, Vol. 9, No. 1, pp. 79-97.
- [9] Jackson, P., Wuerz, R., Hariskos, D., Lotter, E., Witte, W., & Powalla, M. (2016). Effects of heavy alkali elements in Cu (In, Ga) Se₂ solar cells with efficiencies up to 22.6%. *physica status solidi (RRL)–Rapid Research Letters*, 10(8), 583-586.
- [10] Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, 2(1), 79-87.
- [11] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.