

スキャンパス理論に基づく移動ロボットの手先動作時の関心領域特性と検出特性  
 Positional Features and Algorithmic Predictability of Regions-of-Interest in Robot Hand  
 Movement based on Scanpath Theory

藤田 豊己<sup>†</sup>  
 Toyomi Fujita

### 1. はじめに

ロボットが人間や他のロボットと協調して作業を行う場合、他者の行動を観察してその動作内容を認識することが必要である。そのとき、ロボットに人間と同様な注視機能があれば有効であると考えられる。そこで、他者の動作観察時の人間のスキャンパス特性を解析し、応用することができないかと考えている。スキャンパスとは人間の視覚走査における固視点の一連の軌跡であり、この固視によって捉えられる関心領域は、ROI (region-of-interest) と呼ばれている[1]。筆者らはこれまでに被験者や画像種類によるスキャンパス特性を解析し、関心領域が脳内で作用するトップダウン的な認知処理およびボトムアップ処理との度合に依存するとの知見を示した[2]。人間同士の協調作業においても、自ずと相手を観察して他者の動作を理解する目的がはたらき、そのときの関心領域はその情報獲得に有効となるような領域となることが予想される。例えば、どの程度トップダウン的な処理が優位にはたらくのか、などの特性がわかれば他者動作理解にむけたロボットの注視機能のための有益な知見が得られることに期待が持てる。

そこで本研究では、ロボットの注視機能の実現に向けて、ロボット動作観察時の人間のスキャンパス特性を検証する。具体的には、協調相手のロボットが対象物を操作しようとしている場面を一方のロボットが観察している場面を想定する。最初に、観察ロボットが捉えた相手ロボット手先動作のシーンを記録し、その中の画像を使用し、それを被験者が観察したときの関心領域を計測する。本研究では、この関心領域を hROIs (human regions-of-interest) と呼ぶ。そして、計測された hROIs を位置的類似度による指標を用いて比較し、関心領域特性を検証する。さらに、画像処理手法を応用してロボット動作時の関心領域の検出を試みる。この関心領域を aROIs (algorithmic regions-of-interest) と呼ぶ。そして同様に位置的類似度による指標を用いて aROIs と hROIs とを比較し、検出した関心領域の有効性および妥当性を検証する。

### 2. 関心領域 (hROIs) の計測

被験者がディスプレイ上に表示される画像を観察したときの関心領域 hROIs (human regions-of-interest) を計測した。ここでは、呈示した画像上の関心のある領域にマウスポインタを移動してクリックする実験を行い、その指定された領域を関心領域とみなす。先行研究で、同様の方法によって得られた関心領域は、眼球運動計測による実際のスキャンパスと同程度の位置的類似度を持つことが報告されており[3]、本手法もその知見に基づいた。

視覚刺激となる画像には、ロボットが作業のために手先

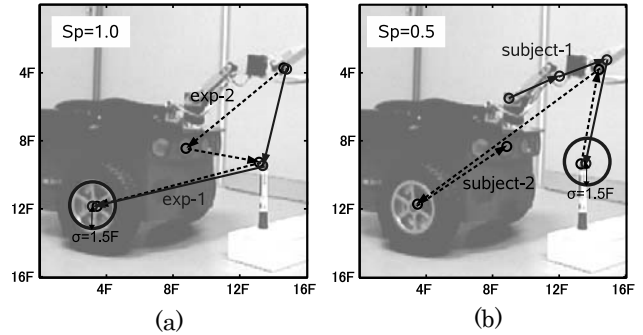


図1 hROIsの実験結果とそれらの比較例

を動作させているシーンの連続画像の中から3枚を選んだ。それらの一部を図1に示す。さらに無関係な画像を7枚加え、計10枚の画像を使用した。それらを1枚ずつランダムに表示し、各画像について計測を行った。被験者には何も指示を与えずに観察させたときの関心領域を計測した。7人の被験者が実験に参加し、各被験者について、この計測を3回以上おこなった。各画像の表示サイズは全て  $24^\circ \times 24^\circ$  とした。ここで、中心窩の視野角  $1.5^\circ$  に相当する画像サイズを単位 F で定義し、以下で用いる。

### 3. hROIs の位置的類似度特性

各 hROIs の位置的な類似度を  $Sp$  (positional similarity index) と呼ばれる指標により評価する。2組の hROIs を比較する場合、基準となる hROIs の ROI 数が  $n$  で、その中の  $m$  個が比較されるもう一方の hROIs 中のどれか一つの ROI と位置的に一致するとき、類似度の値を  $Sp=m/n$  とする。位置的な一致を判断するための許容距離  $\sigma (=1.5F)$  を与え、ある ROI の中心が別の ROI の中心から半径  $\sigma$  内の領域にあれば両者の位置は一致するとした。

この方法で、同一画像における同一被験者での異なる実験の hROIs 間と、異なる被験者から得られた hROIs 間で比較した。各  $Sp$  値を  $SpRhh$ ,  $SpLhh$  と呼ぶ。図1はそれらの結果例を示す。図1(a)は  $SpRhh$  の例で、exp-1 の hROIs を基準にし、その3つの ROI がいずれも exp-2 の ROI のいずれかと  $\sigma$  の範囲内で一致するため、 $SpRhh=1.0$  となる。図1(b)は  $SpLhh$  の例で、subject-1 の4つの ROI のうち2つが subject-2 の ROI のいずれかと距離  $\sigma$  内にあるため、 $SpLhh=0.5$  となる。

この  $SpRhh$  と  $SpLhh$  について、被験者7人において可能な全ての hROIs の組み合わせから平均値を求めた。各画像および3画像全体における各平均値を図2に示す。 $SpRhh$  は非常に高い値となった。また、 $SpLhh$  についても  $SpRhh$  ほどではないが高い値となった。

<sup>†</sup> 東北工業大学工学部知能エレクトロニクス学科  
 Tohoku Institute of Technology

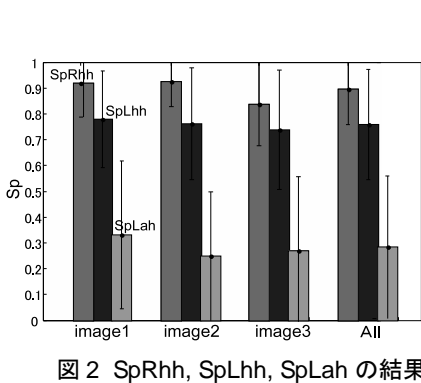


図2 SpRhh, SpLhh, SpLahの結果

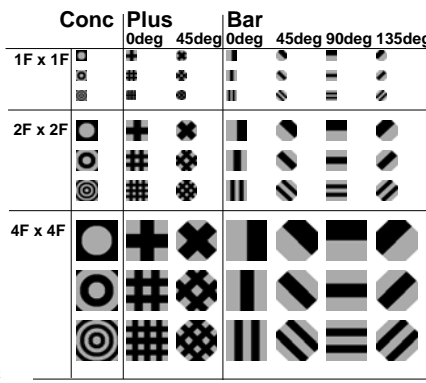


図3 画像処理パターン

	Concentric		Plus		Bar			
		0deg	45deg	0deg	45deg	90deg	135deg	
1F x 1F (mean 0.50)	0.69	0.61	0.58	0.28	0.26	0.19	0.27	
	0.59	0.50	0.61	0.53	0.68	0.58	0.54	
	0.85	0.45	0.34	0.45	0.64	0.31	0.52	
2F x 2F (mean 0.23)	0.13	0.18	0.16	0.18	0.17	0.13	0.17	
	0.21	0.09	0.42	0.27	0.27	0.20	0.23	
	0.56	0.07	0.26	0.17	0.52	0.00	0.42	
4F x 4F (mean 0.12)	0.01	0.03	0.06	0.04	0.04	0.00	0.01	
	0.08	0.12	0.20	0.11	0.16	0.00	0.27	
	0.44	0.03	0.16	0.29	0.31	0.00	0.23	

図4 各パターンにおける SpLah

#### 4. 画像処理手法による関心領域(aROIs)の検出

対象画像に対して画像処理を適用して得られる特徴領域を aROIs (algorithmic regions-of-interest) と定義する. 本研究では, 基本特徴のサイズ, 方向, 周期を変化させた計 63 通りの同心円型 (Concentric), 十字型 (Plus), 棒型 (Bar) の正規化パターンを用いた (図 3). 先の実験で使用した画像の正規化したグレイレベルデータとの相関を求め, それから生成された特徴マップの局所最大値をとる領域を値の高い順に 7 個抽出したものを aROIs とした [4].

この方法によって得られた aROIs と hROIs の位置的な類似度を 3 章で述べた Sp の指標によって同様に求め, 関心領域の検出性能を評価した. この Sp を SpLah と呼ぶ.

そして, 全てパターンにおける aROIs と hROIs との組み合わせから得られた SpLah 値の平均値を求めた. 各画像および 3 画像全体における SpLah の平均値を 3 章で述べた SpRhh, SpLhh の結果とともに図 2 に示す. この結果では, SpLah は SpRhh, SpLhh と比べて非常に低い値となっている. 次に, 各パターンにおける SpLah (3 画像平均値) を図 4 に示す. 各表示位置は図 3 のパターン位置と対応している. いくつかで SpLah 値が高くなり, 最高値は 0.85 と全平均と比べて非常に高い値となった. また, サイズ 1F x 1F の場合で高い SpLah 値となる傾向が得られた.

#### 5. 考察

位置的類似度 Sp は, 2 つの ROI セットにおける位置に関する一貫性を示す. 図 2 の結果から, SpRhh 値は 3 画像平均で 0.90 となり, 同一被験者での高い一貫性が得られた. スキャンパス理論では, 脳内の内部認知モデルがトップダウンの能動的視覚を制御すると論じており [1], この結果は, 同一の内部認知モデルによるプロセスが脳内ではたらく, 一貫性の高い関心領域を生成したということの意味すると考えられる.

また, SpLhh 値は, 3 画像全体の平均で 0.76 と高くなった. 本実験では被験者に特定のタスクを与えておらず, その場合でも対象物を操作しようとするようなロボット動作の観察では自ずとトップダウンの能動的視覚が強くなると思われ. また, SpRhh の値より低い値となったのは, 異なる被験者では脳内の認知モデルによるプロセスの一貫性が同一被験者間より低いためであると考えられ, 従来の知見 [5] を支持する結果となった.

4 章での aROIs と hROIs の比較による SpLah 値は, 画像処理手法による ROIs の検出性能を示すと考えることがで

きる. 図 2 に示した全てのパターンの平均値が低くなったのは, これらのパターンがボトムアップ的な特徴を主に示すからであろう. SpLhh の結果からトップダウン能動的視覚が強くなるかと考察されたロボット動作画像においては, これらの手法だけでは ROI 検出には不十分であるかもしれない. しかしながら, 図 4 の結果から, サイズ 1F x 1F のものは高い SpLah 値を示すことがわかった. SpLhh の結果が示す通り, 異被験者間でもその関心領域は完全には一致しない. そこで, SpLhh の値を検出性能の最高基準値と考えると, 一部の SpLah 値には, 3 画像の SpLhh 値 (0.76) に近い値を示したものもあり, それらの検出性能は高いものであると評価できる. 1F x 1F のパターンは中心窩サイズのボトムアップ処理をよく表すと考えられ, その点でも妥当性があると言える.

#### 6. おわりに

本研究では, ロボット手先動作画像を観察したときの人間のスキャンパス関心領域を計測し, 位置的類似度 Sp を用いて特性を検証した. さらに, 画像処理手法により関心領域を検出し, その妥当性を評価した.

本研究では, 心理実験により計測されたスキャンパス特性を解析し, ロボットの注視機構への応用を試みるという点に特徴がある. 現段階ではまだ対象画像の数などで不十分であるため, 今後さらに様々な画像を使用するなどして特性を検証していく必要がある.

#### 謝辞

本研究の一部は日本学術振興会科研費基盤研究 (C)(20500182) の支援を受けた.

#### 参考文献

- [1] D.Noton and L.W.Stark, "Scanpaths in eye movements during pattern perception," Science, Vol.171 pp.308-311, (1971).
- [2] Claudio M. Privitera, Toyomi Fujita, Dimitri Chernyak, Lawrence W. Stark, "On the discriminability of hROIs, human visually selected regions-of-interest," Biological Cybernetics, Vol.93, No.2, pp.141-152, (2005).
- [3] Lawrence W. Stark, Claudio M.Privitera, Huiyang Yang, Michela Azzariti, Yeuk Fai Ho, Ted Blackmon, Dimitri Chernyak, "Representation of human vision in the brain: How does human perception recognize images?," Journal of Electronic Imaging, Vol.10, No.1, pp.123-151, (2001)
- [4] L.Itti and C.Koch, "Computational modeling of visual attention," Nature Reviews, 2(3):194-203, (2001).
- [5] L.W.Stark and Y.Choi, "Experimental metaphysics: the scanpath as an epistemological mechanism," Visual Attention and Cognition, Edited by W.H.Zangemeister, H.S.Stiehl, and C.Freksa, Eds., pp.3-69, Elsevier, Amsterdam (1996).