

機関番号：31303
 研究種目：基盤研究（C）
 研究期間：2008～2010
 課題番号：20500182
 研究課題名（和文） Scanpath 理論に基づくロボットの他者動作認識のための注視機能の研究
 研究課題名（英文） A Study on Robotic Attention for Action Detection based on Scanpath Theory
 研究代表者
 藤田 豊己（FUJITA TOYOMI）
 東北工業大学・工学部知能エレクトロニクス学科・准教授
 研究者番号：90293141

研究成果の概要（和文）：ロボットが作業のために手先を動作している画像を観察したときの人間の注視による関心領域を計測し、位置的類似度を用いて特性を検証した。その結果、対象とした動作においては、トップダウン的な処理が優位な Scanpath（視線走査）が生じるがわかった。さらに、画像処理手法により関心領域を検出し、その妥当性をその特性結果を利用して評価した。その結果、ボトムアップ処理とトップダウン処理を表す特徴を統合することで有効な注視領域検出が可能となり、基本的な動作検出も可能となることを示した。

研究成果の概要（英文）：This study investigated human visual scanpath features in a scene of robot hand movement. Human regions-of-interest (hROIs) were measured in psychophysical experiments and compared using a positional similarity index on the basis of scanpath theory. Results showed consistent hROI loci due to dominant top-down active looking in such a scene. This study also presented a method for detecting hROIs which used both bottom-up and top-down image processing algorithms (IPAs). Algorithmic regions-of-interest (aROIs) generated by this method were compared to the hROIs. The results showed that the IPAs have a high ability to predict hROIs and to detect actions of a partner robot.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008 年度	2,100,000	630,000	2,730,000
2009 年度	500,000	150,000	650,000
2010 年度	600,000	180,000	780,000
年度			
年度			
総計	3,200,000	960,000	4,160,000

研究分野：ロボット工学，情報科学

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：ロボット注視機能，Scanpath，関心領域

1. 研究開始当初の背景

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し、それに対応した行動をとることができれば効率的である。例えば、図1に示すように、複数の移動マニピュレータ型ロボットが軸と車輪型円盤の組み立て

作業を行う場合を考える。円盤は黄緑色のロボットの周辺にある。軸は両者のロボットの中間にあり、水色のロボットは既にその方向に向っている。このとき、黄緑色のロボットにとって、相手のロボットがその軸をどのように掴もうとしているのかによって対応すべき行動は大きく異なってくる。軸を上から

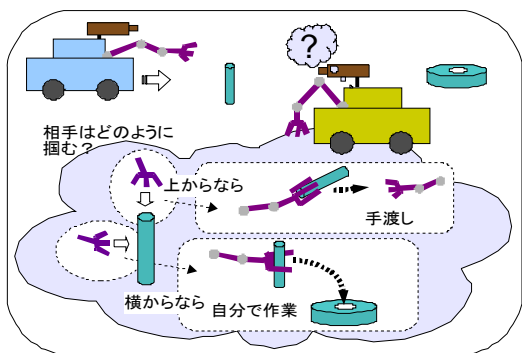


図1 協調行動の例

掴めば相手はそれを自分に渡して次の作業に移ると推測できるので、相手から軸を受け取るよう行動するべきである。一方、軸の中心付近を横から掴めば相手はそのまま移動して自ら円盤へのはめ込み作業を行うと推測できるので、速やかに次の作業に移ればよい。このように、円滑な協調作業のためにはロボットが他者の動作を認識することが重要となる。

他者の動作を認識するためには、動作の動画の中からある特徴領域を注視点として注目し、連続する部分動作とその遷移を検出する必要がある。従来研究では、限定されたタスクに対して事前に注視点や動作内容を設定する手法がとられている。しかし、協調作業の場合は、周囲の環境や対象物の位置関係を考慮する必要があり、全ての状況において事前に注視点を与えて部分動作や遷移の時点を定義することは難しい。そこで、本研究では、人間がロボットの動作を観察したときの注視点を解析し、その知見を応用することで注視機能を実現する手法を提案する。

人間は固視とサッケードから成る眼球運動を繰り返し、環境を認識している。この固視点の一連の軌跡は Scanpath と呼ばれており、特定領域(関心領域)への注視を繰り返すことがわかっている。申請者は、先に行った研究において、Scanpath による関心領域は、脳内で作用するトップダウン的な認知処理と、画像特徴から生じるボトムアップ処理との度合に依存するとの知見を得た。つまり、人間は、意図や目的に応じて Scanpath を変え、環境認識のための情報を効果的に獲得していると言える。「協調作業における他者動作の理解」という明確な目的がある状況では、脳内でトップダウン的な処理が強く作用し、そのときの Scanpath は非常に高い類似性を持つものと予想される。そこで、他者ロボットの動作観察時の人間の Scanpath 特性を解析することで動作理解のための注視特性が得られることが期待できる。さらにその結果を応用してロボットの注視機能を実現することが可能となると考えられる。

2. 研究の目的

本研究の目的は、背景で述べた Scanpath 特性の知見を応用するとともに、画像処理に基づくボトムアップ的手法とトップダウン的手法とを統合することで Scanpath 関心領域を検出し、ロボットの他者動作認識のための注視機能を実現することである。さらに、その注視機能に基づいて他者ロボットの動作検出の実現を目指す。

(1) ロボット動作観察時の人間の Scanpath 特性の検証

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。

(2) 画像処理手法によるロボット動作時の関心領域の検出

画像処理手法により得られる関心領域を 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。また、トップダウンに与えた特徴も加味して生成された関心領域についても同様に比較し、その有効性を検証する。

(3) ロボット動作の検出への応用

本手法により生成された関心領域をロボット動作の検出に応用する。連続的に得られた関心領域から動作を検出する。

3. 研究の方法

(1) 実験用ロボットシステム構築

実験用ロボットシステムとして、2台の車輪型移動ロボットを使用し、一台は観察用、もう一台は動作用とした。観察用ロボットにはパン、チルト、ズーム制御が可能なカメラを搭載し、画像処理用 PC に無線で送信できるようにした。画像処理 PC では画像入力ボードを用いて動画画像処理を行うビジョンシステムを構築した。動作用ロボットには、5自由度マニピュレータを搭載し、対象物のピックアンドプレイス操作を可能にした。

(2) Scanpath 計測実験および特性解析

構築したロボットシステムにおいて、動作ロボットが物体操作のためにハンドを動作し

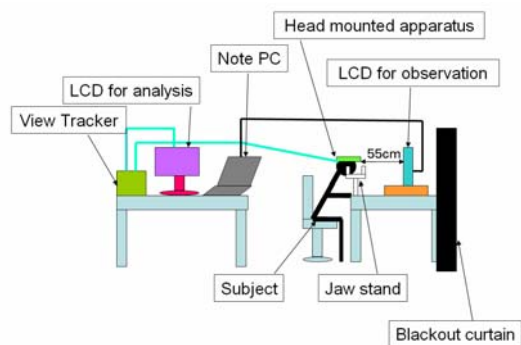


図2 Scanpath計測実験システム

ているシーンを観察ロボットにより記録し、その中の3枚の画像を被験者が観察したときの関心領域を計測した。計測システムを図2に示す。被験者は計測用ヘッドマウントを装着し、目の高さがディスプレイ中央に合うようにアゴ台で頭を固定して画像を観察する。被験者とディスプレイ間の観察距離は55cmとした。ディテクト社のView Trackerを使用して視線を計測し、停留点を抽出してhROIsとした。

各被験者について、この計測を3回以上行った。ここで、中心窩の視野角 1.5° に相当する画像サイズを単位Fで定義する。

次に、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した。2組のhROIsを比較する場合、基準となるhROIsの関心領域数がnで、その中のm個が比較されるもう一方の関心領域中のどれか一つ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、ある関心領域の中心が別の関心領域の中心から半径 σ 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間で比較し類似度Spを求めた。このSp値をSpRhhと呼ぶ。また、異なる被験者から得られたhROIs間でも比較した。このSp値をSpLhhと呼ぶ。

図3にSpRhh, SpLhhの例を示す。図3(a)はSpRhhの例で、exp-1のhROIsを基準にし、その3つの関心領域がいずれもexp-2の関心領域のいずれかと σ の範囲内で一致するため、 $SpRhh=1.0$ となる。図3(b)はSpLhhの例で、subject-1の4つの関心領域のうち2つがsubject-2の関心領域のいずれかと距離 σ 内にあるため、 $SpLhh=0.5$ となる。

(3) 画像処理によるScanpath検出

注視機能実現のために、画像処理カーネルを利用してアルゴリズムによる関心領域を検出し、Scanpathを生成した。それと被験者

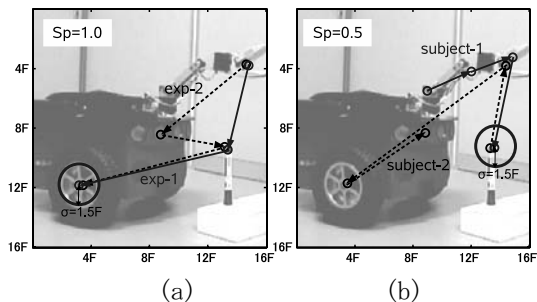


図3 hROIsの実験結果とそれらの比較例

データとの位置的類似度特性を被験者データ間の比較と同様の手法で求めることで検出性能を評価した。

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。それらを図4に示す。それぞれのパターンについて、先の実験で使用した画像を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。

さらに、ロボットや対象物の領域をトップダウン処理による特徴パターンとして与え、カーネルによるボトムアップ処理と統合してScanpathを検出した。トップダウンによるパターンとして、ロボットの特徴領域 (車輪、ロボットアームのベース、肘、手先) と対象物 (ペン) の領域を与えた。手先については反時計回りに 45° 、 90° 回転した領域も使用した。これらと各画像との相関をとることでマッチングによる特徴マップが得られる。それらと各カーネルパターンによるマップとを線形結合したものから、上記と同様に各パターンにおけるaROIsを求めた。それらはトップダウンとボトムアップ処理の統合により検出されたScanpathを表す。

(4) 注視機能による動作検知

Scanpath検出で使用したカーネルをロボットのハンド動作の動画像に適用して得られた関心領域より動作検出を行った。動画像内の各画像に対して上記の方法でaROIsを検出し、その変化から動作ベクトルを抽出した。ある画像のaROIs内の1つの関心領域について、次の画像においてその近傍領域にある関心領域を同一関心領域とみなして動作ベクトルを求めた。2画像間での領域間距離がある閾値を超える場合は同一関心領域とはみなさないようにした。

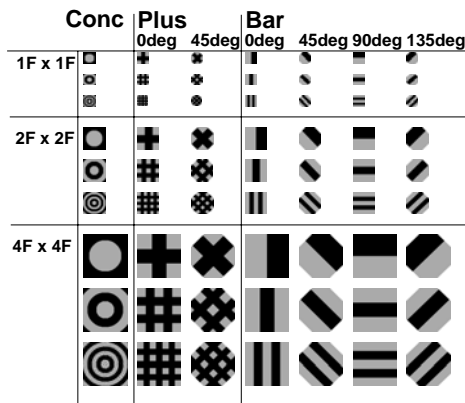


図4 画像処理パターン

4. 研究成果

(1) Scanpath の位置的類似度特性

先に述べた SpRhh と SpLhh について、被験者7人において計測された hROIs から可能な全ての組み合わせにおける値を求め、それらの平均値を求めた。ロボット動作順の各画像 (Image1~3) および 3 画像全体における各平均値を図5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。

位置的類似度 Sp は、2つの関心領域のセットにおける位置に関する一貫性を示す。この結果から、同一被験者での高い一貫性が得られたことがわかる。Scanpath 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、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。

また、SpLhh 値も高い値となった。本実験では対象物を操作しようとするようなロボットハンドの動作を観察したが、このような観察対象では自ずとトップダウンの能動的視覚が強いはたらくからであると考えられる。これが SpRhh の値より低くなったのは、異なる被験者では脳内の認知モデルによる

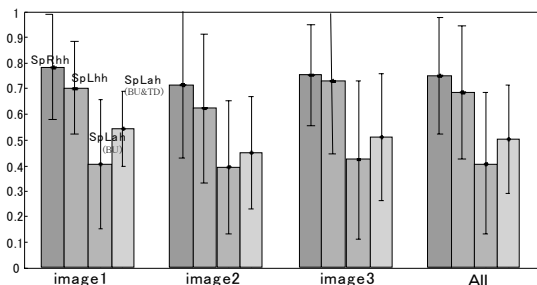


図5 Scanpath比較結果

	Concentri	Plus		Bar			
		0de	45de	0de	45deg	90deg	135de
0.5F x 0.5F (mean0.47)	0.6	0.4	0.6	0.2	0.4	0.3	0.5
	0.3	0.4	0.5	0.3	0.5	0.4	0.6
	0.4	0.4	0.3	0.4	0.5	0.5	0.6
1F x 1F (mean0.42)	0.6	0.3	0.5	0.2	0.3	0.2	0.3
	0.5	0.4	0.4	0.2	0.5	0.2	0.6
	0.5	0.4	0.3	0.3	0.6	0.1	0.6
2F x 2F (mean0.32)	0.5	0.3	0.2	0.2	0.5	0.1	0.1
	0.1	0.1	0.5	0.2	0.5	0.2	0.4
	0.4	0.0	0.5	0.2	0.4	0.0	0.5

図6 各パターンにおけるSpLah

プロセスの一貫性が同一被験者間より低いためであり、従来の知見を支持する結果である。

(2) 画像処理手法による関心領域 (aROIs) の検出

前述の方法によって得られた aROIs と hROIs の位置的類似度を先に述べた 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。

図5のそれぞれ左から3番目の "SpLah (BU)" の棒グラフが各画像および3画像全体における SpLah の平均値の結果であるが、その値は SpRhh, SpLhh と比べて非常に低くなっている。

図6は各カーネルパターンにおける SpLah (3画像平均値) を示す。各表示位置は図4のパターンの位置に対応している。いくつかで SpLah 値が高くなるものがあることがわかった。また、サイズ 0.5F x 0.5F と 1F x 1F の場合で高い SpLah 値となる傾向が得られた。

さらに、ロボットや対象物の領域をトップダウン処理による特徴パターンとして与え、カーネルによるボトムアップ処理と統合して Scanpath を検出したときの aROIs を同様に hROIs と比較して SpLah を求めた。図5のそれぞれ右端の "SpLah (BU&TD)" の棒グラフはその結果である。これらは各カーネルにおいて得られた値の平均値を示す。

なお、特徴マップ線形結合の重み係数について、ここではトップダウン処理の影響を知るために、トップダウン特徴全体と各カーネルパターンとで 9:1 の比率とした。各トップダウンのパターンの中の各係数も各画像において異なる値とした。それは以下の相手ロボ

ット動作の理解のための推測に基づいて設定した。

- ・ 観察の最初 (Image1) では、まず環境内のロボットと対象物を検出する必要がある。そのためにロボットおよびペンの特徴パターンに均等に重みを与えた。ただし、手先の回転は考慮しなかった。
- ・ その後の観察 (Image2) では、アームがどのように動作し、対象物に向かっているかを知る必要がある。そのために車輪の特徴領域は考慮せず、代わりに手先の回転に他より大きい重みを与えた。ただし、ここではまだ手先の 90° 回転は考慮していない。
- ・ 手先が対象物に接近したときの観察 (Image3) では、どのように対象物を操作するかを知ることが重要となるため、手先と対象物との幾何学的関係を知る必要がある。よって対象物と手先の回転のみに重み係数を与え、その他の特徴は考慮しなかった。

本実験で得られた SpLah 値は、画像処理手法による hROIs の検出性能を示すと考えることができる。図 5 の“SpLah (BU)” のグラフで示した SpLah 値が低くなったのは、使用したカーネル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。トップダウンを示す特徴も考慮した“SpLah (BU&TD)”の結果はその値より大きく、SpLhh 値に近い結果となった。これより、トップダウン処理を加えることで有効な Scanpath 検出が可能であることがわかる。また、図 6 の結果から、サイズの小さいカーネルは高い SpLah 値を示しており、カーネルのサイズが重要であることがわかる。1F × 1F のパターンは中心窩サイズのボトムアップ処理を表すものであり、その点でも妥当性があると言える。

(3) ロボット動作検出への応用

本研究で使用したカーネルパターンをロボット手先動作の検出に応用した。図 7 にその結果の一例を示す。この結果は観察ロボットにより記録された手先動作動画の各画像に対して、“Concentric 1F×1F 3cycles” のカーネルを用いて aROIs を検出し、その aROIs の変化から移動ベクトルを抽出したものである。図 7 では、動画像の中の 6 枚の画像についての aROIs と移動ベクトルを示している。いくつかの移動ベクトルはロボットの腕および手先の動作を示しており、有効な動作検出が行われたことがわかる。

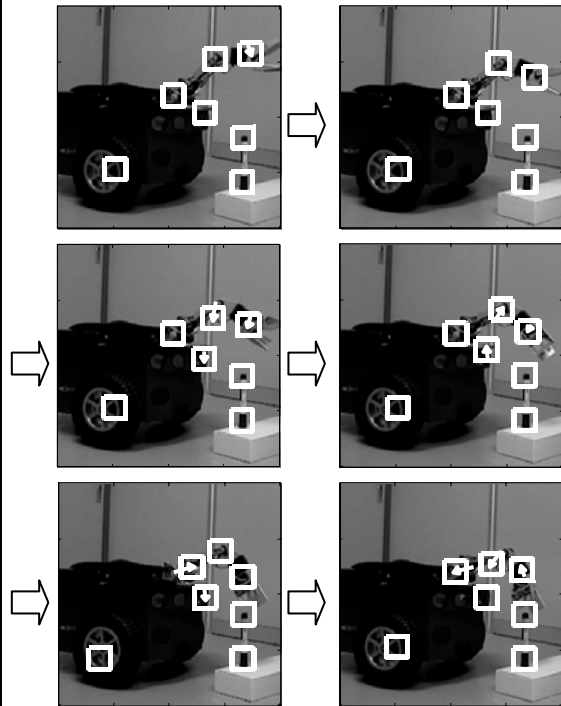


図 7 ロボットハンド動作の検出

(4) まとめ

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 Scanpath 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。その結果、ボトムアップ処理とトップダウン処理を表す特徴を統合することで有効な Scanpath 検出が可能となり、基本的な動作検出も可能となることを示した。この手法を発展させることで、今後のロボットの注視機構の実現およびその応用による他者動作認識の実現が期待できる。そのために、今後さらに様々な画像を使用するなどして人間の視覚特性を検証するとともに、その結果を活用してボトムアップ・トップダウン処理手法を発展させていきたいと考えている。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕 (計 2 件)

① Toyomi Fujita, Claudio M. Privitera, “Positional Features and Algorithmic Predictability of Visual Regions-of-Interest in Robot Hand Movement”, Journal of Robotics and Mechatronics Vol.21 No.6, pp.765--772, 2009, 査読有

②藤田豊己, ``作業移動ロボットの手先動作時の関心領域の特性と検出'', 東北工業大学紀要 I 理工学編 第29号, pp. 31-37, 2009, 査読有

[学会発表] (計5件)

① Toyomi Fujita, "Bottom-up Regions-of-Interest in Observation of Robot Hand Movement: Comparisons with Human Experiments", The 2010 IEEE International Conference on Systems, Man, and Cybernetics(SMC 2010), 2010年10月12日, Istanbul, Turkey

② Toyomi FUJITA, "Loci-Similarity and Algorithmic Predictability of Visual Regions-of-Interest in Robot Hand Movement", The 2010 International Symposium on Intelligent Systems (iFAN 2010), 2010年9月26日, Tokyo Metropolitan University, Tokyo

③藤田豊己, "移動ロボット手先動作時の関心領域特性と検出", 日本機械学会ロボティクス・メカトロニクス講演会 2010 (ROBOMEC2010), 2010年6月16日, 旭川大雪アリーナ(旭川市)

④千葉和也, 藤田豊己, "ロボット手先動作画像観察時の注視位置特性", 第19回インテリジェント・システム・シンポジウム (FAN Symposium '09), 2009年9月17日, 会津大学

⑤藤田豊己, "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", FIT2009 (第8回情報科学技術フォーラム), 2009年9月4日, 東北工業大学

[その他]

ホームページ等

http://www.tohtech.ac.jp/~elecs/ca/fujitalab_hp

6. 研究組織

(1) 研究代表者

藤田 豊己 (FUJITA TOYOMI)

研究者番号 : 90293141