

共同注視における意図理解の構成論的モデル

金野 武司, 橋本 敬

北陸先端科学技術大学院大学 知識科学研究科

概要: コミュニケーションの原型的能力として注目される乳幼児の行動の1つに共同注視行動がある。この行動の捉え方として、乳幼児が他者の意図を理解し、対象を主体的に見ることが重要であるという指摘がある。そこで本論では、共同注視行動の主体的側面の理解を目的として、構成論的アプローチを用いた計算論的なモデルの構成を試みる。このモデルは、親の顔やオブジェクトを注視する行動を強化学習によって獲得する。その獲得過程で、体験する親の視線方向とオブジェクトの配置方向を頻度分布に蓄積し、この頻度分布を用いて親の視線方向からオブジェクトの配置方向を連想する。このモデルをコンピュータシミュレーションに実装した結果、親の視線方向からオブジェクトを連想することで、親の視線が指し示す方向にある視界外のオブジェクトを注視できるようになることが確認された。

キーワード: 共同注視, 他者の意図理解, 乳幼児の主体性, 構成論的アプローチ

Developmental Construction of Infant's Joint Visual Attention as Understanding Intentionality

Takeshi Konno and Takashi Hashimoto

School of Knowledge Science,

Japan Advanced Institute of Science and Technology

Abstract: Joint visual attention is thought of as a basis of infant's communication ability. The importance of subjectivity of the joint attention is pointed out. Namely, infants understand others' intentions and subjectively gaze at the others' gazing points. In this paper, for understanding the subjective aspect of the behavior, we constructed a computational model with constructive approach. The model acquires behavior of gazing at mother's face and objects through reinforcement learning. In the process of acquisition, the model stores the time series relations between mother's eye directions and placement direction of objects as a frequency distribution. Using the distribution, the model associates mother's eye direction with the placement direction of objects. As results of computer simulations, we confirmed that the model could acquire the behavior of joint visual attention. Furthermore, the model can gaze at an object beyond sight following the association.

Keywords: Joint Visual Attention, Understanding Others' Intentions, Infant's Subjectivity, Constructive Approach.

1 背景

乳幼児が見せるコミュニケーションの原型的能力の1つに共同注視行動がある。これは、他者が見る方向を乳幼児も見るという単純な行動である。しかしこの行動を、他者が注意を向けるものと同じく注意を向ける能力の前身であると捉えるとき、それはコミュニケーションの成り立ちを理解する上での重要な行動になる。Butterworth と Jarrett はこの共同注視行動を統制さ

れた状況下での行動観察から理論化し、それを生態学的 幾何学的 空間表象的段階を経る発達過程として記述した^[1]。この記述からは、乳幼児が空間的な認知能力を形成していく側面のみを捉えがただが、コミュニケーションの原型的能力として共同注視行動を見るときには、その背後にある主体性の形成過程にこそ注目する必要がある。つまり、共同注視行動が生態学的な段階として生得的に特徴付けられた反射的な行動で

あるとするところから、やがて自ら見ようとするものを見るという主体性によって発する行動へと変化していく中で、幾何学的あるいは空間表象的段階が表出すると捉えるのである。Tomasello はこの過程を、その主体性を基盤とした他者の意図理解の過程として捉えることを試みながら、心の理論 (Theory of Mind) へとつながる発達過程を描き出している [2]。ただし、他者の意図を理解する基盤が、すでに存在する心的原形によるものであるとするところには異論があり、Corkum と Moore は、生態学的な共同注視行動が、結果的に心的基盤を形作っていくのだと主張している [3]。このように、乳幼児の内部モデルとして主体性や意図の存在を仮定することの重要性は認識されながらも、それを検証する心理学的な実験においては、乳幼児の内部状態を直接覗き見ることができないことに起因して、客観的な考察を進めることが難しくなっている。

こういった状況を補い得る手法に構成論的アプローチがある [4]。これは、対象とする現象のモデルを実際に作って動かす手法である。これにより、構成するモデルを乳幼児の行動と比較することはもとより、乳幼児の内部に相当する振る舞いを観察することが可能になる。この構成論的アプローチの中で長井らは認知発達ロボティクス [5] の実現として共同注視ロボットを開発している [6]。この共同注視ロボットには、生得的な反射回路 (比例制御器) から教示を受ける学習器が、親の視線と自分の運動方向の対応関係を学習していくモデルが実装されている。しかし、親の視線が直接的に運動方向を導くモデルでは、ロボットが見ようとするもの、つまりはその主体性が表現できないのではないかと考える。そこに必要なのは、生得的な行動傾向によって共同注視状況を経験した結果として、例えば親が向いた先に何かおもしろいものがあると予想するようになり、その予想したものを見ようとしてその視線方向に自らの視線を向けるというようなモデルなのではないだろうか。

そこで本論では、まず視界に映る親の顔やオブジェクトを捉えようとする生得的な知覚傾向を仮定し、それによって捉えた親の顔やオブジェクトを視界の中央に収める運動方略を試行錯誤的な学習によって獲得するモデルを構成する。ここで試行錯誤的な学習を用いるのは、何かおもしろいものというのが、経験によって形成される非常に任意性の高い信号と推測されるため、そういった信号に結び付いた運動方略の獲得には、

反射回路からの教示的な学習よりも試行錯誤的な学習の枠組みが適していると考えるからである。次に、その運動中に体験する時系列関係性を頻度分布に蓄積しながら、その蓄積した頻度分布を用いた確率的な連想のモデルを構成する。これによって、構成するモデルが親の視線方向に置かれたオブジェクトを連想し、それを見ようとして共同注視行動を表出することを期待する。

具体的には、2 節において構成するモデルを説明し、3 節でそのモデルを使った実験および結果を示す。4 節で得られた結果を考察し、さらに 5 節において本論の結論と今後の課題を明らかにする。

2 モデル

共同注視行動に主体性を見るモデルがどのように構成できるかを検討するために、柔軟な環境設定を可能にするコンピュータシミュレーションを利用する。また、現実との接点を考慮しながらも、できるだけ抽象化された単純な環境およびモデルを構成する。そしてこの環境の中に子どもや親のエージェントとオブジェクトを用意し、その子エージェントに対して果たすべきタスクを設定する。

以下では、まず構成する環境とそこで設定するタスクを説明する。その上で、子エージェントのモデルをどのように構成するかを説明する。

2.1 環境とタスク

2次元仮想世界での情報表現：

子エージェントの網膜を中心とした半径 $1[m]$ の球を考える。その球面上の一点を視界中心としたとき、球の経線に沿って x 軸を、緯線にそって y 軸を置く。さらにその視界中心から x, y 軸両方向に $\pm 0.5[m]$ の幅で子エージェントの視界を設定する (Fig.1)。親の顔やオブジェクトはその球上に配置され、子エージェントから見る奥行き情報は考慮しない¹。よって、子エージェントが見る視界情報は 2次元で表現される。また、この視界は一定速度でのみ動かすことが可能であると、その速度を $0.5[m/sec]$ とする。これは視界の端から端までを $2[sec]$ で移動する速度である。また、コン

¹奥行き情報の座標変換は大きな問題だが、今は子エージェントが扱う状況の簡略化を優先する。

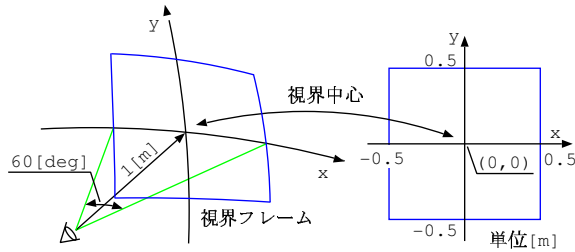


Fig.1: 視界の設定.

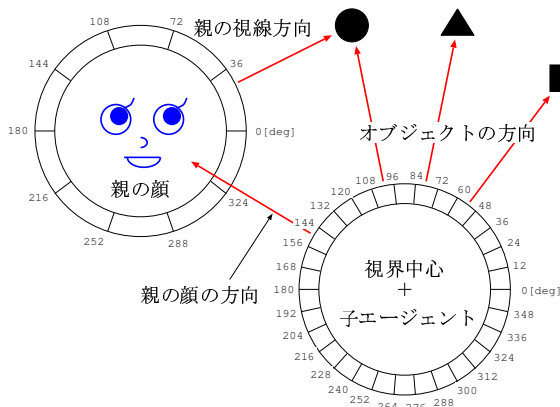


Fig.2: 視界から得る情報. 子エージェントは、特徴とその配置方向の情報を視界から得る。例として、親の顔からは、特徴として $0 \sim 36[\text{deg}]$ の視線方向、配置として $144 \sim 156[\text{deg}]$ の配置方向を得る。また、オブジェクトからは特徴としての形状、配置として $96 \sim 108[\text{deg}]$ の配置方向を得る。

コンピュータシミュレーションは離散時間で進行するとし、そのサンプリング時間を $10[\text{msec}]$ とする²。これにより、1 サンプリングの移動距離は $d = 5[\text{mm}]$ となる。

子エージェントが視界から得る情報：

子エージェントが視界から得る情報は、親の顔、もしくはオブジェクトの特徴とその配置である。親の顔の特徴を視線方向とし、オブジェクトの特徴を形状³とする。配置として取得できるのは、子エージェントの視界中心に対する極座標での方向とする (Fig.2)。それぞれの情報の取得分解能は Table.1 に示すとおりである。

²眼球運動と頭部回転の中間的な特徴を実現した結果としてこのような設定にする。実際に眼球運動だけを考えれば $0.5[\text{m/sec}]$ は遅すぎるし、遠心性信号のフィードバック時定数を考えれば $10[\text{msec}]$ のサンプリング時間はひと桁短い。

³便宜上、形としただけで、これは色でも大きさでも良い。

Table.1: 視界から得る情報

特徴	親の視線方向 (極座標)	$36[\text{deg}]$ の分解能
	オブジェクトの形状	の3つ
配置	見える方向 (極座標)	$12[\text{deg}]$ の分解能

る。また、ヒトの網膜情報は視界の端と中央でその解像度が異なることが知られているが^[7]、子エージェントの視界内ではその特徴および分解能は変化しないものとする。

ここで設定する情報表現は、主に形式的な処理に対する必要性から決定されている。よって、神経生理学的にどのような情報表現が妥当かといったことなどに関しては別途検討が必要である。

子エージェントのタスク：

子エージェントのタスクは、オブジェクトや親の顔を視界の中央に収めることとする。オブジェクトや親の顔が視界の中央に収められたと判断するのは、それぞれのターゲットが視界中心の半径 $25[\text{mm}]$ の円内に収まったときとする。

2.2 子エージェントのモデル

子エージェントは、視界から得る親の顔やオブジェクトの特徴とその配置から、それらを視界の中央に収めるべく運動方向を試行錯誤する。この試行錯誤を実現するために、リファレント報酬付きテーブル型 Sarsa 方式^[8]として知られる強化学習の枠組みを用いる。強化学習は基本的に自分の状態 (s) に対する行動 (a) に価値 (Q) を割り当てて、ある行動を起こしたときに得られる報酬 (r) によってその行動の価値を形成する枠組みである。このとき、その報酬の定義には、見ようとするものを視界の中央に収めようとする傾向と、それが実現できているかどうかを評価する機能が必要になる。そこで本論では、これを実現するために2つの生得的機能を仮定し、前者を偏向器、後者を評価器として実装する。

以下ではその生得的仮定と、その実装である偏向器および評価器を説明し、その後で強化学習の実装を説明する。

偏向器：

視界の端で動くものには反射的に反応してしまうような生得的な傾向が我々には備わっている^[9]。こういったレベルで既に獲得した機能として、ここでは単純に視界内で最も端にあるものが見るものになる機能を仮定する。この実現として、視界に表示される親やオブジェクトは、偏向器を通すことによって一番端にあるものが視界中心に置かれるものとする (Fig.3)。

評価器：

ももとの視界 (Fig.3:左) と、見るものが視界中心に移動した視界 (Fig.3:右) を比較したとき、その画像的な違いが一致度の情報を含んでいることになる。そこで、その一致度を評価値として認識する機能を生得的に持つことを仮定する。ここで、ある時刻 t での評価値 (E_t) は、視界から得るオブジェクトや親の顔の配置方向の変化 $\Delta\theta_t$ が運動方向の最小分解能である $\pm 12[\text{deg}]$ 以内に入っていれば正の値 (1) とし、そうでなければ負の値 (-1) とする。

$$E_t = \begin{cases} 1 & \text{if } -12 \leq \Delta\theta_t \leq 12[\text{deg}] \\ -1 & \text{otherwise} \end{cases} \quad (1)$$

運動器 (強化学習)：

偏向器によって見るものが決まるとき、その見るものの特徴と配置の情報を合わせて状態 s とする。それが視界の中央に移動しているかどうかを評価器が (1) 式によって判断するとき、その報酬を (2) 式に従って決定する。

$$r_t = E_t - \tilde{r}_t. \quad (2)$$

\tilde{r}_t がリファレント報酬であり、これは (3) 式によって更新される。

$$\tilde{r}_{t+1} = \tilde{r}_t + \alpha_r [r_t - \tilde{r}_t]. \quad (3)$$

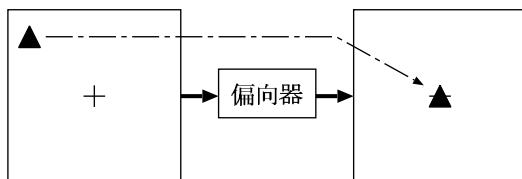


Fig.3: 偏向器の機能

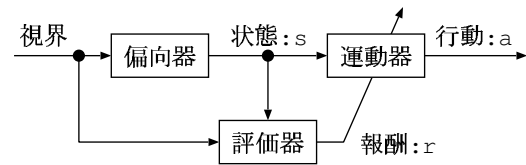


Fig.4: 子エージェントのシステムブロック図

ここで α_r はステップサイズパラメータである ($\alpha_r = 0.01$)。

状態 s に対する行動 a は、その価値 Q で形成される確率分布から (4) 式によって選択され、その分布は (5) 式によって更新される。

$$p_t(a|s) = \frac{e^{Q_t(s,a)/\tau}}{\sum_{a'=1}^{N_a} e^{Q_t(s,a')/\tau}}, \quad (4)$$

$$Q_{t+1}(s, a) = Q_t(s, a) + \alpha_Q [r_t + \gamma Q_t(s_{t+1}, a_{t+1}) - Q_t(s_t, a_t)]. \quad (5)$$

ここで τ は温度係数、 α_Q は学習率、 γ は割引率である ($\alpha_Q = 0.1, \gamma = 1$)。 τ は試行回数の経過と共に減少させるパラメータなので、その設定は実験の中で詳述する。また、 N_a は極座標で表現される運動方向の分割数である ($N_a = 360/12 = 30$)。

子エージェントのシステム：

ここまでに説明した子エージェントのモデルをシステムブロック図にしたのが Fig.4 である。

3 実験と結果

3.1 実験 1：運動の形成

子エージェントの視界の中で、オブジェクトと親の顔を交互に表示し、その中で見るもの (状態 s) に対する視界中心の移動方向 (運動 a) がどのように形成されるかを確認する。また、このときのオブジェクトは親の見ている方向に置く。さらに、親の顔は最初に子エージェントの視界中心に置き、その後は常に同じ位置に置く。そしてオブジェクトは親の顔を中心とする半径 200[mm] の円周上に配置する。このオブジェクトの配置は、子エージェントがオブジェクトを視界中心で捉えたときにも、親の顔が視界から外れない位置である⁴。

⁴200[mm] の半径：視界幅 (=1000[mm]) の 20%

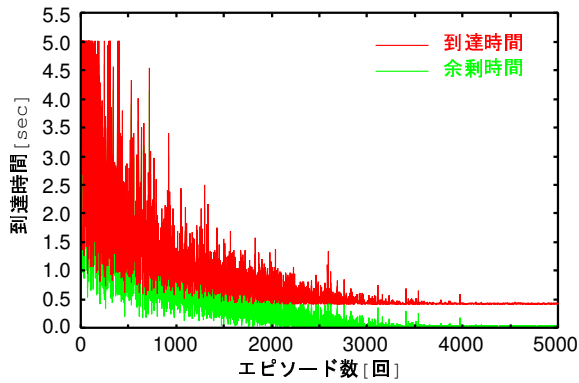


Fig.5: 移動時間の変化特性

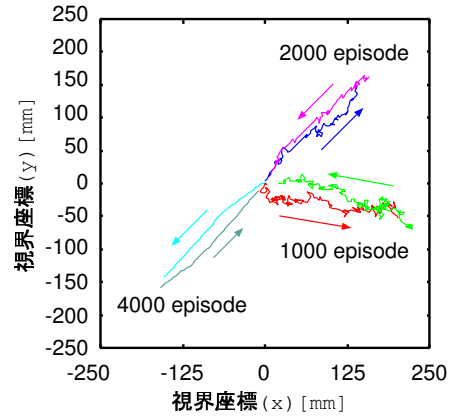


Fig.6: 視界中心の移動軌道例

3.2 結果 1

親の顔またはオブジェクトを視界中心で捉えようとする試行を 1 エピソードとして、それを 5000 回繰り返し続けたとき、その移動時間は Fig.5 のように変化する。Fig.5 においては、5[sec] を 1 エピソードの制限時間とした。制限時間に達したときには、その試行はうまくいかなかったものとして次の試行に移行させた。余剰時間として示したのは、掛かった時間から視界中心を最短距離で移動させた場合の時間を引いたものである。ここから、5000 エピソード後にはほぼ最短時間で親の顔やオブジェクトを視界中心に収めることができるようになってきていることが分かる。ここで、(4) 式で説明を保留した温度係数 τ は、初期エピソードで 0.8 からスタートし、5000 エピソードで 0.1 となるように線形的に減少させた。また、各エピソードによる段階的な視界中心の移動軌道は Fig.6 のようになる。ここから学習が進むにつれて、移動軌道が徐々になめらかになる過程が確認できる。

3.3 エージェントモデルへの頻度分布形成機能の追加

親が見る方向にオブジェクトが配置されている状況を体験し続けるとき、運動器に入力される状態 (s) の遷移過程を記録すると、その内部情報として親の視線方向の次に来るオブジェクトの配置方向は、親の視線方向に対応した方向になるはずである。よって、体験する状態 (s) の遷移過程を頻度分布として蓄積して、親の視線方向から次に来るだろうオブジェクトの配置方

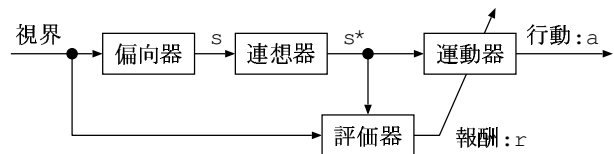


Fig.7: 子エージェントシステムへの連想器の追加

向を連想的に予期すれば、運動器は連想したオブジェクトの配置方向に対応した方向へ動き出すことができるはずである。その結果として、視界内にオブジェクトが存在しない場合でも、親の視線方向から視界の外へオブジェクトを探しに行く行動が実現できるのではないかと推測する。

そこで、この考えを連想器として実装し、視界外へのオブジェクトの追視が可能になるかどうかを確認する。

連想器：

見るものが切り替わったときの状態 (s) の遷移過程 ($s \rightarrow s^*$) を頻度分布として $F(s, s^*)$ に蓄積する。そして、その頻度分布を使って、状態 (s) から運動器に入力する状態 (s^*) を (6) 式によって確率的に決定する。

$$p(s^*|s) = \frac{F(s, s^*)}{\sum_{s'=1}^{N_c} F(s, s')} \quad (6)$$

ここで N_c は親の顔およびオブジェクトの特徴と、その配置方向を合わせた状態数の最大値である。

この連想器の中に、試行中には見るものを変更しない機能を仮定する。これにより連想を行なうのは、見るものが視界中心に収められるか、もしくは設定する試行限界時間 (= 5[sec]) に達したときとなる。これは本来であれば、感覚器が受け取る刺激が刻一刻と変化

する中で、その変化に直ちに反応して見ようとするものを変えるか、もしくは邪魔が入っても見ようとするものを見続けるかを判断するような機能として考えるべき問題である。しかし、本論では上述のような仮定を置き、まずは視界外への追視が可能になるかどうかの検証を優先する。また、この実装には、見ようとするものが同じものであるかどうかを判断する必要があり、これも本来であれば体験する時系列データから形成されるべきものだが、ここではそれを既知っているものとして扱う。

3.4 実験 2 :

頻度分布の形成と視界外への追視

実験 1 と同じ状況を体験するとき、Fig.5 に見る運動結果は、その初期段階では運動方策が未熟なために、設定する試行限界時間に達してしまう場合があることを示している。そこで、親が見る方向に配置されたオブジェクトを見ることができた関係のみを蓄積するために、その試行限界時間に達することがなくなる 1000 エピソード以降の状態 (s) の遷移過程を頻度分布として蓄積する。その後、5000 エピソードまで蓄積した頻度分布を用いて、視界外への追視を試みる。このとき、オブジェクトは視界の対角線距離の半分の 1.2 倍 ($= 1.2 \times 500\sqrt{2}[mm]$) の円周上に配置する⁵。

3.5 結果 2

親の顔を視界の中央に収めてからオブジェクトを見ると、状態 (s) は、親の顔の視線方向に対応したオブジェクトの方向特徴に遷移するはずである。これを 1000 ~ 5000 エピソードの間に蓄積された頻度分布で確認したのが Fig.8 である。親の見る方向にオブジェクトを配置しているので、その頻度分布は親の視線方向に対応したオブジェクトの配置方向に集中する。しかし、オブジェクトから親への頻度分布はこのようにはならない。なぜなら、オブジェクトの形という特徴に、親を指し示す情報が存在しないことを仮定しているからである。もし、今の枠組みでオブジェクトから親へ

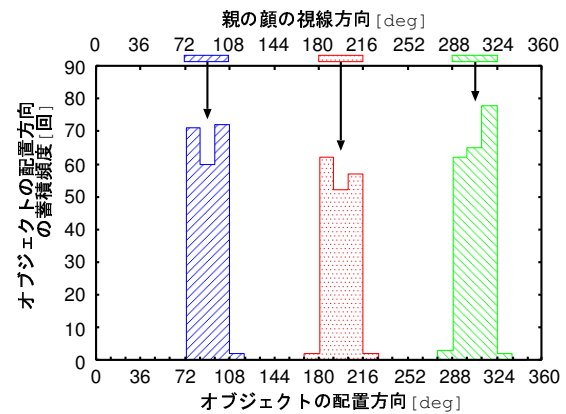


Fig.8: 親からオブジェクトを見たときの状態の遷移頻度分布の例。親の顔の視線方向 (72 ~ 108, 180 ~ 216, 288 ~ 324[deg]) に対して、その次に見たオブジェクトの方向特徴をカウントしたときの頻度分布を示している。

の連想を働かせるのなら、連想器はオブジェクトを見るに至るまでの経過を保存する必要がある。この必要性は、ある場所までの道順を覚えることに良く対応する。親の視線からオブジェクトへ到達するのは、目的地がランドマーク的な対応関係を持っている状況に相当し、オブジェクトから親へ到達するのは、自分の移動方向と距離の履歴から、目的地に着いたときに、出発点からの方向を計算する状況に相当する。これが共同注視行動においてそのまま使われているとは考えにくい。現状のシステムでオブジェクトから親へ正しく戻ることを実現するには、移動履歴を保存して、オブジェクトへ向かったときの親からの正しい方向を知る必要がある。これに対して、視界中心の移動軌道が Fig.9 の経路を辿った場合を考える。ここで、親の顔

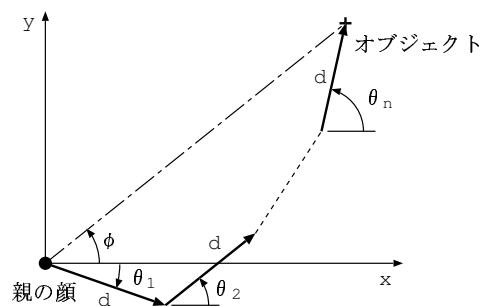


Fig.9: 視界中心の移動軌道

⁵親の視線方向の分解能である 36[deg] に沿って視界中心を移動させたとき、0[deg] 方向に置かれたオブジェクトは視界の対角線距離の半分の 1.68 倍に置かれたときに、一度も視界に入らない状態になる。これを考慮して 1.2 倍のオブジェクト配置半径とした。

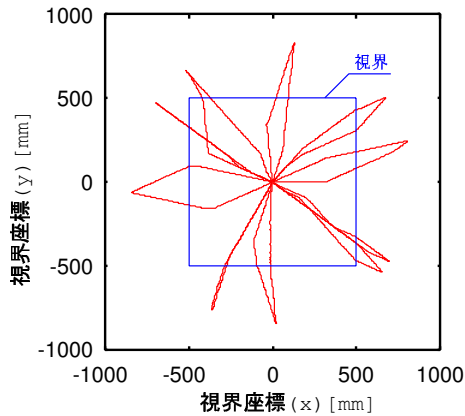


Fig.10: 視野外への追視軌道例. 親の視線方向内に一様乱数でオブジェクトを配置した場合の視界中心の移動軌道例。

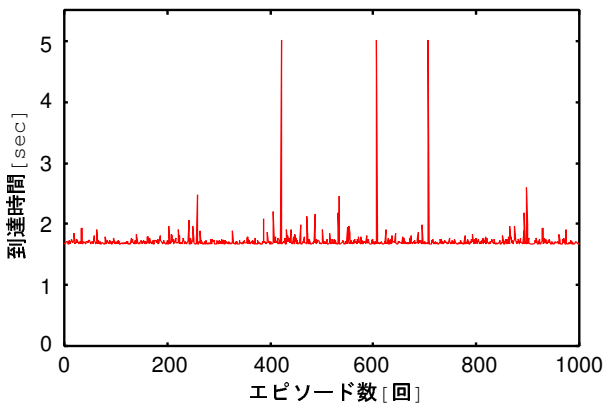


Fig.11: 視野外への追視時間特性

からオブジェクトへ視界中心が移動したとき、子エージェントは移動中の $\theta_1, \theta_2, \dots, \theta_n$ を保持していることを仮定する。この状態でオブジェクトを視界中心に収めると、親から見たオブジェクトの方向 (ϕ) は (7) 式で計算できる⁶。

$$\phi = \tan^{-1} \left(\frac{\sum_i \sin \theta_i}{\sum_i \cos \theta_i} \right). \quad (7)$$

これをモデルに実装して視野外へオブジェクトを配置すると、Fig.10 のような交互追視軌道が実現できる。また、これを 1000 エピソード実施したところ、追視時間は Fig.11 のように変化した。

この中で失敗する 3 回の事例を解析すると、その原因は 2 つあることが分かった。1 つは視線方向の分解

⁶距離情報が不要なのは、視界中心の一時刻あたりの移動距離が $d = 5[mm]$ と一定のためである。

能の中に運動軌道が収まらないケースがあるためであり⁷、もう 1 つが、4000 回のエピソードでは必要な頻度分布が一度も埋まらない場合が有り得るためである。前者は視線方向の認識に、後者は頻度分布の形成にそれぞれ汎化の能力が必要なのではないかと考えている。

4 考察

視界に映る親の顔やオブジェクトをその中央で捉えるというタスクを子エージェントに課したとき、試行錯誤的な学習によって、正しい運動方策を獲得できることが確認された (3.2 節)。また、この運動方策を獲得する過程で体験する共同注視関係を頻度分布として蓄積し、親の視線方向から次に見るだろうオブジェクトを確率的に連想することで、視野外への追視が可能になることが確認された (Fig.10)。この結果から、親の視線方向に形状の異なる複数のオブジェクトが存在する場合には、親との関係の中で高い頻度で現れた特徴を持つオブジェクトを見ようとするのが推測される。また、その状態は見ようとするオブジェクトを子エージェントが明確に持っていることを表現しているため、これが主体性の一部を表現し得るのではないかと考えられる。

しかし、逆にオブジェクトから親の顔を連想する場合には、オブジェクトの特徴が親を指し示す情報を持っていないので、正しい方向への運動には何らかの情報が必要となる。本論ではシステムの枠組みを変えずに行動経路を保存することによってこの問題に対処したが、本来は親からの働き掛けとして声が掛けられたり、視界に介入したりすることの方が頻繁に起こることだと考えられる。また、親を見に戻る場合には、それまでの行動履歴よりは筋肉の状態などの自己受容感覚を用いる方が共同注視行動のレベルでは現実的な対処方法だと考えられる。さらに、見ようとしたオブジェクトが見つけれない状態は、単なる制限時間によって知るものではなく、その見ようとするオブジェクトがどれくらいの移動量および時間で見えるはずだ、という予期を含むことが必要になるものと思われる。また、大きな問題として、連想器には一度見ようとしたものを試行中には変化しない機能や、その見ようとするものの同一性 (カテゴリ) を既に知っている状態を

⁷この可能性は Fig.8 の視線方向範囲から外れた頻度の存在から理解できる。

仮定しているが、これは本来は、経験によって形成すべきものである。

全体として、偏向器および評価器に導入した生得的仮定と試行錯誤的な学習によって、オブジェクトや親の顔を視界の中央で捉えるような定位の能力を実現したとき、その時系列関係性の蓄積と利用が子エージェントの主体性を表現し、その中で共同注視行動を表出し得ることが示唆された。これは共同注視行動を主体性の中で獲得する過程で意図を推論する視点が芽生えることによって、他者の意図を理解するようになり、ひいてはそれが心的基盤の形成へつながるといふ Corkumらの発達因果方向性^[3]を支持するモデルになっていけるのではないかと考えている。そして、こういった発達観を基礎とする中で、モデル内の観察を可能にする利点は、他者の意図を理解していく過程を解明しようとする乳幼児の心理学的実験による知見に、相補的な貢献を可能にするのではないかと期待している。

5 結論

本論では、生得的に持つと仮定する知覚傾向とその評価能力から、行動方略を試行錯誤的に獲得するモデルを構成した。さらに、その過程で得る時系列関係性を頻度分布に蓄積し、親の視線からオブジェクトを連想することで、視界の外への追視を実現するモデルを構成した。またこのモデルにより、オブジェクトの注視から再度、親の顔を見に戻るような交互凝視の行動を実現できることも確認できたが、現段階で提案したモデルでは、連想器におけるカテゴリ形成の問題や、奥行き知覚に対する運動器での座標変換モデルの構築、あるいは視界内で見ることのできる特徴分解能の問題など、いくつかの課題が浮彫りになっている。また、抽象化モデルが持つ基本的な環境設定などの妥当性に関しても、より細かな検討が必要である。

こういった問題点の解決として、ひとつには形式的な枠組みでの実装に問題の原因を見て取れることから、運動器や連想器をニューラルネットワークで実装し、その汎化機能を利用することを検討している。また連想器に関しては、リカレントニューラルネットワークを用いた時系列予測機能の利用を併せて検討している。さらに、こういった汎化能力の利用に伴って、今後は実画像を扱うような、より実環境に近い状況への拡張を行なっていく必要がある。

参考文献

- [1] Butterworth,G.E. and Jarrett,N.L.M., “What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy,” *British Journal of Developmental Psychology*, Vol.9,pp.55-72,(1991).
- [2] Tomasello,M., “Joint attention as social cognition,” In Moore,C. & Dunham,P.J.(Eds.), *Joint Attention: Its Origins and Role in Development*, Lawrence Erlbaum, pp.103-130,(1995), (訳書:ジョイント・アテンション, ナカニシヤ出版,pp.93-117,(1999)).
- [3] Corkum,V. and Moore,C., “Development of joint visual attention in infants,” In Moore,C. & Dunham,P.J.(Eds.), *Joint Attention: Its Origins and Role in Development*, Lawrence Erlbaum, pp.61-83,(1995), (訳書:ジョイント・アテンション, ナカニシヤ出版,pp.57-76,(1999)).
- [4] 金子 邦彦, 津田 一郎, 複雑系のカオスのシナリオ, 朝倉書店,(1996).
- [5] Asada,M., MacDorman,K.F., Ishiguro,H.,and Kuniyoshi,Y., “Cognitive developmental robotics as a new paradigm for the design of humanoid robots,” *Robotics and Autonomous Systems*,Vol.37,pp.185-193,(2001).
- [6] Nagai,Y., Hosoda,K., Morita,A., Asada,M., “A constructive model for the development of joint attention,” *Connection Science*,Vol.15,No.4,pp.211-229,(2003).
- [7] 喜多 伸之, Rougeaux,S., 國吉 康夫, 坂根 茂幸, “仮想ホロボタを用いた実時間両眼追跡,” 日本ロボット学会誌, Vol.13, no.5, pp.683-690, 1995.
- [8] Sutton,R.S. and Barto,A.G., *Reinforcement Learning*,A Bradford Book,MIT Press,cambridge,MA,(1998), (強化学習, 三上貞芳, 皆川 雅章共訳, 森北出版,(2000)).
- [9] Atkinson,J., Hood,B., Wattam-Bell,J.,and Braddick,O.J., “Changes in infants’ ability to switch visual attention in the first three months of life,” *Perception*,Vol.21,pp.643-653,(1992).