

# 知能化空間におけるサービスロボットのための 顔の向きを利用した視野特性と視覚特性に基づく人の視野推定

## Visual Field Estimation for Service Robot Based on Physical and Cognitive Characteristics of Human Vision Using Human Face Direction in Intelligent Space

精密工学専攻 43号 福島拓志

Hiroshi Fukushima

### 1. はじめに

近年、人と共存するロボットが普及している。このようなロボットは、情報を人へ伝達するメディアの一つとして位置づけられる。人からロボットへアプローチすることによって、人はロボットから情報を得ることができる。人からロボットへアプローチする場合、人は自らロボットに対して情報を求めているため、ロボットに注意が向いている。一方で、道に迷った人への道案内サービスなど、人の補助や支援を行うために、ロボットから人へアプローチする場合も想定される。しかし、ロボットから人へアプローチする場合、はじめ人は必ずしもロボットに注意を向けているとは限らない。そのため、何かしらの活動をしている人の目の前からロボットがアプローチすることによってその人の活動を阻害する、人の後ろからアプローチすることによって人がロボットの接近に気づかないという問題が生じる。以上のことから、人と共存するロボットには、人の活動に合わせて適切に人の注意を獲得することが求められる。

ロボットが人の活動を知ることができれば、人の活動に基づきロボットが最適な方法によって人へアプローチすることが可能になる。しかし、ここで重要なことは、ロボットがなるべく人の活動を阻害することなく、ロボットへ対象となる人の注意を向けさせることである。人が何に注意を向けているかをロボットが知ることができれば、ロボットが人の目の前から直接接近して人の活動を阻害する、ロボットが適切に接近しないために人がロボットの存在に気付かないという問題は解決できる。そこで、ロボットが人の注意を適切に獲得するために、人の視野に着目した。

人は視線方向を中心に幅広い視野を持つ。中心視野では文字や物体の識別に優れ、周辺視野では動きの知覚に優れることが知られている<sup>(1)</sup>。以上のことから、ロボットが人の視線方向を知ることができれば、人の視野を推定することができる。さらに、人の視野を推定することができれば、ロボットは上で述べた人の視野特性の広がりやを推定することができる。ロボットがこの視野特性に基づいて人へアプローチすることができれば、人は活動を阻害されることなく自然とロボットの存在を知覚できると考え、本研究では人の活動を阻害することなく視野を推定することを目的とする。

人の視野の位置は視線によって変化するため、視線方向を推定しなければ視野を推定することはできない。そこで、視線方向を推定することを考える。環境に設置されたセンサを使って視線方向を推定することによって、人の活動を阻害することなく視野を推定することが可能である。

現在までに環境に設置されたカメラから得られる目の画像を利用して視線推定を行う研究が数多くある。代表的なものとして、一つは目画像から眼球モデルを作成し、幾何学的に視線方向を推定する研究である<sup>(2)</sup>。この研究はカメラの前に座った人の目画像から取得される虹彩の位置を利用し、幾

何学的に視線方向を推定することを実現している。もう一つは学習ベースで視線情報を得る研究である<sup>(3)</sup>。この研究はノートパソコンを使用する人の顔画像から取得される目や顔の向きと視線方向の関係を学習したデータを使うことによって、照明変動が大きい日常環境において回帰的に視線方向を推定することを実現している。しかし、これらの手法では視線を向ける対象となるディスプレイ上の位置を事前にシステムへ与える必要があるため、視野推定を行う環境が変化するたびにキャリブレーションを行う必要がある。

この課題を解決するために本研究では、動きの知覚に優れた周辺視野内に人の注目を集める要素が出現した場合、人の視線はその要素へ向くという視覚特性と、顔の向きから予測される視野特性に基づいて視線方向を求め、視野を推定するシステムを提案する。この手法では、人が視線を向ける対象となる物体をシステムが能動的に判別するため、事前に視線を向ける対象の情報をシステムへ与えることなく、視野を推定することが可能である。

しかし、視覚特性について定性的な知見は存在するが、定量的に明らかにされている知見はほとんど見当たらない。そこで本研究では、動物体がどのように動くとき人は動物体に対して視線を向けるのかを定量的に評価し、動物体に対する人の視覚特性を明らかにする。そして得られた知見に基づき、人の活動を阻害しない視野推定システムを実現する。

### 2. 視野特性と視覚特性を利用した視線推定

#### 2.1 人の顔の向きの計測と視野モデルの作成

人の視野を推定するためには、その対象となる人が環境内のどの位置にいるのかを知る必要がある。そこで本研究では、上野山ら<sup>(4)</sup>の手法を用いて環境内の人の検出およびトラッキングを行う。計測にはレーザ測域センサ（北陽電機株式会社 UTM-30LX）を使用する。この手法では、検出した人にそれぞれ ID を付加し、自律移動ロボットが人の道案内を可能にするためにその人が道案内を必要としている度合を同時に算出する。道案内を必要としている人の視野を推定することによって、ロボットがその人へ適切にアプローチし、道案内をすることができると考えられる。そこで、本研究では環境内で道案内の度合が最も高い人を視野推定の対象とする。視野推定の対象と判定された人の顔の向きの計測には PTZ カメラ（アクシスコミュニケーションズ社 P5514）を使用する。顔の向きの計測には Cascade Regression Framework による顔特徴点の位置合わせを利用した Zface<sup>(5)</sup>を使用し、カメラ映像における顔の位置、顔の姿勢を得る。

本研究の先行研究<sup>(6)</sup>では、心理学的知見を基に、計測された顔の向きから推定される物体の識別能力の強度（識別能力強度）と動きの知覚能力の強度（知覚能力強度）の分布をモデル化した。顔の向きを視野の中心として、識別能力の強度は視野の中心付近で値が 1 となり、周辺に向かうにつれて 0

に近づく．一方，知覚能力の強度は中心から左右 35 deg，上 25 deg，下 30 deg 付近で値が 1 となり，視野の中心付近や周辺に向かうにつれて 0 に近づく．視野特性を推定したモデルを Fig. 1 に示す．

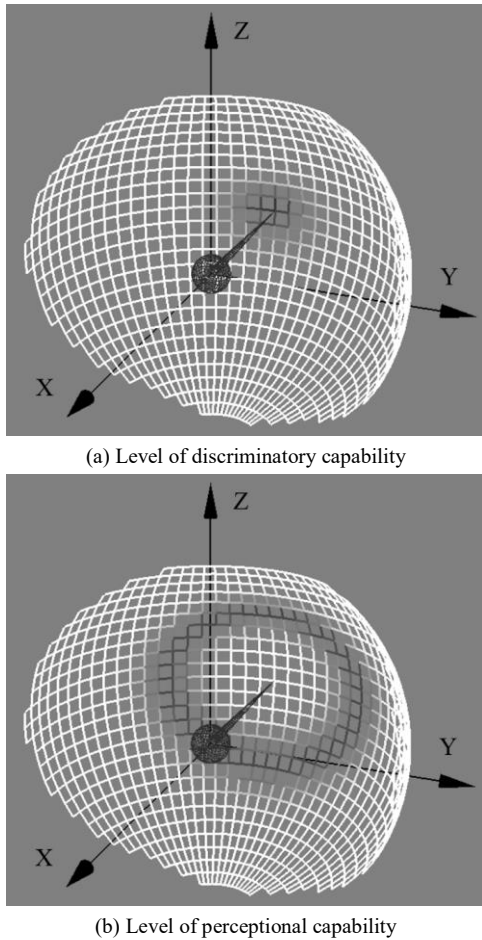


Fig. 1 Visualization of the levels of physical characteristics of vision: (a) Level of discriminatory capability in the central visual field, (b) Level of perceptual capability in the peripheral visual field

図中の球体は頭部を，三角錐は頭部姿勢を表しており，そのときの視野は網目状の領域である．その領域中で色が濃い部分ほど強度が高いことを表している．識別能力強度，知覚能力強度は対象者の位置と顔の向きに基づき，逐次推定される．推定された識別能力強度，知覚能力強度は次の節で説明する視線推定のための注目要素判定に使用される．

## 2.2 動物体に対する人の視覚特性

人は動物体に注意を向けることが可能であり，また動物体は人の注意をひくことが知られている<sup>(1)</sup>．また，人はすでに見たことがある動物体には視線を向けない傾向があるといわれている<sup>(7)</sup>．さらに，人の視線方向を変化させる要因は，人の視野内に出現した動物体の進行方向<sup>(6)</sup>や速度<sup>(8)</sup>であることがわかっている．本研究では，以上のような視覚的な刺激に対する人の視覚的反応の特性を視覚特性と呼ぶ．視覚特性に基づき，2.1 節と同様の手法によって検出された動物体が，人が注意を向け，視線を向ける対象となる注目要素であるかを判定する．

まず，人の視線方向を変化させる要因が動物体の進行方向であることを利用し，動物体の進行方向について評価する．Fig. 2 のように，動物体  $i$  の進行方向を表すベクトルを  $\mathbf{V}_i$ ，動物体  $i$  の位置と視野推定対象者の位置とがなすベクトルを  $\mathbf{D}_i$ ，動物体  $i$  が計測範囲内に現れてから経過したステップ数

を  $t_i$  ( $t_i = 0, 1, 2, \dots$ ) とするとき，二つのベクトルがなす角度  $\theta(t_i)$  を式(1)から求める．そして，式(2)によって動物体の進行方向に関する評価値  $E_a(t_i)$  を算出する．

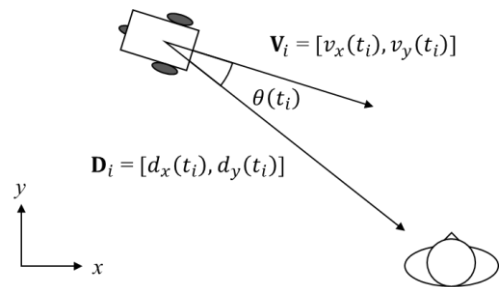


Fig. 2 Relationship between the direction of a moving object and a target person

$$\theta(t_i) = \cos^{-1} \left[ \frac{v_x(t_i)d_x(t_i) + v_y(t_i)d_y(t_i)}{\sqrt{v_x(t_i)^2 + v_y(t_i)^2}} \right] \quad (1)$$

$$E_a(t_i) = \frac{180 - \theta(t_i)}{180} \quad (2)$$

ここで， $v_x(t_i)$ ， $v_y(t_i)$  は  $t_i$  における動物体  $i$  の  $x$  方向， $y$  方向の速度， $d_x(t_i)$ ， $d_y(t_i)$  は  $t_i$  における動物体  $i$  と対象者との  $x$  方向， $y$  方向の距離を表す．

次に，人の視線方向を変化させる要因が動物体の速度であることを利用し，動物体の速度を式(3)によって評価する．

$$E_v(t_i) = 1 - \alpha \exp(-\beta \|\mathbf{V}_i\|) \quad (3)$$

ここで， $E_v(t_i)$  は  $t_i$  における動物体  $i$  の速度に関する評価値であり， $\alpha$ ， $\beta$  は設計パラメータである．

検出された動物体を人の注目要素とする判定と，その注目要素と判定された動物体へ対象者が視線を向けるときの視線方向の導出方法を説明する．まず，視野推定対象者と判定された人へカメラを向け，顔の向きを計測する．計測された顔の向きから作成された視野モデルを付加し，対象者の視野特性の分布を表す環境地図を構築する．次に，推定された対象者の視野内に存在する動物体について，注目要素の判定を行う．本研究の先行研究<sup>(6)</sup>における実験結果と和田らの報告<sup>(8)</sup>より，以下の知見が明らかになっている．

- 一度視線を向けた動物体には視線を向けない傾向がある．
  - 人の周辺視野内で  $\theta(t_i)$  が 10 deg 以下の角度で物体が動いたとき，人はその物体に視線を向ける．
  - 人の周辺視野内での物体の速度が上昇するにつれて，その物体へ視線を向ける人の割合が高くなる傾向がある．
- 以上の知見を考慮し，以下の条件を一つでも満たす動物体を注目要素と判定する．
- (I) 対象者の中心視野の識別能力強度が 0.5 以上の位置に存在し，対象者が一度も視線を向けていない．
  - (II) 対象者の周辺視野の知覚能力強度が 0.5 以上の位置に存在する．
    - (A) 対象者が一度も視線を向けておらず，進行方向評価値  $E_a(t_i)$  が 0.6 以上である．
    - (B) 進行方向評価値  $E_a(t_i)$  が 0.95 以上である．
    - (C) 速度評価値  $E_v(t_i)$  が 0.7 以上である．

そして，式(4)から視線方向を算出する．なお，注目要素と判定される動物体が存在しない場合，対象者の注目を集める要素がないとみなし顔の向きを視線方向とする．

$$\theta_g(t_i) = \begin{cases} \theta_f(t_i) - \tan^{-1} \left[ \frac{x_t(t_i) - x_a(t_i)}{y_t(t_i) - y_a(t_i)} \right] & \text{if I or II-A or II-B} \\ \theta_f(t_i) & \text{or II-C is true,} \\ \theta_f(t_i) & \text{otherwise.} \end{cases} \quad (4)$$

ここで  $\theta_f(t_i)$  は  $t_i$  における対象者の体の正面と顔の方向がなす角度,  $x_t(t_i)$ ,  $y_t(t_i)$  は  $t_i$  における対象者の位置,  $x_a(t_i)$ ,  $y_a(t_i)$  は  $t_i$  における注目要素と判定された動物体の位置である。

### 3. 検証実験

提案した視野モデルと視線推定手法の有用性を検証すること, 動物体が人の注目を集める速度で動くとき, その速度を維持する時間が人の視覚特性に影響を与えるかどうかを検証することを目的として実験を行う。

#### 3.1 実験設定

被験者から 2.5 m 離れた位置を左から右へ横切るようにロボット (オカテック株式会社 MEC Robot #001) が動く状況を用意し, その状況下で被験者がロボットへ視線を向けるかを検証する。使用したロボットを Fig. 3 に示す。



Fig. 3 Appearance of the robot used in the experiment

提案した視野モデルがどのような環境下で適用できるかを確認するために, 被験者を頭部の動きを限定したタスクに取り組むグループと頭部の動きが自由であるタスクに取り組むグループの二つに分ける。今回は, 前者のグループにはノートパソコンに表示された文章の黙読, 後者のグループにはトランプを使った七並べをタスクとして課す。また, 音による視線誘導を避けるため, 被験者はヘッドホンを着用する。施行ごとにロボットの速度が上昇するタイミングと速度上昇の継続時間を変更する。各動作を Fig. 4 に示す二箇所の速度変更位置によって決定する。

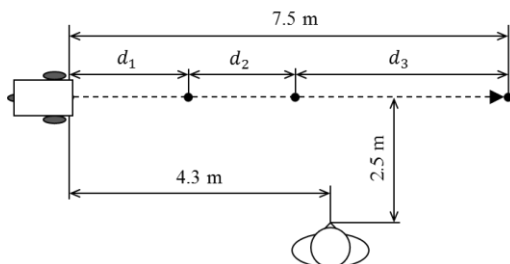


Fig. 4 Condition of the experiment. The robot started the movement from the left front position of the subjects (4.3 m to the left, 2.5 m to the front) and changed the speed at the position  $d_1$  and  $d_1 + d_2$  away from the start position. Then, it stopped at the position 7.5 m away from the start position and backed to the start position.

速度変更位置を変えた計 5 パターンを用意し, パターン 1, 2, 4, 3, 5, 1, 3, 5, 2, 4 の順に走行する。速度変更位置は被験者の視野特性の分布に基づいて決定する。各走行パターンにおけるロボットの速度変更位置を Fig. 5 に示す。

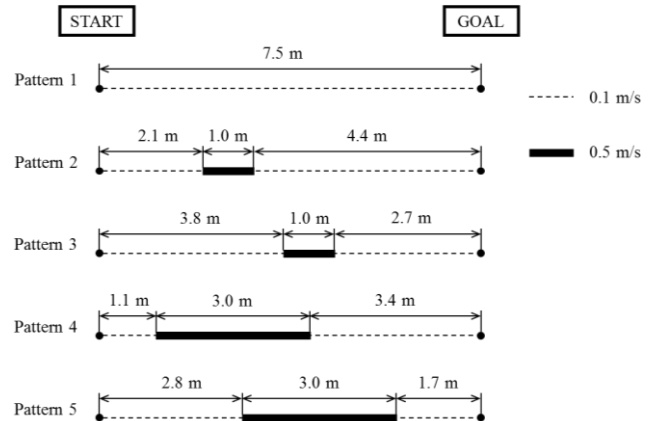


Fig. 5 Condition of speed-changing positions of the robot in each pattern

実験中の被験者の視野特性を確認するために, パターンごとにロボットの右側面に数字が書かれた紙を貼り, 実験後にその数字とロボットの動き方について問うアンケートを用意する。パターンごとの提案手法による推定結果と実際の被験者の視線を比較し, 評価する。実際の視線は, 被験者の顔を撮影した映像と被験者の頭に載せたカメラ映像から求める。式(3)における設計パラメータは  $\alpha=0.41$ ,  $\beta=0.00079$  とする。被験者は 20 代の男性 8 名である。

#### 3.2 実験結果と考察

本稿では, 提案システムによって視線を向けたと推定された被験者の割合, 実際にロボットへ視線を向けた被験者の割合, ロボットが加速してから被験者が視線を向けるまでの時間について考察する。文章の黙読をしたグループにおける結果を Table 1 に, 七並べをしたグループにおける結果を Table 2 に示す。

Table 1 Rate of the subjects who turned their eyes to the robot in each movement pattern when they read sentences on a laptop computer

Movement pattern	Proposed method [%]	Actual gaze [%]
1	38	88
2	100	88
3	0	100
4	100	88
5	38	88

Table 2 Rate of the subjects who turned their eyes to the robot in each movement pattern when they played cards

Movement pattern	Proposed method [%]	Actual gaze [%]
1	38	100
2	88	100
3	13	100
4	100	100
5	13	100

Table 1, 2 から, 両タスクに共通して, ほとんどの被験者がすべての走行パターンにおいてロボットへ視線を向けるという結果が得られた。また, 提案手法による推定結果について, 周辺視野でロボットが加速するパターン 2, 4 において視線を向けると判定され, 中心視野で加速するパターン 3, 5 においては視線を向けないと判定されたシーンが多く見られた。

パターン 2, 4 において, 実際の視線と提案手法による推定結果が一致した。提案手法では, 知覚能力強度が高いと推定された領域においてロボットが加速したために, そのロボ

ットが注目要素と判定され、視線を向けると推定された。実際、被験者はそれまでノートパソコンやトランプに視線を向けていたが、ロボットが加速するとロボットへ視線を向けた。これは、被験者がロボットの加速を周辺視野で知覚してロボットに注意を向けたためであると考えられる。このことから、周辺視野における知覚能力強度の推定結果と視線の推定結果は適切であるといえる。

一方、中心視野でロボットが加速するパターン 3, 5 においても被験者はロボットへ視線を向けた。中心視野と推定された領域においてロボットが加速するシーンにおいて被験者はロボットへ視線を向けたことから、実際は中心視野と推定された領域に知覚能力があったことがわかる。このことから、今回の実験環境では本研究で提案した知覚能力強度の分布と実際の被験者の知覚能力の分布に差異があったと考えられる。提案手法によって推定される視線方向を原点としたときの水平方向の知覚能力強度の分布を Fig. 6 に示す。

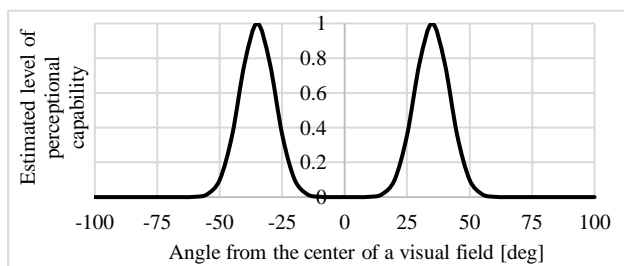


Fig. 6 Distribution of the estimated level of perceptual capability. The angle is expressed with right direction as positive.

提案手法では、Fig. 6 に示す知覚能力強度の分布は視線方向の移動のみに依存して変化するものとしてモデル化した。しかし、タスクに取り組んでいない状態と比較して、黙読や七並べをする場合、被験者は画面やカードに焦点を合わせていると想定される。それに伴い焦点距離が短くなった結果、中心視野が狭まり、知覚能力強度の分布が中心視野側に移動した、もしくは拡大したと考える。そのため、今回の実験環境においては、被験者の中心視野における知覚能力強度がタスクに取り組んでいない状態と比較して高くなり、被験者は中心視野と推定された位置でロボットの加速を知覚することにつながったと考えられる。以上のことから、知覚能力強度の分布は人の焦点距離によって変化すると予想される。したがって、この結果を考慮し、人の顔と物体との距離によって推定される焦点距離に応じて、視野モデルの知覚能力強度の分布を調整することによって、動物体が中心視野に存在するシーンにおける視野推定の精度は向上するといえる。

黙読をしたグループにおけるパターン 5 や七並べをしたグループにおけるパターン 3, 5 では、シーンによって視線推定結果に差が見られた。これは実験中に被験者が頭部を動かしたために、周辺視野における推定された知覚能力強度の位置が変動したことが原因として挙げられる。パターン 3, 5 では、顔が正面を向いている場合は視線を向けないと判定される。しかし、被験者が頭部を動かしたことによって周辺視野でロボットが加速するシーンとなり、視線を向けるという判定結果になった。実際に被験者はそのシーンにおいて視線を向けたことから、適切な視野推定が行えていたといえる。

Fig. 7 に各走行パターン各タスクにおける、ロボットの速度が上昇してから被験者が視線を向けるまでの反応時間を示す。

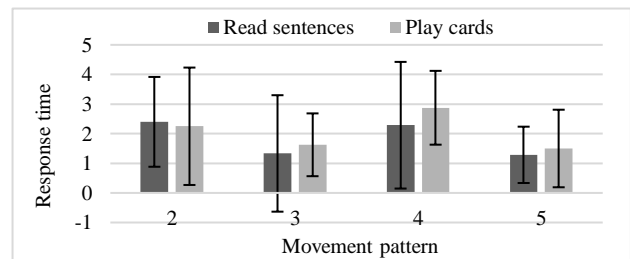


Fig. 7 Response time from the start of increasing speed of the robot until the subject turns his or her eyes

Fig. 7 から、周辺視野において速度が上昇しそれを維持する時間が異なるパターン 2 と 4、中心視野において速度が上昇しそれを維持するパターン 3 と 5 の間に大きな差は見られない。このことから、視覚特性は速度が上昇しそれを維持する時間に影響を受けないことが確認された。

#### 4. おわりに

本研究では人の視野特性と視覚特性を利用して、人の視野を推定する手法を提案した。検証実験により、顔の向きが固定されないタスクに取り組む人に対しても、周辺視野における知覚能力強度が正しく推定されることが示された。中心視野においては実際の知覚能力と提案した視野モデルの間に差が見られたが、人の焦点距離に基づいて知覚能力強度の分布を調整することによって改善できると考える。

今後は推定対象者が静止せず、移動するシーンにおける視野推定の実現を目指す。

#### 参考文献

- (1) 日本視覚学会編, 視覚情報処理ハンドブック, 朝倉書店, 東京 (2000).
- (2) Xiaolong Zhou, Haibin Cai, Youfu Li, and Honghai Liu, Two-Eye Model-Based Gaze Estimation from A Kinect Sensor, 2017 IEEE International Conference on Robotics and Automation, (2017), pp.1646-1653.
- (3) Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling, Appearance-Based Gaze Estimation in the Wild, 2015 IEEE Conference on Computer Vision and Pattern Recognition, (2015), pp.4511-4520.
- (4) N. Uenoyama and M. Niitsuma, Temporal Segmentation of Environmental Map Based on Changing Rate of Activity Areas and Improvement of Tracking Method using LRF, 2015 IEEE/SICE International Symposium on System Integration, (2015), pp.57-62.
- (5) L. A. Jeni, Jeffrey F. Cohn, and T. Kanade, Dense 3D Face Alignment from 2D Videos in Real-Time, 11th IEEE International Conference on Automatic Face and Gesture Recognition, (2015), pp.1-8.
- (6) H. Fukushima and M. Niitsuma, Investigation of Cognitive Characteristics of Human Vision to Moving Objects for Visual Field Estimation, 2017 IEEE/SICE International Symposium on System Integration, (2017).
- (7) M. I. Posner and Y. Cohen, Components of visual orienting, Attention and Performance, (1984), pp.531-556.
- (8) K. Wada, M. Yageta and M. Tooyama, Preliminary Investigation of Attention Responses against Behavior Observation Robot ' s Movements, The 24th IEEE International Symposium on Robot and Human Interactive Communication, (2015), pp.202-205.