ショートノート

姿勢推定のための環状放射ネットワークモデル†

早瀬 光浩*・渡邊 昌平*・加納 政芳*

本稿では、楕円を用いた概括的な物体姿勢推定モデルを提案する。本モデルは、楕円を回転させたものを環状方向に、長短軸比を変化させたものを放射方向に配置し、それぞれを同心円状のネットワークとして接続することで構築されている。同様の考え方に基づくモデルとして、VTNモデルがある。このモデルでは、楕円体や直方体を回転移動・平行移動させて、2次元平面上に投影することで、物体の見え方図形を作り、それらを格子ネットワーク上に接続している。これを用いて、物体の姿勢推定を行うが、回転角・形状ともに同一な見え方図形がネットワーク内に複数存在するという問題がある。また、投影時の回転・平行移動度合い（離散化度合い）が一定のため、3次元物体の歪度が高くなると、十分な精度が得られないという問題もある。一方、提案モデルでは、これらの問題は発生しないため、高速かつ正確に物体の位置・姿勢を推定できると考える。

キーワード：人間共生システム、姿勢推定、環状放射ネットワークモデル、VTNモデル

1. はじめに

近年、物体認識技術は、工場での自動検査装置に見られるビジョンシステムや、医療分野における病変の自動認識システム、歩行者や車両の認識を行う高度交通システムだけではなく、デジタルカメラの顔抽出など、一般のユーザが利用する機器類にも広く応用されている。これらのシステムは、様々な物体を認識するものではなく、ある特定の物体を対象としており、それらのうちの多くは、2次元画像を用いて3次元物体やシーンを推定するものである。

2次元画像から3次元物体を推定する手法は、アピランススペースとモデルベースに大別できる。

アピランススペースで知られているものにパラメトリック固有空間法がある[1]。これは、パラメトリックに変化する対象の見かけ画像集合を主成分分析することで生成した辞書を利用し、姿勢推定を行う手法である。この手法は動画認識[2]や照明設計[3]などに応用され、有効性が確認されている。しかし、一般的に教示データを多数用意する必要があることや、類似した形状ごとにシステムを構築しなければならないといった欠点がある。

他方、モデルベースの物体認識の研究として、複数の円柱による物体認識[5]～[8]や、可変する円柱グリミティブによる姿勢推定、3次元エッジモデルを用い

たエッジによる推定[9]などが挙げられる。また、楕円体などの3次元物体を2次元平面上に投影して2次元見え方図形を作成し、それを格子ネットワーク状に接続することで、物体の姿勢推定を行うVTNモデルがある。同モデルでは、一様な離散度合いによって見え方図形が作成されているため、(1)ネットワーク内に回転角・形状ともに同一な図形が複数存在する、(2)3次元物体の歪度が高くなると十分な精度が得られない、という問題がある。

上記の問題を解消するために、本稿では、楕円を用いた環状放射ネットワークモデルを提案する。本モデルは、楕円を回転させたものを環状方向に、楕円の長短軸比を変化させたものを放射方向に配置し、それぞれを同心円状のネットワークとして接続するモデルである。これにより同一形状の重なり問題を解消できる。また、VTNモデルでは、3次元物体を一定角度で回転しているため、楕円体の歪度に比例してマッチングの精度が低下するが、本モデルでは、歪度に応じて回転角度を制御しているため、さまざまな形状に対して高精度にマッチングが行える。

2. 環状放射ネットワークモデル

図1に環状放射ネットワークモデルの概念図を示す。本節では、まず2.1節でモデル上に配置される無回転の楕円（楕円標準形）について説明し、2.2節で楕円の回転・接続を定義し、2.3節で放射方向の接続について述べる。

† A Radial-Ring Network Model for Object Posture Estimation
Mitsuhiro HAYASE, Shohei WATANABE and Masayoshi KANO

* 中京大学
Chukyo University

図1 環状放射ネットワークモデルの概念図

2.1 楕円標準形

i 番目の環に配置される楕円標準形(i -楕円標準形)を次のように定義する.

$$\frac{x^2}{a^2} + \frac{y^2}{\sigma^{2i}a^2} = 1. \quad (1)$$

ただし, σ は重なり率(後述)である($\sigma < 1$). i -楕円標準形は, 長径 $2a$, 短径 $2a\sigma^i$ の楕円となり, 外環ほど(i が大きいほど)扁平率の高い楕円が配置される. また, その面積 S_i は, $S_i = \pi a 2\sigma^i$ であり, i -楕円標準形と $\{i+1\}$ -楕円標準形を, 重心を重ねて配置したときの重なり率 σ は i によらず一定となる.

2.2 環状方向の接続

環 i において, 配置される楕円数を N_i , 楕円の回転角を $\theta_i = \pi/N_i$ [rad], i -楕円標準形を ϑ [rad]回転させた楕円を $o_i(\vartheta)$ とする. このとき, N_i 個の楕円 $o_i(j\theta_i)$ ($j=0, \dots, N_i-1$)において, $o_i((j \bmod N_i)\theta_i)$ と $o_i(((j+1) \bmod N_i)\theta_i)$ を接続する.

具体的な N_i としては以下が考えられる.

$$N_i = \begin{cases} 1 & (i=0) \\ N_1 & (i=1) \\ N_1 2^{i-1} & (i \geq 2). \end{cases} \quad (3)$$

この例で, 例えば $N_1 = 4$ とすると, 図1のように中心周りに4個の楕円が配置され, その後, 順次, 2の倍数で楕円数が増える環が構築される.

環 i における隣り合う楕円において, 両楕円の重心を重ねて配置したときの共通面積 D_i は, 以下の式で表現できる.

$$D_i = 2\sigma^i a (\pi - \arctan A_i - \arctan B_i). \quad (4)$$

ただし,

$$A_i = \sqrt{\frac{a^2}{r_i^2(\theta_i/2) \cos^2 \theta_i/2} - 1}, \quad (5)$$

$$B_i = \sqrt{\frac{a^2}{r_i^2((\pi - \theta_i)/2) \sin^2 \theta_i/2} - 1}, \quad (6)$$

$$r_i(\vartheta) = \frac{\sigma^i a}{\sqrt{(\sigma^{2i} - 1) \cos^2 \vartheta + 1}}. \quad (7)$$

したがって, 環 i における隣り合う楕円間の重なり率 δ_i は, 以下の式で表される.

$$\delta_i = \frac{D_i}{S_i} \neq \text{一定} \quad (0 < \delta_i < 1). \quad (8)$$

2.3 放射方向の接続

環 i と環 $i+1$ において, 同一角度だけ回転させた楕円が両環に存在すれば, それらの楕円の重なり率は σ となる. したがって, N_i は2を底とする指数関数で表現されることが望ましい. そこで, N_i ($i \geq 2$)を, $N_1 2^{xi}$ で表現される関数とし, $j\theta_i = k\theta_{i+1}$ なる楕円 $o_i(j\theta_i)$, $o_{i+1}(k\theta_{i+1})$ を接続する. ただし, x_i は, $\sigma \leq \delta_i$ なる最小の整数とする^{*}.

3. マッチングアルゴリズム

3.1 対象図形の正規化

環状放射ネットワークに配置される楕円は, 長短軸の比率を変化させて半径 a の円に内接させているため, 面積の異なるものである. したがって, 対象領域とマッチングを行うには何らかの方法でスケールリングを行う必要がある. 本稿では, 楕円のサイズを変更するのではなく, 対象領域の図形をスケールリングすることでマッチングを行う. 対象領域図形のスケールリングには最小包含円を用いる.

最小包含円とは, 平面上に散布するある点集合を全て含む最小の円のことであり, これを用いることで, 適切に楕円とのマッチングを行える. マッチングは, 楕円と対象図形の両者の重心を重ね合わせることで行うため, 対象図形を, その重心から最遠画素までの距離が a となるようにスケールリングすればよい. この場合の最小包含円は, 重心と重心からの最遠画素を含む長短軸比 a の楕円となる.

3.2 探索

本稿では, 環状放射ネットワークの探索のために, 勾配法に基づく探索アルゴリズムを考える. 図2にア

^{*} x_i は1以上の整数であり, i によって単調増加($x_i \leq x_{i+1}$)するものであれば定義上問題はない.

入力：対象図形 I

出力：探索結果 t

1. **begin**
2. $o := o_0(0);$ % o は探索中の楕円
3. $t := o_0(0);$
4. $L := \phi;$ % L は探索候補
5. $M := o;$ % M は探索済集合
6. **repeat**
7. o に接続する楕円 v のうち, $v \notin M$ かつ $m_I(o) \leq m_I(v)$ なる楕円を L に追加;
8. $o = \arg \max_u m_I(u) | u \in L;$
9. **if** $m_I(t) < m_I(o)$
10. **then** $t := o;$
11. $L := L \setminus o;$
12. $M := M \cup o;$
13. **until** $L = \phi$
14. **end.**

図2 探索アルゴリズム

ルゴリズムを示す。同図において, $m_I(o)$ は, スケーリング済みの対象図形 I と楕円 o とのマッチング率を求める関数である。同関数は, I の最小包含円の中心と楕円の中心を重ね合わせることで, 以下の式で計算できる。

$$m_I(o) = \frac{|o \cap I|}{\max(|o|, |I|)} \quad (9)$$

ここで, $|x|$ は画像領域 x の画素数を表す。

本アルゴリズムでは, 環状放射ネットワークの中心に配置された正円 $o_0(0)$ を探索ポイント o として探索を開始する(同図中 11.2-5)。まず, o に接続された楕円のうち, 未探索の楕円に注目し, 対象図形とのマッチング率が o よりも高いものを探索候補の集合 L に加える(1.7)。つぎに, L 内の楕円のうち, 最大のマッチング率を持つものを選択する(1.8)。選択された楕円がこれまでの探索で最大のマッチング率を有している場合は, それを解候補 t とする(11.9 and 10)。探索ポイント o を探索候補集合 L から除外し, 探索済の集合 M に加える(11.11 and 12)。上記の探索を探索候補が無くなるまで繰り返すことで, 解を求める(1.13)。

4. 評価実験

ロボットの物体把持動作を考慮した実験として, 卓上の文具や工具などの各部位の姿勢推定を行う。出力結果についてはVTNモデル[10]との比較を行う。実験パラメータを表1に示す。同表のように, σ , N_i , Enlargement degree, およびRotation degreeを設定し, 提案手法とVTNモデルのネットワーク内の楕円数(Ellipses)をほぼ同一にすることで, 姿勢推定の精度にどの程の度差が生じるかを調査する。具体的なパラメータの設定手順は以下の通りである。

(a) 入力画像

(b) 対象図形

図3 入力画像3 (a)と対象図形3 (b)。3 (b)内の全セグメントが探索の候補となる。

表1 実験パラメータ

Proposed model	
σ	0.75
N_i	$\begin{cases} 4 & (i = 1) \\ N_1 2^{i-1} & (i \leq 9) \end{cases}$
Ellipses	2045

VTN model	
Enlargement degree	0.1
Rotation degree	9[deg]
Ellipses	2000

VTNモデルのEnlargement Degreeは、楕円サイズを変更するパラメータであり、楕円の扁平率変化が10段階、最大扁平率が0.9となるように0.1に設定した。一方、環状放射ネットワークモデルでは、 $\sigma=0.75$ 、 $i \leq 9$ より、楕円の扁平率変化が10段階、最大扁平率は0.925となるため、VTNモデルとほぼ同一条件といえる。この条件を満たしつつ、ネットワーク内の楕円数をほぼ同一にするために、VTNモデルのRotation degreeを9[deg]と設定した。

入力画像(図3(a))を画像ピラミッド法によってセグメンテーションし(図3(b))、その画像内の全セグメントを対象図形とした。また、探索終了時において $m_i(t) < 0.5$ のセグメントは姿勢推定の対象外とした。

図4(a)に提案手法による推定結果を、図4(b)にVTNモデルによる姿勢推定結果を示す。VTNモデルでは、楕円の扁平率とは無関係に同一角度で回転させてモデルを生成し、姿勢推定を行っている。すなわち、回転対称な楕円の数は、不変としている。そのため、扁平率の高い物体を認識する場合には大きな角度誤差が生じる。実際に、図4(d)を見るとわかるよう

に、中央付近にあるハンマーの柄の推定角度には5[deg]程度の誤差が生じていた。一方で提案手法の場合は1[deg]未満の誤差となっており(図4(c))、正確な姿勢推定が行えているといえる。

従来方法で姿勢推定の精度を高める場合には、扁平率の低いモデルまで不要に生成してしまう。これに対して、本稿で提案するモデルの場合には、扁平率にあわせて指数的にモデル数を増加させているため、少ないモデル数で高精度の姿勢推定が行える。これを確認するために、表1のパラメータを利用して、提案モデルとVTNモデルの処理速度と精度の比較を行った。表2に示す環境で、ランダムに描画された欠損楕円に対して、姿勢推定を1万回行った。楕円の欠損は、楕円の上に楕円を描画することで作成した。実験結果を表3に示す。同表において、Matching rateはマッチ

表2 実験環境

CPU	Intel®Core™i7 2.67GHz
OS	Windows7 64bit
RAM	4.00GB

(a) 提案モデルによる推定結果

(b) VTNモデルによる推定結果

(c) 入力画像との比較 (提案モデル)

(d) 入力画像との比較 (VTNモデル)

図4 姿勢推定結果

表 3 実験結果

	Matching rate	Runtime [ms]	Number of searched ellipses
Proposed model	0.780 ± 0.130	0.40 ± 0.61	34.2 ± 40.0
VTN model	0.786 ± 0.123	8.13 ± 3.79	640.8 ± 185.3

ング率, Runtimeは実行時間, Number of searched ellipsesは対象図形とのマッチング回数である. 同表から, 提案モデルは, VTNと比較して, 同程度のマッチング精度を維持しつつ, 約20.3倍の高速化が行えていることが確認できる. VTNでは, 図5に示すように縦横拡大によって扁平率の異なるスケールの楕円が生成されるため, 複数の正円が生成される. 同様に, マッチング率が大きく異なる低扁平率の楕円も大量に生成される. この結果, マッチング回数が必然的に多くなり, 実行時間が増加する. 一方, 提案手法では, これらの問題は生じないため, 高速かつ正確にマッチングを行えたと考える.

図 5 VTN モデルにおける類似図形の生成

図 6 提案手法による推定結果例

図6に, 他の画像の実験結果を示す. 同図からも, 提案モデルによって物体の大まかな姿勢が推定できていることが確認できる

しかしながら, 提案手法では, オクルージョンがある場合や, 適切な領域分割が行えなかった場合については, 正確な姿勢推定を行うことができない. オクルージョンの問題に対しては, 楕円検出手法[13, 14]を取り入れること, 領域分割の問題に対しては, 距離画像を用いることを検討している.

5. おわりに

本稿では, 楕円を構成要素とする環状放射ネットワークモデルを提案した. また, 楕円プリミティブの組合せで多様なシーンを表現することを示した. このことから, 物体認識のみならず形状から推測認識するようなシーンへの応用の可能性が示唆されたと考えられる.

本稿で提案したアルゴリズムは, 対象図形が楕円であれば最適解を求めることができるが, 楕円以外の図形に対しては必ずしも最適解を求めることはできない. この問題に対しては, 対象図形の複雑さを考慮に入れたコスト関数などを導入し, 探索範囲を動的に変化させることで, 準最適解を求めることを考えている.

参考文献

- [1] H. Murase and S. K. Nayer: "Visual Learning and Recognition of 3-D Objects from Appearance," International Journal of Computer Vision, vol.14, no.1, pp.5-24 (1995)
- [2] H. Murase and R. Sakai: "Moving Object Recognition in Eigenspace Representation: Gait Analysis and Lip Reading," Pattern Recognition Letters, vol.17, no.2, pp.155-162 (1996)
- [3] H. Murase and S. K. Nayer: "Illumination Planning for Object Recognition Using Parametric Eigenspace," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.16, no.12, pp.1219-1227 (1994)
- [4] A. Laubenheimer, S. Richter and K. Kroschel: "3D Pose and Shape Estimation with Deformable Models in Life-like Scenes," IEEE/RAS International Conference on Humanoid Robots, pp.159-166 (2007)
- [5] T.O. Binford: "Visual Perception by Computer," IEEE Conference on Systems and Control (1971)

- [6] G.J. Agin and T.O. Binford: "Computer Description of Curved Objects," International Joint Conference on Artificial Intelligence, pp.629-640 (1973)
- [7] D. Marr and H.K. Nishihara: "Representation and Recognition of the Spatial Organization of Three-dimensional Shapes," Proceedings of the Royal Society of London B, vol.200, no.1140, pp.269-294 (1978)
- [8] R.A. Brooks: "Model-Based Three-dimensional Interpretations of Two-dimensional Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.5, no.2, pp.140-150 (1983)
- [9] N. Payet and S. Todorovic: "From Contours to 3D Object Detection and Pose Estimation," IEEE International Conference on Computer Vision (2011)
- [10] M. Hayase and S. Shimada: "Posture Estimation of a Human Body from Thermal Images of 2D Appearance Model of 3D Ellipsoidal Model," Journal of Advanced Computational Intelligence and Intelligent Informatics, vol.13, no.3, pp.172-177 (2009)
- [11] A. Kosir and J.F. Tasic: "Pyramid Segmentation Parameters Estimation Based on Image Total Variation," EUROCON, vol.2, pp.179-183 (2003)
- [12] R. Marfil, L. Molina-Tanco, A. Bandera, J.A Rodriguez and F. Sandoval: "Pyramid Segmentation Algorithms Revisited," Pattern Recognition, vol.39, no.8, pp.1430-1451 (2006)
- [13] K. Kanatani and N. Ohta: "Automatic Detection of Circular Objects by Ellipse Growing," International Journal of Image and Graphics, vol.4, no.1, pp.35-50 (2004)
- [14] S. Tsuji and F. Matsumoto: "Detection of Ellipses by A Modified Hough Transformation," IEEE Transaction on Computers, vol.27, no.8, pp.777-781 (1978)
- [15] G. Hetzel, B. Leibe, P. Levi, B. Schiele, "3D Object Recognition from Range Images Using Local Feature Histograms," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol.2, pp.394-399 (2001)
- (2012年4月2日 受付)
(2012年6月10日 採録)
- [問い合わせ先]
〒470-0393 愛知県豊田市貝津町床立101
中京大学 情報理工学部 機械情報工学科
加納 政芳
TEL: 0565-46-6910
FAX: 0565-46-1299
E-mail: mkanoh@sist.chukyo-u.ac.jp

著者紹介

はやせ みつひろ
早瀬 光浩 [学生会員]

2006年中京大学情報科学部情報科学学科卒業。2008年中京大学大学院情報科学研究科博士前期課程修了。同年、同大学大学院情報科学研究科博士後期課程進学。コンピュータビジョンの研究に従事。特に、物体認識に興味を持つ。情報処理学会会員。

わたなべ しゅうへい
渡邊 昌平 [非会員]

2010年中京大学情報理工学部情報システム工学科卒業。2012年中京大学情報科学研究科情報科学専攻修士課程修了。在学中、コンピュータビジョンの研究に従事。

かのう まさよし
加納 政芳 [正会員]

2004年名古屋工業大学大学院工学研究科博士後期課程修了。同年中京大学講師。2010年同大学准教授。現在に至る。博士(工学)。感性・知能ロボティクス、インタラクションの研究に従事。特に、人の感性に主導されるヒューマン・ロボット・インタラクションに興味を持つ。2006年日本感性工学会技術賞。2010年日本知能情報ファジィ学会論文賞。日本ロボット学会、人工知能学会、日本感性工学会、日本公衆衛生学会、情報処理学会、電子情報通信学会、電気学会、日本顔学会、IEEE各会員。

A Radial-Ring Network Model for Object Posture Estimation

by

Mistuhiko Hayase, Shohei Watanabe and Masayoshi Kanoh

Abstract :

A model is described that estimates object posture by using a number of ellipses. First, the ellipses are rotated and arrayed in the annular direction. Next, the ellipses that change long-and-short axis ratio are arrayed in the radial direction. Finally, a radial-ring network model is built by connecting each of these ellipses to form a concentric network. A variable template network (VTN) model is based

on the same concept: an ellipsoid is rotated, enlarged and reduced, and then several a 2D appearance figure is projected onto each 2D plane surface. The VTN model is constructed by connecting these figures to form a grid. Although object postures can be estimated using this model, there are multiple figures with the same rotation angle and form in the VTN model. Similarly, since the discrete degree at the time of projection is constant, if the skewness of a 3D object is high, this can not estimate its posture accurately. The proposed model does not suffer these problems, enabling it to estimate object position and posture quickly and accurately.

Keywords : human symbiotic systems, posture estimation, radial-ring network model, VTN model

Contact Address : **Masayoshi KANOH**

*Department of Mechnics and Infformation Technology School of Information Science and Technology
101 Tokodachi, Kaizu-cho, Toyota, 470-0393, JAPAN*

TEL : 0565-46-6910

FAX : 0565-46-1299

E-mail : mkanoh@sist.chukyo-u.ac.jp