

LoG フィルタを用いた AEH モデルによる人物顔認識 Face Recognition Using AEH Model Combined With LoG Filters

山崎 勇樹[†]
Yuuki Yamazaki

高野 博史[†]
Hironobu Takano

中村 清実[†]
Kiyomi Nakamura

1. はじめに

近年、情報化社会の発展に伴い、セキュリティに対する関心が非常に高まっており、顔、指紋、虹彩などの生体情報を利用した個人認証技術（バイオメトリクス）の研究が盛んに行われている。特に顔による個人認証は、普段から人と人の間で自然に行われている認証方法なので、他のバイオメトリクスに比べ心理的負担が少ないなどの利点がある。主な顔認証技術として、主成分分析による手法や Gabor Wavelet を用いた手法などがある。これらの手法では、あらかじめ顔画像を登録してから本人認証を行っている。

一方、脳の頭頂連合野と側頭連合野の相互関係に Jones の仮説を組み込んだ「連合野一内嗅野一海馬系（AEH）モデル」が考案され、顔画像による人物認識への応用について研究が行われてきた[1]。AEH モデルによる顔認識システムは、逐次的に人物顔を学習し、同じ人物顔が入力されるとその人物の記憶を強化するという特徴を持っている。しかし、AEH モデルによる顔認識には、正面顔であっても微少な変化（時間経過や表情変化など）に脆弱であるという問題があった。そこで本研究では、網膜にある神経節細胞の受容野の感度分布に近似できる LoG（Laplacian of Gaussian）フィルタを AEH モデルに導入することにより、顔認識率を向上させることができ、顔画像の微小な変化による影響を低減することができた。

2. 連合野一内嗅野一海馬系（AEH）モデル

AEH モデルとは、連合野一内嗅野一海馬系ニューラルネットモデル（Association cortex – Entorhinal cortex – Hippocampal formation Model）の略称であり、脳の頭頂連合野と側頭連合野の相互関係に Jones の仮説を組み込んだニューラルネットモデルである[1]。図 2 に AEH モデルの構造を示す。このモデルは大きく分けて視覚野、側頭連合野、頭頂連合野、内嗅野一海馬系の四つの部分から構成されている。これらの部分を以下では便宜上、視覚野を VC、側頭連合野を IT、頭頂連合野を PG、内嗅野一海馬系を EC-HF とそれぞれ略す。

網膜に入ってきた入力情報（顔画像）が、外側膝状体を通り V1 層へ入力される。V1 層では、入力顔画像を顔の構成要素である部分顔（眉・目・鼻・口）に分離し、さらにコントラスト強調を行い V2・V3 層へ出力する。V2・V3 層では、顔全体の位置情報は位置情報認識機能をもつ PG 層へ、各部分顔の形状情報は特徴抽出機能をもつ Layer S へ出力する。Layer S では各部分顔の形状情報を集約し、Layer C（各 V4・TEO 層）に出力する。Layer S からの出力パターンを階層的ネットワークによりペースコード情報として認識を行うことで、入力画像の位置ずれを許容する。そして、V4・TEO 層からのペースコード情報は部分顔ごとに認識され、それらの認識結果は LTM 層で一つにまとめられ顔全体の人物認識が行われる。

† 富山県立大学大学院工学研究科。
Graduate School of Engineering, Toyama Prefectural University

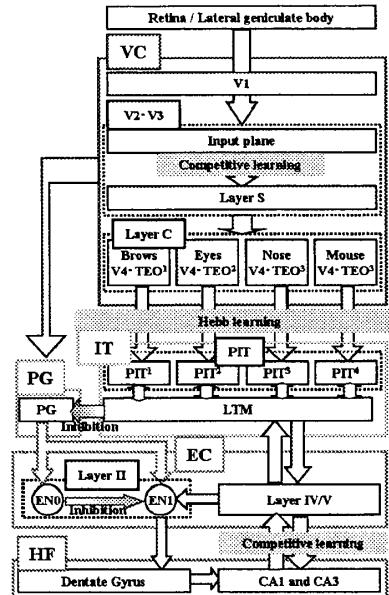


図 1 AEH モデルの構造

入力された顔が未学習であるか既学習であるかによって、以降の動作は異なる。未学習の顔が入力された場合は学習を行い、既学習の顔が入力された場合は想起及び再学習を行う。

まず、未学習顔が入力された場合を説明する。未学習顔の場合、LTM 層はいずれのニューロンも弱い出力を示し、その結果、PG 層に対して抑制は働く。PG 層は EC 層 Layer II の抑制性ニューロン EN0 と興奮性ニューロン EN1 に対して、共に強い出力を与える。この 2 つがそれぞれ違う閾値をもち、EN1 は EN0 から抑制を受けている。EN1 は EN0 から抑制を受けているが、PG 層からの強い入力があるため EN1 は活性化する。EN1 からの出力が一定値以上に達すると HF の Dentate Gyrus を経由して CA1 and CA3 で学習に必要なスペースコードが生成される。そのスペースコードは CA1 and CA3 から EC 層の Layer IV/V に出力され、CA1 and CA3 と Layer IV/V 間で競合学習が行われる。この競合学習で勝ち残った出力は EN1 と LTM 層に、さらに、LTM 層から PIT 層にフィードバックされる。V4・TEO 層と PIT 層間のヘブ学習で部分顔ごとの人物記録、V2・V3 層と V4・TEO 層間の競合学習でスペースコードの記録が行われる。一方、Layer IV/V から EN1 に強い出力がフィードバックされることで EC – HF の反響閉回路が動作し、一定期間、反響閉回路の活動が維持される。この間に Layer IV/V – CA1 and CA3 間の競合学習でスペースコードの記録も行われる。

次に、既学習顔が入力された場合を説明する。既学習顔の場合、LTM 層のいずれかのニューロンが強い出力反応を示し人物認識を行う。その結果、PG 層に対して強く抑制が働く。この抑制のため PG 層は EN0 と EN1 に対し弱い出力を与える。EN1 は EN0 から抑制を受けているため

活性化しない。一方、LTM層のニューロン出力はLayer IV/Vを活性化させる。これにより、入力情報に対応するスペースコードがCA1 and CA3で誘起し、スペースコードの強化が起きる。そのスペースコードはCA1 and CA3からLayer IV/Vに再度フィードバックされ、CA1 and CA3 – Layer IV/V間で競合学習が行われる。そして、Layer IV/Vの出力はEN1とLTMにフィードバックされる。しかし、EN1はEN0から抑制を受けているため活性化されず、EC – HFの反響閉回路は動作しない。一方、LTM層からPIT層にフィードバックされた出力によって、V4・TEO層とPIT層間のヘブ学習、V2・V3層とV4・TEO層間の競合学習が行われ、記憶の長期増強が起きる。

3. V1層によるLoGフィルタ処理

AEHモデルはV1層で入力顔画像のコントラスト強調を行う。従来研究ではコントラスト強調としてラプラシアンフィルタを用いていた。しかし、ラプラシアンフィルタではエッジのみを検出するので、顔画像の微小な変化に脆弱であった。そこで、顔画像の微小な変化を吸収するばかりの処理を含み、網膜にある神経節細胞の受容野特性にも近似できるLoGフィルタによってコントラスト強調を行った。

人間の視覚系の初期段階である網膜には、刺激の光強度分布を検出するX細胞、及び、強度の時間的変化を検出するY細胞がある。特にX細胞に関しては、その受容野の感度分布が式(1)で表される二次元LoG関数に近似されている[2]。

$$\nabla^2 G(r) = -\frac{1}{\pi\sigma^4} \left(1 - \frac{r^2}{2\sigma^2}\right) \exp\left(-\frac{r^2}{2\sigma^2}\right) \quad (1)$$

但し σ^2 はガウス関数の分散を示す。

一方、自然なエッジ検出を目的とした人間の視覚特性に基づくLoGフィルタスケールの選択法が提案されている[3]。人間の知覚は異なる大きさの受容野を持つ複数の空間周波数チャンネルを有し、大きな受容野で外界の大まかな構造を知覚し、小さな受容野で細部を知覚していることが示唆されている[3]。Wilsonは心理実験結果から、人間の視覚系には6個の空間周波数チャネルがあることを示している。各チャネルにおける周波数のピークは、0.8, 1.7, 2.8, 4.0, 8.0, 16.0c/d(cycle/degree)である。これらの周波数ピークに基づきLoGフィルタのフィルタスケールが算出されている。

また、Boutetらは帯域通過フィルタを通した低、中、高周波数画像による人物弁別課題により、顔の部分特徴の異なる人物弁別成績が5~20c/ow(cycle/object width)の中周波数画像で最も良い結果を得ている[4]。本稿では8~20c/owに相当する16.0c/dを周波数ピークとしたLoGフィルタを用いた。

4. 人物顔による認識実験

LoGフィルタを用いたAEHモデルによる人物顔認識実験を行った。人物顔認識実験にはソフトピアジャパンDBの正面顔を用いた(図2)。AEHモデルは逐次的に人物を学習(登録)する特徴があるので、AA'BB'CC'…と顔画像を提示する方法により顔認識実験を行った。A'はAとは同一人物であるがAとは時間間隔があり、Aと比べて微小な変化がある顔画像である。認識実験では100人分である合計200枚の顔画像を使用した。図2の上段はABC…の顔画

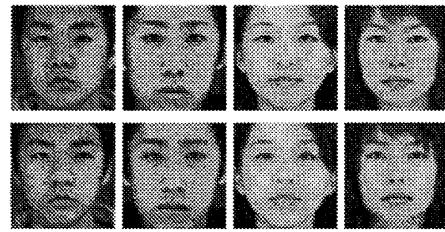


図2 認識実験に使用した顔画像
(上段:ABC…,下段:A'B'C'…)

像、下段はA'B'C'…の顔画像を表す。また、AEHモデルには想起を行う際にその人物の記憶を強化するという特徴をもっているので、呈示順序による認識率の偏りが起こる可能性がある。認識率の偏りを防ぐために、10通りの呈示順序により認識実験を行い、平均認識率を求めた。

LoGフィルタを用いたAEHモデルによる顔認識実験の結果を表1に示す。従来方法はラプラシアンフィルタを用いたAEHモデルによる認識結果を示し、提案方法はLoGフィルタを用いたAEHモデルによる認識結果を示す。新奇顔受入率は新奇人物(ABC)を学習する割合を示す。そして、認識率は新奇人物顔を学習し、二度目の入力顔画像で本人(直前に入力された顔画像)を想起(再認)する割合を示す。従来方法では新奇顔受入率が48.7%であるのに対して、提案方法では95.4%となり、LoGフィルタにより人物弁別性能が向上した。また、従来方法では顔認識率が34.5%であるのに対して、提案方法では顔認識率が81.4%となり、46.9%の顔認識率の向上を得ることができた。

表1 顔認識実験結果

	従来方法	提案方法
新奇顔受入率	48.7%	95.4%
顔認識率	34.5%	81.4%

5. 終りに

本稿では、LoGフィルタを用いたAEHモデルによる顔認識システムを提案した。そして、顔認識実験により従来のAEHモデルに対して46.9%の顔認識率の向上が確認できた。よって、LoGフィルタにより微小な変化による影響を抑えることができた。今後の課題として、眼鏡による本人認識率の低下への対応や、より精度の良い人間の知覚に基づいた顔認識システムの構築を考えている。

謝辞

実験に使用した顔画像データは、財団法人ソフトピアジャパンから使用許諾を受けたものであり、無断で複写、利用、配布等を行うことは禁じられています。

参考文献

- [1] 山崎 勇樹, 高野 博史, 中村 清実, “光環境変化に対するAEHモデルによる顔認識特性”, 信学技報, NC2007-188, p.p.451–456, (2008).
- [2] デビッド・マー著, 乾 敏郎, 安藤 広志 訳, “ビジョン—視覚の計算理論と脳内表現ー”, 産業図書, p.p.58–70, (1987).
- [3] 福島 和恵, 鳥谷 明, 鎧沢 勇, “LOGフィルタのスケール選択法の提案”, テレビジョン学会技術報告, Vol.16, No.84, p.p.31–38, (1992).
- [4] I. Boutet, C. Collin, J. Faubert, “Configural face encoding and spatial frequency information”, Perception & Psychophysics, Vol.65, No.7, p.p.1078–1093, (2003).