

PDPモデルに基づく画像理解システムの構成

3D-10 -ICE System(Image CEntered System)の提案-

鶴田直之 谷口倫一郎 雨宮真人
九州大学大学院総合理工学研究科

1. はじめに

ニューラルネットワークを用いた人工知能の研究では、行動や認知にはさまざまなレベルで拘束条件が働いており、その拘束条件を同時に満足する最適解として行動や認知が決定づけられているという計算理論が展開されている。例として、エネルギー最小化原理に基づく視覚モデル^{[1][2][3]}や言語理解の分野の意識下パラダイム^[4]が挙げられる。しかし、これらの試みではそれぞれ高レベル、低レベルの拘束条件の設定方法に関する研究がまだ不十分である。

我々の研究の目的は、視覚が信号処理のレベルと記号処理のレベルの柔軟なインターフェイスを果たしているという観点に立ち、前述の二つのアプローチの利点を併用して高レベルの拘束条件を利用可能な画像理解システムを構成することにある。

本稿では、上記の観点から画像理解システムの構成方法を提案し、その特徴と利点について述べる。さらに、システムを実現する際の問題点とその解決のための戦略を我々の研究の今後の計画と共に報告する。

2. 信号処理と記号処理のインターフェイス

エネルギー最小化原理に基づく視覚のモデルは、一般に、データ駆動の認識の結果が外界に関する事前知識としての拘束条件を満たしていない度合およびモデル駆動の予測に基づく再構成画像と入力画像との誤差の二つのエネルギーの和を最小にするように動作する。しかし高レベルの拘束条件に関しては、どのように記号化してネットワーク内に表現するか研究が不十分である。

一方、言語理解の分野では、文法的な拘束条件だけでなく実世界に関する拘束条件を利用しようとする意識下パラダイムが重要視されつつある。ただし、記号間の関係をそれぞれ一つのユニットで表現すると関係の組み合わせの爆発に対応できない。そこで、複数の関係を複数のユニットで符号化(パターン化)して表現する必要がある。しかし、その符号化方法はまだ十分に研究されていない。

エネルギー最小化原理に基づく視覚モデルと意識下パラダイムに基づく言語理解は、それぞれの課題を

補完し合うものであると捉えることができる。すなわち、視覚モデルで用いる高レベルの拘束条件を言語的な単語で表現することができ、単語間の関係の符号化された表現として予測に基づく再構成画像を用いることができる。したがって、視覚を言語的な記号処理と初期視覚の信号処理のインターフェイスとして捉えることによって、高レベルの知識を利用可能な画像理解システムが構成可能になる。

次節では、我々の提案する画像理解システムの構成と動作について述べる。我々のシステムは、次節で述べるように、中心に配置した再構成画像と周囲に配置した単語を中間視覚によって放射状に接続した構造を持つ。そこで、以下では我々のシステムをICE System(Image CEntered System)とよぶ。

3. ICE Systemの構成と動作

3-1. ICE Systemの構成

図1に、ICE Systemの構成の概略図を示す。最も外側の層は記号レベルの層で、構成要素は単語に対応して名詞、動詞、前置詞、数詞などの品詞にグループ化されている。各品詞内は抑制性の結合を持ち、一つ(あるいは2、3個)のユニットしか一度に活性化できない。品詞間には直接の結合はない。

中間層は信号と記号の双方向の変換が可能なネットワークである。この層で処理される情報は、単語とは直接対応付けできない準記号である。以後、この層を準記号層と呼ぶ。ただし、複数の名詞から一つの名詞が構成される場合があるので準記号層と名詞層には明確な境界はない。

最も内側の層は信号レベルの層で、筆者らがすでに提案している群化と正則化を併用した初期視覚モデル^[7]である。扱う情報は直接入力画像との誤差を計算するための再構成画像である。

3-2. 認識過程と想起過程

ネットワーク内では、信号から記号へ変換する認識過程と記号から信号へ変換する意識的想起過程^[8]が並行して実行される。

外部から画像が入力されると、信号層で画像の領域分割と画像の正則化、奥行きなどの検出が並行して行なわれる。準記号層では、信号層で得られた領域の輪郭等を統合しながら各品詞で対応した単語をそれぞれ活性化する。単語の活性化値は、対応した対象物だけでなく、同時発生確率の高い付随的な情報にも影響される。さらに、記号層内の抑制性接続によって最も信頼性の高い単語が最終的に活性化される。

また、活性化した記号からは認識過程と逆の経路をたどって画像の想起が進む。このとき、付随的な情景も想起するように情報が伝搬する。画像が完全に想起されるのは、記号層のすべての品詞から想起の情報が得られた時である。

ネットワークが安定するのは、記号層の各品詞で記号が活性化し、それらによって想起された再構成画像と入力画像との誤差が十分に小さくなった時である。

以上の機能は、再構成画像と一つの品詞層の間だけをみると選択的注意機構をもつネオコグニトロン^{[5][6]}とほぼ同じである。

3-3. 画像の逐次的理解

ICE System では、記号によって記述できる関係は一時刻に一つ（または2、3個）だけである。したがって、画像の理解は逐次的な処理となる。

ここで、関係を一つずつ安定に抽出するのに品詞層の抑制接続と活性化した単語からの選択的な想起^[5]が有効に働くが、画像が複雑になると抽出される特徴が多すぎて選択機構が安定しないことが予想される。そこで、認識過程では信号層のスケールの値を大きな状態から初めて、注意を引く領域に対してのみスケールを局所的に小さくする。これにより、マルチスケールによるトップダウン型の画像理解が可能になる。

また、一旦認識された記号を不活性化状態にする必要がある。このため、ユニットの疲労機能あるいはより上位のシステムからの抑制信号が必要になる。

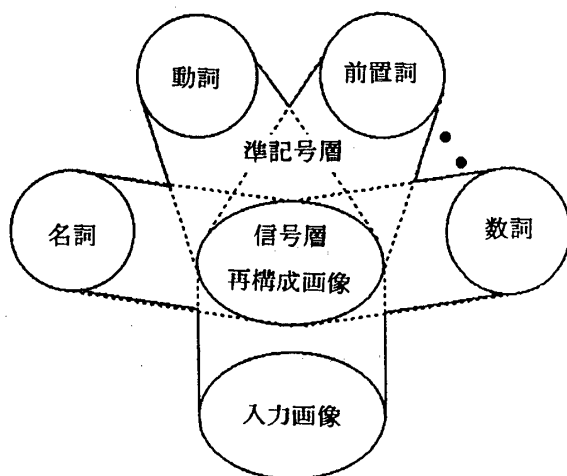


図 1. ICE System の構成図

4. ICE System の利点

ICE System では、記号層の結合は同一品詞内の抑制結合と名詞層で複数の名詞から抽象的な名詞が構成される場合の興奮性結合だけであり、品詞間の関係や一単語で表現できないような複雑な多項関係などを表現するためのユニットは必要としない。異なった品詞に属する記号は全て再構成画像を通じて結合されており、複雑な多項関係や一過性の関係は再構成画像として表された実態によって同時に活性化したということによって表現される。

したがって、異なった品詞間の記号の関係は全て再構成画像の形で符号化表現されているためにユニット数の組合せの爆発を起こすことなく任意の関係を表すことができる。

5. 今後の計画と問題点

我々は、最終的には動詞を拘束条件に用いて動画像の理解を可能にすることも目標としているが、時系列データの学習や認識過程と想起過程の同時性の維持など原理的な課題が多く残されている。

そこで今後の計画としては、例として、積木の世界の濃淡の静止画像を対象とし、記号を幾何学的な形状の名詞、および位置関係を表す前置詞に限定したプロトタイプを並列計算機上で試作する。

6. まとめ

本稿では、PDP モデルを応用した高次の拘束条件を利用可能な画像理解システム ICE System を提案し、その構成と利点について述べた。

【参考文献】

- [1] 角所, 北橋ら: "エネルギー最小化原理による画像の3次元解釈", 信学技法 PRU89-108(1990.1)
- [2] 川人, 乾: "視覚大脳皮質の計算理論", 信学論 D-II, vol. j73-D-II, No. 8, pp. 1111-1121(1990.8)
- [3] J. A. Feldman, M. A. Fandy, N. H. Goddard: "Computing with Structured Neural Networks", Computer, March(1988)
- [4] D. L. Waltz, J. B. Pollack: "Massively Parallel Parsing: A Strongly Interactive Model of Natural Language Interpretation", cognitive science, 9, pp. 51-74(1985)
- [5] 福島邦彦著: "神経回路と情報処理", 朝倉書店(1989)
- [6] 今川, 福島: "選択的注意機構のモデルによる英字筆記体連結文字列認識", 信学技法 NC90-50(1990.11)
- [7] 鶴田, 谷口, 雨宮: "群化と正則化を併用したニューラルネットワークによるエッジ検出", 信学技法 AI90-48(1990.7)
- [8] 山田, 天満: "階層ニューラルネットによる意識的想起方法と想起学習の検討", 信学技法 NC90-22