

大学の研究・動向

確率的画像処理と生命・知性のモデル

情報学研究科 システム科学専攻 システム情報論講座 論理生命学分野

教授 石 井 信

講師 大 羽 成 征

助教 前 田 新 一

1. はじめに

われわれ人間の脳の情報処理機械としての機能を考えるとき、「学習するシステム」の観点が不可欠です。情報処理機械としての脳の機能とは、周りの状況を入力し、これに対する適切な理解に基づいて、適切な行動判断を出力することです。適切な判断のために、脳は入力毎にリアルタイム性の高い計算を瞬時に行い続けています。これは多数の神経細胞が同時並列に働いていることによる演算性能の賜物とも言えますが、脳の機能を理解するために演算性能の高さに着目するのみでは足りません。入力される情報の意味・意義の理解は、生まれてから何年何十年をかけて積み重ねてきた経験に基づく意味世界と照らし合わせて初めて生まれるものだからです。経験の蓄積から意味世界を創りだしてゆく「学習するシステム」があるからこそ、目の前の入力を瞬時に位置づけ、瞬時に理解し、瞬時に適切な判断につなげることができるのです。

我々の具体的な研究対象は、脳、脳を構成する神経網、神経網を構成する個々神経細胞、脳によって構成される視聴覚、運動制御、高次行動決定などのシステム、さらには癌細胞その他の挙動を調べるシステム生物学から、これらを扱ううえで必要な理論基盤の研究、そしてその工学応用に及びます。広範な対象のさまざまな階層のどこを考える際にも、「学習するシステム」の観点が欠かせません。

「生命」と「知性」はまわりの環境に応じて自らを作り変える「学習するシステム」です。論理生命学分野では「生命」・「知性」・「学習するシステム」に関わる広範にわたる研究を、ボトムアップの実験科学のアプローチと、トップダウンの理論的アプローチの2つの方向から行なっています（図1）。

本稿では、我々が行なっている研究のなかでも視覚システムと画像情報処理に関わるものについていくつかスポットを当てて紹介いたします。

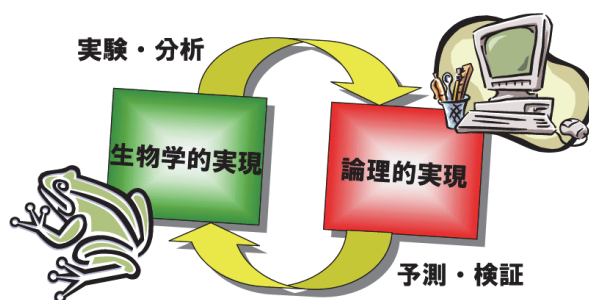


図1：論理生命学の枠組み

2. 生成モデルと因子化

視覚的情報の蓄積によって意味世界を構成するための理論は、画像統計学の名前で知られています。画像統計学では、意味世界の構築をデータ生成モデルという考え方をを用いて説明します。

一連の画像データを蓄積した結果に基づいて、一枚一枚の画像の意味を知りたいとします。生成モデルの考え方では、隠された意味特徴（たとえば一人の具体的な顔）を想定し、確率的なノイズを加えるプロセスを経て、実際の観測画像が「生成」されるプロセスを仮定します。このプロセスを生成モデルと呼びます。次に、一連の画像データの蓄積から、このプロセスの詳細を決める変数を決定します。すると、この生成モデル逆に使うことによって、一枚の画像からその裏にある隠された意味特徴を瞬時に計算して決めることができるようになります。生成モデルの応用のなかでも、理論的に重要であり応用上の展開が広いものとして、因子化と再構成とがあります。

因子化とは、観測の一つ一つを少数の因子の重ね合わせとして理解するモデルのことを言います。この場合、重ねあわせの重み係数が「生成モデルにおいて隠された意味特徴」となります。この際、何を因子とするかを経験に基づいてあらかじめ適切に決定する必要があります。例えば千差万別の人間の個性に対して、良い因子化を行えば優れた人間理解に繋がるかもしれませんが、悪い因子化を行えば無根拠な差別に繋がるかもしれません。画像処理においても、良い因子化は優れた画像理解・画像処理につながります。

因子化は、人間の脳においても重要な計算原理のひとつであると考えられており、その証拠として、ヒトや類人猿の脳の視覚野において画像パッチの因子として反応する神経細胞が見つかっています。これらは、視野に入力される一定の形状に対して特異的に反応するフィルターとして活動します。計算機上で一般の自然画像を大量に用意して画像パッチの因子を求めると、哺乳類の神経細胞が示す「因子」機能と酷似した因子が得られます。

図2は画像パッチの因子化の一例です。画像から、小さな（例えば 15×15 ピクセル程度の）領域を切り出します。これを画像パッチと呼びます。元の画像が白～灰色～黒のグレースケール画像であった場合、各ピクセルの輝度を適当な範囲内の実数値で表すことができますので、ひとつのパッチは数値ベクトル（この例では225次元実数値ベクトル）で表すことができます。画像パッチの因子化では、ひとつひとつのパッチ画像を、少数の基底パッチ $\mathbf{u}_k$ の線形和 $\sum_k v_k \mathbf{u}_k$ で表します。基底パッチを適切に求めることができれば、ひとつひとつの観測パッチを少ない個数の比例係数 $v_1 \cdots v_K$ で表すことができます。

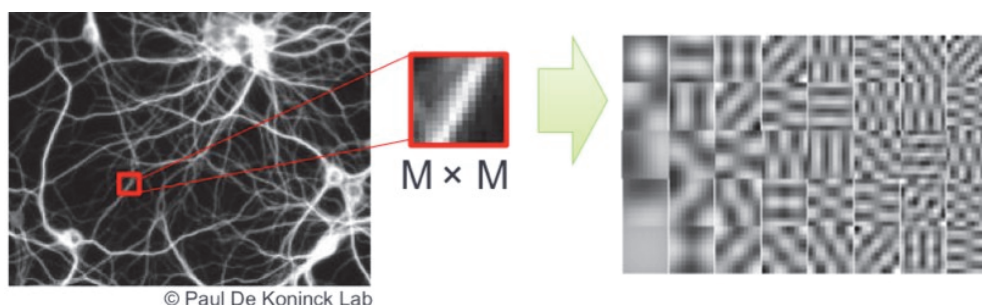


図2：画像パッチの因子化。画像の一部を取り出したパッチを、少数因子パッチの和で表す。

画像パッチ因子化の典型的な応用は、画像復元です。あらかじめ適切に定めた因子パッチがあれば、これをノイズの大きな画像や、傷のある画像のパッチに適用することで比例係数 $v_1 \cdots v_K$ を推定し、そして推定した比例係数に基づいて因子パッチの線形和を求めることで、傷の無いパッチを復元することができます。画像中の全ての場所でパッチを復元すれば、画像全体でノイズや傷の除去ができます。

最近の研究では、不完全なデータばかり、しかも少量しか得られないような状況から、よい因子化を行う方法を工夫しています。こういう状況は実際に珍しくありません。例えば顕微鏡画像データから稀少な構造を検出したいとき、探索テンプレートとして使える典型例は少数しか得られません。そこで、少数のパッチを学習する「視点」を様々に増やすことで実質的な学習の効率を高める工夫が功を奏する場合があります。神経細胞の繊維構造を検出する特殊用途のための因子化モデルは、一般用途の因子化モデルよりも高いノイズ除去性能を示しました。[1]

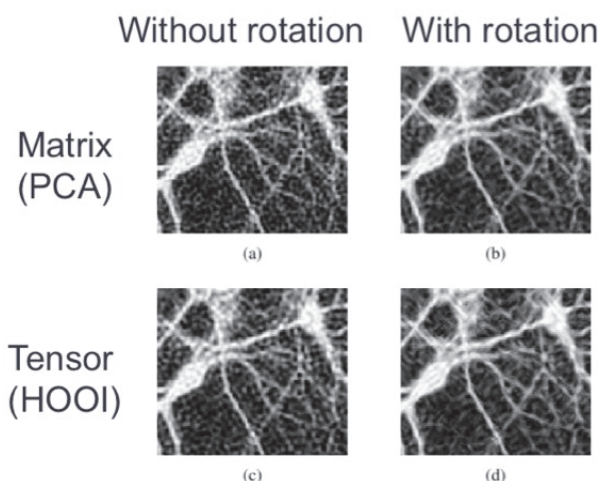


図3：画像因子化による画像ノイズ低減法の比較。パッチ回転とテンソル因子化法の組み合わせ（右下）によってノイズ低減の精度が増した。

3. 再構成と超解像とX線CT

生成モデルの応用において再構成も重要な考え方です。直接観測できない隠れた真の対象と、ノイズ付加プロセスとを想定し、これらに基づいて限られた観測が「生成」されると仮定します。ノイズを含む有限の観測に基づいて、直接観測できない隠された真の対象の構造を「再構成」するのです。

超解像は、解像度の低い観測画像を多数集めることによって、直接観測できない高解像度画像を再構成する、という画像情報処理技術です。初めて見る人には魔法のようですが、基本的な仕組みは難しいものではありません。

超解像は、解像度の低い観測画像を多数集めることによって、直接観測できない高解像度画像を再構成する、という画像情報処理技術です。初めて見る人には魔法のようですが、基本的な仕組みは難しいものではありません。

同じ対象を複数回撮影するとき、これを細かく見ると撮影位置はほんの少しずつだけお互いにズレています。図4の(a)は小さな○は、数字の8を3x3の低解像度で2度撮影したときの、各ピクセルの中心点を示しています。撮影結果は図(b)と図(c)のようになりました。どちらを見ても一枚の画像から撮影対象の形は全く読み取れませんが、大量の撮影がお互いにほんの少しずつでもズレていてくれさえすれば対象の形を知るための情報が増えてゆくことが分かります。超解像処理では、このように少しずつ異なる観測の結果を集めて再構成することで、解像度の高い画像を得ます。

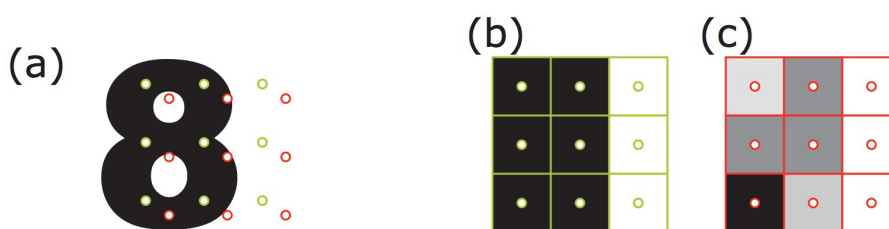


図4：複数画像に基づく超解像の仕組み。

私たちは生成モデルによる再構成を改善するために、「背景知識利用」の方法を工夫しています。図5は、雲に遮蔽された衛星画像9枚に基づく超解像の結果です。(a)を典型例とする衛星画像9枚から地上の様子を知ることが目的です。雲に遮られているせいで、地上の一部しか見えていません。また、一枚一枚の解像度は100ピクセル×100ピクセルと粗いものです。これらを「超解像処理」することで、(e)のようなクッキリした地上画像を得ることができます。単なる平均(b)や中央値(c)などの統計処理ではぼんやりした映像しか得られません。超解像(d)によって400ピクセル×400ピクセルの詳細

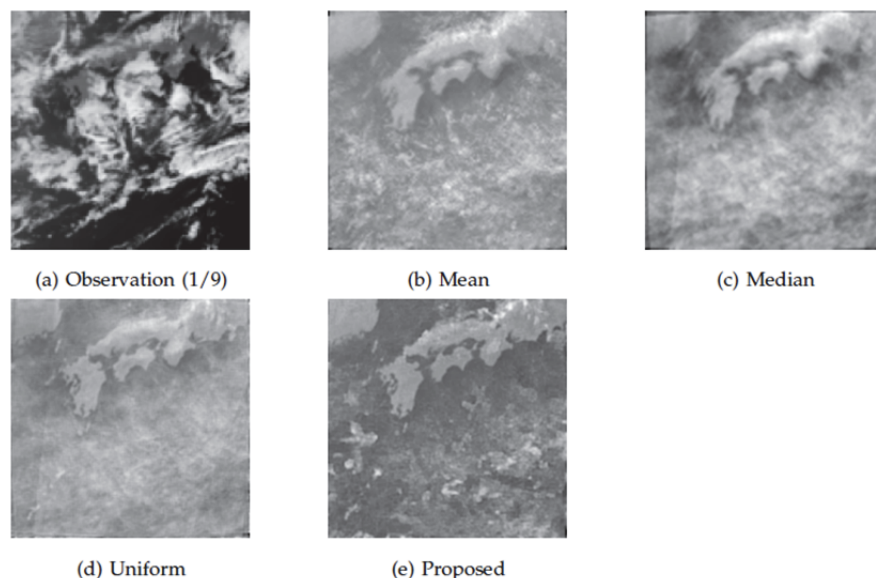


図5：雲に遮蔽された衛星画像9枚に基づく4×4超解像の結果。(a)使用した9枚の画像のうちの一枚。(b)9枚の平均をとったもの(c)9枚の中央値をとったもの(d)雲による遮蔽を考慮せずに超解像を試みた結果(e)提案法。遮蔽を推定しつつ超解像を行った結果。

細画像が得られますが、雲による遮蔽の影響があり得ることを背景知識として利用する特別な解析を行うと(e)のように全体的にハッキリします[2]。

X線CTは、様々な角度から対象物にX線を照射して得られる投影像をもとに、その対象物の内部をX線吸収係数の違いによって可視化します。非侵襲で人体内部の組織構成を推定することができるため、医用画像診断の重要なツールとなっています。この手法も原理は超解像と同様であり、複数の方向から照射されたX線の吸収量データを集めて、内部構造を「再構成」しています。

X線CTによる再構成の解像度を高めるために最も簡単な方法は、より多くの角度から投影像を得ることや、X線照射量を増やして投影像の信号雑音比を向上させることです。しかしこれらによって、X線被曝量を増大させてしまうのは人体を対象とするときの大きなデメリットとなります。より少ない被曝量でより良い再構成結果を得るために、「背景知識利用」の工夫が効きます。図6は、撮像対象である人体の各組織のX線吸収係数のおおよその分布を背景知識として利用することで、性能を向上させた例です。とても高いX線吸収係数をもつ義歯が含まれるとき、従来法(FBP法, 最尤推定法[MLCT])では強いゴーストが発生しますが、「高いX線吸収係数を持つ箇所があり得る」こと「同様の吸収係数を持つ箇所は一塊をなす傾向にあること」を背景知識として盛り込んだベイズ法[3]を適用することで推定を改善させました。

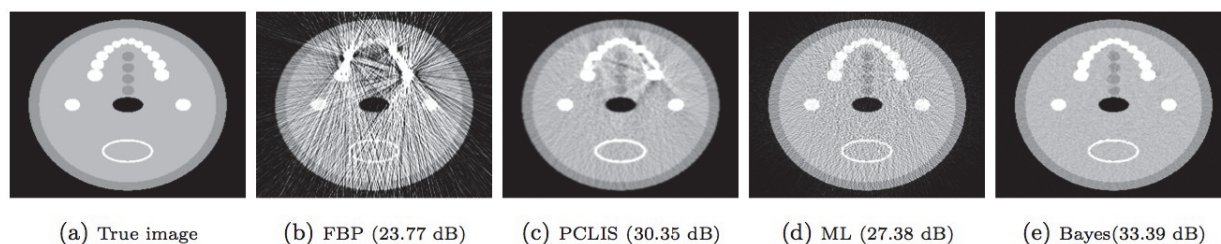


図6：3つの義歯の入った頭部を模したファントムデータに対するX線CT画像再構成の結果。数字はピーク信号雑音比を表し、真の画像との誤差が小さいほど大きな値をとる。

4. 今後のチャレンジ

今後我々がチャレンジしてゆきたい話題の一つに、顕微鏡下動画像データの解析があります。いま、システム神経生物学の最先端では、神経細胞の線維構造と小突起状器官が形成する局所回路の構造と機能が大変にホットな話題となっており、これを調べるべく顕微鏡による撮影手法の工夫も急ピッチで進んでいます。なかでも神経細胞やその小器官を染色する分子蛍光を用いた撮影が盛んに研究されていますが、X線CTと同様に撮影時の光量を増すことで対象を痛める問題があり、信号雑音比を一定以上に増すことができない中で、画像処理によって高い解像度を求めたいニーズがあります。また、こうした撮影を時間に沿って行うことで、神経細胞やその各種細部構造体の動的挙動を知りたいニーズがあります。

ここで「背景知識」を用いた「再構成」を用いた超解像技術や、動画像の「因子化」技術の発展が求められています。本稿2章と3章では因子化と再構成とを別個のものとして説明いたしましたが、生成モデルの立場から考えれば2つを密接に関連付けるモデルは自然な拡張です。動画像は形を持ったオブジェクトと、各オブジェクトの時間的変動から成りますが、動画像中から複数オブジェクトを検出する問題は、画像統計の立場から言えば動画像の因子化に他なりません。限られた光量による撮影画像から、神経細胞の動的挙動を高い時空間解像度で求めるのは、再構成の問題です。すべてを盛り込んだら何が見えるのでしょうか？

ここまで我々の画像処理技術の原理について簡単にご紹介しましたが、技術的詳細をさらにお知りになりたい方は [4] をご覧ください。この先のチャレンジに対しても、同様の原理のとおりデザインした手法が思うように働けば簡単ですがなかなか一筋縄ではいかないのが常です。モデルの妥当性、計算技術の問題、撮影時に混入する思いも寄らないアーティファクトの問題、などなどひとつひとつクリアしてゆくのはなかなか大変です。しかし、開発した技術がピタリとはまることで、誰も見たことのない映像を世界で最初に目の当たりにする瞬間が我々のものになるかもしれないのです。それを夢みながら研究を進めてゆけるのは、学習するシステムによる高度な画像処理研究の世界の大きな魅力と言えるのではないのでしょうか。

- [1] Kouno, M., Nakae, K., Oba, S., and Ishii, S. (2011) Microscopic image restoration based on tensor factorization of rotated patches. The Seventeenth International Symposium on Artificial Life and Robotics 2012, 902-905.
- [2] Kanemura, A., Maeda, S., Fukuda, W., and Ishii, S. (2010) Bayesian image superresolution and hidden variable modeling. Journal of Systems Science and Complexity, 23 (1), 116-136.
- [3] Fukuda, W., Maeda, S., Kanemura, A., and Ishii, S. (2010) Bayesian x-ray computed tomography using material class knowledge, in ICASSP, pp. 2126-2129.
- [4] 前田新一, 兼村厚範, 石井信. (2011) 確率システムの立場からの画像情報処理技術. システム制御情報学会誌, 55 (12), pp.532-538.