ML EM画像再構成における並列計算

OGAWA, Koichi / 池本, 浩幸 / IKEMOTO, Hiroyuki / 尾川, 浩一

(出版者 / Publisher)
法政大学計算センター
(雑誌名 / Journal or Publication Title)
Bulletin of Computer Center, Hosei University / 法政大学計算センター研究報告
(巻 / Volume)
8
(開始ページ / Start Page)
49
(終了ページ / End Page)
54
(発行年 / Year)
1995-03-31
(URL)
https://doi.org/10.15002/00024696

ML EM 画像再構成における並列計算

池本浩幸
法政大学工学部 †
尾川 浩一
法政大学工学部 †

ML EM アルゴリズムによる画像再構成は不完全な投影データからでも良質な再構成画像を得ることができるが、そのアルゴリズムは反復計算を含むものであるために再構成に要する時間は長くなる。本研究では超並列コンピュータを用いて二つの ML EM 画像再構成の並列化を行なった。一つはノードにおける処理の単位を画素とする手法、もう一つはノードにおける処理の単位を投影経路とする手法である。この二つの手法によって ML EM 画像再構成の実行時間を短縮することができた。また、解析的画像再構成を並列化する場合の有効性や問題点についても述べた。

1. はじめに

エミッション CT(ECT : Emission CT) の映像化の方法には解析的手法と統計的手法とがある。解析的手法は投影定理に基づいて再構成を行なうものであり、統計的手法は確率の理論を応用して再構成を行なうものである。統計的手法の一つとして、期待値最大化 (EM : Expectation Maximization) 法 [1] を用いた最尤 (ML : Maximum Likelihood) 推定 [2] による画像再構成法 (ML EM アルゴリズム) がある。ML EM アルゴリズムやその他の統計的手法は、投影データが不完全なものであっても良質な画像を再構成することができるという長所をもつが、そのアルゴリズムが反復計算を行うため、解析的手法に比べかなり長い再構成時間を要する。そのため、統計的画像再構成法の高速化 [3] に関するさまざまな方法が考えられているが、近年、超並列コンピュータの普及にともない、統計的手法の並列処理による高速化 [4] の研究が行なわれるようになった。

一方超並列コンピュータは、一つのプロセッサの処理能力を高め逐次処理の高速化を図る従来のコンピュータとは異なり、多数のプロセッサ (ノード) に処理を分散し並列に処理を行なうことによって高速化を図るコンピュータである。超並列コンピュータを用いて ML EM アルゴリズムのような比較的負荷の軽い計算を何度も繰り返すような処理を並列化する場合には、データ通信やノード間での負荷の偏りがボトルネックとなり、並列性が大きく損なわれることが多い。このような問題に対処するために、本研究では超並列コンピュータ CM-5E による ML EM アルゴリズムの並列処理を行なうための処理の分散方法やノード間の通信方法などについて検討した。そしてノードの負荷を小さくすることによって全体の処理時間を短縮するアルゴリズムまたはデータ通信量を減らすことによって並列性を高めるアルゴリズムを提案し ML EM アルゴリズムの並列化を実現した。

2. 最尤推定法

パラメータ θ を持つ母集団より抽出された一つの確率変数 X が x という実現値をとる確率密度が $f(x|\theta)$ で与えられるとする。この関数を x の関数ではなく x を固定して θ の関数と見た場合、これを $L(x|\theta)$ と書き、θ の尤度 (Likelihood) という。

いま、帰属不明の確率変数 X の一つの実現値 x_a があり、それに対して既知である二つの分布の母集団があるものとする。この時、それぞれの母集団に対するパラメータの値は θ_1、θ_2 となり、x_a がそれぞれの母集団に属すると考えられる確からしさ (尤度) $L(x_a|\theta_1)$、$L(x_a|\theta_2)$ は Fig. 1に示すようになる。x_a がどちらの分布に帰属するかを考えた場合、尤度が大きい方であると考えるのが自然である。このような原理に基づいた推定値の決定法が最尤推定法である。

確率変数 $\mathbf{X} = \mathbf{X}^n (= X_1, X_2, ..., X_n)$ の実現値 $\mathbf{x} = \mathbf{x}^n (= x_1, x_2, ..., x_n)$ に対して、パラメータ $\theta = \theta^k (= \theta_1, \theta_2, ..., \theta_k)$ のもとで、ある一組の確率密度関数が $f(x|\theta)$ で与えられ $\mathbf{x}$ が独立であるとすると、結合確率密度は

$$L(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) \tag{1}$$

† 〒184 東京都小金井市梶野町 3 − 7 − 2

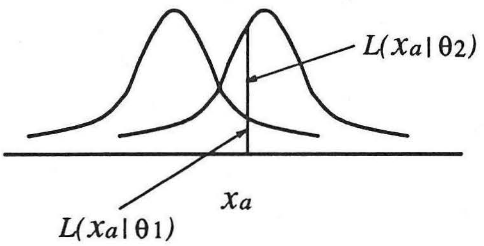

Fig. 1 最尤推定法

となる。この式を θ の関数と考えた時、これを θ の尤度関数（Likelihood Function）という。この値を最大にする $\theta = \hat{\theta}(x)$ を最尤推定値（Maximum Likelihood Estimate）といい、最尤推定値を用いて x_i の点推定を行なうことを最尤推定法（Maximum Likelihood Method）という。$L(\mathbf{x}|\theta)$ と $\ln L(\mathbf{x}|\theta)$ とは最大値をとる θ は同じなので、簡単のため (1) 式の対数

$$\ln L(\mathbf{x}|\theta) = \sum_{i=1}^{n} \ln f(x_i|\theta) \tag{2}$$

を考え、これを最大にする θ を求める。したがって、$L(\mathbf{x}|\theta)$ が θ に関して微分可能ならば微分された関数はもとの関数において最大値をとる θ を与えれば 0 になるので

$$\frac{\partial \ln L(\mathbf{x}|\theta)}{\partial \theta} = 0 \tag{3}$$

として θ を求めることができる。

3. 期待値最大化法

期待値最大化（EM : Expectation Maximization）法とは不完全データ $\mathbf{y}$ に対して尤度関数 $L(\mathbf{y}|\theta)$ を最大にするパラメータ θ を決定する際に、$\mathbf{y}$ の替わりに $\mathbf{y}$ に付随して得られるデータ $\mathbf{x}$ を用いて $L(\mathbf{y}|\theta)$ を最大にする θ を決定する方法である。

X から Y への many-to-one 写像であるサンプル空間 X、Y があり、X における完全データ $\mathbf{x}$ を観測する代わりに Y における不完全データ $\mathbf{y}$ が観測できるとする。$\mathbf{x}$ の密度関数がパラメータ θ として $f(\mathbf{x}|\theta)$ で与えられるとすると $\mathbf{y}$ の密度関数は

$$g(\mathbf{y}|\theta) = \int_{X(\mathbf{y})} f(\mathbf{x}|\theta) d\mathbf{x} \tag{4}$$

で与えられる。ただし、$X(\mathbf{y}) = \{\mathbf{x} : \mathbf{y}(\mathbf{x}) = \mathbf{y}\}$ である。パラメータ θ は最尤推定（Maximum Likelihood Estimation）を用いて $g(\mathbf{y}|\theta)$ を最大にすることによって推定される。

ここで、$k(\mathbf{x}|\mathbf{y}, \theta)$ を $\mathbf{y}$ と θ が与えられたときの $\mathbf{x}$ の条件付き密度関数であるとすると、ベイズの定理

$$f(\mathbf{x}|\theta) = g(\mathbf{y}|\theta) \cdot k(\mathbf{x}|\mathbf{y}, \theta) \tag{5}$$

より、$\mathbf{y}$ の密度関数は

$$g(\mathbf{y}|\theta) = \frac{f(\mathbf{x}|\theta)}{k(\mathbf{x}|\mathbf{y}, \theta)} \tag{6}$$

である。両辺の対数をとれば

$$\ln g(\mathbf{y}|\theta) = \ln f(\mathbf{x}|\theta) - \ln k(\mathbf{x}|\mathbf{y}, \theta) \tag{7}$$

となる。ここで、推定するパラメータを θ' としたとき、$\mathbf{y}$ および任意の θ に対して $\ln g(\mathbf{y}|\theta')$ の条件付き期待値は

$$\begin{aligned} E\{\ln g(\mathbf{y}|\theta')|\mathbf{y}, \theta\} &= \int \ln g(\mathbf{y}|\theta')\mathbf{k}(\mathbf{x}|\mathbf{y}, \theta)\mathbf{dx} \\ &\triangleq \ln g(\mathbf{y}|\theta') \end{aligned} \tag{8}$$

となる。これを尤度関数 $L(\theta')$ として

$$L(\theta') \triangleq \ln g(\mathbf{y}|\theta') \tag{9}$$

と定義すると

$$L(\theta') = E\{\ln f(\mathbf{x}|\theta')|\mathbf{y}, \theta\} - E\{\ln k(\mathbf{x}|\mathbf{y}, \theta')|\mathbf{y}, \theta\} \tag{10}$$

となる。ここで $\mathbf{y}$ と θ が与えられたときの $\ln f(\mathbf{x}|\theta')$ の期待値を $Q(\theta'|\theta)$、$\mathbf{y}$ と θ が与えられたときの $\ln k(\mathbf{x}|\mathbf{y}, \theta')$ の期待値を $H(\theta'|\theta)$ とすると

$$Q(\theta'|\theta) = E\{\ln f(\mathbf{x}|\theta')|\mathbf{y}, \theta\} \tag{11}$$

$$H(\theta'|\theta) = E\{\ln k(\mathbf{x}|\mathbf{y}, \theta')|\mathbf{y}, \theta\} \tag{12}$$

となり、尤度関数は

$$L(\theta') = Q(\theta'|\theta) - H(\theta'|\theta) \tag{13}$$

となる。

ここで Jensen の不等式

$$H(\theta'|\theta) \leq H(\theta|\theta) \tag{14}$$

によって θ の更新により $H(\theta'|\theta)$ が減少することが保証される。したがって $L(\theta')$ を最大にする θ' を求めるためには $Q(\theta'|\theta)$ を最大にする θ' を求めればよ

い。したがって、EM アルゴリズムは以下に示すように定義される。

EM アルゴリズム

Step1
パラメータの初期値 θ_0 の決定

Step2
Expectation-step $Q(\theta|\theta_n)$ の計算

Step3
Maximization-step $Q(\theta|\theta_n)$ を最大にする $\theta=\theta_{n+1}$ の選択

Step4
収束条件を満足していなければ $n=n+1$ として Step2 へ
満足していれば終了

4. ML EM アルゴリズムによる画像再構成理論

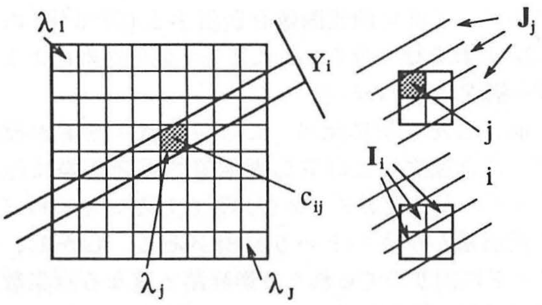

Fig. 2 各記号の定義

Fig. 2のように i を投影、j を画素、$\mathbf{I}_i$ を投影 i に関わる画素の集合、$\mathbf{J}_j$ を画素 j に関わる投影の集合、Y_i を投影 i で検出される実測値、c_{ij} を画素 j の情報が投影 i に検出される確率、λ_j を画素 j の画素値と定義すると、ML EM アルゴリズムによる画像再構成理論は次のようになる。

最尤（ML : Maximum Likelihood）推定による画像再構成では、任意の画像 λ が与えられたときの投影データ $\mathbf{Y}$ が得られる条件付き確率 $f(\mathbf{Y}|\lambda)$ を最大にするような画像 λ を推定する。

光子の放出はポアソン過程にともなう現象であるから、投影データもポアソン分布に従う。したがって $f(\mathbf{Y}|\lambda)$ は

$$f(\mathbf{Y}|\lambda)=\prod_i\{\frac{\exp(-\sum_{j\in\mathbf{I}_i}c_{ij}\lambda_j)(\sum_{j\in\mathbf{I}_i}c_{ij}\lambda_j)^{Y_i}}{Y_i!}\} \tag{15}$$

となり、対数をとり λ について無関係な項をはずすと

$$\begin{aligned}&\max_\lambda \ln f(\mathbf{Y}|\lambda)\\&=\max_\lambda\sum_i\{-\sum_{j\in\mathbf{I}_i}c_{ij}\lambda_j+Y_i\ln(\sum_{j\in\mathbf{I}_i}c_{ij}\lambda_j)\}\end{aligned} \tag{16}$$

となる。これを解くために EM アルゴリズムを用いる。ここでは、可観測不完全データである積分投影データ $\mathbf{Y}$ に対して $\ln f(\mathbf{Y}|\lambda)$ を最大にする λ を決定する際に、$\mathbf{Y}$ の替わりに非観測完全データである $\mathbf{Y}$ の各要素 $\mathbf{X}$ を用いて $\ln f(\mathbf{Y}|\lambda)$ を最大にする λ を決定する。ここで X_{ij} が画素 j から放射された光子が投影 i で検出される投影データであるとすると

$$Y_i=\sum_{j\in\mathbf{I}_i}X_{ij} \tag{17}$$

となる。つまり、（16）式は

$$\begin{aligned}\ln f(\mathbf{X}|\lambda)\ =\ &\sum_i\sum_{j\in\mathbf{I}_i}\{-c_{ij}\lambda_j\\&+\ X_{ij}\ln(c_{ij}\lambda_j)-\ln X_{ij}!\}\end{aligned} \tag{18}$$

となる。Expectation-step により、投影データ $\mathbf{Y}$ と現在の画像データ λ^n が与えられたときの $\ln f(\mathbf{X}|\lambda)$ の期待値は

$$\begin{aligned}&E\{\ln f(\mathbf{X}|\lambda)|\mathbf{Y},\lambda^n\}\\&=\sum_i\sum_{j\in\mathbf{I}_i}\{-c_{ij}\lambda_j+N_{ij}\ln(c_{ij}\lambda_j)\}+R\end{aligned} \tag{19}$$

となる。ここで、R は定数項、N_{ij} は

$$N_{ij}=\frac{c_{ij}\lambda_j^nY_i}{\sum_{k\in\mathbf{I}_i}c_{ik}\lambda_k^n} \tag{20}$$

である。次に、Maximization-step によって

$$-\sum_{i\in\mathbf{J}_j}c_{ij}+\sum_{i\in\mathbf{J}_j}\frac{c_{ij}\lambda_j^nY_i}{\sum_{k\in\mathbf{I}_i}c_{ik}\lambda_k^n}\cdot\frac{1}{\lambda_j}=0 \tag{21}$$

となり、$\lambda_j = \lambda_j^{n+1}$ として解くと再推定式は

$$\lambda_j^{n+1} = \frac{\lambda_j^n}{\sum_{i\in\mathbf{J}_j} c_{ij}} \sum_{i\in\mathbf{J}_j} \frac{c_{ij} Y_i}{\sum_{k\in\mathbf{I}_i} c_{ik}\lambda_k^n} \tag{22}$$

となる。

5. ML EM アルゴリズムの並列化

(22) 式はおもに更新中の画素値の投影データ ($\sum_{k\in\mathbf{I}_i} c_{ik}\lambda_k^n$) と、注目している画素と投影経路との距離からその画素を投影経路が通る面積の比率 (c_{ij}) の二つの値を計算するものである。本研究では反復回数や投影角を保持、管理し、データをまとめる役割を持つホストを設け、二つの計算を超並列コンピュータが持つプロセッシングノード (以下ノード) に分散することによって ML EM アルゴリズムの並列化を図った。各ノードに割り当てる処理の実行範囲 (以下、処理単位) は、コンピュータで行なう画像再構成における離散的な処理の区分を考えると画素か投影経路のいずれかを用いるのが適当であるので、並列化の手法は以下のようになる。

5.1. 画素を処理単位とする手法

画素をノードの処理単位とし、各ノードが再構成画像の投影の計算と画素値の更新を並列に行なうことのできるアルゴリズムは次のような流れになる。

まず、各ノードは割り当てられた画素の集合における不完全な投影データを計算する。そして、各ノードは計算した投影データをホストへ送信する (Fig. 3 ; 矢印 (a) の方向)。次に、ホストはリダクションによって集めた投影データの投影経路ごとの総和を計算し、完全な投影データを作る。ホストはブロードキャストによって各ノードに投影データを送信する (矢印 (b) の方向)。ブロードキャストされた投影データをもとに各ノードは割り当てられた画素の画素値の更新を行なう。そして各ノードは更新した画素値をホストへ送信する (矢印 (a) の方向)。ホストは更新された画素値をもとに再構成画像を作り、ブロードキャストによって投影データの計算に必要なデータである再構成画像を各ホストに送信する (矢印 (b) の方向)。これを繰り返すことによって並列処理による画像再構成が行なわれる。

後述する手法に比べ、この手法には各ノードに割り当てられた画素数が等しいために負荷にほとんどバラつきがなく、処理にかかる時間が短いという長所があるが、その反面、データ通信量は大きくなっている。

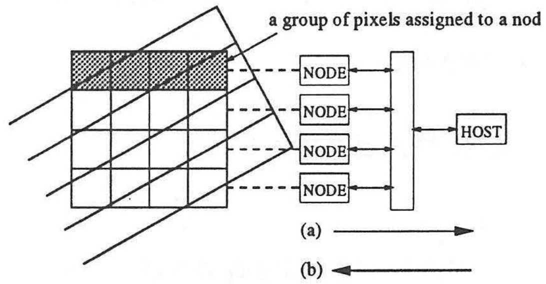

Fig. 3 画素を処理単位とする手法

5.2. 投影経路を処理単位とする手法

投影経路をノードの処理単位とし、各ノードが再構成画像の投影の計算と画素値の更新を並列に行なうことのできるアルゴリズムは次のようになる。

まず、各ノードは割り当てられた投影経路における投影データを計算する。投影データの計算を終了したノードはすぐに割り当てられた投影経路と重なっている部分の画素値の更新を行なう。そして、各ノードは更新した画素値をホストへ送信する (Fig. 4 ; 矢印 (a) の方向)。ホストは集めた画素値の画素ごとの総和を計算し、完全な画素値 (再構成画像) を作る。再構成画像は投影データの計算を行なうために必要なので、ホストはブロードキャストによって各ノードに再構成画像を送信する (矢印 (b) の方向)。これを繰り返すことによって並列処理による画像再構成が行なわれる。

前述した手法に比べ、この手法はノードが投影データを独立して計算し画素値の更新の際に他のノードがもつ投影データを必要としないためにデータ通信量が小さいという長所がある。しかし、各ノードに割り当てられた投影経路と重なる画素数が投影経路の位置や角度によって変化するために、負荷にバラつきが多く、処理にかかる時間も長くなっている。

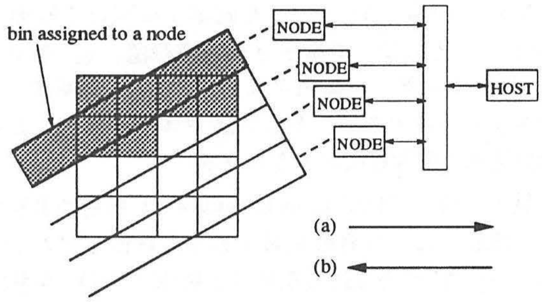

Fig. 4 投影経路を処理単位とする手法

6.　シミュレーション結果

前節の通り、画素をノードの処理単位とする並列化(並列1)および投影経路をノードの処理単位とする並列化(並列2)によってML EMアルゴリズムによる画像再構成の並列処理を考え、その有効性を確かめるためにシミュレーションを行なった。シミュレーション条件はTable 1の通りである。また、再構成画像をFig. 5に、実行時間、1回あたりの平均時間および逐次処理との実行時間の比率をTable 2に示す。

Table 1 シミュレーション条件

画素数	64 × 64
投影数	45(0～180°)
投影経路数	64
原画像	Shepp phantom
反復回数	1、5、10回
ノード数	32

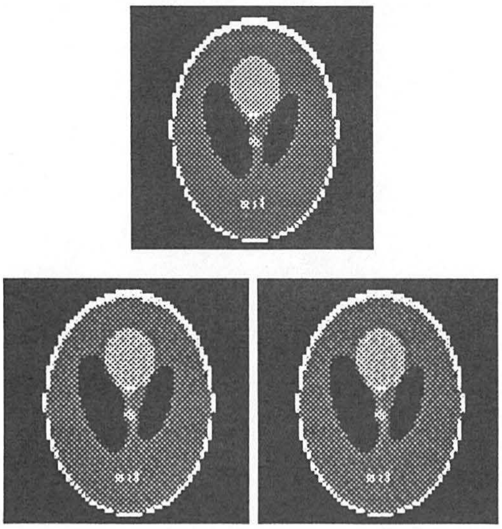

Fig. 5 上；原画像、左；並列1、右；並列2

Table 2 シミュレーション結果

繰り返し回数	逐次処理	並列1	並列2
1	3.10	0.82	1.10
5	15.7	4.12	5.48
10	31.4	8.24	11.0
平均	3.14	0.82	1.10
比率	1.00	0.26	0.35

7.　検討

画素をノードの処理単位とする並列化の手法と投影経路をノードの処理単位とする並列化の手法とを比較すると、前者は後者に比べて40%ほどノードの負荷が小さく、後者は各ノードが計算した投影データを他のノードに送ったり、画素値の更新を行なうために他のノードから投影データを集めたりする必要がないためにデータ通信量が前者の約66%である。その結果、処理全体では画素をノードの処理単位とする並列化の手法のほうが投影経路をノードの処理単位とする並列化の手法より約26%実行時間が短縮されている。

両手法において処理時間を大幅に短縮することができなかったのは、逐次処理自体の処理の負荷が小さかったために並列化の有効性が低くなってしまったことと、逐次処理にはなかったコミュニケーションに比較的時間がかかってしまったことが原因であると思われる。

8.　まとめ

超並列コンピュータを用いてML EM画像再構成を並列化した結果、画素をノードの処理単位とする並列化の手法では約26%に、投影経路をノードの処理単位とする並列化の手法では約35%に処理時間を短縮することができた。

参考文献

1) A. P. Dempster , N. M. Laird and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm (with discussion), " *J. Royal Statist. Soc. B*, vol. **39**, pp. 1-38, 1977.

2) K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography, " *J. Comput. Assist. Tomog.* , vol. **8**, 2, pp. 306-316, Apr. 1984.

3) T. R. Miller and J. W. Wallis, "Fast Maximum-Likelihood Reconstruction, " *The Journal of Nuclear Medicine*, vol.**1**, 33, pp. 1710-1711, Sep. 1992.

4) M. I. Miller and C. S. Butler, "3-D Maximum *A Posteriori* estimation for single photon emission computed tomography on massively-Parallel computers, " *IEEE Trans. Med. Imaging*, vol. **MI-12**, 3, pp. 560-565, Sep. 1993.

キーワード

画像再構成、並列処理、アルゴリズム

Summary

Parallel Calculation Algorithm in ML-EM Image Reconstruction

HiroyukiIKEMOTO
College of Engineering, Hosei University †
Koichi OGAWA
College of Engineering, Hosei University†

Abstract

The image reconstruction method using ML-EM(Maximum Likelihood method using Expectation Maximization) algorithm is a kind of stochastic approach, and is useful when acquired projection data is incomplete. However the method is not practical because of computation time. In this paper, we proposed two parallel computation algorithms using a massively parallel computer CM-5E. In this study we developed two message-passing programs; one is a pixel oriented program and the other is a bin oriented one. These two methods were evaluated and the validity of parallel processing in image reconstruction was also discussed.

Key Words

image reconstruction, parallel processing, algorithms

†3 – 7 – 2, Kajino-cho, Koganei-shi, Tokyo 184, Japan