Image Restoration based on Maximum Entropy Method with Parameter Estimation by Means of Annealing Method

Kohei Arai 1
Graduate School of Science and Engineering
Saga University
Saga City, Japan

Abstract—Image restoration based on Maximum Entropy Method (MEM) with parameter estimation by means of annealing method is proposed. The proposed method allows spatial resolution enhancement. Using overlap sampling with a low resolution of sensor, high spatial resolution (corresponding to the sampling interval) can be achieved through ground data processing with image restoration methods. Through the experiments with simulation imagery data derived from Advanced Very High Resolution Radiometer (AVHRR) data, it was found that spatial resolution enhancement can be achieved, MEM is superior to the others when S/N ratio is poor (less than 33) while Conjugate Gradient Method (CGM) is superior when the S/N ratio is higher than 33. It was also found that the CGM is superior to the proposed method for the existing sampling jitter.

Keywords—Image restoration; Maximum Entropy Method (MEM); annealing; Advanced Very High Resolution Radiometer (AVHRR); Conjugate Gradient Method (CGM)

I. INTRODUCTION

Image restoration methods can be roughly classified into linear restoration filters and nonlinear restoration filters [1]. The former starts with classical Wiener filters and parametric Wiener filters that only restore the best approximation image on average, and evaluates the difference between the restored image and the original image on the observed image, not on the space of the original image. A general inverse filter, a least-squares filter with constraints, a projection filter and a partial projection filter that may be greatly affected by noise in the restored image have been proposed [2]. However, the former is insufficent for optimization of evaluation criteria, etc. and is under study.

On the other hand, since the latter is essentially a method for finding a nonlinear solution, it can only take a method based on an iterative method, and various methods based on iterative methods have been tried. Although there are various iterative methods, there are a stationary iterative method represented by the successive over-relaxation method (SOR method) and a non-stationary iterative method represented by the conjugate gradient method [3].

In general, the former requires a large number of iterations, but the accuracy is high, and the latter has excellent convergence, but accumulation of rounding errors is a problem. When applied to image restoration, it is necessary to pay attention to noise resistance.

On the other hand, the maximum entropy method has been proposed as an image restoration method because it can take into account constraints (or prior knowledge) and resistance to noise [4]. However, this parameter estimation requires solving transcendental equations, and here again, iterative methods are needed [5]. As a method for estimating this parameter, methods using stationary iterative methods such as Newton's method and quasi-Newton's method and non-stationary iterative methods such as the conjugate gradient method have already been proposed. However, these methods are essentially local depending on the initial value. The solution may fall or the solution may diverge depending on the noise.

The method proposed in this article is based on image restoration based on the maximum entropy method, and in its parameter estimation, an annealing method was introduced to avoid the local solutions and divergence described above and obtain a stable minimum solution. Annealing is an iterative method of searching for the minimum solution of an objective function. In this case, it has been theoretically proved that the solution converges to a constant value regardless of the initial value [6]. This image restoration based on the maximum entropy method is a type of nonlinear restoration filter, and is the only image restoration that can be applied when the observation process is nonlinear, the constraints on the original image are nonlinear, and the evaluation criterion indicating the quality of the restored image is nonlinear. The problem of restoring high spatial resolution images from densely sampled data can be reduced to the problem of image restoration [7].

The author takes up the conjugate gradient method and the maximum entropy method, which are famous for reaching a solution in a finite number of times, and try to clarify the adaptive limits of image restoration based on these methods.

The superiority of the proposed method is compared with the image restoration based on the conjugate gradient method, taking the case of restoring a high spatial resolution image from a high-density sampled sequence as an example. In addition, the author has confirmed the robustness against noise compared with the most general least-squares method.
The following section describes research background followed by theoretical background. Then the proposed method is described followed by experiment. After that conclusion is described together with some discussions.

II. RESEARCH BACKGROUND

Inverse Problem in Super-Resolution and Its Solution is well introduced [8]. The image resolution using subpixel motion is proposed [9]. Also, a method for image compression with a cosmetic restoration is proposed [10]. On the other hand, DEM estimation with Simulated Annealing: SA based on surface reconstruction method is proposed [11]. Meanwhile, a method for estimation of Sea Surface Temperature: SST, wind speed and water vapor with microwave radiometer data based on simulated annealing is proposed [12].

Simultaneous estimation of geophysical parameters with microwave radiometer data based on accelerated Simulated Annealing: SA is proposed [13]. On the other hand, category decomposition method for un-mixing of Mixed Pixel: Mixels acquired with spaceborne based visible and near infrared radiometers by means of Maximum Entropy Method: MEM with parameter estimation based on Simulated Annealing is proposed [14]. Moreover, method for estimation of refractive index and size distribution of aerosol using direct, diffuse and aureole by means of simulated annealing is proposed [15]. Also, estimation of SST, wind speed and water vapor with microwave radiometer data based on simulated annealing is proposed [16].

III. PROPOSED METHOD

A. Image Restoration Method

In image restoration, when evaluating the quality of the estimated image \( f' \) of the original image \( f \), \( f \) cannot be directly used for evaluation because \( f \) is unknown. Then, the observed image \( g \) that observed \( f' \) was compared with the estimated image \( Hf' \), and minimize the following L2 norm.

\[
\|Hf'-g\|^2 \rightarrow \text{min.} \tag{1}
\]

It is assumed that the most likely original image \( f \) has been estimated when. Here, \( H \) is the transfer function of the image degradation operator. Such a method is called a least squares method.

Taking the variation of \( f' \) in equation (1) and setting it to 0

\[
\delta f' H (Hf'-g) = 0 \tag{2}
\]

\[
H' (Hf'-g) = 0 \tag{3}
\]

Solving equation (3) for \( f' \) gives:

\[
f' = (H'H)^{-1}Hg \tag{4}
\]

(Moor-Penrose generalized inverse matrix).

For equation (4), if \( H \) is regular and there is an inverse matrix,

\[
f' = (H'H)^{-1}Hg = H^+g \tag{5}
\]

That is, when an inverse matrix exists in \( H \), the image restoration filter is the inverse matrix itself. However, such cases are rare. In addition to these, starting with the Wiener filter and the parametric Wiener filter that assume that the frequency spectrum of the signal-to-noise ratio is known, the difference between the restored image and the original image is evaluated and approximated on the observed image, not in the space of the original image.

Generalized inverse matrix filters, least-squares filters with constraints, and projection filters and partial projection filters that may be significantly affected by noise in restored images have been proposed, but should not be dealt with here.

B. Maximum Entropy Method

Suppose an image with non-negative elements is normalized. Since the value \( P_i \) of each pixel is \((1) \sum_{i=1}^{M} P_i = 1 \) and \((2) \forall i: 1>P_i>0 \), it can be regarded as a random variable. The entropy of the image can be calculated by considering the pixel value as the probability.

\[
M \times N \begin{align*}
En &= -\sum_{i=1}^{M} f_i \log f_i \tag{6}
\end{align*}
\]

where \( En \): image entropy, \( M \times N \): is the number of elements in the image vector \( f_i \). However, in equation (6), the maximum value is taken when the image is flat, so a constraint condition is attached to this formula. The following restraint condition is considered,

\[
F: Hf'-g|^2 = \rho^2 \tag{7}
\]

Then, the following Lagrange's undetermined multiplier \( \lambda \) is considered with MEM,

\[
J = -f_i \log f_i + \lambda [(Hf'-g)-\rho^2] \tag{8}
\]

It can be said that it is equivalent to maximizing the above equation. Therefore, let \( f' \) which maximizes this energy \( J \) be the restored image in MEM. However, since this formula is a one-to-many mapping from vector to constant, and image restoration by filtering is difficult to realize, so an iterative method must be used to obtain the restored image.

That is, transcendental equations must be solved for this parameter estimation. As the parameter estimation method, methods using steady-state iterative methods such as Newton's method and quasi-Newton's method and non-steady-state iterative methods such as conjugate gradient method have already been proposed. May fall into a local solution or the solution may diverge depending on noise.

C. Parameter Estimation of Maximum Entropy Method by Annealing

The method proposed in this text is based on the image restoration based on the maximum entropy method, and in the parameter estimation, an annealing method was introduced in order to avoid the above local solution and divergence and obtain a stable minimum solution. It is a thing. Annealing is an iterative method for finding the minimum solution of an objective function. In this case, it is theoretically proved that the solution converges to a constant value regardless of the initial value [6].

www.ijacsa.thesai.org
The algorithm that embodies this is called simulated annealing, and is abbreviated here as annealing. The annealing algorithm is shown below.

1: Initial state \( x \) is set arbitrarily.
2: for \( t = 0 \) to \( \infty \) do begin
3: Perturb the state \( x \) and select the next state candidate \( y \). Generate random numbers according to the probability distribution \((..., P(Y, y), ...) \equiv (...), P(x, y), ...)\).
4: Set the temperature \( T(t) = \Delta / \ln (t + 2) \) and decide whether or not A is applied to \( y \). The acceptance matrix is \( 1/(1 + \exp \Delta E / T) \).
5: Generate random numbers according to the probability distribution \((P(Z = 1), P(Z = 0)) \equiv (A(x, y), 1-A(x, y))\). If \( Z = 1 \), swap \( x \) and \( y \). If \( Z = 0 \), leave \( x \) unchanged.

Where \( t \) is the time, \( \Delta \) is a large constant for temperature control, A is the acceptance matrix, and \( P \) is the perturbation matrix. Also, as shown in Fig. 1, the state after change is selected as the next state with a probability of \( 1/(1 + \exp \Delta E / T) \).

The probability of transition can be reduced by gradually lowering the temperature \( T \) as the number of calculations increases. In other words, even if the temperature falls to a local value while the temperature is high, the transition to a higher energy is stochastically allowed, so that the local value can be escaped.

As a result, the optimal distribution of the following Boltzmann distribution defined by the objective function \( F \) and the positive constant \( T \) is determined, and the objective function reaches the minimum solution.

\[
q(x) = \left(\frac{\exp(-F(x)/T)}{Z}\right)
\]

where, in the case of image restoration, the target of annealing is an image, and there is no pixel that cannot be selected when determining the next state \( y \) from a certain state \( x \).

From this, a uniform random number is generated instead of the perturbation matrix, and the pixel to be changed next is selected according to the random number. Also, as shown in Fig. 2, instead of the acceptance matrix, the state after change is selected as the next column value with the probability of \( 1/(1 + \exp \Delta E / T) \).

The probability of transition can be reduced by gradually lowering the temperature \( T \) as the number of calculations increases. In other words, even if the temperature falls to a local value while the temperature is high, the transition to a higher energy is stochastically allowed, so that the local value can be escaped. Also, the value of the undetermined constant \( \lambda \) is changed and the Root Mean Square (rms) error is calculated to obtain the optimum value. For example, in the case of the image used in the experiment, the optimum \( \lambda \) is obtained as shown in Table I.

**D. Image Restoration Method Based on the Conjugate Gradient Method**

The conjugate gradient method is an effective solution for simultaneous linear equations with symmetric positive definite matrices as coefficients. If the observed image is a symmetric positive definite matrix and it is regarded as a system of linear equations represented by \( g = Hf \), the image restoration can be performed using the conjugate gradient method.

The conjugate gradient method is performed by generating a vector sequence of iterative solutions (sequential approximations to the solution), a residual vector corresponding to the iterative solution, and a search direction.
vector used for updating the iterative solution and the residual vector.

Also, for each iteration, two inner product operations are performed in order to calculate the updating scalar value that is defined so that the vector sequence satisfies the orthogonal condition. In a system of linear equations with a symmetric positive definite matrix as a coefficient, these conditions mean that the distance to the true solution is minimized under a certain norm.

The $i$-th iterative solution $f'^i$ is updated in each iteration by multiplying the search direction vector $p_i$ by a multiplier $\alpha^i$.

$$ f'^i = f'^{i-1} + \alpha^i p_i $$

The corresponding residual vector $r'$ is

$$ r' = r'^{i-1} - \alpha^i \cdot H p_i $$

$$ a = ai = r'^{(i-1)^t} \cdot r^{(i-1)} p^{(i-1)^t} H p_i $$

Among all $a$ expressed by equation (12), $r^{(i)^t} H^{-1} r^{(i)}$ is selected to be minimized. The search vector group is updated as follows using the residual vector.

$$ p^i = r' \cdot p^i / p^{(i)^t} r^{(i)} $$

$$ \beta i = r'^i / p^{(i)^t} r^{(i)} $$

By choosing and $\beta$, $p_i$ and $H p_i$, the equivalent condition $r'$ and $r^i$ are orthogonal. By selecting such $\beta i$, $p_i$ and $t_i$ become orthogonal to all previous $H p_j$ and $r_j$, respectively. ($j = 1, 2, 3, \ldots$) Next, the convergence condition of the conjugate gradient method will be described. In the iteration using the conjugate gradient method, noise may be increased in one iteration due to its fast convergence.

Therefore, it is necessary to introduce some noise suppression processing. Regularized Least Square Method is frequently used as this method.

$$ e = lgn \cdot [H] g^2 + \gamma lg |t|^2 $$

is used as the error function. Where $gn$ and $\gamma$ are the measured data including noise and the power ratio of noise to signal. Here, assuming that noise is additive to the observed data $g$, such as $gn = g + n$, the first term on the right side of equation (14) corresponds to the noise power, and the second term is 1.

Since it is the product of $l / (S / N)$ and the power of the signal, this also corresponds to the power of noise. In the minimization of the error function in equation (11), the first term and the second term are kept at the same magnitude, and as a result, a solution with $S / N$ of 1 / $r$ is obtained.

If the error function of equation (11) is partially differentiated with respect to $g$ and the result is set to 0, the following is obtained.

$$ [H]^t fn = ([H]^t [H] + \gamma [I]) \cdot g $$

Comparing equations (4) and (15), it can be seen that the matrix $[H]^t [H]$ is replaced by $[H]^t [H] + \gamma [I]$. In other words, by replacing $[H]^t [H]$ with $[H]^t [H] + \gamma [I]$ in the equation expressing the iteration of the conjugate gradient method, an image restoration algorithm based on the regularized least squares criterion is obtained. That is, it can be said that the image restoration by the conjugate gradient method is an optimization method that constrains $S / N$ to be $1/\gamma$.

IV. EXPERIMENT

A. Principle of Improving Resolution

By shortening the sample interval of the low-resolution sensor (this is called high-density overlap sampling), the resolution equivalent to the sample interval can be obtained. The direction $A$ in the figure is the scanning direction of the satellite, and the direction $B$ is the traveling direction of the satellite. For example, if the overlap sample ratio is $mo$ in the satellite advancing direction and no in the satellite scanning direction (Fig. 3), the high-resolution image can be expressed by the following equation.

$$ y_{ij} = \frac{1}{mo \cdot no} \sum_{k=1}^{mo} \sum_{l=1}^{no} x_{i+k,j+l} + e_{ij} $$

where, $mo$: satellite scan direction overlap sample ratio, $no$: satellite advancing direction overlap sample ratio, $x$: observation target, $y$: overlapped sampled image, $l$: miscellaneous $j$ sound. This is expressed as a matrix as follows.

$$ Y = AX + E $$

![Fig 3. Spatial Resolution Enhancement with an Overlap Sampling.](image)

The problem of restoring a high spatial resolution image from data sampled at high density can be reduced to the problem of image restoration [7]. That is, the problem of estimating high-resolution data from data with a sample interval of less than one pixel is a problem of restoring the original image $X$ from the observed image $Y$ in consideration of the deterioration operator $A$ and noise $E$.

B. Simulation Images

In the experiment, National Oceanic and Atmospheric Administration: NOAA / Advanced Very High Resolution of Radiometer: AVHRR images (14 × 14 pixels) were used.

First, this image (original image) was multiplied by a matrix of high-density overlap samples to create an overlap sample image (observation image). Image restoration was applied to this image by the method described above to estimate the original image (estimated image or restored image), and the rms error from the original image was evaluated. Also, in that case, Table II is shown to examine the resistance to noise.
TABLE II. GAUSSIAN NOISE USED

<table>
<thead>
<tr>
<th>NOISE</th>
<th>MEAN</th>
<th>VARIANCE</th>
<th>S/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>30</td>
<td>46.7</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>60</td>
<td>33.0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>150</td>
<td>20.8</td>
</tr>
</tbody>
</table>

TABLE III. GAUSSIAN NOISE USED FOR JITTER SIMULATION

<table>
<thead>
<tr>
<th>NOISE</th>
<th>MEAN</th>
<th>VARIANCE</th>
<th>REMARKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0</td>
<td>1</td>
<td>Approx.±1% of jitter</td>
</tr>
<tr>
<td>b</td>
<td>0</td>
<td>4</td>
<td>Approx.±1% of jitter</td>
</tr>
<tr>
<td>c</td>
<td>0</td>
<td>9</td>
<td>Approx.±1% of jitter</td>
</tr>
</tbody>
</table>

The author prepared three types of noise, added them to the overlapped sample images, and then performed image restoration, and evaluated the noise characteristics of each restoration method. Furthermore, as the nonlinearity of the observation process, sampling jitter was simulated and added to the overlapped sample images.

This assumes that the electronic jitter of the sampling pulse is randomly mixed. The sampling interval is divided into 128 equal parts, the normal random numbers shown in Table III are generated, quantized in 256 (-127 to +128) steps, and the pixel position is shifted to be restored as an observed image.

C. Experimental Result

Fig. 4 to 7 show the original image, the overlapped sample image (overlap sample ratio is 4), and the restored images when the S/N is 33, respectively.

At this time, in the maximum entropy method with annealing (proposed method), as shown in Fig. 8, the rms error continues to decrease as the number of iterations increases. This is because it is theoretically guaranteed to approach the minimum solution infinitely. However, in the case of image restoration based on the conjugate gradient method, as shown in Fig. 9, when the noise is large (S/N = 33 or more), increasing the number of iterations may rather increase the rms error.

Fig 4. The Original Image (Left) and the observed Image (Right).

Fig 5. Restored Image with Noise based on MEM with the Parameter Estimation by Means of Annealing.

Fig 6. Restored Image with Noise based on Least Square Method: LSQ.
The author was able to confirm that again. Table IV shows the results of evaluating the RMS error of each method. When the S/N is 33 (Noise Type B) or higher, the restoration effect of the proposed method (image restoration based on the maximum entropy method with annealing) is superior to that based on the conjugate gradient method in terms of rms error. The author found out that from this, it can be seen that the maximum entropy method with parameter estimation by annealing is superior in terms of resistance to noise.

It was also confirmed that noise resistance cannot be expected by the least-squares method in which the constraint conditions are not optimized.

Next, in the case of no additional noise, we tried to recover from the observed image with jitter added and evaluated the rms error with the original image. The results show that, as shown in Table V, the image restoration of the proposed method is the most tolerant to jitter. As mentioned in the foreword, this is due to the fact that the maximum entropy method is the only method that considers the nonlinearity of the observation system.

This jitter type c is equivalent to the jitter based on the normal random number of ±3% of the pixel interval considering the standard deviation. In the image restoration based on the conjugate gradient method, even if the jitter is 1%, the rms error is more than doubled compared to when it is not, and it can be said that it is sensitive, but the proposed method is relatively insensitive in that respect.

V. CONCLUSION

It is confirmed that the image restoration based on the maximum entropy method with parameter estimation by annealing is effective in the noisy case. Furthermore, in the reconstruction problem of high spatial resolution images from high density sample images from NOAA / AVHRR data used in this experiment, if the S / N is better than 33, the image based on the conjugate gradient method is used. When restoration is sufficient, but when it is worse than that, it was found that image restoration based on the maximum entropy method with parameter estimation by annealing is superior.

Furthermore, it was confirmed that the resistance to noise cannot be expected by the least squares method in which the constraint conditions are not optimized. Comparing the

**TABLE IV. RMS ERROR FOR THE PROPOSED AND EXISTING METHODS (AGAINST NOISE WITHOUT JITTER)**

<table>
<thead>
<tr>
<th>Type</th>
<th>LEAST SQUARE</th>
<th>MEM WITH ANNEALING</th>
<th>CONJUGATE GRADIENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td>39</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>B</td>
<td>79</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>C</td>
<td>99</td>
<td>23</td>
<td>24</td>
</tr>
</tbody>
</table>

**TABLE V. RMS ERROR FOR THE PROPOSED AND EXISTING METHODS (AGAINST SAMPLING JITTER WITHOUT NOISE)**

<table>
<thead>
<tr>
<th>Type</th>
<th>LEAST SQUARE</th>
<th>MEM WITH ANNEALING</th>
<th>CONJUGATE GRADIENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>a</td>
<td>69</td>
<td>19</td>
<td>22</td>
</tr>
<tr>
<td>b</td>
<td>88</td>
<td>21</td>
<td>28</td>
</tr>
<tr>
<td>c</td>
<td>98</td>
<td>27</td>
<td>33</td>
</tr>
</tbody>
</table>
resilience of the sampling pulse to electronic jitter, we find that the image restoration of the proposed method is the most tolerant to jitter. This is because the maximum entropy method can consider the nonlinearity of the observation system.

VI. FUTURE RESEARCH WORKS

The proposed image restoration method is applicable to all of images. Although this time only NOAA/AVHRR remote sensing satellite images are used for the experiment, other general images have to be used for the future research works. Furthermore, it is confirmed that the proposed method takes much longer time in comparison to the other conventional image restoration methods, such as least square method, Conjugate Gradient Method: CGM. Not only restoration performance, but also required computation resources have to be evaluated in the future research works.

ACKNOWLEDGMENT

The author would like to thank Mr. Yasuhiro Minematu and Dr. Yashushi Terayama of former students of Saga University for their experimental research works. Also, the author would like to thank Prof. Dr. Hiroshi Okumura and Prof. Dr. Osamu Fukuda of Saga University for their valuable comments and suggestions.

REFERENCES


AUTHOR’S PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Technology of the University of Tokyo Press, 1991.

2013. He is now award committee member of ICSU/COSP4. He wrote 55 books and published 620 journal papers as well as 450 conference papers. He received 66 of awards including ICSU/COSPAR Vikram Sarabhai Medal in 2016, and Science award of Ministry of Mister of Education of Japan in 2015. He is now Editor-in-Chief of IJACSA and IJISA. (http://teagis.ip.is.saga-u.ac.jp/index.html)