The Astrophysical Journal, 636:1159–1165, 2006 January 10

SPATIO-SPECTRAL MAXIMUM ENTROPY METHOD. I. FORMULATION AND TEST

Su-Chan Bong,[1,2] Jeongwoo Lee,[3] Dale E. Gary,[3] and Hong Sik Yun[2]

Received 2003 October 24; accepted 2005 September 20

ABSTRACT

The spatio-spectral maximum entropy method (SSMEM) has been developed by Komm and coworkers in 1997 for use with solar multifrequency interferometric observation. In this paper we further improve the formulation of the SSMEM to establish it as a tool for astronomical imaging spectroscopy. We maintain their original idea that spectral smoothness at neighboring frequencies can be used as an additional a priori assumption in astrophysical problems and that this can be implemented by adding a spectral entropy term to the usual maximum entropy method (MEM) formulation. We, however, address major technical difficulties in introducing the spectral entropy into the imaging problem that are not encountered in the conventional MEM. These include calculation of the spectral entropy in a generally frequency-dependent map grid, simultaneous adjustment of the temperature variables and Lagrangian multipliers in the spatial and spectral domain, and matching the solutions to the observational constraints at a large number of frequencies. We test the performance of the SSMEM in comparison with the conventional MEM.

Subject headings: Sun: radio radiation — techniques: image processing — techniques: interferometric — techniques: miscellaneous — techniques: spectroscopic

1. INTRODUCTION

The maximum entropy method (MEM) is a technique for extracting as much information from a measurement as is justified by the signal-to-noise ratio of the data. The MEM is used in a variety of fields, including radio astronomy (Wernecke & D'Addario 1977), tomography (Mottershead 1995), cellular structure (Dubertret et al. 1995), and magnetohydrodynamics (Jordan & Turkington 1995). When this is applied to astronomical radio observations, where Fourier-transform imaging is used, the MEM deals with a model in the *x*-*y* plane so that its Fourier-transform pair in the *u*-*v* plane is compared with the measured visibilities for agreement. Based on information theory, the MEM uses a priori knowledge in reconstructing spatial images in addition to the quantities available from the direct measurement. Typically, the apriorities are uniformity and positivity of the map (Wernecke & D'Addario 1977; Gull & Daniell 1978; Skilling & Bryan 1984; Cornwell & Evans 1985; Narayan & Nityananda 1986; Sivia 1996), and use of such a priori information helps to reduce the ambiguity in imaging caused by insufficient *u*-*v* sampling due to the limit in the number of baselines and array configuration (Rohlfs & Wilson 1996; Cornwell et al. 1999).

In many astrophysical problems, however, it is more desirable to retrieve spatial and spectral information together. This scientific need has not been exploited by the existing radio interferometers. However, the use of multifrequency techniques in synthesis interferometry will increase with the newer generations of radio telescopes equipped with wideband receivers and back ends (e.g., the Frequency-Agile Solar Radiotelescope; Bastian 2003; White et al. 2003). The traditional multifrequency synthesis (MFS, e.g., Sault & Conway 1999), which assumes little change of source with frequency, would not fully address this issue. There are other types of MFS that are designed to address the frequency-dependent source structure using the well-established CLEAN deconvolution method (Conway & Stannard 1975; Braun et al. 1987; Murphy 1988; Conway et al. 1990). These techniques are, however, targeted to a relatively narrow range of frequencies (typically 10%). More ideally, we seek an algorithm that retains the spectral variation of intensity at a spatial point and uses that information in determining the brightness temperature at the point to achieve wideband (extending over an octave) imaging spectroscopy. In this regard, we note that Komm et al. (1997, hereafter KHG97) presented an algorithm of spatio-spectral maximum entropy method (SSMEM), developed in an application to the multifrequency interferometric data gathered with the Owens Valley Solar Array (OVSA), which extracts the spectral, as well as spatial, information under the principle of the MEM.

The goal of this paper is to establish a working SSMEM algorithm suitable for wideband imaging spectroscopy by refining the original formulation by KHG97. The essence of the KHG97 algorithm lies in so-called spectral entropy, which is explicitly designed to force the spectrum at a given spatial position to run smoothly with frequency by communicating with intensities at neighboring frequencies as a local power law. Since it is not the smoothness but the uniformity that the traditional entropy ensures, the spectral entropy does not really conform to the MEM principle and hence requires justification. We argue that astronomical continuum spectra usually appear in a power-law function of frequency (e.g., Dulk 1985), and therefore this knowledge can serve as additional information in spectral reconstruction, whereas there is no such specific a priori information known for spatial images. The power-law index itself is, of course, known only after the multifrequency imaging is completed. Since it is still an ad hoc assumption that the spectral entropy term can add to the traditional entropy term, we have to justify our approach based on the test result of the proposed algorithm using a model of choice.

The plan of this paper is as follows: In § 2 we present the formulation of the object function of the SSMEM. We discuss optimization techniques and convergence criteria for use within the SSMEM in § 3. We present the result of a performance test in § 4, and finally conclude in § 5.

[1] Korea Astronomy and Space Science Institute, Hwaam-dong Yuseong-gu, Daejeon 305-348, South Korea; scbong@kasi.re.kr.

[2] Astronomy Program, School of Earth and Environmental Science, Seoul National University, Seoul 151-747, South Korea; yun@astrosun.snu.ac.kr.

[3] Physics Department, New Jersey Institute of Technology, 161 Warren Street, Newark, NJ 07102; leej@njit.edu, dgary@njit.edu.

2. AN EXTENSION OF MEM TO SSMEM

The conventional MEM solves an imaging problem by maximizing the spatial entropy subject to any constraints involved with the imaging. In the Lagrangian formulation, this is equivalent to maximizing the object function, which contains the entropy and the constraint terms. After Cornwell & Evans (1985), the object function, J, in the following form has often been used:

$$J = H - \alpha\chi^2 - \beta F, \tag{1}$$

where α and β are the Lagrangian multipliers and the quantities H, χ^2, and F are defined as

$$H = -\sum_j T_j \ln\left(\frac{T_j}{m_j e}\right), \tag{2a}$$

$$\chi^2 = \sum_i \frac{\left|V_i - V_i'\right|^2}{\sigma_i^2} - n_V, \tag{2b}$$

$$F = \sum_j T_j - F'. \tag{2c}$$

Here T_j and m_j are, respectively, the map temperature and default temperature at the jth pixel and V_i' is the visibility measured at the ith baseline. Here j represents an index over map pixels, running in the range of $1 \le j \le n_T$, and i denotes an index over interferometric baselines in $1 \le i \le n_V$, where n_T is the total number of the map pixels and n_V is the total number of the baselines. In the case of rotational synthesis, i includes both baselines and times. Under this formulation, T_j values are adjusted until the corresponding object function J reaches its maximum. At the same time, the Lagrangian multipliers, α and β, are also updated to keep the resulting χ^2 and F in agreement with their expectation values, i.e., zero. As a result, we end up with a solution map T_j that has the maximum entropy while satisfying the constraints.

To extend the MEM to the SSMEM, we define a new object function that represents the entropy of the system in both spatial and frequency space. We compare this situation to the grand canonical ensemble as opposed to the canonical ensemble in statistical mechanics. Namely, the above object function J at an individual frequency that we now call the spectral object function J_k is like a canonical ensemble, and the grand object function is defined in the grand canonical ensemble. We thus use two indices, the usual index j for the spatial grids and the other k for frequency,

$$\sum_j \rightarrow \sum_{j=1}^{n_T}\sum_{k=1}^{n_\nu} \quad \text{and} \quad T_j \rightarrow T_{jk},$$

where n_ν is the number of frequencies and n_T is again the number of map pixels, assumed to be independent of the frequency. Now under this index convention, we represent the grand object function modified to the form

$$\begin{aligned} J &= \sum_k J_k \\ &= \sum_k H_k - \sum_k \alpha_k \chi_k^2 - \sum_k \beta_k F_k + \gamma \sum_k S_k. \end{aligned} \tag{3}$$

The last term, the new feature in the SSMEM, consists of the spectral entropy S and the associated constant γ, a measure of the relative importance of the spectral entropy to the spatial entropy. Since S_k is an entropy term, we assign the positive sign to it and take the dimensionless constant γ as a user-supplied parameter rather than a Lagrangian multiplier.

What makes this extension of the MEM to the SSMEM nontrivial is the spectral entropy that KHG97 introduced in the following form:

$$S = \sum_k S_k = -\sum_{k=1}^{n_\nu}\sum_{j=1}^{n_T} \tau_{jk} \ln\left(\frac{\tau_{jk}}{m_{jk} e}\right), \tag{4}$$

where

$$\tau_{jk} = m_{jk} + \left|T_{jk} - \left\langle T_{jk}\right\rangle\right|. \tag{5}$$

In this expression, m_{jk} is the default map at position j and frequency k, for which we use a spatially flat map satisfying the flux constraint at each frequency, and $\langle T_{jk}\rangle$ is the logarithmically interpolated temperature from adjacent frequencies, and therefore τ_{jk} is a measure of how much a temperature T_{jk} at a given frequency and spatial point differs from the one that smoothly connects temperatures at nearby frequencies at the same spatial point. Because the microwave spectrum from the solar active region is expected to follow a power law in most of the frequency range, we can expect that the difference $\left|T_{jk} - \langle T_{jk}\rangle\right|$ at each frequency should be minimized in the reconstruction. Were there no other constraint, S_k would be maximized when $T_{jk} = \langle T_{jk}\rangle$, and the spectrum at a given spatial position would follow a power-law function of frequency as we intended (see § 1).

Suppose $\tilde{T}_{jk}^{\pm}$ denotes brightness temperatures belonging to the two neighboring frequencies $\nu_{k\pm1}$ at the same spatial location $\boldsymbol{x}_{jk}$. The quantity $\langle T_{jk}\rangle$ is then found by logarithmic interpolation along the frequency axis,

$$\ln\left\langle T_{jk}\right\rangle = a_k^+ \ln\tilde{T}_{jk}^+ + a_k^- \ln\tilde{T}_{jk}^-, \tag{6}$$

with

$$a_k^+ = \frac{\ln(\nu_k/\nu_{k-1})}{\ln(\nu_{k+1}/\nu_{k-1})},$$
$$a_k^- = 1 - a_k^+,$$

which will ensure power-law spectral behavior. As a complication, $\tilde{T}_{jk}^{\pm}$ are not generally contained in the $T_{jk'}$, but must be interpolated to the positions $\tilde{\boldsymbol{x}}_{jk}^{\pm}$, which match the spatial point $\boldsymbol{x}_{jk}$ but belong to $k' = k \pm 1$ frequency slices. We illustrate how to compute $\langle T_{jk}\rangle$ in Figure 1 using a temperature T_{12} at position x_1 and frequency ν_2 as an example. As Figure 1*a* shows, this complication arises only because the pixel size of radio maps, to be optimal for mapping, should vary continuously in proportion to the inverse of the frequency, i.e., $\tilde{\boldsymbol{x}}_{jk}^{\pm} \neq \boldsymbol{x}_{jk'}$. KHG97 simplified this problem by dividing the whole frequency into three spectral regimes within each of which the same spatial resolution is shared among different frequencies. We here maintain the continuous variation of map resolution, and therefore the temperatures $\tilde{T}_{jk}^{\pm}$ at neighboring frequencies should be interpolated from spatially neighboring temperatures at each frequency plane $\nu_{k\pm1}$. To determine $\tilde{T}_{jk}^{\pm}$, we introduce a bilinear interpolation factor $R_{jkj'k'}^{\pm}$ such that

$$\tilde{T}_{jk}^{\pm} = \sum_{j'k'} R_{jkj'k'}^{\pm} T_{j'k'}, \tag{7}$$

where $R_{jkj'k'}^{\pm}$ is nonzero only when $k' = k \pm 1$ and at four points j' surrounding the point $\tilde{\boldsymbol{x}}_{jk}^{\pm}$, which matches $\boldsymbol{x}_{jk}$. Figure 1*b* shows

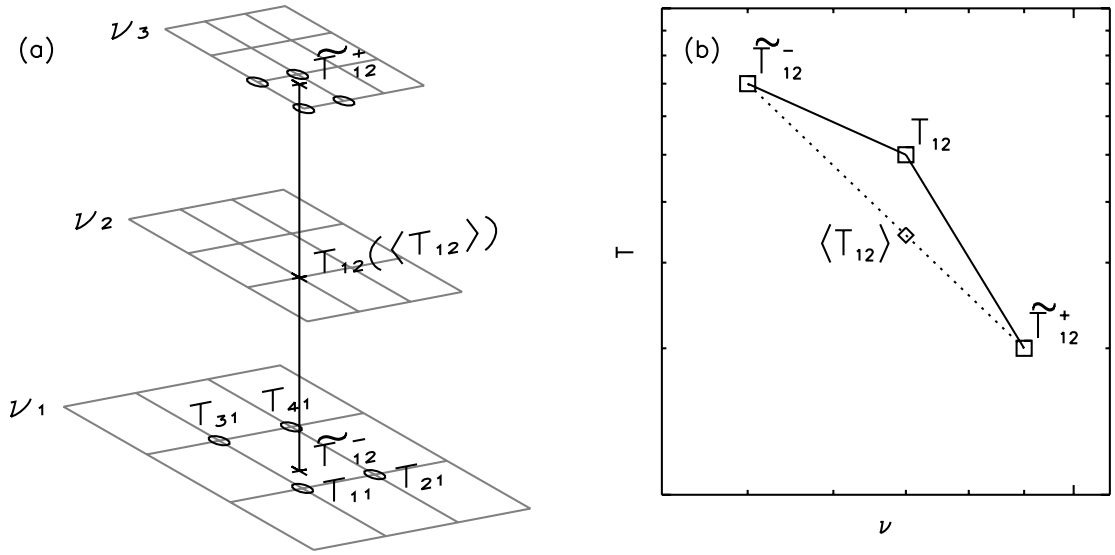

FIG. 1.—Illustration of the reference temperature $\langle T \rangle$. (*a*) Map grids at three frequencies (ν_1, ν_2, and ν_3). We denote the temperature at position grid point 1 at frequency ν_2 as T_{12}, for instance. While T_{12} is immediately found on the map grid at ν_2, $\langle T_{12} \rangle$ should, however, be interpolated from $\tilde{T}_{12}^{-}$ at ν_1 and $\tilde{T}_{12}^{+}$ at ν_3 at the position of $\boldsymbol{x}_{12}$. Since the latter two quantities are not always found on their grids, they should be interpolated from neighboring temperatures. Here $\tilde{T}_{12}^{-}$ is bilinearly interpolated from four adjacent grid points (T_{11}, T_{21}, T_{31}, and T_{41}), and likewise $\tilde{T}_{12}^{+}$ is from four adjacent grid points (*circles in the map grids at* ν_3). (*b*) After $\tilde{T}_{12}^{\pm}$ are determined, $\langle T_{12} \rangle$ is logarithmically interpolated from them (*dotted line*). The difference $|T_{12} - \langle T_{12} \rangle|$ is now a measure for the deviation of the spectrum from a power law and used in defining the spectral entropy.

$\langle T_{jk} \rangle$ relative to T_{jk} in the spectrum, which is used as a measure of the spectral entropy.

It is important that we are able to express the spectral entropy in an analytic form using the constant factors $R_{jkj'k'}^{\pm}$. This allows us to express the first and second derivatives of the spectral entropy term also in analytic forms even when a frequency-dependent map grid is used. The expressions for the gradients are given in the Appendix.

3. OPTIMIZATION TECHNIQUES

Optimization of the entropy consists of the maximization of the object function and adjustment of the Lagrangian multipliers so that the constraints are satisfied. We express these two steps in terms of two sets of equations,

$$\nabla_T J = 0 \quad \text{and} \quad \left\{\chi_k^2, F_k\right\} = 0. \tag{8}$$

As way to solve these equations, Sault (1990) suggested combining the above two sets into the following single set of equations:

$$\nabla_\xi J = 0. \tag{9}$$

Here $\xi = \left\{T_{jk}, \alpha_k, \beta_k\right\}$ and equation (8) could be converted to equation (9) because the second set of equations in equation (8) can alternatively be written as $\{\partial J/\partial \alpha_k, \partial J/\partial \beta_k\} = 0$. Equations such as equation (9) can conveniently be solved using the Newton-Raphson method as suggested by Sault (1990),

$$\delta\xi = -\left(\nabla_\xi \nabla_\xi J\right)^{-1} \cdot \nabla_\xi J. \tag{10}$$

Note that up to this point, only a trivial difference between Sault (1990) and the present study is found in the additional frequency index k.

A technical issue in using the Newton-Raphson method is that the Hessian matrix $\nabla_\xi \nabla_\xi J$ is so huge in size as to make full calculation undesirable. Sault (1990) suggested, following Cornwell & Evans (1985), using a modified Newton-Raphson method in which only the diagonal elements of the matrix are used. The effect of missing off-diagonal elements can be accommodated by a constant boosting factor. We use the same approach for the SSMEM. Note, however, that this approximation can be justified for the Hessian matrix of the χ^2 term but is yet to be verified for the spectral entropy term present in our SSMEM. At this stage, we proceed without any justification and rely on the resulting solution map to see whether such diagonalization approximation can be used with the present SSMEM.

To set the criterion for map convergence, the magnitude of $\nabla_T J$ needs to be defined. Since this is a vector quantity and the criterion should also be met at each frequency, we define the magnitude of the vector in the row corresponding to the kth frequency as

$$|\nabla_T J|_k = \sqrt{\sum_{j=1}^{n_T} \left(\frac{\partial J}{\partial T_{jk}}\right)^2}.$$

We want this quantity to be smaller than a small fraction of that of the constituent terms,

$$|\nabla_T J|_k \le \epsilon \sqrt{|\nabla_T H_k|^2 + \alpha_k^2 |\nabla_T \chi_k^2|^2 + \beta_k^2 |\nabla_T F_k|^2 + \gamma^2 |\nabla_T S|_k^2}, \tag{11}$$

where we set ϵ to a small value, $\epsilon = 0.03$.

For the criteria of the constraints, we can consider the constraints met if the absolute values $\left\{\left|\chi_k^2\right|, |F_k|\right\}$ are small compared with $\left\{n_{Vk}, F_k'\right\}$ or comparable to their uncertainties. We thus set the criteria for the constraints as

$$\left|\chi_k^2\right| \le \max\left(\epsilon n_{Vk}, \delta\chi_k^2\right), \tag{12a}$$

$$|F_k| \le \max\left(\epsilon F_k', \delta F_k\right), \tag{12b}$$

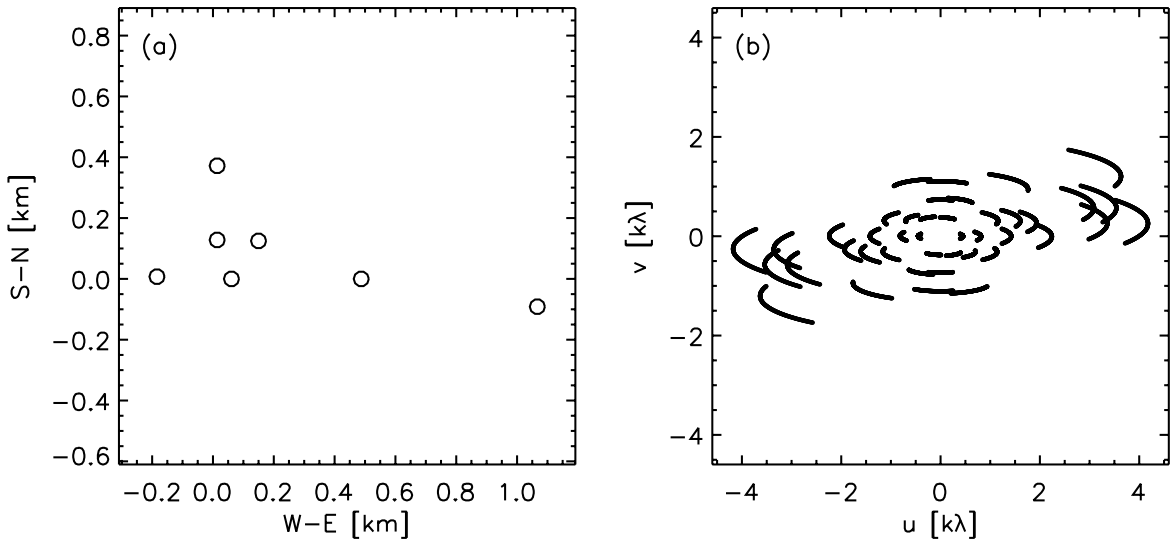

FIG. 2.—Models used in the test run. (*a*) Model array set in resemblance with the actual antenna configuration of OVSA. Geographic north is up, and east is to the right. Each circle represents an antenna position. (*b*) The *u*-*v* coverage during 4 hr rotational synthesis resulting from the antenna configuration. Here *u* points west and *v* points north.

for every frequency k. Now the uncertainties in the constraints χ^2_k and F_k would simply be

$$\delta\chi^2_k = \sqrt{2n_{Vk}}, \tag{13a}$$

$$\delta F_k = \delta F'_k, \tag{13b}$$

where $(2n_{Vk})^{1/2}$ is the standard deviation of the χ^2 distribution and $\delta F'_k$ is the uncertainty of F'_k given as an input parameter.

4. PERFORMANCE TEST

In this section we test the efficiency of the algorithms and the quality of the resulting maps and spectra using models for radio interferometric array configuration and radiation sources.

4.1. *Procedure*

We choose a model interferometric array that resembles the current OVSA system as shown in Figure 2*a*. A total of 21 *u*-*v* points are obtained from such a configuration in a snapshot mode at one frequency. For the rotational synthesis, we take *u*-*v* points every minute obtaining 1260 *u*-*v* points per frequency for 1 hr. Figure 2*b* shows the corresponding *u*-*v* distribution at a single frequency (∼5000 points) for 4 hr rotational synthesis. Note that the model array has a very sparse and asymmetric *u*-*v* distribution, and we create the model visibilities with frequency-dependent noise, making the spatio-spectral reconstruction more challenging. Our main concern is how each method handles this situation.

For our purpose the model source should have a frequency-dependent morphology. We create the radiation source in close resemblance to a real source in the Sun. The microwave radiating electrons in the Sun are trapped in coronal magnetic loops connecting two footpoints. At low frequencies the whole loop is optically thick, and thus the source appears as single large object with a temperature close to the coronal temperature. Toward high frequencies the loop top becomes optically thin, and thus the source size decreases with frequency and shrinks to two sources located at footpoints where magnetic fields are stronger (see, e.g., Kundu et al. 1989; Gary & Hurford 1994). We simulate this frequency-dependent behavior using a combination of two Gaussian sources whose size and temperature vary with frequency. The model sources at selected frequencies are shown in the top panels of Figure 3. At intermediate frequencies, the source morphology gradually changes from one to another.

To simulate observed visibilities, we need to add noise. We first generate noiseless model temperatures t_{jk}, which are then fast Fourier-transformed into the *u*-*v* plane and interpolated into the above *u*-*v* points shown in Figure 2 to obtain the true observed visibilities, V^t_{ik}. We then create the observed visibilities V'_{ik} by adding noise at each baseline i and frequency k, according to the rule

$$V'_{ik} = V^t_{ik} + N\sigma_{ik}, \quad \sigma_{ik} = 0.03\left|V^t_{ik}\right| + 0.02, \tag{14}$$

where N is random noise of unit variance and σ_{ik} is the standard deviation of the data in units of solar flux units (sfu).

4.2. *Performance of SSMEM*

We use 20 frequencies between 1.4 and 18.0 GHz and a map size of typically 128 by 128 pixels. We run the SSMEM for three different values of γ. Setting $\gamma = 0$ in the present formulation makes the SSMEM reduce to the standard MEM. On the other hand, a high value of γ represents a case in which the role of spectral smoothness is more emphasized than that of the usual spatial entropy. This comparison therefore demonstrates how well the spectral entropy term functions in the present SSMEM.

The results are shown in Figure 3. The top panels (Figs. 3*a*–3*c*) show the two-dimensional maps reconstructed with three values of γ (*contours*) at three frequencies in comparison with the true source (*gray-scale image*). In spite of the small number of baselines along the north-south direction, the round shape of the Gaussian sources is fairly well reconstructed at all frequencies regardless of the adopted γ-value. All results thus show little difference in terms of the two-dimensional morphology. To check in more detail, we plot, in the middle panels (Figs. 3*d*–3*f*), the one-dimensional scans of the maps along the $y = 0''$ line. It appears that the maps obtained with $\gamma = 3$ are closest to the true source at some frequencies but give rather higher temperatures than the true source at 5.8 GHz. The overestimation must be an artifact and thus is unwanted. To compare the SSMEM with the MEM, we can see that the temperatures reconstructed under the

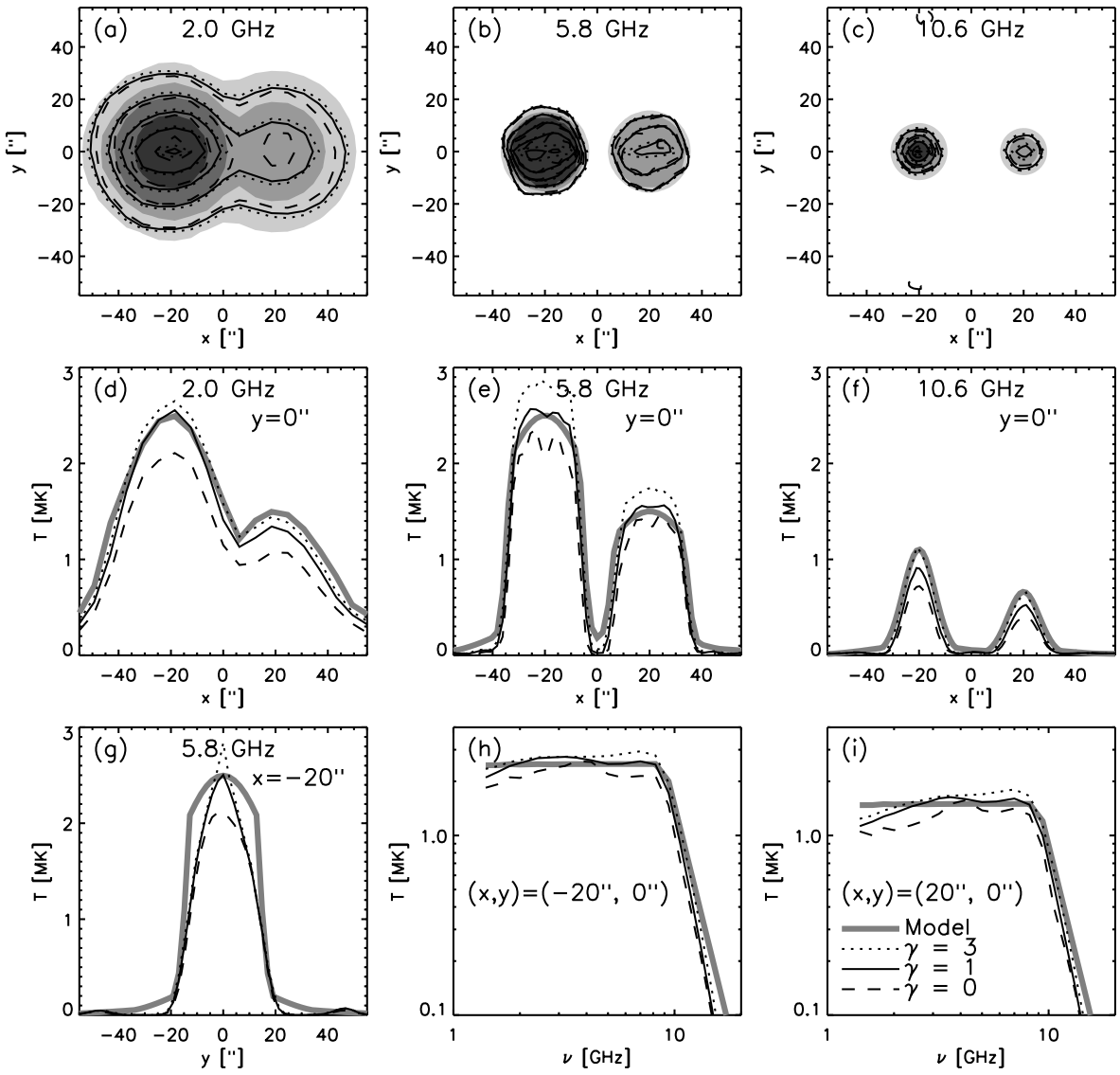

FIG. 3.—Test run of the SSMEM using differing values of γ. The top panels show maps at (*a*) 2.0 GHz, (*b*) 5.8 GHz, and (*c*) 10.6 GHz. The true source is shown as a gray-scale image, and the contours are the reconstructed images: the solid lines now represent the results obtained using $\gamma = 1$, and the dashed and dotted lines show the results for $\gamma = 0$ (MEM) and $\gamma = 3$ (SSMEM with emphasis on spectral entropy). The contour levels increase by 20% of the maximum of the true source. Panels (*d*)–(*f*) show one-dimensional scans of the maps shown in the top panels along $y = 0''$. Panel (*g*) is a one-dimensional scan of the maps at $x = -20''$, while panels (*h*) and (*i*) are the local spectra measured at the position of two local maxima: $(x, y) = (-20'', 0'')$ and $(20'', 0'')$, respectively.

MEM are lower than those of the true source, while the SSMEM produces temperatures closer to the true source at all frequencies. At 5.8 GHz, the MEM map contains more artificial structure than the SSMEM maps, indicating that the reconstructed map deviates more from the true source, and the spectrum also disagrees with the model (see Fig. 3*h*). Figure 3*g* shows the one-dimensional scan along $x = -20''$, where the *u*-*v* points are relatively deficient. In this case, the MEM solution deviates more from the source, whereas the SSMEM solution with $\gamma = 1$ approaches the true source more closely.

The advantage of the SSMEM over the MEM is more obvious in terms of the spectral reconstruction. The last two panels, Figures 3*h*–3*i*, compare the local spectra reconstructed at the center of the two sources with the model spectra. The MEM spectra show more unwanted fluctuations, whereas both SSMEM spectra appear smoother and have closer resemblance to the original spectra at both locations. This demonstrates that the spectral entropy in the SSMEM plays the intended role of information exchange across neighboring frequencies so that the resulting spectra appear as a more power-law–like spectrum than in the MEM. We, however, should note that the spectra reconstructed by all methods commonly fall more rapidly than the original model toward the lowest and the highest frequencies. This is primarily because the locations of the *u*-*v* points at these frequencies are inadequate to fully characterize the visibility distribution in *u*-*v* space. Therefore, the SSMEM is not immune to the problem of insufficient *u*-*v* sampling. We can only expect that the SSMEM provides a relatively better spectrum than the MEM in the frequency range where a moderate amount of *u*-*v* information is available.

KHG97 pointed out that increase of γ may lead to a problem called "spectral oversmoothing." Our SSMEM solutions do not show such behavior, and we suspect that this owes to the frequency-dependent Lagrangian multipliers and convergence criteria, which were underdeveloped in the KHG97 formulation but are now working properly. Instead we see the reconstructed

temperature at strong sources getting higher as we increase γ, which is certainly an unwanted artifact. We thus understand that the expected role of the spectral entropy term is an auxiliary improvement of the local spectrum by sharing information at adjacent frequencies and that the factor γ should be set at a modest value (i.e., $1 \leq \gamma \leq 2$) so as not to interfere with the function of the spatial entropy.

The present calculation was performed using IDL 6.0 on a Windows 2000 PC having a 500 MHz Pentium III CPU and 256 MB of RAM. In the present calculation, we use 20 frequencies between 1.4 and 18.0 GHz and a map size of typically 128 by 128 pixels. When 4 hr synthesis (about 5000 *u-v* points per frequency) is used, the total time required to finish the job is 14,000, 24,000, and 22,000 s when $\gamma = 0$, 1, and 3, respectively. When $\gamma = 0$ the entropy term calculation is faster because the program skips over the calculation related with the spectral entropy. Comparing the $\gamma = 1$ and 3 cases, however, we cannot see any significant change of the required time with larger γ. A similar trend continues when we adopt a much poorer, snapshot *u-v* coverage (21 *u-v* points per frequency), in which it took 990, 2800, and 3400 s for $\gamma = 0$, 1, and 3, respectively.

5. CONCLUSION

We have presented the formulation and implementation of an algorithm called SSMEM that we developed as a tool for astronomical Fourier-transform imaging at multiple frequencies over a broad spectral range. The present formulation of the SSMEM, when compared with the initial formulation of KHG97, has a number of improvements as follows: (1) Computation of spectral entropy is refined by introducing continuous spatial resolution in the *x-y* plane, which is optimized for each frequency for the best use of the frequency variation. (2) The Lagrangian multipliers are also set for individual frequencies, and the flux constraint is introduced for the reason that Cornwell & Evans (1985) raised. These two treatments ensure a map at each frequency that is consistent with the observations. (3) The convergence criteria are set for each object function at each frequency in order to ensure that the correct solution is found. Two quantities yet remain ad hoc in the present SSMEM formulation: the spectral entropy and the dimensionless parameter, γ, introduced to lever the spectral entropy. We have justified the spectral entropy with a practical need, namely, its functioning as complementing the conventional MEM in multifrequency imaging through information exchange across frequencies. We suggest a modest value of $\gamma \sim 1$ for adequate balance between the roles of spatial and spectral entropies because the spectral entropy itself is designed to make a contribution to the solution similar to the one made by the spatial entropy. The optimal value of γ may, however, depend on the data. For this problem, it is a good practice to check the relative contribution of the spectral and spatial entropies to the resulting SSMEM maps by comparison with the MEM maps.

In our test with an uneven and sparse array equipped with 20 frequencies across 1–18 GHz and a modest amount of noise, the SSMEM maps are found to be closer to the true model than the MEM maps in regions where *u-v* information is insufficient. Otherwise similar map qualities are obtained, but the SSMEM reproduces the local spectra of the model better than the MEM. Therefore, the advantage of the SSMEM over the MEM is more obvious in terms of the spectral reconstruction. We expect that such merit of the SSMEM will remain even when *u-v* coverage is significantly better than that adopted in the present study. The reason is that there is always irreducible noise in data, and a systematic treatment of the frequency-dependent constraints will improve reconstruction of local spectra. In addition, wideband imaging spectroscopy with a fixed array will inevitably encounter the problems of inadequate *u-v* coverage at some frequencies (unless the source variation with frequency scales in proportion to the wavelength) and different-sized resolution elements involved with each map. Our results demonstrate that this problem can be addressed by the SSMEM, in that the spectral entropy plays the necessary role of communicating across frequencies. The primary application of the present algorithm could be astronomical imaging spectroscopy in which the imaging is done with an interferometer array at multiple frequencies. Good examples are the solar radio interferometric arrays designed to operate at a wide range of frequencies, such as the Owens Valley Solar Array (Hurford et al. 1984) and the Frequency-Agile Solar Radiotelescope (Bastian 2003; White et al. 2003).

We are very grateful to an anonymous referee who suggested significant modification of our original paper to the present form. We also appreciate the help from Rudi W. Komm, who had initiated this project. One of us (S. B.) was financially supported by the BK21 Project of the Ministry of Education, KOSEF grant R14-2002-043-0100-0, and the MOST Fund of M1-0407-00-0001 in Korea. J. L. has been supported by NSF grants AST 01-38317 and AST 98-7366. The OVSA is supported by NSF grant AST 03-07670 and NASA grant NAG5-11875 to New Jersey Institute of Technology.

APPENDIX

THE GRADIENT OF THE OBJECT FUNCTION IN SSMEM

In order to maximize the object function J we require analytical expressions for ∇J and $\nabla\nabla J$. We here show only those of the spectral entropy term because those of the other terms (spatial entropy, χ^2 constraint, and the flux constraint) are well known. The first derivative of the spectral entropy is

$$\frac{\partial}{\partial T_{jk}}\left[\sum_{k'} S_{k'}\right] = \frac{\partial}{\partial T_{jk}}\left[-\sum_{j'k'} \tau_{j'k'} \ln\left(\frac{\tau_{j'k'}}{m_{j'k'}e}\right)\right] = -\sum_{j'k'} \ln\left(\frac{\tau_{j'k'}}{m_{j'k'}}\right)\frac{\partial \tau_{j'k'}}{\partial T_{jk}}. \tag{A1}$$

The last term is further expanded to

$$\frac{\partial \tau_{j'k'}}{\partial T_{jk}} = \pm\left[\delta_{jj'}\delta_{kk'} - a^{+}_{k'}R^{+}_{j'k'jk}\frac{\langle T_{j'k'}\rangle}{\tilde{T}^{+}_{j'k'}} - a^{-}_{k'}R^{-}_{j'k'jk}\frac{\langle T_{j'k'}\rangle}{\tilde{T}^{-}_{j'k'}}\right]. \tag{A2}$$

Here the upper, plus sign is for $T_{j'k'} > \langle T_{j'k'} \rangle$ and the lower, minus sign is for $T_{j'k'} < \langle T_{j'k'} \rangle$. If $T_{j'k'} = \langle T_{j'k'} \rangle$ then the gradient vanishes, making the gradient continuous.

The second derivative of the spectral entropy is

$$\frac{\partial^2 S}{\partial T_{jk}^2} = -\frac{\partial}{\partial T_{jk}} \left[\sum_{j'k'} \ln\left(\frac{\tau_{j'k'}}{m_{j'k'}}\right) \frac{\partial \tau_{j'k'}}{\partial T_{jk}} \right]$$
$$= -\sum_{j'k'} \left[\frac{1}{\tau_{j'k'}} \left(\frac{\partial \tau_{j'k'}}{\partial T_{jk}}\right)^2 + \ln\left(\frac{\tau_{j'k'}}{m_{j'k'}}\right) \frac{\partial^2 \tau_{j'k'}}{\partial T_{jk}^2} \right], \tag{A3}$$

where the last term becomes

$$\frac{\partial^2 \tau_{j'k'}}{\partial T_{jk}^2} = \mp \frac{\partial^2 \langle T_{j'k'} \rangle}{\partial T_{jk}^2}$$
$$= \mp \langle T_{j'k'} \rangle \left[a_{k'}^{+} \left(a_{k'}^{+} - 1\right) \left(\frac{R_{j'k'jk}^{+}}{\tilde{T}_{j'k'}^{+}}\right)^2 + a_{k'}^{-} \left(a_{k'}^{-} - 1\right) \left(\frac{R_{j'k'jk}^{-}}{\tilde{T}_{j'k'}^{-}}\right)^2 \right]. \tag{A4}$$

REFERENCES

Bastian, T. S. 2003, Adv. Space Res., 32, 2705
Braun, R., Gull, S. F., & Perley, R. A. 1987, Nature, 327, 395
Conway, J. E., Cornwell, T. J., & Wilkinson, P. N. 1990, MNRAS, 246, 490
Conway, R. G., & Stannard, D. 1975, Nature, 255, 310
Cornwell, T. J., Braun, R., & Briggs, D. S. 1999, in ASP Conf. Ser. 180, Synthesis Imaging in Radio Astronomy II, ed. G. B. Taylor, C. L. Carilli, & R. A. Perley (San Francisco: ASP), 151
Cornwell, T. J., & Evans, K. F. 1985, A&A, 143, 77
Dubertret, B., Rivier, N., & Schliecker G. 1995, in Maximum Entropy and Bayesian Methods, ed. K. M. Hanson & R. N. Silver (Dordrecht: Kluwer), 285
Dulk, G. A. 1985, ARA&A, 23, 169
Gary, D. E., & Hurford, G. J. 1994, ApJ, 420, 903
Gull, S. F., & Daniell, G. J. 1978, Nature, 272, 686
Hurford, G. J., Read, R. B., & Zirin, H. 1984, Sol. Phys., 94, 413
Jordan, R., & Turkington, B. 1995, in Maximum Entropy and Bayesian Methods, ed. K. M. Hanson & R. N. Silver (Dordrecht: Kluwer), 347
Komm, R. W., Hurford, G. J., & Gary, D. E. 1997, A&AS, 122, 181 (KHG97)
Kundu, M. R., White, S. M., & Schmahl, E. J. 1989, Sol. Phys., 121, 153
Mottershead, C. T. 1995, in Maximum Entropy and Bayesian Methods, ed. K. M. Hanson & R. N. Silver (Dordrecht: Kluwer), 425
Murphy, D. W. 1988, Ph.D. thesis, Univ. Manchester
Narayan, R., & Nityananda, R. 1986, ARA&A, 24, 127
Rohlfs, K., & Wilson, T. L. 1996, Tools of Radio Astronomy (2nd ed.; Berlin: Springer)
Sault, R. J. 1990, ApJ, 354, L61
Sault, R. J., & Conway, J. E. 1999, in ASP Conf. Ser. 180, Synthesis Imaging in Radio Astronomy II, ed. G. B. Taylor, C. L. Carilli, & R. A. Perley (San Francisco: ASP), 419
Sivia, D. S. 1996, Data Analysis: A Baysian Tutorial (New York: Oxford Univ. Press)
Skilling, J., & Bryan, R. K. 1984, MNRAS, 211, 111
Wernecke, S. J., & D'Addario, L. R. 1977, IEEE Trans. Comp., C-26, 351
White, S. M., Lee, J., Aschwanden, M. J., & Bastian, T. S. 2003, in Proc. SPIE, 4853, 531