IEEE Pacific Rim Conference on Communications, Computers and Signal Processing
June 1st - 2nd, 1989

MAXIMUM ENTROPY IMAGE PROCESSING USING TRANSFORM DOMAIN CONSTRAINTS

Cedric A. Zala, Ian Barrodale, and Carmen E. Lucas

Barrodale Computing Services Ltd.,
Suite 200, 1677 Poplar Ave.,
Victoria, B. C., Canada V8P 4K5

Robert F. MacKinnon

Defence Research Establishment Pacific,
FMO, Victoria, B. C., Canada V0S 1B0.

Abstract

The maximum entropy (ME) method is currently of considerable interest in image processing applications. In this technique, the ME model for the image is constructed by maximizing the entropy function $-\sum_{j=1}^{N} x_j \ln(x_j/p_j)$ subject to constraints provided by the data, where x_j are the unknown model pixel amplitudes and p_j constitute prior knowledge of the pixel amplitudes. We describe here a novel formulation of the ME method, where the data constraints are expressed in the form of fixed bounds on the elements of an orthogonal transform of the model. The bounds are set on the basis of both the observed data and an estimate of the noise statistics in the transform domain; prior knowledge, if available, may also be incorporated. Using a special-purpose conjugate gradient algorithm developed for this problem [2], we present here 1-D examples which illustrate substantial SNR enhancement using the new formulation with both Fourier and Walsh transforms. A simple strategy for selecting an initial feasible solution for the algorithm is also presented: *the transform of the initial feasible solution is set as close to the transform of the prior knowledge as permitted by the constraints.* This is shown to be the least squares solution to the constrained problem with the objective function: $\sum_{j=1}^{N}(x_j - p_j)^2$; it frequently provides a very close approximation to the final ME solution.

INTRODUCTION

Maximum entropy (ME) methods are currently of great interest in signal and image reconstruction applications [1]-[11]. ME methods have achieved widespread use due to certain desirable properties such as consistency and the ability to incorporate prior knowledge effectively. ME provides a unique solution for inverse or reconstruction problems which are ill-posed: i.e., many potential solutions are consistent with the data. The ME solution is appropriate in that it introduces the least bias, or is maximally non-committal relative to the prior knowledge. Thus it may be hoped that artificial structure will be minimized in an ME solution.

In ME methods, the entropy function of the unknown model pixel amplitudes x_j with prior knowledge pixel amplitudes p_j, defined as

$$S = S(\mathbf{x}) = -\sum_{j=1}^{N} x_j \ln(x_j/p_j) \tag{1}$$

is maximized subject to constraints imposed by the data and positivity constraints on x_j and p_j. In the absence of data constraints, the maximum of S occurs when each $x_j = p_j/e$; i.e., the unconstrained output is a scaled version of the prior knowledge. When in addition, no prior knowledge exists (all p_j are constant), the output x_j are all equal and the image is uniform. At the other extreme, as the constraints imposed by the data are tightened, the ME output converges to the data, no matter what the prior knowledge. With data constraints of intermediate tightness, the ME solution is defined by both the data and the prior knowledge. Thus the form and tightness of the constraints determine the balance between the measured data and the prior knowledge in the output ME image. In practice, the noise statistics of the image are used to set the constraints.

In a current widely used formulation of the data constraints [1], the chi-square function

$$\chi^2 = \sum_{j=1}^{N} \left[\frac{(x_j - d_j)}{\sigma_j}\right]^2 \tag{2}$$

is constrained to be less than N, the number of pixel elements, where d_j are the data measurements and σ_j are the standard deviations for the noise on the image. The chi-square may be defined in either the spatial or Fourier frequency domains. This formulation has been successfully applied to image processing problems in a wide variety of fields.

In order to take advantage of signal-to-noise ratio (SNR) enhancement, an alternative formulation which is appropriate for band-limited data involves lower and upper bounds constraints in the transform domain. The bounds may be set on the basis of the transform of the data and an estimate of the noise amplitude spectrum. In this way, transform regions of high SNR are relatively tightly constrained while those of low SNR are relatively loosely constrained. The ME criterion is then used as a means of obtaining a consistent and unique solution subject to these bounds constraints, and incorporating prior knowledge. A preliminary report on this formulation has appeared in [10], where some 2-D illustrations of broadband images are provided.

In this paper, we describe the bounds formulation of the ME problem and outline an algorithm [2] for computing its solution. The performance of the method for 1-D synthetic examples is illustrated using both Fourier and Walsh transforms. The application considered here is that of signal reconstruction with noisy data.

FORMULATION AND SOLUTION OF THE ME BOUNDS PROBLEM

In the bounds formulation of the ME method, the entropy function (Eq. 1) is maximized subject to bounds constraints of the form

$$l_j \le \hat{x}_j \le u_j \quad \text{for } j = 1, 2, \ldots, N, \tag{3}$$

where the l_j and u_j are fixed lower and upper bounds, respectively, and the $\hat{x}_j$ are elements of an orthogonal transform of the model x_j. It may be shown using convexity arguments that provided the constraints are consistent, the ME solution for this formulation is unique [10].

The bounds may be conveniently set on the basis of the estimated spectra of the noise in the transform domain. Given the transform of the data $\hat{d}_j$ and an estimate of the transform spectrum of the noise $\hat{\sigma}_j$, the bounds may be set as

$$\begin{aligned} l_j &= \hat{d}_j - h\hat{\sigma}_j \\ u_j &= \hat{d}_j + h\hat{\sigma}_j, \end{aligned} \tag{4}$$

where h defines the relative tightness of the constraints in terms of the noise spectrum. Thus the bounds would be of the order of h standard deviation units for the noise, which provides an effective way of controlling the tightness of the constraints.

A special-purpose algorithm was used to compute the ME solution to the bounds formulation; a description of this algorithm may be found in [2]. Briefly, the algorithm is based on the conjugate gradient method for unconstrained optimization, requiring only a few N-vectors of storage, and thus can accommodate the large number of pixels in real images. Provided that the variables which are constrained by bounds remain at their bounds, the problem may be regarded as an unconstrained optimization in the remaining components. During the iterations, the algorithm uses a composite function as an approximation to the entropy function to accommodate intermediate values of x_j which might be zero or negative. Additions and deletions to the active set of constraints are made dynamically until eventually the correct active set is determined and the solution is obtained, or the constraints are shown to be inconsistent. Experience with the algorithm shows that the transform computations (fast Fourier and fast Walsh transforms) take about half the time of the entire calculation.

In practice, the number of iterations required by the algorithm depends on the tightness of the constraints and on the choice of initial feasible solution. The tighter the constraints, the more iterations are required in general: with tight constraints the active set of constraints is larger and more likely to change during an iteration. When the active set is changed, the algorithm performs a steepest descent step before a conjugate gradient step is possible (in the next iteration). Thus with tight constraints, the conjugate gradient method may only come into play during the more advanced stages of processing, when the model approaches the final ME solution and there are fewer changes in the active set.

In seeking to reduce the number of iterations, it is advantageous to provide the algorithm with an initial solution which is as close to the final ME solution as possible. Such a solution which is compatible with the data constraints may be obtained by setting the transform of this initial feasible solution as close to the transform of the prior knowledge as permitted by the constraints. (We note that this solution will not necessarily satisfy the ME positivity constraints in the spatial domain, but the ME algorithm [2] was designed to accommodate intermediate non-positive values for the model.)

Several observations may be made about this choice of initial feasible solution. First, it is the solution to a least squares image processing problem incorporating prior knowledge, namely, to minimize

$$\sum_{j=1}^{N}(x_j - p_j)^2, \tag{5}$$

subject to the constraints (3). This may be shown as follows: since an orthogonal transform preserves L_2 norm, we must have

$$\sum_{j=1}^{N}(x_j - p_j)^2 = \sum_{j=1}^{N}(\hat{x}_j - \hat{p}_j)^2. \tag{6}$$

Consider the situation when $\hat{x}_j$ is set as close to $\hat{p}_j$ as permitted by the constraints. There are two possibilities: $\hat{x}_j = \hat{p}_j$ (constraint not active), and $\hat{x}_j = l_j$ or u_j (constraint active). Any change in this solution must either violate a constraint or increase the norm. Thus the solution must be a least squares solution. As discussed below, we have found that this easily computed initial solution is often an excellent approximation to the final ME solution and is a highly suitable choice for the initial estimate for the algorithm. Using this initial solution, we have observed that the difference between the initial and final entropy objective functions was usually very small (often less than 1 part in 1000), and that the initial and final images were often almost indistinguishable. It should be noted that this simple strategy is not restricted to positive pixel values, and may have more general applications than ME in image processing. It provides a fast means of signal restoration, with incorporation of prior knowledge, when the data d_j and model x_j may be non-positive.

RESULTS AND DISCUSSION

The ME method was examined for the application of signal restoration from noisy data. For the synthetic data examples, 1-D images containing various features were generated and Gaussian noise of specified standard deviation was added. The noisy image was transformed using Fourier or Walsh transforms (see [12] for description of the Walsh transform and a Fortran program for fast implementation), and the constraints (3) were set based on the transform $\hat{d}_j$ and the noise standard deviation $\hat{\sigma}_j$ as in (4). Results for two levels of constraint tightness h are presented, corresponding to 1.0 and 2.0 noise standard deviations. The prior knowledge in these examples was either uniform (all $p_j = 1.0$) or consisted of the original signal (all $p_j = x_j$). In some examples, both the initial feasible solutions and the final ME output of the examples are shown for comparison.

Fig. 1 shows examples of ME processing using the Fourier and Walsh transforms. For these tests, three features, all of maximum amplitude 1.0, were added to a background of amplitude 1.0; the features were a boxcar, a Gaussian, and a Gaussian-windowed cosine of frequency 0.39 times the Nyquist frequency. These features had widely different Fourier spectra, with the first being essentially broadband, and the second and third being band-limited Gaussians centered at frequencies zero and 0.39. The noise-free original image is shown in Fig. 1(a), and the noisy image, which contains Gaussian noise of standard deviation 0.2, in Fig. 1(b). The noisy image was processed by the ME algorithm, using one of two levels of tightness constraints ($h = 1.0$ and 2.0) and two types of prior knowledge (uniform and original signal). For the Fourier transform, both the initial feasible solutions (first column) and the final ME solutions (second column) are shown; only the final ME solutions are shown for the Walsh transform (third column).

Several characteristics of ME processing using the bounds constraints are apparent from Fig. 1. First, *the initial feasible solutions were almost indistinguishable from the final ME outputs.* This was observed in a wide variety of tests in addition to the examples shown here. In general, a substantial reduction in the noise levels was achieved by the processing, using either Fourier or Walsh transforms. As the constraints were relaxed, the output images progressively approached the prior knowledge. As expected, signal restoration was far better when appropriate prior knowledge (i.e. the original noise-free data) was provided than when the prior knowledge was uniform.

Both low and high frequency features were reasonably well preserved in the output of ME processing, since this processing depends on SNR in the transform domain and involves no *a priori* band-pass filtering. For the case of uniform prior knowledge,

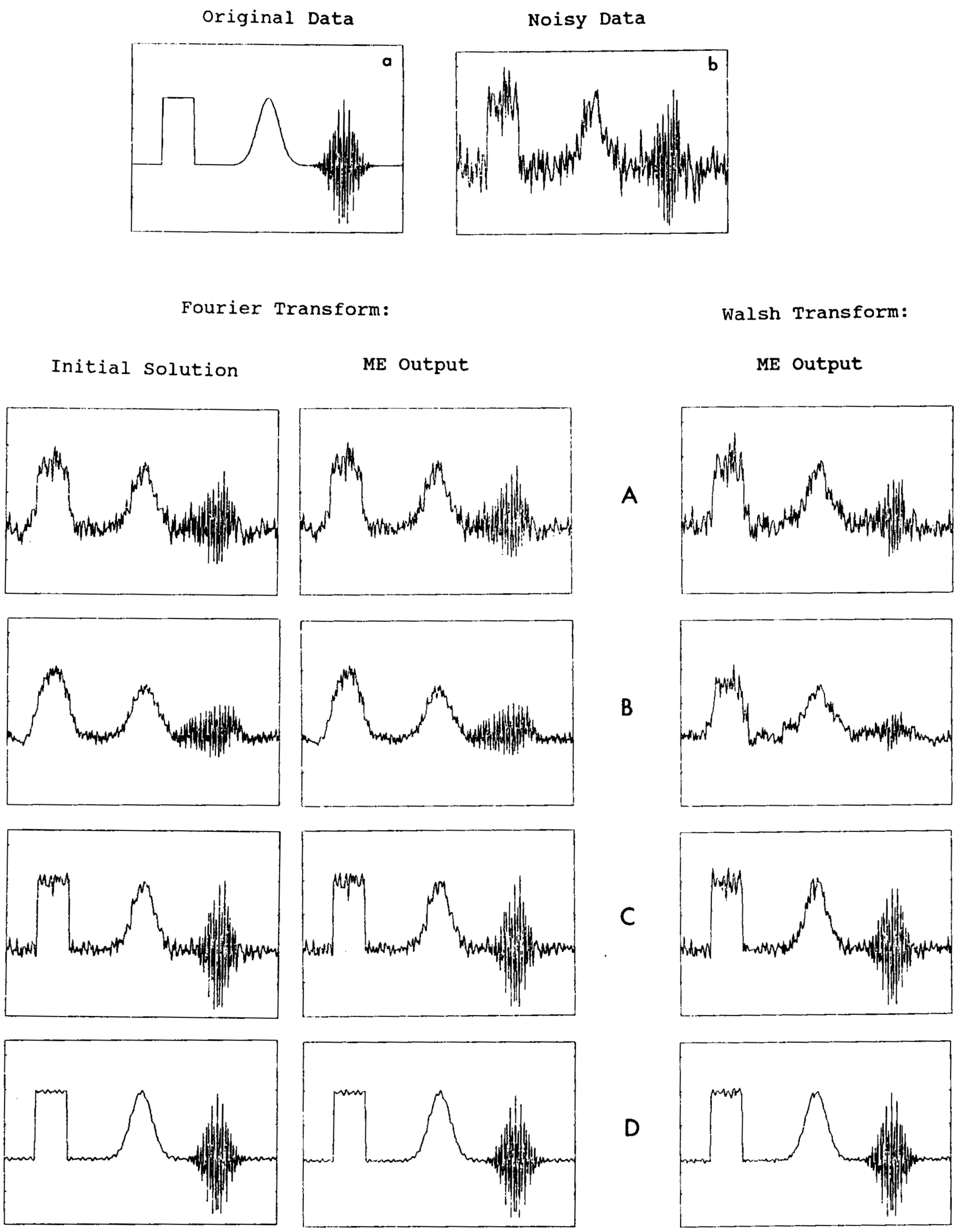

Figure 1. ME processing of a noisy data set with various features, using Fourier and Walsh transforms as indicated. Both the least squares initial feasible solution and the final ME output are shown for the results using the Fourier transform (two left columns); only the ME output is shown for the Walsh transform (right column). In the top row (a) shows the original noise-free data and (b) shows the noisy data. The results of processing are shown in rows A to D: (A) uniform prior knowledge, tightness constraint $h = 1$; (B) uniform prior knowledge, h=2; (C) original data prior knowledge, $h = 1$; (D) original data prior knowledge, $h = 2$.

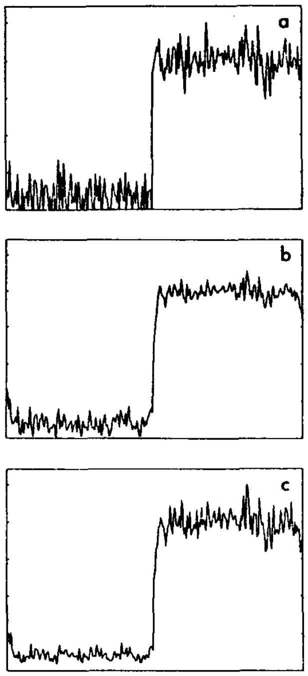

Figure 2.Comparison of initial and final stages of processing an image with large dynamic range, with constraint tightness parameter $h = 1$: (a) noisy data; (b) initial feasible solution; (c) final ME solution.

the recovery of the Gaussian-windowed cosine was better for the Fourier transform than for the Walsh transform. Conversely, the recovery of the boxcar was somewhat better for the Walsh transform. This may be understood in terms of the representation of the signal by the transform. For the ME bounds constraints processing, best results will be achieved when the SNR of the data in the transform domain is as high as possible, i.e. when the data are highly band-limited with respect to that transform. Since the Fourier transform of the Gaussian-windowed cosine is more band-limited than its Walsh transform, recovery is better when the Fourier transform is chosen. The Walsh transform usually provides a more band-limited representation of blocky features, and would be expected to provide better recovery of these features. In general, the more band-limited the signal in the transform domain, the better the signal recovery using ME bounds constraints processing.

In the examples above, the initial feasible solution was observed to be very similar to the final ME solution. This initial solution may be rapidly computed, in contrast to the final ME solution, which may involve large numbers of iterations, especially for tightly constrained problems. In view of its potential as an approximation of the true ME solution, experiments were performed which were directed at identifying situations where the initial and final ME solutions were significantly different. It was found that the similarity between the initial and final ME solutions depended on the dynamic range of the data: the smaller the dynamic range, the better the approximation of the initial solution. To illustrate such a difference, an example with large dynamic range is shown in Fig. 2. Here the original signal consisted of a step function of height 1.0 on a base of 0.1, to which noise of standard deviation 0.1 was added (yielding some negative data values). In the initial feasible solution (Fig 2(b)), the noise levels in the low and high parts of the function were comparable, but in the ME solution (Fig. 2(c)) the noise level was clearly less in the low region than in the high region. This amplitude-dependent processing is a familiar feature of ME [7].

CONCLUSIONS

We conclude that the bounds constraints formulation is a convenient way of taking advantage in ME processing of SNR enhancement in the transform domain, while permitting the incorporation of prior knowledge. In general, the more band-limited the signal in the transform domain, the more appropriate this bounds formulation will be. The algorithm [2] provided an effective means of computing the solution to the ME bounds problem. In the case of data which are not of high dynamic range, the final ME solution was found to be closely approximated by a simple initial least squares estimate obtained by setting the transform as close to the transform of the prior knowledge as permitted by the bounds constraints. No least squares computations are required to generate this initial solution.

References

[1] S.F. Gull and J. Skilling, "Maximum entropy method in image processing," IEE Proc., vol. 131, pp. 646-659, 1984.

[2] M.J.D. Powell, "An algorithm for maximizing entropy subject to simple bounds," Math. Prog., Series B, vol. 42, pp. 171-180 (1988).

[3] J.E. Shore and R.W Johnson, "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," IEEE Tr. Inform. Th., vol. IT-26, pp. 26-37, 1980.

[4] E.T. Jaynes, "On the rationale of maximum entropy methods," Proc. IEEE, vol. 70, pp. 939-952, 1982.

[5] J. Skilling, "The maximum entropy method," Nature, vol. 309, pp. 748-749, 1984.

[6] B.R. Frieden, "Image enhancement and restoration," from Picture Processing and Digital Filtering, (T.S. Huang, ed.), Topics in Applied Physics, vol. 6, (Springer-Verlag, New York), 1979.

[7] R. Narayan and R. Nityananda, "Maximum entropy image restoration in astronomy," Ann. Rev. Astron. Astrophys., vol. 24, pp. 127-170, 1986.

[8] C.R. Smith and W.T. Grandy, (eds.), "Maximum-entropy and Bayesian methods in inverse problems," Dordrecht, Holland: D. Reidel Publishing Company, 1985.

[9] S.F. Burch, S.F. Gull, and J. Skilling, "Image restoration by a powerful maximum entropy method," Comp. Vis. Gr. Image Proc., vol. 23, pp. 113-128, 1983.

[10] I. Barrodale, C.A. Zala, and R.F. MacKinnon, "Image processing by a maximum entropy procedure incorporating frequency domain bounds and prior knowledge," Proc. IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, pp. 163-169, 1987.

[11] R. F. MacKinnon, "Minimum cross-entropy noise reduction in images," J. Opt. Soc., 1989. (to appear)

[12] K.G. Beauchamp, "Walsh functions and their applications," Techniques of Modern Physics Series 3, Academic Press, London, 1975.