Observe that the first integral represents the contribution to the LF of the deterministic component and is classically implemented by means of a matched filter. The second integral represents the contribution due to the random fluctuations of the striated pattern and can be implemented by an estimator-correlator device. Obviously, when $K_{w_l}(x_1, x_2)$ are zero (i.e., the patterns are deterministic), the second term disappears, and the LF canonical form reduces to the one reported in [5]. On the other hand, when the cross-sections are zero-mean stationary signals, the LF does not depend on the offset, and only the pattern orientations can be estimated.

For stationary cross-section deviations $\Delta w_l^o(x) = w_l^o(x) - \overline{w}_l^o(x)$, the covariance functions $K_{w_l}(x_1, x_2)$ depend on the arguments difference only, and the optimal estimate $\Delta \tilde{g}_w^{opt}(s, \gamma_l)$ can be evaluated by means of the Wiener filter. In particular, let $S_{w_l}(\xi) = \mathcal{F}_1\{K_{w_l}(s)\}$ be the power density spectrum of $\Delta w_l^o(x)$; then, the power density spectrum $\tilde{S}_{w_l}(\xi)$ of $\Delta \tilde{g}_w(s, \gamma_l)$ is

$$\tilde{S}_{w_l}(\xi) = \frac{S_{w_l}(\xi)}{|\xi| S_n^p(\xi, \gamma_l)}. \tag{23}$$

Then, since the whitened observation noise has a unitary power density, the frequency response $H^{opt}(\xi, \gamma_l)$ of the Wiener filter is

$$H^{opt}(\xi; \gamma_l) = \frac{S_{w_l}(\xi)}{S_{w_l}(\xi) + |\xi| S_n^p(\xi, \gamma_l)} \tag{24}$$

The corresponding canonical form is illustrated in the scheme of Fig. 4.

Finally we observe that when $n(x, y)$ is white, (16) implies that the operator given by the filtered RT cascaded with the whitening filter reduces to the unitary operator $\tilde{\mathcal{R}}_{s,\theta} = H_s^{1/2} \mathcal{R}_{s,\theta}$, where $\mathcal{H}_s^{1/2}$ is a 1-D filter whose frequency response is $|\xi|^{1/2}$.

The LP canonical form of Fig. 4 has been applied to both simulated and field data. In the example of Figs. 5 and 6, the original synthetic image of two stationary, zero mean, Gaussian planar waves and its noisy version, respectively (SNR $= -10$ dB), are shown. The restored image reported in Fig. 7 has been obtained by thresholding the whitened projections and then applying the backprojection operator. It is evident from Fig. 7 that the residual error lies along the same directions of the detected waves.

As a real-word example we applied the LP to the imaging of moving vehicles with linear arrays for classification at highway gates. With reference to Fig. 8, the imaging system consists of a CCD vertical array performing an auto scan of the vehicles while in motion (see Fig. 9) and a CCD horizontal array providing a space-time image employed for motion estimation (see Fig. 10) [5]. The vertical scan geometry distortion produced by a variable speed is recovered by resampling this scan according to the velocity estimated by the LP (see Fig. 11). Concerning the LP performance, it is worthwhile to mention that the field results obtained with the real-time prototype implementing the processing scheme proposed in [5] for unknown cross-sections that closely approximate the LP at high signal to noise ratios evidenced that the reconstruction accuracy is adequate for reliable vehicle classification.

IV. Conclusions

In this correspondence, we investigated the structural properties of the LP for random straight patterns. In particular, we have shown that for any fixed value of their offsets and orientations, the RT provides sufficient statistics for both signal detection and parameter estimation. In the presence of Gaussian patterns, the RT is cascaded with a whitening filter and an estimator-correlator post processor, whereas for non-Gaussian patterns, the optimal scheme requires more complex statistics.

References

[1] S. R. Deans, "Hough transform from the Padon transform," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 3, no. 2, pp. 185–188, Mar. 1981.

[2] E. R. Davies, *Machine Vision.* San Diego, CA: Academic, 1990.

[3] J. Bernd, *Digital Image Processing: Concepts, Algorithms, and Scientific Applications.* Berlin, Germany: Springer-Verlag, 1991.

[4] B. D. Carlson, E. D. Evans, and S. L. Wilson, "Search radar detection and track with the Hough transform, Part I, Part II, and Part III," *IEEE Trans. Aerospace Electron. Syst.*, vol. 30, no. 1, pp. 102–125, Jan. 1994.

[5] A. Neri, G. Jacovitti, A. Minniti, and T. Tasselli, "Radon transform based velocity estimation of vehicles with linear arrays," in *Proc. SPIE, Advanced Signal Processing: Algorithms, Architectures, Implement. IV*, vol. 2027, pp. 39–50, 1993.

[6] B. Ristic and B. Boashash, "Kernel design for time-frequency signal analysis using the Radon transform," *IEEE Trans. Signal Processing*, vol. 41, no. 5, pp. 1996–2007, May 1993.

[7] H. K. Aghajan and T. Kailath, "Sensor array processing techniques for super resolution multi-line-fitting and straight edge detection," *IEEE Trans. Image Processing*, vol. 2, no. 4, pp. 454–465, Oct. 1993.

[8] A. K. Jain, *Fundamentals of Digital Image Processing.* London, U.K.: Prentice-Hall International, 1989.

[9] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part III.* New York: Wiley, 1971.

Block-Iterative Methods for Image Reconstruction from Projections

Charles L. Byrne

Abstract—**The simultaneous MART algorithm (SMART) and the expectation maximization method for likelihood maximization (EMML) are extended to block-iterative versions, BI-SMART and BI-EMML, that converge to a solution in the feasible case, for any choice of subsets. The BI-EMML reduces to the "ordered subset" EMML of Hudson *et al.* when their "subset balanced" property holds.**

I. Introduction

In its discrete form, the problem of reconstructing a nonnegative image from projections is that of finding a nonnegative J by 1 vector x satisfying the linear system $y = Px$, where the I by 1 vector $y > 0$ is the projection data and the I by J matrix P has nonnegative entries; we shall assume throughout that the columns of P sum to one. We are concerned here with iterative methods for finding x, particularly with the "expectation maximization" method for likelihood maximization (EMML) [1]–[5] and the "simultaneous MART" (SMART) [6], [21].

The "expectation maximization" approach to likelihood maximization (EMML), as it has been applied in emission tomography since

Manuscript received December 20, 1993; revised November 1, 1995. This work was supported by the National Cancer Institute under Grant CA42165 and Grant CA51071, with principal investigators M. King and W. Penney of the Department of Nuclear Medicine, University of Massachusetts Medical Center, Worcester, MA. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Nikolas P. Galatsanos.

The author is with the Department of Mathematics, University of Massachusetts at Lowell, Lowell, MA 01854 USA.

Publisher Item Identifier S 1057-7149(96)03168-5.

about 1980, leads to an iterative algorithm for finding nonnegative solutions of the linear system $y = Px$ [1]–[6], [14]. When there are no nonnegative solutions of $y = Px$, the algorithm produces a nonnegative minimizer of $KL(y, Px)$; here KL denotes the Kullback–Leibler distance between nonnegative vectors; that is, $KL(a, b) = \sum a_n \log(a_n/b_n) + b_n - a_n \geq 0$. The EMML is typically applied in the inconsistent case, in which, because of additive noise, there is no nonnegative solution of $y = Px$. Either $I \geq J$ or $I \leq J$, and both are usually large.

A second algorithm, discovered independently in 1972 by Schmidlin [10] and by Darroch and Ratcliff [11], called the SMART algorithm in [6], leads to a nonnegative minimizer of $KL(Px, y)$. The algorithms are similar in many respects, but the nonsymmetric nature of the KL distance leads to curious divergences in the theoretical development, with the result that certain questions that have been answered for SMART remain unsolved for EMML (see [15] and [23]). Darroch and Ratcliff designed their algorithm to select one probability vector out of the many satisfying a given set of linear constraints; for example, we might want a probability vector on a product space having prescribed marginals. SMART minimizes $KL(x, x^0)$ over all $x \geq 0$ with $y = Px$, in the consistent case, where x^0 is the starting vector. In the inconsistent case, in which there is no nonnegative x for which $y = Px$, SMART converges to the unique nonnegative minimizer of $KL(Px, y)$ for which $KL(x, x^0)$ is minimized [6], [21], [23].

In many applications, both I and J are large, and EMML and SMART can be slow to converge. Recent work by several authors [7], [8], [19], [20] suggests that acceleration of iterative algorithms can be achieved if one uses "block-iterative" methods [9], also called "ordered subset" methods [19], [20], [22], in which only subsets of the data are employed at each step. It is reasonable then to ask whether EMML and SMART have block-iterative versions. The SMART algorithm has a block-iterative version (BI-SMART), similar to, but somewhat simpler than, the block-iterative MART of Censor and Segman [9], that converges in the consistent case, for any partition of the data into subsets. Hudson *et al.* [19], [20] have shown that EMML possesses a block-iterative version ("ordered subset" EMML) for those matrices P that are "subset balanced;" that is, for which a partition of $\{1, 2, \cdots, I\}$ exists such that $\sum^n P_{i,j}$ depends only on j and not on n, for all subsets in the partition, where $\sum^n$ indicates summing over just those indices i in the nth subset, S^n, of the partition. The block-iterative EMML (BI-EMML) given here converges for all choices of subsets and reduces to the ordered subset method when subset balance holds.

The method of proof of convergence of EMML and SMART given in [6], [21], and [23] can be extended to prove convergence of their block-iterative versions; as a special case, we obtain a convergence theorem for MART that extends the result of Lent [12].

II. The Bi-SMART and Bi-EMML Algorithms

The EMML algorithm is usually presented according to the original development by Dempster *et al.* [13], involving conditional expectations. A more geometric approach to both EMML and SMART was given in [23]. The EMML algorithm begins with $x^0 > 0$ and takes as the successor of x^k:

$$x_j^{k+1} = x_j^k \sum_i P_{i,j}\left(\frac{y_i}{Px_i^k}\right) \tag{1}$$

for each $j = 1, \cdots, J$ and for $k = 0, 1, 2, \cdots$. The SMART algorithm defines the successor of x^m

$$x_j^{m+1} = x_j^m \exp\left[\sum_i P_{i,j} \log\left(\frac{y_i}{Px_i^m}\right)\right] \tag{2}$$

for each $j = 1, \cdots, J$ and for $m = 0, 1, 2, \cdots$.

We obtain a block-iterative version of the SMART algorithm (BI-SMART) by partitioning the index set $\{1, 2, \cdots, I\}$ into N disjoint subsets $\{S^1, S^2, \cdots, S^N\}$ and applying (2) to each subset in turn, without renormalizing the columns of P. Beginning with $z^0 > 0$, the BI-SMART iterative scheme is obtained as follows—for $m = 0, 1, 2, \cdots$ and $n = m(\bmod N) + 1$,

$$z_j^{m+1} = z_j^m \exp\left[m_n^{-1} \sum\nolimits^n P_{i,j} \log\left(\frac{y_i}{Pz_i^m}\right)\right], \tag{3}$$

for each $j = 1, \cdots, J$ and for $m = 0, 1, 2, \cdots$, where, again, $\sum^n$ denotes summation over those i in S^n, and $M_n = rmmax\{\sum^n P_{ij}\}$ over j.

We get the MART algorithm of Gordon *et al.* [16] by taking $N = I$ and each set S^n to be the singleton $\{n\}$. The MART iteration is then

$$\begin{aligned} z_j^{m+1} &= z_j^m \exp\left[P_{n,j} \log\left(\frac{y_n}{Pz_n^m}\right)\right] \\ &= z_j^m \left(\frac{y_n}{Pz_n^m}\right)^{P_{n,j}} \end{aligned} \tag{4}$$

for each $j = 1, \cdots, J$, for $m = 0, 1, 2, \cdots$, and $n = m(\bmod N) + 1$.

We can prove the following theorem [24].

Theorem: When there are nonnegative solutions of $y = Px$ the BI-SMART sequence $\{z^m\}$ converges to the unique solution for which $KL(x, z^0)$ is minimized. If z^0 is a constant vector, then $\{z^m\}$ converges to the solution maximizing the Shannon entropy.

Lent [12] has shown convergence to the maximum Shannon entropy solution for the MART case and for z^0 a constant vector. In [9], Censor and Segman generalize the MART algorithm to include block-iterative versions. Their algorithm is similar to the BI-SMART, but includes weighting terms instead of the normalization of the columns of P. In a more recent paper [17], Zenios and Censor study the behavior of the Censor–Segman algorithm, and in [18] Aharoni and Censor consider related block-iterative methods.

The formula (1) for EMML given above is valid so long as we have normalized P to have column sums equal to one. In the absence of such normalization, the EMML iteration becomes

$$x_j^{k+1} = x_j^k \frac{\left[\sum_i P_{i,j}\left(\frac{y_i}{Px_i^k}\right)\right]}{\left(\sum_i P_{i,j}\right)}. \tag{5}$$

The block-iterative version of the EMML (BI-EMML) is analogous to the BI-SMART, but uses arithmetic, rather than geometric, averaging; beginning with $z^0 > 0$, the BI-EMML iterative scheme is obtained as follows: for $m = 0, 1, 2, \cdots$ and $n = m(\bmod N) + 1$,

$$z_j^{m+1} = \left(1 - m_n^{-1} \sum\nolimits^n P_{i,j}\right) z_j^m + z_j^m m_n^{-1} \sum\nolimits^n P_{i,j}\left(\frac{y_i}{Pz_i^m}\right) \tag{6}$$

for each $j = 1, \cdots, J$ and for $m = 0, 1, 2, \cdots$, where, again, $\sum^n$ denotes summation over those i in S^n.

III. A Feedback Method for the Inconsistent Case

In typical cases we do not expect there to be a nonnegative vector x for which $y = Px$. In these inconsistent cases, BI-SMART does not converge to a single vector, but to a cycle of vectors, one for each of the N subsets; how distinct these vectors are is a function of the degree of inconsistency in the system $y = Px$, although we have not succeeded in quantifying this. This is at least what we have always observed; no one has yet succeeded in proving convergence to limit cycles for MART (or for BI-SMART, therefore) in the inconsistent case. The question then arises: What are we to do with these limit cycle vectors when what we want is a single approximate solution? One approach we are actively considering involves the use of the cycle vectors to create new "pseudodata," which is then fed back and the algorithm restarted.

To be precise, the feedback method works as follows: assuming convergence of BI-SMART to a limit cycle of N vectors, $z^1, z^2, \cdots, z^N$, we replace each old data value y_i with the new value $(Pz^n)_i$, where i is in the subset S^{n+1}, and we set $S^0 = S^{N+1}$. The new data vector is then used in exactly the same way the old y was, beginning with the same initial vector and performing BI-SMART until convergence to a second limit cycle. The infinite sequence of data vectors calculated in this way can be shown to converge to a vector y^∞, for which the system $y^\infty = Px$ is (nonnegatively) consistent. In simulations we have noted two interesting phenomena: first, this convergence is remarkably quick, with y^∞ often reached within two or three repetitions of the feedback; and second, the solution x of $y^\infty = Px$ is a strictly positive vector in most cases. This suggests that the repeated feedback is regularizing the problem; that is, effecting a filtering of the original data. The exact criterion implicitly used here is not yet understood. This feedback has been tried in connection with BI-EMML with some success [22].

ACKNOWLEDGMENT

The author wishes to thank Prof. Y. Censor of the University of Haifa for helpful discussions on these matters.

REFERENCES

[1] A. Rockmore and A. Macovski, "A maximum likelihood approach to emission image reconstruction from projections," *IEEE Trans. Nucl. Sci.*, vol. NS-23, pp. 1428–1432, 1976.

[2] L. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imag.*, vol. MI-1, no. 2, pp. 113–122, Oct. 1982.

[3] K. Lange and R. Carson, "EM reconstruction for emission and transmission tomography," *J. Comput. Assist. Tomog.*, vol. 8, pp. 306–312, 1984.

[4] Y. Vardi, L. Shepp, and L. Kaufman, "A statistical model for positron emission tomography," *J. Amer. Statist. Assoc.*, vol. 80, pp. 8–37, 1985.

[5] K. Lange, M. Bahn, and R. Little, "A theoretical study of some maximum likelihood algorithms for emission and transmission tomography," *IEEE Trans. Med. Imag.*, vol. MI-6, no. 2, pp. 106–114, June 1987.

[6] C. Byrne, "Iterative image reconstruction algorithms based on cross-entropy minimization," *IEEE Trans. Image Processing,* vol. 2, no. 1, pp. 96–103, Jan. 1993.

[7] S. Holte, P. Schmidlin, A. Linden, G. Rosenqvist, and L. Eriksson, "Iterative image reconstruction for positron emission tomography: a study of convergence and quantitation problems," *IEEE Trans. Nucl. Sci.*, vol. 37, no. 2, pp. 629–635, Apr. 1990.

[8] G. Herman and L. Meyer, "Algebraic reconstruction techniques can be made computationally efficient," *IEEE Trans. Med. Imag.*, vol. 12, no. 3, pp. 600–609, Sept. 1993.

[9] Y. Censor and J. Segman, "On block-iterative entropy maximization," *J. Inform. Optimizat. Sci.*, vol. 8, no. 3, pp. 275–291, 1987.

[10] P. Schmidlin, "Iterative separation of sections in tomographic scintigrams," *Nucl. Med.*, vol. 15, no. 1, 1972, Schatten Verlag, Stuttgart.

[11] J. N. Darroch and D. Ratcliff, "Generalized iterative scaling for log-linear models," *Annals Math. Statist.*, vol. 43, pp. 1470–1480, 1972.

[12] A. Lent, "A convergent algorithm for maximum entropy image restoration with a medical x-ray application," in *Image Analysis and Evaluation,* R. Shaw, Ed. Washington, DC: SPSE, 1977, pp. 238–243.

[13] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Statist. Soc.: Series B,* vol. 39, pp. 1–38, 1977.

[14] I. Csiszar and G. Tusnady, "Information geometry and alternating minimization procedures," *Statist. Decisions,* suppl. issue no. 1, pp. 205–237, 1984.

[15] I. Csiszar, "A geometric interpretation of Darroch and Ratcliff's generalized iterative scaling," *Annals Statist.*, vol. 17, no. 3, pp. 1409–1413, 1989.

[16] R. Gordon, R. Bender, and G. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography," *J. Theoret. Biol.*, vol. 29, pp. 471–481, 1970.

[17] S. Zenios and Y. Censor, "Parallel computing with block-iterative image reconstruction algorithms," *Appl. Num. Math.*, vol. 7, pp. 399–415, 1991.

[18] R. Aharoni and Y. Censor, "Block-iterative projection methods for parallel computation of solutions to convex feasibility problems," *Linear Algebra Applic.*, vol. 120, pp. 165–175, 1989.

[19] H. Hudson, B. Hutton, and R. Larkin, "Accelerated EM reconstruction using ordered subsets," *J. Nucl. Med.*, vol. 33, p. 960, 1992.

[20] H. Hudson and R. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imag.*, no. 4, pp. 601–609, Dec. 1994.

[21] C. Byrne, "Erratum and addendum to 'Iterative image reconstruction algorithms based on cross-entropy minimization'," *IEEE Trans. Image Processing,* vol. 4, no. 2, 1995.

[22] W. Xuan, "Evaluating a feedback technique in image reconstruction from projections," preprint, Dept. of Statistics, MacQuarie University, 1995.

[23] C. Byrne, "Iterative reconstruction algorithms based on cross-entropy minimization," in *Image Models (And Their Speech Model Cousins),* S. Levinson and L. Shepp, Eds. Institute for Mathematics and its Applications, vol. 80, Dec. 1995.

[24] ——, "Accelerating the multiplicative algebraic reconstruction technique (MART) and related block-iterative methods by equation rescaling," *IEEE Trans. Image Processing.,* to be published.

On the Metric Properties of Discrete Space-Filling Curves

C. Gotsman and M. Lindenbaum

***Abstract*—A space-filling curve is a linear traversal of a discrete finite multidimensional space. In order for this traversal to be useful in many applications, the curve should preserve "locality."**

We quantify "locality" and bound the locality of multidimensional space-filling curves. Classic Hilbert space-filling curves come close to achieving optimal locality.

I. INTRODUCTION

Denote $[N] = \{1, .., N\}$. A discrete m-dimensional space-filling curve of length N^m is a bijective mapping $C : [N^m] \longrightarrow [N]^m$ such that $d(C(i), C(i+1)) = 1$ for all $i \in [N^m - 1]$, where $d()$ is the Euclidean metric. In other words, the curve C of length N^m traverses all N^m points of the m-dimensional grid with side length N, making unit steps and turns only at right angles. For a historical account of classical space-filling curve constructions, see [8].

Space-filling curves are useful in applications where a traversal (scan) of a multidimensional grid is needed. Some algorithms perform

Manuscript received September 14, 1994; revised September 6, 1995. A preliminary version of this work was presented at the IEEE International Conference on Pattern Recognition, Jerusalem, 1994. The associate editor coordinating the review of this paper and approving it for publication was Prof. Charles A. Bouman.

The authors are with the Computer Science Department Technion—Israel Institute of Technology, Haifa, Israel.

Publisher Item Identifier S 1057-7149(96)03178-8.