[19] M. A. Perkowski, "A fundamental theorem for EXOR circuits," in *Proc. IFIP WG 10.5 Workshop on Applications of the Reed-Muller Expansion in Circuit Design*, U. Kebschull, E. Schubert, and W. Rosenstiel, Eds., Hamburg, Germany, Sept. 1993, pp. 52–60.

[20] S. Rahardja and B. J. Falkowski, "Classifications and graph based representations of switching functions using a novel complex spectral technique," *Int. J. Electron.*, vol. 83, no. 6, pp. 731–742, Dec. 1997.

[21] S. Rahardja and B. J. Falkowski, "Detection of Boolean symmetries using Complex Hadamard transformation," in *Proc. 30th IEEE Int. Symp. Circuits and Systems*, Hong Kong, June 1997, vol. 3, pp. 2116–2119.

[22] K. R. Rao and N. Ahmed, "Complex BIFORE transform," *Int. J. Syst. Sci.*, vol. 2, pp. 149–162, 1971.

[23] K. R. Rao, L. C. Mrig, and N. Ahmed, "A modified generalized discrete transform" *Proc. IEEE*, vol. 51, pp. 668–669, May 1973.

[24] K. R. Rao, K. Revuluri, M. A. Narasimhan, and N. Ahmed, "Complex Haar transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 102–104, Feb. 1976.

[25] J. A. Shah, O. Akiwumi-Assami, and B. Johnson, "A chip set for lossless image compression," *IEEE J. Solid-State Circuits*, vol. 26, pp. 237–244, Mar. 1991.

[26] R. S. Stankovic, M. R. Stojic, and M. S. Stankovic, "Recent developments in abstract harmonic analysis with applications in signal processing," in *Science*. Belgrade, Yugoslavia: Science Publisher, 1996.

[27] R. S. Stankovic and Q. S. Zhang, "Complex bridge functions," in *Proc. 5th Int. Workshop Spectral Techniques*, Beijing, China, Mar. 1994, pp. 142–153.

[28] Y. Tadokoro and T. Higuchi, "Discrete Fourier transform via Walsh transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 295–296, June 1979.

[29] C. C. Tsai and M. M. Sadowska, "Detecting symmetric variables in Boolean functions using generalized Reed-Muller forms," *IEEE Symp. Circuits and Systems*, London, U.K., May 1994, pp. 287–290.

[30] W. D. Wallis, A. P. Street, and J. S. Wallis, "Combinatorics: Room squares, sum-free sets, Hadamard matrices," *Lecture Notes in Mathematics*. Berlin: Springer-Verlag, 1972, no. 292.

[31] R. K. Yarlagadda and J. E. Hershey, *Hadamard Matrix Analysis and Synthesis with Applications to Communications and Signal/Image Processing*. Norwell, MA: Kluwer, 1997.

[32] L. P. Yaroslavsky, *Digital Picture Processing—An Introduction*. New York: Springer-Verlag, 1985.

[33] L. A. Zalmanzon, "Fourier, Walsh, and Haar transforms and their application in control, communication and other fields," in *Science*. Moscow, U.S.S.R.: Science Publisher, 1989.

[34] C. J. Zarowski, M. Yunik, and G. O. Martens, "DFT spectrum filtering," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 461–470, Apr. 1998.

Data Clustering Using Hierarchical Deterministic Annealing and Higher Order Statistics

A. N. Rajagopalan, Avinash Jain, and U. B. Desai

***Abstract*— In this brief, we propose an extension to the hierarchical deterministic annealing (HDA) algorithm for clustering by incorporating additional features into the algorithm. To decide a split in a cluster, the interdependency among all the clusters is taken into account by using the entire data distribution. A general distortion measure derived from the higher order statistics (HOS) of the data is used to analyze the phase transitions. Experimental results clearly demonstrate the improvement in the performance of the HDA algorithm when the interdependency among the clusters and the HOS of the data points are also utilized for the purpose of clustering.**

I. Introduction

The problem of data clustering is quite extensively encountered in image processing. Important applications include image segmentation [1], pattern recognition [2], and image compression using vector quantization [3]. Recently, Rose *et al.* proposed a novel clustering method, in which the annealing process with its phase transitions leads to a natural hierarchical clustering [4]–[8]. One does not need to know the total number of clusters in advance. Rather, one has a natural way of deciding on the final number of clusters. Unlike traditional clustering methods [9]–[15], which are basically descent algorithms, the hierarchical deterministic annealing (HDA) clustering algorithm is insensitive to the choice of the initial configuration.

In this brief, we extend the HDA method to include two important additional features.

1) In [6], the Hessian corresponding to a cluster is computed by considering only the points which belong to that cluster, and hence, a split in the cluster is governed by only those points. In our method, the intercluster dependencies are also accounted for by considering the *entire* data distribution (and not just individual clusters) to compute the Hessian. As will be shown, the utility of considering intercluster dependency becomes significant when data points belonging to different clusters overlap.
2) The distortion function for which we analyze the phase transitions is quite general. In most clustering algorithms, a square error or weighted-square-error distortion measure is usually used. In effect, this amounts to assuming the underlying distribution of the data to be Gaussian [2]. But in practical applications, the distribution could be arbitrary. We propose a general distortion measure based on the higher order statistics (HOS) of the data.

Positive definiteness of the Hessian is used as a criterion for splitting the clusters. This criterion was chosen over the perturbation variant of the HDA method [8], because knowing to predict the next critical temperature allows acceleration of the annealing process between transitions, while being more careful during the transition. It may also be mentioned here that the idea of using the entire data distribution appears in [8], in the context of rate-distortion theory. While we express the condition for bifurcation as a general condition on the

Manuscript received January 12, 1998; revised May 6, 1999. This paper was recommended by Associate Editor B. Linares-Barranco.

The authors are with the Department of Electrical Engineering, Indian Institute of Technology, Bombay, 400 076, India.

Publisher Item Identifier S 1057-7130(99)06541-6.

Hessian, the result in [8] solves for the explicit critical temperature assuming squared error distortion.

II. Intercluster Dependency and Phase Transitions

In accordance with [4]–[7], let y_j be the centroid of the cluster Y_j, and $d(x, y_j)$ be the distortion function for representing data vector x by y_j. Let $P(x \in Y_j)$ be the probability that x belongs to the cluster Y_j. Then, the expected distortion is given by $\langle D\rangle = \sum_x \sum_j P(x \in Y_j) d(x, y_j)$. As no *a priori* knowledge of the data distribution is assumed, of all possible distributions that yield a given value of $\langle D\rangle$, we choose the one that maximizes the entropy $\sum_x \sum_j P(x \in Y_j) \log P(x \in Y_j)$. It turns out [6] that the resultant distribution is the Gibbs distribution and is given by $P(x \in Y_j) = (1/(Z_x)) \exp(-\beta d(x, y_j))$, where $Z_x = \sum_k \exp(-\beta d(x, y_k))$ is the partition function and β is the Lagrange multiplier which is determined by the value of $\langle D\rangle$.

When the set of cluster representatives y_j is not fixed and it is desired to estimate the most probable set of cluster parameters, the effective cost to be minimized turns out to be the free energy (a well-known concept in statistical mechanics [4]), which is given by

$$F = -\frac{1}{\beta} \sum_x \log Z_x. \tag{1}$$

In physical analogy of annealing, β is inversely proportional to the temperature. The set of vectors $\{y_j\}$ that minimize F can be obtained by solving

$$\sum_x P(x \in Y_j) \frac{\partial d(x, y_j)}{\partial y_j} = 0.$$

By definition, the sub-Hessian is given by $H_{jk}[i, l] = \left(\partial^2 F\right)/\left(\partial y_j(i) \partial y_k(l)\right)$. Using the partial derivatives of F, we obtain

$$H_{jj} = \sum_x P(x \in Y_j)(\Phi_j - \beta C_{jj}) + \beta P^2(x \in Y_j) C_{jj}$$
$$H_{jk} = \beta \sum_x P(x \in Y_j) P(x \in Y_k) C_{jk}, \quad j \neq k.$$

Here, Φ_j and C_{jk} are square symmetric matrices whose (i, l)th entries are given by

$$\Phi_j[i, l] = \frac{\partial^2 d(x, y_j)}{\partial y_j(i) \partial y_j(l)}$$
$$C_{jk}[i, l] = \frac{\partial d(x, y_j)}{\partial y_j(i)} \frac{\partial d(x, y_k)}{\partial y_k(l)}.$$

Remark: It may be noted here that our expressions for $H_{\rm jj}$ and $H_{\rm jk}$ are slightly different from the ones given in [6]. This is because, unlike in [6] where only data points belonging to individual clusters were considered to compute the Hessian, we consider the *entire* data distribution at every phase transition.

Though each $H_{\rm jk}$ may not be positive semi-definite, the following lemma holds.

Lemma 1: The Hessian $\mathcal{H}$ of the complete data set is positive definite if the matrix $P_j = \sum_x P(x \in y_j)(\Phi_j - \beta C_{\rm jj})$ is positive definite $\forall j$.

Proof: Let m be the total number of natural clusters at a given instant, and n be the dimension of the data vector. Then, $\mathcal{H}$ is an $mn \times mn$ matrix constructed from $n \times n$ sub-matrices. We decompose the Hessian of the complete data set as $\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2$, where

$$\mathcal{H}_1[j, k] = H_{jk} = Q_{jk}$$
$$\mathcal{H}_2[j, k] = \begin{cases} P_j, & \text{if } j = k \\ 0, & \text{if } j \neq k. \end{cases}$$

A careful observation shows that each matrix Q_{jk} can be expressed as the sum $\beta \sum_x (P(x \in Y_j) P(x \in Y_k) u_j u_k^T)$, where the vector $u_j^T = [(\partial d(x, y_j))/(\partial y_j(1)) \cdots (\partial d(x, y_j))/(\partial y_j(n))]$. We now show that $\mathcal{H}_1$ is semi-positive definite. Let vector v be a concatenation of the n-tuples $a_1 \cdots a_m$. Then

$$\begin{aligned} v^T \mathcal{H}_1 v &= \left(a_1^T \cdots a_m^T\right) \mathcal{H}_1 \left(a_1^T \cdots a_m^T\right)^T \\ &= \beta \sum_x P(x \in Y_j) P(x \in Y_k) \left(\sum_{i=1}^m a_i^T u_i\right) \left(\sum_{i=1}^m u_i^T a_i\right) \\ &= \beta \sum_x P(x \in Y_j) P(x \in Y_k) t^2 \geq 0 \end{aligned}$$

where $t = \sum_{i=1}^m (a_i^T u_i)$ is a real number. Hence, H_1 is positive semi-definite. Now

$$v^T \mathcal{H} v = v^T \mathcal{H}_1 v + v^T \mathcal{H}_2 v.$$

Clearly, $\mathcal{H}$ is positive definite if P_j is positive definite $\forall j$. □

Thus, from Lemma 1, cluster Y_j splits when P_j ceases to be positive definite. The above is a generalization of [5]–[6], in the sense that one examines the Hessian derived using the entire data set (and not just the individual clusters) for deciding the splitting of clusters. This is the proposed modified condition for analyzing the phase transitions.

III. HOS for Clustering

In clustering schemes, a simple distortion measure such as the Mahalanobis or the squared error distance is usually used. This is a reasonable measure, as long as the underlying distribution of the data is Gaussian. However, for non-Gaussian distributions, a distortion measure based on just the first- and second-order statistics is unlikely to be satisfactory and it would be desirable to use a distortion measure that also takes the HOS of the data into account while clustering.

In practice, the probability density function $f(x \mid y_j)$ is seldom known. Nevertheless, it may be possible to estimate various joint moments of the random vector X from the given data. In [16], the density function is expressed as a series expansion in terms of the higher order moments by using the orthogonality relations between the Hermite polynomial and the Gaussian function. As explained in [16], if the mean of the random vector X is y_j and the covariance matrix is R, then one can express $f(x)$ as

$$f(x) = N(y_j, R) \left(\sum_{n=3}^{\infty} C_n \underline{H}_n \left(R^{-1/2}(x - y_j)\right)\right)$$

$$\text{where} \quad C_n = \int f(x) \underline{H}_n^T \left(R^{-1/2}(x - y_j)\right) \mathrm{d}x = E\left[\underline{H}_n^T \left(R^{-1/2}(X - y_j)\right)\right]. \tag{2}$$

Here, $N(y_j, R)$ is multivariate normal with mean y_j and covariance R. The vector $\underline{H}_n(\underline{x})$ is a vector whose elements are given by the product $(\prod_{i=1}^N (H_{k_i} x_i)/\sqrt{k_i!})$ for all permutations of k_i, $i = 1, \cdots, N$, such that $\sum_{i=1}^N k_i = n$. The term $H_{k_i}(x_i)$ is the Hermite polynomial of order k_i and is defined as $[(\partial^{k_i}/\partial t_i^{k_i}) \exp(t_i x_i - \frac{1}{2} t_i^2)]_{t_i = 0}$.

In pattern classification, a general decision criterion is defined in terms of the probability density functions as $-\log f(x)$ [2]. In light of the above expansion for $f(x)$, we define the HOS-based distortion measure as $d(x, y_j) = -\log f(x)$. Therefore

$$\begin{aligned} d(x, y_j) = {} & \frac{1}{2}(x - y_j)^T R^{-1}(x - y_j) \\ & - \log\left(1 + \sum_{n=3}^{\infty} E\left[\underline{H}_n^T \left(R^{-\frac{1}{2}}(X - y_j)\right)\right] \cdot \underline{H}_n(R^{-1/2}(x - y_j))\right). \end{aligned} \tag{3}$$

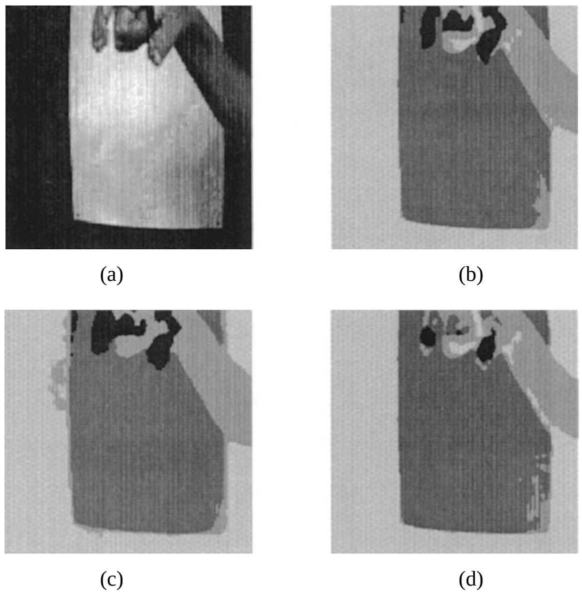


Fig. 1. (a) First frame of a motion image sequence. Segmentation results using the (b) Euclidean, (c) Mahalanobis, and (d) HOS-based clustering algorithms. Clearly, the color markers on the fingertips are best resolved by the HOS-based clustering algorithm.

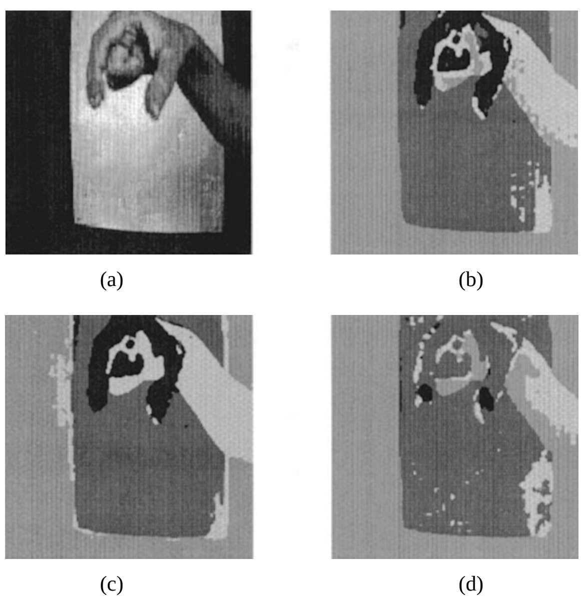


Fig. 2. (a) The 11th frame of the image sequence. Segmentation results using the (b) Euclidean, (c) Mahalanobis, and (d) HOS-based clustering algorithms. The proposed HOS-based algorithm correctly segments the color markers on the fingertips.

When $f(x \mid y_j)$ is Gaussian, (3) reduces to the Mahalanobis metric [16].

We propose to use the HOS-based distortion measure given in (3) for clustering in the HDA algorithm. It is expected to perform better than the traditional Euclidean or Mahalanobis distances, as it uses the HOS of the data for clustering purposes. Of course, this improvement is at the expense of increased computational complexity. However, it must be noted that no inversion is required in the proposed series expansion for the HOS (≥ 3). Hence, the increase in computational complexity is only moderate with the use of HOS. As a compromise between accuracy and computational complexity, we consider terms only up to $n = 3$ in (3) in our experimental studies. For a sound estimation of the statistics of the data, the data set must be sufficiently large.

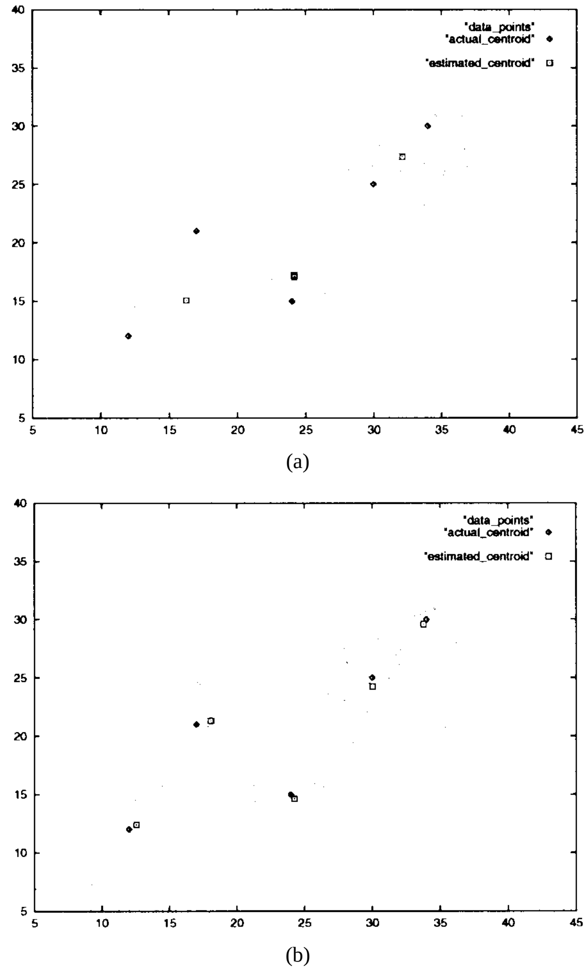


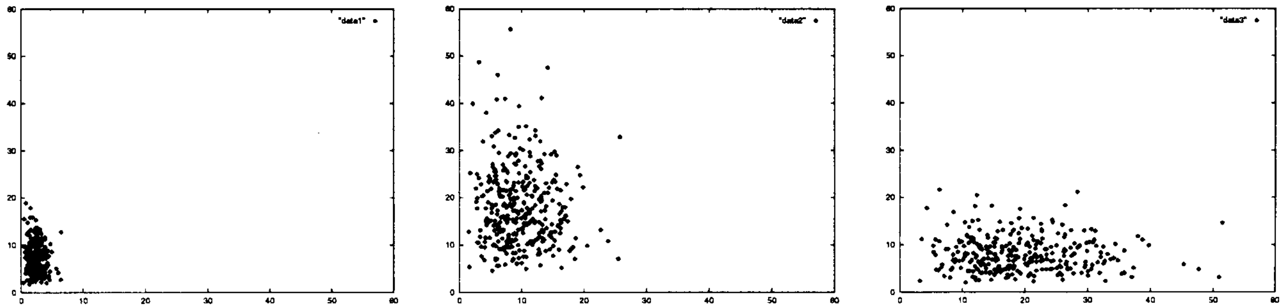
Fig. 3. The data points were generated from five Gaussian distributions. Results of the hierarchical DA clustering algorithm (a) without intercluster dependency and (b) with intercluster dependency.

TABLE I
NUMBER OF CORRECT AND FALSE MATCHES

	Cluster 1	Cluster 2	Cluster 3
Actual number of points in each cluster.	250	350	300
Square distortion: Points correctly classified.	246	224	215
Total number of points in each cluster.	416	240	244
HOS-based distortion: Points correctly classified.	240	238	283
Total number of points in each cluster.	277	258	365

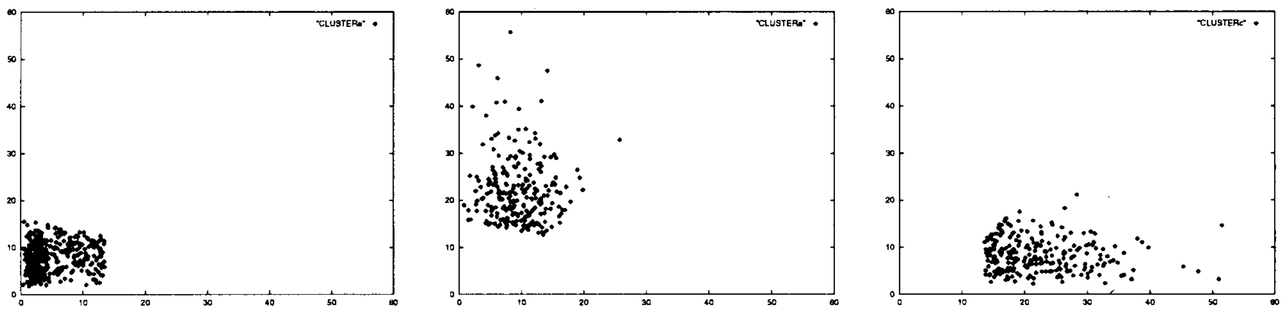
IV. EXPERIMENTAL RESULTS

We first demonstrate the effectiveness of HOS for clustering by using it for performing color segmentation of a real image sequence of a moving hand. Circular color markers were attached on different fingertips and the task was to correctly identify these red color markers on the fingertips. The Euclidean, the Mahalanobis, and the HOS-based distortion measures were each used independently in the k-means algorithm for classification. Figs. 1 and 2 show representative frames of the image sequence and the corresponding segmentation results. From the figures, it is quite clear that the proposed HOS-based clustering scheme has superior color segmenting capability. The Euclidean and Mahalanobis distortion measure based k-means algorithms tend to merge the segment of the color



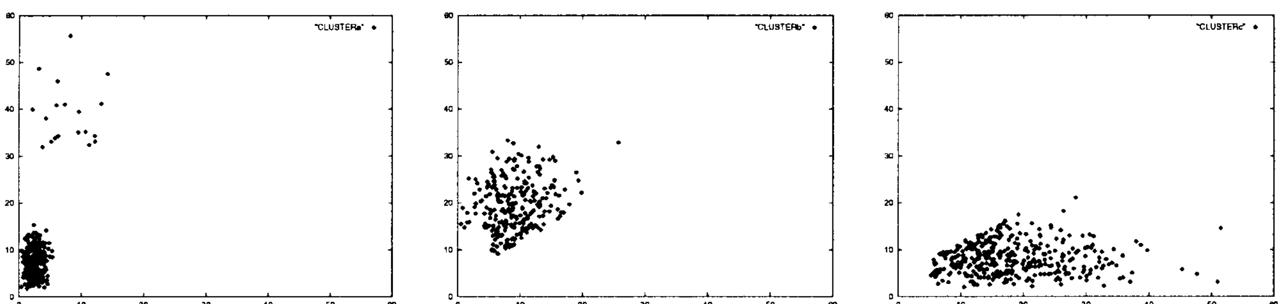
(a) (b) (c)

Fig. 4. Data points corresponding to three different clusters. (a) Cluster 1. (b) Cluster 2. (c) Cluster 3. The underlying distribution for the data points is the gamma distribution.



(a) (b) (c)

Fig. 5. Performance of HDA clustering algorithm using square distortion on the data points given in Fig. 4.



(a) (b) (c)

Fig. 6. Performance of HDA clustering algorithm using HOS-based distortion measure on the data points given in Fig. 4.

markers with that of the fingers. On the other hand, the HOS distortion measure resolves the color blobs on the fingertips quite successfully.

In the experiment on HDA, five clusters were generated, such that there was a significant overlap among the data points of these clusters [see Fig. 3(a)]. The HDA algorithm proposed in [4]–[7] was then used to perform clustering, and the results are shown in Fig. 3(a). The true centroid positions are indicated by solid diamonds, while the estimated values are denoted by open squares. We observe that the estimates of the centroids are not satisfactory. In fact, there are three centroids at almost the same point. The HDA algorithm was next used with intercluster dependency incorporated. The entire data set was used to compute the Hessian at every phase transition. The clustering results for this method are shown in Fig. 3(b). Clearly, the performance now improves significantly. All the centroids have been identified quite accurately. Thus, if the approximate critical temperature is used to accelerate the process of [4]–[7], then the performance would be inferior to that of the proposed method (and [8]). When the clusters are quite apart, then the improvement due to intercluster dependency is only marginal. In the above simulations, the squared error distortion measure (with mass constraints [7]) was used for both cases.

Finally, we demonstrate the utility of the HOS-based distortion measure for clustering in the HDA method. We generated three clusters each governed by the gamma distribution. Each individual cluster is shown in Fig. 4. A difficult clustering example was deliberately chosen. The DA clustering algorithm was used with: 1) the square distortion and 2) the HOS-based distortion measure. The clustering results for these cases are shown in Figs. 5 and 6. Each cluster has been shown separately for better visual depiction. The number of correct and false matches are given in Table I. We note that in the case of square distortion, many points have been wrongly classified. The HOS-based distortion case, on the other hand, has comparatively fewer false matches. Also, the number of correct matches are significantly more for the HOS case (76 more data points are correctly classified using HOS as compared to the square distortion measure). It must, however, be noted that the HOS-based HDA algorithm tends to misclassify some far away points in the first cluster.

V. Conclusion

In this brief, we have demonstrated the advantages of incorporating intercluster dependencies in the HDA clustering algorithm and equipping the algorithm with a general distortion measure that makes use of the HOS of the given data points. We show that, at the cost of moderate extra computation, it is possible to achieve good clustering performance even when the clusters overlap significantly.

Acknowledgment

The authors thank the anonymous reviewers for their insightful comments and suggestions.

References

[1] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
[2] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
[3] E. Yair, K. Zeger, and A. Gersho, "Competitive learning and soft competition for vector quantizer design," *IEEE Trans. Signal Processing*, vol. 40, pp. 294–309, Feb. 1992.
[4] K. Rose, E. Gurewitz, and G. C. Fox, "Statistical mechanics and phase transitions in clustering," *Phys. Rev. Lett.*, vol. 65, pp. 945–948, Aug. 1990.
[5] K. Rose, E. Gurewitz, and G. C. Fox, "A deterministic annealing approach to clustering," *Pattern Recognit. Lett.*, vol. 11, pp. 589–594, 1990.
[6] K. Rose, E. Gurewitz, and G. C. Fox, "Vector quantization by deterministic annealing," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1249–1257, July 1992.
[7] K. Rose, E. Gurewitz, and G. C. Fox, "Constrained clustering as an optimization problem," *IEEE Trans. Pattern Recognit. Mach. Intell.*, vol. 15, pp. 785–794, 1993.
[8] K. Rose, "A mapping approach to rate-distortion computation and analysis," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1939–1952, Nov. 1994.
[9] G. Ball and D. Hall, "A clustering technique for summarizing multivariate data," *Behav. Sci.*, vol. 12, pp. 153–155, 1967.
[10] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *J. Cybern.*, vol. 3, pp. 32–57, 1974.
[11] J. C. Bezdek, "A convergence theorem for the fuzzy ISODATA clustering algorithm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, pp. 1–8, 1980.
[12] S. Z. Selim and M. A. Ismail, "On the local optimality of the fuzzy isodata clustering algorithm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, pp. 284–288, Mar. 1986.
[13] J. C. Bezdek, R. J. Hathaway, M. J. Sabin, and W. T. Tucker, "Convergence theory for fuzzy c-means: Counterexamples and repairs," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-17, pp. 873–877, Oct. 1987.
[14] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, pp. 84–95, 1980.
[15] S. Z. Selim and M. A. Ismail, "K-means type algorithms: A generalized convergence theorem and characterization of local optimality," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, pp. 81–86, Jan. 1984.
[16] A. N. Rajagopalan, "HOS-based metric for face finding," SPANN Lab., Dept. Elect. Eng., IIT Bombay, India, Tech. Rep. SPANN.96.4, 1996.

A New Approach to Synthesize Sharp 2-D Half-Band Filters

Seo How Low and Yong Ching Li

***Abstract*—The frequency-response masking method is an efficient technique for the realization of sharp one-dimensional filters. Recently, this technique has been extended to the synthesis of sharp two-dimensional (2-D) filters. While it has been demonstrated that sharp diamond-shaped filters may be realized efficiently, the techniques previously presented cannot be applied to the synthesis of 2-D half-band filters. In this brief, we propose a modification of the technique for the synthesis of sharp 2-D half-band diamond-shaped filters. In addition, by exploiting the special properties of 2-D half-band filters, the complexities of the band-edge shaping and masking filters are further reduced. This results in a very efficient implementation of 2-D half-band diamond-shaped filters.**

I. Introduction

Two fundamental operations in one-dimensional (1-D) digital signal processing (DSP) are decimation and interpolation by a factor of two. These operations introduce undesirable artifacts into the spectrum of the signal under consideration. To suppress these artifacts, anti-aliasing and anti-imaging filters are required. Finite-impulse response (FIR) filters having the half-band characteristic are natural choices in these applications. An FIR half-band filter has the important property that half of its coefficient values are trivial. Consequently, using a half-band filter in filtering operations yields a significant advantage in terms of computational complexity.

The conversion between signals sampled using the rectangular sampling lattice and the quincunx sampling lattice is an example of sampling rate alteration by a factor of two in the two-dimensional (2-D) case. In such a situation, the associated anti-aliasing and anti-imaging filter is the 2-D diamond-shaped (DS) filter. As in the 1-D case, the 2-D half-band DS filter is particularly attractive for such an application, since it has the desirable property that half of its coefficient values are trivial [4]. For this reason, we find that this filter is used in many video and image processing applications. For example, it can be used in the conversion between progressive and interlaced scanning motion pictures [5]–[7].

The frequency-response masking (FRM) technique [8], [9] is an efficient method for realizing sharp 1-D filters. In [1]–[3], it has been demonstrated that sharp 2-D DS filters may also be synthesized using the FRM technique. However, the partitioning scheme previously presented does not allow the synthesis of 2-D half-band DS filters. Furthermore, the realization of half-band filters would have required additional constraints on the sub-filters. In this paper, we introduce a new approach for the synthesis of sharp 2-D half-band DS filters and discuss the additional constraints required to realize half-band filters. In addition, we further reduce the complexity of the band-edge shaping and masking filters used in the FRM technique by taking advantage of the properties of half-band filters. A very efficient implementation structure of 2-D half-band filters is obtained.

The organization of this paper is as follows. In Section II, a description of the polyphase decomposition of 2-D filters is given. This

Manuscript received May 8, 1998; revised April 16, 1999. This paper was recommended by Associate Editor M. Simaan.

S. H. Low is with the Infineon Technologies Asia Pacific Pte Ltd., Development Center, Singapore 349253.

Y. C. Lim is with the Department of Electrical Engineering, National University of Singapore, Singapore 119260.

Publisher Item Identifier S 1057-7130(99)06542-8.