Short Papers

Segmentation of Random Fields Via Borrowed Strength Density Estimation

Carey E. Priebe, David J. Marchette, and George W. Rogers

Abstract—In many applications, spatial observations must be segmented into homogeneous regions and the number, positions, and shapes of the regions are unknown a priori. Information about the underlying probability distributions for observations in the various regions can be useful in such a procedure, but these distributions are often unknown. Furthermore, while there may be a large number of observations overall, the anticipated regions of interest may be small with few observations from the individual regions. This paper presents a technique designed to address these difficulties. A simple segmentation procedure can be obtained as a clustering of the disjoint subregions obtained through an initial low-level partitioning procedure. Clustering of these subregions based upon a similarity matrix derived from estimates of their marginal probability density functions yields the resultant segmentation. It is shown that this segmentation is improved through the use of a "borrowed strength" density estimation procedure wherein potential similarities between the density functions for the subregions are exploited. The borrowed strength technique is described and the performance of segmentation based on these estimates is investigated through an example from statistical image analysis.

Index Terms—Mixture model, profile likelihood, image analysis, digital mammography.

1 Introduction

RANDOM field applications often require small-area estimates for some aspect of the local statistical structure. For instance, estimates of the local probability density of the observations in an image can be used to improve upon a preliminary segmentation. This is especially useful in cases for which little or no knowledge of the number, spatial structure, or statistics of the underlying image regions is assumed. Unfortunately, the requirement for local estimation implies that there are few observations available for these individual estimation problems and it is difficult to obtain sufficiently accurate estimates. This paper presents an approach using borrowed strength estimators which often can improve small-area estimation, and hence segmentation capabilities.

The purpose of this paper is not to develop a new segmentation scheme. Rather, we wish to show explicitly that the borrowed strength methodology presented in Section 2, which produces improved density estimates by taking advantage of potential similarities between the local densities, can result in improved segmentation. Thus we provide comparisons of a simple segmentation scheme with and without borrowed strength.

Let $\xi(x): R^0 \to \Re$ be a random field, with domain of definition $R^0 \subset \Re^d$. For many applications it suffices to assume that the sampling is performed on a regular lattice. Let $I^d = [0,1]^d$ be the unit cube in d-space and let $R^0 = L_N^d$ be the N-pitch lattice in I^d;

$$R^0 = L_N^d = \left\{ (l_1 / N, \ldots, l_d / N), l_i \in \{0, \ldots, N\} \right\} \tag{1}$$

N corresponds to the *resolution*. For instance, in digital image processing one may consider $\xi(x): L^2_{M_1 \times M_2} \to \Re$ where $L^2_{M_1 \times M_2}$ represents an $M_1 \times M_2$ lattice of pixel locations and the value of the field observations represents pixel intensity.

Assuming for simplicity that the image is made up of r disjoint regions $R^0 = \cup R^i (i = 1, \ldots, r)$ with associated probability density functions $\xi(x) \sim \alpha^i(\xi)$ for $x \in R^i$, then the goal is a segmentation of the image into disjoint regions each of which is homogeneous. For the purposes of this paper the simple definition of segmentation is a partition of R^0.

$$J = \left\{ J^1, \ldots, J^\rho \right\} \text{is a segmentation of } \xi(x) \Leftrightarrow J^i \cap J^j = \varnothing \text{ for } i \neq j \text{ and} \cup J^k = R^0 \tag{2}$$

Thus a segmentation is a clustering, and the goal is to determine which $x \in R^0$ are to be grouped together. Fig. 1 gives an example of this idea from digital mammography; the mammogram consists of (subjectively) $r = 5$ regions: healthy tissue, tumorous tissue, edge of breast, off breast, and the calcified artery in the lower right corner. This example will be used throughout to illustrate the assumptions and approach.

It is often the case that the original image pixel values are inappropriate for segmentation analysis. In mammography, much attention has been given to texture features. It is well-established that gray level alone is insufficient to characterize mammographic tissue, and that local texture is relevant to the analysis (see, for instance, [1], [2], [3]). For this paper we consider one of the simplest versions of a local texture, the coefficient of variation σ / μ. For each pixel x we calculate the mean μ and the variance σ^2 for the pixel values in a window of radius s centered at x. This yields a derived field whose observations represent a local roughness characterization of the original image, normalized for intensity level. Fig. 2 shows the local coefficient of variation field associated with the mammogram from Fig. 1 for $s = 3$.

There are a large number of simple, low-level segmentation algorithms available ([4], [5, chapter 3]). These algorithms provide an initial partitioning of an image. For simplicity and concreteness we will consider the watershed algorithm ([6], [7]) which automatically provides a segmentation according to definition (2) above. (A general feeling for the philosophy of watershed segmentation can be gained from the name itself, the idea being a partition of the image into local basins of attraction.) Fig. 3 depicts a watershed segmentation of the coefficient of variation field derived from the example mammogram.

Many low-level segmentation algorithms, including the watershed approach, give good localization of boundaries but generate extraneous regions. The goal of a second stage segmentation routine is to refine this initial segmentation. Our approach is to cluster the watershed regions, which will necessarily yield a segmenta-

- *C.E. Priebe is with the Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, MD 21218. E-mail: cep@jhu.edu.*
- *D.J. Marchette and G.W. Rogers are with the Naval Surface Warfare Center, Code B10, Dahlgren, VA 22448.*

Manuscript received 14 July 1995; revised 25 Nov 1996. Recommended for acceptance by K. Bowyer.
For information on obtaining reprints of this article, please send e-mail to: transpami@computer.org, and reference IEEECS Log Number P96125.

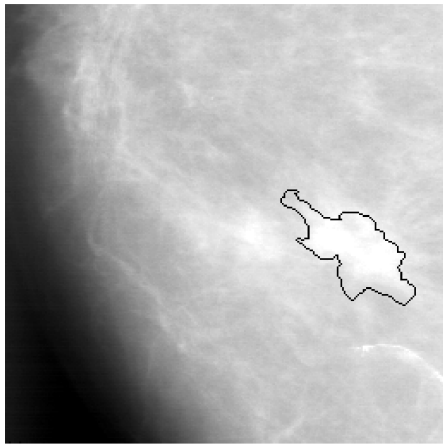

Fig. 1. Digitized mammogram (enhanced for display) and radiologist's boundary for biopsy-proven malignant tumor.

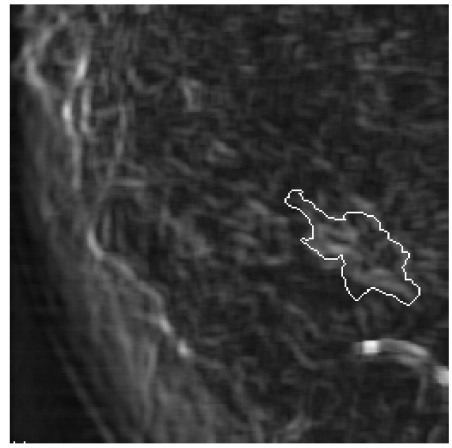

Fig. 2. Local coefficient of variation field and radiologist's boundary for the mammogram depicted in Fig. 1.

tion. (Other approaches to refinement, such as adjustment of region boundaries, are also possible.) The watershed segmentation depicted in Fig. 3 can be considered as a regional structure on R^0 which defines the local regions. Thus we have $R^0 = \cup \tilde{R}^i \left(i = 1, \ldots, \tilde{r}\right)$, where the $\tilde{R}^i$ are the watershed regions. For the example shown in Fig. 3, $\tilde{r} = 101$. Estimates of the probability density for these regions can be used to cluster the regions, yielding a final segmentation.

The segmentation refinement procedure requires a similarity matrix $T^{i,j} = \left\| \hat{\alpha}^i, \hat{\alpha}^j \right\|$ based on the local density estimates $\hat{\alpha}^i \left(i = 1, \ldots, \tilde{r}\right)$. One choice is the integrated squared error given by

$$T^{i,j} = ISE\left(\hat{\alpha}^i, \hat{\alpha}^j\right) = \int_{-\infty}^{\infty} \left(\hat{\alpha}^i(\xi) - \hat{\alpha}^j(\xi)\right)^2 d\xi .$$

In addition, it is often necessary to incorporate spatial proximity into the procedure. For context-free segmentation, this clustering is not based upon spatial dependencies among the regions. Such an approach is often unrealistic ([8, chapter 13], [9, chapter 7]), ignoring useful spatial information. Fig. 4 presents the irregular lattice derived from the watershed segmentation and used to represent spatial neighborhoods. The lattice is obtained by considering two sites, or watershed regions, to be neighbors if the Euclidean distance between their centers of mass is less than a prescribed constant c. This is analogous to the neighborhood structure imposed in [9] on North Carolina counties. Fig. 4 shows the neighborhoods associated with $c = 0.075$. As discussed in [9], the justification for $c = 0.075$ is based on a trial-and-error analysis of a lattice's ability to characterize the spatial dependence.

In this paper we consider the case in which the initial low-level segmentation and the spatial neighborhood structure are given and the goal is to refine the segmentation via region clustering. Furthermore, we consider the clustering algorithm to be given (see the Appendix). The only aspect of the procedure which is altered for comparison is the method of obtaining the estimates $\hat{\alpha}^i$, and it is shown that the second stage segmentation benefits from using borrowed strength.

2 Density Estimation

To improve segmentation of the image (Fig. 1) we suggest that a refinement of the initial segmentation produced by the watershed algorithm can be obtained by clustering the watershed regions (Fig. 3), and that this clustering be based on spatial proximity (Fig. 4) and local probability density estimates $\hat{\alpha}^i$ for the marginal densities of the coefficient of variation (Fig. 2) in the various regions $\tilde{R}^i$. The implicit hypothesis is that the $\hat{\alpha}^i$ are sufficiently different for different tissue types (healthy and tumorous) to aid in the clustering.

An initial investigation of this hypothesis, given in Fig. 5, is promising. Fig. 5 depicts kernel density estimates (see, for instance, [10]) for the coefficient of variation in the true regions R^H (healthy tissue) and R^T (tumorous tissue) in the example mammogram. For this image the numbers of observations per region are $n^H = 63{,}505$ and $n^T = 2{,}031$. This plot serves as first-order verification that the probability density functions for tumorous versus healthy tissue differ and estimates can be useful in clustering the watershed regions.

Unfortunately, the lack of knowledge of the location of differing regions (tumor and healthy) necessitates that the regions $\tilde{R}^i$ obtained in the initial segmentation (Fig. 3) be small compared to the anticipated size of the true but unknown R^i. These region sizes are too small to allow accurate nonparametric estimation of the densities α^i, especially in light of the even smaller effective number of observations due to dependence.

The competing requirements of estimation of an unknown density and local investigation to determine the segmentation regions lead to an impasse which cannot easily be overcome. However, in many applications it is reasonable to assume that the underlying local densities are (potentially complex) mixture models. Furthermore, it may be the case that the underlying mixture components can be considered to be invariant across class with the probability density functions differing only in their mixing coefficients.

In order to perform parametric estimation of the α^i, assumptions must be made as to the form of the densities. As a working assumption, consider the α^i to be finite mixture models (e.g., [11], [12]). For simplicity we assume mixtures of normals, but the treatment can be generalized to mixtures of any absolutely continuous exponential family density. Assume $\alpha^i(\xi) = \alpha\left(\xi; \psi^i, \lambda^i\right)$ is a mixture of m^i component normals. That is,

$$\alpha^i(\xi) = \sum_{t=1}^{m^i} \pi_t^i \varphi\left(\xi; \mu_t^i, v_t^i\right) \tag{3}$$

where $\psi^i = \left(\mu_1^i, v_1^i, \ldots, \mu_{m^i}^i, v_{m^i}^i\right)$ and $\lambda^i = \left(\pi_1^i, \ldots, \pi_{m^i}^i\right)$.

We let $\hat{\alpha}^i$ represent a maximum likelihood estimate of the marginal densities of the coefficient of variation for the n^i observations $\xi\left(x_1\right), \ldots, \xi\left(x_{n^i}\right)$, where $\left\{x_1, \ldots, x_{n^i}\right\} \subset \tilde{R}^i$. These estimates are termed "local" throughout because they are based only on the

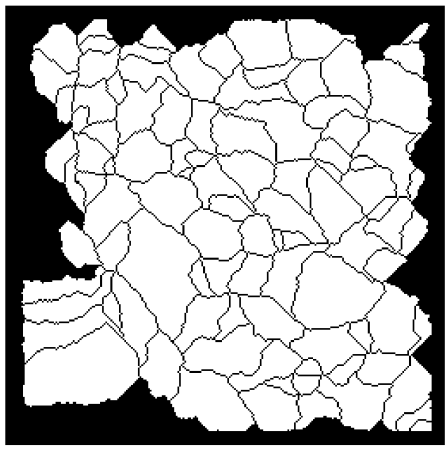

Fig. 3. Watershed algorithm yields an initial segmentation of the example mammogram into $\tilde{r} = 101$ local regions.

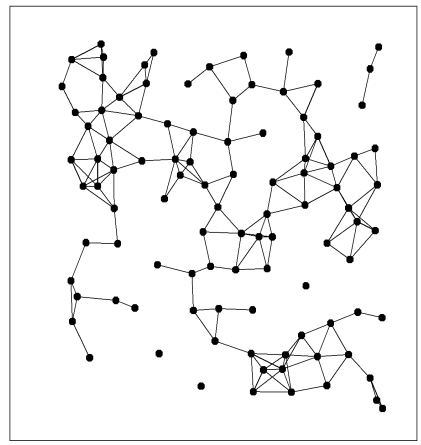

Fig. 4. Lattice derived from watershed regions ($c = 0.075$) depicted in Fig. 3 represents spatial neighborhood structure.

local sample from region $\tilde{R}^i$, unlike the borrowed strength estimates described below. Once the number of terms has been determined, the $\hat{\alpha}^i$ are conventional maximum likelihood estimates obtained via the EM algorithm (see, for instance, [11]). Given m^i, an estimate for $\left(\psi^i, \lambda^i\right)$ is obtained by maximizing the regional likelihood

$$L^i\left(\psi^i, \lambda^i\right) = \prod_{j=1}^{n^i} \alpha\left(\xi\left(x_j\right); \psi^i, \lambda^i\right) \tag{4}$$

It should be noted, however, that the determination of the number of terms in a mixture model, required for both the local and borrowed strength estimators considered herein, is no mean feat ([12, Section 5.2], [13]). The complexity of the density estimates used here is chosen using the adaptive mixtures procedure of [13]. The requirement in this work for a data-driven determination of the number of terms in the models makes it necessary to use a flexible mixture-based estimator such as adaptive mixtures. Priebe et al. [14] investigate semiparametric borrowed strength.

Consider the estimates $\hat{\alpha}^H$ and $\hat{\alpha}^T$ shown in Fig. 6. These estimates have complexity $m^H = 9$ and $m^T = 4$, and compare favorably with the kernel estimates. As one would expect, the moderate size of the tumorous sample translates to a less accurate estimate. Nevertheless, Fig. 6 suggests that, at least for this example, the mixture model assumption may indeed be reasonable, and normal mixture models can be used to estimate the densities.

The borrowed strength assumption is that the local estimates can be improved upon when there are fundamental similarities between the α^i and α^j, even when these densities differ. In particular, when the α^i are finite mixture models whose underlying mixture components ψ^i are invariant across the entire field domain R^0 in terms of their location in parameter space, then the local mixture probability density functions differ only in their mixing coefficients λ^i. An estimator which exploits this invariance by using all the observed data ("borrowing strength" from potentially dissimilar densities) to develop an estimate of the invariant parameters ψ^i and imposing this estimate as a constraint on the estimation of the λ^i (and hence the local probability densities α^i) can produce superior local estimates and hence superior segmentation. This procedure can be seen to be a profile likelihood technique [15].

The additional assumption is made that $m^i = m^0$ and

$$\left(\mu_1^i, \nu_1^i, \ldots, \mu_{m^i}^i, \nu_{m^i}^i\right) = \psi^i = \psi^0 = \left(\mu_1^0, \nu_1^0, \ldots, \mu_{m^0}^0, \nu_{m^0}^0\right)$$

for all i. Thus for normal mixtures

$$\alpha^i(\xi) = \sum_{t=1}^{m^0} \pi_t^i \varphi\left(\xi; \mu_t^0, \nu_t^0\right) \tag{5}$$

ψ^0 is common to all of the densities α^i and the difference between the densities is encompassed entirely in the mixing coefficients $\lambda^i = \left(\pi_1^i, \ldots, \pi_{m^0}^i\right)$. The borrowed strength estimate is obtained by obtaining a maximum likelihood estimate for ψ^0 based on all the data (the n^0 observations $\xi\left(x_1\right), \ldots, \xi\left(x_{n^0}\right)$, where $\left\{x_1, \ldots, x_{n^0}\right\} \subset R^0$). The profile likelihood estimate $\tilde{\lambda}^i$, given this estimate $\tilde{\psi}^0$ of ψ^0, is then obtained based on the local sample of n^i observations $\xi\left(x_1\right), \ldots, \xi\left(x_{n^i}\right)$. That is, we combine the joint likelihood estimate $\tilde{\psi}^0$ obtained by maximizing

$$L^0\left(\psi^0, \lambda^0\right) = \prod_{j=1}^{n^0} \alpha\left(\xi\left(x_j\right); \psi^0, \lambda^0\right) \tag{6}$$

with the regional profile likelihood estimate $\tilde{\lambda}^i$ obtained by maximizing

$$L^i\left(\lambda^i \middle| \tilde{\psi}^0\right) = \prod_{j=1}^{n^i} \alpha\left(\xi\left(x_j\right); \lambda^i \middle| \tilde{\psi}^0\right) \tag{7}$$

Estimating $\tilde{\alpha}^i(\xi) = \alpha\left(\xi; \tilde{\psi}^0, \tilde{\lambda}^i\right)$ using (6) and (7) is termed *borrowed strength maximum likelihood*. The improvement in these estimates that can be gained through the use of the borrowed strength methodology is potentially significant.

For model (5) we have consistency of the borrowed strength estimators and their superiority to the local approach for independent observations. These results follow in a straightforward manner from the standard maximum likelihood results in finite mixtures of exponential family densities as given in, for instance, [11]. A more detailed presentation of borrowed strength mixture models can be found in [16]. Under spatial dependence the above equations represent M-estimators (see, for example, [17, chapter 3]) and additional considerations are necessary to establish consistency.

Consider an eight-term approximation $\left(m^0 = 8\right)$ to

$$\alpha^0 = \left(n^H / n^0\right)\alpha^H + \left(n^T / n^0\right)\alpha^T$$

where $n^0 = n^H + n^T$. (As before, adaptive mixtures [13] is used to determine the number of terms in the model.) This approximation also yields the estimate $\tilde{\psi}^0$ for the means and variances of the borrowed strength estimates $\tilde{\alpha}^H$ and $\tilde{\alpha}^T$; these approximation are shown in Fig. 7 to compare favorably with the kernel estimates indicating that the borrowed strength assumption may indeed be warranted.

3 Example Results

When employing any statistical procedure it is necessary to validate the assumptions to the extent possible. For borrowed strength the key assumption is that the different local densities can be modelled as finite mixtures with the same means and variances. This assumption is investigated for digital mammography in [18]. In summary, the ISE between borrowed strength mixture estimates and their associated kernel estimates for healthy and tumorous tissue $(ISE(\alpha\ ,\ \alpha_{KE})$ and $ISE(\alpha\ ,\ \alpha_{KE}))$ is negligible compared to the *ISE* between estimates for different tissue types $(ISE(\alpha\ ,\ \alpha\ ))$.

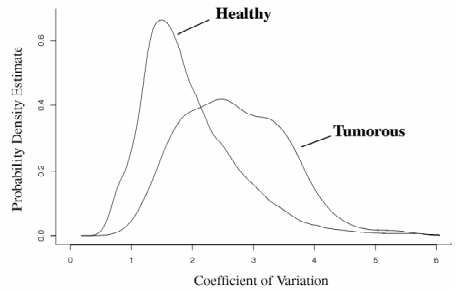

Fig. 5. Kernel density estimates for healthy versus tumorous tissue coefficient of variation.

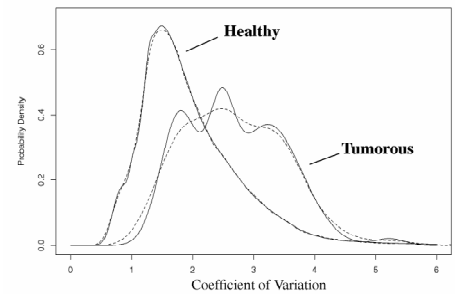

Fig. 6. Comparison of local mixture model estimates $\hat{\alpha}^H$ and $\hat{\alpha}^T$ with kernel estimates (dashed) for the probability density of healthy versus tumorous tissue coefficient of variation.

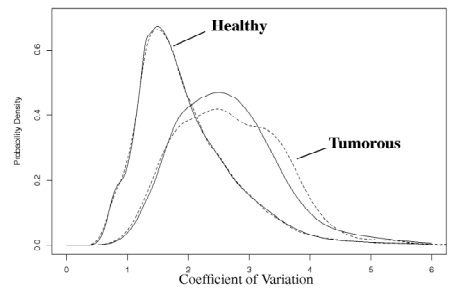

Fig. 7. Comparison of borrowed strength estimates $\hat{\alpha}^H$ and $\hat{\alpha}^T$ with kernel estimates (dashed) for the probability density of healthy vs. tumorous tissue coefficient of variation.

Nevertheless, it is difficult to conclude that the densities for the watershed subregions $\tilde{R}^i$ are being accurately modeled based on the estimates $\tilde{\alpha}^H$ and $\tilde{\alpha}^T$, and it is these subregion densities which are at the heart of the borrowed strength assumption. Figs. 6 and 7 could be misleading. An investigation of subregion density estimation performance is necessary, and will now be presented. Actual estimates for selected watershed regions in the example mammogram are investigated to indicate the improvement in local density estimates afforded by the borrowed strength methodology and how this improvement impacts the eventual agglomerative segmentation of the initial regions.

Four regions, two tumorous and two healthy, have been selected. For each of these four regions, both local and borrowed strength estimates have been obtained. The two tumorous estimates are more alike when using borrowed strength than when using local estimation. Similarly for the two healthy estimates. These within-class results, taken together with the between-class results, show that borrowed strength yields tumorous estimates more distinguishable from healthy estimates than does the local estimation procedure. These results are presented numerically, in terms of *ISE*, in Table 1. The tumorous regions are $T = \{63, 69\}$ and the healthy regions are $H = \{70, 73\}$. These results involve only the *ISE* distances and do not take into account spatial proximity. Nevertheless, in terms of the most simple-minded clustering, we can see that the borrowed strength estimates will allow the four regions to be separated into the correct two clusters, while the local estimates will not.

TABLE 1
ISE Results for Selected Regions From the Example Mammogram

	Local Estimates	Borrowed Strength Estimates
Within Tumorous Class	$ISE\left(\hat{\alpha}^{63},\ \hat{\alpha}^{69}\right) = 0.16$	$ISE\left(\tilde{\alpha}^{63},\ \tilde{\alpha}^{69}\right) = 0.04$
Within Healthy Class	$ISE\left(\hat{\alpha}^{70},\ \hat{\alpha}^{73}\right) = 0.12$	$ISE\left(\tilde{\alpha}^{70},\ \tilde{\alpha}^{73}\right) = 0.02$
Between Classes	$\min_{\substack{h\in H\\ t\in T}} ISE\left(\hat{\alpha}^{t},\ \hat{\alpha}^{h}\right) = 0.13$	$\min_{\substack{h\in H\\ t\in T}} ISE\left(\tilde{\alpha}^{t},\ \tilde{\alpha}^{h}\right) = 0.11$

Actual segmentation of the image depicted in Fig. 1 is quite successful. Figs. 8 and 9 give the results of applying the segmentation algorithm described previously to local and borrowed strength probability density estimates, respectively, for each watershed region. A synopsis of the approach employed includes:

1) Given an image I (Fig. 1), feature maps, consisting of scalar or vector-valued observations at each $x \in R^0$, are obtained. $F = f(I)$ used here is shown in Fig. 2, where f yields a local coefficient of variation.
2) $C = c(F)$ produces local regions via a low-level operation. The scheme employed here uses the watershed algorithm, with results shown in Fig. 3.
3) For each of these regions we produce density estimates $\hat{\alpha}^i$ obtained solely on the data in region $\tilde{R}^i$ and $\tilde{\alpha}^i$ using the borrowed strength approach.
4) Consideration of some distance (*ISE* is used here) between these estimates together with the spatial information provided by a lattice based on region location (Fig. 4) yields a distance $d^{i,j}$ between $\tilde{R}^i$ and $\tilde{R}^j$.

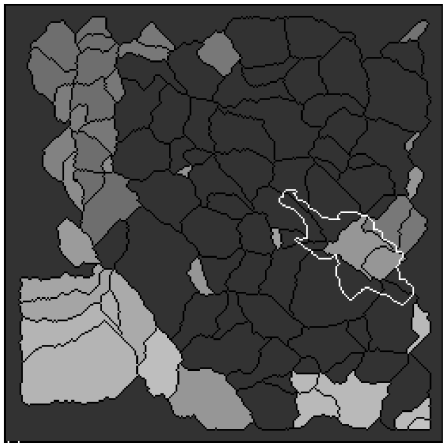

Fig. 8. Resultant segmentation without borrowed strength.

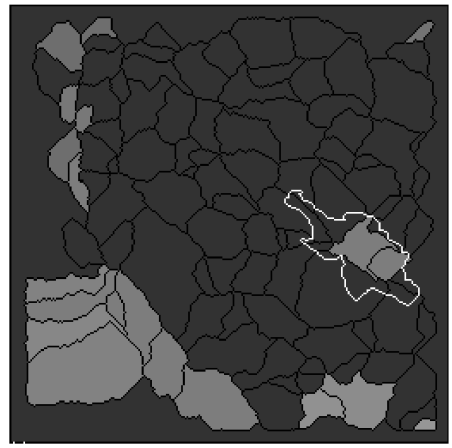

Fig. 9. Resultant segmentation using borrowed strength.

5) The symmetric $\tilde{r} \times \tilde{r}$ similarity matrix $D = \{d^{i,j}\}$ thus obtained, and a clustering parameter δ, are all that is required to produce a clustering of the original regions, yielding a final segmentation. The simple clustering algorithm used in this paper is given in the Appendix.

The segmentation algorithms have produced distinct clusters, or classes, with the two regions corresponding to the location of the tumor making up one of these. Thus the segmentor has successfully indicated the tumorous tissue as distinct from the rest of the image. The two watershed regions entirely within the tumorous region have been correctly segmented as distinct from the majority of the tissue. Separate segmentation clusters have been obtained for tissue on the edge of the breast. The lower right corner corresponds to the calcified artery. It is seen from comparing the figures that the borrowed strength procedure produces fewer false positives - regions of healthy tissue which are not clustered with the majority of the healthy tissue. Monte Carlo simulation results and quantitative results from a set of mammograms are presented in [19] indicating that this conclusion genralizes beyond this example image.

4 Conclusions

A borrowed strength approach to estimating local probability density functions in a random field has been presented as a technique to improve upon an initial low-level segmentation routine. Given an initial region map, probability density estimates, and spatial proximity information, the initial regions are clustered to produce a final segmentation.

This algorithm is seen to yield superior results when employing the borrowed strength density estimation technique, as compared to local estimation. The assumption is that some subset of the original segmentation boundaries are acceptable and we wish to delete those which are extraneous through clustering. Thus borrowing strength has the potential to improve present performance in segmentation applications for which no simple parametric assumptions can be made and there is a limited number of observations available for the required estimation of local densities. Using specific algorithms for the various components of the overall approach, we have presented a detailed example from digital mammography which indicates how and why the borrowed strength technique yields superior results

Issues which need to be addressed include the effect of within-region dependencies on the density estimation procedures, the incorporation of class allocation dependency assumptions, the choice of low-level segmentation algorithm, the choice of texture feature, and particulars concerning the clustering routine.

Appendix: Clustering Routine

Given a distance $\delta > 0$ and a similarity matrix $D = \{d^{i,j}\}$ for regions $\tilde{R}^i (i = 1, \ldots, \tilde{r})$, for $\tau \in \{1, \ldots, \tilde{r}\}$, let $T^\tau(\delta) = \{\tau' : d^{\tau,\tau'} < \delta\}$. Let T denote the set of ρ' *unique* $T^\tau(\delta)$; $T = \{T^1, \ldots, T^{\rho'}\}$. Note that for $\delta > 0$, $\tau \in \cup T^i \forall \tau$ and T^i, T^j are not necessarily disjoint.

Let $\hat{J}' \equiv \{\hat{J}'^1, \ldots, \hat{J}'^{\rho'}\}$ where

$$\hat{J}'^i \equiv \left\{ \tau \middle| i = \arg\max_{j \, s.t. \, \tau \in T^j} Card(T^j) \right\}$$

(with ties broken arbitrarily).

Note that $\hat{J}'^i$ might be null, $\tau \in \cup \hat{J}'^i \forall \tau$, and the $\hat{J}'^i$ are disjoint. Let ρ be the number of nonnull $\hat{J}'^i$ and $\hat{J} \equiv \{\hat{J}^1, \ldots, \hat{J}^\rho\}$ be the set of these "clusters." Then $\hat{J}$ is a segmentation of R^0.

Acknowledgments

This work is partially supported by Office of Naval Research Grant N00014-95-1-0777, Office of Naval Research Grant R&T 4424314, and the Naval Surface Warfare Center Independent Research Program. The authors wish to acknowledge an associate editor and three anonymous referees for many useful suggestions that greatly enhanced the quality of the paper.

References

[1] J.Y. Hsiao and A.A. Sawchuck, "Supervised Texture Image Segmentation Using Feature Smoothing and Probabilistic Relaxation Techniques," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 1,279-1,292, 1989.

[2] P. Miller and S. Astley, "Classification of Breast Tissue by Texture Analysis," *Image and Vision Computing*, vol. 10, pp. 277-282, 1992.

[3] C.E. Priebe, J.L. Solka, R.A. Lorey, G. Rogers, W. Poston, M. Kallergi, W. Qian, L.P. Clarke, and R.A. Clark, "The Application of Fractal Analysis to Mammographic Tissue Classification," *Cancer Letters*, vol. 77, pp. 183-189, 1994.

[4] R.M. Haralick and L.G. Shapiro, "Image Segmentation Techniques," *Computer Vision, Graphics and Image Processing*, vol. 29, pp. 100-132, 1985.

[5] R. Jain, R. Kasturi, and B.G. Schunck, *Machine Vision*. New York: McGraw-Hill, 1995.

[6] J. Serra, *Image Analysis and Mathematical Morphology*. London: Academic Press, 1982.

[7] L. Vincent and P. Soille, "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 583-598, 1991.

[8] G.J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*. New York: John Wiley, 1992.
[9] N.A.C. Cressie, *Statistics for Spatial Data*. New York: John Wiley, 1993.
[10] B.W. Silverman, *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall, 1986.
[11] R.A. Redner and H.F. Walker, "Mixture Densities, Maximum Likelihood and the EM Algorithm," *SIAM Rev.*, vol. 26, pp. 195-239, 1984.
[12] B.S. Everitt and D.J. Hand, *Finite Mixture Distributions*. London: Chapman and Hall, 1981.
[13] C.E. Priebe, "Adaptive Mixtures," *J. Am. Statistical Assoc.*, vol. 89, pp. 796-806, 1994.
[14] C.E. Priebe, D.J. Marchette, and G.W. Rogers, "Semiparametric Nonhomogeneity Analysis," *J. Statistical Planning and Inference*, vol. 59, no. 1, pp. 45-60, 1997.
[15] D.R. Cox and N. Reid, "Parameter Orthogonality and Approximate Conditional Inference (with discussion)," *J. Royal Statistical Soc.*, B, vol. 49, pp. 1-39, 1987.
[16] C.E. Priebe, "Nonhomogeneity Analysis Using Borrowed Strength," *J. Am. Statistical Assoc.*, vol. 91, pp. 1,497-1,503, 1996.
[17] P.J. Huber, *Robust Statistics*. New York: John Wiley, 1981.
[18] C.E. Priebe and D.J. Marchette, "Characterizing Mammographic Tissue via Borrowed Strength Density Estimation," *Digital Mammography 1996: Proc. Third Int'l Workshop Digital Mammography*, Chicago, June 1996, K. Doi, M.L. Giger, R.M. Nishikawa, and R.A. Schmidt, eds. Amsterdam: Elsevier, 1996.
[19] C.E. Priebe, D.J. Marchette, and G.W. Rogers, "Segmentation of Random Fields via Borrowed Strength Density Estimation," The Johns Hopkins Univ., Dept. of Mathematical Sciences Technical Report #547, 1996b.