Model-Based Methods for Textile Fault Detection

J. G. Campbell,[1] C. Fraley,[2,3] D. Stanford,[2] F. Murtagh,[4] A. E. Raftery[2]

[1] Faculty of Informatics, University of Ulster, Londonderry BT48 7JL, Northern Ireland

[2] Department of Statistics, Box 354322, University of Washington, Seattle, WA 98195-4322

[3] MathSoft Inc., 1700 Westlake Avenue N., Suite 500, Seattle, WA 98109

[4] School of Computer Science, The Queen's University of Belfast,
Belfast BT7 1NN, Northern Ireland

ABSTRACT: Addressing the problem of automatic fault detection in woven and dyed fabric, we discuss a number of new statistical model-based methods and relate them to a first stage of point/local detection and a second stage of extended pattern detection. One model-based method defines a maximum likelihood binarization of the image. In another model-based method, we describe a discrete Fourier transform-based texture analysis technique that is highly effective for woven textiles in discriminating subtle flaw patterns from the pronounced background of repetitive weaving pattern and random clutter. Finally, we describe a model-based clustering method that can be employed to aggregate perceptual groupings of point and local detections. © 1999 John Wiley & Sons, Inc. Int J Imaging Syst Technol, 10, 339–346, 1999

I. FLAW DETECTION IN TEXTILE FABRIC

Obstacles to machine inspection of textile fabrics include the difficulty of characterizing defects and the high data rate. In the manual inspection process, the flaws are marked using chalk or metallic tape, to be later handled via cutting and excising. In the context of automated manufacturing, there is clearly significant scope for introducing intelligence to these phases. If automatically detected flaw location can be supplied to an automatic cutter, then an optimal cutting plan may be followed, i.e., allowing flaws to be avoided with minimal waste. Moreover, manual inspection is known to be inefficient due to its repetitive and tedious nature (Newman and Jain, 1995; Sanders and McCormick, 1987).

Obvious flaws can be detected by sizable and abrupt tone deviations, perhaps using low-pass filtering to absorb small normal spots. In this article, we pay attention to more subtle flaws—those that present a minor local deviation but which, due to their occurrence in an extended spatial pattern, attract the attention of the human eye.

Flaws may fall into one of the following classes:

1. Local point flaws, like those mentioned above, are characterized by a severe tone change over only a few pixels or very few millimeters (e.g., holes). Because these are easy to detect by simple variations on thresholding, they are not further discussed here.
2. Medium scale flaws are characterized by change of *texture* over a number of millimeters.
3. Extended flaws (e.g., exhibiting a linear pattern extending over a number of centimeters). There is a space-intensity duality here. A very intense point flaw is easy to detect by the inspector (and the consumer). At the other extreme, a very extended pattern of very faint local flaws is quite perceptible by a consumer, but presents significant inspection challenge.

Many extended flaws are difficult to detect locally, due to their low intensity with respect to a strong background spatial pattern. Human perception can pick them up as an anomalous pattern and it is this ability that we attempt to replicate.

At a basic level (i.e., considering local and medium-scale flaws), we note that flaw detection differs from mainstream pattern classification in that the range of possible flaws is not easily characterized, but using a model, "normal" can be characterized and flaws then defined as significant deviation from "normal." In addition, considering extended patterns (clusters) of local and medium-scale flaws, these clusters *do* form identifiable shapes.

This leads to a general two-stage detection process (see Table I): first, detection of local and medium-scale flaws using an anomaly criterion, which for some fabrics may be a simple binarization; and second, a model-based clustering method.

Section II describes local anomaly detection through a novel binarization method. Section III proceeds to medium-scale flaw detection through texture analysis. Thus, Sections II and III address the first stage of our flaw detection program. Finally, Section IV considers detection of extended flaw patterns by addressing the second stage of our detection program.

The model-based clustering method described in Section IV is novel in considering the flaws in terms of a superimposed linear point pattern and a background uniform point pattern. The task is, therefore, a distribution mixture one. The linear point pattern is

Correspondence to: F. Murtagh

© 1999 John Wiley & Sons, Inc. CCC 0899–9457/99/040339-08

Table I. Preprocessing and extended pattern detection.

Stage 1: preprocessing
- P1. Binarization using a mixture model, suitable for knitted textiles.
- P2. Anomalous texture detection (texture binarization) using DFT features; suitable for woven textiles.

Stage 2: extended pattern detection
- E1. Model-based spatial-pattern clustering; suitable for arbitrarily shaped defects.
- E2. Linear pattern detection via Hough transform (Campbell et al., 1998); suitable for extended linear defects.
- E3. Point pattern fusion via morphological filtering (Campbell and Murtagh, 1998); suitable for medium-scale defects.

In stage 1, we select P1 or P2. In stage 2, we select E1 or E2 or E3. DFT, discrete Fourier transform.

modeled as a very linear, highly elliptical Gaussian shape. This is made operational, mathematically, by appropriately fine-tuning the variances and covariances of the Gaussian distribution. The goodness of a given parameterization of the distribution mixture at any time is judged on the basis of a Bayesian criterion, which allows one such parameterization (or model) to be assessed relative to another.

II. FAST IMAGE BINARIZATION

For this model-motivated approach, we take the image intensity distribution as the basis for creating a binary version of the image. We deal with a combined mixture density of two *univariate* Gaussian distributions $f_k(x;\ \theta) \sim \mathcal{N}(\mu_k, \sigma_k)$. The overall population has the mixture density

$$f(x;\ \theta) = \sum_{k=1}^{2} \pi_k f_k(x;\ \theta)$$

where the mixing or prior probabilities, π_k, sum to 1.

When the mixing proportions are assumed equal, the log-likelihood takes the form

$$l(\theta) = \sum_{i=1}^{n} \ln\left[\sum_{k=1}^{2} \frac{1}{2\pi\sqrt{|\sigma_k|}} \exp\left\{-\frac{1}{2\sigma_k}(x_i - \mu_k)^2\right\}\right]$$

The expectation-maximization (EM) algorithm is then used to iteratively solve this (see Celeux and Govaert, 1995, and Section IV below).

Figure 1 shows a selection of eight images of fabric with various thread or tone imperfections or faults. The corresponding binarized images are shown in Figure 2. The image in the lower row, second from the left, is now taken as a typical case. Figure 3, bottom, shows the horizontal and vertical marginal profiles based on the original image (Fig. 1). Little of significance can be noted. Figure 3 (top) shows the horizontal and vertical marginal profiles based on the binarized image (Fig. 2). The upper left profile in Figure 3 shows little of interest. The upper right profile, corresponding to the horizontal distribution in our fabric image, is clear-cut in its portrayal of the thread faults in this image. Typically up to 100 EM iterations are required, these being carried out on a fixed number of histogram bins.

The model-based binarization (adaptive thresholding) addresses well-known data quality problems—sensing problems such as temporal calibration drift, coupled with spatial variability of sensitivity and illumination. As such, it is necessary to estimate parameters only infrequently.

III. TEXTURE ANALYSIS USING A WINDOWED FOURIER TRANSFORM

Denim fabric yields a very sparse frequency domain, associated with the highly harmonic pattern due to weaving (see Figs. 5 and 8, discussed below). A 32 × 32 windowed two-dimensional DFT is used to provide a local feature extractor. This space-dependent Fourier transform bears some similarity to the short-time Fourier transform (STFT) (Oppenheim and Schafer, 1979). This method is particularly well suited to woven textile, because it allows the underlying woven pattern to be especially parsimoniously modeled (i.e., characterizable by very few features). By using the power

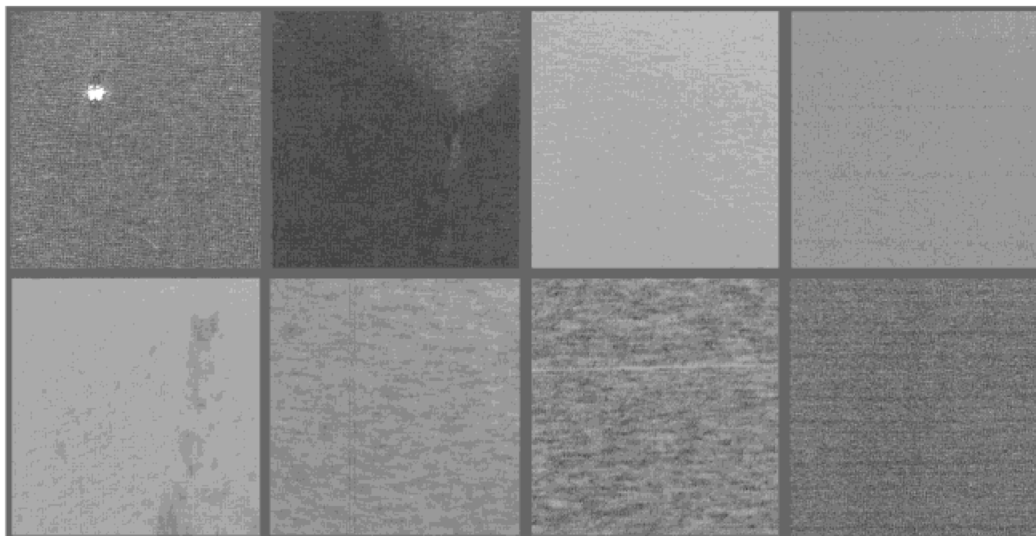

Figure 1. A selection of eight fabrics showing different types of thread and dyeing flaws.

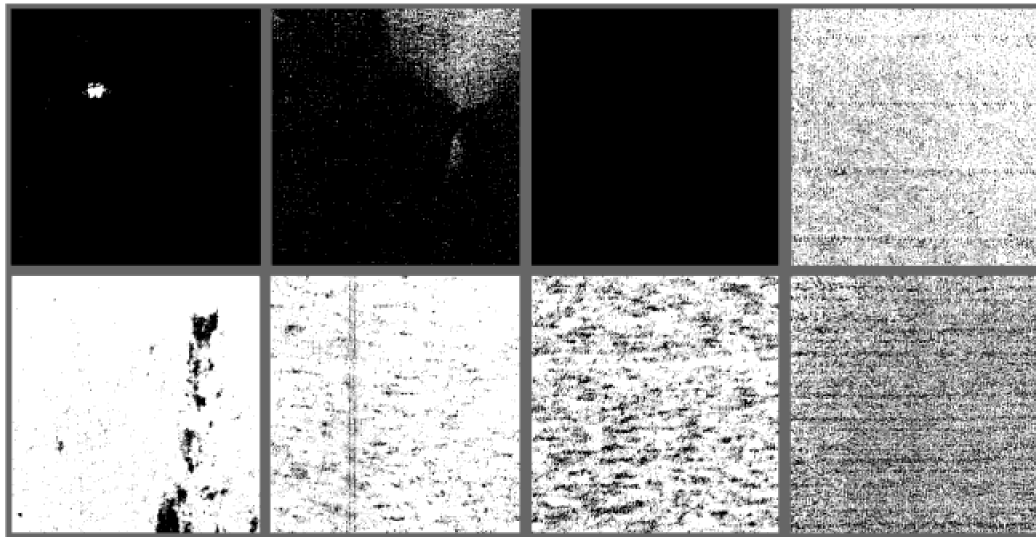

Figure 2. Binarized versions of the fabrics shown in the previous image. Method: distribution mixture based on intensity distributions.

spectrum component, we note that we assure shift invariance, i.e., features are invariant with respect to the phase of the moving window with respect to weaving pattern.

As an illustration, Figure 4 demonstrates the effectiveness of the DFT for fixed-pattern characterization. The fixed-pattern (middle) image results from only 0.5% of the Fourier components. The variance of the input image decreased nearly threefold, following removal of the fixed pattern (see image on right hand side in Fig. 4).

The effectiveness of this use of windowed DFT, as a local texture flaw detector, has been described by Campbell and Murtagh (1998). Assuming that the distribution of the texture feature (defined as the DFT of the 32×32 subimage) can be modeled as Gaussian, the Mahalanobis distance in feature space follows a χ^2 distribution. Thus, a threshold on flaw detection can be easily specified which is based on prescribed false-alarm rates. That is, this procedure is based on an ROC (receiver operating characteristic) analysis of the detection problem using a Neyman-Pearson criterion. Campbell and Murtagh (1998) check the validity of the false-alarm rates empirically and find considerable agreement with theory. Detailed discussion and formulae are given in Campbell and Murtagh (1998).

The DFT-Mahalanobis distance detector is especially amenable to real-time implementation. Its implementation in a (possibly analog) neural network structure is discussed in Campbell et al. (1995). Moreover, the components (DFT, quadratic form, threshold) are easily and efficiently implemented using DSP processor technology.

Recent work by Sari-Sarraf and Goddard (1998) has successfully applied a wavelet transform to similar (woven textile) data. We note the close analogy between the space-dependent Fourier transform used here, the Gabor transform, and the wavelet transform (Masters, 1993). Other recent work includes the application of Gaussian Markov random field modeling to detect flaws in woven textiles including denim (Cohen et al., 1991), the use of the wavelet transform (Jasper et al., 1996), and of neural networks (Hoffer et al., 1996).

IV. MODEL-BASED CLUSTERING

As indicated in Table I, stage 2 of our processing involves any of a selection of extended spatial pattern detection techniques. The application of the Hough transform for linear flaws is described in Campbell et al. (1998). The application of morphological filtering for false-alarm elimination in *medium-scale* defects is described in Campbell and Murtagh (1998).

In this section, we discuss the modeling of the (preprocessed) two-dimensional point pattern data as a (highly elliptical) Gaussian, subject to Poisson background clutter. To begin, we discuss the modeling in the general context of distribution mixtures.

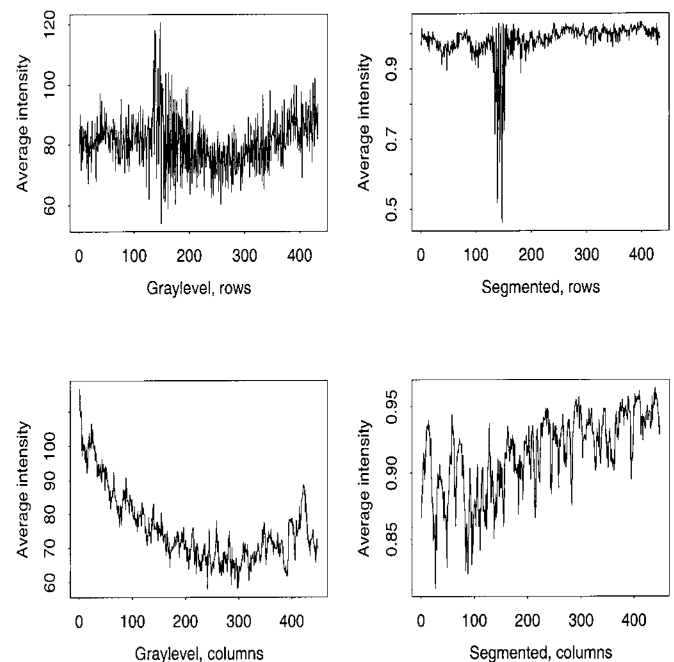

Figure 3. Marginal distributions for the image shown in the lower of the two rows, second from left, in the previous two images. Top: Marginals based on binarized image. Bottom: Marginals based on original image.

A. The Distribution Mixture Problem. Consider data that are generated by a mixture of $(G - 1)$ bivariate Gaussian densities, $f_k(x;\ \theta) \sim \mathcal{N}(\mu_k, \Sigma_k)$, for clusters $k = 2, \ldots, G$, and with Poisson background noise corresponding to $k = 1$. The overall population thus has the mixture density

$$f(x;\ \theta) = \sum_{k=1}^{G} \pi_k f_k(x;\ \theta)$$

where the mixing or prior probabilities, π_k, sum to 1, and $f_1(x;\ \theta) = \mathcal{A}^{-1}$, where $\mathcal{A}$ is the area of the data region. This is the basis for *model-based clustering* (Banfield and Raftery, 1993; Dasgupta and Raftery, 1998; Murtagh and Raftery, 1984; Banerjee and Rosenfeld, 1993).

The parameters, θ and π, can be estimated efficiently by maximizing the mixture likelihood

$$L(\theta,\ \pi) = \prod_{i=1}^{n} f(x_i;\ \theta),$$

with respect to θ and π, where x_i is the i-th observation.

In this article, we assume the presence of two clusters, one of which is Poisson noise, the other Gaussian. This yields the mixture likelihood

$$L(\theta,\ \pi) = \prod_{i=1}^{n}\left[\pi_1 \mathcal{A}^{-1} + \pi_2 \frac{1}{2\pi\sqrt{|\Sigma|}} \times \exp\left\{-\frac{1}{2}(x_i - \mu)^T \sum^{-1} (x_i - \mu)\right\}\right],$$

where $\pi_1 + \pi_2 = 1$.

B. Implementation: Iterated E- and the M-Steps. An iterative solution is provided by the EM algorithm of Dempster et al. (1977). Let the "complete" (or "clean" or "output") data be $y_i = (x_i, z_i)$ with indicator set $z_i = (z_{i1}, z_{i2})$ given by (1, 0) or (0, 1). Vector z_i has a multinomial distribution with parameters $(1;\ \pi_1, \pi_2)$. This leads to the *complete data log-likelihood:*

$$l(y, z;\ \theta,\ \pi) = \sum_{i=1}^{n} \sum_{k=1}^{2} z_{ik}[\log \pi_k + \log f_k(x_k;\ \theta)]$$

The E-step then computes $\hat{z}_{ik} = E(z_{ik} \mid x_1, \ldots, x_n, \theta)$, i.e., the posterior probability that x_i is in cluster k. The M-step involves maximization of the *expected complete data log-likelihood:*

$$l^*(y;\ \theta,\ \pi) = \sum_{i=1}^{n} \sum_{k=1}^{2} \hat{z}_{ik}[\log \pi_k + \log f_k(x_i;\ \theta)].$$

The E- and M-steps are iterated until convergence.

For the two-class case (Poisson noise and a Gaussian cluster), the complete data likelihood is

$$L(y, z;\ \theta,\ \pi) = \prod_{i=1}^{n}\left[\frac{\pi_1}{\mathcal{A}}\right]^{z_{i1}}\left[\frac{\pi_2}{2\pi\sqrt{|\Sigma|}} \times \exp\left\{-\frac{1}{2}(x_i - \mu)^T \sum^{-1} (x_i - \mu)\right\}\right]^{z_{i2}}$$

The corresponding expected log-likelihood is then used in the EM algorithm. This formulation of the problem generalizes to the case of G clusters, of arbitrary distributions and dimensions.

C. Assessment: The Bayes Information Criterion. In order to assess the evidence for the presence of a defect, we use the *Bayes factor* for the mixture model, M_2, which includes a Gaussian density as well as background noise, against the "null" model, M_1, which contains only background noise. The Bayes factor is the posterior odds for the mixture model against the pure noise model, when neither is favored *a priori.* It is defined as $B = p(x|M_2)/p(x|M_1)$, where $p(x|M_2)$ is the *integrated likelihood* of the mixture model M_2, obtained by integrating over the parameter space. For a general review of Bayes factors, their use in applied statistics, and how to approximate and compute them, see Kaas and Raftery (1995).

We approximate the Bayes factor using the Bayesian Information Criterion (BIC) (Schwarz, 1978). In the present context, this takes the form:

$$2 \log B \approx BIC = 2 \log L(\hat{\theta},\ \hat{\pi}) + 2n \log \mathcal{A} - 6 \log n,$$

where $\hat{\theta}$ and $\hat{\pi}$ are the maximum likelihood estimators of θ and π and $L(\hat{\theta},\ \hat{\pi})$ is the maximized mixture likelihood. Any value of BIC greater than zero corresponds to evidence for a defect. Conventionally, BIC values between 0 and 2 correspond to weak evidence, values between 2 and 6 correspond to positive evidence, values between 6 and 10 correspond to strong evidence, and values greater than 10 correspond to very strong evidence (Kaas and Raftery,

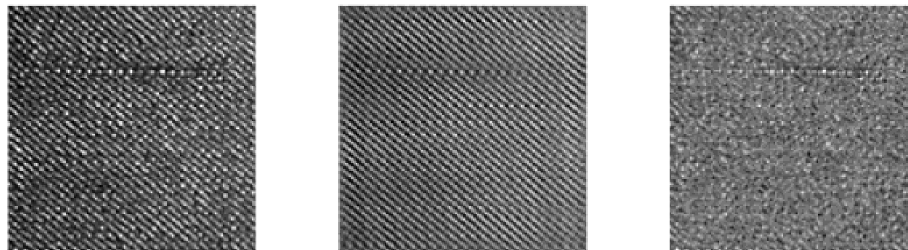

Figure 4. Analysis and removal of fixed pattern.

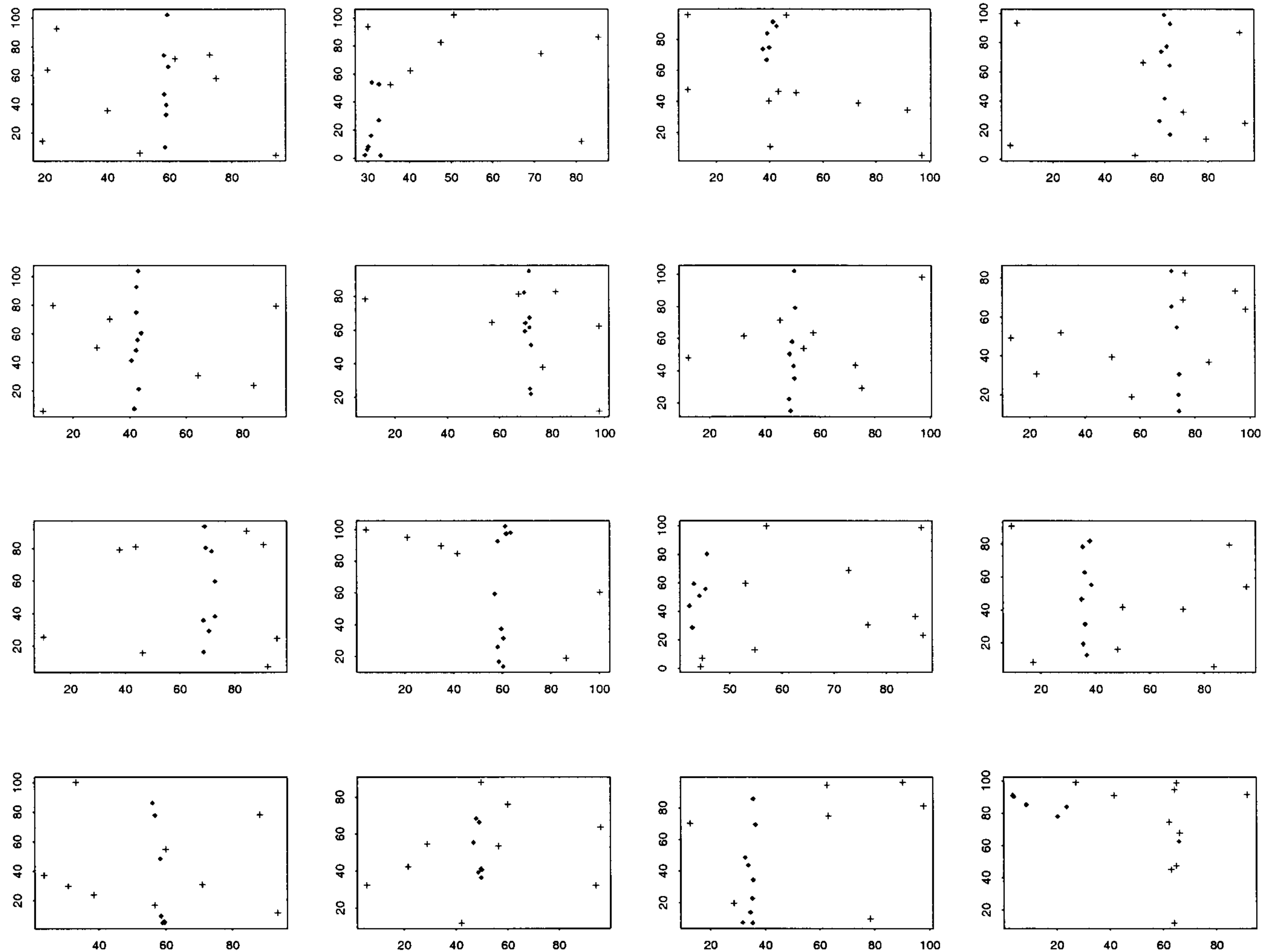

Figure 5. A sample set of 16 results of model-based clustering. Black diamonds show aligned points. Crosses show background noise. See text for discussion.

1995). The BIC is prone to false positives but performs very well compared to other testing criteria (Titterington et al., 1985; Leroux, 1992). This is backed up by experimental results (Dasgupta and Raftery, 1998).

D. Automated Linearity Finding and Testing. The method described so far does not incorporate any explicit mechanism for linearity or alignment seeking. When there is only one flaw in the image, corresponding to a single Gaussian cluster, this does not seem to matter. The unconstrained Gaussian density tends to adapt to what is in the image, finding a feature that is highly linear (i.e., long and thin) if it is present. However, if there are several flaws, perhaps intersecting one another, a more explicit incorporation of linearity might be advantageous. We now indicate briefly how this can be done.

In model-based clustering, the covariance matrix Σ associated with a cluster is parameterized (Banfield and Raftery, 1993) as $\Sigma = \lambda DAD^T$, where λ is the largest eigenvalue of Σ, D is the matrix of eigenvectors, and $A = diag\{1, \alpha\}$. Each of the three components of this decomposition of the covariance matrix corresponds to a geometric and visually intuitive property of the cluster that it describes. Thus, λ corresponds to the *volume* of the cluster, D to its *orientation,* and A (or equivalently here, α) to its *shape.* The value α is the ratio of second to first eigenvalues. For α close to 1, clusters will be spherical; while for values approaching 0, the clusters will be very linear (i.e., their members will be highly aligned).

The user (or program, e.g., using Bayes factors as described below) can set or determine values of λ to control the cluster volume, D to control orientation, and A to control shape. By constraining some or all of λ, D, and A to be equal across clusters, the finding of clusters of different types can be prioritized. In this article, we are interested only in letting the data determine the best value for A, which amounts to determining the best value for α. Murtagh and Raftery (1984) assumed user specification of α. A maximum likelihood estimate of the clusters (using EM) may additionally be used automatically to determine an optimal value of α in the following way.

Take a set of n points comprising a Gaussian cluster (n_1 points), with spatially homogeneous Poisson background (n_0 points), and let the sample covariance matrix for the cluster have spectral, or singular value, decomposition $\hat{\Sigma} = L\Omega L^T$. The maximized classification log-likelihood of the data, with α assumed known, is

$$2l = -(n - n_0)(2 \log(2\pi) + 2(1 - \log 2)) + \log(|A|) - 2n_1 \log(\mathrm{tr}(\Omega_k A^{-1})/n_1) - 2n_0 \log(\mathcal{A}).$$

The "profile likelihood" with respect to α is then maximized. This results in a likelihood equation that yields an estimate of α as the ratio of eigenvalues of the estimated sample crossproduct matrix for the Gaussian component (Dasgupta and Raftery, 1998). This reinforces the approach of Murtagh and Raftery (1984) by casting this problem in a likelihood framework.

To summarize, we seek a highly elliptical Gaussian cluster superimposed on a homogeneous Poisson background. Furthermore, we use the BIC quality criterion for the fit of this two-cluster mixture model to the data.

E. Linear Pattern Detection through Model-Based Clustering. Figure 5 shows a range of point patterns randomly generated in line with the point patterns representing textile flaws, which will be looked at in the next subsection. In each case, in a region of dimensions [0, 100] × [0, 100], we generated (1) eight linear points, approximately vertical, and (2) an additional set of eight background uniformly distributed points. The linear points were all generated at a horizontal pixel that was uniformly distributed in [25, 75]; each point then had a uniformly distributed value in [0, 5] added. The vertical component of these points was uniformly distributed in [1, 100] with a uniform value in [0, 5] added. The background points were each uniformly distributed in [0, 100] with a uniform perturbation from [0, 5].

Figure 5 shows the result of model-based clustering, where the points represented as black diamonds are members of the linear cluster and the crosses are points not assigned to the linear class. The corresponding BIC values are revealing. From upper left, they are 2.4, −5.1, 1.0, −6.2 (comprising row 1); 3.2, 7.1, 3.0, −4.3 (comprising row 2); −9.4, −5.3, −8.5, −3.7 (comprising row 3); and 0.7, 3.3, −2.0, −15.1 (row 4). A few comments on these follow.

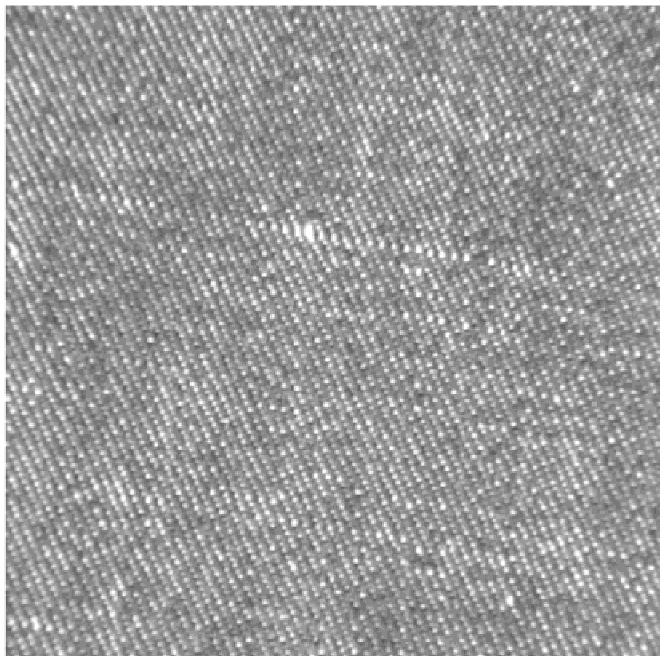

Figure 6. Image used for experimentation referred to as d10.

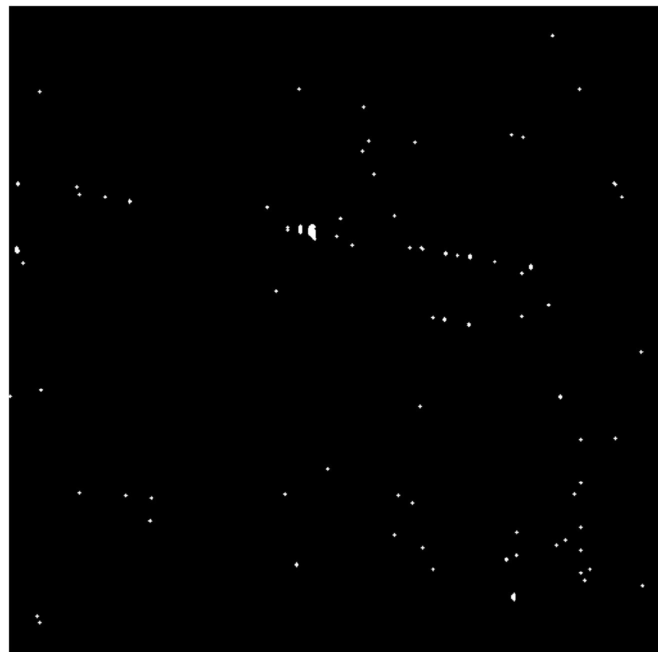

Figure 7. Image d10 following thresholding and a morphological opening.

In Figure 5, row 1, result 1 is excellent; result 2 is mediocre and is indicated as such by the low BIC value; result 3 is acceptable, even if distant aligned points are lost; and result 4 is good, with the low BIC value indicating some lack of tight linearity. In row 2, all results are very good. In row 3, results 1, 2, and 4 are very good and result 3 is acceptable. In row 4, results 1, 2, and 3 are reasonable to very good. Result 4 shows one very poor result, which is clearly indicated by the very low BIC value.

We see that excellent results can be obtained and that the BIC value can valuably show up problematical results. Results on these small point pattern sets, including plotting, were practically instantaneous.

F. High-Level Analysis. For model-based clustering, image preprocessing consists of cropping image boundary; simple thresholding using a three-sigma detection limit; a morphological opening with a 3 × 3 cross element; labeling contiguous regions; rejecting regions of low cardinality; and replacing retained regions with their centroids. The morphological operation is being used here as a decision filter; intuitively it can be seen to eliminate a particular class of false alarms. Figure 6 shows a sample denim fabric image, Figure 7 shows the thresholded and opened image, and Figure 8 (left) shows the point pattern retained for analysis. Figures 9 and 10 show another example. To explain the automatic detector of aligned point patterns, we apply a distribution mixture methodology in an unusual way. This methodology is both effective and computationally highly efficient.

Figures 8 and 10 (left) show the random starting configurations and the right panels show the final, optimized configurations. The BIC is highly indicative of the best solution, indicating strong evidence for the presence of a defect. The clustering method can produce different results for different starting values, although the results are highly consistent and the BIC value indicates when the solution is not a good one. The ability of the method to detect flaws

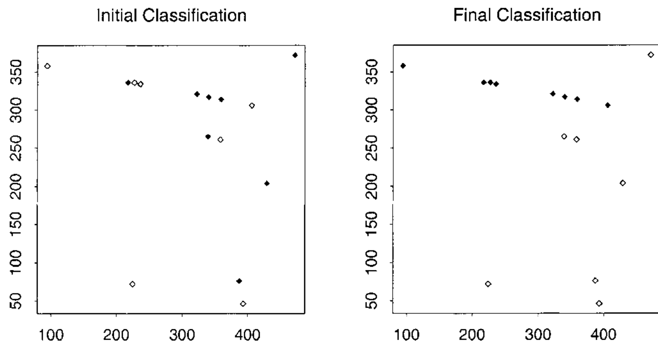

Figure 8. Analysis of image d10. The point detections derived from the image shown in Figure 5 are shown in pixel coordinates. The initial classification consists of random assignments. The final classification is a fit of an elongated ellipse (black) with background noise (white).

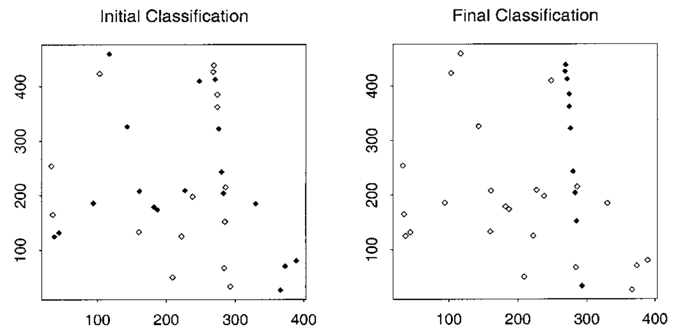

Figure 10. Analysis of image d8. The point detections derived from the image shown in Figure 8 are shown in pixel coordinates. The initial classification consists of random assignments. The final classification is an extremely elongated elliptical cluster (black) with a Poisson noise background (white).

declines if too many points are used. The BIC's value was very reliable for treating data without any apparent aligned set of points. The clustering and BIC approach offers advantages in terms of threshold selection and robustness.

Image preprocessing was carried out in IDL. The analysis of the point patterns was carried out in S-Plus. The code used for the latter is available in the S archive at the Statlib server, http://lib.stat.cmu.edu/S/mclustem.one and also at http://www.stat.washington.edu/fraley/software.html.

The computation time for this model-based clustering is dominated by the formation of the Cholesky factor of the weighted sample crossproduct matrix, for the calculation of $\det(\Sigma_k)$ in the M-step, which is $O(np^2)$. For a given data dimensionality, p, this is linear in the number of observations. The latter number, n, is also small for most applications. A few EM iterations are required to reach convergence.

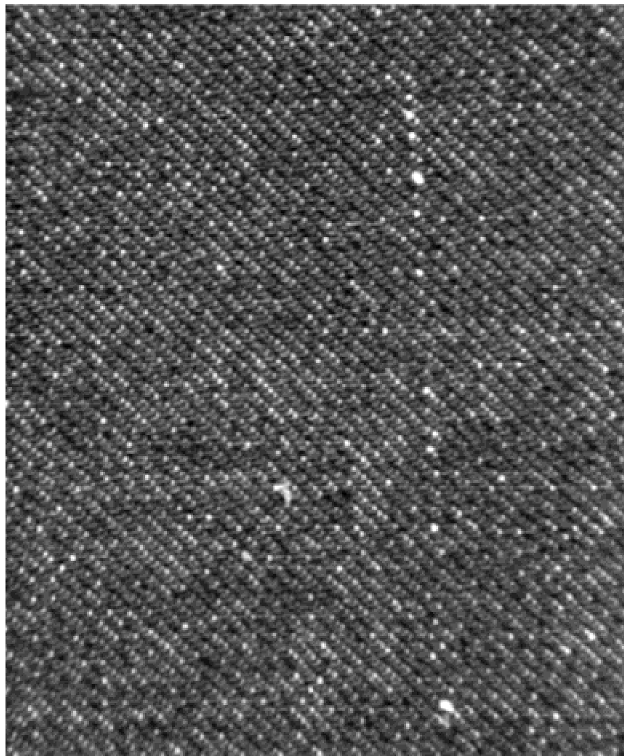

Figure 9. Image used for experimentation referred to as d8.

V. CONCLUSION

We have introduced model-based methods to counter the difficulties that surround the characterization of flaws. Generally, a flaw is defined by its departure from normality, so that characterization of normality is an important part of any model. This is typical of "data mining" type problems.

We described a range of model-based methods for the detection of flaws in knitted and woven textiles and discussed the bases of these methods. These ranged over image thresholding or binarization, the subtraction of "background" fixed pattern, and model-based linearity finding.

We have indicated the effectiveness of these model-based methods in practical case studies. We concentrated on examples chosen for their detection difficulty. We stressed the *perceptual grouping* nature of the extended fault detection problem. Our methods are effective and efficient in dealing with this problem.

ACKNOWLEDGMENTS

This work was supported in part by Office of Naval Research Grants N-00014-96-1-0192 and N-00014-96-1-0330. We are also grateful to the referees for a number of suggestions.

REFERENCES

S. Banerjee and A. Rosenfeld, Model-based cluster analysis, Pattern Recognition 26 (1993), 963–974.

J.D. Banfield and A.E. Raftery, Model-based Gaussian and non-Gaussian clustering, Biometrics 49 (1993), 803–821.

J.G. Campbell, C. Fraley, F. Murtagh, and A.E. Raftery, Linear flaw detection in woven textiles using model-based clustering, Pattern Recognition Lett 18 (1998), 1539–1548.

J. Campbell, A. Hashim, T.M. McGinnity, and T.F. Lunney, Flaw detection in woven textiles by neural network, J.G. Keating (Editor), Fifth Irish neural networks conference, National University of Ireland, Maynooth College, 1992, pp. 92–99.

J.G. Campbell and F. Murtagh, Automatic visual inspection of woven textiles using a two-stage defect detector, Optical Eng 37 (1998), 2536–2542.

G. Celeux and G. Govaert, Gaussian parsimonious clustering models, Pattern Recognition 28 (1995), 781–793.

F.S. Cohen, Z. Fan, and S. Attali, Automatic inspection of textile fabrics using textural models, IEEE Trans Pattern Anal Machine Intelligence 13 (1991), 803–808.

A. Dasgupta and A.E. Raftery, Detecting features in spatial point processes with clutter via model-based clustering, J Am Stat Assoc 93 (1998), 294–302.

A.P. Dempster, N.M. Laird, and D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, J R Stat Soc Series B 39 (1977), 1–22.

L.M. Hoffer, F. Francini, B. Tiribilli, and G. Longobardi, Neural networks for the optical recognition of defects in cloth, Optical Eng 35 (1996), 3183–3190.

W.J. Jasper, S.J. Garnier, and H. Potlapalli, Texture characterisation and defect detection using adaptive wavelets, Optical Eng 35 (1996), 3140–3149.

R.E. Kass and A.E. Raftery, Bayes factors, J Am Stat Assoc 90 (1995), 773–795.

B.G. Leroux, Consistent estimation of a mixing distribution, Ann Stat 20 (1992), 1350–1360.

T. Masters, Signal and Image Processing with Neural Networks, Wiley, New York, 1993.

F. Murtagh and A.E. Raftery, Fitting straight lines to point patterns, Pattern Recognition 17 (1984), 479–483.

T.S. Newman and A.K. Jain, A survey of automated visual inspection, Comput Vision Image Understanding 61 (1995), 231–262.

A.V. Oppenheim and R.W. Schafer, Discrete-time signal processing, Prentice-Hall, Englewood Cliffs, NJ, 1979.

M.S. Sanders and E.J. McCormick, Human factors in engineering and design, McGraw-Hill, New York, 1987.

H. Sari-Sarraf and J.S. Goddard Jr., Robust defect segmentation in woven fabrics, Proc Comput Vision Pattern Recognition 1998 (1998), 938–944.

G. Schwarz, Estimating the dimension of a model, Ann Stat 6 (1978), 461–464.

D.M. Titterington, A.F.M. Smith, and U.E. Makov, Statistical analysis of finite mixture distributions, Wiley, New York, 1985.