©Copyright 2008 Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

Gururajan A., H. Sari-Sarraf, E. F. Hequet, “ Statistical Approach to Unsupervised Defect Detection and Multi-Scale Localization in Two-Texture Images,” *Optical Engineering*, 47(2), 027202-1--027202-10, (2008).

This article can also be found online at:

<http://spiedl.aip.org/getpdf/servlet/GetPDFServlet?filetype=pdf&id=OPEGAR000047000002027202000001&idtype=cvips&prog=normal>

DOI:

10.1117/1.2868783_

Statistical approach to unsupervised defect detection and multiscale localization in two-texture images

Arunkumar Gururajan
Hamed Sari-Sarraf
Texas Tech University
Department of Electrical and Computer
Engineering
Lubbock, Texas 79403
E-mail: arunkumar.gururajan@ttu.edu

Eric F. Hequet
Texas Tech University
International Textile Center
Lubbock, Texas 79403

Abstract. We present a novel statistical approach to unsupervised detection and localization of a chromatic defect in a uniformly textured background. The test images are probabilistically modeled using Gaussian mixture models, and consequently defect detection is posed as a model-order selection problem. The statistical model is estimated using a modified Expectation-Maximization algorithm that aids in faster convergence of the scheme. A test image is segmented only if a defective region/blob has been declared to be present, and this improves the efficiency of the entire scheme. This work places equal emphasis on defect localization; hence, an elaborate statistical multiscale analysis is performed to accurately localize the defect in the image. The underlying idea behind the multiscale approach is that segmented structures should be stable across a wide range of scales. The efficacy of the proposed approach is successfully demonstrated on a large dataset of stained fabric images. The overall detection rate of the system is found to be 92% with a specificity of 95%. All of these factors make the proposed approach attractive for implementation in online industrial applications.

© 2008 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2868783]

Subject terms: defect detection; Gaussian mixture models; Expectation-Maximization; multiscale localization; fabric stain detection.

Paper 070575RR received Jul. 7, 2007; revised manuscript received Oct. 4, 2007; accepted for publication Oct. 8, 2007; published online Feb. 28, 2008. This paper is a revision of a paper presented at the SPIE conference on Machine Vision Applications in Industrial Inspection XV, Jan. 2007, San Jose, California. The paper presented there appears (unrefereed) in SPIE Proceedings Vol. 6503.

1 Introduction

Defect detection constitutes an important part of quality control in industry. There are a wide range of defects that occur during industrial production. Of greater concern and cost are the defects that affect the aesthetics and/or the functionality of the end product. When these defects are not controlled, yields decrease, manufacturing costs rise, and the ability to compete in the market place is compromised. In many industrial applications, defect detection is done visually through trained human inspectors. The main drawback of this approach is that manual inspection is always subjective and the obtained measurements are usually not repeatable. However, the development of fast and specialized computing hardware has facilitated the emergence of image processing algorithms to real-world industrial inspection problems. Currently, machine vision algorithms are used in a broad range of applications, some of which are defect detection in textiles (fabrics),¹⁻³ semiconductors (silicon wafers),⁴ wood,⁵ and steel and ceramic tiles.^{6,7}

Before describing our proposed approach, we briefly survey the taxonomy of defects and defect detection algorithms in the literature. In the context of machine vision, a defect can be broadly defined as a region that locally breaks the homogeneity of texture, and hence causes the container

image to be two-textured.⁸ Based on this definition, a defect can be classified into two categories: chromatic defects and structural defects. As an example, the presence of an oily stain on a fabric can be considered to be a chromatic defect, while a local disruption in the weave pattern of a fabric can be classified as a structural defect. Also, a defect-free background can be either uniformly textured¹⁻³ or can be comprised of a nonstationary texture.⁶ A typical image processing algorithm for defect detection is comprised of two parts—*detecting* the presence of a defect and *localizing* (*segmenting*) the defect in the image.

Defect detection algorithms can be broadly divided into two categories: supervised and unsupervised. Most of the defect detection approaches rely on a supervised scheme wherein the background (defect-free region) is learned through the presentation of appropriate training samples to the machine vision system. Consequently, any part of the test image that does not look similar to the background (learned during training) gets classified as a defect. The recent work in Ref. 6 is an example of a supervised approach for detecting defects in a nonstationary background. Supervised approaches work extremely well in applications where the defect-free background does not change significantly from one image to another. However, in certain online industrial applications, as in Ref. 8, the intersample background variability may be high, which would require the algorithm to be retrained on a new set of images. This,

in turn, can affect the productivity and is one of the main limitations of using a supervised scheme. In such situations, unsupervised algorithms that need minimal knowledge of the defect and the background are preferred.

There are a number of technical approaches for defect detection in the literature, among which probabilistic models^{6,9} are becoming increasingly popular due to their ability to 1. model texture and 2. successfully capture the noisy background variations. In this framework, the basic assumption is that the image pixels are identical and independently distributed (iid) random variables characterized by a parametric density function whose parameters are unknown. Gaussian mixture models (GMM) have been widely used in computer vision problems due to the isotropic nature of the Gaussian functions, and more importantly, mathematical tractability.¹⁰ Maximum likelihood estimates of the parameters of the GMM are generated using the Expectation-Maximization (EM) algorithm or several variants.^{11–13} However, a major drawback of using a GMM of image intensities in defect detection is that the spatial information in the image is completely neglected. These issues have been addressed in the past using Markov random fields (MRFs), wherein a conjugate prior is used to capture the spatial correlation among image pixels during the estimation procedure.^{14–16} For instance, in Ref. 16, Cohen, Fan, and Attali present an online fabric inspection system that models the defect-free textile web as a Gauss-Markov random field (GMRF). These approaches, however, tend to be computationally intensive—typical MRF algorithms visit all the lattice points in a specific order and perform a local computation at each site. This is an iterative process that ends only when some form of convergence is reached.¹⁶

This work presents a fully automated, unsupervised statistical scheme for detection and localization of blob-like defects caused due to chromatic anomalies in a uniformly textured image. The statistical modeling part is based on a GMM of the image intensities. Defect detection is posed as a model order selection problem, and if a defect is declared to be present, it is subsequently segmented. The important aspects of this work that separate it from the existing statistical defect detection approaches are as follows.

1. A modified version of EM¹⁷ that takes into account the spatial continuity of the image pixels is implemented. This modification also ensures faster convergence in likelihood, and makes the algorithm relatively insensitive to initialization. Moreover, this step is computationally inexpensive compared to approaches described in Refs. 14–16, thus, making it more attractive to implementation in online industrial applications.
2. The detection step is performed upfront, and the image is segmented only if a defect is identified. This is in contrast to many existing approaches, including the recent work⁶ where detection and segmentation are performed simultaneously. Since defects are not very frequent in industrial applications, we make the case that having a detection step prior to segmentation improves the efficiency of the overall approach.
3. The detection threshold is based on a function of an information theoretic criterion termed Minimum De-

scription Length (MDL),^{18,19} that takes into account both model fit and model complexity. The formulation of this statistic provides a good tradeoff between the true positive rate and the false positive rate.

4. The proposed approach places equal emphasis on defect localization, because an accurate segmentation is crucial to the subsequent application-specific measurements. In Ref. 6, Xie and Mirmehdi argue that multiscale analysis alleviates the problems of scale selection, and for similar reasons, we propose a novel scheme for multiscale image segmentation based on receiver operating characteristic (ROC) curves. Here, regional information about the defect is obtained at multiple scales, and is optimally fused to yield the final segmentation map.

Finally, a textile application involving a large dataset of stained fabric images is used to evaluate the proposed defect detection approach with suitable measures for validating the accuracy of the detection and localization techniques. A part of this work has been briefly presented in Refs. 20 and 21.

The rest of the work is organized as follows. Section 2 discusses the proposed statistical modeling approach using GMM and its estimation based on the modified EM algorithm. Section 3 elaborates on the mathematical setup for the defect detection scheme. Section 4 describes the Bayesian decision rule for obtaining the segmented image (localized defect) from the estimated density function, and subsequently presents the statistical generation of the best segmentation map using the proposed multiscale approach. This is followed by the experiments and results on the previously mentioned dataset of fabric images. Finally, we end the work with some concluding remarks.

2 Statistical Modeling

In this section, we present the statistical modeling procedure involved in the proposed defect detection scheme. Before outlining the model, we would like to make some brief but important remarks on the preprocessing steps involved prior to modeling. First, it should be noted that though this approach is applicable to color images, all the ensuing statistical schemes in this work are developed for gray-level images. This is because a chromatic defect in a color image would certainly be enhanced in an appropriately chosen color plane. Several methods exist in the literature to tackle the problem of color plane selection,^{22–24} and the recent work by Stokman and Gevers²⁴ describes how different color models could be optimally fused together. In our work, however, the main focus is on defect detection and localization, and we proceed with the assumption that an optimal color plane (gray-level image) has been obtained from the raw test image. The reader should note that, from this point on, the term test image refers to a grayscale image. The second issue is the presence of noisy variations in the background of the test image. This factor becomes significantly intrusive in the defect detection scheme, especially in complex industrial environments with dirt and unfavorable lighting conditions.²⁵ Hence, the traditional approach in defect detection applications^{26,27} is to detect a defect at a coarse scale and localize (segment) the same at a finer scale. Typically, a coarse scale image is obtained by

Fig. 1 (a) Synthetic image containing a defect. (b) Defect-free image.

convolving the test image with a Gaussian filter. It should be mentioned that the scale parameter for the Gaussian filter has to be chosen with care—too coarse a scale might lead to deterioration of the object, and too fine a scale might not be successful in mitigating the effect of noise. Hence, scale selection is highly contextual in defect detection applications, and one would need to know the size of the defect relative to the background to select an appropriate scale. Therefore, we assume that the test image has been suitably filtered at a coarse scale prior to the statistical modeling scheme. Figure 1 displays two synthetic test images, one containing a defect (made up of 5 blobs) [Fig. 1(a)] and the other comprising a noisy defect-free background [Fig. 1(b)].

In keeping with our assumption of a chromatic defect, it would be logical to expect that the histogram of the test image would be bimodal, and hence, we model the image intensities in the object (defective region) and background (defect-free region) as a GMM. We now proceed to define a GMM and introduce suitable notation for the same.

2.1 Gaussian Mixture Models

A mixture model with model order $K > 1$ in $\mathfrak{R}^n$ for the data dimensionality $n \geq 1$ can be mathematically represented as:

$$p(x/\Theta) = \sum_{k=1}^K \alpha_k p(x/\theta_k), \quad \forall x \in \mathfrak{R}^n, \quad (1)$$

where $\alpha_1, \dots, \alpha_k$ are the mixing proportions or the priors, each θ_k is the set of parameters defining the k 'th mixture component, and $\Theta = \{\theta_1, \dots, \theta_k, \alpha_1, \dots, \alpha_k\}$ is the complete superset of parameters needed to specify the finite mixture model. Since the priors are essentially probabilities, $\alpha_k = P(\theta = \theta_k)$ must satisfy

$$0 \leq \alpha_k \leq 1 \quad \text{and} \quad \sum_{k=1}^K \alpha_k = 1. \quad (2)$$

For the mixture to be a Gaussian mixture (finite normal mixture), each component conditional density $p(x/\theta_k)$ should follow a normal probability distribution, that is,

$$p(x/\theta_k) = \frac{1}{(2\pi)^{\frac{n}{2}} \det(\Sigma_k)^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) \right\}, \quad (3)$$

with the parameter $\theta_k = (\mu_k, \Sigma_k)$, where μ is the mean vector and Σ is the covariance matrix of the multivariate Gaussian distribution.

As mentioned previously, we try to model the foreground and background intensities as a mixture of two univariate normal distributions, and hence, the mean vector and covariance matrix reduce correspondingly to a scalar mean μ and variance σ^2 . In our setup, the parameters of the GMM are unknown and need to be estimated before making a decision about the presence/absence of a defect. In the following section, we provide a brief overview of the estimation of the parameters of the GMM via EM, and then proceed to outline the modified EM algorithm adopted in the proposed approach.

2.2 Estimating the Gaussian Mixture Models Via the Expectation-Maximization Algorithm

A statistical model is complete only if both its functional form and its parameters are determined. The commonly used approach for determining the parameter set Θ from a given dataset is through maximum-likelihood estimation. Since we have a multidimensional optimization problem when estimating mixture models, numerical optimization techniques such as EM are preferred over classical calculus-based approaches. The EM algorithm is a general iterative technique for computing the maximum likelihood when the observed data can be regarded as incomplete.¹¹ EM approximates the nonconvex likelihood function by a convex function, and then tries to increase the value of the modified likelihood function at every iteration. The iterative procedure stops when there is convergence in likelihood. If there are N data points, the likelihood function is given by

$$L = \prod_{n=1}^N p(x_n/\Theta), \quad (4)$$

where the right-hand side of this equation is as defined in Eq. (1). EM consists of two steps, an expectation step, where the posterior probabilities are estimated, and a maximization step, where the parameters are updated. This is illustrated as shown next.

2.2.1 Initialization

As with most greedy numerical optimization schemes, EM is extremely sensitive to the initialization of the parameter values. In this work, the K-means algorithm²⁸ is used to initialize the means, variances, and prior values of the GMM.

2.2.2 E-step

Computing the posterior probabilities is as follows. By Baye's rule, the posterior at iteration t can be written as

$$P^{(i)}(\theta_k/x_n) = \frac{p^{(i)}(x_n/\theta_k) \cdot P(\theta = \theta_k)}{p^{(i)}(x_n/\Theta)}, \quad (5)$$

where $p^{(i)}(x_n/\Theta)$ and $p^{(i)}(x_n/\theta_k)$ are computed according to Eqs. (1) and (3), respectively.

2.2.3 M-step: parameter update

The prior, mean, and variance of the GMM are updated as follows:

$$\alpha_k^{(t+1)} = \frac{\sum_{n=1}^N P^{(t)}(\theta_k/x_n)}{N}, \quad (6)$$

$$\mu_k^{(t+1)} = \frac{\sum_{n=1}^N P^{(t)}(\theta_k/x_n) \cdot x_n}{\sum_{n=1}^N P^{(t)}(\theta_k/x_n)}, \quad (7)$$

$$\sigma_k^{2(t+1)} = \frac{\sum_{n=1}^N P^{(t)}(\theta_k/x_n) \cdot \|x_n - \mu_k^{(t+1)}\|^2}{\sum_{n=1}^N P^{(t)}(\theta_k/x_n)}. \quad (8)$$

These steps are repeated iteratively until there is convergence in the likelihood, i.e.,

$$|L^{(t+1)} - L^{(t)}| < \epsilon, \quad (9)$$

where ϵ is a very small and predetermined number.

2.3 Modified Expectation-Maximization Algorithm

It should be noted that the estimation procedure in the previous section does not take the spatial structure of the image pixels into account. A principled approach to incorporate spatial information into the GMM would be to impose spatial constraints by modeling the priors as a hidden Markov random field (HMRF). However, as mentioned in Sec. 1, the computational complexity of such a model is prohibitive. In Ref. 17, Silverman et al. demonstrate that this can be achieved by introducing a smoothing operation in the E step. They also prove that this modified EM converges faster and that it is also relatively insensitive to initialization as compared to the original EM algorithm. We further note that this modification is computationally inexpensive compared to the MRF approach, and we adopt this version of EM in our defect detection scheme. The mathematical scheme for the modified EM is presented next.

Let us consider an image of size M rows by L columns that is modeled by a GMM. In the modified EM algorithm, we represent the posteriors as a 3-D matrix $\mathbf{P}$ of size $M \times L \times K$, where K is the model order of the GMM. Any element $p_{i,j,k}$ is given by

$$p_{i,j,k} = P(\theta_k/x_{i,j}). \quad (10)$$

Hence, the probability value at the point (i,j,k) gives the posterior probability of occurrence of class “ k ” given the data point $x_{i,j}$. Now, if the image is made up of distinct smooth regions, this will directly be translated into the structure of the posterior matrix. For example, if the image is made up of K smooth regions, then it would be logical to expect each of the K posterior matrices to have distinct

smooth regions as well. However, the presence of noise causes perturbations in the posterior values, and since the original EM treats every pixel as an iid random variable, there is a definite bias in the values of the estimated parameters. The bias in the estimation procedure translates into a noisy segmentation result. The modified EM algorithm takes care of the presence of noise by smoothing the posterior matrix that was computed during the expectation step. Smoothing is achieved by Gaussian filtering each of the K matrices consisting of the posterior probability values. Within this mathematical framework, we introduce the method for defect detection.

3 Defect Detection

The defect detection problem is framed as a decision theoretic scheme by posing it as a binary hypothesis testing problem as follows. We assume the absence of a defect under the null hypothesis. Mathematically, the hypotheses can be written in terms of the underlying probability density as

$$H_0: p(x/\Theta) = p(x/\theta_1), \quad (11)$$

$$H_1: p(x/\Theta) = \sum_{k=1}^2 \alpha_k \cdot p(x/\theta_k), \quad (12)$$

where the conditional probability distributions are modeled by univariate Gaussians. The alternate hypothesis H_1 , which indicates the presence of a defect, is represented by a second-order GMM. The problem of deciding between the competing hypothesis reduces to one of model order selection. There are two issues to be kept in mind when deciding the best model: 1. the chosen model should provide a reasonably accurate fit to the observed data, and 2. the model complexity should not be high. It should be noted that the two issues are intimately related, as an increased model complexity leads to a better fit of the model to the data.

Rissanen¹⁸ proposed the concept of using an information theoretic criterion termed minimum description length (MDL) for selecting the appropriate model order. MDL chooses a model that trades off goodness of fit on the observations and model complexity. Basically, if we have two models that fit the data equally well, MDL chooses the simpler model, and hence, avoids the problem of overfitting. The goodness of fit to the data D under a hypothesis H can be expressed by the likelihood function $L(D/H)$, and the model complexity translates into the parametric complexity $C(H)$ of the assumed probability density function (GMM in our case). MDL can be operationalized as follows^{19,29}: the optimal model order k_{opt} is given by the argument k that maximizes the following expression.

$$\log L(\Theta/X) - \frac{m_k}{2} \log N, \quad (13)$$

where m_k denotes the model complexity (number of means, covariances, and priors) of the k 'th order GMM. In the context of the defect detection problem, the MDL under the null and alternative hypotheses can be written as

Table 1 Relative change in MDL values for the selected images.

Image	Relative percent change in MDL
Fig. 1(a)	5.5
Fig. 1(b)	0.6

$$\text{MDL}_{H_0} = \log L_0(\Theta/X) - \frac{m_{k_0}}{2} \log N, \quad (14)$$

$$\text{MDL}_{H_1} = \log L_1(\Theta/X) - \frac{m_{k_1}}{2} \log N, \quad (15)$$

where $\log L_0(\Theta/X)$ and $\log L_1(\Theta/X)$ are the log-likelihood functions defined for the density functions (under the null and alternate hypotheses) as given in Eqs. (11) and (12). These likelihood functions are evaluated by estimating the density of the mixture model (under each of the hypotheses) using the modified EM algorithm that was described in Sec. 2. The k 'th order univariate GMM has k unknown means, k unknown variances, and $k-1$ unknown priors. Hence, the expression for model complexity m_k is given by

$$m_k = k + k + (k-1) = 3k-1. \quad (16)$$

In the current scenario, m_{k_0} and m_{k_1} take the values 2 and 5, respectively. If the model under the alternate hypothesis fits the data well and has lesser complexity as compared to the model under the null hypothesis, we would have

$$\text{MDL}_{H_1} > \text{MDL}_{H_0}. \quad (17)$$

However, in most cases, images are corrupted by noise (white or colored), and this leads MDL to always favor the formation of a second- (or higher) order GMM. Therefore, we choose the test statistic T as follows

$$T = \left| \frac{\text{MDL}_{H_1} - \text{MDL}_{H_0}}{\text{MDL}_{H_1}} \right|. \quad (18)$$

Essentially, the term T denotes the relative change in MDL value between the null and the alternate hypotheses. We then compare T with a preset empirical threshold t , and reject the null hypothesis if $T > t$. It should also be mentioned that in cases of images where multiple (>2) Gaussians are present (violation of the assumption of a single object and background), this statistic again becomes very small and leads to the acceptance of the null hypothesis. Table 1 shows the values of the test statistic based on Eq. (18) for the images displayed in Fig. 1.

We can see that there is a significant difference between the corresponding values, making it easy to set a suitable threshold value t to classify the defective sample.

4 Defect Localization

If a defect is detected, a Bayesian decision rule is applied to localize the same in the test image. We label each pixel x_{ij}

Fig. 2 (a) Test image containing a defect. (b) Segmented image at the coarse scale used for detection.

in the image into two classes: ω_1 (defective) and ω_2 (defect free) by applying the following decision rule:

$$x_{ij} = \begin{cases} \omega_1 & \text{if } \alpha_{\omega 1} p(x_{ij}/\theta_{\omega 1}) > \alpha_{\omega 2} p(x_{ij}/\theta_{\omega 2}) \\ \omega_2 & \text{otherwise} \end{cases}. \quad (19)$$

It should be mentioned that direct application of this rule on the coarsely filtered image leads to poor object localization in the segmentation map. Figure 2 shows the case of the synthetic test image segmented at the coarse scale (used for detection). The challenge here lies in estimating the characteristic scale of the foreground object. Lindeberg³⁰ proposed a technique for estimating the optimal scale in an image. However, in most cases, a single scale does not suffice to localize the object accurately. Hence, we propose to take the image through a range of scales, segment the image at each scale, and generate the optimal segmentation map using statistical receiver operating characteristics. This approach is based on a similar idea³¹ that has been proposed for adaptive edge detection.

4.1 Multiscale Analysis

In this procedure, the image is taken through a number of scales, varying from a fine scale (application dependent) to a coarse scale (used for object detection). The block diagram of the process is shown in Fig. 3. The modified EM algorithm is used in each case to estimate the density function, and the filtered image at each scale is segmented using the Bayesian segmentation rule as described before. If the image is taken through N_s scales, there are N_s resulting segmentation maps. For the sake of clarity, we term these N_s maps as original segmentation maps (OSMs). It is assumed here that the labels corresponding to the defect (ω_1) and the background (ω_2) are known. The segmentation results are then tested for correspondence. A pixel that has been labeled as belonging to the defect class in all the N_s maps has the highest correspondence, while a defect class pixel that has been labeled in only one map would have the lowest correspondence. Therefore, we can apply N_s correspondence thresholds to the N_s segmentation maps to yield N_s projected segmentation maps (PSMs). For example, a correspondence threshold of $n_s < N_s$ would yield a PSM containing pixels that occur in at least n_s of the segmentation maps. Each of the N_s PSMs are now compared to the N_s OSMs to yield the true positives and false positives. True positives and false positives are defined as follows: if

Fig. 3 Block diagram of the data-driven multiscale segmentation.

a pixel in a particular PSM labeled as belonging to class ω_1 (defective) corresponds to a similar labeling (ω_1) in the OSM, then it is termed as a true positive and can be expressed as follows,

$$TP = P(\omega_{1_{\text{PSM}}} / \omega_{1_{\text{OSM}}}), \quad (20)$$

where $P(\cdot)$ refers to the probability, and $\omega_{1_{\text{PSM}}}$ and $\omega_{1_{\text{OSM}}}$ refer to the pixel having a label of ω_1 in PSM and OSM, respectively. On the other hand, if a pixel in PSM labeled as belonging to class ω_1 corresponds to a contradictory label (ω_2) in OSM, then it is termed as a false positive, that is,

$$FP = P(\omega_{1_{\text{PSM}}} / \omega_{2_{\text{OSM}}}). \quad (21)$$

In a similar manner, we can also define and compute the true negatives (TN) and the false negatives (FN) as follows.

$$TN = P(\omega_{2_{\text{PSM}}} / \omega_{2_{\text{OSM}}}), \quad (22)$$

$$FN = P(\omega_{2_{\text{PSM}}} / \omega_{1_{\text{OSM}}}). \quad (23)$$

Therefore, the average values of the true positives ($\overline{TP}$), false positives ($\overline{FP}$), true negatives ($\overline{TN}$), and false negatives ($\overline{FN}$) can be obtained by dividing the respective sums by N (total number of pixels in the image). The true positive rate (TPR) and the false positive rate (FPR) can now be defined as follows³²:

$$TPR = \frac{\overline{TP}}{p}, \quad (24)$$

$$FPR = \frac{\overline{FP}}{1 - p}, \quad (25)$$

where

$$p = \overline{TP} + \overline{FN}. \quad (26)$$

The TPR and FPR are also referred to as the *sensitivity* and $(1 - \text{specificity})$, respectively. The ROC is a monotonically increasing curve representing the TPR versus the FPR . The optimal segmentation map corresponds to the point on the ROC that maximizes the TPR and minimizes the FPR . This is because the defective regions are salient structures in the image and would be expected to appear across most of the scales as true positives. Noisy background variations, however, tend to vanish at higher scales and add to the FPR . Therefore, the step of optimizing the TPR and FPR values essentially translates into maintaining structures (corresponding to the defect) in the segmentation map that are stable across scales. The selection of the optimal operating point on the ROC curve is achieved using the chi-square measure defined in Ref. 32 as

$$\chi_{\text{ROC}}^2 = \frac{(\text{sensitivity} - Q) [\text{specificity} - (1 - Q)]}{1 - Q} \cdot \frac{1}{Q}, \quad (27)$$

where $Q = \overline{TP} + \overline{FP}$.

A higher χ^2 indicates a better operating point.³² Therefore, the desired point on the ROC is chosen to be that value that maximizes the expression in Eq. (27), yielding the optimal segmentation map. Figure 4 shows the results of segmentation on the synthetic image using just a coarse scale versus using the proposed multiscale analysis. It can be seen visually that the latter provides a better localization of the defect. We also perform a quantitative evaluation of the proposed scheme, which is described in the following section.

4.2 Quantitative Assessment of Defect Localization

To validate the segmentation results of the proposed scheme on a set of test images, we need to have the following: 1. the ground truth segmentations for the test images, and 2. a suitable measure for quantifying the algorithmic segmentation compared to the ground truth.

Under the assumption that the ground truth is available, there are several performance evaluation measures^{33–35} for assessing segmentation algorithms. Since the proposed localization scheme is a region-based technique, we need to use a suitable measure to compare two region segmented images. Huo and Giger³⁵ proposed a region-based measure that takes into account the area of overlap between the regions in the automatically segmented image and the ground truth. We adopt this measure to evaluate the segmentation results in our approach.

Let us denote the set of all regions segmented by the algorithm and the ground truth as A and G , respectively. Then the segmentation measure in Ref. 35 is given by

Fig. 4 (a) The original test image containing the defect. (b) Ground truth segmentation for the corresponding image. (c) Result of segmentation at the coarse scale used for detection. (d) Segmentation result obtained using the proposed localization method.

$$S = \frac{|A \cap G|}{|A \cup G|}, \quad (28)$$

where $|\cdot|$ denotes the cardinality of the set. If the algorithmic segmentation is identical to the ground truth, the numerator and denominator of the prior expression equal each other and S becomes unity. If the algorithmic segmentation is completely wrong, the area of overlap (numerator) becomes zero, and hence, S reduces to zero. Table 2 shows the values of S for the segmentation results on the synthetic image (Fig. 4) using just a coarse scale versus that obtained using the proposed localization scheme (multiple scales).

It can be seen that the proposed scheme improves accuracy by around 20%, and in many applications, this difference can significantly impact subsequent measurements.

5 Experiments and Results

We present in this section an interesting application of this algorithm in the arena of fabric quality inspection. The textile industry is particularly interested in quantitatively evaluating the ability of fabrics to release stains after laundering. This involves the visual examination of a stained and laundered fabric by a technician, and assignment of a

stain release grade (from 1 to 5) based on a reference set of stained images.³⁶ In this application, the intersample background variation is high between different fabrics, and hence an unsupervised approach is more suited for the same. In this context, we apply our algorithm to a set of digitized fabric images, detect if a stain is present, and subsequently localize the stain if it has been detected. We report the detection results in terms of true positives (sensitivity) and false positives (1 - specificity). Also, we report the quantitative results of the proposed multiscale segmentation scheme in terms of the segmentation measure described in Sec. 4.2.

5.1 Test Dataset

The test dataset was comprised of a total of 360 fabric images consisting of three fabric colors, namely khaki, white, and yellow. Six types of commonly occurring stains were used in the experiment, and the fabric samples were prepared according to a standard method, as prescribed in Ref. 36. It should be mentioned that some of the selected stains were those that could be easily removed by laundering, whereas the other stains were significantly harder to remove. This was done to test the performance of the system with respect to the generation of false positives.

The fabrics were digitized using a color scanner at a resolution of 75 dots per inch. As the fabric images were in color (multiplane), a preprocessing scheme was developed to select the best plane for the defect detection algorithm. The color image was decomposed into the red, green, blue, hue, and saturation channels,³⁷ and subsequently the principal components analysis³⁷ was applied to the five color channels. Since the first principal component accounts for most of the total variance of the data (around 75 to 85% in this application), we selected the same for further processing. To mitigate the effect of noise, the first principal component was smoothed using a Gaussian filter at a coarse scale. It should be mentioned that the typical size of a test image was around 200 pixels by 200 pixels, and the defect (stain) size varied from 1500 pixels to about as high as 15,000 pixels.

5.2 Parameter Value Selection

There are three important parameters in the proposed algorithm that can be customized for any application. They are: 1. detection threshold t , 2. the range of scales for the multiscale localization, and 3. epsilon parameter in the EM algorithm. For the prior application, the detection threshold was set at 2%. This value was chosen empirically, and it should be mentioned that this parameter should be chosen with care, as a high threshold value could lead to a reduced true positive rate, while a low value could lead to a high false positive rate. The range of scales selected for the Gaussian filter were from 0.5 (fine scale) to 4 (coarse scale) in steps of 0.5, which amounted to a total of eight different scales. Again, as mentioned in Sec. 2, the scale values for the Gaussian filter are application dependent. The parameter epsilon in our application was chosen to be 0.01. In other words, the expectation-maximization (EM) algorithm will be terminated if the percent difference in the likelihood values between successive iterations is less than 1%. If the value of epsilon is too high, it is possible that EM has not converged, and hence the estimated parameters (mean, vari-

Table 2 Segmentation accuracy for the defective image at different scales.

Image in Fig. 1(a)	Segmentation accuracy S
Coarse scale	0.71
Multiscale analysis	0.92

Fig. 5 (a), (b), and (c) Fabric images containing stains. (d), (e), and (f) Corresponding segmented images using the proposed approach.

ance, and priors) may not correspond to the maxima of the likelihood function. On the other hand, if epsilon is too low, the algorithm might take too long to converge, and would computationally prohibit its implementation in online systems. Therefore, a tradeoff in the value of epsilon is usually desirable. Finally, it should be noted that the previously mentioned parameter values have been used to generate all the results presented in this work.

5.3 Results

Before going to the detection and segmentation results, we outline the generation of the ground truth. The stained fabrics were examined by three trained technicians, and stain release grades from 1 through 5 were assigned manually to the test images. According to Ref. 36, a grade of 1 indicates a strong presence of a stain, whereas a grade of 5 indicates the absence of a stain. In our experiments, the grades assigned by the three technicians were averaged, and an average stain release grade of above 4 was taken to indicate the absence of a stain. For evaluating the segmentation results, a set of 85 images containing stains (defects) were selected and hand segmented by two experts.

The overall results of detection on the 360 images are presented in terms of true positive percent and false positive percent. A true positive occurs when the proposed system detects a stain (rejects the null hypothesis) and is in agreement with the ground truth. Similarly, a false positive is defined as a case where the system detects a stain and is not in agreement with the ground truth. We obtained very good detection results with a true positive rate (sensitivity) of 92% and a false positive rate (1 - specificity) of 5%. The segmentation results for the selected 85 images were generated and compared to the corresponding region-segmented ground truth images using the segmentation measure S discussed earlier. The average localization accuracy (S) was found to be 0.80 with a standard deviation of 0.12. These values are very acceptable for this application, and are an indicator of the robust performance of the system. It should be mentioned that in certain cases, stains are usually broken into a main blob and a series of smaller blobs due to the action of the detergent. However, during the ground truth generation, only the main blob was land-

Fig. 6 (a) Mass in a mammogram image. (b) Defect on a SEM image of a semiconductor wafer. (c) Segmented mammogram. (d) Segmented semiconductor wafer.

marked by the experts. The algorithmic segmentation, however, generated both the main and smaller blobs. If we considered only the region corresponding to the main blob in the segmented image and the ground truth for computing the localization accuracy, the segmentation measure S increases to 0.91. Figure 5 displays sample fabric images and the corresponding segmentation results. The final step of this application is to extract suitable measurements from the segmented stains, and assign an appropriate grade to the stained fabric. This is currently being developed by the authors, and it should be mentioned that initial results have been promising.

As an ending note, we would like to emphasize that the proposed scheme can be customized effectively to other industrial and medical applications that suit the assumption of a chromatic defect in a uniformly textured background. Figure 6 shows two cases where the proposed approach is used to successfully detect and segment: a mass in a mammogram [Fig. 6(a)] and a defect in a scanning electron microscope (SEM) image of a semiconductor wafer [Fig. 6(b)].

6 Discussions and Conclusion

We address the problem of detecting and optimally segmenting a chromatic defect in a uniformly textured image. The proposed method is fully automatic, and hence, can be used for online inspection problems, where the test image background changes often. There are only three main parameters to be tuned in the defect detection scheme, and this adds to the ease of customization. Defect detection is performed upfront, and the algorithm proceeds to the localization stage only if the test image is reported to contain a defect. In a practical scenario, defects occur less frequently, and therefore, this step improves the overall efficiency of

the scheme. Multiscale segmentation plays an important part in the scheme, as accurate defect localization is crucial for application-specific measurements that can be made from the test image. The computational complexity of the proposed scheme depends mainly on the number of scales selected for the multiscale segmentation. This is because at each scale, we perform a parameter estimation step using the modified EM algorithm. Finally, it should be mentioned that this algorithm has been validated on a large dataset of two-texture images and has been found to yield excellent detection and segmentation results.

Acknowledgment

The authors would like to thank Cotton Incorporated for partially funding this research project.

References

1. J. Escofet, R. Navarro, M. Millan, and J. Pladellorens, "Detection of local defects in textile webs using Gabor filters," *Opt. Eng.* **37**(8), 2297–2307 (1998).
2. A. Serdaroglu, A. Ertuzun, and A. Ercil, "Defect detection in textile fabric images using wavelet transforms and independent component analysis," *Patt. Recog. Image Anal.* **16**(1), 61–64 (2006).
3. Y. F. Zhang and R. R. Bresee, "Fabric defect detection and classification using image analysis," *Text. Res. J.* **65**(1), 1–9 (1995).
4. R. Jain, A. Rao, A. Kayaalp, and C. Cole, "Machine vision for semiconductor wafer inspection," in *Machine Vision for Inspection and Measurement*, H. Freeman, Ed., pp. 283–314, Academic Press, New York (1989).
5. O. Silven, M. Niskanen, and H. Kauppinen, "Wood inspection with non-supervised clustering," *Mach. Vision Appl.* **13**, 275–285 (2003).
6. X. Xie and M. Mirmehdi, "Texems: texture exemplars for defect detection on random textured surfaces," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(8), 1454–1464 (2007).
7. C. Boukouvalas, J. Kittler, R. Marik, and M. Petrou, "Automatic color grading of ceramic tiles using machine vision," *IEEE Trans. Ind. Electron.* **44**(1), 132–135 (1997).
8. H. Sari-Sarraf and J. S. Goddard, "Vision system for on-loom fabric inspection," *IEEE Trans. Ind. Appl.* **35**(6), 1252–1259 (1999).
9. Z. Liang, R. J. Jaszczak, and R. E. Coleman, "Parameter estimation of finite mixtures using the EM algorithm and information criteria with application to medical image processing," *IEEE Trans. Nucl. Sci.* **39**(1), 1126–1133 (1992).
10. S. Medasani and R. Krishnapuram, "A comparison of Gaussian and Pearson mixture modeling for pattern recognition and computer vision applications," *Pattern Recogn. Lett.* **20**(1), 305–313 (1999).
11. M. M. Dempster, N. M. Laird, and D. B. Jain, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. Ser. B (Methodol.)* **39**(2), 1–38 (1977).
12. R. A. Redner and H. F. Walker, "Mixture densities, maximum likelihood, and the EM algorithm," *SIAM Rev.* **26**, 195–239 (1984).
13. M. A. T. Figueredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(3), 381–396 (2002).
14. S. Z. Li, *Markov Random Field Modeling in Computer Vision*, Springer-Verlag, Berlin, Germany (1995).
15. Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm," *IEEE Trans. Med. Imaging* **20**(1), 45–57 (2001).
16. F. Cohen, Z. Fan, and S. Attali, "Automated inspection of textile fabrics using textural models," *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(8), 803–809 (1991).
17. B. W. Silverman, M. C. Jones, J. D. Wilson, and D. W. Nychka, "A smoothed EM approach to indirect estimation problems, with particular reference to stereology and emission tomography," *J. R. Stat. Soc. Ser. B (Methodol.)* **52**, 271–324 (1990).
18. J. Rissanen, "Modeling by shortest data description," *Automatica* **14**, 465–471 (1978).
19. G. Schwarz, "Estimating the dimension of a model," *Ann. Stat.* **6**, 461–464 (1978).
20. A. Gururajan, H. Sari-Sarraf, and E. F. Hequet, "Statistical modeling, detection, and segmentation of stains in digitized fabric images," *Proc. SPIE* **6503**(650304), 1–8 (2007).
21. A. Gururajan, E. F. Hequet, and H. Sari-Sarraf, "Automatic measurement of soil release through image analysis," *Proc. Beltwide Cotton Conf.*, New Orleans, LA (2007) (abstract only).
22. T. Gevers and A. W. M. Smeulders, "Color based object recognition," *Pattern Recogn.* **32**, 453–464 (1999).
23. L. Macaire, N. Vandenbroucke, and J. G. Postaire, "Color image segmentation by pixel classification in an adapted hybrid color space: application to soccer analysis," *Comput. Vis. Image Underst.* **90**(2), 190–216 (2003).
24. H. Stokman and T. Gevers, "Selection and fusion of color models for image feature detection," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(3), 371–381 (2007).
25. R. T. Chin, "Automated visual inspection algorithms in computer vision," in *Computer Vision: Theory and Industrial Applications*, Torras, Ed., pp. 377–404, Springer-Verlag, New York (1992).
26. C. Neubauer, "Segmentation of defects in textile fabric," *Proc. 11th Intl. Conf. on Patt. Recogn.*, pp. 688–691 (Aug. 1992).
27. K. Ratcliff, A. Harris, C. Allen, and R. Bicker, "Three-dimensional datuming and inspection of mechanical parts, using wire-frame modeling," *Proc. SPIE*, **1197**, 230–236 (1989).
28. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Wiley Interscience, New York (2001).
29. J. Rissanen, *Stochastic Complexity in Statistical Inquiry*, World Scientific, New York (1989).
30. T. Lindeberg, "Feature detection with automatic scale selection," *Int. J. Comput. Vis.* **30**(2), 77–116 (1998).
31. Y. Yitzhaky and E. Peli, "A method for objective edge detection evaluation and detector parameter selection," *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(10), 507–512 (2003).
32. H. C. Kraemer, *Evaluating Medical Tests: Objective and Quantitative Guidelines*, Sage Publications, Newbury Park, CA (1992).
33. Y. J. Zhang and J. J. Gerbrands, "Objective and quantitative segmentation evaluation and comparison," *Signal Process.* **39**(3), 43–54 (1994).
34. M. Borsotti, P. Campadelli, and R. Schettini, "Quantitative Evaluation of color image segmentation results," *Pattern Recogn. Lett.* **19**(8), 741–747 (1998).
35. Z. M. Huo and M. L. Giger, "Evaluation of an automatic segmentation method based on performances of an automated classification method," *Proc. SPIE* **3981**, 16–21 (2000).
36. *AATCC Technical Manual 130-1995*, American Association of Textiles (1995), pp. 217–219.
37. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ (2002).

Arunkumar Gururajan received a BE degree with distinction in electronics and communication from Bharathidasan University in 2001, and a MS degree in electrical engineering from Texas Tech University in 2003. He is currently a PhD candidate in the Department of Electrical and Computer Engineering at Texas Tech University. He is also a research associate at the International Textile Center, developing machine vision algorithms for the textile industry. Prior to his enrolment in the PhD program, he worked as a researcher at the Bioacoustics Research Program at Cornell University. His research interests include statistical inference in computer vision and pattern recognition.

Hamed Sari-Sarraf received his BS, MS, and PhD in electrical engineering from the University of Tennessee, Knoxville, in 1984, 1986, and 1993, respectively. He was a research and development staff member at Oak Ridge National Laboratory until 1999. He is currently an associate professor in the Department of Electrical and Computer Engineering at Texas Tech University. His research interests include applied computer vision.

Eric Hequet is Research Associate Professor in the Department of Plant and Soil Science and Associate Director of the International Textile Center (ITC) at Texas Tech University. He holds a joint appointment with the Texas AgriLife Research in the Texas A&M system. He holds a PhD from the University of Haute Alsace in France. Before joining the ITC, his experience progressed from cotton breeding at experiment stations in Africa, to head of CIRAD's Cotton Technology Laboratory in Montpellier - France, then to Director of

the international cotton program for CIRAD. He has generated 36 refereed journal publications, 2 books, and more than 100 conference papers. He has provided international leadership in research on the measurement of cotton fiber properties and contaminants, including the impacts of these on textile processing performance. He is currently focused on collaborative research with the cotton breeding and cotton biotechnology community to develop improved properties in cotton fibers.