

Rochester Institute of Technology

RIT Digital Institutional Repository

Articles

Faculty & Staff Scholarship

2006

Automated Algorithm for the Identification of Artifacts in Mottled and Noisy Images

Onome Augustine Ugbeme
Rochester Institute of Technology

Eli S. Saber
Rochester Institute of Technology

Wencheng Wu
Xerox Corporation

Kartheek Chandu
Rochester Institute of Technology

Follow this and additional works at: <https://repository.rit.edu/article>

Recommended Citation

Onome Augustine Ugbeme, Eli Saber, Wencheng Wu, Kartheek Chandu, "Automated algorithm for the identification of artifacts in mottled and noisy images," *Journal of Electronic Imaging* 16(3), 033015 (1 July 2007). <https://doi.org/10.1117/1.2761920>

This Article is brought to you for free and open access by the RIT Libraries. For more information, please contact repository@rit.edu.

Copyright 2007 Society of Photo-Optical Instrumentation Engineers.

This paper was published by SPIE and is made available as an electronic reprint (preprint) with permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

Automated algorithm for the identification of artifacts in mottled and noisy images

**Onome Augustine Ugbeme
Eli Saber**

Rochester Institute of Technology
Department of Electrical Engineering
Rochester, New York 14623
E-mail: essee@rit.edu

Wencheng Wu
Xerox Corporation
800 Phillips Road
Webster, New York 14580

Kartheek Chandu
Rochester Institute of Technology
Department of Electrical Engineering
Rochester, New York 14623

Abstract. We describe a method for automatically classifying image-quality defects on printed documents. The proposed approach accepts a scanned image where the defect has been localized a priori and performs several appropriate image processing steps to reveal the region of interest. A mask is then created from the exposed region to identify bright outliers. Morphological reconstruction techniques are then applied to emphasize relevant local attributes. The classification of the defects is accomplished via a customized tree classifier that utilizes size or shape attributes at corresponding nodes to yield appropriate binary decisions. Applications of this process include automated/assisted diagnosis and repair of printers/copiers in the field in a timely fashion. The proposed technique was tested on a database of 276 images of synthetic and real-life defects with 94.95% accuracy. © 2007 SPIE and IS&T. [DOI: 10.1117/1.2761920]

1 Introduction

Image quality stands among one of the most important attributes of image acquisition and printing devices. In the past decade, we have seen a tremendous improvement in the quality of printed documents due to significant technological advances in non-impact printing. Present-day print engines are required to meet consistent and stable image-quality requirements as measured by various metrics and ultimately evaluated by customers. The current marketplace demands the best image quality at competitive costs with minimum downtime. Hence, the ability for print engine vendors to reliably achieve the highest levels of quality will ensure them a leadership role in the printing industry.

Even though the quality of printed documents has im-

proved significantly over the past decade, current print engines still produce a variety of image defects and artifacts. Figure 1 illustrates sample defects that often result from faults or degradations in the underlying image development and electrophotographic processes. These artifacts are manifested in a variety of ways and can occur in different locations on the printed document. Therefore, current print environments (e.g., Print shops) utilize trained quality assurance personnel (QAP) to visually inspect a subset of the output documents to “ensure” that customer hard copies are free of defects. Once a defect has been spotted, the QAP is required to render an initial classification (i.e., deletion, spot, debris centered deletion (DCD), etc.). This is usually done in a visual fashion and as such is prone to errors and subjective judgments. Based on the QAP’s assessment, the artifact’s signature is then utilized to search the diagnostic documents for potential corrective actions (i.e., clean the charge devices, replace the developer, etc.). This is expensive, and time-consuming and occasionally results in false actions due to incorrect assessments. Hence, algorithms that are capable of identifying artifacts in an objective fashion in order to render appropriate actions by QAPs or less experienced operators are needed.

One approach for tackling this problem is by utilizing Content-Based Information Retrieval (CBIR) techniques. In an effort to subdue manual annotations of large image databases, research interests in this area have abounded since the early 1990s.^{1,2} CBIR employs visual properties of a query image—*color, shape, texture, frequency, and regions of interest (ROI)*—to traverse image databases according to user’s interests. This is achieved by utilizing highly descriptive multidimensional feature vectors that can be extracted from global or local positions within the image us-

Paper 06126RR received Jul. 22, 2006; revised manuscript received Mar. 21, 2007; accepted for publication Apr. 21, 2007; published online Aug. 13, 2007.

1017-9909/2007/16(3)/033015/11/\$25.00 © 2007 SPIE and IS&T.

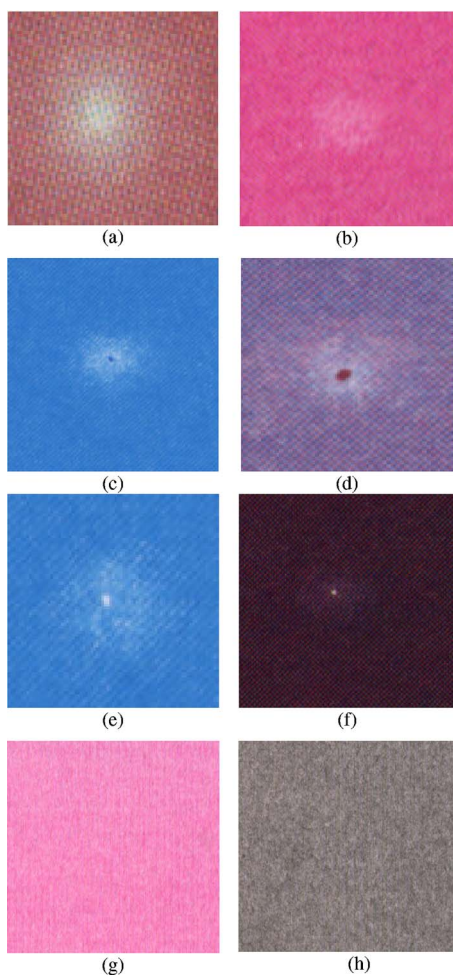


Fig. 1 A sample manifestation of defects of interest: (a) and (b) deletion; (c) and (d) debris centered deletion; (e) and (f) debris-missing DCD; (g) and (h) mottle.

ing one or more of the above visual content properties. Some applications require features that are minimally sensitive to changes—affine, reflections, perspective, deformations, or luminance³—in the image. For instance, shape features can be represented with moment invariants,⁴ Fourier descriptors,⁵ or chain-codes.⁶ Similarity measures between the feature vector of a query sample and archived features corresponding to database images are then computed using distance metrics such as Euclidean,^{6,7} Minkowski,⁸ Mahalanobis,⁹ and Kullback–Leibler (KL) divergence.¹⁰ The decision of a “matched” image is achieved by searching through the similarity values. This can be performed using an exhaustive search or more efficiently via indexing schemes. Popular indexing schemes include R^* -trees,¹¹ quad-trees,¹² or self-organizing maps (SOMs).^{13,14} Thus, an automatic defect identification system can be devised to compare a defective sample with images in an already existing online database from which top matches can be selected. Additionally, the system can be designed to dynamically perform updates when new defects are encountered.

Some systems counteract the defect recognition problem via a series of image processing and pattern classification steps. Iivarinen and Visa¹⁴ employed unsupervised SOMs for the classification of base paper surface defects such as

holes and spots. Their classification is based on features extracted from the internal structure, shape, and texture traits of the defective areas. Segmentation and morphological operations are among the preprocessing steps carried out prior to feature extraction and classification. Additionally, weld defect detection has been accomplished by Li and Liao¹⁵ via Gaussian distribution functions of horizontal line profiles from the defect images. They also employ background removal, dark image enhancement, and image normalization before properties of the profiles are exploited as features. Mery and Berti¹⁶ used textural features from co-occurrence matrices and two-dimensional Gabor functions. Prior to the feature extraction stage, they explore Laplacian of Gaussian edge detection to segment “key” defective areas. More recently, Ng¹⁷ proposed a novel histogram-based thresholding scheme to assist in the segmentation of “small-sized” sheet metal defects. In general, the algorithms described above are designed to handle defects that possess a high-level contrast to the surrounding background and are not well suited for recognition of artifacts in printed documents.

In this paper, we propose a new classification algorithm for identifying artifacts resident on scanned copies of printed documents. The proposed algorithm utilizes primarily size-, shape-, and/or region-based features to effectively classify local artifacts. These artifacts tend to have large intraclass variations, possess low contrast, and/or exhibit illumination non-uniformities. The classes of interest [see Figs. 1(a)–1(h)] include deletions, DCDs, debris-missing DCD, and mottle. Hence, our proposed technique employs appropriately designed image processing steps such as de-screening, gray-scale conversion, and local normalization to ensure uniformity among samples of the same class and to facilitate subsequent shape, size, and region feature analysis. The effectiveness of our algorithm is demonstrated on a large database of scanned electrophotographic and synthetic images. These images possess significant variations within each class that serve to thoroughly test the sensitivity of the method to intraclass differences.

The advantages of our proposed algorithm lie in its ability to objectively classify a given artifact, once localized by a QAP or more importantly a less experienced operator in a practical print shop environment. To this effect, once a defect has been visually detected, the operator/QAP scans the region of interest and proceeds to classify/identify the defect using our proposed algorithm. Once identified, a set of correctives actions is subsequently followed to eliminate the culprit and place the device back into operation as quickly as possible, thereby minimizing downtime and maximizing profits. The remainder of this paper is organized as follows. Section 2 provides a brief description of the defects. Section 3 describes the proposed artifact analysis and classification algorithm. Experimental results are presented in Section 4, and conclusions are drawn in Section 5.

2 Overview of Image-Quality Defects

Image defects are undesirable qualities of an electrophotographic copy produced by a printer or copier. Even though print engines have been thoroughly tested during the manufacturing process, the occurrences of defects remain inevitable due to the volume and diversity of printed material.

Among the most common are deletions, mottle, DCD, streaks, bands, and moiré patterns. Deletions [Figs. 1(a) and 1(b)] are usually manifested as elliptical regions containing a localized group of pixels lighter than the uniform background and are, in general, a result of an error in the charging process. DCDs [Figs. 1(c) and 1(d)], on the other hand, resemble a deletion with the added presence of a centralized collection of localized dark pixels (called debris). A special case of DCDs exists, where the debris is missing due to accidental “rub off” by the electrophotographic process revealing the paper color [see Figs. 1(e) and 1(f)]. This yields abnormally bright gray-level values compared to its immediate neighboring pixels. Mottle defects [Figs. 1(g) and 1(h)] refer to non-uniformity in the perceived print density (i.e., reflectance) and can be gauged by the relationship between light and dark regions. The ISO-13660 standard^{18,19} defines mottle as non-uniformity occurring on a spatial scale between 1.27 mm and 12.7 mm. It characterizes it as large-area print-quality attributes possessing aperiodic fluctuations of density at a spatial frequency less than 0.4 cycles per millimeter in all directions. Even though mottle has received quite a bit of attention in the industry, a quantifiable universal measure has yet to be developed. The best achieved benchmark so far has been introduced in the ISO-19751 image-quality standards for systems.^{20,21} Other image defects (e.g., streaks and bands) are certainly of significant interest but are not handled in this paper, since their shape properties differ greatly from the elliptical shapes identified by the proposed method.

3 Proposed Artifact Analysis and Classification Algorithm

Our proposed method is summarized in Fig. 2. It is composed of two major steps: image preprocessing for de-screening and normalization, and ROI identification and classification. The input to the algorithm is a scanned color sample (i.e., digital image) containing the defect. The flow-chart, shown in Fig. 2, represents a classification approach based on the expected statistics of the ROI in the images. It depicts five leaf nodes and four decision nodes that are used to represent the classes and decisions, respectively. Each of the decision nodes employs size and/or shape features to classify the defect by utilizing empirically selected thresholds. These procedures are described in the following subsections, starting with the preprocessing steps.

3.1 Image Preprocessing

Locating ROIs involves applying a segmentation routine to separate the image into foreground and background segments. The foreground region consists of pixels that represent the defective areas. However, direct thresholding of the images without preprocessing generally leads to inaccurate ROI selections. Scanned images tend to contain halftone screens (or marking screens) used to produce the image on the substrate during the electrophotographic, ink-jet, or lithographic marking process. Researchers have developed de-screening procedures to counteract this problem. Dunn and Mathew²² treated the halftone screens as textures capable of being extracted using a single circularly symmetric filter. Sharma²³ developed a process responsible for determining the identity of the underlying marking process by analyzing the power spectra of a digital image for the

presence of significant energy at high frequencies. This information is then utilized for scanner or copier recalibrations in order to produce a high-fidelity document with minimal screens.

We adopt the same general frequency-domain de-screening approach to minimize the effect of the screens on the classification process. In particular, we analyze the image power spectrum to determine the existence of pertinent high-frequency energy content that represents the signature of the underlying halftone screens. The peaks corresponding to the halftone structure are quite distinguishable in the Fourier domain. Hence, the effect of the screen is significantly reduced via a repetitive “notch” frequency-domain filtering operation to yield a sufficiently smooth image for further analysis. In every iteration, the high energy content in the frequency domain is located and filtered out. This process is performed individually on each channel for a predefined number of iterations using a Butterworth notch filter⁵ of order $n=15$ given by

$$F_{\text{notch}} = \frac{1}{1 + \left[\frac{D_0^2}{D_1(u,v)D_2(u,v)} \right]^n}, \quad (1)$$

where

- D_0 =radius of filter,
- $D_1(u,v) = \sqrt{(u-M/2-u_0)^2 + (v-N/2-v_0)^2}$,
- $D_2(u,v) = \sqrt{(u-M/2+u_0)^2 + (v-N/2+v_0)^2}$.

The variables u and v are the frequency-domain coordinates. The origin of F_{notch} has been shifted to the center frequency coordinates [i.e., $F_{\text{notch}}(0,0)$ is located at $u=M/2$ and $v=N/2$]. Thus, the notch locations are symmetrically located at (u_0, v_0) and $(-u_0, -v_0)$. The radius of the notch filter is chosen, through empirical testing, as the minimum of $(M/5, N/5)$. The order is selected as $n=15$ in order to preserve the contrast of debris pixels in DCD and debris missing samples.

An illustration of this application on a DCD image is shown in Fig. 3. Figure 3(a) shows a halftone scanned image that contains a DCD. The frequency spectrum [Fig. 3(b)] shows significant energy (outlined by the boxes) at high frequencies that correspond to the halftone screens. The black spots [Figs. 3(c)–3(e)] are indications of successive frequency-domain filtering operations. The final de-screened image shown in Fig. 3(f) is thus obtained from the inverse Fourier transform of the filtered versions of all three channels.

Special instances arise where the “debris” is only visible in one channel. Hence, multiple channel processing (i.e., 3-channel processing of the image) is not desirable due to the added computational complexity and potential for inaccurate results. We therefore transform the image to a single gray-scale channel. There are numerous RGB to gray-scale methods used in the computer vision literature. Some simply employ the average of the RGB channels as a corresponding gray-scale image. A better approximation of the brightness can be derived by summing weighted versions of the R, G, and B channels.²⁴ Other methods handle the problem via de-saturation,²⁵ i.e., removal of the saturation infor-

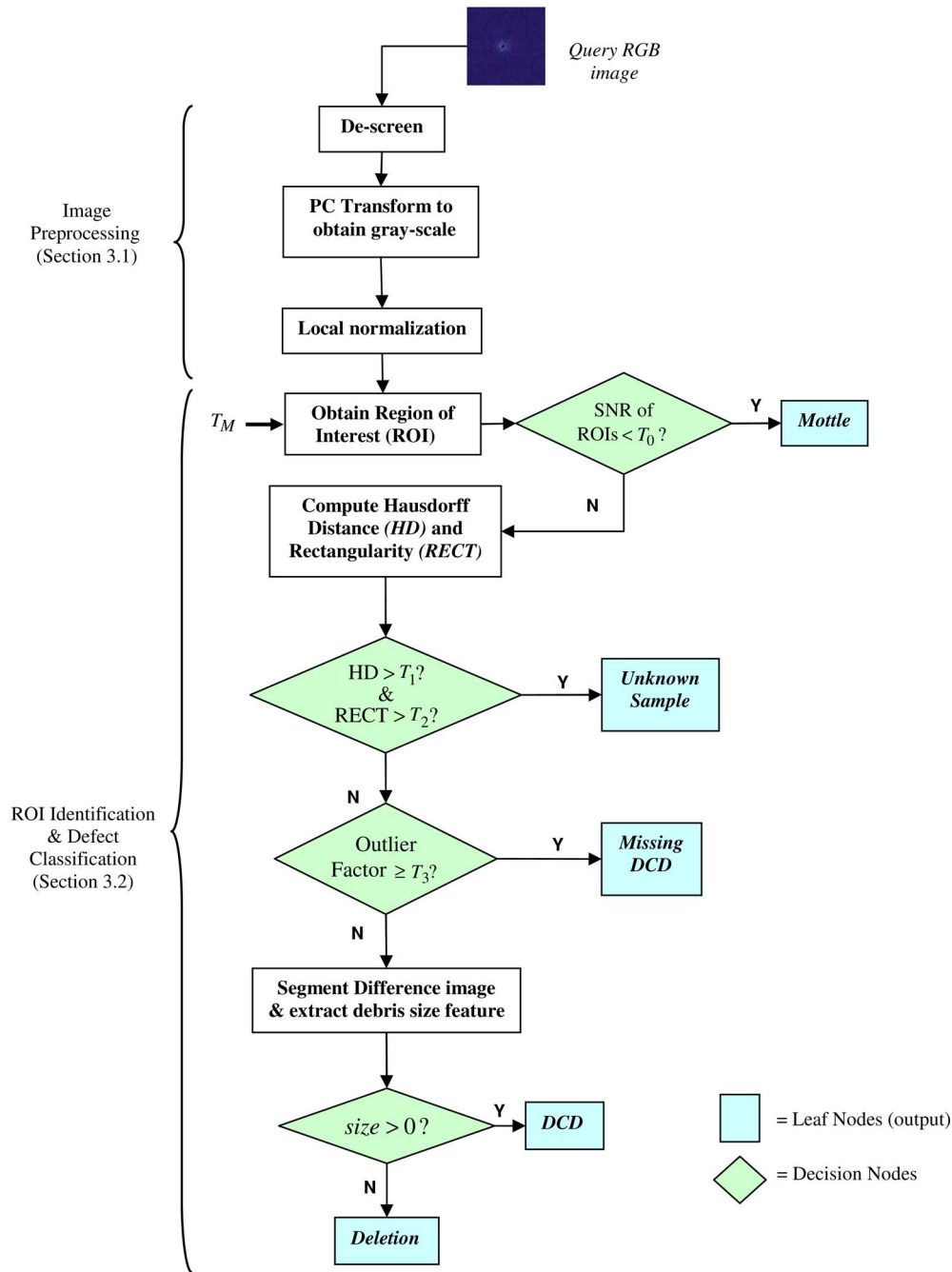


Fig. 2 Defect detection and identification framework.

mation of the image. In particular, Matlab’s²⁶ gray-scale conversion (rgb2gray) computes a gray-scale image by summing weighted versions of the R, G, and B channels using the coefficients 0.2989, 0.5870, and 0.1140, respectively. For our data set, these conventional methods often fail to preserve the saliency of the centered debris. Rather, they tend to be successful when applied to normal images (e.g., portraits, landscape scenes, road scenes, etc.) that have an appreciable contrast and exhibit a multimodal histogram property. To avoid this limitation, we perform a principal component (PC) analysis on the RGB image in order to convert it to a single channel in an optimum fashion. Figure 4 illustrates the advantages of utilizing a PC-

based conversion versus a weighted standard summation²⁶ of channels. The DCD image [Fig. 4(a)] bears “questionable” debris, which is only visible in the blue channel [Fig. 4(d)], and is more appropriately “highlighted” in the first principal component [see Fig. 4(f)] as compared to a standard RGB-to-gray conversion [see Fig. 4(e)].

Principal component analysis²⁷ is a linear data reduction approach that optimally projects a d -dimensional data set onto a lower-dimensional subspace in a mean-squared error sense. It does this by performing a coordinate rotation that aligns the transformed axes with the directions of maximum variance of the data. Let Ψ_1 , Ψ_2 , and Ψ_3 represent the

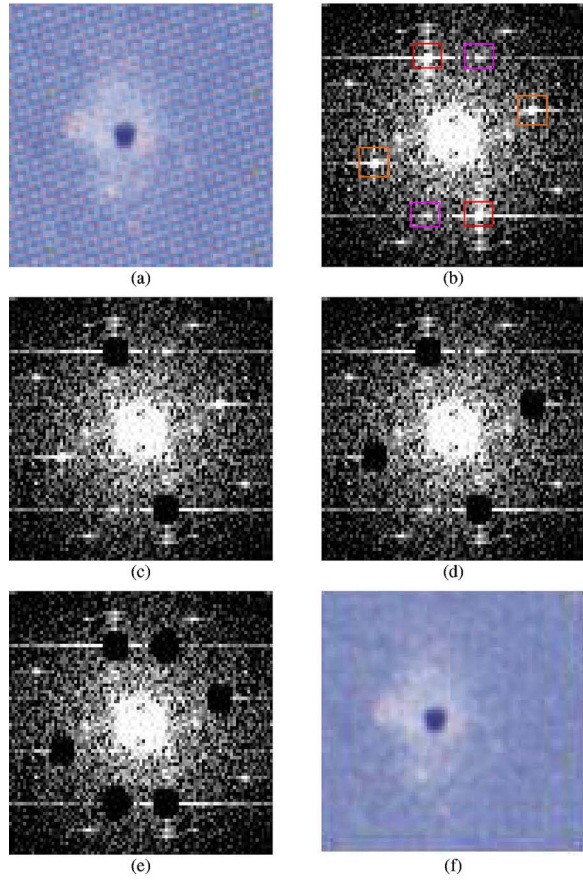


Fig. 3 De-screening process: (a) original scanned document; (b) frequency spectrum of red channel showing presence of abnormally high energy at high frequencies; (c)–(e) successive notch filtering; (f) de-screened version of Fig. 3(a).

R, G, and B channels in lexicographic ordering and $\Psi = [\Psi_1 \Psi_2 \Psi_3]^T$ be the corresponding $3 \times MN$ concatenation. The mean vector $m = [m_1 \ m_2 \ m_3]^T$ and covariance matrix C are computed as follows:

$$\sigma_{\Psi_k}^2 = E[(\Psi_k - m_k)^2] = \frac{1}{MN-1} \sum_{l=1}^{MN} (\Psi_k(l) - m_k)^2, \quad (2)$$

$$C = \begin{bmatrix} \sigma_{\Psi_1}^2 & \sigma_{\Psi_1}\sigma_{\Psi_2} & \sigma_{\Psi_1}\sigma_{\Psi_3} \\ \sigma_{\Psi_2}\sigma_{\Psi_1} & \sigma_{\Psi_2}^2 & \sigma_{\Psi_2}\sigma_{\Psi_3} \\ \sigma_{\Psi_3}\sigma_{\Psi_1} & \sigma_{\Psi_3}\sigma_{\Psi_2} & \sigma_{\Psi_3}^2 \end{bmatrix}, \quad (3)$$

where

- $\sigma_{\Psi_k}^2 = E[(\Psi_k - m_k)^2] = 1/MN-1 \sum_{l=1}^{MN} (\Psi_k(l) - m_k)^2$, and
- $\sigma_{\Psi_j}\sigma_{\Psi_k} = E[(\Psi_j - m_j)(\Psi_k - m_k)^T]$
 $= \frac{1}{MN-1} \sum_{l=1}^{MN} (\Psi_j(l) - m_j)(\Psi_k(l) - m_k).$

C is a real, symmetric matrix that can be expressed as

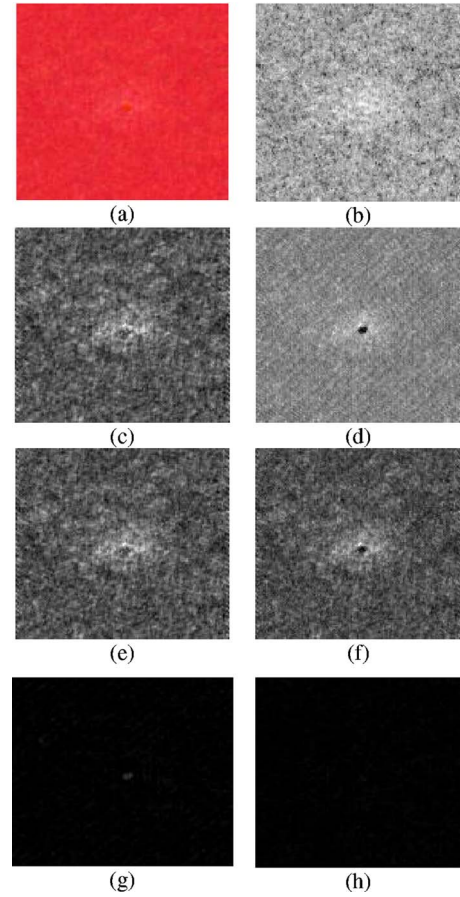


Fig. 4 Comparison of PCs with alternative gray-scale method: (a) original DCD image with "invisible" debris; (b)–(d) corresponding R, G, and B channels, respectively; (e) Matlab's rgb2gray procedure; (f) first PC; (g) second PC, (h) third PC.

$$C = U\Lambda U^T, \quad (4)$$

where U is a 3×3 orthonormal matrix of eigenvectors corresponding to the ordered eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ contained in the diagonal matrix $\Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$. The principal components of Ψ are calculated by

$$Y = U^T\Psi = [Y_1 \ Y_2 \ Y_3]^T. \quad (5)$$

Hence, the variance of the original information is distributed among the eigenvalues, with the first eigenvalue (λ_1) representing the largest variance along the corresponding PC Y_1 . The first PC is selected as the medium for classification under the hypothesis that the corresponding eigenvalues bears more than 75% of the overall variance. In the rare cases where this is not true, the channel with the highest variance is chosen. Sample re-ordered versions of Y_1 , Y_2 , and Y_3 are shown in Figs. 4(f)–4(h).

Once a high-contrast gray-scale image has been obtained, a local normalization (LN) procedure is employed to compensate for non-uniform background situations. The LN process is designed to handle large illumination variations [see Figs. 5(a) and 5(c)] characteristic of a number of samples in the database. This approach is given by

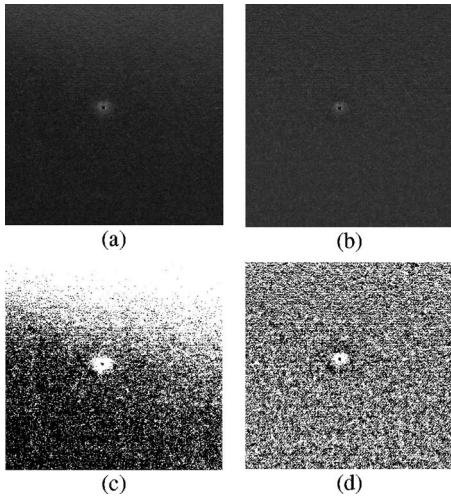


Fig. 5 Local normalization: (a) non-uniform background sample; (b) local normalized version of Fig. 5(a); (c) binarized result of Fig. 5(a); and (d) binarized result of Fig. 5(b).

$$g(i,j) = \frac{f(i,j) - m_f(i,j)}{\sigma_f(i,j)}, \quad (6)$$

where

- $f(i,j)$ is the selected/transformed gray-scale image,
- $m_f(i,j)$ is an estimation of the local mean of $f(i,j)$,
- $\sigma_f(i,j)$ is an estimation of the local standard deviation,
- $g(i,j)$ is the output image.

The above-outlined approach efficiently removes variations in the image [see Fig. 5(b)]. The block diagram in Fig. 6 depicts the implementation of the LN procedure. An estimation of the image’s local mean is obtained by filtering with a 70×70 spatial Gaussian low-pass filter, $h_1(i,j)$ of standard deviation, $\sigma_1=21$. The $[\cdot]^2$ and $\sqrt{[\cdot]}$ symbols (see Fig. 6) represent the “square” and “square root” operations, respectively, and are used to complete the computation of the standard deviation. The second smoothing filter, $h_2(i,j)$ is equivalent to $h_1(i,j)$.

3.2 ROI Identification and Defect Classification

3.2.1 ROI identification

We process $g(i,j)$ by utilizing a median-based thresholding approach to select the ROIs. Other thresholding schemes tend to immediately separate the debris without providing any deletion boundary. The median threshold was observed to be more robust to noise and outliers within the image

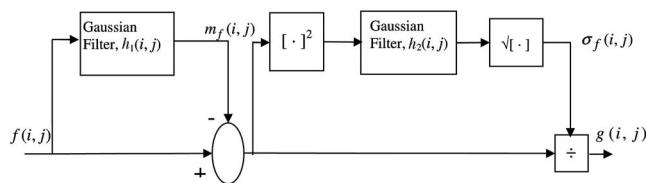


Fig. 6 Local normalization block diagram.

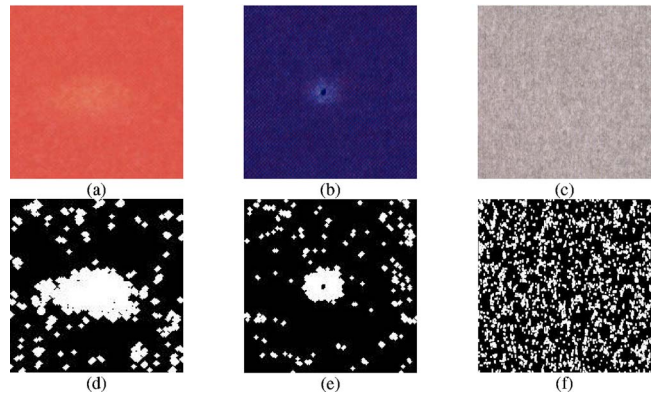


Fig. 7 ROI feature extraction for (a)–(c) deletion, DCD, and mottle images, respectively; (d)–(f) respective morphologically opened images.

when compared to a mean-based thresholding. If the median is given by T_M , then the binary image $b(i,j)$ is computed as

$$b(i,j) = \begin{cases} 1, & \text{if } g(i,j) \geq T_M, \\ 0, & \text{if } g(i,j) \leq T_M. \end{cases} \quad (7)$$

The images in Figs. 5(c) and 5(d) were obtained by utilizing Eq. (7).

An *opening* morphology operation²⁸ is then employed to remove noisy objects from the binary image, $b(i,j)$. Figure 7 depicts a comparison of typical morphologically “opened” images obtained from a deletion, DCD, and mottle images, respectively. The *opening* morphological operations yield a binary image, where the pixels that differ from the specified shape and size of the 2×2 structural element are assigned the value of 1 and are displayed as “white.” The corresponding largest group of connected white pixels is selected as the ROI. The relationship between the largest region and its neighbors is employed as a discriminatory feature. Note that the size of the largest region is not necessarily equivalent to the area of the defect but is utilized to provide its corresponding location for classification purposes. Specifically, the ratio of the root mean square (RMS) of the largest ROI to the RMS of other ROIs is employed as the feature of interest. This feature is thereafter compared to an empirically determined threshold T_0 to yield class decisions (at the first decision node) between mottle and deletion-type (or non-mottle) images. Figure 8 shows typical distributions of the number of pixels in each region for a mottle type and a non-mottle type. If the signal of interest is given by the major ROI, then we expect a large signal-to-noise ratio (SNR) for the non-mottle image and a low SNR for a mottle image.

In order to ensure that only defects bearing a deletion-type signature are considered for further processing, we impose a shape contour test by utilizing the Hausdorff distance.²⁹ In particular, the contour of the major ROI is compared to a fitted ellipse’s contour, which is created using a similar approach to that found in Saber *et al.*³⁰ Additionally, the rectangularity³¹ of the major ROI measured by the ratio of the area of the region to the area of its minimum bounding rectangle (MBR) is employed as another feature.

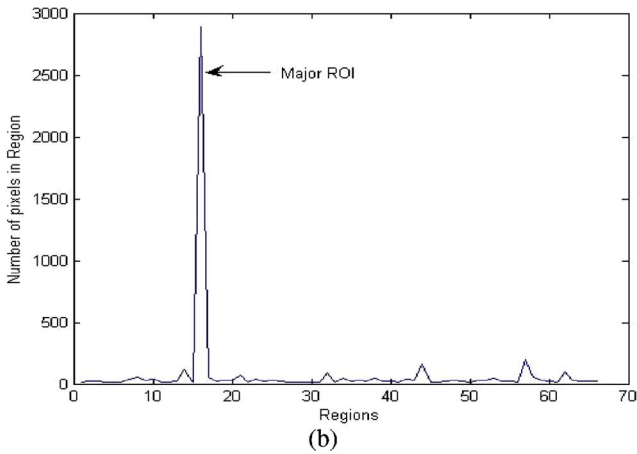
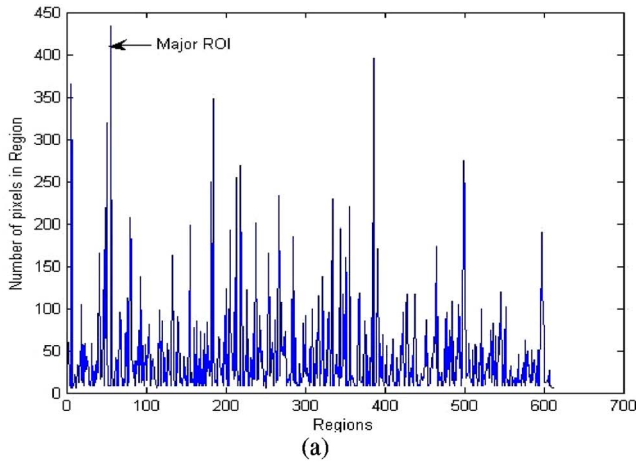


Fig. 8 Study of SNR of (a) typical mottled image; (b) typical non-mottle type.

These two features (i.e., Hausdorff distance and rectangularity) ensure that only defects exhibiting an elliptical shape configuration are regarded for further processing. They are compared with the thresholds T_1 and T_2 to separate unknown defects from deletion-type samples. This completes the second decision node.

3.2.2 Outlier processing (deletion vs. DCD)

The first step toward processing only deletion-type images utilizes the convex hull of the major ROI as a mask [Fig. 9(a)] for the replacement of potential missing debris pixels. To obtain these pixels, a threshold is automatically computed and applied to the masked region. Let p represent a given gray level. This threshold is obtained by first generating the histogram $h(p)$ for $p_{min} \leq p \leq p_{max}$, as shown in Fig. 9(b). Then, the set containing possible bright outlier pixels within the mask is defined by

$$\mathfrak{R}^1 = \{p: h(p) > 0 \text{ and } p > p_{min} + 0.75^*(p_{max} - p_{min})\}, \quad (8)$$

where the constraint $p > p_{min} + 0.75^*(p_{max} - p_{min})$ is designed to limit the desired threshold to only 25% of the bright pixel values. Let \mathfrak{R}^D be a new set of elements, which are backward differences between adjacent elements of \mathfrak{R}^1 .

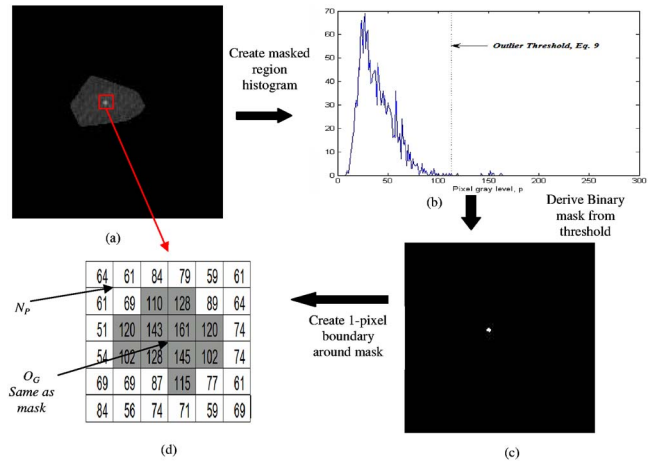


Fig. 9 Missing debris identification procedure.

The desired threshold can thus be obtained as

$$\mathfrak{R}^T = \{\mathfrak{R}^1: \mathfrak{R}^D > \theta\}, \quad \text{Outlier threshold} = \min\{\mathfrak{R}^T\}, \quad (9)$$

where $\theta = 5$ (i.e., minimum number of pixels) for our purpose. To avoid conflicts (i.e., two or more bright outlier regions), the largest region is simply chosen. An outlier factor (OF) (i.e., a measurement of how significantly a group of pixels deviates from other neighboring pixels) is compared with a threshold T_3 as

$$OF \geq T_3 \quad \text{and} \quad OF = \text{median}\{O_G\} - \text{median}\{N_p\}, \quad (10)$$

where O_G = potential outlier group and N_p = neighborhood pixels contained in a bounding box around O_G . This helps to quantify the outlier measure of the region in question in order to determine its significance to its neighbors. Figure 9 shows a successful identification of the missing pixels, where the O_G region [Figs. 9(c) and 9(d)] has been obtained via thresholding with the desired outlier threshold. The N_p region [Fig. 9(d)] is created from a 1-pixel boundary around O_G .

To efficiently differentiate between deletion and DCD defects, we exploit the major difference—the presence of a group of dark pixels in an approximate center of the elliptical region. This is a difficult task, as demonstrated in our previous procedure,³² wherein information from the histogram was utilized to devise a threshold that localizes possible debris pixels. Due to the noisy nature of deletion samples, additional steps involving the acceptance or rejection of segmented dark pixels are needed. For the low-contrast DCD images, the debris pixels are not successfully identified after thresholding. Figure 10 shows intermediate results obtained by applying several thresholding mechanisms to pixels within the created mask [Fig. 10(b)] for a DCD sample. The best result [Fig. 10(e)] is achieved by a hole-emphasizing routine, which utilizes morphological reconstruction to “fill holes” in the image and thereafter compute a difference image. This was shown to provide superior results when compared with standard thresholding techniques such as Otsu’s method,³³ as illustrated in Figs. 10 and 11. Otsu’s method [Fig. 10(c)] does not provide a desirable outcome, while the modified Otsu’s approach^{17,33} [Fig. 10(d)] is close to the ground-truth result but is noisier

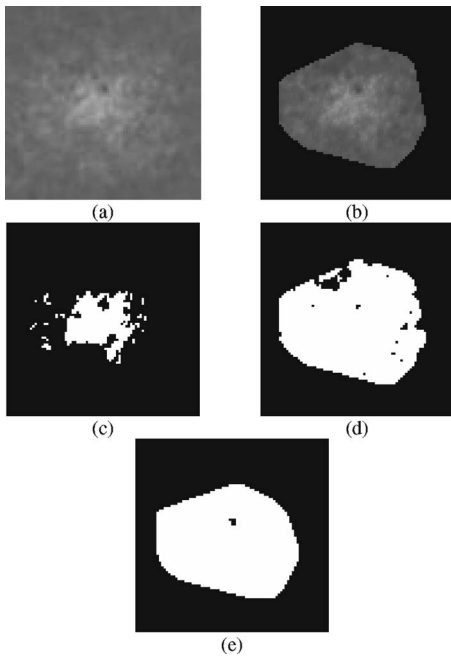


Fig. 10 Within-mask segmentation: (a) original image; (b) masked image; (c) Otsu's thresholding; (d) modified Otsu's thresholding;^{17,24} (e) morphological reconstruction approach.

compared to the morphological reconstruction approach. Figure 11 illustrates a similar situation with a noisy deletion image.

The “fill-hole” process is a special morphological transformation²⁸ (called a geodesic transformation) that accepts two images—a marker and a mask. The mask is used to restrict the growth (or decay) of the marker image during

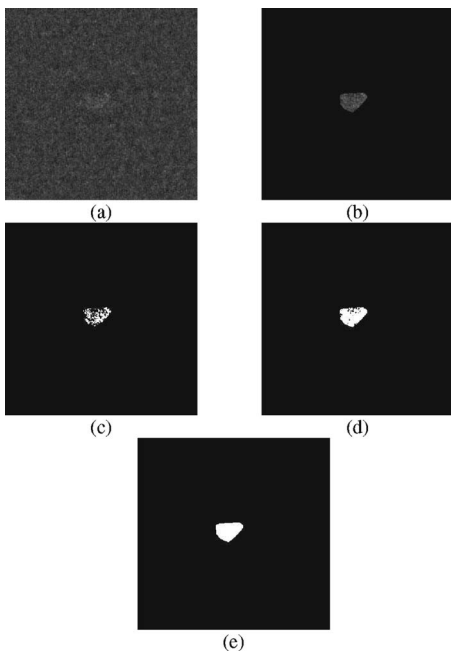


Fig. 11 Within-mask segmentation for noisy image: (a) original image; (b) masked image; (c) Otsu's thresholding; (d) modified Otsu's thresholding; (e) morphological reconstruction approach.

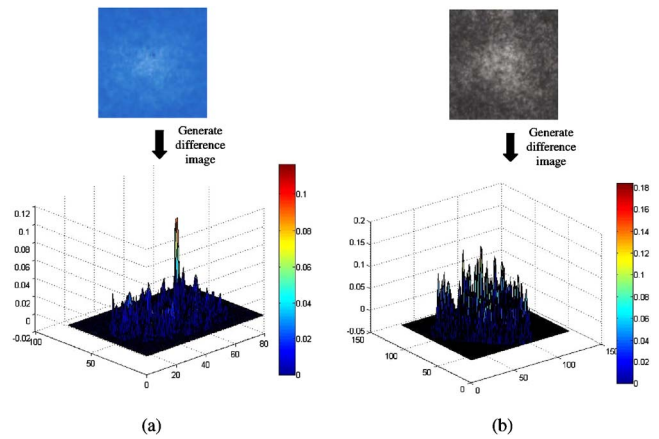


Fig. 12 Difference images: (a) DCD, SNR=22 dB; (b) deletion, SNR=15 dB.

regular morphological operations. For this process, the marker is set to the maximum value of the image except along its border, where the values of the original image are kept, while the mask is represented by the image itself. An erosion of the marker is then iteratively performed until a stable result is achieved. The final eroded marker constitutes the *filled image*, and the holes can be obtained by subtracting the input image from the filled one (see Ref. 28 for a detailed explanation). This morphological reconstruction ensures only prominent dark pixels are emphasized, as shown in Fig. 12. The gray-scale SNR of the difference image is used to quantify the disparity between DCD and deletion samples. If the peak points denote the debris of interest, then the DCD are expected to possess a higher SNR level when compared to the deletion. A low SNR is an indication that no further processing is required and therefore that no prominent outliers exists.

The number of pixels representing the debris is utilized as the discriminating feature between deletion and DCD samples. The Mahalanobis distance,⁸ in a normalized range [0,100] of the segmented debris from the center of the mask [i.e., convex hull of major ROI as in Fig. 9(a)], is employed for classification between DCDs and deletions. This results in a final classification of the artifact into one of five possible categories.

4 Experimental Results

We tested the performance of our proposed algorithm on a database of 276 images of artifacts (see Fig. 1 for a representative sample) that consists (1) 264 scanned images of RGB format provided by Xerox Corporation with accompanying ground-truth labels provided by a QAP, and (2) 12 non-defective and synthetic images (Fig. 13) comprising logos and “rectangular” shape defects. The scanned images were comprised of 68 DCDs, 23 debris-missing DCDs, 80 deletions, and 93 mottle images (see Table 1). The synthetic images were introduced to test the algorithm’s robustness and to ensure that only regions satisfying the “elliptical” ROI property are processed as deletion-type defects. The region of interest portraying the artifacts is scanned at 600 dots per inch (dpi) using a flat-bed scanner typical to the one available at any print shop to ensure high quality and sufficient pixel detail. The thresholds utilized in our classi-

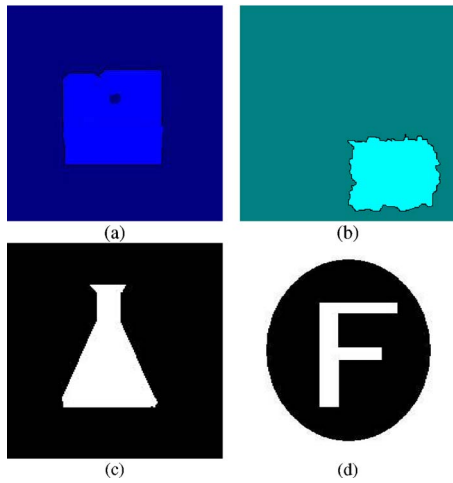


Fig. 13 Synthetic samples: (a) DCD; (b) deletion; (c) and (d) logos.

fication were $T_1=25$, $T_2=0.8$, and $T_3=0.158$. These were obtained by empirical testing. Of the 23 debris missing DCD images, a few samples exist that are very challenging, due to their manifestation, and as such can easily be misclassified as deletions by an operator or QAP.

The performance of our proposed algorithm is documented in Table 1. The classification accuracy of the DCD class is 86.76%. The misclassified samples are a result of low-contrast images [see Figs. 14(a) and 14(b)], where the bead is barely visible or missing entirely, resulting in a low signal-to-noise ratio. Given the lack of contrast, it is entirely possible that these samples would also be misclassified by operators and/or QAPs in a production-type environment.

Deletion classification yields 97.5% accuracy. Misclassified samples appear to possess a potential missing debris group of pixels [Figs. 14(c) and 14(d)] and thereby are classified as debris-missing DCDs by our proposed algorithm. Once again, this misclassification can be attributed to potential confusion from the ground-truth information. The results for the remaining categories are also shown in Table 1, where the mottle and arbitrary images were classified with 99% and 100% accuracy, respectively. The total cor-

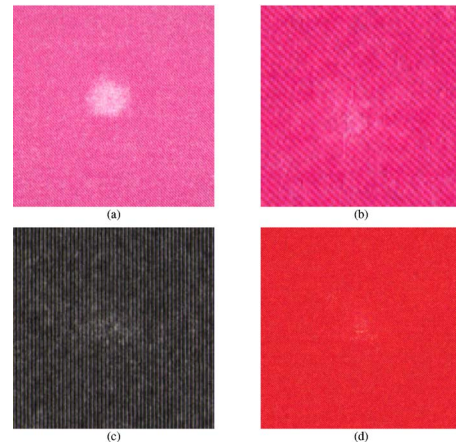


Fig. 14 Misclassified samples: (a) DCD with low contrast; (b) DCD with an invisible bead; (c) and (d) deletions with possible debris-missing regions.

rect classification rate is 94.95%, which is obtained from a weighted average of the given individual accuracies.

The effectiveness of the principal component transformation for gray-scale conversion is also demonstrated. Table 2 shows the resultant classification obtained using a standard RGB-to-gray conversion, similar to the one found in Matlab²⁶ (see `rgb2gray` function). It can be seen from Table 2 that the Matlab `rgb2gray` function results in lower classification accuracy especially for the case of debris-missing DCDs.

The robustness of the algorithm against varying degrees of random Gaussian noise was tested. First, the noise present in the normal images was quantified by selecting random samples and extracting arbitrary-sized cropped regions from various locations in the image. We found the images to have an intrinsic noise level with standard deviation $\sigma \approx 0.04$; thus, any additional noise tends to further deplete the contrast. Repeated tests with white Gaussian noise ($\sigma=0.01, 0.02, \text{ and } 0.05$) are summarized in Fig. 15, from which it can easily be seen that the classification accuracy is inversely proportional to the level of noise added, due to a significant drop in contrast. Mottle tends to hold up

Table 1 Classification results using our proposed method.

True Class	Number of Images	Classification results					Accuracy DCD
		DCD	Debris Missing DCD	Deletion	Mottle	Unknown	
DCD	68	59	3	5	1		86.76%
Debris-missing DCD	23	2	21				91.3%
Deletion	80			78			97.5%
Mottle	93			1	92		99%
Other	12					12	100%

Table 2 Classification results using Matlab's rgb2gray approach.

True Class	Number of Images	Classification Results					Accuracy
		DCD	Debris Missing DCD	Deletion	Mottle	Unknown	
DCD	68	51	1	13	3		75%
Debris missing DCD	23	1	3	18	1		13%
Deletion	80	3		77			96.25%
Mottle	93				93		100%
Other	12					12	100%

well against the noise due to its inherently noisy nature. The misclassified deletion images are classified as Mottle and tend to worsen with increasing noise. As the level of noise is increased from $\sigma=0.01$, 0.02, and 0.05, the total classification (average of all the classifications) is reduced from 93.51%, 86.22%, and 78.68%. Since the acquisition of the images will normally take place under non-noisy conditions, this does not represent a major limitation of the algorithm.

5 Conclusions

In this paper, an algorithm for automatically classifying a specific set of image-quality defects in printed documents is proposed. The algorithm accepts scanned versions of defected printed media, where the defect has been localized a priori by a customer service engineer, and provides accurate classifications of defect type (deletion, DCD, debris-missing DCD, or mottle). Due to large variations between elements of the same class, several preprocessing techniques were carried out on each image to attain some level of uniformity among the samples. Using a custom tree classifier, binary decisions were made by employing simple shape and size constraints at each node. The use of principal component analysis to obtain a gray-scale image preserves the contrast of the original RGB sample. Additionally, the use of local normalization procedures helps to

avoid misclassifications by making the background illumination uniform wherever possible. An accuracy of 94.95% was still attained despite the noisy nature and low contrast of several samples. However, this accuracy tends to depreciate with increasing noise levels. Since this procedure has proved to be quite successful, the next step involves an automation of the defect localization process by possibly incorporating the original electronic document along with an online scan of the printed output to help in automatically localizing and classifying the artifact.

Acknowledgments

This research was funded in part by the Center for Electronic Imaging Systems (CEIS), a NYSTAR-designated Center for Advanced Technology, Xerox Corporation, and the Electrical Engineering Department at the Rochester Institute of Technology. The authors wish to thank Dr. Daniel Phillips and Dr. Sohail Dianat for their helpful comments.

References

1. Y. Rui, T. S. Huang, and S. Chang, "Image retrieval: Current techniques, promising directions, and open issues," *J. Visual Commun. Image Represent* **10**(1), 39–62 (1999).
2. A. M. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(12), 1349–1380 (2000).
3. H. Burkhardt and S. Siggelkow, "Invariant features for discriminating between equivalence classes." In *Nonlinear Model-based Image/Video Processing and Analysis*, John Wiley and Sons, New York (2000).
4. M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inf. Theory* **IT-8**, 179–187 (1962).
5. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ (2002).
6. B. M. Mehre, M. Kankanhalli, and W. F. Lee, "Shape measures for content based image retrieval: A comparison," *Inform. Process. & Manage.* **33**(3), (1997).
7. Y. Rui, T. S. Huang, and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," *Proc. of IEEE Int. Conf. Image Process.*, Vol. 2, pp. 815–818, (1997).
8. W. Y. Ma and B. S. Manjunath, "Netra: A toolbox for navigating large image databases," *Multimedia Syst.* **7**(3), 184–198 (1999).
9. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed., John Wiley and Sons, New York (2001).
10. N. Vasconcelos, "On the complexity of probabilistic image retrieval," *IEEE Int. Conf. Comp. Vis. Patt. Recog.*, Vol. 2, pp. 400–407 (2001).
11. N. Beckmann, H. Kriegel, R. Schneider, and B. Seeger, "The R²-tree: An efficient robust access method for points and rectangles," *ACM SIGMOD Int. Conf. Manage. of Data*, Atlantic City, NJ (1990).
12. H. Samet, "The quadtree and related hierarchical data structures,"

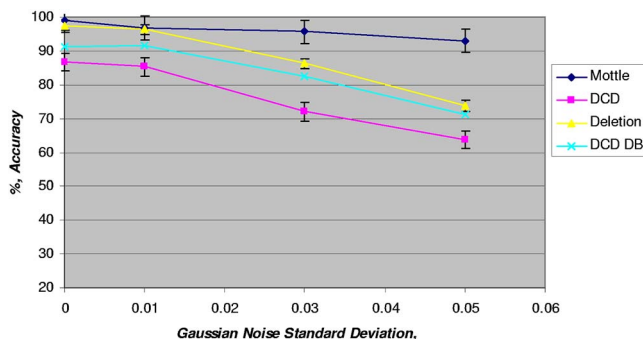


Fig. 15 Addition of increasing levels of Gaussian noise and resulting performance.

- ACM Comput. Surv. **16**(2), 187–260 (1984).
13. H. J. Zhang and D. Zhong, "A scheme for visual feature-based image indexing," *Proc. of SPIE Conf. on Storage and Retrieval for Image and Video Databases III*, pp. 36–46, San Jose, CA (1995).
 14. J. Iivarinen and A. Visa, "An adaptive texture and shape-based defect classification," *Proc. of the 14th Int. Conf. Patt. Rec.*, Vol. 1, pp. 117–122 (1998).
 15. Y. Li and T. W. Liao, "Weld defect detection based on Gaussian curve," *IEEE Proc. 28th Southeastern Symp.*, pp. 227–231 (1996).
 16. D. Mery and M. A. Berti, "Automatic detection of welding defects using texture features," *Insight* **45**(10), 676–681 (2003).
 17. H. Ng, "Automatic thresholding for defect detection," *Proc. 3rd Int. Conf. Image and Graphics*, pp. 532–535 (2004).
 18. J. C. Briggs, D. J. Forrest, A. H. Klein, and M. Tse, "Living with ISO-13660: Pleasures and perils," *Proc. IS&T's NIP15, Int. Conf. Digital Printing Tech.*, Orlando, FL (1999).
 19. ISO/IEC DIS 13660 Draft International Standard, "Office equipment—Measurement of image quality attributes for hardcopy output—Binary monochrome text and graphic images," *Int. Org. for Standardization, ISO/IEC JTC1 SC28* (1996).
 20. D. R. Rasmussen, K. D. Donohue, Y. S. Ng, W. C. Kress, F. Gaykema, and S. Zoltner, "ISO 19751 macro-uniformity," *Image Quality and System Performance III*, Vol. 6059 (2006).
 21. S. P. Farnand, E. N. Dalal, and Y. S. Ng, "Recent progress in the development of ISO 19751," *Image Quality and System Performance III*, Vol. 6059 (2006).
 22. D. F. Dunn and N. E. Mathew, "Extracting color halftones from printed documents using texture analysis," *Proc. Int. Conf. Image Process.*, Vol. 1, pp. 787–790 (1997).
 23. G. Sharma, "Methods and apparatus for identifying marking process and modifying image data based on image spatial characteristics," U.S. Patent #6353675B1 (2002).
 24. W. K. Pratt, *Digital Image Processing*, John Wiley & Sons, New York (1991).
 25. C. Bunks, *Grokking the GIMP*, 1st ed., New Riders Publishing, New York (2000).
 26. R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*, 1st ed., Prentice Hall, Englewood Cliffs, NJ (2003).
 27. J. E. Jackson, *A User's Guide to Principal Components*, 1st ed., John Wiley & Sons, New York (1991).
 28. P. Soille, *Morphological Image Analysis: Principles and Applications*, Springer-Verlag TELOS, New York (1999).
 29. D. Huttenlocher, G. Klanderman, and W. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(9), 850–863, (1993).
 30. E. Saber and A. M. Tekalp, "Frontal-view face detection and facial feature extraction using color, shape and symmetry based cost functions," *Pattern Recogn. Lett.* **19**, 669–680, (1998).
 31. P. L. Rosin, "Measuring shape: Ellipticity, rectangularity, and triangularity," *Proc. 15th Int. Conf. on Patt. Recog.* (2000).
 32. O. A. Ugbeme, E. Saber, and W. Wu, "An automated defect classification algorithm for printed documents," *Proc. International Cong. Imaging Science*, pp. 317–320 (2006).
 33. N. Otsu, "A threshold selection method from gray-level histogram," *IEEE Trans. Syst. Man Cybern.* **SMC-8**, 62–66 (1978).



Onome Augustine Ugbeme received his MS degree in electrical engineering from the Rochester Institute of Technology (RIT), Rochester, NY, in 2006 with a focus in image and video processing. He received his BS degree in electrical engineering from the University of New Orleans, Louisiana, in 2004. He is currently an audio/video software engineer at Advanced Micro Devices (AMD) Digital TV unit. He is a member of the IEEE, the National Society of Black Engineers (NSBE), and the Tau Beta Pi and Eta Kappa Nu Honor Societies.

He is a member of the IEEE, the National Society of Black Engineers (NSBE), and the Tau Beta Pi and Eta Kappa Nu Honor Societies.



Eli Saber is an associate professor in the Electrical Engineering Department at the Rochester Institute of Technology (RIT). Prior to that, he worked for Xerox Corporation from 1988 until 2004 in a variety of positions, ending as product development scientist and manager at the Business Group Operations Unit. He received his BS degree in electrical and computer Engineering from the University of Buffalo in 1988 and his MS and PhD degrees in the same discipline from the University of Rochester in 1992 and 1996, respectively. From 1997 until 2004, he was an adjunct faculty member at the Electrical Engineering Department of RIT and at the Electrical & Computer Engineering Department of the University of Rochester. His research interests are in the areas of digital image and video processing, including video segmentation, object tracking, content-based video analysis and summarization, multicamera surveillance video processing, and image understanding. He is an associate editor for the *IEEE Transactions on Image Processing*, the *Journal of Electronic Imaging*, the *IEEE Signal Processing Magazine Applications Column*, and guest editor for the special issue on color image processing for the *IEEE Signal Processing Magazine*. He is currently serving as the chair of the IEEE Technical Committee (TC) on Industry DSP Technology and is a member of the Image and Multi-dimensional Digital Signal Processing TC.



Wencheng Wu received his PhD degree in electrical engineering from Purdue University, Indiana, in 2000. He is currently a member of Wilson Center for Research and Technology at Xerox Corporation working on the development of print quality metrics, image simulation, and modeling for print quality evaluation, etc. He is a member of the Institute of Electrical and Electronic Engineers (IEEE), the Phi Tau Phi Scholastic Honor Society, and the Imaging Science and Technology Society. He is also the current section chair of IEEE Signal Processing Society, Rochester, NY.



Kartheek Chandu is currently pursuing his MS degree in electrical engineering at the Rochester Institute of Technology (RIT), Rochester, NY. He obtained his BS degree in electronics and communication engineering from the Nagarjuna University, A.P., India, in 2005. He is a member of the IEEE. At the Rochester Institute of Technology, he is a research assistant working on online defect analysis.