Surface inspection of web materials using the self-organizing map

Jukka Iivarinen[a] and Juhani Rauhamaa[b]

[a]Helsinki University of Technology, Laboratory of Computer and Information Science
P.O. Box 2200, FIN-02015 HUT, Finland

[b]ABB Industry Oy, Pulp & Paper Division, P.O. Box 94, FIN-00381 Helsinki, Finland

ABSTRACT

A surface inspection problem is divided into three parts, into an image acquisition part, into a defect detection part that is suitable for hardware implementation, and into a defect classification part that is done in a user's terminal. In the defect detection part extraction of texture features is done and potential defect areas are marked. The proposed scheme is taught only with examples of fault-free surface. In the defect classification part features describing the shape and internal structure of defects are extracted and defects are classified into different defect classes. Examples of defects are used to train the classification system. Use of the self-organizing map (SOM) in defect detection and in defect classification makes the proposed method adaptable to different types of surfaces and to different types of defects. Only reselection of features may be necessary to cope with different surface and defect characteristics. The results of experiments with base paper samples are encouraging.

Keywords: surface inspection, defect detection, defect classification, self-organizing map

1. INTRODUCTION

Surface inspection is important in many maintenance tasks but also in many industrial production processes, for example, in timber grading, in steel manufacturing, and in paper and textile industries. Here we are mainly interested in inspection of web materials. Web materials refer to materials that are produced in the form of continuous rolls. Metals, papers, plastics, and textiles are examples of web materials.

The purpose of a web inspection system is to detect and classify those defects which impair product quality with regard to the requirements set by the user. The requirements mostly deal with the suitability of the product to the intended use of it. In the worst case defects may make a product functionally deficient or even unusable. Critical defects are also those which cause production disturbances, e.g web breaks. The most common defects of web products are holes and different types of light or dark spots. In addition, in different manufacturing lines there appear defects which are specific to those processes. Accordingly, inspection systems designed for different processes have properties which are common to all of them, but they also contain diverging solutions for process specific defect types.

In web surface inspection real-time processing is typically needed. Some basic requirements for an inspection system are: 100% inspection coverage, resolution even less than 1 mm^2, capability to handle web speeds up to 30 m/s. In order to fulfill these requirements, inspection systems have been mainly based on gray level thresholding. Recently, more sophisticated methods have been proposed. Use of texture and color information and more complicated classifiers has been possible due to increased calculation capability of computers and specialized hardware. Some commercial solutions for surface inspection can be found in Refs. 1–4. A review on texture defect detection was given in Ref. 5, and a survey of automated visual inspection in Ref. 6. The complete defect classification system is described in a recent paper by Brzakovic and Vujovic.[7] Other studies on surface inspection can be found in Refs. 8–17.

Further author information–
J.I.(correspondence): E-mail: Jukka.Iivarinen@hut.fi
J.R.: E-mail: Juhani.Rauhamaa@fidri.mail.abb.com

2. THE PROPOSED SYSTEM MODEL FOR WEB INSPECTION

The proposed system model for web inspection has three phases: an image acquisition phase, a defect detection phase, and a defect classification phase (Figure 1). Defect detection is done in a specialized hardware since it is computing intensive. More time can be spend on classifying defects, because defects are usually rare. Defect classification is done on a user's terminal.

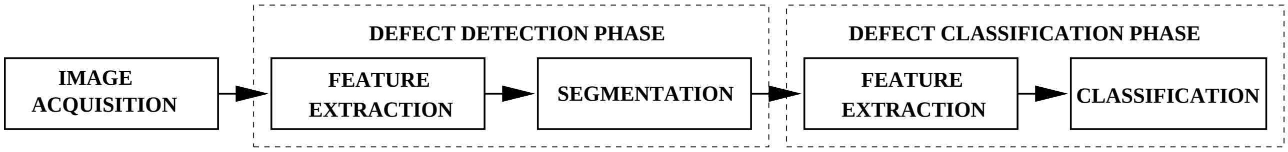

Figure 1. The proposed system for defect detection and classification consists of three phases.

The self-organizing maps (SOMs)[18] are used both in defect detection and in defect classification. The SOM is an unsupervised, self-organizing neural network. In defect detection a modified SOM, called the statistical SOM,[14,19] is used to separate defects from a fault-free web. It is taught only with samples taken from a fault-free web. This makes the SOM easily adaptable to different surfaces. In defect classification the SOM is used to cluster unknown defects. It thus finds the classes (clusters) that are inherent to defect samples. These classes are then given an explanation (or a label) by hand.

In the following the requirements and solutions for web inspection are discussed from the viewpoint of paper making. However, most of the problems and the proposed solutions are common to other web materials as well.

2.1. Image Acquisition

Most web defects are optically discernible. Therefore, today the detection principle of modern web inspection systems is based on the use of optical techniques. Especially linear CCD-array cameras are common as the sensors of these systems. As an example, a typical paper inspection system consists of a camera beam, a light source beam, defect information processing electronics, and an operator station (Figure 2).

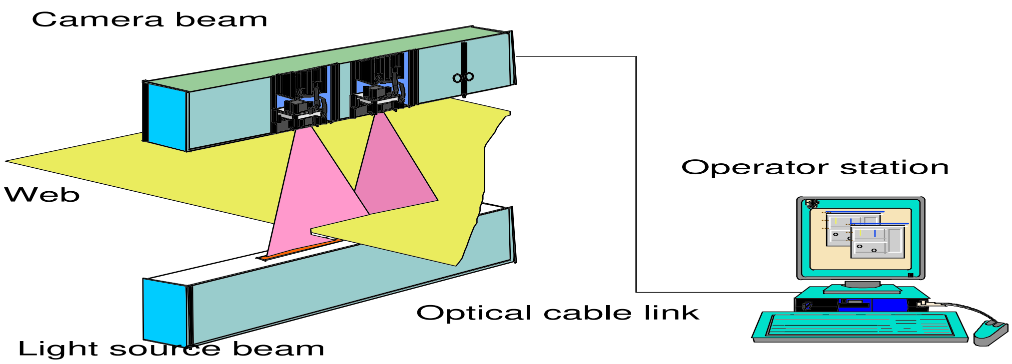

Figure 2. A typical paper inspection system.

A specific property of paper manufacturing is the high production rate. Modern paper machines are wide and fast. The speed of a paper machine can be even 30 m/s. At the same time the width of a paper may exceed 9 meters. This means that about 300 million square millimeters of paper is produced each second, and the same amount of production has to be inspected with 100% coverage. The problem is made even more difficult by the fact that very often the size of the smallest defects which have to be detected is smaller than a square millimeter. This gives rise to use of many parallel CCD-cameras so that the whole paper width in cross direction of the web is covered and the resolution requirement is met. Standard off-the-shelf cameras are commonly used, but also special digital cameras containing defect processing capabilities have been designed for this purpose.[20] Also the pixel (the light sensitive cell of a CCD array) readout frequency should be as high as possible to diminish the effect of the so-called pixel stretch to the imaging resolution in the machine direction.

A typical pixel clock frequency is 20 MHz. When using a 1024 pixel sensor, a 20 kHz scan frequency is achieved. This results in a pixel stretch of 1.5 mm at the web speed of 30 m/s. For shorter scan times, either sensors with smaller number of pixels or the so-called multitapped sensors have to be used. A shorter scan time means shorter exposure. How short this can be done, depends ao. on the amount of available light. These alternatives, however, make the system more complicated and raise its cost.

The information flow available from each camera is 20 Mbytes or more depending on the resolution of the A/D-conversion. A typical paper inspection application may require tens of cameras. Therefore, it is obvious that a distribution of the digital processing is necessary. A solution is the above-mentioned digital camera designed specifically for the web inspection. Such a camera contains two main parts: a CCD-sensor module and an electronic processing unit. The analog video signal available from the sensor is not ideal one and it needs some compensations. Typically a removal of a DC-offset, a linearization of the videocurve and an AGC of the signal levels are necessary. After the A/D-conversion the digital signal processing methods can be used.

The light source beam is needed to achieve sufficiently strong and even illumination on the paper surface to be inspected. The type of the light source varies according to the specific needs of the paper grade. E.g. wide spectrum halogen lamps and fluorescent tubes are typically used. The need for ripple-free illumination calls for special arrangements in lamp power supply units. Depending on the paper grade, either transmitting or reflecting illumination is used. Transmitting illumination is a typical solution for papers with basis weight less than 200 g/m^2. Reflecting illumination is more suitable for heavy or coated papers with high opacity.

2.2. Defect Detection

Paper manufacturers have designed hundreds of paper grades for different purposes. Variation of paper grade properties makes defect inspection more difficult. Most of the paper belongs to newsprint, printing and writing, and packaging grades. When considering paper inspection, important properties are the paper grammage, and the optical properties, such as opacity and gloss. The lightest paper grades weigh less than 20 g/m^2. These light grades have often a low opacity value, i.e., they are fairly transparent, but e.g. for printing grades, high opacity values are desirable. In the other end of the grammage scale there are paper boards whose grammage is 300 g/m^2 or more. Paper can be coated, supercalendered, or processed otherwise to make reproducing of text and color pictures better. These treating methods aim at changes in optical properties of paper and have accordingly effects on inspection. The optical properties are not constant along the paper surface. E.g. paper formation causes variations to it. Therefore, when imaging normal defectless paper, the signal received is not even. On some paper grades this variation may be considerable, making the detection of defects even more difficult.

The proposed defect detection procedure can be shortly stated as follows (Figure 3). Scan an image with a small gliding window. For each window area repeat the following steps:

1. **Feature Extraction**
   Calculate a set of texture features for a given image window. Collect these features into a feature vector **f**.
2. **Segmentation**
   Compare **f** with the SOM and find the best-matching unit c. Show the minimum error e_c as the output. If e_c is larger than or equal to the specified limit T, the pixel is classified as a defect. Otherwise the pixel is fault-free.

The limit T is given to the process in advance. It depends on the features and on the desired accuracy of segmentation. It is used to determine on how many features the given feature vector **f** may differ from the best-matching unit.

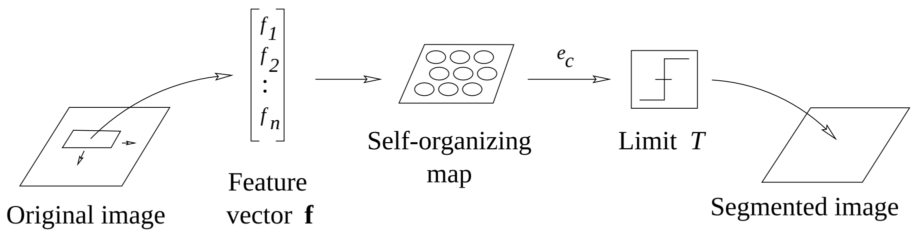

Figure 3. The proposed defect detection procedure.

2.2.1. Features for defect detection

Images of web surfaces can often be considered as stochastic textures, hence the co-occurrence matrices,[21] also known as the spatial gray level dependence matrices, are used for the texture description. The co-occurrence matrix is a second order statistical measure of gray level variation. It indicates the joint probability of gray level occurrence of two pixels at a certain displacement in an image. The co-occurrence matrix method has been found to be good in classifying stochastic textures.[22,23] The co-occurrence matrix is reduced to a set of texture features to make calculation time and memory requirements smaller.[21] These features are energy, contrast, and mean.

2.2.2. Segmentation procedure

The main idea is to train a classifier with fault-free samples only. Defect detection is then based on the following idea: an unknown sample is classified as a defect if it differs enough from the estimated distribution of texture features extracted from fault-free paper samples. The self-organizing map (SOM)[18] is used to estimate this distribution. For the complete description of the segmentation scheme see Ref. 14,19. In Fig. 4 there are some examples of paper images and their segmentations.

2.3. Defect Classification

Due to the 2-dimensional nature of paper, most of its defects can be regarded as optical surface flaws, which are detectable by a human eye. Therefore, classification is commonly based on their visual appearance. When defining defect classes, not only the defect appearance and its effects on a final product are considered, but also those features which reveal the cause of a defect are important.

The number of defect types is large because many of them are paper grade or paper machine specific. The most common defects of base papers are holes, light spots, dark spots, edge cracks, calender cuts, and wrinkles. Typical defects of coated papers are streaks, coating skips, and coating backing roll spots (the so-called butterflies). These defect types can be further divided into subcategories. For example, holes can be classified into wire holes, slime holes, clean holes, etc. Light spots are normally thin locations, or spots caused by water or oil drops. Often different basic defect types appear combined in a same defect. Examples are slime holes and the so-called fish eyes. These add difficulty to the defect classification. Size of a defect is an important classification characteristic. For certain defect types also shape and orientation may be important. For example, wrinkles are formed if there is excess of paper in a cross direction location of the paper web. To the inspection system they appear as repeating dark spots of certain size, shape and orientation.

There are several aspects when considering importance of paper defects. Firstly, it is paid attention to their effects on the quality of paper from the viewpoint of its final use. It is obvious that the quality of e.g. art printing paper gets worse if dirt spots are present in it, although this can be regarded sometimes only as an esthetical problem. On the other hand, white food wrapping paper is not allowed to have small dark spots, because they are regarded then as a hygienic issue. The extreme cases are those where certain defects spoil the product made of it, e.g. holes in a capacitor or in a cigarette paper. Secondly, the defects may cause quality problems in products made by further processing (finishing or converting) of paper. For example, in the coating process defects of the base paper may prevent local adhering of coating material to it. Similar problems occur also when printing coated magazine paper. Thirdly, defects may cause runnability problems in the paper making process itself (in a paper machine or in a coater) or in a customer process such as a printing press. The problems mentioned are typically due to holes, edge cracks, or wrinkles. In these cases also defect size and shape are important issues.

The proposed defect classifier is depicted in Figure 5. There are three stages: a pre-processing stage, a classification stage, and a combiner stage. The classifier stage has three branches, one branch for each feature set. Each branch has a feature extraction unit and a classification unit. The outputs from the classification stage are combined in a combiner to produce the final classification.

2.3.1. Features for defect classification

Shape is an important characteristic of an object. The goal in shape recognition is to find a description which contains sufficient information to distinguish between different target shapes. Five simple shape descriptors are used to characterize the shape of a defect. The descriptors are: convexity, principal axis ratio, compactness, circular variance, and elliptic variance. They are translation, rotation, and scale invariant shape descriptors. Variations of most of them have been widely used in object recognition.[24,25] Each descriptor alone is insufficient for a complex

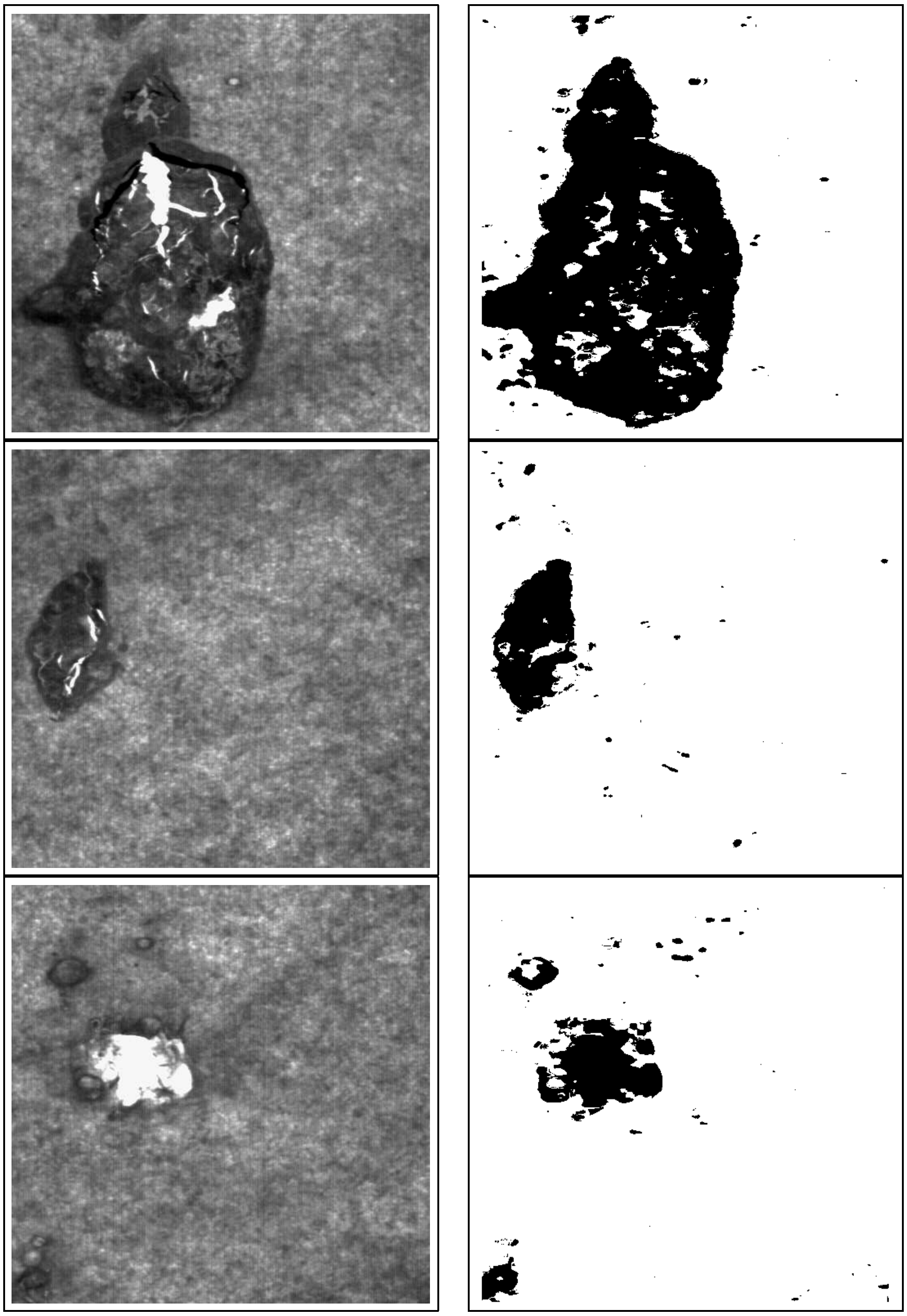

Figure 4. Surface images and their segmentations. Defects are marked with black color.

recognition task, but the combination of them is shown to have good recognition capabilities and low computation costs.[26,27]

In addition to shape a defect has some kind of internal structure that consists of dark and light areas, holes, etc. This internal structure can be regarded as a distribution of gray levels in a defect. This distribution can be characterized as a gray-level histogram. A set of texture features calculated from the co-occurrence matrix[21] is also used to characterize defect's internal structure. The co-occurrence matrix is formed to a defect and four texture

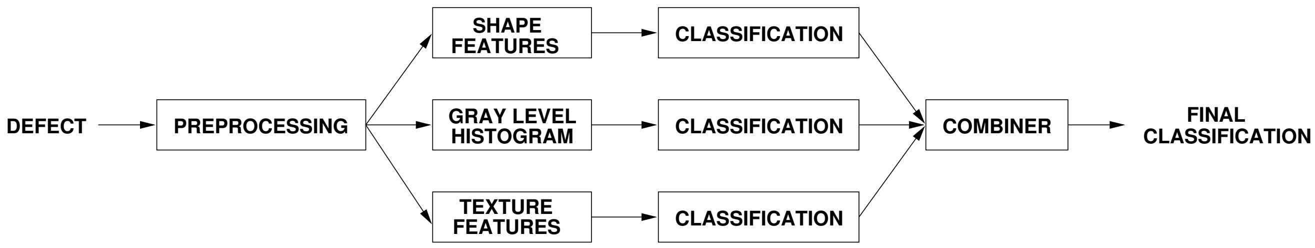

Figure 5. The defect classifier.

features, energy, contrast, entropy, and mean, are calculated from it.

2.3.2. Classification procedure

The codebooks are formed with the self-organizing map (SOM) algorithm.[18] The advantage of using the SOM in codebook generation lies in the fact that the codebooks can be formed in an unsupervised way, simply by clustering unclassified training samples together, and then labeling the SOM units (or clusters) to represent different classes. It is easier to give a label to a cluster of training samples than to each training sample alone, especially if the classes are hard to determine and they are overlapping. In addition to manually given labels each SOM unit can be assigned a few representative defects. They act as examples of typical defects that belong to a SOM unit.

After feature extraction the best-matching unit (BMU) is searched for the given feature vector. The BMU is the SOM unit with the smallest distance between unit's weight vector and the given feature vector. The output of the classification unit consists of two items: the error between the weight vector of the BMU and the given feature vector, and the label of the BMU. The error is used to determine if the given feature vector belongs to the BMU or not. If the latter is the case then we are having a case that has not been encountered previously. A new unit can then be added to the SOM to represent a new class. The weight vector of the new unit is set equal to the feature vector, and the new unit is given a label by hand.

In the combiner the labels from each classifier are combined to produce the final classification. The labels should be verbal descriptions so that the final classification would also be a verbal description.

An example defect and its features are depicted in Figure 6(a). The shape features (SSD) are calculated from the Gaussian smoothed contour of a defect and normalized. Notice that the SSD features result in rotation invariant classification. The gray-level histogram (GLH) and the texture features (TEX) are calculated using only 8 gray levels. In Figure 6(b)-(d) are the classification results. Each column shows the best-matching codebook defect when using different features. In the bottom row are the numbers and the weight vectors of the best-matching SOM units. The first and the second row show one of the typical defects (with its feature vector) that belongs to the best-matching unit. The labels of the best-matching units are *elongated, smooth shape* (SSD), *light spot* (GLH), and *light spot* (TEX). The final classification is then *elongated, smooth, light spot.*

3. CONCLUSIONS

The proposed two-stage classification procedure is a general one and can be used in different classification problems. Reselection of features may be necessary to adapt the proposed classifier to work with different types of surfaces and defects. In defect detection the classifier is taught with examples of fault-free surface while in defect classification shape and internal structure characteristics of defects are learned from examples. The self-organizing maps (SOMs) are used as classifiers. The defect classifier allows new, previously unknown, defects to be added to the codebook gradually during a long period. This is necessary since collecting samples of all possible defects is a time-consuming task and thus it is not reasonable to do it before initial training.

The results of experiments with base paper defects are encouraging. The evaluation of the proposed method with more data and the refinement of the defect classification scheme are the main objectives for future work.

ACKNOWLEDGMENTS

The financial support of the Technology Development Centre of Finland (TEKES's grant 4172/95) is gratefully acknowledged.

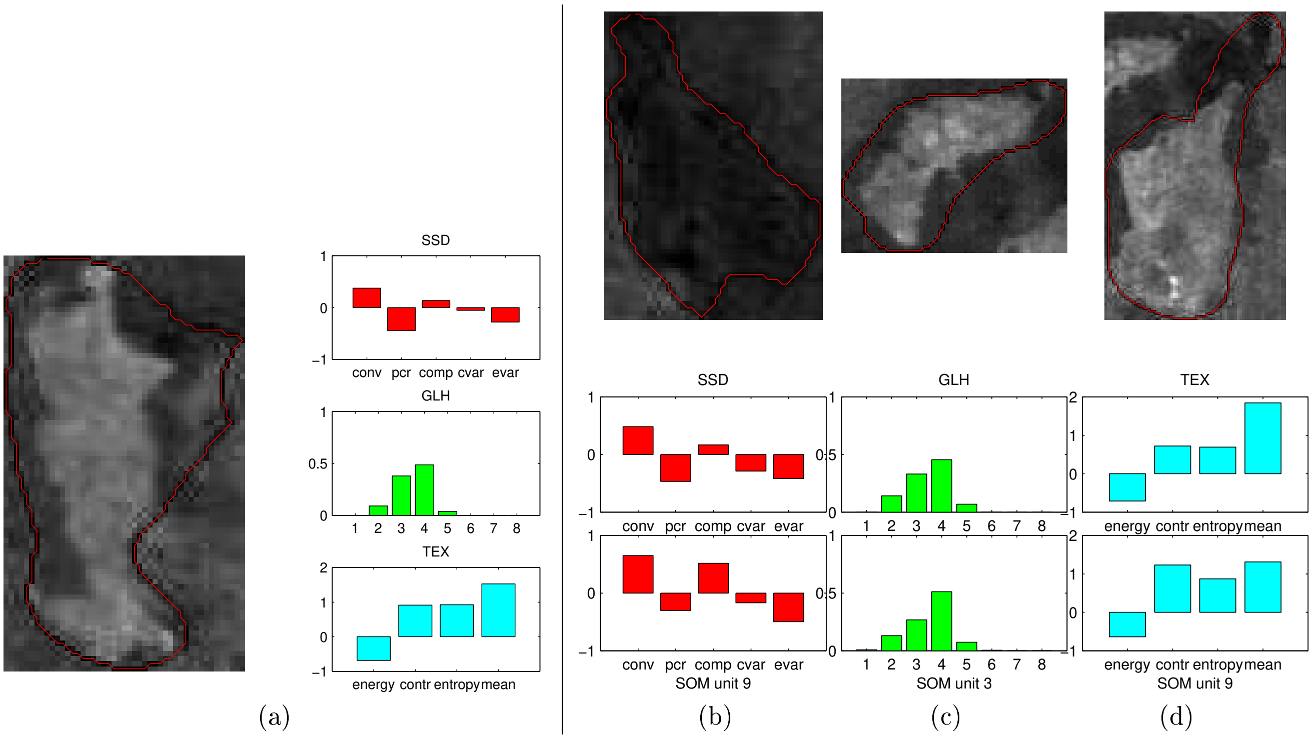

Figure 6. (a) A defect and its shape features (SSD), gray level histogram (GLH), and texture features (TEX). The best matching defects to the defect in (a) with respect to (b) the shape, (c) the gray level histogram, and (d) the texture features.

REFERENCES

1. J. C. Badger, "An automatic web inspection system with advanced defect detection and classification capabilities," in *Proceedings of 1992 Nonwovens Conference*, pp. 69–80, (Marco Island, Florida), May 10–14 1992.
2. J. E. Graf, S. T. Enright, and S. I. Shapiro, "Automated web inspection ensures highest quality nonwovens," *Tappi Journal* **78**(9), pp. 135–138, 1995.
3. H. Ruuska and I. Akerberg, "Practical inspection systems for nonwoven diaper linerfabrics," *Tappi Journal* **78**(9), pp. 135–138, 1995.
4. L. Mäkelin, J. Moisio, and J. Järvinen, "From hole detection to image processing and information technology," in *Proceedings of 50th Appita Annual General Conference*, vol. 1, pp. 817–822, 1996.
5. K. Y. Song, M. Petrou, and J. Kittler, "Texture defect detection: A review," in *Applications of Artificial Intelligence X: Machine Vision and Robotics*, Proc. SPIE 1708, pp. 99–106, 1992.
6. T. S. Newman and A. K. Jain, "A survey of automated visual inspection," *Computer Vision and Image Understanding* **61**, pp. 231–262, Mar. 1995.
7. D. Brzakovic and N. Vujovic, "Designing a defect classification system: A case study," *Pattern Recognition* **29**(8), pp. 1401–1419, 1996.
8. M. Borghesi, V. Cantoni, and M. Diani, "An industrial application of texture analysis," in *Proceedings of the 7th International Conference on Pattern Recognition*, vol. 1, pp. 420–423, (Montreal, Canada), July 30–Aug. 2 1984.
9. M. Unser and F. Ade, "Feature extraction and decision procedure for automated inspection of textured materials," *Pattern Recognition Letters* **2**, pp. 185–191, Mar. 1984.
10. J.-L. C. Guay, "A machine vision system for real time inspection of moving web," in *Intelligent Robots and Computer Vision VIII: Systems and Applications*, Proc. SPIE 1193, pp. 204–228, 1989.
11. F. S. Cohen, Z. Fan, and S. Attali, "Automated inspection of textile fabrics using textural models," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**, pp. 803–808, Aug. 1991.

12. C. Fernandez, S. Fernandez, P. Campoy, and R. Aracil, "On-line texture analysis for flat products inspection. neural nets implementation.," in *Proceedings of the 20th International Conference on Industrial Electronics, Control and Instrumentation*, pp. 867–872, (Bologna, Italy), Sept. 6–9 1994.
13. S. Sardy and L. Ibrahim, "Experimental medical and industrial applications of neural networks to image inspection using an inexpensive personal computer," *Optical Engineering* **35**, pp. 2182–2187, Aug. 1996.
14. J. Iivarinen, J. Rauhamaa, and A. Visa, "Unsupervised segmentation of surface defects," in *Proceedings of the 13th International Conference on Pattern Recognition*, vol. IV, pp. 356–360, (Wien, Austria), Aug. 25–30 1996.
15. J. Iivarinen, J. Rauhamaa, and A. Visa, "An adaptive two-stage approach to classification of surface defects," in *Proceedings of The 10th Scandinavian Conference on Image Analysis*, vol. I, (Lappeenranta, Finland), June 9–11 1997.
16. A. Meier and U. Priber, "On-line surface inspection with fuzzy-asic modules," in *5th European Congress on Intelligent Techniques and Soft Computing*, vol. 3, pp. 2474–2479, (Aachen, Germany), Sept. 8–11 1997.
17. J. Iivarinen and A. Visa, "An adaptive texture and shape based defect classification," in *Proceedings of the 14th International Conference on Pattern Recognition*, (Brisbane, Australia), Aug. 16–20 1998.
18. T. Kohonen, *Self-Organizing Maps*, Springer-Verlag, Berlin, 1995.
19. J. Iivarinen and A. Visa, "Unsupervised image segmentation with the self-organizing map and statistical methods," in *Intelligent Robots and Computer Vision XVII: Algorithms, Techniques, and Active Vision*, D. P. Casasent, ed., Proc. SPIE 3522, 1998.
20. J. Rauhamaa and S. Rantapuska, "Single sensor architecture for inspection of fast moving paper webs," in *XIV IMEKO World Congress*, vol. XA, pp. 25–30, (Tampere, Finland), June 1–6 1997.
21. R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-3**, pp. 610–621, Nov. 1973.
22. J. S. Weszka, C. Dyer, and A. Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-6**, pp. 269–285, Apr. 1976.
23. R. Conners and C. Harlow, "A theoretical comparison of texture algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2**, pp. 204–222, May 1980.
24. S. Marshall, "Review of shape coding techniques," *Image and Vision Computing* **7**, pp. 281–294, Nov. 1989.
25. M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*, Chapman & Hall Computing, London, 1993.
26. M. Peura and J. Iivarinen, "Efficiency of simple shape descriptors," in *Advances in Visual Form Analysis*, C. Arcelli, L. P. Cordella, and G. Sanniti di Baja, eds., pp. 443–451, World Scientific, Singapore, 1997.
27. J. Iivarinen, M. Peura, J. Särelä, and A. Visa, "Comparison of combined shape descriptors for irregular objects," in *Proceedings of the 8th British Machine Vision Conference*, vol. 2, pp. 430–439, (University of Essex, UK), Sept. 8–11 1997.