Content-Based Image Retrieval in Surface Inspection

Jukka Iivarinen, Jussi Pakkanen, and Juhani Rauhamaa*

Helsinki University of Technology
Laboratory of Computer and Information Science
P.O. Box 5400, FIN-02015 HUT, Finland
Email: jukka.iivarinen@hut., jussi.pakkanen@hut.

Abstract

In this paper a prototype system is described for the management and content-based retrieval of defect images in huge image databases. This is a real problem in surface inspection applications, since modern inspection systems may produce up to thousands of defect images in a day. The retrieval is based on shape and internal structure characteristics of defects, so no manual labeling nor annotation needs to be made. We are using a noncommercial, generic content-based image retrieval (CBIR) system called PicSOM that is modified to fit to the special requirements of our application. The results of experiments show that the prototype system works fast with good retrieval results.

1 Introduction

The need for efficient and fast methods for content-based image retrieval (CBIR) has increased rapidly during the last decade in many fields of industry and research. For example, modern surface inspection systems produce huge amounts of image data that has to be stored, managed, browsed, searched, and retrieved efficiently. The CBIR systems seem to be ideal tools for management of this huge image data.

Surface inspection of web materials is a challenging problem that consists of several subproblems, e.g. image acquisition, defect detection, and defect classification. For example, for running web processes there exist commercial solutions [12, 15, 2]. In this paper we are dealing with management and content-based retrieval of defect images that are obtained from a real, online paper inspection system. We have designed and realized a prototype CBIR system that is based on a noncommercial, generic content-based image retrieval system called PicSOM [9, 10].

2 Surface Inspection Problem

Surface quality is one of those paper properties that are constantly watched by paper makers. This is because various anomalies of paper surface not only spoil the quality of final product, but they may be critical to the runnability of paper machines and converting or finishing equipment as well. Especially harmful are large holes, but also other types of weak areas in the paper may induce paper breaks during production.

3 Content-Based Image Retrieval

In content-based image retrieval, similar images are searched from a database based on the similarity of their visual features. These features may measure different visual properties of images, e.g. their color, tex-
ture, or edgeness. No manual annotation nor labeling of images is needed. This is a desired property in surface inspection since, as discussed in the previous chapter, labeling of defect images is hard or even impossible.

One of the first and maybe still the best-known CBIR system is QBIC (Query by Image Content) [1], but several other commercial and noncommercial CBIR systems have been introduced during the last decade [14]. The problem with these kinds of ready-made CBIR systems is that they are designed to be generic ones. Thus applying them to a specific problem at hand is not always feasible nor efficient. There usually exists some problem-specific knowledge that could be utilized in the CBIR process, but this knowledge cannot be embedded to a generic system. Also these generic CBIR systems are usually not efficient enough to be used in demanding, real applications, such as surface inspection.

3.1 PicSOM Retrieval System

PicSOM has been developed in Laboratory of Computer and Information Science at Helsinki University of Technology to be a generic content-based image retrieval (CBIR) system for large, unannotated databases [9, 10]. It builds on the concepts of unsupervised clustering, self-organizing maps, and relevance feedback. The clustering property of PicSOM is very useful in our application since it makes it possible to automatically find relevant classes within defect images. This helps to detect the most severe defects or any other defect types of interest. Since PicSOM has been developed in our laboratory, it has been easy to modify it to meet the special requirements of our surface inspection problem.

**PicSOM Training**

PicSOM works by first calculating a chosen number of feature sets for each image in the database. Several types of features can be used e.g. color, shape, texture, and structural features. Then a tree-structured self-organizing map (TS-SOM) [8] is trained with each feature set. TS-SOM is a basic building block of the PicSOM system. It is a tree-structured vector-quantizer that has self-organizing maps (SOMs) at each of its hierarchical levels. The SOM sizes grow from top to bottom (see Figure 3). This kind of tree-structure makes training and searching much faster when compared to a normal, nonhierarchical SOM of the same size. After training, each image is associated to the closest map unit in each TS-SOM. Then the database can be searched.

**Relevance Feedback in PicSOM**

The relevance feedback means that the PicSOM system learns during the user query what kind of images the user wishes to retrieve and what features are important in each query. First the user selects a set of images that, in his opinion, are similar to the image(s) he is searching for. These images are marked on the TS-SOMs with positive weights and the unselected images with negative weights. Then the TS-SOMs are low-pass filtered to enhance the effect of the positive weights (or the user-selected images). These low-pass filtered response maps are seen on the bottom levels of the TS-SOMs in Figure 3. Then PicSOM finds the images that are close to the selected images in TS-SOMs using these response maps. Seek results from all three TS-SOMs are combined (see a detailed description on how this is done in [10]) and the best-matching images are returned. These images are then shown to the user who again selects the ones that are similar to his target image. When this procedure is repeated, the selected images should form cluster(s) in TS-SOMs (see the bottom levels of the TS-SOMs in Figure 3). Since these clusters are formed by the user, they should be similar to his target images. Eventually the system shows the user the desired image and the search is completed.

4 A Prototype System for Defect Image Retrieval

In this paper we have applied the PicSOM system to content-based retrieval of paper defect images in a large unannotated defect image database. The problem at hand is now the following one: Given a new defect or a set of defects, retrieve similar defects that might have appeared previously. The retrieval is based
on different features that are extracted from defect images, so there is no need for manual annotation nor labeling.

4.1 Overview of our CBIR System

We have made some modifications to the original PicSOM system that affect mostly feature extraction and visualization parts of PicSOM. As an extra problemspecific knowledge we have binary segmentation masks for each defect image. This information is utilized in PicSOM so that feature extraction is only done for defect areas in each defect image. PicSOM had three TS-SOMs, one for each feature set. The TS-SOMs had three levels, $4 \times 4$, $16 \times 16$, and $64 \times 64$.

The basic structure of our CBIR system is shown in Figure 2. Given a new defect image, three different feature sets are extracted (shape features, gray level histogram, texture features), similar defects according to each feature set are retrieved from the database, and these retrieved images are combined to produce the final, combined set of similar defect images.

4.1.1 Feature Extraction

Two types of features are of interest when considering defect images: shape features and internal structure features. Shape features are used to capture the essential shape information of defects in order to distinguish between differently shaped defects, e.g. spots and wrinkles. Internal structure features are used to characterize the gray level and textural structure of defects.

**Shape features** Five simple shape descriptors are used to characterize the shape of a defect [5, 7]. 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 these shape descriptors have been widely used in object recognition [11, 16]. Each descriptor alone is insufficient for a complex recognition task, but a combination of them has been shown to have good recognition capabilities and low computation costs [5]. We are also using the angle information as a sixth feature since we do not want to have a rotation invariant feature set (it is vital to distinguish between horizontally and vertically oriented defects).

**Internal structure features** In addition to shape a defect has some kind of internal structure that consists of dark and light areas, holes, etc. The internal structure can be regarded as a distribution of gray levels in a defect. This distribution can be characterized as a simple gray-level histogram. In addition to a gray-level histogram, a set of texture features calculated from the co-occurrence matrix [3] is also used to characterize the defect’s internal structure. The co-occurrence matrix is a second order statistical measure of gray level variation that indicates the joint probability of gray level occurrences at a certain displacement in an image. The co-occurrence matrix is formed to each defect and four texture features, energy, contrast, entropy, and mean, are calculated from it. The texture features were calculated from the $256 \times 256$ co-occurrence matrix of a defect.

4.2 Defect Database

The defect database has 13000 defect images that were obtained from a real, online paper inspection system. The images have different kinds of defects, e.g. dark and light spots, holes, and wrinkles. Their sizes vary according to the size of a defect, and they have 256 gray levels. After acquisition, the images are segmented so that each defect image has a gray level image and a binary segmentation mask that indicates defect areas in the image. Some example images with their segmentation masks are shown in Figure 1. Segmentation of these kind of images is a very hard problem, and it cannot be done with simple gray-level thresholding techniques [6, 4]. The image database with the defect segmentation masks were provided by ABB Oy. All further processing is done only for defect areas; we are thus omitting the noninteresting background (or normal, intact surface). It is important to get rid of the background since it usually covers most of the image. Thus it may spoil feature extraction and then the whole CBIR process; the point here is to find similar defects, not similar images. Especially the shape features need to have the segmentation information, otherwise they would be impossible to extract. On some occasions there was more than one defect in
Figure 3: This figure shows the user interface of PicSOM. On the top are three TS-SOMs (one for each feature set) and then the images selected by the user. Below these are the best-matching images found by the PicSOM system.

an image. In this case we only calculated the features of the largest defect.

4.3 Preliminary evaluation results

Measuring our CBIR system objectively is difficult. We do not have a standardized set of test images. The images come from a real on-line process, and since they are not pre-classified, we must measure the search results subjectively through visual inspection.

The example query in Figure 3 shows that the system works quite well. Under the three TS-SOMs are the images selected by the user (or the query images), and on the bottom are the images returned by the PicSOM system. All returned images are visually similar to the query images. The system retains a similar level of success when queried with different types of defects. This is illustrated in Figure 4. In it each row represents a single query. On the left is a single query image and to their right are the best matches returned by the system. It can be seen that the system works with several different kinds of images, such as pale blemishes, white streaks and so on.

The shape and texture maps work especially well, with the histogram map having slightly worse performance. However the true power comes from combining the maps. The PicSOM engine combines the various maps in a powerful manner, yielding good results. The PicSOM system is also very fast. Its tree-based structure yields efficient indexing and querying engines. With the database of 13000 defect images, queries are completed almost in real-time.

4.4 Benefits of our system

We have shown that our system clusters similar defects quite well. The search times are also very low. These features are very useful in on-line web inspection. An expert can classify clusters in the TS-SOMs according to their severity. The system can then compare newly found defects to the images in the database. If the new image goes near the severe clusters, the system can issue a warning and prepare for possible complications.

Another problem of web inspection is that different paper machines may produce different kinds of defects. This is because paper mills utilize different technical solutions for raw material processing, actual paper web forming, drying, and surface treatment. In addition, various raw materials and variations in their quality cause different types of surface defects to emerge. This is not a problem for our system. It can be easily retrained with image data gathered from the current paper machine. Therefore our system is very efficient and flexible.

5 Conclusions

In this paper we have discussed the surface inspection problem and especially the content-based retrieval of defect images in huge image databases that are common to such problems. We outlined the basic structure of our prototype CBIR system that builds on the non-commercial CBIR system called PicSOM. Some modifications were made to the original PicSOM system that affected mostly feature extraction and visualization parts of PicSOM. Experiments were done with the image database that has 13000 surface defect images taken from a real, online process. These images were automatically segmented beforehand to aid the retrieval process. The results of experiments are very promising; our prototype CBIR system works fast with good retrieval results.

Acknowledgments

The authors wish to thank the PicSOM group (J. Laaksonen, M. Koskela, S. Laakso, E. Oja) at Helsinki University of Technology. The financial support of the Technology Development Centre of Finland (TEKES’s grant 40397/01) and the Academy of Finland (project New Information Processing Principles, 44886) is gratefully acknowledged.
Figure 4: Shown here are a few example queries. On the left are the query images and next to them are the best results found.

References


