A Content Based Retrieval Engine for Circuit Board Inspection

Badrinarayan Raghunathan and Scott T. Acton

Oklahoma Imaging Laboratory
School of Electrical and Computer Engineering
Oklahoma State University
Stillwater, OK 74078 U.S.A.
E-mail: sacton@okstate.edu

Abstract

In this paper we apply content based retrieval (CBR) to the automated inspection of printed circuit boards. Manual probing of faulty circuit boards is expensive and time consuming for electronics manufacturers. The proposed CBR system allows groups of similarly faulted boards to be identified and repaired simultaneously. We improve upon current matching techniques used in CBR, exploiting relative pairings of features rather than simple distance measures. The novel matching techniques are seen to improve upon the traditional approach in cases of varying background conditions, environmental factors and inter-chip interactions.

1. Introduction

There is an increasing trend toward digital image databases in photography, medicine, engineering, science and the entertainment industry. Given these growing amounts of digital imagery, content based retrieval (CBR) is becoming a subject of intense research in the image processing community [1], [2]. In this paper, we demonstrate a CBR application that is profitable to the industry and extends beyond conventional web-searching CBR applications.

We propose a CBR engine for the automated inspection of printed circuit boards. Typically in CBR, the similarity between two images is assessed by means of a simple distance measure between two corresponding feature sets. We propose an approach that is quite different. In our technique, pairs of features in the query image and in the candidate match image are compared to the same pair of features in a set of prototype images. Here, the prototype images are functional circuit boards that are of the same model as the query board. Since the proportionality between feature pair vectors of the prototypes is maintained under differing imaging conditions, the feature pair values exhibit a linear relationship in feature space. Therefore, similarity between a query image and a candidate image is assessed by means of a distance measure between feature pair vectors and a linear path through feature space. The linear path is computed by regression on the feature pair vectors of the prototypes.

At present technicians spend a considerable amount of time probing circuit boards that fail conventional electrical tests. It is our goal to reduce this waste of technician effort. As part of our inspection system, the circuit boards are imaged by an infrared digital imaging system. The thermal signature from the operating chips is captured by an infrared video sequence. All circuit boards that fail electrical tests are designated faulty circuit boards (FCB's) while the others are designated known good boards (KGB's). The proposed CBR system is intended to be an integral part of the overall inspection system. Presently data from the last frame of the video sequence is used for CBR, as the chips have reached thermal equilibrium.

Given a query in the form of specific intensity features of a board image, the CBR system retrieves similar images from the board image database. The ultimate goal of this CBR system is to retrieve similar FCB's so that they can be repaired simultaneously. Figure 1 shows a sample query and the corresponding matches generated by the CBR engine.

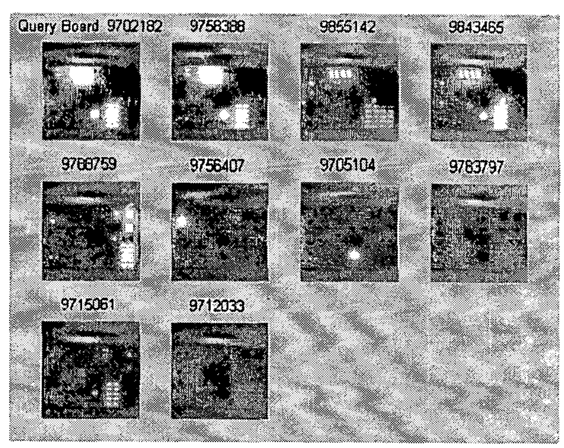


Figure 1. Sample query and CBR results

0-7803-5467-2/99/ $10.00 © 1999 IEEE

The proposed CBR system can be broken down into two parts – the actual search engine and the off-line processing that generates the system libraries. Figure 2 shows the configuration of the system library. The first section of the paper explains the off-line processing of the inspection system and the consequent generation of the system library. This section is followed by the proposed methodology for formulation of higher level features to be used in CBR. Experimental results that demonstrate the successful application of CBR to manufacturing are given.

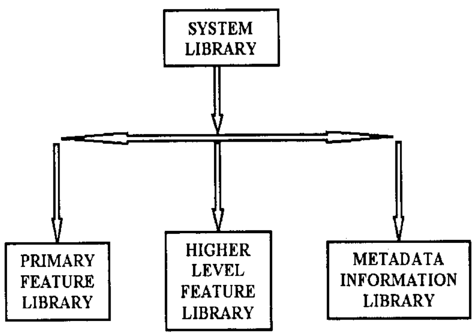


Figure 2. Configuration of system library

OFF-LINE PROCESSING

Figure 3 summarizes the off-line processing of the circuit board video sequences.

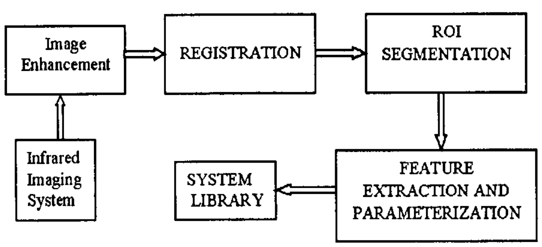


Figure 3. Off-line processing

Each of the digital image sequences is first subjected to an image enhancement routine by which image contrast is improved. Image registration is performed to remove the effects of translation and rotation during the imaging process. Various regions of interest (ROI's) are identified by an image segmentation routine (see Figure 4). The ROI template delineates regions corresponding to important integrated circuit packages on the board.

Feature extraction through the video sequence is performed within image segments specified by the ROI template. Some of the primary features extracted are the average intensity of chip, the maximum intensity of chip, the minimum intensity of chip, Haar wavelet coefficients, image granulometry and the *chip isolation intensity*. The chip isolation intensity is the height of a Gaussian model fit to the 2-D intensity signature of a particular chip. In addition, each circuit board in the sequence database library has some associated *metadata*. The purpose of using metadata in CBR is to facilitate rapid retrieval of possible matches. In this application the metadata are quite limited. We are given metadata that reveal the class of a board and its type (KGB/FCB).

Simple extraction and comparison of primary features, such as the average intensity of a chip, is not appropriate for CBR, because background conditions, environmental factors and various other inter-chip interactions render the singular comparisons ineffective and undependable.

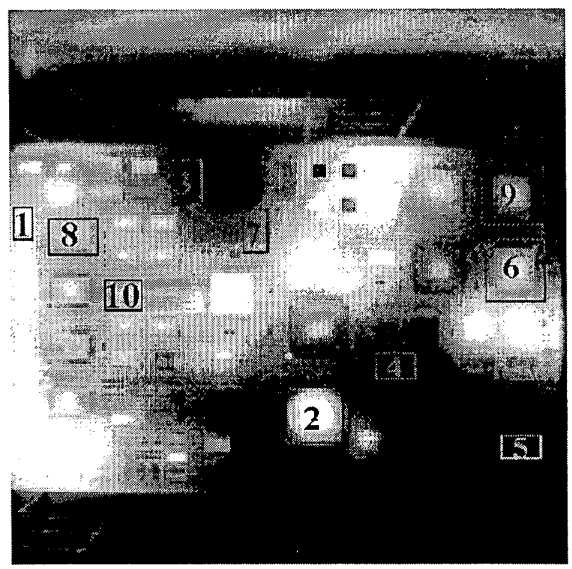


Figure 4. A segmentation template with ROI's (chips) numbered

2. Formulation of Higher Level Features

The video feature extraction process generates an overabundance of data that necessitates reduction to higher level features. The pairwise feature analysis technique is used for formulation of higher level features from primary features. The pairwise relationship of a single primary feature such as average chip intensity can be visualized as in Figure 5. The various points in the scatterplot correspond to various FCB's and KGB's. A line can be fit to the KGB's (prototype instances) to the feature pair coordinates via linear regression. The line is defined by

$$y = \beta_0 x + \beta_1$$

where β_0 corresponds to the slope and β_1 corresponds to the y-intercept of the regression line. The residual γ of a particular feature pair vector coordinate (x_i, y_i) with respect to the linear regression result is given by

$$\gamma = | \ y_i - \beta_0 x_i - \beta_1 \ |.$$

This residual is computed for all pairs of prototype features on the feature plot. The mean of all these residuals is referred to as γ_{KGB}, the standard deviation for the prototype set. γ_{iFCB} is defined as the residual of a feature pair corresponding to a FCB i.

A tolerance factor Γ_p is computed from the standard deviation of the features pair values in the prototype set. If $\gamma_{iFCB} \leq \Gamma_p$ is not satisfied, the corresponding feature pair on FCB i is considered to be an outlier. In other words, the point corresponds to an outlier board. An outlier board has an associated residue factor ε given by

$$\varepsilon = | \ \gamma_{iFCB} - \Gamma_p \ |.$$

This residue factor is a measure of the distance of the outlier board from the estimated regression line.

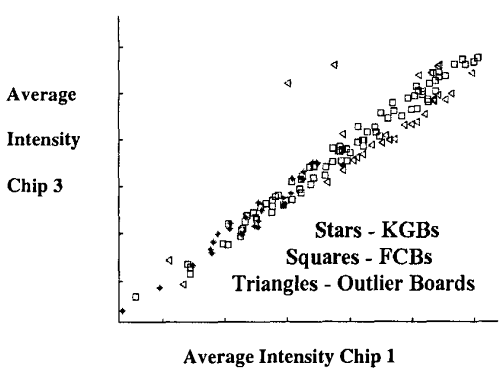


Figure 5. Pairwise feature scatterplot with linear regression

A list of outlier boards is computed for all pairwise feature sets (each scatterplot). This list of outlier boards for all pairwise features is reduced to a list of outlier boards associated with each ROI. This list is computed using the aggregate residue factors corresponding to a particular board in various feature pairs. Hence these boards have an aggregate residue factor associated with each ROI. The aggregate residue factors are designated as higher level features. These results are used to prioritize the various ROI on a board in order of most likely cause of defect to least likely cause of defect and to perform similarity assessment in CBR.

3. Query Formulation and CBR

In this application, we have tested two techniques for CBR. One approach uses primary features, another exploits higher level features. The primary feature used in the CBR examples in this paper is the chip isolation intensity.

The query can be formulated by specifying the query FCB. The objective of CBR in this context is to find FCB's with a similar set of faulty chips. Based on the metadata information associated with the query board, the CBR engine examines only boards within the target class. The candidate boards are selected from the target database, and a similarity factor is computed by a least squares matching between higher level features generated by our pairwise methodology as compared to CBR using primary features. The matches are displayed in order of their similarity factors.

One of the most difficult problems in CBR is the designation of a cutoff parameter that limits the pool of possible matches. In this application, the cutoff parameter is specified automatically by the first retrieved KGB in the pool of possible matches. Figure 6 shows the query FCB and the best match by CBR using primary features, and Figure 7 shows the query FCB and the best match for CBR using higher level features generated by our pairwise feature analysis.

CBR by way of higher level features has shown the advantage of increased match relevance. For the circuit board application, this relevance is quantified in terms of overlapping faulty chips between the boards with respect to the query FCB.

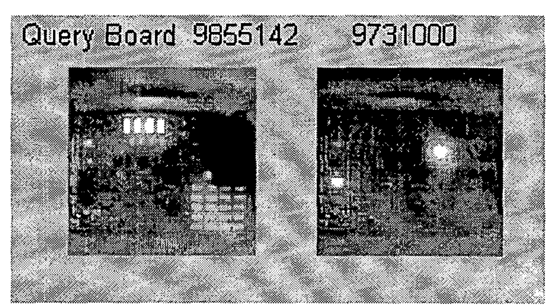


Figure 6. CBR using primary features

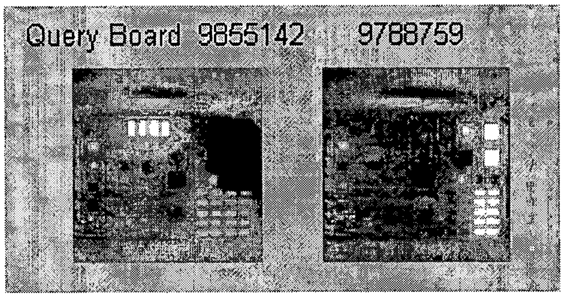


Figure 7. CBR using higher level features

4. Analysis

We now introduce two terms: the intra-board correlation factor and the inter-board correlation factor in the analysis of our methodology and our results. The intra-board correlation is a correlation measure computed for average intensity data sets for various chip pairs. The Karl-Pearson correlation coefficient is used for this purpose:

$$\rho = \frac{Cov(x,y)}{\sqrt{Var(x)Var(y)}}.$$

Given n chips on each board, each chip has n-1 possible chip pairs. On the whole there are $n(n$-$1)$ possible chip pairs. Table I is indicative of the intra-board correlation factor behavior over all $n(n$-$1)$ chip pairs.

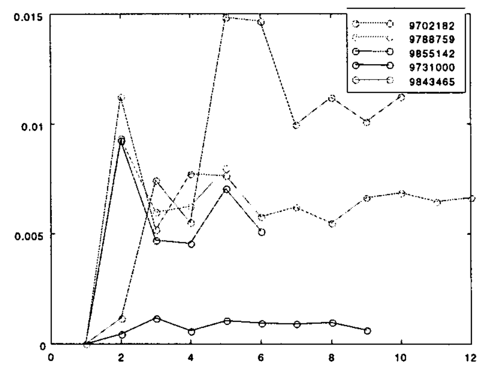
The standard deviation and the average of the intra-board correlation factor ρ from Table I clearly portray the linear behavior of the given data sets. This validates the application of linear regression to the given data sets.

Table I. Intra-board correlation factor behavior for 90 possible chip combinations (*i.e.* 10 chips)

Minimum ρ	**0.92768**
Maximum ρ	**0.99912**
Mean ρ	**0.97765**
Standard Deviation ρ	**0.020323**

The inter-board correlation factor is computed by means of a proportionality factor. Given boards B_1 and B_2 with feature sets $(f_1 \ldots\ldots f_n)$ and $(f_1^{'} \ldots\ldots f_n^{'})$. The proportionality factors for B_1 and B_2 are $(\frac{f_1}{f_1^{'}} \ldots\ldots \frac{f_n}{f_n^{'}})$. Let the average proportionality factor be $\frac{f}{f^{'}}$. Let us now study the inter-board correlation between the query board and each of the matches. One would expect the correlation to be maximal for a perfect match and then decrease as the order progresses from best match to worst. Thus, the standard deviation of the associated proportionality factors should increase from the best to worst matches. Figure 8 shows this relationship, which confirms the CBR results.

SD of $\frac{f}{f^{'}}$



Match number (from best to worst)

Figure 8. Standard deviation of inter-board correlation from best to worst matches for different query boards

The fact that the behavior of the inter-board measure is not monotonic indicates that simple distance measures or simple proportionality measures do not guarantee a proper relevance of matches in terms of a decreasing order of similarity.

5. Comments

We have developed a working prototype of the CBR engine described in this paper. We have also developed tools to formulate queries for CBR by specific chip(s) on a specific board. We are in the process of introducing new models of boards to the board database library so as to observe the scalability of the entire CBR system. Our experimental results show that the cutoff is reached earlier using CBR by higher level features generated by our pairwise methodology, as compared to CBR using primary features. The pairwise feature analysis technique has proven to be an effective tool in the formulation of higher level features. This novel technique overcomes the shortcomings of conventional primary feature-based techniques by use of relative pairings of feature points rather than simple distance measures. The simple distance measures used with primary features are ineffective when imaging conditions are variable or when inter-chip interactions are significant.

We also strive to improve the flexibility of the query formulation, which is important from the perspective of this application. Currently, our focus is on automating the extraction of board features and consequent generation of the system library. The advantage of our methodology is associated with the pairwise technique used to generate a significantly simplified, yet comprehensive on-line database library, lending to the increased relevance of matches. This bodes well for extension of our pairwise feature analysis by prototype modeling methodology to more general CBR applications such as digital mammography and E-commerce.

6. References

[1] W. Y. Ma and B. S. Manjunath, "NeTra: A Toolbox for Navigating Large Image Databases," *Proc. of the IEEE International Conf. on Image Processing*, Santa Barbara, CA, pp. 568-571, October 1997.

[2] A. Pentland, R. W. Picard, and S. Sclaroff, "Photobook: tools for content based manipulation of image databases," *Proc. of the SPIE Conf. On Storage and Retrieval for Image and Video Databases –II*, No. 2185, San Jose, CA, pp. 34-47, February 1994.

[3]. W. Niblack et al., "The QBIC project: querying images by content using color, texture, and shape," *Proc. of the SPIE Conf. On Storage and Retrieval for Image and Video Databases*, vol. 1908, San Jose, CA, pp. 173-187, February 1993.

[4]. J. S. Milton and Jesse C. Arnold, *Probability and Statistics in the engineering and computing sciences*, McGraw-Hill Book Company, USA, 1986.