A Multiscale Assembly Inspection Algorithm

An important aspect of robust automated assembly is an accurate and efficient method for the inspection of finished assemblies. This novel algorithm is trained on synthetic images generated using the CAD model of the different components of the assembly. Once trained on synthetic images, the algorithm can detect assembly errors by examining real images of the assembled product.

At a time when quality and cost are becoming even more important in the manufacturing process, accurate and efficient inspection is critical. However, the complexity of electrical and mechanical assemblies has reached a point where human inspection can be fatiguing, unreliable, and expensive. This has prompted many manufacturers to implement automated inspection systems, particularly for the inspection of two-dimensional components such as printed circuit boards [1]. These systems are typically driven by the information available in the CAD model of the circuit to be implemented.

Unfortunately, efforts to achieve the advantages of CAD-driven inspection systems for three-dimensional assemblies have been largely unrealized. While a large body of research exists in the area of CAD-driven vision [2], this work has been primarily focused on the recognition of three-dimensional objects as opposed to their inspection. For object recognition tasks, one typically needs to index into a large or complex database of possible object models. Discrete object features, such as edges and corners, have been widely used in these recognition approaches because they allow for efficient indexing into the model database [3].

In contrast to feature-based techniques, template-based approaches directly compare gray-scale image data to a predefined model or template. For inspection, the additional information present in gray-scale data makes template-based approaches potentially more sensitive. In fact, direct comparisons have shown that template-based methods can outperform feature-based methods in some applications [4]. Traditionally, template matching has been viewed as being computationally expensive; however multiresolution processing can make these approaches computationally efficient [5].

Our approach to automated inspection is grounded in multiresolution template matching, however, the templates are stochastic models of the expected gray-scale variations in the image [6, 7]. These models are built by examining computer-generated images that exhibit the allowable variations in a correct assembly. The process of obtaining a model is conceptually illustrated in Figure 1 using the example assembly shown in Figure 2. The process starts by using information from the CAD model to generate multiple ray-traced images of the assembly within the range of acceptable tolerances [8, 9]. Connectivity information calculated from the CAD model is used to create an object tree that describes how assembly errors can manifest themselves in the image [10, 11]. The synthetic images and object tree are then used by the training algorithm to determine a stochastic model of the acceptable gray-scale and positional variations in an image of a real assembly. The training is performed using the Expectation Maximization (EM) algorithm [12] with the resulting parameters normalized using a single real image.

A key feature of our approach is the utilization of computer-generated synthetic images in the training process. This approach has the advantages of avoiding the manufacture of physical prototypes while providing a wider variety of acceptable variations. This not only results in a more robust inspection algorithm but also
We approach automated inspection as a problem in object detection, where we assume the inspection algorithm must make decisions based on a monochrome image of the object. This facilitates concurrent design of the assembly and its inspection system. Since each step in the inspection design process is automated, inspection becomes feasible for small batches.

In the next section we present an overview of the multiscale inspection. We will next describe the CAD database and its role in guiding the inspection algorithm and address the synthetic image generation process. We will then present an example with our experimental results and summarize our conclusions.

**MULTISCALE OBJECT DETECTION**

We approach automated inspection as a problem in object detection, where we assume the inspection algorithm must make decisions based on a monochrome image of the object. Our multiscale detection algorithm is based on a stochastic object model, which is tailored to a specific object by adjusting the model structure and changing model parameters. The model generation and parameter estimation is driven by a CAD model of the object as described in the following sections.

Our inspection algorithm models an object as a stochastic tree, where the nodes of the tree represent various components, or subassemblies, of the object. These subassemblies contain the key features for discrimination and error detection. Nodes near the root of the tree typically model larger structures that aid in locating the object while nodes further down "zoom in" on the critical areas where assembly errors are likely to occur. We represent the two-dimensional position and orientation of each subassembly as a state vector $X$. Since the position of a subassembly varies from image to image we model it as a random vector. The state density function for a node depends only on the state of its parent node in the object tree and on a set of node specific parameters $\phi$. Thus, the states form a Markov chain along any path from the root of the object tree to a leaf node.

Figure 3 shows an object tree, where the superscript $(i)$ is used to denote quantities specific to node $i$. The image data $Y$ associated with each node is modeled as a set of random variables with density functions parameterized by a template $\theta$ that indicates the expected appearance of the subassembly as well as the expected data variability. The data values will also depend on the position of the subassembly in an image, so the overall object model is as shown in Figure 3, with the arrows
from the root to the leaves, performing a sequential MAP (SMAP) estimate of the state at each node [6] and passing this estimated state to the child nodes. The object passes inspection only if all subassemblies are located and found to match the model. This procedure hinges on our ability to compute a log likelihood ratio at each state and resolution, which in turn relies upon a knowledge of the parameters $\theta$ and $\phi$ for each node. These parameters, as well as the structure of the object tree, are determined using the CAD model of the object. These parameters are adjusted at the actual inspection site by using a single real image to compensate for variations in lighting environments.

We first use the CAD model to identify the important subassemblies and generate the object tree. The CAD model then generates a series of training images, each of which contains a properly assembled object, with all subassemblies and viewing conditions within their allowed tolerances. The object tree is “overlaid” on one of the training images, identifying the location and orientation of each subassembly in that image. This information is used to initialize the model parameters, which are then determined from the full set of training images via the iterative expectation maximization (EM) algorithm [6, 12].

We use the multiresolution Haar transform of each image as our data (see Figure 4) and use a corresponding multiresolution template at each node of the object tree (see Figure 5). This allows us to model each node at resolutions appropriate to the important features in the subassembly. It also permits us to search for the subassemblies via a fast multiscale search technique. We use recent results from the theory of multiscale random processes to aid in the analysis and construction of the model [13, 6].

The search for the most likely position $X$ of a subassembly begins at a coarse resolution and progresses to finer resolutions. For a given resolution and candidate state we use the image data and templates at that and coarser resolutions to compute the log likelihood ratio between the hypothesis that the subassembly is present and the hypothesis that it is not. The states with the largest log likelihood ratio are investigated at the next finer resolution. The search continues in this fashion until the largest log likelihood ratio exceeds a predefined decision threshold $\beta$, at which point the search returns the associated state as the position of the subassembly. The search will terminate in a “no match” condition if it reaches a point where all remaining candidate states have log likelihood ratios less than a rejection threshold $\alpha$. Thus, the search takes the form of a sequential likelihood ratio test, where the search progresses in scale rather than time. A more detailed discussion of the algorithm can be found in [7].

The inspection algorithm searches for the object in an image by using this fast multiscale search technique at each node of the object tree. The search traverses the object tree from the root to the leaves, performing a sequential MAP (SMAP) estimate of the state at each node [6] and passing this estimated state to the child nodes. The object passes inspection only if all subassemblies are located and found to match the model. This procedure hinges on our ability to compute a log likelihood ratio at each state and resolution, which in turn relies upon a knowledge of the parameters $\theta$ and $\phi$ for each node. These parameters, as well as the structure of the object tree, are determined using the CAD model of the object. These parameters are adjusted at the actual inspection site by using a single real image to compensate for variations in lighting environments.

We first use the CAD model to identify the important subassemblies and generate the object tree. The CAD model then generates a series of training images, each of which contains a properly assembled object, with all subassemblies and viewing conditions within their allowed tolerances. The object tree is “overlaid” on one of the training images, identifying the location and orientation of each subassembly in that image. This information is used to initialize the model parameters, which are then determined from the full set of training images via the iterative expectation maximization (EM) algorithm [6, 12].
**CAD DATABASE ANALYSIS**

As discussed in the preceding section, data obtained from an analysis of the CAD model of the assembly is used to guide the inspection algorithm in the training process by addressing the different variations possible and identifying an appropriate object tree in the training images. To be able to produce a large number of synthetic images with the required variations, a suitable CAD model has to be generated for the desired assembly and its components. The relationship among the different components of the assembly must be known to provide the inspection algorithm with an appropriate object tree.

For purposes of illustration, the pattern wheel assembly pictured in Figure 2 is used as a simple example of a realistic assembly used in a small batch manufacturing environment. For each component of the pattern wheel assembly, a CAD model was created using the TWIN Solid Modeling Package developed by the CADLAB of Purdue University. TWIN is a boundary representation solid modeler but also accepts Constructive Solid Geometry (CSG) models as input. Therefore, components are created in the CSG format and then converted to boundary representation for internal calculations [8].

Three CSG models are created for every component to account for component variations that naturally occur in a manufacturing environment. One model represents the component at its maximum tolerance limit while another represents it at its minimum tolerance limit. Both models are created by using the maximum and minimum dimensions that allow the component to function as designed. These are the Maximum Material Condition (MMC) and Least Material Condition (LMC) of the component, with the region in between forming the component’s tolerance zone [9]. Creating a CAD model of the component that lies within its tolerance zone generates an instance of the component that is within its allowed tolerance. Figure 6 shows an example of the tolerance zone of a component. A new random instance of the component is generated by tracing the CSG tree of the LMC and the CSG tree of the MMC models of the component and selecting from a uniform distribution of the geometric parameters of the primitives in the CSG trees. The third CSG model represents the component at its optimal dimensions which can then be used, along with the other two CSG models, for generating a nonuniform distribution of random components that are within tolerance.

The locations of all components within the nominal assembly are specified by homogeneous transformation matrices that are used to accomplish two tasks. First, the variations between the relative locations of different components is specified in terms of a tolerance zone for these homogeneous matrices [10] so that different instances of properly functioning assemblies can be created. Second, when the components are transformed as specified by these homogeneous matrices, the contact relationship between any two components can be found by querying the modeler. The identification of the bounding faces between any components that are mated is used to form a graph in which the vertices represent the different components of the assembly and the edges between the vertices represent liaisons between the components, as connections or as contacts. This graph is called the liaison diagram [11] and it shows the interaction among the different components of the assembly. This process is demonstrated using Figure 7 which shows an exploded view of the pattern wheel assembly. This figure illustrates the order of operations required to complete the assembly and highlights the single common axis of insertion for the pins and the shaft. Using this information along with the correct final state of the assembly, the system determines which surfaces will be in contact. For example, the single cylinder pins (the high density and the unlatch pins) are inserted only into wheel-a1 and wheel-a2 and therefore form a close relationship with the wheels through their side surfaces. The shaft is only inserted into the center hole of the gear. The multi-cylinder pins (the alignment pins), however, are in contact with the holes in the gear, the locking ring, and the wheels and functionally are required to maintain a precise separation between the two wheels as well as between wheel-a1 and the locking ring which rests on the gear. This last piece of information can not
be clearly deduced from the exploded view but is clear in the resulting liaison diagram shown in Figure 8 which is formed from the contact information explained above. Each circle in the diagram represents a surface on an individual component with the arrows indicating a contact between surfaces of different components. The liaison diagram forms the basis from which the object tree is generated. Contacts between different components imply that they were brought together by the assembly process and that errors in that process would result in misaligned or missing contacts. Thus, to check for errors in the insertion of an individual component, the liaison diagram is traversed starting with the vertex associated with that component. This process is illustrated for the high density pin in Figure 9, resulting in a data structure that we refer to as an error tree. The error trees for all components of interest can then be combined to create the complete object tree (see Figure 3) that guides the inspection process.

All relevant information needed by the inspection algorithm that can be gleaned from the CAD model is now available. The remaining step in the training process is the actual creation of realistic images that include the object tree and the different environmental variations that may occur in the actual assembly workcell. This is the topic of the next section.

SYNTHETIC IMAGE GENERATION

There are two image generation algorithms used to create synthetic images from the CAD model of the assembly. The first is a fast rendering technique that uses only a simple local illumination model and takes advantage of special purpose VLSI hardware for performing geometrical calculations. These draft images are used to identify an optimal camera location and to further refine the object trees used by the inspection algorithm [14]. The location of the camera relative to the assembly under inspection will greatly affect the sensitivity of identifying possible errors. Our approach to computing an optimal camera location is to try and maximize the amount of information in the resulting image that is relevant to evaluating contacts between components. To do this we tag each of the contact surfaces in the error trees with a unique ambient color, with all other surfaces set to black. We then render the assembly using a graphics workstation equipped with a Z-buffer using only the ambient intensity of the polygons. The resulting image contains the number of visible pixels for each surface of interest. Since with hardware support this procedure takes only a fraction of a second, it can be repeated for a large number of potential camera locations. The optimal camera location is selected as that which maximizes the number of visible pixels and surfaces. The information from this image created from the optimal camera location is also used to identify the location and size of the object nodes required by the inspection algorithm. To simplify processing, all object nodes are rectangular, however, a mask is used to identify the regions within the node that correspond to error tree surfaces. Only this region is used in building the statistical model of the node. This prevents irrelevant background information from affecting the sensitivity of the inspection process. An illustration of creating this mask information is provided in Figure 10. These masks are also visible in the templates depicted in Figure 5.

The second rendering technique is used to create more physically realistic synthetic images that are required to build the statistical model of what a correctly assembled product should look like (see Figure 1). To obtain sufficiently realistic
images, light-object interaction must be modeled. Although graphics workstations available today can generate shaded images at video rates, the illumination models used to generate these images typically only deal with the very first reflection from an object’s surface. These so called first-order or local models do not include the effects of light reflecting from several objects or being transmitted through objects. These global models need to be considered in order to obtain more realistic images that can model different types of materials, particularly those that are highly reflective such as polished metals [15].

The only established rendering techniques that attempt to model global lighting effects are ray tracing and radiosity. Because specular effects are very hard to model with radiosity, metallic parts are hard to simulate in images that are rendered using radiosity techniques. As a result, we selected ray tracing as the rendering technique for this application.

**AN EXAMPLE**

We will now illustrate the implementation of the entire assembly error inspection system for the simple example of the pattern wheel presented in Figure 2. First, the CAD model of the assembly is used to create instances of the different components that have simple variations within their allowed tolerances. This is required in order to produce training images that include acceptable levels of variations. Next, as explained in our discussion of CAD database analysis, the liaison diagram of the pattern wheel assembly is generated automatically from surface contact information (see Figure 8). The exploded view of Figure 7 suggests that the pattern wheel assembly can be completely assembled with only a single common axis of insertion. This identifies the pins as potential sources of errors that can occur due to misalignment. From the liaison diagram the system automatically generates the error trees of the pins. These error trees, like the example in Figure 9, are used as the object trees (see Figure 3) which identify the important locations in the images for various types of assembly errors. The number of levels included in the tree is determined by the visibility of the root node, i.e., if the single view used by the inspection algorithm results in the root node being partially occluded then evidence of misassembly should be corroborated by its child nodes to which the error may be propagated.

Ray tracing is then used to generate a group of synthetic images from the assembly CAD model. Determining the optimal number of images to use is the subject of ongoing research, however, for this particular example six images were used. Parameters for the illumination model are specified from the material types of the different pattern wheel assembly components. Light and camera placements are automatically evaluated to provide maximum visibility of contact surfaces representing potential locations of assembly errors. The synthetic images along with the information from the object trees are then passed to the multiscale inspection algorithm for training. Figure 11 illustrates one such synthetic image where the wheel-a2 error tree is shown superimposed onto the image (see Figure 8). The object tree consists of the “boxed in” areas. The upper left corners of the nodes are connected to one another to indicate the connectivity of the tree, and the number of boxes around each node indicates that node’s level in the tree. Note that up to this point no real images of the assembly are required so that this system can also be used to evaluate prospective designs for their inspectability without even building a prototype. Once the system has been trained it is ready to perform its function of inspecting real assemblies. Prior to automated inspection, a single real image of a correctly assembled product must be obtained from the actual inspection station to compensate for...
The algorithm then "zooms in" to look for the pin, guided by the information in the error trees. At this point it detects a mismatch which is indicated in the image by an "X" in the area where a problem was detected.

CONCLUSIONS
This article has discussed the implementation of an assembly inspection system that uses a multiscale algorithm to detect errors in assemblies after being trained on images of correctly assembled products. It has been shown that synthetic images generated by using the CAD models of the assembly and computer graphics ray tracing rendering techniques can be effectively used to train the algorithm. The use of synthetic images has the advantages of simplifying the training process and automating the image selection and the object allocation procedure. In addition, the problem of addressing the different variations that can occur in the assembly workcell is simplified. The use of synthetic images generated directly from CAD models also allows the fine tuning of the inspection algorithm early in the design process thus allowing the assembly and inspection processes to be designed concurrently.

ACKNOWLEDGMENTS
This work was supported by the NEC Corporation, National Science Foundation grant number CDR 8803017 to the Engineering Research Center for Intelligent Manufacturing Systems, National Science Foundation grant number MIP93-00560, and an AT&T Bell Laboratories Ph.D. Scholarship.

REFERENCES
[9] A.A.G. Requicha and S.C. Han, "Representation of geometric features, tolerances, and attributes in solid models based on construc-

Figure 12. A real video image of a defective assembly with the error correctly identified by the inspection algorithm. The "X" in the box identifies the location in the image in which a mismatch with the training images was found thus indicating the missing alignment pin. Note that the slightly different position, orientation, and scale of the real image as opposed to the synthetic training images as indicated by the large bounding box does not adversely affect the robustness of the algorithm.

Figure 13. A real video image of a defective assembly resulting from misplacing the top wheel (wheel-a2). The algorithm detected no errors in trying to locate the assembly and the first level of the wheel-a2 error tree as indicated by the single and double line squares. A defective assembly was identified when the inspection algorithm descended to the second level of the error tree.
Charles A. Bouman received a B.S.E.E. degree from the University of Pennsylvania in 1981, and a MS degree in electrical engineering from the University of California at Berkeley in 1982. From 1982 to 1985, he was a staff member in the Analog Device Technology Group at the Massachusetts Institute of Technology, Lincoln Laboratory. In 1987 and 1989, he received MA and Ph.D. degrees in electrical engineering from Princeton University under the support of an IBM graduate fellowship. In 1989, he joined the faculty of Purdue University where he currently holds the position of Associate Professor in the School of Electrical and Computer Engineering.

Professor Bouman’s research interests include statistical image modeling and analysis, multiscale processing, and the display, printing of images. He is particularly interested in the applications of statistical signal processing techniques to problems such as fast image search and inspection, tomographic reconstruction, color halftoning, and multispectral image segmentation. Professor Bouman has performed research for numerous government and industrial organizations including National Science Foundation, US Army, Hewlett-Packard, NEC Corporation, Apple Computers, Xerox, and Eastman Kodak. From 1991-1993, he was also an NEC Faculty Fellow.

Professor Bouman is a member of IEEE, SPIE, and IS&T professional societies. He has been both chapter chair and vice chair of the IEEE Central Indiana Image Processing and Computer Engineering Chapter. Currently, he is an associate editor of the IEEE Transactions on Image Processing, and a member of the IEEE Image and Multidimensional Signal Processing Technical Committee.

Daniel Tretter received BS degrees in electrical engineering and mathematics from Rose Hulman Institute of Technology in 1987, and a MSE degree from Purdue University in 1988. From 1989 to 1990, he was a member of the technical staff at The Aerospace Corporation. He received his Ph.D degree in electrical engineering from Purdue University in 1994. Dr. Tretter is currently employed at Hewlett-Packard Laboratories in Palo Alto, CA. His current research interests include multiscale image processing and image modeling. He is a member of the IEEE.

Khalid W. Khawaja received the B.S.E.E., M.S.E.E., and Ph.D degrees in electrical engineering from Purdue University, West Lafayette, Indiana, in 1989, 1990, and 1995 respectively. He is currently employed at Structural Dynamics Research Corporation (SDRC) in Cincinnati, Ohio. His primary research interests are in CAD/CAM modeling and image synthesis. He is a member of the IEEE.

Anthony A. Maciejewski received the B.S.E.E., M.S., and Ph.D. degrees in electrical engineering from The Ohio State University, Columbus, in 1982, 1984, and 1987 respectively. In 1985-1986 he was an American Electronics Association Japan Research Fellow at the Hitachi Central Research Laboratory in Tokyo, Japan. Since 1988 he has been with the School of Electrical and Computer Engineering at Purdue University, West Lafayette, Indiana, where he is currently an Associate Professor. His primary research interests center on the simulation and control of kinematically redundant robotic systems.