Scale-Space Methods for Image Feature Modeling in Vision Metrology

C.S. Fraser and J. Shao

Abstract

A vision-based measurement system applicable for use in two-dimensional (2D) industrial inspection is described. The system embodies a five-stage approach to feature modeling comprising the use of multi-scale functions for filtering and feature location, image structure representation through a fusing of filter responses, extraction of edge information, edge building, and a determination of geometric feature attributes. The purpose of feature modeling is to describe and represent feature geometry and dimensions. In an inspection context, information on conformance to design can be provided by using this approach. The application of the system to the tasks of on-line steel plate and wheel rim inspection are described and, following an explanation of the use of scale-space methods for feature modeling, results of experimental tests are summarized.

Introduction

Automated process control and inspection within a manufacturing environment have long been priorities for developments in machine vision, and numerous vision-based assembly line inspection systems are now enjoying routine industrial use. The focus of the overwhelming majority of these systems is upon scene analysis to support object recognition and feature classification, rather than on the recovery of dimensional characteristics of the objects being imaged. Yet, digital photogrammetry offers the potential of extending the capabilities of machine vision systems to encompass the recovery of metric information to high precision through feature modeling.

In this paper the authors describe a quasi real-time vision-based dimensional inspection system, the computational basis of which fundamentally comprises a five stage approach to feature modeling. These stages are (1) convolution of the original image with multi-scale functions to locate feature structures; (2) representation of image structures by fusing multi-resolution filter responses in the form of amplitude; (3) extraction of image edges through a squared gradient approach; (4) edge building using chain codes; and (5) determination of geometric attributes. The motivation for the development of the system was provided by the desire to offer a solution for two practical dimensional inspection problems: the determination of critical dimensional characteristics of rolled steel plating, and the extraction of circularity and concentricity information for automobile wheels.

In the following sections, the measurement tasks are briefly outlined along with their essential characteristics, the imaging configuration for the vision-based inspection system is described, the five-phase feature modeling approach is discussed, and experimental results obtained with the system are summarized. Although the scope of the practical applications in the paper is confined to two specific inspection tasks, the feature modeling approach developed is thought suitable for broader application with only the final phase of the operation — the recovery of geometric attributes — requiring significant task-specific tailoring.

Dimensional Inspection Tasks

Rolled Steel Plating

Figure 1 shows an image of rolled steel slabs in a plate mill. The rolling and cutting process for hot steel plating is by its nature an imprecise operation, yet the final price of the plating is based upon exact dimensional characteristics, that is, parameters such as "good length" (shown as *a* in Figure 1) and "camber," which expresses the degree of curvature. If
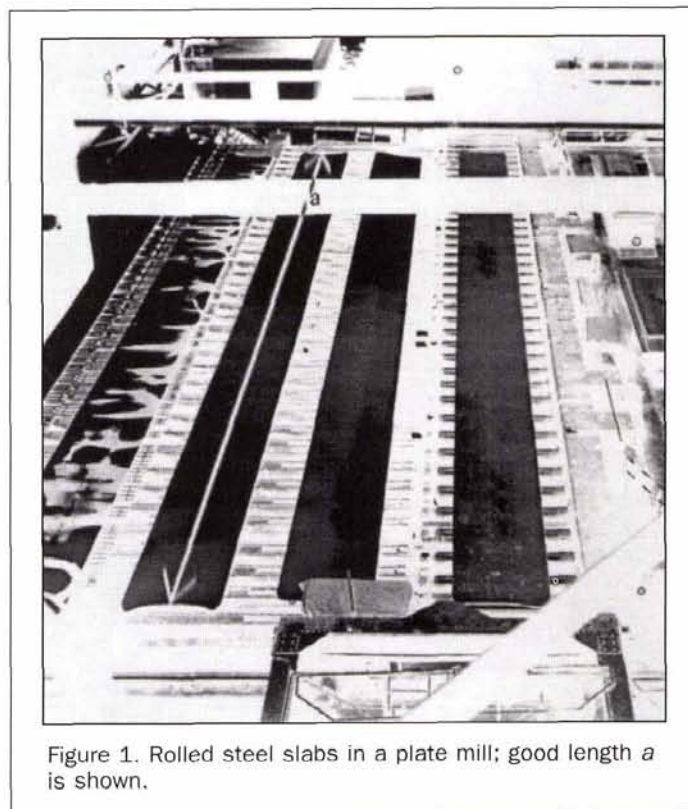

Figure 1. Rolled steel slabs in a plate mill; good length *a* is shown.

Department of Geomatics, The University of Melbourne, Parkville, Victoria 3052, Australia (c.fraser@engineering.unimelb.edu.au).
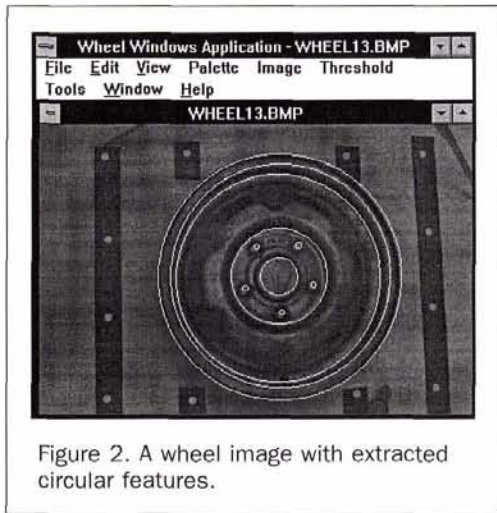
Figure 2. A wheel image with extracted circular features.

shape parameter values fall outside specified tolerances, the slab is withdrawn from production and recycled. The traditional manual inspection of cut steel to the accuracy required (sub-centimetre over as much as 20 m or so) displays a number of shortcomings.

Rolled steel is routinely measured by contact linear measurement methods, but simple dimensions of width and length are by no means sufficient. Rather than being precisely rectangular, the hot steel slabs have an end "profile" which can form after the cut, and they are generally curved to some degree. Quality parameters such as the good length and camber take into account non-rectangular shape, yet determination of these parameters requires that a labor-intensive and time-consuming set of contact measurements be carried out. This off-line process does not facilitate any immediate feed-back to the slab rolling process, and therefore adjustments to the shaping process cannot be effected on line. Plan-view monitors are available but these are very imprecise.

A vision-based approach to dimensional inspection of steel plating would seem to offer a number of advantages, most notably the ability to automatically and fully characterize the essentially two-dimensional (2D) shape of the object and to provide an on-line feedback of dimensional information to support process control.

Wheel Inspection

Figure 2 shows an image of an automobile wheel rim with its prominent circular features located using edge detection and edge building. Dimensional inspection of the wheel needs to encompass a determination of the circularity and diameter of the outer rim and flange surfaces, the circularity of the bolt hole centers, and the concentricity of the best-fit circles.

Once again, the dimensional inspection process is one of extracting 2D information because, although the wheel has depth, this dimension is of less interest. As is suggested by the image in Figure 2, a vision-based inspection approach offers considerable benefits because an automated mapping is afforded for all circular features, with the circularity and concentricity being determined by means of a highly overdetermined least-squares estimation process involving tens to hundreds of edge points rather than from a few isolated contact measurements.

Imaging Configuration

With the emphasis of the inspection tasks being on the recovery of 2D feature information, a basic imaging configura-

tion of a single CCD camera can be adopted, as shown in Figure 3. The feature modeling approach to be described could also be implemented for 3D object extraction, but only following inclusion of an appropriate image matching and stereo model restitution strategy. The present discussion will be restricted to the 2D case in recognition of the nature of the two practical inspection tasks being addressed.

Fundamentally, XY object space coordinate information is obtained from xy image coordinate measurements through the standard relationships of perspective geometry. Projective transformations can be directly applied for planar objects whereas, for the wheel inspection with its different "depth" planes, a collinearity approach is warranted. In both cases, the CCD camera is assumed to be photogrammetrically calibrated and the sensor exterior orientation is assumed to have been determined. Under these conditions, the precision of edge point determination, σ_{XY}, is essentially given as $S\sigma_{xy}$ where S is the scale number (object distance/focal length) and σ_{xy} is the edge detection precision within the image. In general, σ_{xy} is the limiting factor in determining measurement precision, leaving the imaging scale as the main parameter to be optimized.

The precision of recovery of the parameters of the feature modeling can be expected to exceed σ_{XY} as a consequence of the estimation processes involved, with the attainable accuracy being chiefly dependent on the number of located edge points. The emphasis of this paper is mainly upon the feature modeling approach itself and not so much on an optimization of the imaging geometry. Thus, this issue will not be further discussed, although accuracy aspects will be briefly addressed in conjunction with the outline of the practical tests conducted.

Feature Modeling by Scale-Space Methods

Multi-Scale Functions

Robust edge extraction has been carried out using the three steps of (1) a multi-scale representation by convolution of the
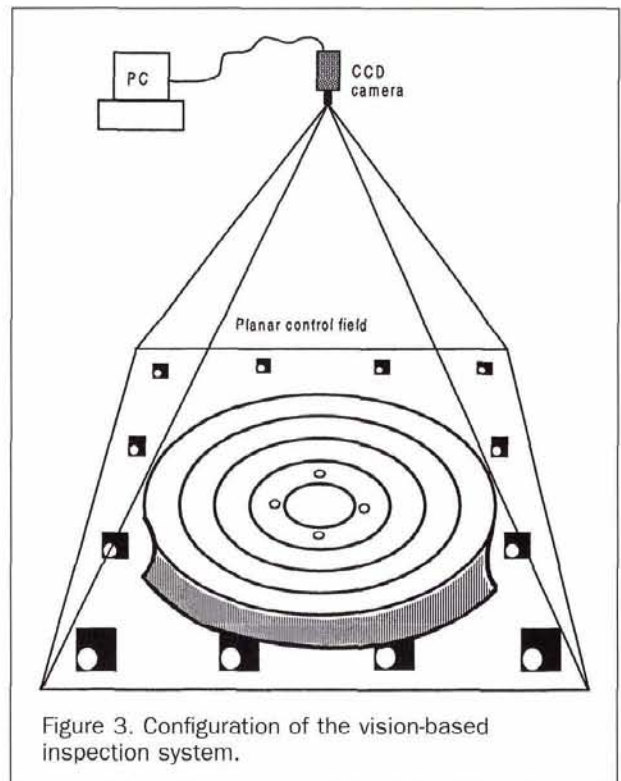

Figure 3. Configuration of the vision-based inspection system.

image with multi-scale functions, (2) a normalization of the intermediate response in each scale and a fusion of the normalized outcomes, and (3) edge extraction using the squared gradient approach (Shao and Foerstner, 1994). Of the range of edge detection techniques proposed over the years, the strategy of employing multi-scale functions is physiologically quite plausible as it is consistent with the idea that filter banks for human vision contain filters which are both orientation and scale sensitive. To provide an insight into the operation of multi-scale edge detection, we consider the simple case of an image of a computer keyboard. Identifiable edges at the coarsest scale comprise the outline of the keyboard. As the resolution becomes progressively finer, first the edges of the different groupings of keys are apparent, and finally the edges of individual keys are identifiable.

The scale of the edge detection can be tuned to highlight specific resolutions, and responses from different scales can be combined. An efficient approach to multi-resolution analysis is using wavelets (e.g., Mallat, 1989) which afford image decomposition by means of a recursive process of low-pass and high-pass filtering. Adoption of the multi-scale approach can lead to a more efficient suppression of noise, and various categories of smooth edges can thus be more efficiently detected (e.g., Witkin, 1983). Moreover, the method avoids the shortcoming of delocalization of edges at higher scales which can occur with single-scale filter functions.

The first derivatives of the Gaussian function (Li and Shao, 1994) and the Gabor function (Biguen and du Buf, 1994) have been employed for this project as multi-scale filters for both untextured and textured image edges, with the cardinal frame of multi-scale edge detection (Shao and Foerstner, 1994) being applied for the representation of local image structures. More emphasis has been given to the Gaussian function because it has some notable advantages. These include the fact that convergence is rapid; the integral of the first derivative is equal to zero; the scaling behavior is perfect, thus allowing most types of edges to be efficiently detected; the energy is concentrated, which is an essential property for an excellent filter; and it is a computationally elegant, explicit expression. With the Gaussian filter, different scales have been adopted for the variances in the two image-coordinate axes to provide an improved mask-image matching effect for elliptical image features. Also, the filter response normalization at the different scales is carried out using a partially analytical expression rather than through the use of simulated noise which was the approach adopted in Shao and Foerstner (1994).

Use of the Gaussian filter for edge extraction usually follows one of two approaches, the first utilizing an image convolution by means of second derivatives of the function (i.e., zero crossings), and the second employing first derivatives (Li and Shao, 1994). In the latter approach, which has been adopted here, first derivatives of the Gaussian function in the x and y image coordinate directions are employed for the computation of edge gradients which are then translated to strength of normal edges represented by a squared gradient (Shao and Foerstner, 1994). The first derivatives, Gs_x in x and Gs_y in y are expressed by

$$
\begin{aligned}
Gs\_x &= -\frac{x}{\sigma_x \sigma_y} Gs, \\
Gs\_y &= -\frac{y}{\sigma_x \sigma_y} Gs,
\end{aligned}
\tag{1}
$$

where σ_x and σ_y are scales in the x and y directions, and Gs is the Gaussian function

$$
Gs = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{x^2 + y^2}{2\sigma_x\sigma_y}\right).
\tag{2}
$$

Robust edge detection has to be taken into account in cases where heavy texture and excessive noise are present in the image. For this purpose, the Gabor function Gb, which is orientation and scale sensitive, is also adopted: i.e.,

$$
Gb = Gs \times e^{j2\pi(Ux+Vy)}
\tag{3}
$$

where $U = \cos\theta/\sigma_x$ and $V = \sin\theta/\sigma_y$, with θ being an arbitrary angle in the x-y plane.

After the filter functions are constructed, the filter response R can then be derived through an image transform operation: i.e.,

$$
R = \iint_{x,y} G(x - u, y - v) f(x,y) \, dxdy
\tag{4}
$$

where G is the filter, Gs_x, Gs_y, or Gb, with the amplitude or magnitude A of the filter response being further used for edge representation: i.e.,

$$
A = |R| \text{ or } A = \sqrt{(R\_x)^2 + (R\_y)^2}
\tag{5}
$$

where R_x and R_y are the filter responses of the first derivatives of the Gaussian function in the x and y directions.

Fusion of Filter Responses and Edge Extraction

The squared gradient (SG) has been employed for the measurement of local image characteristics in view of its demonstrated performance for line and point feature extraction (e.g., Shao and Foerstner, 1994): i.e.,

$$
\Gamma A_i = \nabla_s A_i \cdot \nabla_s^T A_i
\tag{6}
$$

where ΓA_i expresses a filter response for scale number i, with the gradient function $\nabla_s A_i$ of scale i being given for a smoothing kernel g_s by the convolution

$$
\nabla_s A_i = \nabla g_s * A_i = \left(\frac{\partial}{\partial x} g_s \, \frac{\partial}{\partial y} g_s\right)^T * A_i.
\tag{7}
$$

In order to make full use of responses resulting from the SG at each of the multiple scales, information fusion is adopted whereby the SG is fused through the following weighted sum, which avoids further non-linearity:

$$
\Gamma A = \sum_i \Gamma A_i \bullet W_i
\tag{8}
$$

Two options for the weights W_i are implemented; the first expressed in terms of the noise gradient response of the ith Gabor filter: i.e.,

$$
W_i = 1/\sigma^2_{\Gamma A_i}.
\tag{9}
$$

The first part of Equation 5 is tuned to be

$$
A = \frac{N}{\sigma_x\sigma_y} |R|, \text{ with } N = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |R^2| \, dxdy
\tag{10}
$$

so that the normalization $\text{Var}(\sigma^2_{\Gamma A})$ leads to $\text{Var}(\Gamma A_i) = $ constant. For the first derivatives of the Gaussian filter, the value $\text{Var}(\sigma^2_{\Gamma A})$ is evaluated analytically. This approach illustrates that any filter response is natural, i.e., the weak responses are not enlarged and those that are strong are not reduced. All weak responses will thus be merged into the strong responses.

This situation is different from the second weighting scheme for fusion, under which the weight is chosen according to the range of ΓA: i.e.,

$$
W_i = 1/max_{xy} \{\Gamma A_i(x,y)\}
\tag{11}
$$

assuming that $min \{\Gamma A_i(x,y)\} = 0$. The method is more straightforward, because it is independent of the different types of filters applied (e.g., Gabor and Gaussian). Compared
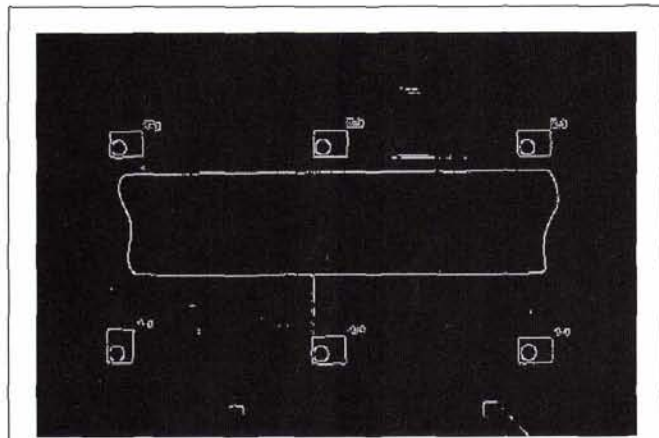
Figure 4. Edges extracted from an image of a model steel plate.

Figure 5. Edges extracted from a wheel image.

with the first approach, the second will enlarge the filter response for both edges and low noise. In the case of both texture and non-texture in the same image, fusion of the responses of both the Gaussian filter and the Gabor wavelets through another weight W_t can be performed in accordance with the noise response for the SG, indicated as

$$\Gamma A = \sum_i \Gamma A_i \bullet W_i \bullet W_t. \qquad (12)$$

The weight W_t can be determined depending on image content. For instance, W_t may be assigned a value less than 1 if the image is free of normal discontinuities, or it may be allocated a large value where normal discontinuities are present and texture is lacking. Where texture and non-texture exist simultaneously, W_t can be assigned a value of unity (e.g., for the Gabor filter). The edge strength E can be secured from the final SG value

$$E = tr\Gamma A \qquad (13)$$

where tr is the trace operator which is employed to obtain the total energy of the image function near the edges. The energy maximum $E \to max$ is used to finally derive edge pixels, with the detected edges having the property of being one-pixel wide, which is necessary for the subsequent recognition process.

Shown in Figures 4 and 5 are the edges extracted from an image of a model steel plate and from a wheel image, respectively, using the approach of convolution by means of multi-scale functions followed by a fusing of filter responses and edge extraction by the SG approach. The samples shown are two of a total of 20 test images with varying degrees of contrast, texture, and imposed noise which were examined as part of the project. Overall, the edge detection approach adopted proved to be quite robust and reasonably fast, consuming about four seconds on a Pentium PC for each of the 1.5K by 1K Kodak DCS420 digital camera images, with the time being dependent on image size and the number of scales used.

Edge Point Connection

Once all candidate edges have been detected, the next task in the feature modeling process is to "build" the one-pixel wide edges through point connection by chain coding. The polygons so generated can then be examined through shape analysis to determine whether they form a feature of interest, for example, circles in the wheel inspection case. Actual implementation of the chain coding can be either automatic or
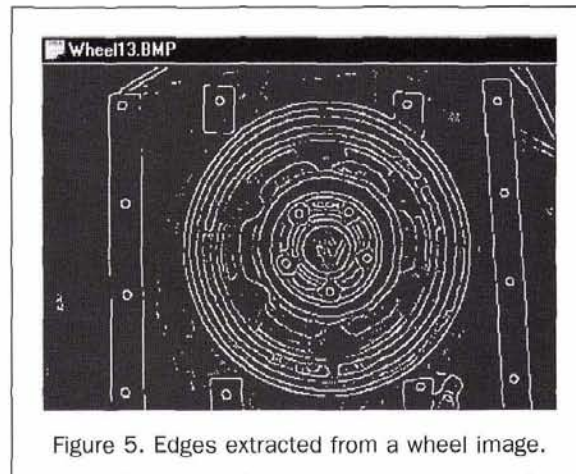
semi-automatic, depending on the nature of both the edges and the polygons. With both approaches, uninterrupted edges are built using an edge linking that considers an 8-connected search path.

The search path commences at the position a in Figure 6 which is adjacent to the edge point o. The edge can then lead through any of the eight positions a, b, ..., h, depending upon the position of the next detected edge point, which is selected using a sequential search. Once the next edge point is found, the position o is reclassified as a non-edge pixel, the next edge point assumes that o position and the process continues with the result that connected points form several chains of codes.

In the case of the steel plates, which are assumed to have only perimeter edges, a region growing approach can be used to complement the edge-linking process. Figure 7 gives a representative example of linked edges for an image of a steel plate, whereas Figure 2 provides a corresponding image for a wheel rim.

Feature Geometry

The final process in the feature modeling operation is the extraction of critical dimensions and shape characteristics for the feature(s) of interest. In the case of the steel plates, geometric attributes of interest center on the camber and good length, which necessitate extraction of both the near parallel, straight (or gently curved) long edges of the rolled steel plate, as well as the shape profiles at each end. For the detection of "straight" lines in the 2D image, the Hough Transform (e.g., Pao et al., 1992) is employed followed by
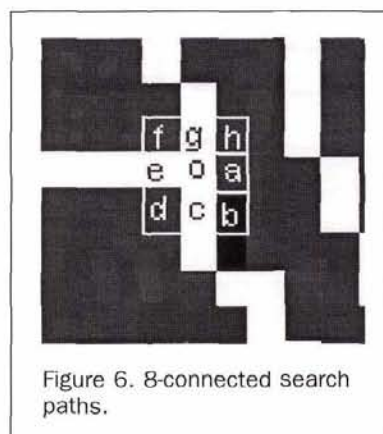
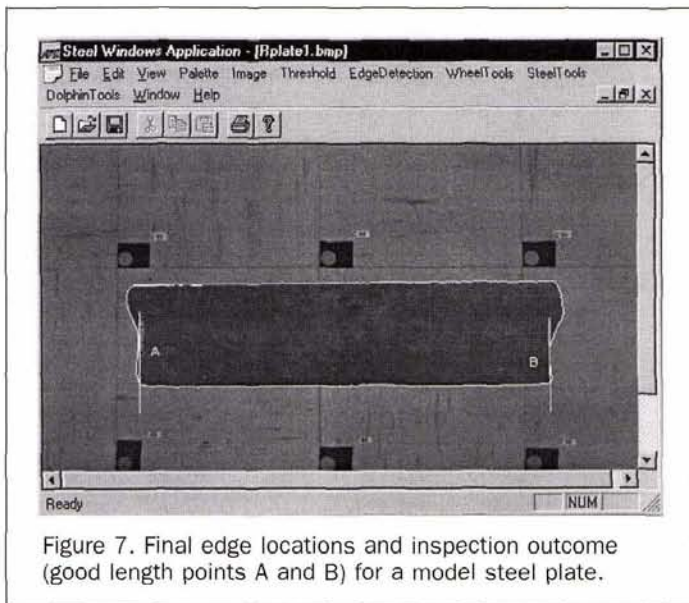

Figure 6. 8-connected search paths.

Figure 7. Final edge locations and inspection outcome (good length points A and B) for a model steel plate.

computation of the largest rectangle that can be fully contained within the extracted edge polygon. The length of the rectangle then defines the good length, and the camber is determined by the mean lateral displacement of the two edges from the best-fitting rectangle at the mid-point of the steel plate.

In the process of determining wheel circularity and concentricity, a best-fitting shape determination is again employed, this time utilizing a standard least-squares circle fitting procedure with *a priori* constraints to assist in the recovery of only the desired circular features. The parameters of interest for each circle, including the one through the bolt holes, are the coordinates of the center point and the mean radius. Concentricity can then be quantified in terms of differences in the coordinates of the best-fitting center points, while size and shape information comes directly from the computed radii and departures from circularity at all measured edge points. Biases in bolt-hole positions can also be determined and compared to specified tolerances. Figure 8 illustrates graphically both the best-fitting circle results and the outcomes of the concentricity determination (for the outer rim and bolt holes). The accuracy of recovery of these geometric parameters is naturally a function of both the accuracy of edge-point determination and the number of edge points.

Summary of Experimental Testing

In order to examine the robustness, accuracy, and general practicability of the multi-scale feature modeling approach adopted, an initial series of test measurements was performed using 15 images of three different wheels and five images of steel plates, both model plates and actual slabs in a plate mill (Figure 1). Experimental testing confirmed edge detection accuracy of 1 pixel, though it must be recalled that sub-pixel parameter estimation is possible in the subsequent least-squares computation of dimensional characteristics. Figure 7 depicts the final edge locations extracted for one of the model steel plates. This plate is 602 mm long ("good length") and it was imaged at a scale of 1:180 with a Kodak DCS420 still-video camera fitted with a 20-mm lens. A precision (RMS 1 sigma) of 1.6 mm could be anticipated for a single edge point in this application, yet the accuracies of dimensional data extracted turned out to significantly exceed this level. Multiple determinations of good length and plate

width revealed accuracies of 0.5 mm or better, from which edge modeling to the 0.3-pixel level could be inferred. In any application of the vision-based inspection approach to full size plates of 10 m or more in length, it is clear that, if sub-centimetre accuracy levels were to be maintained, a high resolution CCD camera would be necessary.

In order to quantify the accuracy of the wheel inspection process, results obtained in the vision-based inspection were compared to actual dimensions measured to 0.05-mm accuracy with the V-STARS/M real-time digital photogrammetry system from Geodetic Services, Incorporated (Brown and Dold, 1995). Anticipated edge-point precision from the 1:70-scale DCS420 imagery for the wheel rim shown in Figure 8 was $\sigma_{XY} = 1$ pixel in object space, or close to 0.6 mm. The precision of circle parameters derived from the edge points is then a function of the number of such points. For a sample size of 24, for example, the center-point coordinates and radius can be determined to a precision (RMS 1 sigma) of 0.3 pixels or 0.2 mm.

A comparison of the results of concentricity determination and of computations of radii yielded accuracies consistent with expectations. For example, the "true" offset between the centers of the best fitting circles to the bolt holes and the outer rim was 0.44 mm, whereas the feature modeling approach yielded a figure of 0.68 mm. The level of agreement between determinations of best-fitting radius is influenced by the degree of circularity of the rim in question. In the test of the wheel shown in Figure 8, an RMS departure of circularity of 1.6 mm was obtained from 24 edge points, yet the agreement between the computed radius values from the feature modeling approach and from direct measurement with V-STARS/M still amounted to only 0.4 mm or close to 0.7 pixels.

Concluding Remarks

The results of the practical testing of the vision-based dimensional inspection system developed have shown that the *multi-scale feature modeling* approach is both robust and accurate, even in the presence of heavy noise. The emphasis of the work conducted so far on scale-space methods for image feature modeling by the authors has been primarily on the extraction and building of edges and the subsequent feature modeling, and secondarily upon a full metric evaluation
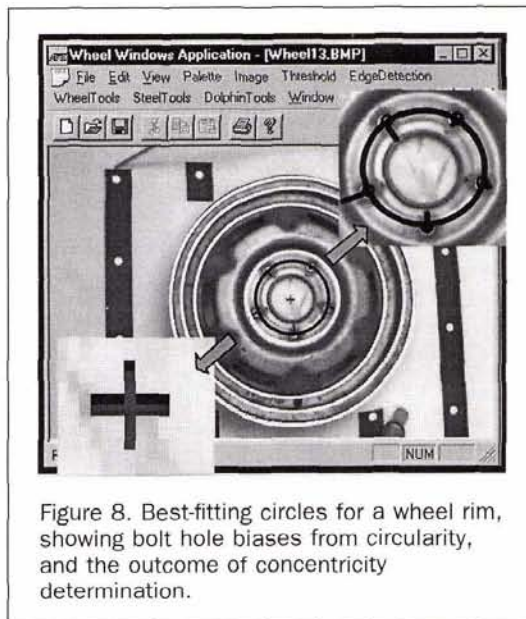

Figure 8. Best-fitting circles for a wheel rim, showing bolt hole biases from circularity, and the outcome of concentricity determination.

of the approach. Nevertheless, the experimental test program conducted has indicated that, in both the applications areas of steel plate and wheel rim inspection, accuracies attained have been consistent with design expectations.

What has been demonstrated in this paper is that multi-resolution modeling and approximation provide a viable foundation upon which to develop efficient methods of image feature modeling. Multiresolution approximation techniques are, coincidently, also currently attracting research attention in other areas of geomatics, notably in gravity field modeling, digital elevation model estimation, and ionospheric mapping (Li and Schwarz, 1996).

Acknowledgment

References

Biguen, J., and J.M.H. du Buf, 1994. N-folded Symmetries by Complex Moments in Gabor Space and Their Application to Unsupervised Texture Segmentation, *IEEE Trans. PAMI*, 16(1):80–87.

Brown, J.D., and J. Dold, 1995. V-STARS — A System for Digital Industrial Photogrammetry, *Optical 3-D Measurement Techniques III* (A. Gruen and H. Kahmen, editors), Wichmann Verlag, Heidelberg, pp. 12–21.

Li, D., and J. Shao, 1994. Wavelet Theory and Its Application in Image Edge Detection, *ISPRS Journal of Photogrammetry and Remote Sensing*, 49(3):4-11.

Li, Z., and K.P. Schwarz, 1996. Multiresolution Approximation in Geomatics Engineering, *Proc. Geoinformatics '96 Wuhan*, Wuhan, China, 16–19 October, 1:264–271.

Mallat, S.G., 1989. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation, *IEEE Trans. PAMI*, 11(7):674–693.

Pao, D., H. Li, and R. Jayakumar, 1992. Shape Recognition Using the Straight Line Hough Transform, Theory and Generalization, *IEEE Trans. PAMI*, 14(11):1076–1089.

Shao, J., and W. Foerstner, 1994. Garbor Wavelets for Texture Edge Extraction, *International Archives of Photogrammetry and Remote Sensing*, 30(3):745–752.

Witkin, A.P., 1983. Scale Space Filtering, *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, Karlsruhe, 8-12 August, pp. 1019–1022.

Call for Nominations
for the Silver Anniversary Presentation of the
WILLIAM T. PECORA AWARD

The William T. Pecora Award is presented annually to recognize outstanding contributions by individuals or groups toward the understanding of the Earth by means of remote sensing. This year marks the 25th anniversary presentation of the award, which is sponsored jointly by the Department of the Interior (DOI) and the National Aeronautics and Space Administration (NASA).

The award was established in 1974 to honor the memory of Dr. William T. Pecora, former Director of the U.S. Geological Survey, and Under Secretary, Department of the Interior. Dr. Pecora was a motivating force behind the establishment of a program for civil remote sensing of the Earth from space. His early vision and support helped establish what we know today as the Landsat satellite program.

The award consists of a citation and plaque which are presented to the winner at an appropriate public forum by the Secretary of the Interior and the NASA Administrator or their representatives. The name of the recipient is also inscribed on permanent plaques which are displayed by the sponsoring agencies.

ELIGIBILITY

Any individual or group working in the field of remote sensing of the Earth is eligible to receive the William T. Pecora Award.

An **individual award** recognizes scientific and technical remote sensing achievements, as well as contributions leading to successful practical applications. Consideration will be given to sustained career achievements or singular contributions of major importance to the field of remote sensing.

A **group award** recognizes an organization, or part of an organization, that has made major breakthroughs in remote sensing science or technology, or developed an innovative application, that has a significant impact on the user community or national and international policies. Group achievements should be documented in contributions to the open literature.

NOMINATION PROCEDURE

Nominations may be made by an individual, organization, or professional society. Nominations for the 1998 award must be received by the Committee no later than **June 30, 1998**. For nomination procedures, please contact:

PECORA AWARD COMMITTEE
U.S. Geological Survey
590 National Center
Reston, VA 20192-0001
703-648-4519