Universität des Saarlandes

Fachrichtung 6.1 – Mathematik

Preprint Nr. 192

Robust Automated Multiple View Inspection

Luis Pizarro, Domingo Mery,
Rafael Delpiano and Miguel Carrasco

Saarbrücken 2007

Fachrichtung 6.1 – Mathematik
Universität des Saarlandes

Preprint No. 192
submitted: April 19, 2007

Robust Automated Multiple View Inspection

Luis Pizarro
Saarland University
Department of Mathematics
P.O. Box 15 11 50
66041 Saarbrücken
Germany
pizarro@mia.uni-saarland.de

Domingo Mery
Pontificia Universidad Católica de Chile
Departamento de Ciencia de la Computación
Av. Vicuña Mackenna 4860
Santiago
Chile
dmery@ing.puc.cl

Rafael Delpiano
Pontificia Universidad Católica de Chile
Departamento de Ciencia de la Computación
Av. Vicuña Mackenna 4860
Santiago
Chile
delpiano@ing.puc.cl

Miguel Carrasco
Pontificia Universidad Católica de Chile
Departamento de Ciencia de la Computación
Av. Vicuña Mackenna 4860
Santiago
Chile
mlcarras@puc.cl

Edited by
FR 6.1 – Mathematik
Universität des Saarlandes
Postfach 15 11 50
66041 Saarbrücken
Germany

Fax: + 49 681 302 4443
e-Mail: preprint@math.uni-sb.de
WWW: http://www.math.uni-sb.de/

Abstract

Recently, Automated Multiple View Inspection (AMVI) has been developed for automated defect detection of manufactured objects. That framework was successfully implemented for calibrated image sequences. However, it is not easy to implement in industrial environments because the calibration is a difficult and unstable process. To overcome these disadvantages, we propose the robust AMVI strategy which assumes that an unknown affine transformation exists between each pair of uncalibrated images. This transformation is estimated using two complementary robust procedures: a global approximation of the affine mapping is computed by creating candidate correspondences via B-splines and selecting those which better satisfy the epipolar constraint for uncalibrated images. Then, we use this approximation as initial estimate of a robust intensity-based matching approach, which is applied locally on each potential defect. The result is that false alarms are discarded, and the defects of an industrial object are actually tracked along the uncalibrated image sequence. The method is successful as shown in our experiments on aluminum die castings.

Keywords: Automated visual inspection, uncalibrated images, image matching, sequence tracking, robustness, X-ray imaging, radioscopic imaging system.

1 Introduction

Recently, the *Automated Multiple View Inspection* (AMVI) approach was developed for automated defect detection [1]. This method is able to detect defects in two steps. In the first step called *identification*, potential defects are automatically identified in each image of the sequence using a single filter without any prior knowledge of the test object. The second step, called *tracking*, attempts to track the identified potential defects along the image sequence. As a result, only existing defects (and not the false detections) are successfully tracked in the image sequence because they are located in positions dictated by the motion of the test object. The preliminary results obtained using AMVI methodology are promising for calibrated image sequences. However, this approach is not suitable in all industrial applications, because calibration is a difficult process, and vibrations of the imaging system may induce inaccuracies in the estimated parameters of the multiple view geometric model. Thus, calibration is not stable and the imaging system must be re-calibrated periodically. A simple method was proposed in [2] to inspect objects using uncalibrated image sequences, where structural

points are used to track the potential defects in the sequence using bifocal constraints. The method achieves good performance in some sequences, but fails when the structure points cannot be matched. In this case the estimation of the fundamental matrix is incorrect, and therefore the tracking also fails.

Following the concept of camera multiplicity or multiple views, a reconfigurable array for machine inspection (RAMVI) was proposed in [3], where the calibration process requires manual intervention. The authors remark the importance of the calibration for accurate inspection and propose a methodology to perform it automatically [4]. The goodness of using multiple views is also described in [5], where a visual inspection system that uses a single camera and mirrors for simulating multiple cameras is proposed. A suitable pattern object is used to find the camera parameters before combining all views.

Calibration might be a extremely complicated procedure for real-time applications and manufacturing systems that cannot be halted for calibration purposes. Therefore, we aim to perform visual inspection avoiding an impracticable, expensive, and/or time consuming calibration process. To overcome these drawbacks, in this paper we propose a new approach for automated visual inspection that can be directly applied on uncalibrated image sequences of multiple views. To deal with the geometric distortions we assume that an unknown affine transformation exists between every pair of consecutive or non-consecutive images. We formulate the search for this affine mapping as a robust local estimation problem using an intensity-based matching approach. In implementing a good tracking algorithm, it is important to put special effort into finding the first (global) matching between every pair of images. This match is used to provide the initial estimate of the local optimisation process applied on each potential defect, which is crucial for convergence. This is why another robust procedure is introduced at this step, which takes advantage of the geometric characteristics of the object being inspected. Using the RANSAC algorithm [6] we select the best candidate correspondences created via B-splines which better satisfy the epipolar constraint for uncalibrated images

The rest of the paper is organised as follows: Section 2 explains our robust approach for uncalibrated AMVI. Section 3 shows preliminary results obtained with the proposed methodology. Finally, Section 4 delineates the concluding remarks and perspectives for future works.

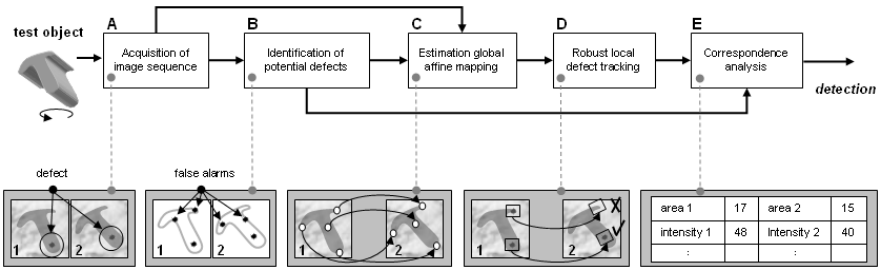

Figure 1: Block diagram of the proposed *robust automated multiple view inspection* system.

2 Proposed Method

Our proposed framework for automated visual inspection consists of 5 steps (A to E), which are outlined in Fig. 1. Before describing each step in detail, a brief introduction to each one is given a continuation.

X-ray imaging systems are widely employed in nondestructive testing. They are particularly useful in automotive and aerospace industries for detecting different types of flaws: porosity, cracks, corrosion, inclusions, debris, rivets, thickness variations, among others [7, 8, 9, 10, 11]. Employing X-rays exploits the fact that most of the material defects are not visible. However, even in radioscopic images the signal-to-noise ratio (SNR) is low, so as the flaw signal is slightly greater than the background noise. For this reason, the identification of real defects with poor contrast can involve detection of false alarms as well. In some applications[1] one view is probably enough for examining material defects. However, the robustness of the inspection process can be increased when redundant information is used to validate flaw detection. Thus, two or more views of the same object taken from different viewpoints can be used to confirm and improve the diagnostic done by analysing only one image. This is a useful and powerful alternative for examining complex objects where uncertainty leads to misinterpretation. A similar idea is also used by radiologists that analyse two different X-ray views of the same breast tissue to detect cancer at early stages. See for example [12], where the proposed method finds automatically correspondences in two views. Detection of delamination defects in rocket boosters is another example recommending the use of radiographic sequences [13]. Section 2.1 explains how the uncalibrated radioscopic image sequence used in our experiments is obtained (Fig. 1 - block A).

[1]for instance in printed circuit board (PCB) inspection.

Once the image sequence is acquired, we search for potential defects on each view. Due to the low SNR of the images, detection of false alarms is likely. However, the detection of the real defects must be ensured in order to make the subsequent tracking possible. Potential defects are segmented and their features extracted in order to match them in a posterior correspondence analysis stage. Section 2.2 shows how the identification of potential defects is performed (Fig. 1 - block B).

Going one step further, we postulate that only real defects can be followed along the image sequence, and logically false alarms discarded. Nevertheless, the uncalibrated imaging system generates images perturbed by geometric distortions, what makes any attempt to search for corresponding defects in two or more views cumbersome. To deal with this problem, we model geometric distortions as affine transformations[2]. Let $\mathbf{H}$ be a non-singular 3×3 matrix defining an affine mapping from all the homogeneous points m_i in one view to the points m'_i in another view, i.e. $m'_i = \mathbf{H}\ m_i$. Three non collinear corresponding points form the following linear system of equations [14, chap.9]

$$\underbrace{\begin{bmatrix} x'_1 & x'_2 & x'_3 \\ y'_1 & y'_2 & y'_3 \\ 1 & 1 & 1 \end{bmatrix}}_{\mathbf{M'}} = \mathbf{H} \underbrace{\begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix}}_{\mathbf{M}}, \tag{1}$$

from which $\mathbf{H}$ can be computed as $\mathbf{H} = \mathbf{M'M^{-1}}$. Additional corresponding points allow a more accurate approximation of the mapping. For n points we compute

$$\mathbf{H} = \mathbf{M'}\,\mathbf{M}^T(\mathbf{M}\,\mathbf{M}^T)^{-\mathbf{1}}. \tag{2}$$

If some unequivocal corresponding points between two consecutive or non-consecutive views were known, the problem of matching potential defective regions between these images would be solved by applying the mapping $\mathbf{H}$ to find corresponding coordinates for those regions, and comparing their extracted features obtained in the previous identification step. However, such corresponding points are not known. To find a reliable subset of such points, we apply the robust methodology RANSAC in selecting the best candidate points created via B-splines that satisfy the epipolar constraint. With the resulting points, a global mapping between two views is estimated using the equation (2). Section 2.3 describes this procedure in detail (Fig. 1 - block C).

The preceding stage presents a reliable mechanism to obtain a global approximation of the geometric mapping between two uncalibrated views. Further-

[2]In this paper we use affine transformations, although it is also possible to implement perspective transformations.

more, it is feasible to consider that the geometric distortion is nonuniform over the entire image. In fact, potential defects may be located in different parts of the image, where a slightly different distortion was induced by the uncalibrated imaging system. Thus, we want to estimate these local deformations considering the previous global computation as an initial local approximation. Therefore, we formulate the search of each potential defect from one view to another as an intensity matching problem, where the intensities of the potential defect in the first view are to be iteratively tracked in the second view. Starting from the global affine transformation a local affine transformation for every potential defect is refined in each iteration. Moreover, it is possible to strengthen this process against illumination variations and partial occlusions[3] looking at the robust formulation of the visual matching problem. Section 2.4 details how the tracking of potential defects is performed (Fig. 1 - block D).
The coordinates of every potential defective region are tracked from the first view to the second one where, in the best case, another potential defective region with similar feature was also found during the identification stage. As a direct result of the tracking process, three criteria must be fulfilled to consider a region as defective: i) identifiability, ii) spatial proximity, and iii) feature proximity. Section 2.5 specifies the correspondence analysis carried out to verify the fulfillment of such requirements (Fig. 1 - block E).
The proposed approach seems to be complicated, which is true from the computational point of view. However, the inspection itself is quite simple because the test object does not require placement accuracy; we only need to place and rotate the object, the rest being done by computer automatically. The bottom part of Fig. 1 shows a synthetic example where any two views of a test object are inspected. The object contains only one defect, but false alarms may appear. Ideally, the inspection system detects the flaw and discards all false alarms.

2.1 Acquisition of the Image Sequence

In order to facilitate the defect tracking over the images, similar projections of the inspected object must be registered along the sequence. For this reason, and for simplicity, different views are taken by rotating the casting at small angular intervals (e.g. 5°, see Fig. 2). Each captured scene consists of only one rigid object in motion, whose 2D trajectories are smooth because there is no significant frame-to-frame motion, the velocity of the test object

[3] Occlusions appear when small flaws move in front (or behind) of a thick cross section of the casting, where X-rays are highly absorbed; and when flaws are located in the outer limits of the visible area of the casting.

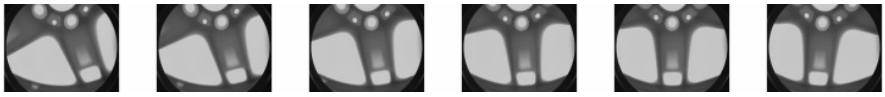

Figure 2: Segment of a real image sequence used in the experiments. Each frame is rotated 5° in the die casting.

is constant, and the motion of the test object is generally only rotational or translational. Since many images are captured, the time of the data acquisition is reduced by taking the images without frame averaging[4]. The usual set up utilised for *automatic visual inspection* (AVI) on aluminum die castings is detailed in [1].

2.2 Identification of Potential Defects

The identification of potential defects aims at segmenting regions that may correspond to real defects. Two general characteristics of the defects are used for identification: i) a defect can be considered as a connected subset of the image, and ii) the gray level difference between a defect and its neighborhood is significant. The potential defects are identified without prior knowledge. First, a Laplacian-of-Gaussian (LoG) kernel and a zero crossing algorithm are used to detect edges on the X-ray image. In real defects, the resulting binary edge image should produce closed and connected contours which outline *regions*. However, a defect may not be perfectly enclosed if it is located at one edge of a regular structure as shown in Fig. 3c. In order to complete the remaining edges of these defects, a thickening of the edges of the regular structure is performed as follows: a) the gradient of the original image is calculated (see Fig. 3d); b) by thresholding the gradient image at a high gray level, a new binary image is obtained; and c) the resulting image is added to the zero crossing image (see Fig. 3e). Afterwards, each closed region is segmented. In order to identify the potential defects, features are extracted from crossing line profiles of each segmented region. Crossing line profiles are gray level profiles along straight lines that cross each segmented region in the middle. If the variance of the crossing line profiles is high, the segmented region is classified as potential defect [15, 16]. Later on, the extracted features are used in the stage of correspondence analysis (Section 2.5) to match tracked potential defects. This is a very simple detector of potential defects,

[4]Digital radioscopic images are generated using a frame grabber, which averages n samples of the scene taken at infinitesimal time intervals in order to reduce noise and improve the signal-to-noise ratio.

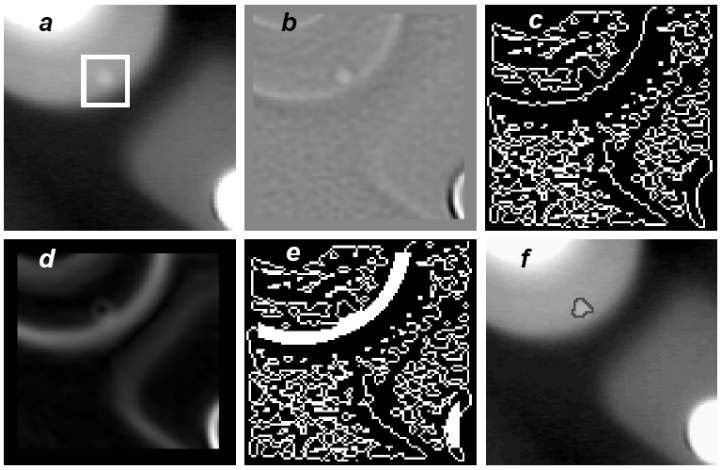

Figure 3: Detection of flaws: a) radioscopic image with a small flaw at an edge of a regular structure, b) Laplacian-filtered image with $\sigma = 1.25$ pixels (kernel size $= 11 \times 11$), c) zero crossing image, d) gradient image, e) edge detection after adding high gradient pixels, and f) detected flaw using the variance of the crossing line profile.

whereby current detection rate of real defects is more than 85%.

2.3 Estimation of the Global Affine Mapping

At this stage we look for a global approximation of the affine mapping **H** between any two different views. As suggested in equation (2), such a transformation can be accurately estimated from a set of n corresponding points. This set of points can be found by performing the following five steps:

1. **Segmentation**. It consists of isolating object parts in which the intensity values are clearly distinguishable from the background. We use the Otsu's segmentation method [17] for this task, which estimates the best separation for bimodal histograms. See Fig. 4.

2. **Feature extraction**. For every segmented region three features are extracted: area; centre of mass $(\bar{i}, \bar{j}) = (\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}})$, in terms of the statistic moment of order $(r+s)$ $m_{rs} = \sum_{(i,j)\in\Omega} i^r j^s$, where Ω is the set of pixels of the segmented region; and the group of four affine moment invariants derived by Flusser and Suk [18].

3. **Region matching**. This step establishes correspondences among segmented regions by measuring their similarity. The smallest norm of

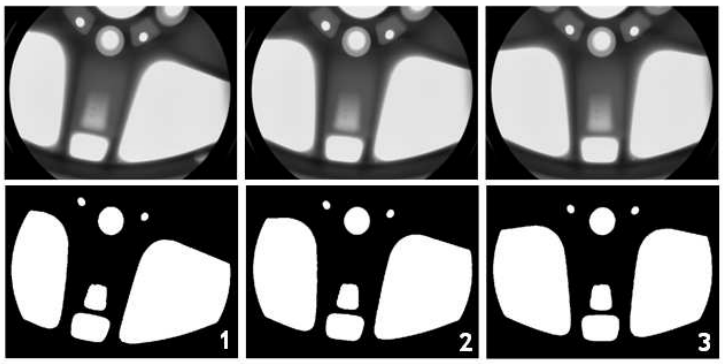

Figure 4: *Top:* Three views of our real image sequence. *Bottom:* Otsu's segmentation method applied on each view.

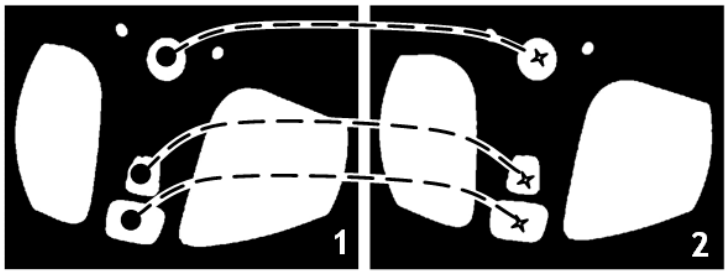

Figure 5: Result of matching regions according to their similarity.

the difference between the normalised feature vectors of two regions in different images is used to label those regions as corresponding. See Fig. 5. In accordance with how the image sequence is generated, it is plausible to consider that corresponding segmented regions in two consecutive frames have similar shapes, except for correspondences that run out of the limits of the visible area of the casting.

4. **Introducing artificial points**. The corresponding centres of mass found in the previous step can be used to compute the mapping $\mathbf{H}$ as in (2). However, in practice we need more correspondences to improve the accuracy of such a computation. We increase the number of matches by interpolating artificial points among the centres of mass via B-splines[5]. The Cox-de Boor's recursive formulation of B-splines can be found in [19]. We use cubic B-splines for knots $t_i \in [0, 1]$ with four control points

[5]B-splines are invariant under affine transformations. In practice, linear splines can also be utilised with enough number of knots.

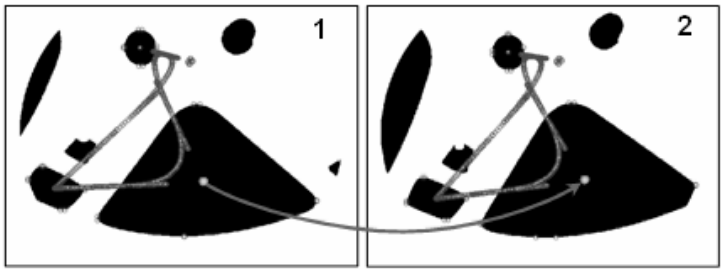

Figure 6: The corresponding points are taken from the centres of mass of complete regions. Artificial corresponding points are added by using B-spline curves that join the determined centres of mass.

$\{P_{-2}, P_{-1}, P_0, P_1\}$. Its matrix form is given by

$$B(t_i) = \left[\begin{array}{cccc} t_i^3 & t_i^2 & t_i^1 & 1 \end{array}\right] \frac{1}{6} \left[\begin{array}{cccc} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{array}\right] \left[\begin{array}{c} P_{-2} \\ P_{-1} \\ P_0 \\ P_1 \end{array}\right]. \quad (3)$$

Varying the number of knots $(t_1, \ldots, t_k)$ among the control points (centres of mass) we regulate the set of artificial points which act as candidate corresponding points. See Fig. 6.

5. **Selection of corresponding points**. According to the *principle of multiple view geometry* [6], all corresponding coordinates between two views are related by the *fundamental matrix*[6] $\mathbf{F}$, such that

$$m_i'^T \ \mathbf{F} \ m_i = 0. \quad (4)$$

This relation is known as *epipolar constraint for uncalibrated images* and indicates that the point m_i' can only lie on the epipolar line $l' = \mathbf{F}\ m_i$ of the point m_i. Then, from the k candidate points created via B-splines we choose the set of n $(n < k)$ correspondences that allow the most accurate computation of the fundamental matrix. This is done by the well-known RANSAC [6] algorithm, which is robust against noise perturbations of the data. This algorithm requires three parameters: the number N of samples/iterations, the threshold t that measures the maximum distance at which a pair of correspondences satisfy (4), and the number n of expected correspondences. We use the Sampson

[6]Do not confuse the fundamental matrix $\mathbf{F}$ with the affine mapping $\mathbf{H}$.

distance [20] and set $t = 2$ pixels. N can be computed as

$$N = \frac{log(1-p)}{log(1-(1-\epsilon)^s)}, \tag{5}$$

using a probability $p = 0.99$ to ensure that at least one sample of s points is free from outliers, being $s = 7$ points necessary to compute the matrix $\mathbf{F}$; and the pessimistic case of having a fraction $\epsilon = 0.5$ of contaminated correspondences in the input data. From the set of k potential matching points generated by B-splines, we expect to find $n = (1-\epsilon)k$ pair of correspondences. For $k = 1000$ knots, $n = 500$ correspondences are expected. See [6] for implementation details. Finally, the selected n points are used to compute the global approximation of the affine mapping $\mathbf{H}$ via equation (2). This approximation of the affine distortion is iteratively refined at every potential defect found in the identification stage (Section 2.2), as we will see next.

2.4 Robust Local Defect Tracking

Once potential defective regions have been identified in two consecutive or non-consecutive uncalibrated images, and given a preliminary estimation of the global geometric distortion between them, we attempt to track the intensities of each potential defect from the first view onto the second view. Only real flaws should be tracked, while false alarms must be consequently discarded. Here we face the well-known *visual matching problem*, which has been dealt with in literature by means of two approaches: feature-based matching (e.g. [21, 22]) and intensity-based matching (e.g. [23, 24]). Our inspection system combines both strategies.

Using the notation presented in [25], the goal of our intensity-based matching algorithm is to align a template image $T(\mathbf{x})$ with another image $I(\mathbf{x})$, where $\mathbf{x} = (x, y, 1)^T$ is a column vector of homogeneous pixel coordinates. A template $T(\mathbf{x})$ represents a potential defective region in the first view and $I(\mathbf{x})$ is the second view where the template has to match. Classical formulations aim to minimise the *sum-of-squared-differences* (SSD) of the intensities between the template T and image I warped onto the coordinate frame of the template, which is known as the *least-squares* (LS) formulation

$$\sum_{\mathbf{x}} \left[I(\mathbf{W}(\mathbf{x};\mathbf{p})) - T(\mathbf{x}) \right]^2, \tag{6}$$

where the sum is performed over all pixels in the template image, and $\mathbf{W}(\mathbf{x};\mathbf{p})$ is the warping map obtained by applying the affine transformation

$\mathbf{H}$ to the template coordinates, i.e.

$$\mathbf{W(x;p)} := \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1+p_1 & p3 & p5 \\ p_2 & 1+p_4 & p_6 \\ 0 & 0 & 1 \end{pmatrix}}_{\mathbf{H}} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \tag{7}$$

The affine mapping $\mathbf{H}$ is parameterised by an unknown vector $\mathbf{p} = (p_1, ..., p_6)^T$. In literature there are several methods to minimise (6). In particular, the Lucas-Kanade algorithm [26] assumes that a current estimation of $\mathbf{p}$ is known and then it solves iteratively for additive increments $\Delta\mathbf{p}$:

$$\sum_{\mathbf{x}} \left[I(\mathbf{W}(\mathbf{x}; \mathbf{p} + \mathbf{\Delta p})) - T(\mathbf{x}) \right]^2, \tag{8}$$

updating the parameter vector as $\mathbf{p} \leftarrow \mathbf{p} + \mathbf{\Delta p}$. In general, the equation (8) is not robust in presence of *outliers* like occlusions, illumination changes and non-gaussian noise, since its quadratic error measure assigns a high influence to gross errors, i.e. large deviations cause undesirable distortions in the resulting matching process. In order to downweigh the effect of outliers in the minimisation process, we derive a *robust* formulation of the matching problem. We seek for the M-estimator of $\mathbf{\Delta p}$ as the minimum of a global energy function

$$\mathbf{\Delta \hat{p}} = \arg\min_{\mathbf{\Delta p}} \mathrm{E}(\mathbf{\Delta p}), \tag{9}$$

where the energy function $\mathrm{E}(\mathbf{\Delta p})$ is defined in terms of a symmetric, positive-definite robust loss function[7] ρ, which has an unique minimum at zero, and it is chosen to be less increasing than square [28], i.e.

$$\mathrm{E}(\mathbf{\Delta p}) = \sum_{\mathbf{x}} \rho(z_{\mathbf{x}}), \tag{10}$$

where $z_{\mathbf{x}}$ is the normalised residue given by

$$z_{\mathbf{x}} = \frac{r_{\mathbf{x}} - \mathit{Median}(\mathbf{r})}{\hat{\sigma}}. \tag{11}$$

$\hat{\sigma}$ is the robust standard deviation of the residual vector $\mathbf{r} = \mathbf{I}(\mathbf{W}(\mathbf{x}; \mathbf{p} + \mathbf{\Delta p})) - \mathbf{T}(\mathbf{x})$, and it is computed through the *median absolute deviation* (MAD) [29] as

$$\hat{\sigma} = \zeta \, \mathit{Median}(|\mathbf{r} - \mathit{Median}(\mathbf{r})|). \tag{12}$$

[7] Alternatives to choose ρ, for instance, are: Cauchy, Huber, Tukey, Geman-McClure and Lorentzian robust functions [27]. In our experiments we use the Geman-McClure one.

The factor $\zeta =: 1/\phi^{-1}(0.75) = 1.4826$ (where ϕ is the cumulative distribution function of the standard normal distribution) is introduced in the equation (12) to obtain a consistent estimator of σ, which reaches the same efficiency as the least-squares estimator when only Gaussian noise exists. Moreover, it has been statistically proven that the median is more robust against outliers than the mean as estimator of the central tendency [29].
To solve the robust estimation problem we use the *iteratively reweighted least squares* (IRLS) algorithm proposed in [30]. Performing a first order Taylor expansion, the residual vector is linearised as

$$\mathbf{r} = \mathbf{I}(\mathbf{W}(\mathbf{x};\mathbf{p})) + \frac{\partial \mathbf{I}}{\partial \mathbf{p}}\boldsymbol{\Delta}\mathbf{p} - \mathbf{T}(\mathbf{x}), \tag{13}$$

and setting the partial derivative of the expression (10) with respect to $\boldsymbol{\Delta}\mathbf{p}$ to zero, we obtain

$$\sum_{\mathbf{x}} w(z_{\mathbf{x}}) \left[\frac{\partial I}{\partial \mathbf{p}}\right]^T \left[I(\mathbf{W}(\mathbf{x};\mathbf{p})) + \frac{\partial I}{\partial \mathbf{p}}\boldsymbol{\Delta}\mathbf{p} - T(\mathbf{x})\right] = \mathbf{0}, \tag{14}$$

where $\psi(u) = \frac{\partial \rho(u)}{\partial u}$ and $w(u) = \frac{\psi(u)}{u}$ are the first partial derivative and the weight of the robust loss function $\rho(u)$, respectively. Finally, the solution of the equation (9) is given by

$$\boldsymbol{\Delta}\hat{\mathbf{p}} = -\mathbf{H}^{-1} \sum_{\mathbf{x}} w(z_{\mathbf{x}}) \left[\frac{\partial I}{\partial \mathbf{p}}\right]^T \left[I(\mathbf{W}(\mathbf{x};\mathbf{p})) - T(\mathbf{x})\right], \tag{15}$$

being the Jacobian and the Hessian defined respectively as

$$\begin{aligned} \frac{\partial I}{\partial \mathbf{p}} &= \frac{\partial I}{\partial \mathbf{W}}\frac{\partial \mathbf{W}}{\partial \mathbf{p}} = \nabla I \frac{\partial \mathbf{W}}{\partial \mathbf{p}}, \\ \mathbf{H} &= \sum_{\mathbf{x}} w(z_{\mathbf{x}}) \left[\frac{\partial I}{\partial \mathbf{p}}\right]^T \left[\frac{\partial I}{\partial \mathbf{p}}\right]^T. \end{aligned}$$

As commented before, we consider the geometric distortions induced by the uncalibrated imaging system over the image domain as nonuniform. Therefore, it is necessary to apply the intensity-matching algorithm on each potential defect in order to estimate more accurately the local deformation at that location. Together with this, the set of features extracted from each potential defect during the identification stage are now used to distinguish between true and false flaws. The following section describes such a procedure.

2.5 Correspondence Analysis

Once individual projections for each hypothetical flaw have been found in both views by applying the local matching algorithm, a correspondence analysis is carried out to determine which of them are real and which are false alarms. A region will be classified as defective if three criteria are fulfilled:

1. **Identifiability**. The detection of existing defects must be ensured in the stage of identification of potential defects (Section 2.2). If we do not segment the defects at that step, we cannot detect them later on[8]. Thus, to be considered as flaw, a potential defect must be detected on both views being analysed.

2. **Spatial proximity**. A discontinuity in the first image must be projected to a position in the second image near a hypothetical defect with similar characteristics. To be considered in the vicinity of a flaw in the second image, the projected centre of mass of the defect can be at most 5 pixels apart -on each coordinate- from its candidate correspondence.

3. **Feature proximity**. To be considered similar to a flaw in the second image, at least four out of six shape characteristics of the projected region might differ by at most 30%, which is measured by taking the norm of the difference between the two normalised vectors of features. The following characteristics are taking into account: area of the segmented defect, average gray value, second derivative, and three different values for contrast.

To overcome the identifiability problem, correspondences in more views can be investigated. For instance, even if we identify a defect in the frames 1 and 3 (but not in frame 2), we can track it if we check the correspondences between views 1 and 3. This strategy was implemented by Mery and Filbert in [1] under the calibrated approach. They were able to track correspondences between a frame i and the following frames $i+1$, $i+2$ and $i+3$. However, the existing trade-off between the computational time these calculation demand and the performance requirements impose by a particular application must be carefully analysed.

[8]Inspection approaches which make use of only one view are also affected by this problem.

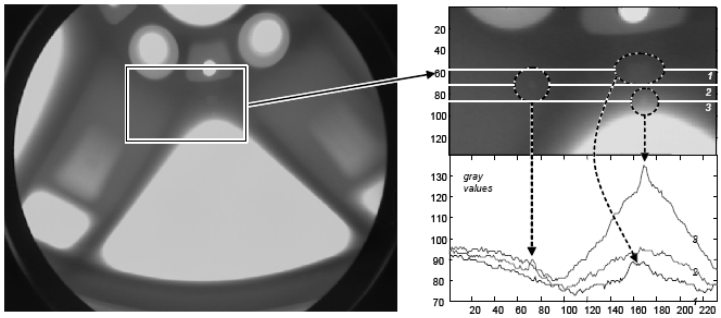

Figure 7: Radioscopic image of a casting with gray level profiles of three defects.

3 Experimental Results

In this section we apply our approach for automated defect detection in a sequence of 72 uncalibrated radioscopic images of an aluminum wheel. A segment of six of views was shown in Fig. 2. The dimensions of the wheel are 470 $[mm]$ diameter and 200 $[mm]$ height. The image size is 572 $\times$ 768 pixels with a dynamic range of 8 bits. The wheel has twelve known flaws. Three of these defects are existing blow holes with diameter $\emptyset = 2.0 - 7.5$ $[mm]$ (see Fig. 7). They were initially detected by a visual (human) inspection. The remaining nine flaws were produced by drilling small holes ($\emptyset = 2.0 - 4.0$ $[mm]$) in positions of the casting which were known to be difficult to detect (see Fig. 3). A pattern of 1 $[mm]$ in the middle of the wheel is projected as a pattern of 3 pixels in the image, i.e. the defects are actually very small. In addition, since the signal-to-noise ratio in our radioscopic images is low, the flaws signal is slightly greater than the background noise, as illustrated in Fig. 7. In our experiments, the mean gray level of the flaw signal (without background) ranges from 2.4 to 28.8 gray values with a standard deviation of 6.1. Analysing a homogeneous background in different areas of interest we obtain a noise signal within ± 13 gray values with a standard deviation of 2.5. For this reason, the segmentation of real defects with poor contrast can as well involve the detection of false alarms.
The results of the segmentation stage are summarised in Table 1 and partially shown in the Fig. 8. One observes that there are 7.74 false alarms per image. Nevertheless, the detection performance in this experiment is still good, because it is possible to identify 86% of all projected flaws along the sequence, whereas 14% of the existing 238 flaws are not identified due to their poor contrast with the background or because they are located at edges

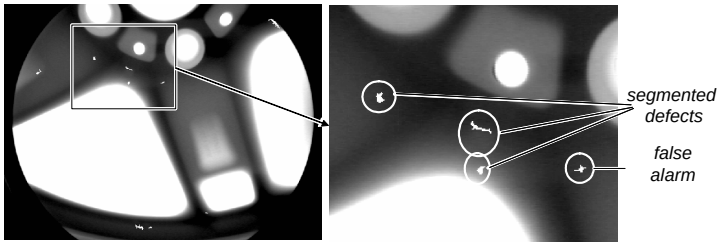

Figure 8: Segmentation of flaws. The existing defects were successfully detected, however, there are also false alarms.

Table 1: Performance of the identification step.

Existing defects (D)	238
Segmented potential defects	761
Detected defects (TP)	204
False alarms (FP)	557
Detection performance (TP/D)	85.71%
False alarms per image	7.74

of regular structures.
The performance of both least squares (6) and robust matching (10) algorithms are detailed in Table 2. Notice that the least-squares method might be equivalently obtained by setting $\rho(z) = z^2$ in the robust formulation (10). Each column subdivides the set of detected potential flaws into separate categories according to their actual condition and the classification given by each algorithm. The first four rows show the false negatives, i.e. real defects that could not be matched in the second frame. The first row enumerates the real defects that were not detected in the segmentation stage; the second row reflects the real defects that were impossible to register because they ran out of the image, which made looking for correspondences unfeasible. In both cases the errors do not count as bad performance of either the algorithms or the correspondence analysis; but they are considered as segmentation problems. The third and fourth rows show those defects in which the matching method either diverges or converges to a wrong location, respectively. In both cases the errors count as limitations of the matching algorithm. Finally, the fifth, sixth and seventh rows enumerate the number of detected defects, false alarms remaining after the matching, and the number of false alarms of the segmentation eliminated through the matching process, respec-

Table 2: Performance of the matching algorithms.

Method	Least-squares	Robust
No segmentation	13	13
Occlusion	10	10
Divergence	14	2
Wrong convergence	22	22
Detected defects	145	157
False alarms	52	55
Eliminated	505	502
Total	761	761
Detection performance	80.1%	86.7 %
False alarms per image	0.72	0.76

Table 3: (C)alibrated versus (U)ncalibrated approaches.

Method	Detected defects	False alarms
C - 3 views	100%	25%
C - 5 views	83%	0%
U - Least-squares	80.1%	26.4%
U - Robust	86.7%	25.9%

tively. The detection performance of the matching algorithms is computed as the ratio number of detected defects to the number of detectable defects, excluding those not identified by the segmentation. The robust algorithm detects almost 87% of the flaws with only 0.76 false alarms per image. The computation time required to process one image pair was in average 24.3 $[s]$ on a Pentium 4, 2.8 GHz desktop computer.

We have used the same image sequence as in [1], where flaws were tracked over previously calibrated images. Since our approach implements tracking over uncalibrated images, the results might not be fairly comparable. However, table 3 outlines the comparative performance of both approaches. The methodology based on calibrated images detects 100% of the defects when using 3 views, whereas 83% when using 5 views. This is because it is more probable to find flaws correctly segmented in 3 views than in 5. Nevertheless, using more views helps to reduce the percentage of false alarms (relative to

the number of potential defects). On the other hand, our robust approach for uncalibrated images makes use of only 2 views and achieves an acceptable compromise between detected defects and false alarms. Ways to improve these results are commented further in the concluding section.
As an extended step of our method, postprocessing of detected flaws could be introduced. For example, defective objects can be automatically taken out from the manufacturing line if the quality control system requires the production of flaw free pieces only; or, before removing them, a human inspector could verify the identified flaws by means of a computer-assisted tool. Moreover, in automated inspection of castings we should identify flaws with diameter greater than 2 $[mm]$, which were imaged as regions of approximately 12 pixels in our experiments on . This allowed us to segment them correctly together with many other false alarms that should be discarded. Our application on aluminum wheels requires that every flaw be detected, i.e. no defects of certain size should remain at any particular location of the casting. However, for other inspection tasks this requirement might be relaxed.

4 Conclusions and Future Work

The multiple view strategy is opening up new possibilities for nondestructive testing by taking into account correspondences between different views of a test object. In this paper, we present the *Robust Automated Multiple View Inspection* strategy for tracking potential defects on uncalibrated image sequences. Modelling the geometric distortion between each pair of consecutive or non-consecutive views as an unknown affine mapping, this framework introduces two complementary robust procedures to accurately estimate such a transformation. First, a global approximation of the mapping is computed through a set of selected corresponding points of the inspected object. Secondly, the intensities of each potential defect in the first view are iteratively matched onto the second view. As a result, only real defects are successfully tracked and false alarms are discarded. The practical importance of our method lies in avoiding the calibration process. The defect detection is carried out directly on the distorted views produced by an uncalibrated imaging system. This might help in manufacturing processes or in real-time applications that cannot be halted for calibration purposes, or it entails a difficult, unstable and time consuming process.
In our experimental results on aluminum die castings we have shown that flaw detection in uncalibrated images is promising. Our framework recognises 86.7% of all existing defects with only 0.76 false alarms per image. The

utilised image sequence is truly representative of those employed in labs to test algorithms for detecting potential flaws. Each image along the sequence contained 12 physical defects synthetically placed in a way that their detection is not evident due to their form, miniature size, location, intensity and deepness; some of them are almost imperceptible. The proposed methodology is straightly generalisable to any manufacturing system of regular structures. Indeed, this framework is not limited to X-rays, and it can be employed in uncalibrated sequences acquired from other imaging systems.
Our tracking scheme is based merely on two views because the quote of false alarms is low, but it can be run for three or more views. In the future we plan to extend this framework from image-pairs to image-triplets by means of trifocal tensors. Instead of using centres of mass to generate artificial points, we will consider structural edges of the objects to avoid lack of closed regions at the outer limits of the visible casting area, where most of the mismatches took place.
To improve the computational performance of our approach it would be worth to consider more efficient algorithms that solve the visual matching problem. In particular, the *inverse compositional* approach proposed in [31] is an interesting alternative that pre-computes the jacobian and hessian matrixes, which are updated at each iteration in our implementation. On the other hand, it is also valuable to look into coarse-to-fine strategies like *multigrid methods*. These allow the implementation of highly efficient real-time applications. See for example [32], where speedups of several hundreds are reached in estimating real-time motion.

Acknowledgement: This work was supported by FONDECYT – Chile under grant no. 1040210.

References

[1] D. Mery and D. Filbert. Automated flaw detection in aluminum castings based on the tracking of potential defects in a radioscopic image sequence. *IEEE T Robotic Autom*, 18(6):890–901, Dec 2002.

[2] D. Mery and M. Carrasco. Automated multiple view inspection based on uncalibrated image sequences. *Lect Notes Comput Sc*, 3540:1238–1247, 2005.

[3] G. Abramovich, J. Barhak, and P. Spicer. Reconfigurable array for machine vision inspection (RAMVI). In *Proceedings of the 3rd Interna-*

tional Conference on Reconfigurable Manufacturing, Ann Arbor, Michigan, 2005.

[4] P. Spicer, K. Bohl, G. Abramovich, and J. Barhak. Robust calibration of a reconfigurable camera array for machine vision inspection (RAMVI) using rule-based colour recognition. In *Proceedings of the First International Conference on Computer Vision Theory and Applications*, pages 131–138, Setúbal, 2006.

[5] S. Gumustekin. A visual inspection system using a single camera and mirrors. In *Proceedings of the 12th IEEE Signal Processing and Communications Applications Conference*, pages 257–260, Kusadasi, 2004.

[6] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, 2002.

[7] H. Boerner and H. Strecker. Automated X-ray inspection of aluminum casting. *IEEE Trans Pattern Anal Mach Intell*, 10(1):79–91, 1988.

[8] D. Filbert, R. Klatte, W. Heinrich, and M. Purschke. Computer aided inspection of castings. In *IEEE-IAS Annual Meeting*, pages 1087–1095, Atlanta, USA, 1987.

[9] W. Heinrich. *Automatische Röntgenserienprüfung von Gußteilen*. PhD thesis, Institut für Allgemeine Elektrotechnik, Technische Universität Berlin, 1988.

[10] G. Mohr and T. Fock. X-ray inspection in the aerospace industry - state of the art, challenges and emerging techonologies. In *Proceedings of the 16th World Conference on Nondestructive Testing*, Montreal, 2004.

[11] Md. Mahi Uddin Khan. Non-destructive testing applications in commercial aircraft maintenance. In *Proceedings of the 7th European Conference on Non-destructive Testing*, Copenhagen, 1998.

[12] Y. Kita, R. Highnam, and M. Brady. Correspondence between different view breast X-rays using curved epipolar lines. *Comput Vis Image Und*, 83(1):38–56, 2001.

[13] V. Rebuffel, S. Pires, A. Caplier, and P. Lamarque. Automatic delamination defects detection in radiographic sequences of rocket boosters. In *Proceedings of the International Symposium on Computed Tomography and Image Processing for Industrial Radiology*, Berlin, 2003.

[14] B. Jähne, H. Haußecker, and P. Geißler. *Handbook of Computer Vision and Applications*, volume 2. Signal processing and pattern recognition. Academic Press, San Diego, CA, 1999.

[15] D. Mery. Crossing line profile: a new approach to detecting defects in aluminium castings. *Lect Notes Comput Sc*, 2749:725–732, 2003.

[16] D. Mery, R. da Silva, L.P. Caloba, and J.M.A. Rebello. Pattern recognition in the automatic inspection of aluminium castings. *Insight*, 45(7):475–483, 2003.

[17] R.M. Haralick and L.G. Shapiro. *Computer and Robot Vision.* Addison-Wesley Publishing Co., New York, 1992.

[18] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis, and Machine Vision.* PWS Publishing, Pacific Grove, CA, 2 edition, 1998.

[19] C. de Boor. *A Practical Guide to Splines.* Springer Verlag, New York, 1978.

[20] P.D. Sampson. Fitting conic sections to 'very scattered' data: An iterative refinement of the bookstein algorithm. *Comput Vis Graph Image Process*, 18:97–108, 1982.

[21] W.M. Wells III. Statistical approaches to feature-based object recognition. *Int J Comput Vis*, 21(1–2):63–98, 1997.

[22] J.W. Hsieh, H.Y.M. Liao, K.C. Fan, M.T. Ko, and Y.P. Hung. Image registration using a new edge-based approach. *Comput Vis Image Und*, 67:112–130, 1997.

[23] M.J. Black and A.D. Jepson. Eigen-Tracking: Robust matching and tracking of articulated objects using a view-based representation. *Int J Comput Vis*, 26(1):63–84, 1998.

[24] G.D. Hager and P.N. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. *IEEE Trans Pattern Anal Mach Intell*, 20(10):1025–1039, 1998.

[25] S. Baker and I. Matthews. Lukas-Kanade 20 years on: A unifying framework. *Int J Comput Vis*, 56(3):221–255, 2004.

[26] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, pages 674–679, Vancouver, 1981.

[27] M.J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Comput Vis Image Und*, 63(1):75–104, 1996.

[28] C.V. Stewart. Robust parameter estimation in computer vision. *SIAM Rev*, 41(3):513–537, 1999.

[29] P.J. Rousseeuw and A.M. Leroy. *Robust Regression and Outlier Detection*. Wiley, New York, 1987.

[30] P.W. Holland and R.E. Welsch. Robust regression using iteratively reweighted least-squares. *SIAM Rev*, 6:813–827, 1977.

[31] S. Baker and I. Matthews. Equivalence and efficiency of image alignment algorithms. In *Proceedings of the 8th IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 1090–1097, Vancouver, 2001.

[32] A. Bruhn, J. Weickert, T. Kohlberger, and C. Schnörr. A multigrid platform for real-time motion computation with discontinuity-preserving variational methods. *Int J Comput Vis*, 70(3):257–277, 2006.