Automated visual inspection using trifocal analysis in an uncalibrated sequence of images

Miguel A. Carrasco, Domingo Mery(*)

Departamento de Ciencia of the Computación, Pontificia Universidad Católica de Chile.
Av. Vicuña Mackenna 4860 (143). Santiago de Chile

e-mail:{ mlcarras@puc.cl, dmery@ing.puc.cl}

(*) Corresponding Author

Abstract: Automated inspection using multiple views (AMVI) has been recently developed to automatically detect flaws in manufactured objects. The principal idea of this strategy is that, unlike the noise that appears randomly in images, only the flaws remain stable in a sequence of images because they remain in their position relative to the movement of the object being analyzed. AMVI has been successfully applied in sequences of calibrated images for which the 3D → 2D transference function for the projection of the views, is known precisely. Nonetheless, its application in industrial environments is a complex task because of the instabilities that are inherent to the system. This investigation proposes a new strategy, based on the detection of flaws in a non-calibrated sequence of images. The methodology designed consists in constructing a model and carrying out a trifocal analysis that allows the determination of the real position of a flaw using corresponding control points in the sequence. Experimental results obtained on radioscopic images of die castings illustrate the potential in the detection of defects in non-calibrated images, detecting the totality of the flaws in the sequence.

Keywords: Computer vision, multiple view geometry, automated visual inspection, defect detection.

1. Introduction

The quality of manufactured goods is one of the principal objectives of the productive process. In order to evaluate quality, there are a variety of inspection and analysis tools that can be carried out during the manufacturing process, however, all of these depend on security standards set by the manufacturer, or by some regulatory agency. While it is true that some manufacturers tolerate production with flaws, for others safety is a critical issue. Among the latter are high pressure equipment, chemical containers, aluminum wheels, etc. The existence of flaws in these products can cause serious accidents. Generally the inspection is carried out visually by trained personnel due to the fact that human visual inspection is flexible and adaptable to new situations not originally considered. However this process has grave problems such as: deficiency due to the time it takes to inspect an object which depends on fatigues and monotony; and inconsistency because the process depends on the capacity and experience of the inspectors (Newman & Jain, 1995). For this reason the process of analysis is a target for automation which allows an improvement in the quality of inspection, as well as a reduction in costs of production.

In recent years Automated Visual Inspection (AVI) has solved a number of problems in the area of quality control through the establishment of precise and objective control policies (Davis, 2005). Through image processing techniques and pattern recognition protocols, AVI allows the detection of flaws ensuring adherence to two basic conditions in the productive process, namely, efficiency and speed. Nonetheless, the majority of these techniques require the individual processing of each flaw in the image, which implies a subsequent analysis of the individual characteristics of each flaw. Subsequently, through the process of pattern recognition, we determined if the flaw is real or a false alarm.

Recently, a new methodology for automated flaw detection has been developed; Automated Multiple View Inspection, (AMVI) (Mery & Filbert, 2002). Equivalent in form to the flaw detection process carried out by an inspector, AMVI detects flaws using the following two steps. The first step, named *identification*, consists of detecting all the anomalous regions or hypothetical flaws in each image by means of a sequence of movements of the object, without any a priori knowledge of the object's structure. The next step, called *tracking*, consists in a follow-up of the hypothetical flaws found in each image in the sequence during the first

step. If the hypothetical flaws continue the length of the image sequence, the hypothetical flaw is tagged as a real flaw, and the object is catalogued as defective. On the other hand if the hypothetical flaws do not show correspondence in the sequence, they are considered to be false alarms. AMVI methodology's founding principle is that only real flaws, (and not false alarms), can be observed throughout the sequence of images because these remain stable relative to the movement of the object. Therefore, with two or more views of the same object, from different points of view, it is possible to improve performance with regards to real flaw detection.

This original strategy, presented in (Mery & Filbert, 2002), requires a previous calibration of the image sequence acquisition system. In calibration we seek to establish the transference function that projects a 3D point in the object onto a 2D point on the image (Mery, 2003a). Unfortunately, the calibration process is difficult to carry out in industrial environments due to the vibrations and random movements that vary in time and are not considered in the original estimated transference function. An alternative method for the carrying out of the AMVI strategy in non-calibrated sequences was presented in (Mery & Carrasco, 2005) for sequences with two images. Nonetheless, because the robustness of the AMVI methodology increases with the number of images analyzed in the sequence, in the present work we propose a modification to the robust system designed in (Mery & Carrasco, 2005), processing three images instead of two. Additionally, in this work we increase the number of corresponding control points that related the images uses B-Spline curves, thus significantly improving the estimation of the multi-focal model necessary for carrying out the tracking.

The remainder of this document is divided into the following sections. Section 2 includes background information on AMVI methodology. Section 3, which is dedicated to the proposed method, includes a description of the methodology used to segment hypothetical flaws, estimate the fundamental matrix robustly, generate artificial control points and estimate the trifocal tensors. Section 4 presents the experimental results. Finally, Section 5 presents the conclusions.

2. Background

The principal objective of AMVI is to follow only the hypothetical flaws, and not to estimate the structure of the object. Initially the methodology was implemented to automate the inspection of aluminum wheels, using a sequence of images in a calibrated system (Mery & Filbert, 2002). In this case the calibration of the object was generated off-line. A projection model was then generated to track flaws throughout the sequence of images, using the principles of multiple view geometry (Hartley & Zisserman, 2000), (Mery, 2003b). The results obtained demonstrated the technical feasibility of detecting the totality of real flaws, together with a high rate of detection of false positives. Nonetheless, in industrial environments, calibration is a complex process due to the vibrations during the acquisition of images of the object, which carries with it a lack of precision for the estimation of the parameters necessary for the multiple view geometric model.

The investigation carried out in (Mery & Carrasco, 2005) puts forth a robust alternative model which requires no calibration of the image acquisition system. One of the principal factors for estimating the movement of the object is that fact that it is a rigid body that has a rotational and/or translational movement with constant velocity and smooth trajectory. The method presented in (Mery & Carrasco, 2005) searches for significant regions of the object to be analyzed that are present throughout the image sequence. Once the correspondence between points in these regions has been established, a two-view model is constructed that serves for establishing correspondence between hypothetical flaws.

The proposed model initially used the following methodology: first, identify the structural points in each image in the sequence; second, find the correspondence, in consecutive images, of the points identified in the first step; third, generate a robust estimate of the fundamental matrix (Hartley & Zisserman, 2000), (Mery, 2003b) of possible corresponding points. Upon finishing this process a mathematical model is used to relate motion between pairs of images. The next phase relates hypothetical flaws in both images. Using epipolar geometry (Hartley & Zisserman, 2000), (Mery, 2003b), we search for flaws that may agree with the properties of the hypothetical flaws in the previous image. This evaluation is carried out by looking at Euclidean distances over the set of area and intensity properties.

Should the flaw have no correspondence in the next image, the latter is rejected, and is considered as a false positive as it does not fulfill with the epipolar constraint.

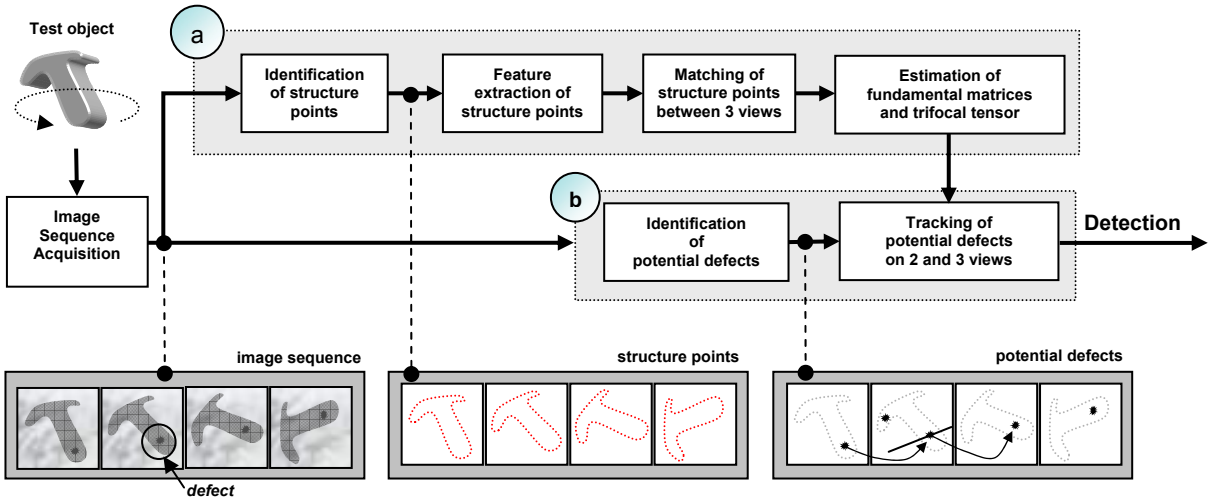


Fig. 1. Block diagram of the uncalibrated automated multiple view inspection: a) estimation of motion model, b) detection of defects.

There are two relevant factors regarding the process developed in (Mery & Carrasco, 2005). First, the segmentation phase is designed to detect the majority of defects without any *a priori* knowledge of the material and/or position of the object. Within the image processing phases, segmentation occurs in an initial stage, however it has a preponderant role in the entire process because poor performance in this phase can cause poor prediction in the detection of real flaws (Castleman, 1996). Secondly, the system of correspondence must evaluate only those pairs of points related in both views. Nonetheless, there is not always a perfect correspondence due to geometric distortions or other anomalies in the capture process. Therefore, a robust algorithm has been used that rejects positions that contain a projection error, and uses only a set of related pairs, thus minimizing the error between the real and projected position, according to a Euclidean distance metric. This method is known as a RANSAC approximation (Hartley & Zisserman, 2000). The term robust refers to the flexibility in the determination of a minimum set of related coordinates, and the rejection of those that do not fulfill with the minimum error allowed between the real and projected position.

3. Proposed Method

Laboratory results have shown that AMVI performs very well in the detection of flaws in aluminum die castings in calibrated environments (Mery & Filbert, 2002). Nonetheless, in industrial environments calibration is a complex process and of high cost for the manufacturer. This section presents a new AMVI method proposed for the automated detection of flaws using trifocal analysis of non-calibrated images, perfecting thus the method designed by the same authors in (Mery & Carrasco, 2005). The improvement is due to the fact that we have extended the analyses from two to three images, estimating the trifocal tensors robustly, and increasing the points of control artificially in order to establish correspondence between (Fig.1). The steps in this new method are presented below:

1. For each of the three images in the sequence (I, J and K), find the hypothetical flaws using the crossing line profile segmentation algorithm (Mery, 2003c), which searches in small high contrast areas.

2. For all the structures in the object being analyzed in images I, J, and K, search for relationships between structures and generate artificial control points with B-Spline curves (Bartels et al., 1998).

3. Estimate the fundamental matrix between images I and J using the RANSAC (Hartley & Zisserman, 2000) method with the structural and artificial points found in Step 2.

4. For all the hypothetical flaws in I found in Step 1, generate the epipolar projection using the fundamental matrix (Hartley & Zisserman, 2000), (Faugeras et al., 2001), (Mery, 2003c) in image J, and determined which flaws are closest to the epipolar line in search of

hypothetical flaws found in Step 1 in image J using the practical bifocal restriction:

a. If there is more than one hypothetical flaw on the epipolar line, find the best relationship on the basis of the properties of area and intensity by means of the smallest Euclidean distance, and store said relationship.
b. Should the epipolar line not pass through a hypothetical flaw in image J, this means that there is projected flaw in image J and the hypothetical flaw in image I is rejected.
c. If there is only one flaw on the epipolar line which meets the Euclidean distance criterion mentioned, then the relationship between the flaws in I and J is stored.

5. Estimate the trifocal tensors between images I, J and K using RANSAC (Hartley & Zisserman, 2000) with the structural and artificial points found in Step 2.

6. For the flaw relationships between I and J found in Step 4, find the position of the center of mass of the hypothetical flaws, and re-project these positions using the trifocal tensor:
   a. If there exists a projected flaw that is a minimum distance from the hypothetical flaw in image K from Step 1, assign this position as matching in three views and thus determine that a flaw in the sequence has been detected.
   b. If there is no flaw in image K, related to the projection, then eliminate the hypothetical flaw and catalogue it as a false positive.

An explanation of these steps is presented below.

3.1 Identification of hypothetical flaws

The segmentation of hypothetical flaws allows the identification of regions in each image of the sequence which may correspond to real flaws (Fig.2). There are two general characteristics used to identify them: i) a flaw is considered as a connected subset in the image, ii) the differences between the gray levels of the flaw and its neighbors is considerable. Initially, identification takes place through a process with no *a priori* knowledge using the convolution of the image with a Laplacian-of-Gaussian (LoG) kernel, and then a zero crossing algorithm (Castleman, 1996). The LoG operator intrinsically uses a Gaussian low-pass filter to reduce noise levels in the image. The results of the operator and the zero cross are a binary image that contains real flaws with connected surroundings. Nonetheless, these surroundings are not always closed. This happens when they are close to the edges of a regular structure (Fig.2c). The solution consists of augmenting the borders of the regular structures. This procedure consists of calculating the gradient of the image in order to identify these positions (Fig.2d) and later generate a binary image which employs only the levels with the most energy in the gradient. (Fig.2e). Once each closed region is segmented, characteristics are extracted through the grey tone profile of a straight line that passes through the center of the segmented region. Those that present a high variance profile are identified as hypothetical flaws (see details in Mery, 2003c). This hypothetical flaw contains a high number of false positives, it has however the following advantages: i) the same detector is applied to all images, ii) it allows the identification of hypothetical flaws, independently of the position or the structure of the object under study, in other words without *a priori* knowledge of the design of the structure, iii) the detection of real flaws is very high (higher than 90%).

Once the segmented regions have been determined, the next step is to determine the position of the center of mass for each hypothetical flaw. For each image $\mathbf{m}_i$ will be used to denote the center of mass of the segmented region r_i. In homogenous coordinates, $\mathbf{m}_i$ represents the 2D spatial position of a point i

$$\mathbf{m}_i = \begin{bmatrix} x_i & y_i & 1 \end{bmatrix} \quad (1)$$

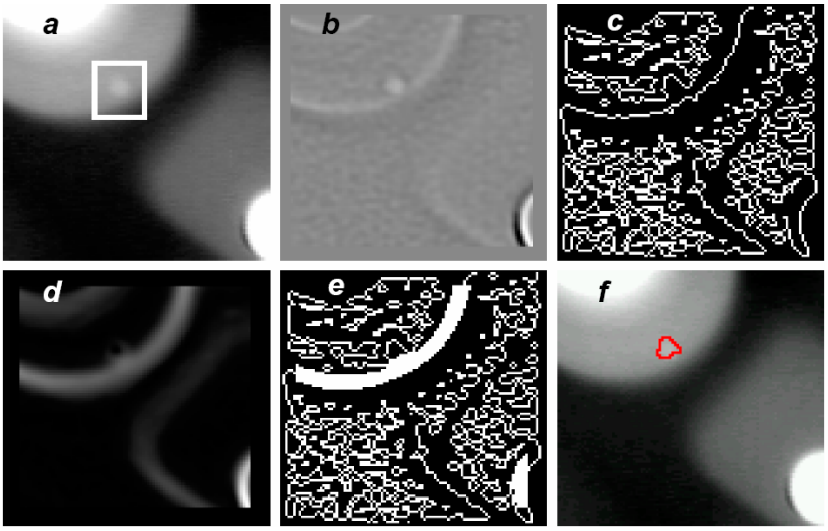
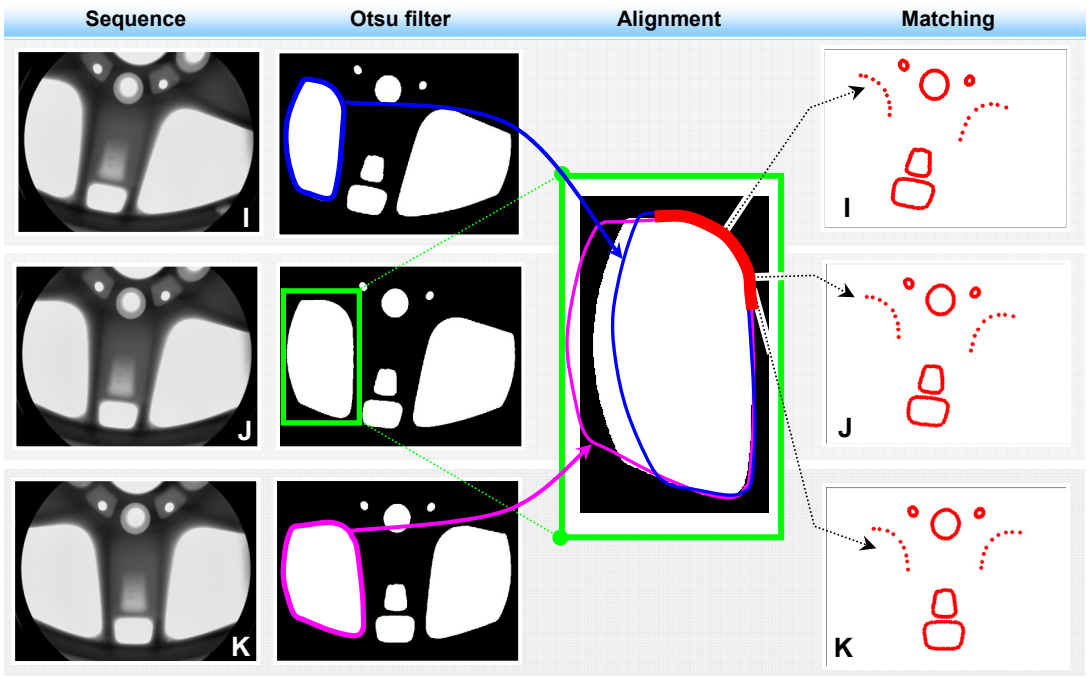


Fig. 2. Flaw detection: a) section of a radioscopic image with a flaw inscribed on the edge of a regular structure, b) application of the Laplacian filter on an image with $\sigma = 1.25$ pixels (kernel = 11 × 11), c) zero crossing image, d) gradient of the image, e) detection of edges after increasing the edges to the highest levels in the gradient, and f) detection of flaws using the variance of the crossing line profile (Mery, 2003c).



(a) (b) (c) (d)

Fig. 3. Non-linear matching algorithm for corresponding points in a sequence of three images: a) sequence of original images, b) application of Otsu's filter to separate the valid tracking structures and generation of border with B-Spline curves, c) alignment of each valid structure with respect to image J, and d) result of alignment with normalized correlation of the pattern of the *x*-coordinate position.

3.2 Identification of control points

The principal problem in the correspondence of the control points in the sequence of images used, is the low number of structures that have a valid projective transformation matrix, **H**. This means that some structures do not have a linear transformation due to the occlusion that results from being near the edge of the image. In these structures, it is possible to see only some regions where there is a correspondence, especially in the interior edges.

Using the rotation information of the sequence analyzed, we designed a practical method for finding a correspondence (Fig.3). The procedure uses the following steps:

a) Construct a binary image using Otsu's method (Haralick & Shapiro, 1992), thus isolating the structures from the image background

b) Generate a quadratic B-Spline curve using the edges of each segmented structure as control points (Bartels et al, 1998).

c) Align the structures of the first and third view with the second view. Generally, the movement of the piece is known *a priori* when its inspection is carried out. This information allows the determination of the angle and initial displacement in order to estimate a possible alignment.

d) Determine a common pattern for the rotation of the structures from the first and third view, using the *x*- coordinates positions from the second view as a base (Jain, 1989).

e) Calculate the position of the pattern relative to the curve of image J, so as to minimize the error by means of a normalized correlation (Fig.4).

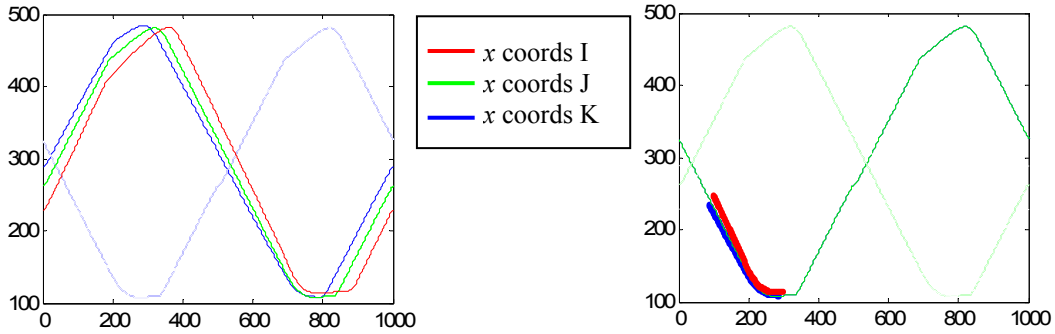


Fig. 4. Determination of *x*-coordinate position for each the B-Spline curve: a) the *x*-coordinate position for the first structure of images I, J, K, b) estimation in the position of the pattern of structures I and K over structure J using the pattern correlation over curve J.

3.3 Estimation of the fundamental matrix

The fundamental matrix is vital for the AMVI process as it relates any position in pairs of images. The precision of this matrix allows the correct determination of hypothetical flaws along the length of the epipolar line (Hartley & Zisserman, 2000), (Mery, 2003b). Nonetheless, if fundamental matrix is not robust, the epipolar line will be incorrect and the subsequent trifocal tensor process will be failed. In this case, if the point $\mathbf{m}_p$ of the first view corresponds to $\mathbf{m}_q$, in the second view, the following relationship is established,

$$\mathbf{m}_q{}^{\mathrm{T}}\mathbf{F}_{pq}\cdot\mathbf{m}_p=0\,, \qquad (2)$$

where $\mathbf{F}_{pq}$ is the fundamental matrix of the projection of points $\mathbf{m}_p$ and $\mathbf{m}_q$ in homogenous coordinates.

Once the set of corresponding positions has been generated in each region in both views, we use the robust RANSAC algorithm to estimate the fundamental matrix (Haralick & Shapiro, 1992). It should be remembered that there is a probability of error between the position of one region its corresponding pair, nevertheless, RANSAC minimizes this error as it uses the set of pairs of points that generates the best estimate of the fundamental matrix.

The RANSAC algorithm uses seven areas of points to determine the fundamental matrix. For this reason

there must be a minimum number of pairs of corresponding points that have been correctly estimated. Should there be an error in the correspondence of control points the fundamental matrix would be incorrectly estimated. Fortunately, it is not necessary for the process that all the points correspond exactly, this being the principal advantage of the robust algorithm in the estimation of the fundamental matrix.

3.4 Evaluation in two images

Once the robustly generated fundamental matrix has been constructed, it is necessary to calculate the epipolar line for each segmented region of the first view. Using the centers of mass for each hypothetical flaw generated in section 3.1 we generate the epipolar line thus

$$\mathbf{l}_{qi} = \mathbf{F}_{pq}^{\mathrm{T}} \mathbf{m}_{pi} = \begin{bmatrix} l_x & l_y & l_z \end{bmatrix}_i , \qquad (3)$$

where $\mathbf{l}_{qi}$ is the epipolar line of flaw i in the second view, and $\mathbf{m}_{pi}$ is the center of mass of flaw i in the first view. The result of $\mathbf{l}_{qi}$ is a line in the xy-plane, and has an equation as follows

$$\mathbf{l}_{qi} = A_i x + B_i y + C_i , \qquad (4)$$

where A_i=l_{xi}, B_i=l_{yi} and C_i=l_{zi} of flaw i, are the coefficients of the epipolar line.

Once the epipolar line of flaw i of the first view has been generated, it is necessary to determine the distance between the corresponding flaw in the second view and the epipolar line. This distance is determined through the practical bifocal restriction (Faugeras et al., 2001). Given that the epipolar constraint is applied to points and not to regions, we consider the center of mass of each hypothetical region to be the corresponding points between pairs of images This simplification is subject to error as it supposes that hypothetical regions have their center of mass in a corresponding point in both images. Nonetheless, we use this restriction because the majority of hypothetical flaws are small and the angles of rotation and deformations are small for each image.

For any flaw i in the first view and flaw j in the second view, we define $\mathbf{m}_{pi}$ and $\mathbf{m}_{qj}$ to be the centers of mass of the regions r_{pi} and r_{qj} in each view respectively. If the Euclidean distance between $\mathbf{m}_{qj}$ and the epipolar line of $\mathbf{m}_{pi}$ is less than a given ε, this implies that the hypothetical flaw in the second view is related to $\mathbf{m}_{pi}$. If the hypothetical flaw is found in both images, then it is considered to be a flaw in the bifocal correspondence, if this is not the case, the region is discarded.

$$d\left(\mathbf{m}_{pi}, \mathbf{F}, \mathbf{m}_{qj}\right) = \frac{\left|\mathbf{m}_{qj}^{T} \mathbf{F} \mathbf{m}_{pi}\right|}{\sqrt{l_x^2 + l_y^2}} < \varepsilon \qquad (5)$$

Experimentally, a given epipolar can contain more than one hypothetical flaw. In this case, we establish a degree of similarity for each hypothetical flaw, under the assumption that they all fulfill the epipolar constraint. For each flaw the characteristics of area and intensity are compared (Fig.5).

Similarity is established when two flaws (one in image I, and the other in the epipolar line of image J) are at a minimum distance in the normalized space of the properties, using the Euclidean distance as a similarity metric. If there are still false positives after the analysis of the two views, it is possible to use the trifocal tensor to eliminate the remaining false positives.

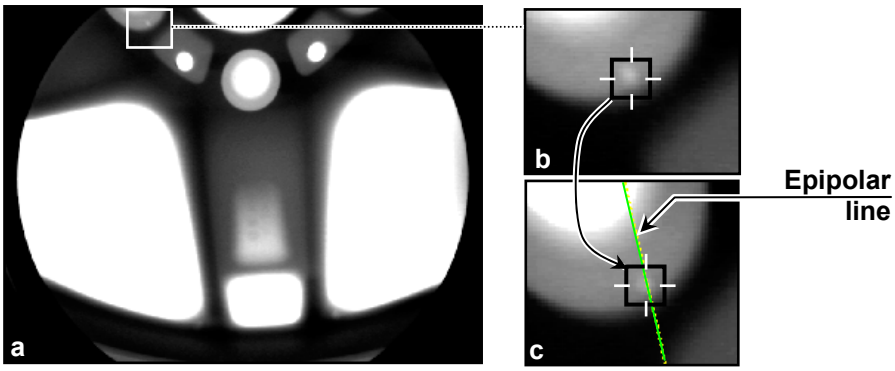


Fig. 5. Epipolar line generated automatically from the fundamental matrix: a) first view, b) identification of a hypothetical flaw, c) intersection of the epipolar line in the second view with one or more corresponding hypothetical flaws.

3.5 Estimation of the trifocal tensors

Trifocal analysis allows the modeling of all the geometric relationships in three views, and is independent of the structure contained in each image (Hartley & Zisserman, 2000), (Mery, 2003b). The tensor, a matrix structure similar to the fundamental matrix, only depends on the movement between images and the internal parameters of the cameras. It can be computed directly from the projection matrices of the views. Nonetheless, is can be calculated from the correspondence of the images without any *a priori* knowledge of the movement or calibration of the object. This characteristic justifies its, because the estimation of the fundamental matrix does not always eliminated the totality of false positives (Mery & Carrasco, 2005). An analysis in three views increases the probability of forming triplets that fulfill with the trifocal condition (Shashua & Werman, 1995).

In order to determine the trifocal tensors, we must have correspondence between the control points of the three views. Nevertheless, this assumption is not met precisely with the non-linear method designed in section 3.2. It is therefore necessary to use the robust algorithm RANSAC (Haralick & Shapiro, 1992) to determine the best triplets of points in order to minimize the projection error in the third view.

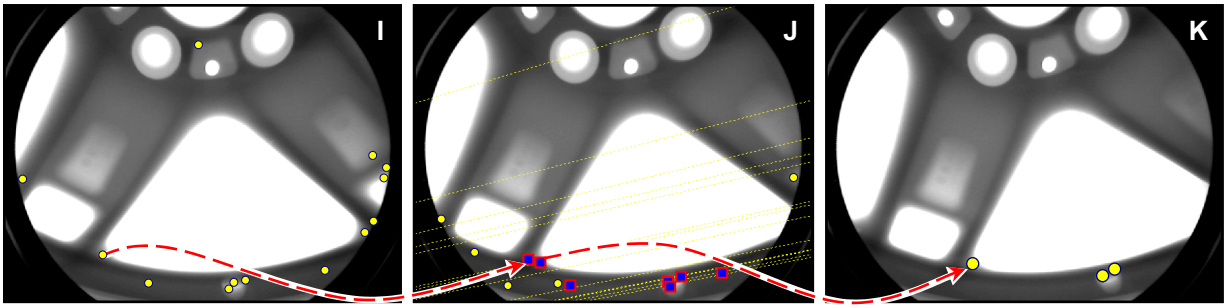
The initial estimation of the tensors is carried out with Shashua's four trilinearities (Shashua & Werman, 1995), (Mery, 2003b). This makes it possible to verify if three corresponding points $\mathbf{m}_p$, $\mathbf{m}_q$ and $\mathbf{m}_s$ in the first, second, and third view respectively, satisfy the trilinearities, in which case they are corresponding points in the three views, and they depend on the projection matrices (Hartley & Zisserman, 2000), (Mery, 2003b).

3.6 Evaluation in three views

To determine of the flaw in the third view corresponds to the trifocal relation, we use the center of mass $\mathbf{m}_p$ and $\mathbf{m}_q$ of regions r_p and r_q of the first and second view respectively. For this we use the re-projection of the trifocal tensor in the third view using the positions $\mathbf{m}_p$ and $\mathbf{m}_q$ in the two first views (Hartley & Zisserman, 2000), (Mery, 2003b). We use only the centers of mass of the two first views which fulfill with the bifocal relationship from section 3.4.

Let us define $\mathbf{m}_s$ as the center of mass of region r_s from the third view. If the Euclidean distance between the real position of the hypothetical flaw $\mathbf{m}_s$ and that which is estimated with the trifocal tensors, $\hat{\mathbf{m}}_s$ is less than some value ε, we take the hypothetical flaw to be a real flaw, as it complies with the correspondence in three views. Should the hypothetical flaw in the third view not agree with the projection of the tensor, it is discarded as it does not fulfill with the trifocal condition (Shashua & Werman, 1995), (Mery, 2003b).

$$d_r = \left\| \hat{\mathbf{m}}_r - \mathbf{m}_r \right\| < \varepsilon \qquad (6)$$



a) segmentation of hypothetical flaws (see 3.1 section)

b) proyection of epipolar line and matching in two views

c) projection of the tensor trifocal using I and J flaws.

Fig. 6. Generalized flaw estimation process: a) segmentation of hypothetical flaws in the first view, b) projection of the epipolar line in the second view using the robust fundamental matrix, c) projection of the coordinates of image 1 and 2, using trifocal tensors robustly over the third view.

Table 1. Performance in the identification of hypothetical flaws in the segmentation phase

Sequence	Number of images	Detected defects/image	False alarms/image
Left images	70	210/70 = 3	205/70 = 2.92
Center images	70	209/70 = 2.98	205/70 = 2.92
Right images	70	209/70 = 2.98	206/70 = 2.94
All images	210	628/210 = 2.99	616/210 = 2.93

Table 2. Performance in the detection of real flaws with two and three views in sequence.

Step	Detected defects in sequence	Real defects in sequence	False alarms in sequence	Dectection Performance	False alarm rate
Bifocal	190	190	93	100%	32.9%
Trifocal	170	172	19	98.8%	9.9%

In general, given that the trifocal condition is analyzed for the sequences that fulfill with the bifocal condition, we reduce the number of false positives generated in two views. Therefore, the number of false positives in three views can only be less than or equal to the number that exists in two views.

4. Experimental Results

This section presents the results of experiments carried out on a sequence of 72 radioscopic images of aluminum wheels (Mery & Filbert, 2002) (Fig.6). There are twelve known real flaws in this sequence. Three of these are impact flaws detected by human visual inspection, (de Ø = 2.0 ~ 7.5 mm), the remaining nine were generated by a drill which made small orifices (Ø = 2.0 ~ 4.0 mm) in positions which would difficult their detection.

We separated analysis into two steps. In the first step, called *identification* potential defects are automatically identified in each image of the sequence using a single filter and no apriori knowledge of the structure of the test object (Mery, 2003c). The results indicate that it exists 2.99 real flaws in each image, and 2.93 false alarms (Table 1). In the second step, called *tracking*, an attempt is made to track the identified potential defects in the image sequence. In this last step, we separate the analysis in two phases: first, the detection of pairs of flaws using the estimation of the fundamental matrix in two views, throughout epipolar constraint; second, using the previous results, we re-projected the pairs of hypothetical flaws in the third view using the trifocal tensor estimation. Both last phases are detailed below.

4.1 Performance with two views

The first phase is to evaluate the performance of the algorithm in two views using the bifocal method. This consists of determining the corresponding flaws between two images in a sequence through the search for flaws in the epipolar line. The method was applied to 70 pairs of radioscopic images (578 x 768 pixels) of aluminum wheels generated in (Mery & Filbert, 2002) for which the angle of rotation 5° is known for each sequence in the image. This information is used in order to align the segmented structures (see details in section 3.4)

The results indicate that the model detects 100% of the real flaws that have correspondence (Table 2, bifocal). This validates the assumption of correspondence between the positions of the real flaws and implies that automated detection with a fundamental matrix allows the detection of corresponding flaws contained in the epipolar line, which is in agreement with the results obtained in (Mery & Filbert, 2002), (Mery & Carrasco, 2005). The study showed a rate of 32.9% of false positives which have correspondence in image pairs. Although this percentage is high, we do not penalize false positives as these can be reduced using a third image.

With respect to the study carried out by the same authors in (Mery & Carrasco, 2005), we have extended our analysis to the entire test image sequence generated in (Mery & Filbert, 2002). The previous study used twelve sequences selected specifically as the matching system was limited to conditions where correspondence was feasible. The present investigation however, includes a non-linear correspondence system based on control points (see details in section 3.2).

The results show an increase in the detection of flaws through the use of the fundamental matrix, using a modified version of the original procedure presented in (Mery & Carrasco, 2005) (Table 3). This is due to the significant increase in the number of control points correlated in pairs and triplets of images through a non-linear correspondence system. The previous investigation used the centers of mass of regions with a variance of less than 7% in the area to be used as corresponding points. The present scheme uses the relation between the edges of the corresponding regions. For this reason RANSAC increases the precision of the calculation of control points and minimizes the estimation error of the fundamental matrix.

Table 3. Comparison of the present investigation and the study carried out in (Mery & Carrasco, 2005).

Technique		Sequence analyzed	Detection performance	False positives
Bifocal AMVI (2005)		12	92.3%	10%
Proposed	Bifocal AMVI	70	100%	32.9%
	Trifocal AMVI	70	98.8%	9.9%

In general, it is possible to determine precisely the fundamental matrix for each pair of images for the following reasons: i) the majority of the positions of each region agree with the rotation and/or translation of the following image; ii) only those regions that have a variation of less than 4% are considered, thus eliminating possible matching errors: iii) RANSAC uses only the best seven pairs from the total of positions in all the regions.

4.2 Performance with three views

The second phase uses the algorithm proposed in section 3.6. After completing the matching of possible pairs of flaws in both images, we extend the detection of flaws to the third image in the sequence. In the present case the study used 70 triplets of images.

Within the tests performed, the best performance of the trifocal tensor did not achieve perfect results. However, 98.5% of the real flaws were detected (Table 2, trifocal). Moreover, the results indicate a reduction in the quantity of false positives from 32.9% with bifocal AMVI to 9.9% with trifocal AMVI. This agrees with the assumption that in the measure that the number of sequenced images is increased, the number of false positives decreases as in general real flaws are in correspondence.

Our experiments have demonstrated that there is a greater sensitivity in the estimation of the tensor for the re-projection of the flaw in the third view. This phenomenon can be explained by the lower precision in correspondence for the regions in the three views, and by an error in the position of the center of mass for each segmented region. Furthermore, in some cases we were able to show that the estimation of the projection trifocal was correct although in the following image the process of segmentation did not generate flaws that corresponded with previous regions. Thus it was not possible to generate a triplet of flaws in correspondence. Let us remember that our trifocal analysis only allows pairs of flaws contained in the two first images of each sequence in order to estimate the projection in the third view.

5. Conclusions

This investigation presents the development of a new flaw detection algorithm in manufactured goods, using a non-calibrated sequence of images. Using new AMVI methodology (Mery & Filbert, 2002) we have designed a novel system of automatic calibration based only on the spatial positions of the structures. The proposed approach uses the projection of the epipolar line, generated by the fundamental matrix and the trifocal tensors in a robust manner, with the purpose of building a movement model without any *a priori* knowledge of structure. We based our investigation on the assumption that hypothetical flaws are real flaws if their positions, in a sequence of images, are in correspondence, because these remain stable in their position relative to the movement of the object.

With respect to the investigation carried out in (Mery & Carrasco, 2005), we have extended the analysis from two to three images per sequence through the estimation of trifocal tensors. Furthermore we have introduced new control points generated artificially through the use of B-Spline curves due to the low quantity of structures that remain stable in three images of a sequence.

Our results indicate that it is possible to generate an automatic model for a sequence of images which represent the movement between the points and the regions contained in these. The possibility of introducing non corresponding control points in

triplets of images, is the principal advantage of the RANSAC algorithm for estimating, robustly, the fundamental matrix and the trifocal tensors. In this way we can use as reference points the edges of the structures or areas with no loss of information using a non-linear method. The principal advantage of our model is the automatic estimation of movement. Nonetheless, in some cases we saw some projection errors due to geometric distortions in the acquisition of images. Our future work is to reduce the number of false positives using supervised classification techniques through the use of neuronal networks (Mery et al, 2003) and to add to the model geometrical distortion correction methods.

Acknowledgements

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

References

Bartels, R.H.; Beatty, J.C.; Barsky, B.A.: "Bézier Curves" Ch. 10 in An Introduction to Splines for Use in Computer Graphics and Geometric Modelling. San Francisco, CA: Morgan Kaufmann, pp. 211-245, (1998).

Castleman, K.: "Digital image processing". Prentice-Hall. Englewood Cliffs, New Jersey, (1996).

Davis, E.R.: "Machine Vision", 3er Edition, Morgan Kaufmann Publishers, Amsterdam, (2005).

Faugeras, O. et. al: "The Geometry of Multiple Images: The Laws That Govern the Formation of Multiple Images of a Scene and Some of Their Applications", The MIT Press, Cambridge MA, London, (2001).

Haralick, R.; Shapiro, L.: "Computer and robot vision". Addison-Wesley Publishing Co., New York, (1992).

Hartley R.; Zisserman A.: "Multiple View Geometry in Computer Vision". Cambridge University Press, Cambridge, UK, (2000).

Jain, A.K..: "Fundamentals of Digital Image Processing". Information and Systems Science Series. Prentice Hall, (1989).

Mery, D.: "Explicit Geometric Model of a Radioscopic Imaging System". NDT & E International, Elsevier, 36(8):587-599, ISSN 0963-8695, (2003a).

Mery, D.: "Exploiting Multiple View Geometry in X-ray Testing: Part I, Theory". Materials Evaluation, 61(11):1226-1233, (2003b).

Mery, D.: "Crossing line profile: a new approach to detecting defects in aluminium castings". Lecture Notes in Computer Science 2749. 725-732, (2003c).

Mery, D.; Carrasco, M.A.: "Automated multiple view inspection based on uncalibrated image sequences". Lecture Notes in Computer Science, LNCS 3540: 1238-1247, (2005).

Mery, D.; da Silva, R.; Caloba, L.P.; Rebello, J.M.A.: "Pattern Recognition in the Automatic Inspection of Aluminium Castings". Insight, 45(7):475-483. ISSN 1354-2575, (2003).

Mery, D.; Filbert, D: "Automated Flaw Detection in Aluminum Castings Based on the Tracking of Potential Defects in a Radioscopic Image Sequence". IEEE Transactions on Robotics and Automation, 18(6):890-901, (2002).

Newman, T.S.; Jain, A.K.: "A survey of automated visual inspection". Computer Vision and Image Understanding, 61(2): 231-262, (1995).

Shashua A.; Werman, M.: "Trilinearity of three perspective views and its associated tensor". In 5th International Conference on Computer Vision (ICCV-95), Boston MA, Jun. (1995).