Automatic Multiple Visual Inspection on Non-calibrated Image Sequence with Intermediate Classifier Block

Miguel Carrasco and Domingo Mery

Departamento de Ciencia de la Computación
Pontificia Universidad Católica de Chile
Av. Vicuña Mackenna 4860(143), Santiago de Chile
mlcarras@puc.cl, dmery@ing.puc.cl

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. This investi- gation proposes a new strategy, based on the detection of flaws in a non- calibrated sequence of images. The method uses a scheme of elimination of potential flaws in two and three views. To improve the performance, intermediate blocks are introduced that eliminate those hypothetical flaws that are regular regions and real flaws. Use is made of images captured in a non-calibrated vision system, so there are no optical, geometric and noise disturbances in the image, forcing the proposed method to be robust, so that it can be applied in industry as a quality control method in non-calibrated vision systems. the results show that it is possible to detect the real flaws and at the same time decrease most of the false alarms.

Keywords: computer vision, multiple view geometry, automated visual inspection, defect detection, industrial applications.

1   Introduction

Since the early 1980s various authors have shown the need to introduce Automatic Visual Inspection (AVI) systems in production processes [1,2,3]. According to them, there is no methodology applicable to all cases, since development is an ad hoc process for each industry. However, there is clear consensus that the use of AVI technologies can reduce significantly the cost and time spent in the process of inspection, allowing the replacement of a large number of inspectors of average training by a limited group of highly trained operators [4]. This has led in recent years to increased productivity and profitability, and to a reduction in labor costs [5]. In spite of their advantages, AVI systems have the following problems: i) they lack precision in their performance, since there is no balance between undetected flaws (false negatives) and false alarms (false positives); ii)

they are limited by the mechanical rate required for placing the object in the desired position; iii) they require a high computer cost for determining whether the object is defective or not; and iv) they generate high complexity in the configuration and lack of flexibility for analyzing changes in parts design. For those reasons, AVI remains as a problem open to the development of new applications.

To counteract the difficulties mentioned above, in recent years a new methodology has been developed to detect flaws automatically making use of the potential of multiple views called Automatic Multiple View Inspection [6,7,8,9] (AMVI). The main objective of AMVI is to exploit the redundancy of information in multiple views that have corresponding parts of the object that is being analyzed, so the information captured from different viewpoints can reinforce the diagnosis made with a single image. This original strategy, presented in [6], requires prior calibration of the image sequence acquisition system. In the calibration we seek to establish the transfer function that projects a 3D point in the object onto a 2D point on the image. 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 transfer function. An alternative method for carrying out the AMVI strategy in non-calibrated sequences was presented in [7] for sequences with two images, and in [8] for sequence with three images. In order to achieve an adequate performance, the number of false alarms to be tracked must be reduced. For that reason, the objective of our research is to improve the performance of the original AMVI scheme by introducing intermediate classifiers between the changes of views in order to reduce the number of false alarms and increase the performance in the detection of real flaws, and additionally to perfect the method of detection of control points to avoid calibration.

The remainder of this document is organized as follows: Section 2 includes background information on AMVI methodology; Section 3 deals with the proposed method, includes a description of the methodology used to generate artificial control points, and to estimate the fundamental matrix and trifocal tensors robustly; Section 4 includes the new intermediate classifier methodology; Section 5 shows the experimental results; and finally, Section 6 presents the conclusions and future work.

2   Background of Multiple Automatic Visual Inspection

Geometric analysis with multiple views represents a new field of analysis and development of machine vision [11,12]. The main idea is to get more information on the test object by using multiple views taken from different viewpoints. Using this idea, a new methodology for detecting flaws automatically, called Automatic Multiple View Inspection (AMVI) was developed in [6]. AMVI methodology is based mainly on the fact that only real flaws and not false alarms can be seen in the image sequence because their position remains stable relative to the object's motion. Therefore, having two or more views of the same object from different viewpoints makes it possible to discriminate between real flaws and false
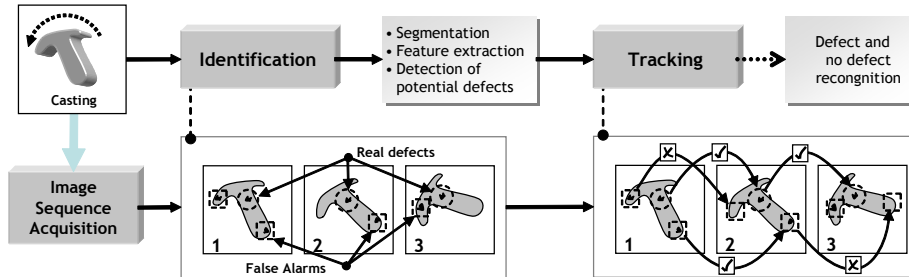
Fig. 1. Visualization of the general model of AMVI for the identification and tracking of hypothetical flaws in two and three views
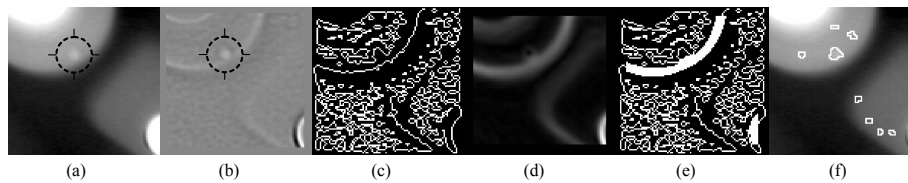

Fig. 2. Flaw detection: a) section of a radioscopic image with a flaw inscribed on the edge of a regular structure; b) application of the Laplacian filter on an image with $\sigma = 1.25$ pixels ($kernel = 11$x11); c) zero crossing image; d) image gradient; e) detection of edges after increasing them to the highest levels in the gradient; and f) detection of flaws using the variance of the crossing line profile (see details in [10])

alarms by means of a geometric tracking process in multiple views. AMVI has been developed under two schemes: calibrated and non-calibrated. Both methods detailed below share the following two steps: *identification* and *tracking*.

Identification: It consists in detecting all the anomalous regions or hypothetical flaws in each image of a motion sequence of the object, without a priori knowledge of its structure. The segmentation of hypothetical flaws allows the identification of regions in each image of the sequence which may correspond to real flaws (Fig.2) (see details in [10]). The process that follows is to extract the characteristics of each hypothetical flaw after identifying the regions by the previous procedure. Various investigations of the extracted characteristics have been described in [10,13]. This information makes it possible to determine if a flaw is corresponding in the multiple view analysis, according to the new intermediate classification system described in Section 4.

Tracking: It consists in following in each image of the sequence the hypothetical flaws detected in the first step, using the positions forced by the geometric restrictions in multiple views. If the hypothetical flaws continue through the image sequence, they are identified as real flaws, and the object is classified as defective. On the other hand, if the hypothetical flaws do not have correspondence in the sequence, they will be considered as false alarms (Fig.1). The AMVI

methodology has as its main foundation the fact that only real flaws, and not false alarms, can be seen throughout the image sequence, because their position remains stable relative to the object's motion.

AMVI is currently a useful tool and a powerful alternative for examining complex objects. It provides two independent approaches: those based on the calibration of a 3D→2D transfer function in the multiple view projection [6], and those based on the estimation of the motion of the control points in correspondence for pairs [7] and triplets of views [8] without prior calibration. A brief description of each is given below. **i) Calibrated method:** The calibrated image sequence flaw tracking method was first used as a quality control method for aluminum castings [6]. This approach consists in the estimation of the 3D→2D model through an off-line process called calibration [14], which is the process that allows the determination of the model's parameters to establish the projection matrix of a 3D point of the object at a 2D point of the digital image. Unfortunately, the model's parameters are usually nonlinear, which implies that the optimization problem does not have a closed solution. For that reason, it is finally impractical in industrial environments, where there are vibrations and random movements that are not considered in the original transfer function, i.e., the calibration is not stable, and the computer vision system must be calibrated periodically to avoid this error.

ii) Non-calibrated method: To avoid the problems involved in the calibrated method, a new system was developed using a sequence of non-calibrated images for two views [7]. This system does not require prior calibration. On the contrary, it can estimate the model of the motion using the images of the sequence in a procedure that can be carried out in line with the computer vision system.

In general, to achieve high precision in the motion model it is necessary to determine a large number of correspondences of control points in pairs and triplets of images in sequence. Many times this condition is difficult to achieve, and for that reason the RANSAC algorithm [11] was used in [8].

In our work, as will be seen in following sections, the attempt is made to improve the non-calibrated approach. First, our proposed method does not involve affine transformations, on the contrary, it aims at estimating correspondences through a geometric process in multiple views. Second, we use intermediate blocks that eliminate those hypothetical flaws that are regular regions and real flaws in order to increase the performance.

3   Proposed Method

Below is an explanation of each of the stages of the non-calibrated AMVI process with intermediate classifiers. The proposed scheme has three steps (A, B and C) detailed in Fig.3. They correspond to the stages of identification (A), extraction of control points (B), and tracking (C).

A. Identification of Hypothetical Flaws: The *identification* stage allows the detection of the hypothetical flaws by means of an algorithm without a priori knowledge of the object that is analyzed. Its most important characteristic is

the high percentage in the detection of real flaws, in spite of the existence of false alarms. Using the method of segmentation and extraction of characteristics described in [10] we determine all the regions with a high degree probability of being real flaws. The next step is to determine the position of the center of mass for each hypothetical flaw. For each image, m_i will be used to denote the center of mass of the segmented region r_i. In homogenous coordinates, $\mathbf{m}_i = [x_i, y_i, 1]^{\mathsf{T}}$ represents the 2D spatial position of a point i. This information makes it possible to analyze the trajectories of the hypothetical flaws in the subsequent stages of the proposed method.

B. Robust Control Points: The *control point extraction* stage allows the determination of corresponding points in multiple views. The process has two general steps: identification of control points, and matching of control points. The first step consists in determining the possible regions that can be in correspondence. The second step allows discarding possible combinations of correspondence that have a large error, storing a subset of correspondences with the highest precision. In our investigation we proposes a new curve alignment system by maximizing Pearson's correlation coefficient [15] in the correspondence between 2D curves, using an isometric transformation between the curves. We use this scheme because in the analysis of manufactured products the object that is analyzed is usually not deformable. This premise justifies the use of a rigid transformation method with which, given a rotation and displacement, it is possible to estimate a correspondence between the object's control points. However, due to the object's rotation, some regions can remain occluded, and therefore the proposed system must consider that only some regions retain this transfor-
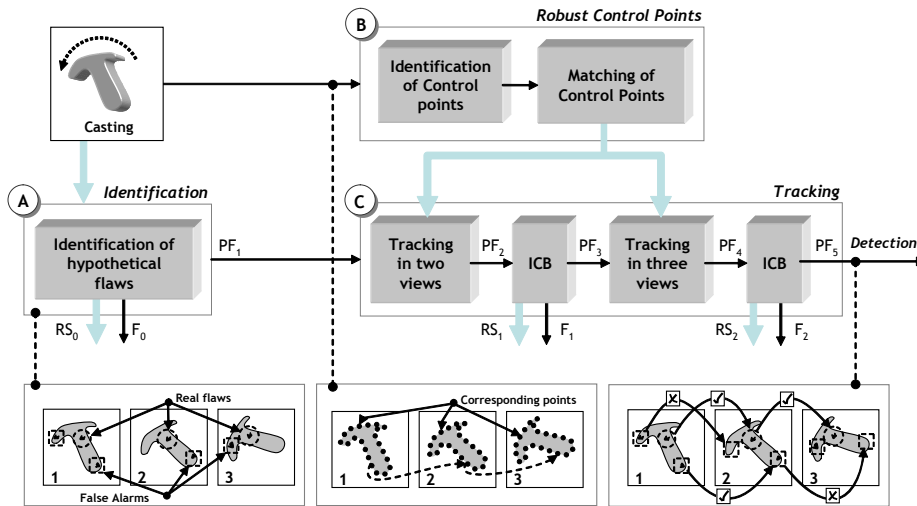

Fig. 3. Visualization of the general model for the identification, control point estimation, and tracking of hypothetical flaws in two and three views

mation. The proposed robust system of control points consists of two stages that are detailed below: matching of regions, and matching of control points.

B.1) Matching of regions: It consists in establishing the correspondence between regions of each view and not the control points. The process designed consists of four stages: First, segmenting those regions in which the intensity of the object is distinguishable from the background, using the method of Otsu [16]. Second, extracting a set of characteristics for each segmented region. This consists in extracting the centers of mass, the area, and the order moments of Flusser-and-Suk [17] of each region in the three views. Third, determining the matching between the segmented regions using the characteristics extracted before, and relating those regions having greater similarity according to a Euclidian distance metric. Fourth, once the correspondences between the regions have been determined, extracting the edges of each region and smoothing them to decrease the noise of each curvature. For that we calculate the perimeter of each segmented region and generate a list in parametric form as $Z_s = [x_s, y_s]$, where $s = 0, \ldots, L-1$ is the index of the list of pixels ordered in a turning direction, and L is the number of pixels of the region's perimeter. Using this parametric form, we generate the Fourier descriptors [18], transforming the Z_s coordinates in a complex value $u_s = x_s + j \cdot y_s$. This signal with period L is transformed into the Fourier domain by means of a discrete Fourier transform (DFT) [19]:

$$F_n = \sum_{s=1}^{L-1} u_s \cdot e^{-j \cdot \frac{2\pi \cdot s \cdot n}{L}} \tag{1}$$

The modulus of the complex Fourier coefficients describes the energy of each descriptor. Therefore, if we choose the highest energy coefficients (above 98%) and return to real space with the inverse discrete Fourier transform (IDFT) we get a smoother curve with less noise. However, when applying the elimination of some Fourier coefficients, the original curve is transformed into a new curve $C_s = [x'_s, y'_s]$, where, $C_s \neq Z_s$.

B.2) Matching of Control Points: It is a process in which the correspondence of points of the curve is established for each view. Using the Fourier processing described above, we define a curve C_1 corresponding to a region in the first view, and a curve C_2 corresponding with C_1 in the second view. For both curves to keep the same distance and be aligned it is necessary to select a section of each list having equal length.

Let P, a section of curve C, be such that $P = C(\delta)$, where $\delta = [s_i, \cdots, s_j]$, where $i, j \in [1, \cdots, n]$. In this way there is a section P_1 in the first view that has the same length as section P_2 in the second view. These sections of the curve do not necessarily have a correspondence, and for that we define a shift operator $\Theta(P, \lambda)$ that displaces list P by λ positions in a turning direction. Operator Θ uses the function "mod" (modulus after division) to determine the λ relative positions that list C, of length P, must turn. Using the above definitions, we design an alignment function as the maximization of Pearson's correlation

coefficient $\rho(\alpha, \beta)$ [15] between the isometric transformation of a section of P_1, with the shift of section P_2 with a jump λ

$$\{\theta, \Delta s_x, \Delta s_y, \lambda\} = \arg\max \|\rho\left([R,t][P_1], \Theta(P_2, \lambda)\right)\| \tag{2}$$

where,

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \qquad t = \begin{bmatrix} \Delta s_x \\ \Delta s_y \end{bmatrix} \tag{3}$$

This maximization function must find parameters $\{\theta, \Delta s_x, \Delta s_y, \lambda\}$ to estimate an alignment between sections P_1 and P_2. The main advantage of using this function is that it does not require a perfect alignment because the correlation coefficient is maximum when the displacement is linear. Another advantage is that curves P_1 and P_2 are open, so the alignment determines only sections that are corresponding, allowing control points to be obtained for curves that have partial occlusion in corresponding regions. Also, the use of parameter λ allows finding a position relation for curve C_2 with P_1, and in this way, while curve P_2 adjusts its shift, curve P_1 adjusts its translation and rotation to become aligned.

C. Tracking of Potential Flaws: The *tracking* stage allows tracking of hypothetical flaws obtained in the identification stage. The method for carrying out the tracking takes place through multiple view geometric analysis [11].

C.1). Two Views: The mathematical formulation that allows relating two points in stereo images is called the fundamental matrix [11]. Within the AMVI field its use is vital because it allows the trajectories of hypothetical flaws to be analyzed in two views and to verify if the flaws are corresponding. In this case, if point m_p of the first view corresponds to m_q, in the second view, the following relationship is established:

$$\mathbf{m}_q^\mathsf{T} \cdot \mathbf{F}_{pq} \cdot \mathbf{m}_p = 0 \tag{4}$$

where $\mathbf{F}_{pq}$ is the fundamental matrix of the projection of points $\mathbf{m}_p$ and $\mathbf{m}_q$ in homogenous coordinates as $[x_p, y_p, 1]^\mathsf{T}$ and $[x_p, y_p, 1]^\mathsf{T}$, respectively. Once the set of corresponding positions has been generated in each region in both views by the method proposed in section 3-B, we use the robust RANSAC algorithm to estimate the fundamental matrix [11].

In our investigation we used the method proposed by Chen et al. [20] to make an initial estimation of the fundamental matrix. The modification of Chen's method consists in choosing a subset of candidate points by means of the epipolar restriction. So in our method we use a combination of the algorithm of Hartley and Zisserman [11] with the normalization of the 2D coordinates, followed by an estimation of the fundamental matrix through the biepipolar restriction [20].

Therefore, using the centers of mass for each hypothetical flaw generated in section 3-A, we generate the epipolar line thus $l_{qi} = \mathbf{F}_{pq}^\mathsf{T} \cdot \mathbf{m}_{pi} = [l_x, l_y, l_z]_i$, 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. 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 constraint [12] as

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

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 $\mathbf{m}_{pi}$ (Fig.4). 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.

C.2) Three views:
The initial estimation of the tensors is carried out with Shashua's four trilinearities [21]. In particular we use an estimation of the tensors that maximizes the number of inliers according to the RANSAC trifocal algorithm. Furthermore, the estimation of the tensors was made with the normalized linear algorithm [11, pp.383]. Once the trifocal tensors have been determined, it is possible to verify whether three points $\mathbf{m}_p$, $\mathbf{m}_q$ and $\mathbf{m}_s$ are corresponding in the first, second, and third view, respectively. For that 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 first two views applying the point-line-point method [11, pp.373]. We use only the centers of mass of the first two views which fulfill the bifocal relationship from section 3-B. 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, since it complies with the correspondence in three views as $d_s = \|\hat{\mathbf{m}}_s - \mathbf{m}_s\| < \varepsilon$. Should the hypothetical flaw in the third view not agree with the projection of the tensor, it is discarded, as it does not fulfill the trifocal condition [21].
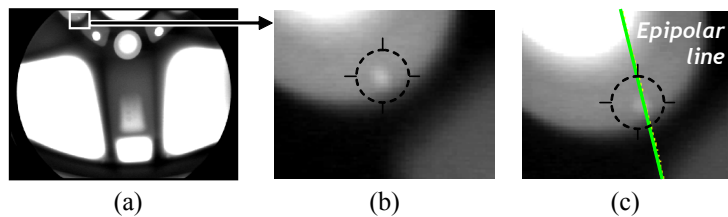

(a)   (b)   (c)

Fig. 4. Epipolar line generated automatically from the fundamental matrix: a) first view; b) zoom-identification of a hypothetical flaw; c) intersection of the epipolar line in the second view with one or more corresponding hypothetical flaws

4   Intermediate Classifier Block Method

The Intermediate Classifier Block (ICB) method proposed uses the classifier ensemble methodology [22], in which different linear classifiers do the classification and then, through the majority of votes technique, the final classification decision is made. The objective of the ICB method is to eliminate those correspondences between hypothetical flaws that have a low probability of being the same flaw in the different views. The ICB method has as input the distribution of two classes: Flaws (F) and Regular Structures (RS). According to this distribution, the classifier must determine the region of space where there are actually flaws only starting from point θ_F, and regular structures from θ_S (Fig.5a). Once these regions are extracted, only the hypothetical flaws contained in the region in which the classifier cannot verify with high probability the kind of class to which they belong are assigned to a new class called Potential Flaw (PF) (Fig.5b). This reduction avoids the analysis of the trajectories of all the flaws in correspondence, thereby improving the performance.

The simplest form of the previous classifier is reflected in the linear separation of the RS, PF and F regions, using the V_1 and V_2 features (Fig.5b). In the case of having three features $[V_1, V_2, V_3]$, the separation between them generates a three-dimensional volume bounded by the cuts of the two-dimensional separations, containing only the hypothetical flaws considered as potential flaws (PF) (Fig.5c). This three-dimensional volume generated from the combination of the two-dimensional features $[V_1, V_2]$, $[V_1, V_3]$ and $[V_2, V_3]$ contains the potential flaws that will be analyzed in the following phases of the multiple views analysis. On the other hand, the regions outside the three-dimensional volume can be flaws or regular structures, depending on the position in which the hyperplanes are projected. Our analysis considers the combination of two to seven features, giving rise to multidimensional section maps generated from the two-dimensional combinations. The methodology used by the ICB consists of a series of stages detailed below.
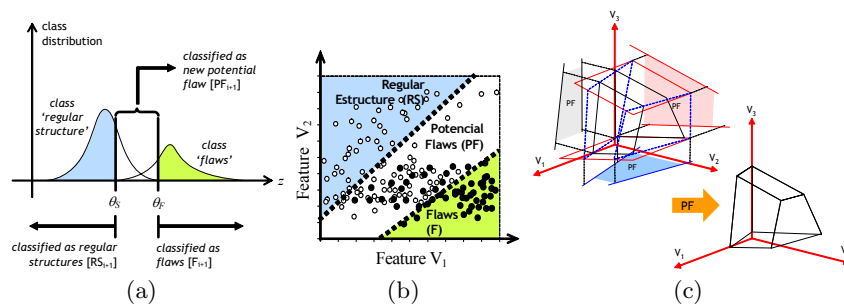

Fig. 5. (a) distribution of classes of hypothetical flaws between the views; b) distribution of classes in two dimensions with the linear separation of the RS, PF and F regions; c) Three dimensional representation of the ICB classification system

i) Assessment method of the ICB classifier: Our problem falls within the framework of supervised classification problems, since the class to which each potential flaw belongs is known. Using this information, the classification model is designed by means of the cross-validation method [23]. To compare the results of the various configurations of the classifier we use two parameters known as the ROC curves [24], *sensitivity* and *1-specificity*, which allow the measurement of the performance of a classification of two classes. The main characteristic of the ROC curve is that it allows the comparison to be independent of the sample.

The objective is for the sensitivity to be maximum (100%) and at the same time the 1-specificity to be minimum (0%), and in this way the classifier guarantees an ideal classification for two classes. In practice this is difficult to achieve, because it depends on the classifier's internal parameters, and it can be quite variable with respect to the noise existing in the data.

ii) Selection of characteristics: The characteristics selected by the ICB classifier are determined automatically using the information contained in each potential flaw, each of which has associated a characteristics vector. To determine the combination of characteristics that separate the classification space we will use the Take-L-Plus-R characteristics selection algorithm [25]. The objective of this algorithm is to determine the best characteristics that allow a greater separation of the space between the classes. In our research we used Fisher's discriminant as the criterion function [26].

iii) Linear Classification: We use a linear classification system that allows finding the hyperplanes that best separate the solution space. For that, the classification process must fit the following linear equation $\mathbf{w}^{\mathsf{T}} \cdot \mathbf{v} + w_0 > 0$, where $\mathbf{w} = \Sigma_{\mathrm{w}}^{-1} \cdot (\bar{v}_1 - \bar{v}_2)$ are the hyperplane parameters, Σ_{w} is the interclass covariance matrix, and $\mathbf{v}$ corresponds to the characteristics vector chosen earlier. Finally, factor w_0 for two classes is determined according to the mean of characteristics $\bar{v}_1$ and $\bar{v}_2$ and the probabilities of each class p_{e1} and p_{e2} according to

$$w_0 = -\frac{1}{2} \cdot (\bar{v}_1 + \bar{v}_2) \cdot \Sigma_{\mathrm{W}}^{-1} \cdot (\bar{v}_1 - \bar{v}_2) - \log\left(\frac{p_{e1}}{p_{e2}}\right) \tag{6}$$

Once an initial solution is obtained for parameters $\mathbf{w}$, the optimization problem tries to fit the hyperplanes so that (7) is maximum, and in that way always ensure that we are obtaining a high performance for each subselection of characteristics.

$$\{\mathbf{w}, w_0\} = \arg\max \{S_n(\mathbf{w}, w_0)\} \ \ s.t. \ S_p(\mathbf{w}, w_0) = 1 \tag{7}$$

This problem has been solved by the Nelder-Mead Simplex method [27]. Then the information from the selected straight lines and characteristics is used to evaluate the performance of the classifier on the test data. At some time each register is used to build the model or to be part of the test. This is necessary because of the low number of registers available at the time of identifying the hypothetical flaws, and for that reason we used the ensemble of classifiers [22].

5 Experimental Results

This section presents the results of experiments carried out on a sequence of 70 radioscopic images of aluminum wheels (see some of them Fig.6). There are twelve known real flaws in this sequence. Three of them are flaws detected by human visual inspection ($\emptyset = 2.0 \sim 7.5$ mm). The remaining nine were generated by a drill which made small holes ($\emptyset = 2.0 \sim 4.0$ mm) positions that would make their detection difficult. The method was applied to 70 radioscopic images (578 x 768 pixels) of aluminum wheels generated in [6] for which the angle of rotation of $5°$ is known for each sequence in the image.

We separated the analysis into three steps. i) *identification*: in this step potential flaws are automatically identified in each image. The result of the identification generates a data base that contains 424 registers with 11 characteristics of the total hypothetical flaws detected in the sequence, 214 registers are real flaws, and 210 registers are regular structures or false alarms that must be reduced. ii) *tracking*: in this step we track the identified potential flaws in the image sequence in two and three views. iii) *ICB method*: finally, we analyzed the performance of the classifiers inserted in two and three views, to filter the hypothetical flaws between the views. The last two steps are detailed below.

i) Performance with two views: The results indicate that the model detects 100% of the real flaws that are corresponding in two views (Table 1, Track 2 Views). This validates the assumption of correspondence between the position of the real flaws and implies that automatic detection with the fundamental matrix allows the detection of corresponding flaws that are contained on the epipolar line, and this agrees with the results given in [6] and [7]. There is, however, a large number of false alarms in sequence (198/388=51%), which must be reduced using a third view.

ii) Performance with three views: 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 this case the performance remains at 100% of real flaws detected in sequence, but it is seen, however, that it has not been possible to eliminate all the false alarms (Table 1, False Alarms). Furthermore, it is seen that the ICB method in two and three views has allowed the detection of a large part of the real flaws (F) and regular structures (RS) with high probability, allowing them to be separated from the multiple views analysis.

Table 1. Performance of the Uncalibrated Tracking

Step	Flaws	Regular Structure	Real Flaws	False Alarms
Track 2-Views	190	198	100%	51.0%
ICB-2	151	94	100%	24.2%
Track 3-Views	137	45	100%	11.6%
ICB-3	18	17	100%	4.4%

Table 2. Comparison between different calibrated and non-calibrated tracking

Method	Tracked	Year-Reference	Analyzed Images	True Positives	False Positives
Calibrated	3	2002 [6]	70	100%	25%
	4	2002 [6]	70	100%	0%
	5	2002 [6]	70	83%	0%
Uncalibrated	2	2005 [7]	24	92.3%	10%
	2	2006 [8]	70	100%	32.9%
	3	2006 [8]	70	98.8%	9.9%
	2	2007 [9]	70	86.7%	14%
	2	2007^{new}	70	100%	24.2%
	3	2007^{new}	70	100%	4.4%

iii) Performance of ICB: The greatest advantage of ICB classifiers for two and three views is the extraction of flaws and regular structures with high probability. The results indicate a clear relation between the performance of the ICB method for two views and the number of characteristics chosen. In this way, with the five best combinations of characteristics the performance in the classification is ideal, but there is a clear decrease in the number of flaws extracted by the ICB method (Fig.7a). In the case of three views, the number of correspondences is drastically reduced because the correspondence of a hypothetical flaw in three images has a lower probability of occurrence (Fig.7b).

iv) Comparison with other methods: Finally, we present a summary of the performance obtained with the calibrated and non-calibrated AMVI approximation (see Table 2). It shows the performances corresponding to the tracking phase of the different investigations carried out with an non-calibrated sequence of X-ray images designed in [6]. According to the results generated in two and three views by the same authors in 2006 [8], it is seen that the intermediate classification block (ICB) technique has allowed a reduction of 8.7% in the correspondence number in two views, and of 5.5% in the case of three views, with a 4.4% remainder that it has not been possible to eliminate by geometric analysis.
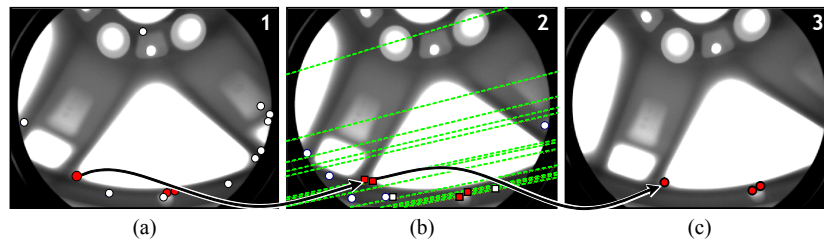

(a)   (b)   (c)

Fig. 6. Generalized flaw estimation process in one sequence of three views: a) segmentation of hypothetical flaws; b) projection of the epipolar line in the second view using the robust fundamental matrix; c) projection of the coordinates of images 1 and 2 using trifocal tensors over the third view
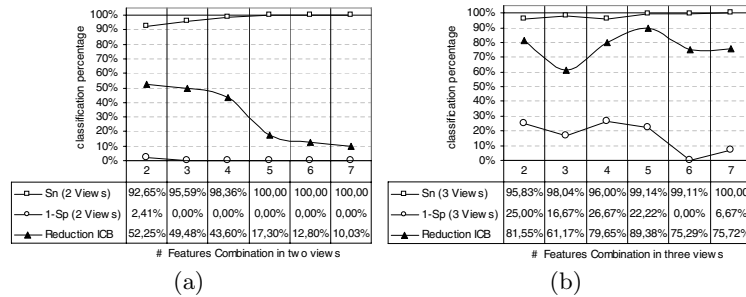
—□— Sn (2 Views)	92,65%	95,59%	98,36%	100,00	100,00	100,00
—○— 1-Sp (2 Views)	2,41%	0,00%	0,00%	0,00%	0,00%	0,00%
—▲— Reduction ICB	52,25%	49,48%	43,60%	17,30%	12,80%	10,03%

Features Combination in two views

(a)

—□— Sn (3 Views)	95,83%	98,04%	96,00%	99,14%	99,11%	100,00
—○— 1-Sp (3 Views)	25,00%	16,67%	26,67%	22,22%	0,00%	6,67%
—▲— Reduction ICB	81,55%	61,17%	79,65%	89,38%	75,29%	75,72%

Features Combination in three views

(b)

Fig. 7. Sensitivity and 1-Specificity performance of the ICB classifier, and percentage of fails reduction of ICB classifier, (a) for two views , b) for three views

6 Conclusions

This investigation presents the development of a new flaw detection algorithm in manufactured goods using an non-calibrated sequence of images. Using new AMVI methodology with the ICB elimination system, we have designed a novel system of automatic calibration based only on the spatial positions of the structures. 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 they remain stable in their position relative to the movement of the object. With respect to the investigation carried out in [8], we have introduced the calculation of corresponding points generated artificially through the maximization of Pearson's correlation coefficient [15] for two curves. 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 them. In this way we can use as reference points the edges of the structures or areas with no loss of information using a nonlinear method. The main advantage of our model is the automatic estimation of movement. Our future aim is to reduce the number of false alarms by means of a method of final verification of the flaws in correspondence, and an analysis of the ICB classification method with other ensemble classification and probabilistic techniques.

References

1. Malamas, E.N., Petrakis, E.G., Zervakis, M.: A survey on industrial vision systems, applications and tools. Image and Vision Computing 21(2), 171–188 (2003)
2. Newman, T.S., Jain, A.K.: A survey of automated visual inspection. Computer Vision and Image Understanding 61(2), 231–262 (1995)
3. Chin, R.T.: Automated visual inspection: 1981-1987. Computer Vision Graphics Image Process 41, 346–381 (1988)

4. Nakagawa, M., Ohnishi, K., Nakayasu, H.: Human-oriented image recognition for industrial inspection system. In: IEEE International Workshop on Robot and Human Interactive Communication, pp. 52–56. IEEE Computer Society Press, Los Alamitos (2000)

5. Brandt, F.: The use of x-ray inspection techniques to improve quality and reduce costs. The eJournal of Nondestructive Testing and Ultrasonics 5(5) (2000)

6. Mery, D., Filbert, D.: Automated flaw detection in aluminum castings based on the tracking of potential defects in a radioscopic image sequence. IEEE Trans. Robotics and Automation 18(6), 890–901 (2002)

7. Mery, D., Carrasco, M.: Automated multiple view inspection based on uncalibrated image sequence. In: Kalviainen, H., Parkkinen, J., Kaarna, A. (eds.) SCIA 2005. LNCS, vol. 3540, pp. 1238–1247. Springer, Heidelberg (2005)

8. Carrasco, M., Mery, D.: Automated visual inspection using trifocal analysis in an uncalibrated sequence of images. Materials Evaluation 64(9), 900–906 (2006)

9. Pizarro, L., Mery, D., Delpiano, R., Carrasco, M.: Robust automated multiple view inspection. Pattern Analysis and Applications

10. Mery, D.: Crossing line profile: a new approach to detecting defects in aluminium castings. In: Bigun, J., Gustavsson, T. (eds.) SCIA 2003. LNCS, vol. 2749, pp. 725–732. Springer, Heidelberg (2003)

11. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision, 1st edn. Cambridge University Press, Cambridge, UK (2000)

12. Faugueras, O., Luong, Q.: The geometry of multiples images. MIT Press, Cambridge (2001)

13. Mery, D.: High contrast pixels: a new feature for defect detection in x-ray testing. Insight 46(12), 751–753 (2006)

14. Mery, D.: Exploiting multiple view geometry in x-ray testing: Part I, theory. Materials Evaluation 61(11), 1226–1233 (2003)

15. Dunn, O.J., Clark, V.A.: Applied statistics: analysis of variance and regression. Wiley, Chichester (1974)

16. Haralick, R., Shapiro, L.: Computer and Robot Vision. Addison-Wesley Publishing Co., New York (1992)

17. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis and Machine Vision, 2nd edn. PWS Publishing, Pacific Grove, CA (1999)

18. Persoon, E., Fu, K.S.: Shape discrimination using fourier descriptors. IEEE Transactions on System, Man and Cybernetics, 170–179 (1977)

19. Castleman, K.R.: Digital Image Processing. Prentice-Hall, New Jersey (1996)

20. Chen, Z., Wu, C., Shen, P., Liu, Y., Quan, L.: A robust algorithhm to estimate the fundamental matrix. Patter Recognition Letters 21, 851–861 (2000)

21. Shashua, A., Werman, M.: Trilinearity of three perspective views and its associated tensor. In: 5th International Conference on Computer Vision (ICCV 1995), Boston MA, p. 920 (1995)

22. Polikar, R.: Ensemble systems in decision making. IEEE Circuits and Systems Magazine 6(3), 21–45 (2006)

23. Mitchel, T.M.: Machine Learning. McGraw-Hill, Boston (1997)

24. Egan, J.: Signal detection theory and ROC analysis. Academic Press, New York (1975)

25. Duda, R.O., Hart, P.E., Stork, D.G: Pattern Classification, 2nd edn. John Wiley & Sons, Inc., New York (2001)

26. Stearns, S.D.: On selecting features for patterns classifiers. In: IAPR International Conference on Pattern Recognition, pp. 71–75 (1976)

27. Lagarias, J.C., Reeds, J.A., Wright, M.H., Wright, P.E.: Convergence properties of the nelder-mead simplex method in low dimensions. SIAM Journal of Optimization 9(1), 112–147 (1998)