Advances on Automated Multiple View Inspection

Domingo Mery and Miguel Carrasco

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

Abstract. Automated visual inspection is defined as a quality control task that determines automatically if a product, or test object, deviates from a given set of specifications using visual data. In the last 25 years, many research directions in this field have been exploited, some very different principles have been adopted and a wide variety of algorithms have been appeared in the literature. However, automated visual inspection systems still suffer from i) detection accuracy, because there is a fundamental trade off between false alarms and miss detections; and ii) strong bottleneck derived from mechanical speed and from high computational cost. For this reasons, automated visual inspection remains an open question. In this sense, *Automated Multiple View Inspection*, a robust method that uses redundant views of the test object to perform the inspection task, is opening up new possibilities in inspection field by taking into account the useful information about the correspondence between the different views. This strategy is very robust because in first step it identifies potential defects in each view and in second step it finds correspondences between potential defects, and only those that are matched in different views are detected as real defects. In this paper, we review the advances done in this field giving an overview of the multiple view methodology and showing experimental results obtained on real data.

Keywords: automated visual inspection, multiple view geometry, industrial applications.

1 Introduction

Visual inspection is defined as a quality control task that determines if a product deviates from a given set of specifications using visual data[1]. Inspection usually involves measurement of specific part features such as assembly integrity, surface finish and geometric dimensions. If the measurement lies within a determined tolerance, the inspection process considers the product as accepted for use. In

---

[1] For an extended overview of automated visual inspection, the reader is referred to excellent review papers by Malamas et al. [1] and Newman and Jain [2]. The information given in this paragraph was extracted from these papers.

industrial environments, inspection is performed by human inspectors or automated visual inspection systems. Although humans can do the job better than machines in many cases, they are slower than the machines and get tired quickly. Additionally, human inspectors are not always consistent and effective evaluators of products because inspection tasks are monotonous and exhausting, even for the best-trained experts. Typically, there is one rejected in hundreds of accepted products. Moreover, human experts are difficult to find or maintain in an industry, require training and their skills may take time to develop. It has been reported that human visual inspection is at best 80% effective. In addition, achieving human 100%-inspection, where it is necessary to check every product thoroughly to ensure the safety of consumers, typically requires high level of redundancy, thus increasing the cost and time for inspection. For instance, human visual inspection has been estimated to account for 10% or more of the total labor costs for manufactured products. Moreover, in some environments (*e.g.*, underwater inspection, nuclear industry, chemical industry, etc.) human visual inspection may be difficult or dangerous. For these reasons, computer vision has been gradually replacing more and more human inspection.

Comprehensive reviews on automated visual inspection are given in [1,2,3,4]. According to these surveys, approaches developed for automated visual inspection are tailored to the inspection task, *i.e.*, there is no general approach applicable to all cases because the development is an *ad hoc* process. Although there are several approaches that have been developed in the last 25 years, automated visual inspection systems still suffer from i) detection accuracy, because there is a fundamental trade off between false alarms and miss detections; and ii) strong bottleneck derived from mechanical speed and from high computational cost. For this reasons, automated visual inspection remains an open question. In this paper, we present recent advances on *Automated Multiple View Inspection*, a robust method that uses redundant views to perform the inspection task. This novel strategy is opening up new possibilities in inspection field by taking into account the useful information about the correspondence between the different views of the test object. It is very robust because in first step it identifies potential defects in each view and in second step it finds correspondences between potential defects, and only those that are matched in different views are detected as real defects. The paper gives an overview of the multiple view methodology and show experimental results obtained on real data.

2   General Overview of the Multiple View Approach

The principle aspects of an automated multiple view inspection system are shown in Fig. 1. Typically, it comprises the following five steps: i) a manipulation system for handling the test piece (manipulator, robot, etc.), ii) an energy source (light, X-ray, etc.), which irradiates the object under test with, iii) image acquisition system (CCD cameras, image intensifier, etc.) that register digital images of the test piece, and iv) a computer to perform the digital analysis of the images and to classify the test piece accepting or rejecting it.
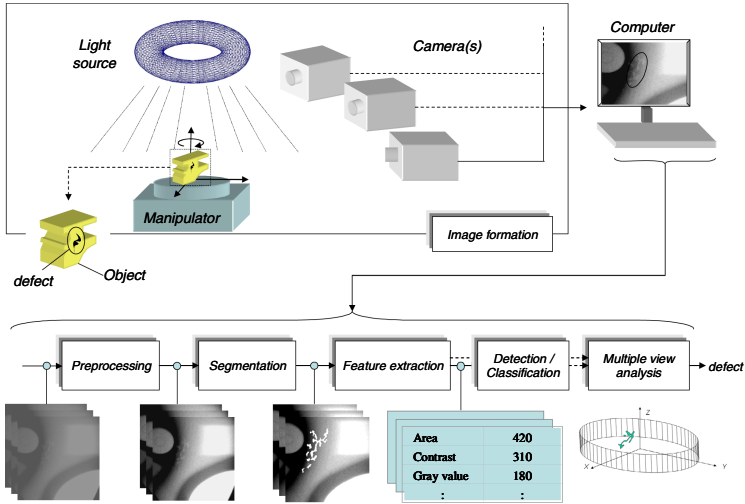
Fig. 1. Computer vision system for automated visual inspection

In the computer-aided inspection, our aim is to identify defects automatically using computer vision techniques. The general automated inspection process, presented in Fig. 1, consists of image formation, preprocessing, segmentation, feature extraction, detection/classification and multiple view analysis [5]. Typically, automated visual inspection using only one view does not follows the last step. The mentioned six steps are explained in further detail:

i) Image formation: Images of the test object are taken and stored in the computer. The human eye is only capable of resolving around 40 grey levels [6], however in automated visual inspection grey level resolution must be a minimum of 2^8 levels. In some applications with X-rays, 2^{16} grey levels are used [7], which allows one to evaluate both very dark and very bright regions in the same image. On the other hand, color image systems are able to capture images in several color spaces with 2^{24} different colors [8,9]. Nowadays, a digital image used in automated visual inspection contains usually more than 2^{20} pixels.

ii) Image preprocessing: The quality of the images is improved using contrast enhancement, noise removal and image restoration techniques. Typically, image enhancement is achieved by histogram manipulation, and noise removal by frame averaging or edge-preserving filtering [6]. Edge-preserving filtering is important for defect detection, because it is desirable to smooth the noise without blurring the edges. Moreover, image restoration involves recovering detail in severely blurred images, which is possible when the causes of the imperfections are known *a-priori* [10,11]. This knowledge may exist as an analytical model, or as *a-priori* information in conjunction with knowledge (or assumptions) of the physical system that provided the imaging process in the first place.

iii) Image segmentation: The digital images are divided into disjoint regions with the purpose of separating the parts of interest from the rest of the scene. Image segmentation plays one of the most important roles in real world computer vision systems. In the last 40 years, this field has experienced significant growth and progress [12][2]. According to [13], monochrome image segmentation techniques are classified into the following categories: histogram thresholding, feature space analysis based methods, edge detection based methods, region based methods, fuzzy logic techniques and neural networks, and pointed out that most of them can be extended to color images by representing color information in appropriate color spaces. However, better performance in segmentation of color images is achieved using vector-value techniques that treat the color information as color vectors in a vector space provided with a vector norm [14]. In image segmentation for detecting defects we aim to separate potential defects from background.

iv) Feature extraction: Since some structural parts of the object could be erroneously segmented as defectively regions, we denoted them as *potential defects*. Subsequently, additional steps are required to eliminate the false alarms of the potential defects. The first of these steps is feature extraction, which is centered principally around the measurement of geometric and chromatic characteristics of regions. Contrast based on crossing line profiles [15] and texture [16] are very helpful to distinct defectively regions from its neighbors. After feature extraction, it is important to know which features provide relevant information about defects. For this reason, a feature selection [17] is performed to find the best subset of the input future set that separates the real defects from the false alarms. Methods based on sequential forward/backward selection achieve effective and fast results but they are suboptimal [18]. On the other hand, a branch and bound method guarantees to find the optimal subset, although the complexity is greater than the mentioned methods, it can be reduced considerably using a fast technique [19].

v) Detection/classification: The extracted (and selected) features of each region are analyzed in order to detect or classify the existing defects. We differentiate between detection and classification of defects. Detection corresponds to a binary classification, because in the detection problem, the classes that exist are only two: 'defects' or 'no-defects', whereas the recognition of the type of defect (*e.g.*, voids, cracks, bubbles, inclusions and slags) is known as classification of defect types [20]. Normally, the 'defect' class constitutes a very small fraction of the total search area. Therefore, the 'defect' class will be either empty or sparsely populated [21]. This implies that there are not sufficient data to train a statistical classifier or statistically evaluate the performance of a detector. In these cases, where the defect probability is very small, minimization of the error probability is not a good criterion of performance, because it can be minimized by classifying every region as 'no-defect' (in a domain where the classes are distributed in a 1:99 ratio (skew $= 10^2$), the maximum likelihood gives 99%

---

[2] Only last year, 194 papers with the word 'image' and 'segmentation' in the title field were indexed by the Web of Science of ISI.

accuracy). For this reason, the probability of detection is typically maximized while keeping the probability of false alarms under a certain predefined value (Neyman-Pearson criterion [22]). Other applications with skewed class distribution can be found in fraud detection [23] or target detection in hyperspectral imaging [24]. In order to increase the samples of the defect class, simulation of defects can be used [25]. The classifier is designed using well known pattern recognition techniques, that can be categorize into generative and discriminative approaches [26]. Generative learning focuses on generative description of samples and tend to synthesize configurations from them. Principal component analysis [17], linear discriminant analysis [17] and hidden Markov models [27] are typical generative classifiers that produce a probability density model for pattern recognition. On the other hand, discriminative learning attempts to compute the mapping for classification from input to output directly without modeling the underlying distributions. It normally achieves superior performance than generative approach in many applications. Traditional neural networks [28] and support vector machines [29] are discriminative classifiers that attempt to maximize the classification boundary margin of classes for recognition. Recent research on combining generative and discriminative learning has shown that proper combinations of two models outperforms pure generative or discriminative models [26,30,31].

vi) Multiple view analysis: Multiple view geometry is increasingly being used in machine vision [32]. It describes explicit and implicit models which relates the 3D coordinates of an object to the 2D coordinates of the digital image pixel, the geometric and algebraic constraints between two, three and more images taken at different projections of the object, and the problem of 3D reconstruction from N views. Since in last step certain 'no-defects' could be classified erroneously as 'defects', we use multiple view geometry as a final discrimination step. The key idea is to gain more information about the test object by analyzing multiple views taken at different viewpoints. Thus, the attempt is made to match or track the remaining potential defects along the multiple views. The existing defects can be effectively tracked in the image sequence because they are located in the positions dictated by geometric conditions. In contrast, false alarms can be successfully eliminated in this manner, since they do not appear in the predicted places on the following images and, thus, cannot be tracked. The tracking in the image sequence is performed using algebraic multi-focal constraints: bifocal (epipolar) and trifocal constraints among others [32,33,34]. Multiple view analysis is a useful and powerful alternative for examining complex objects were uncertainty can lead to misinterpretation, because two or more views of the same object taken from different viewpoints can be used to confirm and improve the diagnostic done by analyzing only one image [33,34].

Finally, the performance of an automated visual inspection method is assessed using a validation technique (*e.g.*, cross-validation, bootstrap and jackknife [17,35]). Usually, some of the collected cases are removed before training begins. Then when training is performed, the cases that were initially removed can be used to test the performance of the inspection method on these test data.

Thus, one can evaluates how well the method will inspect the test objects that has not already examined. Confidence intervals, where the true values of the misclassification error is expected to fall, can be obtained from the test sets.

3   Implemented Multiple View Approaches

Automated multiple view inspection was implemented in the quality control of aluminum castings of the automotive industry using X-ray images [5]. However, the methodology can be used in the inspection of other manufactured products. In this section we present two approaches that were implemented using calibrated and uncalibrated image sequences and their results obtained on real data.

i) Calibrated Approach: In [36] we performed the tracking using a calibrated image sequence , *i.e.*, the model 3D → 2D was *a-priori* known because it was obtained in an off-line process called calibration [37]. The calibration of an imaging system is the process of estimating the parameters of the model, which is used to determine the projection of the 3D test object into its 2D digital image. This relationship 3D→2D can be modeled with a transfer function $F : R^3 \rightarrow R^2$. Using this model the multi-focal tensors can be calculated in order to evaluate the multi-focal constraints for the correspondences of the potential defects in the image sequence [32]. The calibration was performed using the well-known photogrammetric calibration [38], in which a calibration object whose geometry in 3D space is known with high accuracy. Using this technique a true reconstruction of the 3D space without a scale factor is achieved. In the calibration, we estimate the parameters of a geometric model based on n points whose 3D object coordinates $\mathbf{M}_i$ are known, whose 2D image coordinates $\mathbf{w}_i$ are measured, for $i = 1, ..., n$. Using the model we obtain the reprojected points $\mathbf{w}'_i = F(\mathbf{M}_i, \theta)$, *i.e.*, the inferred projections in the digital image computed from the calibration points $\mathbf{M}_i$ and a parameter vector θ. The calibration is performed in each image of the sequence by minimizing the objective function defined as the mean-square discrepancy between measured points $\mathbf{w}_i$ and inferred points $\mathbf{w}'_i$ [32]. Usually, the calibration problem is a non-linear optimization problem. In general, the minimization of the objective function has no closed-form solution. For this reason, it must be iteratively minimized starting with an initial guess θ_0 that can be obtained from nominal values or preliminary reference measurements.

ii) Uncalibrated Approach: The calibration is a very difficult task because the iterative estimation of the parameters is very sensible to the initial guess. In addition, the vibrations of the imaging system induce inaccuracies in the estimated parameters of the model, *i.e.*, the calibration is not stable and the computer vision system must be calibrated periodically (off-line) in order to avoid uncertainty. For this reason, we developed an approach based on the tracking of potential detects in two views [39,40] and in three views [40] using uncalibrated image sequences, in which it was not necessary to calibrate the imaging system. This new approaches track the potential defects based on a motion model estimated from the image sequence itself. Thus, we obtain a motion model by
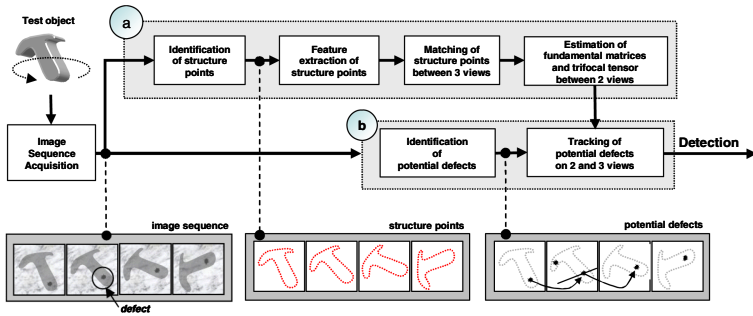
Fig. 2. Block diagram of the uncalibrated automated multiple view inspection: a) estimation of motion model, b) detection of defects [40]

matching structure points of the test object in the images as shown in Fig. 2. The structure points are matched using B-Spline curves and correlated curve sections of the structure (see details in [39] and in [40] respectively). Using RANSAC [32], the matched structure points are employed to estimated the bifocal and trifocal tensors required for the multiple view analysis. In this sense, we do not calibrate the image sequence, we only estimate the bifocal and trifocal tensors required for the tracking. The great disadvantage of this approach is the inherent difficulty in identification of the structure points (and thus the estimation of the motion model) from the test object itself, when the images of the test object do not significantly differ from each other in the sequence, *e.g.*, a glass or a bottle rotating around its vertical axis.

Once the system is calibrated (in the calibration approach) or the motion model is estimated (in the uncalibrated approach) the same algorithm is used to track the potential defects [41]. The tracking algorithm requires the bifocal and trifocal tensors [32] between the views. In the first approach the tensors are obtained from the projection matrices estimated after the calibration, whereas in the second approach the tensors are obtained using corresponding points of the test object in two and three views.

Table 1 summarizes the results obtained on real data using calibrated and uncalibrated approaches. We calculate the performance of the identification and the performance of the tracking separately. *True positives* are the number of defects correctly detected. The true positive percentage is calculated related to the number of the existing defects. *False positives* (or false alarms) correspond to the number of 'no-defects' misclassified as 'defects'. The false positive percentage is given related to the number of detected potential defects. We present three implementations of the calibrated approach. They perform the tracking in three, four and five views (cases C-I, C-II and C-III respectively). We observe that the number of false alarms in the identification is enormous. However, the results are perfect for four views (case C-II) where all defects are detected without any false alarms. The verification of the correspondence on three views flags too many false

Table 1. Performance of calibrated and uncalibrated approaches

Method	Name	Year & Reference	Analyzed Images	Tracked Views	Existing defects	Identification				Tracking			
						True Positives		False Positives		True Positives		False Positives	
	C-I	2002:[36]	72	3	84	71	85%	4310	98%	71	100%	24	25%
Calibrated	C-II	2002:[36]	72	4	84	71	85%	4310	98%	71	100%	0	0%
	C-III	2002:[36]	72	5	84	71	85%	4310	98%	59	83%	0	0%
	U-I	2005:[39]	24	2	39	39	100%	83	68%	36	92%	4	10%
Uncalibrated	U-II	2006:[40]	72	2	233	190	82%	205	52%	190	100%	93	33%
	U-III	2006:[40]	72	3	233	172	74%	205	52%	170	99%	19	10%

alarms. On the other hand, with 5 views we cannot ensure the segmentation of a defect in five views, for this reason some defects cannot be detected. We increase the performance in the segmentation in the uncalibrated approaches reducing the number of false alarms significantly. In case U-I, we perform the tracking in only two views using B-spline curves for the motion model. In case U-II and U-III, the tracking is done in two and three views respectively using correlated curve sections of the structure for the motion model. The results of case U-III are promising because all defects to be tracked, *i.e.*, defects that are present in three views, could be tracked, with only a few number of false alarms. We observe that the performance obtained in calibrated approach is higher, however the calibration is in many cases an excessively difficult and unstable task that can be avoided using an uncalibrated approach.

4   Conclusions

Automated visual inspection remains an open question. Many research directions have been exploited, some very different principles have been adopted and a wide variety of algorithms have been appeared in the literature of automated visual inspection. Although there are several approaches in the last 25 years that have been developed, automated visual inspection systems still suffer from i) detection accuracy, because there is a fundamental trade off between false alarms and miss detections; and ii) strong bottleneck derived from mechanical speed (required to place the test object in the desired positions) and from high computational cost (to determine whether the test object is defective or not). In this sense, Automated Multiple View Inspection offers a robust alternative method that uses redundant views to perform the inspection task. We believe that the method is opening up new possibilities in inspection field by taking into account the useful information about the correspondence between the different views of the test object. Two approaches were developed in the last years: the calibrated and the uncalibrated approaches. Both of them achieve very good performance. However, the calibration of the first approach is a very complicated task, and the identification of structure points in the second approach is inherently difficult when the images of the test object do not significantly differ from each other in the sequence. In order to avoid the mentioned problems, we are working on an on-line calibration of the multiple view system using a calibration object attached to the test object which is imaged in all views. Thus, the images have an enough number of points to calibrate the system.

Acknowledgments

References

1. Malamas, E., Petrakis, E., Zervakis, M.: A survey on industrial vision systems, applications and tools. Image and Vision Computing **21** (2003) 171–188
2. Newman, T., Jain, A.: A survey of automated visual inspection. Computer Vision and Image Understanding **61** (1995) 231–262
3. Chin, R.: Automated visual inspection: 1981-1987. Computer Vision Graphics Image Process **41** (1988) 346–381
4. Chin, R.T., H.C.: Automated visual inspection: A survey. IEEE Trans. Pattern Analysis and Machine Intelligence **4** (1982) 557–573
5. Mery, D.: Automated radioscopic testing of aluminum casting. Materials Evaluation **64** (2006) 135–143
6. Castleman, K.: Digital image processing. Prentice-Hall, Englewood Cliffs, New Jersey (1996)
7. Purschke, M.: IQI-sensitivity and applications of flat panel detectors and X-ray image intensifiers - a comparison. Insight **44** (2002) 628–630
8. Davis, E.: Machine Vision. 3 edn. Morgan Kaufmann Publishers, Amsterdam (2005)
9. Trussell, H., Saber, E., Vrhel, M.: Color image processing. IEEE Signal Processing Magazine **22** (2005) 14–22
10. Nagy, J., Palmer, K., Perrone, L.: Iterative methods for image deblurring: A Matlab object-oriented approach. Numerical Algorithms **36** (2005) 73–93
11. Banham, M., Katsaggelos, A.: Digital image restoration. IEEE Signal Processing Magazine **14** (1997) 24–40
12. Zhang, Y.J.: An Overview of Image and Video Segmentation in the Last 40 Years. In: Advances in Image and Video Segmentation, Zhang, Y.-J. (Ed.). IRM Press, Idea Group Inc. Hershey (2006) 1–15
13. Cheng, H., Jiang, X., Sun, Y., Wang, J.: Color image segmentation: advances and prospects. Pattern Recognition **34** (2001) 2259–2281
14. Koschan, A., Abidi, M.: Detection and classification of edges in color images: A review of vector-valued techniques. IEEE Signal Processing Magazine **22** (2005) 64–73
15. Mery, D.: Crossing line profile: a new approach to detecting defects in aluminium castings. Lecture Notes in Computer Science **2749** (2003) 725–732
16. Mery, D., Berti, M.: Automatic detection of welding defects using texture features. Insight **45** (2003) 676–681
17. Webb, A.: Statistical Pattern Recognition. 2 edn. John Wiley and Sons Ltd., New Jersey (2005)
18. Jain, A., Duin, R., Mao, J.: Statistical pattern recognition: A review. IEEE Trans. Pattern Analysis and Machine Intelligence **22** (2000) 4–37
19. Somol, P., Pudil, P., Kittler, J.: Fast branch and bound algorithms for optimal feature selection. IEEE Trans. on Pattern Analysis and Machine Intelligence **26** (2004) 900–912
20. Mery, D., da Silva, R., Caloba, L., Rebello, J.: Pattern recognition in the automatic inspection of aluminium castings. Insight **45** (2003) 475–483

21. Carvajal, K., Chacón, M., Mery, D., Acuña, G.: Neural network method for failure detection with skewed class distribution. Insight **46** (2004) 399–402
22. Kay, S.: Fundamentals of Statistical Signal Processing: Detection Theory. Prentice Hall Signal, Processing Series, Volume 2, New Jersey (1998)
23. Provost, F., Fawcett, T.: Robust classification for imprecise environments. Machine Learning Journal **42** (2001) 203–231
24. Manolakis, D., Shaw, G.: Detection algorithms for hyperspectral imaging applications. IEEE Signal Processing Magazine **19** (2002) 29–43
25. Mery, D., Hahn, D., Hitschfeld, N.: Simulation of defects in aluminium castings using cad models of flaws and real X-ray images. Insight **47** (2005) 618–624
26. Raina, R., Shen, Y., Ng, A.Y., McCallum, A.: Classification with hybrid generative/discriminative models. In: Advances in neural information processing systems 16, S. Thrun, L. Saul , and B. Schlkopf (Eds.). MIT Press, In Cambridge, MA (2003)
27. Theodoris, S., Koutroumbas, K.: Pattern Recognition. 2 edn. Academic Press, Elsevier, Amsterdam (2003)
28. Bishop, C.: Neural Network for Pattern Recognition. Oxford University Press Inc., New York (1997)
29. Shawe-Taylo, J., Cristianini, N.: Kernel Methods for Pattern Analysis. Cambridge University Press, UK (2005)
30. Kuncheva, L.: A theoretical study on six classifier fusion strategies. IEEE Trans. on Pattern Analysis and Machine Learning **24** (2002) 281–286
31. Mery, D., Chacón, M., Muñoz, L., Gonzalez, L.: Automated inspection of aluminium castings using fusion strategies. Materials Evaluation **63** (2005) 148–153
32. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, UK (2000)
33. Mery, D.: Exploiting multiple view geometry in X-ray testing: Part I: Theory. Materials Evaluation **61** (2003) 1226–1233
34. Mery, D.: Exploiting multiple view geometry in X-ray testing: Part II: Applications. Materials Evaluation **61** (2003) 1311–1314
35. Mitchell, T.: Machine Learning. McGraw-Hill, Boston (1997)
36. 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** (2002) 890–901
37. Mery, D.: Explicit geometric model of a radioscopic imaging system. NDT & E International **36** (2003) 587–599
38. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. on Pattern Analysis and Machine Intelligence **22** (2000) 1330–1334
39. Mery, D., Carrasco, M.: Automated multiple view inspection based on uncalibrated image sequences. Lecture Notes in Computer Science **3540** (2005) 1238–1247
40. Carrasco, M., Mery, D.: Automated visual inspection using trifocal analysis in an uncalibrated sequence of images. Materials Evaluation **64** (2006) In Press
41. Mery, D., Ochoa, F., Vidal, R.: Tracking of points in a calibrated and noisy image sequence. Lecture Notes in Computer Science **3211** (2004) 647–654