Flaw Detection in Aluminum Die Castings Using Simultaneous Combination of Multiple Views

Christian Pieringer-Baeza, Domingo Mery
Department of Computer Science
Universidad Católica de Chile (PUC)
Santiago, Chile
cppierin@uc.cl, dmery@ing.puc.cl

***Abstract*—Recently, X-rays have been adopted as the principal non-destructive testing method to identify flaws within an object which are undetectable to the naked eye. Automatic inspection using radiographic images has been made possible by incorporating image processing techniques into the process. In a previous work, we proposed a framework to detect flaws in aluminum castings using multiple views. The process consisted of flaw segmentation, matching, and finally tracking the flaws along the image sequence. While the previous approach required effective segmentation and matching algorithms, this investigation focuses on a new detection approach. The proposed method combines, simultaneously, information gathered from multiple views of the scene, this does not require searching for correspondences or matching. By gathering all the projections from a 3D point, obtained from a sliding box in the 3D space, we train a classifier to learn to detect simulated flaws using all the evidence available. This paper describes our proposed method and presents its performance record in flaw detections using various classifiers. Our approach yields promising results, 94% of true positives detected with 95% sensitivity in real flaws. We conclude that simultaneously combining information from different points of view is a robust approach to flaw identification.**

***Index Terms*—Automated inspection, flaw detection, saliency, computer vision, multiple views.**

I. Introduction

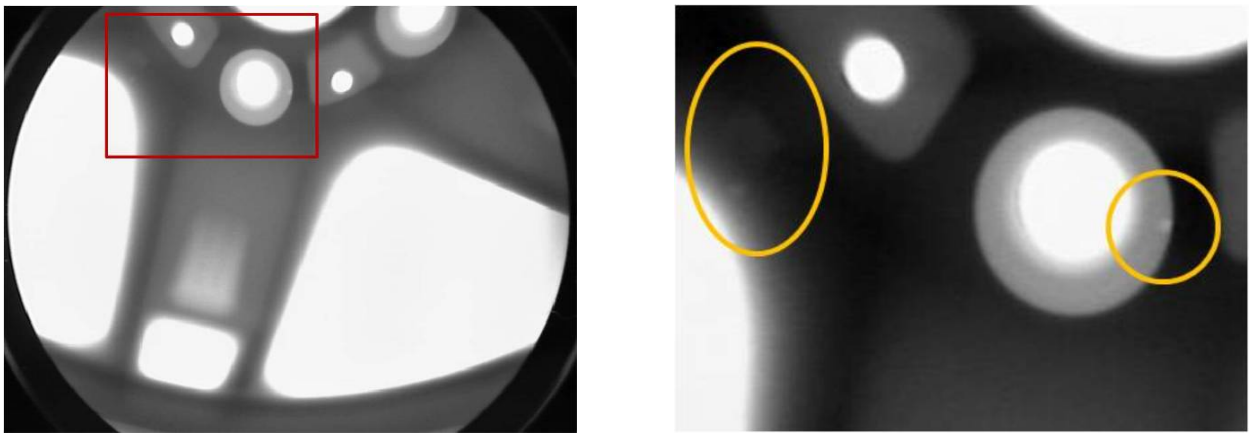
Radioscopy has been embraced as the best tool for non-destructive testing (NDT) in industrial production given that most defects are not visible on the object's surface [1]. The material defects which occur during the casting process must be detected in order to satisfy safety requirements, consequently it is necessary to check 100% of the parts. Even though X-rays detect flaws in cast pieces, they often manifest as small and low contrast objects which are difficult to detect as seen in Fig. 1(a). Due to these difficulties it is necessary to incorporate image processing techniques that accurately highlight flaws, while separating them from the background, then finally classifying the flaws correctly.

A typical automated X-ray system is schematically presented and described in [1]. The process is generally performed in five steps:

- The *manipulator* places the casting in the desired position.
- The *X-ray tube* generates X-rays which pass through the casting.
- The X-rays are detected by a fluorescent entrance screen in the *image intensifier*, amplified and depicted onto a phosphor screen.
- The image intensifier converts the X-rays to a visible radioscopic image.
- The guided and focused image is registered by the *CCD-camera*. The *image processor* converts the analog video signal, transferred by the CCD-camera, into a digital data stream. Digital image processing is used to improve and evaluate the radioscopic image.

New X-Ray techniques utilize flat amorphous silicon detectors as image sensors in industrial inspection systems [2]. These detectors use a semi-conductor to convert energy from the X-ray into an electrical signal without an image intensifier. However due to their high cost, NDT using flat detectors is not as feasible as the use of image intensifiers.

Various approaches to automated flaw detection in aluminum castings can be found in [1], [3]–[6]. Those works proposed flaw detection methods without obtaining a priori information of flaws or the object's structure. They developed methods to manage the lack of a priori information because in real processes flaws are rare, making it is extremely difficult to get samples. However those methods are mainly supported in the pre-processing of the obtained images meaning they are dependent on the parameters of the processing algorithms and therefore



(a) (b)

Figure 1: Flaw example in an aluminum wheel. (a) X-Ray image of aluminum wheel. (b) Magnification of the flaw areas. Circles denote a flaw within the piece.

inflexible to the variations of image intensity. This is where our research differs, we explore a new method of training classifiers with simulated samples of flaws which are easier to obtain than real samples, and then proceed to test the system by applying real flaws [7].

Multiple views drastically improve a systems detection rate by eliminating numerous false alarms, mainly because they provide additional or complementary information about the object being tested. This process involves using a sequence of X-ray images of a casting, all of which are taken from different positions. The next step is the segmentation of a flaw in one view followed by its subsequent tracking throughout the sequence. This approach can be applied to either calibrated or non calibrated sequences [1], [8].

Despite their advantages, the aforementioned methods still require effective segmentation in the first step in order to generate possible flaws to be tracked in following views. In addition, methods based on geometrical constraints to track the hypothetical flaws along the sequence, such as epipolar or trifocal tensors, require robust matching algorithms throughout different views. Motivated to eliminate these disadvantages, we investigate a new approach to combine information from multiple-views which allows for flexible learning without requiring a priori information of the object's structure.

We propose using a sliding box, which moves within 3D space occupied by the casting object, to gather all projections from multiple views of this local space, Fig. 2(a). Our approach allows us to avoid matching because all the projections from the box are locally corresponding. In the following step we create a coarse detection of each flaw from the projections by using a saliency detector as in [9]. Each detection is represented by two kinds of feature descriptors as proposed in [10]: one for shape and one for appearance. Finally, we combine the responses of individual classifiers from all views which results in a final classification of each flaw sequence.

This paper is organized as follows. Section II introduces details of our approach, such as the searching method within the 3D space, the features used to describe aluminum casting flaws, as well as our methodology for combining information from multiple views. Sections III and IV discuss our dataset and results, respectively. Finally, Section V presents conclusions about our approach.

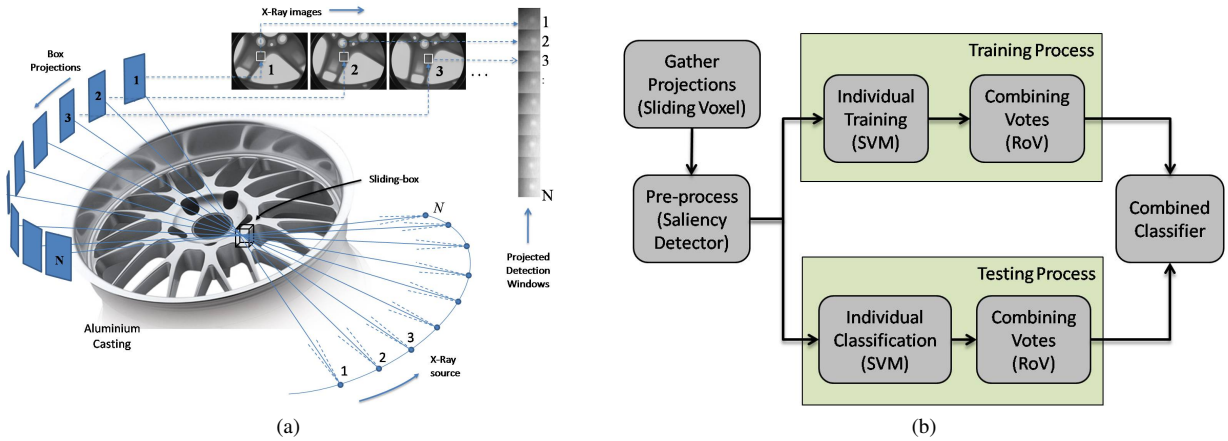
II. Proposed Method

In a previous work, we proposed a framework for tracking flaws in castings [1]. This method has three main parts: flaw segmentation, matching of candidates in different views, and finally tracking of flaws in the image sequence. The system eliminates the flaws that are not tracked in thrird and fourth views.

Our method is based on the principal that it is possible to get results similar to previous works described in the introduction by:

1) replacing the segmentation stage with a more flexible approach based on object detection and recognition, and
2) replacing the matching and tracking stages with a complete and simultaneous analysis of the scene.

An automated flaw detection system first requires samples of the object being tested in order to train a classifier. Once this classifier is fully trained, it is used to detect previously unseen flaws, Fig 2(b). In the next four subsections we explain the main parts of our algorithm.



(a) (b)

Figure 2: Proposed Methodology. (a) Shows process for gathering informaiton from multiple views. (b) Process diagram of the proposed method.

A. *The Sliding Box*

The state-of-the-art approach to machine learning, for object detection and recognition in single images, consists of running a sliding window over an image at different scales and finally detecting the class of interest within the image [11], [12].

We propose a slightly different approach to integrate information, recieved from a sliding box in a multiple view scheme, than methods which use epipolar geometry to establish correspondence and use matching as in [1], [4], [8]. A sliding box is used to scan the 3D space of the scene. Every time the box changes its position, projections from various points of view are simultaneously gathered. Once received, we implement a method to effectively combine the relevant information from multiple views.

Our automatic inspection method uses a calibrated X-Ray system, allowing us to project all 3D points within their corresponding view [13]. In essence, as the box moves through the space occupied by the casting, we project its vertices over each available view in the system. In the following step, we select the regions from each view that encompass the most projected points. Finally, each projected region which is considered valid are grouped together and arranged into a sequence that shows the flaw from all available views, as shown in Fig. 3.

B. *Image Pre-Processing and Saliency Detector*

The proposed method, based on object classification, requires the training of a classifier with flaw samples. However, this is difficult because flaws are very rare occurrences within images. They are generally difficult to detect because they appear within the image as a small section of low contrast pixels, even after being isolated in the flaw sequence. Before training the classifier it is necessary to obtain an extremely detailed sample. To do this, we extract the flaw's structure and remove the majority of background details.

At this stage we apply a salient detector to the inverse image of the selected projection. This detector doesn't use parameters, in fact it can adapt automatically to varying conditions of contrast in the samples. The results are coarser than the segmentation used in [1], [4], but it is still suitable for our approach due to its freedom of parameters. The effects of applying the detector over the original image and inverse image are different, see Fig. 4(c) and Fig. 4(f). Even though both results deliver salient zones, those obtained from the inverse image allow us to correctly and precisely isolate the flaw. Our saliency detector is based on the visual attention systems proposed in [9], [14]. Visual attention systems are inspired by primate visual systems, and accordingly we believe they better emulate inspections made by humans in flaw classification tasks. Saliency detection consists of computing local contrast between a specific region within an image and its surroundings using one or more features such as color, intensity and orientation. Thus the saliency of a region is high when its properties differ considerably from the rest of the image [9].

The detection process begins by obtaining the inverse images of the grey scale patches from each projection, as shown in Fig. 4(d). Then, we apply the saliency detector to the modified patches. This results in a saliency map, see Fig. 4(c) and Fig. 4(f). We then determine a threshold to produce a binary image to remove noise from the edges of the saliency map, Fig. 4(i). Finally, we choose a Region of Interest (ROI) around the maximum value of the saliency map Fig. 4(g) and Fig. 4(h). By trial and error we determined that the ROI served us best when set as 40 pixels. Projections without detections are classified

Figure 3: Flaw sequence constructed with information from multiple views. The display range was extended to facilitate viewing the results

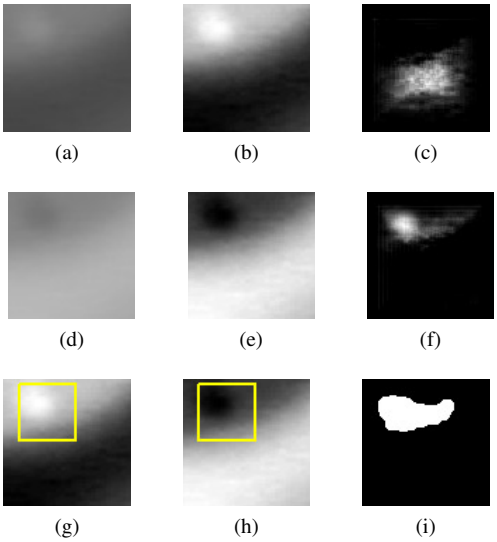


Figure 4: Results of saliency detector applied to original image and its complement. (a) Patch of the original flaw. (d) Complement version of the patch. (b) and (e) Extended display range of the original and complement patches, respectively. (c) and (f) Saliency maps.

as non flaws.

C. Features

Once we have an accurate flaw detected, after applying the saliency detector, we are able to extract three sets of features from the samples: Crossing Line Profile (CLP) [15], Pyramidal Histograms of Oriented Gradients (PHOG) [10], and Histograms of Oriented Gradients based on SIFT descriptor [16].

The selected features are used to represent both shape and appearance as in [10]. The features vector for shape was constructed by concatenating features of CLP and PHOG. These features retrieve the circular shape of the flaws. The dimensions of feature vectors consist of 350 for shape and 128 for appearance.

CLP is defined as the best grey level profile along the length of a straight line within a region of interest (ROI). Eight profiles, distributed every $\pi/8$, are calculated. Each profile is normalized with respect to its average and standard deviation. Fourier Transform is extracted at the best profile and its Fourier coefficients are considered as features vector.

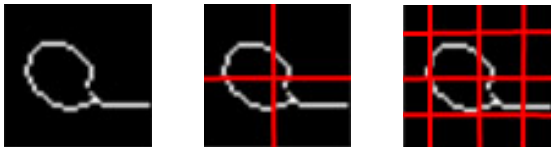


Figure 5: Results of applying PHOG onto saliency maps. (a) Level=0. (b) Level=1. (c) Level=2.

CLP has been proven to effectively represent circular flaws in radioscopic images. We combine the shape features vector and the vector with the best profile inside the bounding box with the flaw detection in the saliency image.

PHOG features have been successfully utilized in investigations dealing with recognition and classification of objects. These features use histograms to encode the gradient information in regards to the defined border of the object at different pyramidal levels. They perform better than the Chamfer Distance, which performs a template matching based on distance transform, because PHOG deals better with rotated images, it is an appropriate compact vector for learning with kernel based algorithms, and is flexible in regards to spatial correspondence. These features were calculated on the saliency map generated by the detector explained in II-B, using only two levels of pyramids to avoid over fitting as suggested in [10], see Fig. 5.

A SIFT descriptor computes locally and projects onto an image the histogram of oriented gradient over a point of interest at a given scale, orientation and position. In our method, we compute the gradient of oriented histograms as local descriptors over the detection bounding box resized to 16x16 pixels with neither scale nor orientation information since that information is already coded in the 3D model.

D. Classification and Combining Multiple-Views

Our classification process consists of two stages: individual classification and joint evaluation utilizing all available views.

First, every element in the flaw sequence as in Fig. 3 was classified individually as a flaw or non-flaw. For this task, we trained two Support Vector Machines (SVMs) with polynomial kernels because SVMs can better manage large feature vectors, as in [10]. These classifiers

	Number of Instances	
Class	Training	Test
Flaws	40	12
Nonflaws	40	37
Total	80	49

Table I: Summary specification of dataset.

were trained with information on appearance and shape, respectively, as we mentioned in section II-C.

Second, the classifiers responses are kept in a vector with values belonging to $\{1, -1\}$. We summarize those values as a Rate of Votes (RoV) computed as the quotient between the sum of flaws and the sum of non flaws (1), where N is the number of views available for each flaw. This rate is used as a feature by a new classifier that then decides whether the sequence represents a true flaw. We calculated two rates: RoV_{shape} and $RoV_{appearance}$, from classification of shape and appearance, respectively. The combination of rates allows us to separate the classes correctly.

$$RoV = \frac{\sum_i^N Flaw\ Vote_i}{\sum_i^N Non\ Flaw\ Vote_i} \tag{1}$$

At this stage of the process we trained six kinds of classifiers: Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Mahalanobis, Artificial Neural Networks (ANN), K-Near Neighbors (KNN), and (SVM) with Linear and Gaussian basis [17]. The best results were achieved by SVM and KNN.

III. Datasets and Methodology

We constructed a training dataset with 80 sequences as shown in Fig. 3, of those sequences 40 were true flaws and 40 non-flaws. The flaws were simulated utilizing the method proposed in [7]. We also constructed a test dataset composed of real flaw sequences previously unseen by the classifier in the training stages. The average number of views for each flaw sequence is 22 out of a possible 72 total views available from the analysis. Table I contains the specifications of each dataset.

The process starts off by extracting the features for shape and appearance, as mentioned in the section II-C, individually from each element of the sequence. For each feature, we get a vector descriptor with U elements of length from every patch. Then, we put all the elements in the training set as a matrix of dimensionality W $\times$ U, where W is the*total number of instances* $\times \sum$ *total number of views available in every trace* and U is the *length of the feature vector*. The same process is applied to the test dataset.

Features were separated into two groups, shape and appearance, as we described in section II-C. We train two classifier SVMs for every group of features using ten fold cross validation [17]. We select the best performing classifier into the training process to use it in the following stages.

Next, each element of the sequence within the training set was classified individually as a flaw or non-flaw. We obtain the respective RoV_{shape} and $RoV_{appearance}$ according to equation (1) using the results of the classifiers. We trained six classifiers mentioned in the section II-D: LDA, QDA, Mahalanobis, ANN, KNN, and SVM.

Finally, we tested our trained classifier with the dataset of real flaws, generating its corresponding RoV.

IV. Experiments and Results

We evaluate our results by applying the standard two class analysis in pattern recognition based on estimating sensitivity (S_n) and specificity (S_p) as defined by equations (2) and (3). There are four possible outcomes of the classifier: TP, TN, FP and FN defined as True Positives, True Negatives, False Positives and False Negatives, respectively. Ideally $S_n = 1$ and $(1 - S_p) = 0$.

As we described in section II-D, the first stage applies an individual assessment of the elements within the flaw sequence. The average performance of the classifiers in the training process was 89.7% for SVM_{shape} and 99% for $SVM_{appearance}$. We use F-Measure to select the best classifier in this stage, as defined in (4). This criterion allows us to characterize the performance in a single measure.

$$S_n = \frac{TP}{TP + FN} \tag{2}$$

$$1 - S_p = \frac{FP}{TN + FP} \tag{3}$$

$$F - Measure = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \tag{4}$$

The final classifier was trained by applying RoV indicators, computed as mentioned in section II-D within the training set. The evaluation was conducted on a new testing set created with real flaws, previously unseen by the training process, both individually and in sequences. The best performance was obtained by the classifier (SVM) Linear, (SVM) with Gaussian kernels and σ = 1, and KNN. Overall, sensitivity is 92% and specificity is 95%. The complete evaluation is summarized in table II.

The features relating to space were constructed with RoV indexes are shown in Fig. 6. It records the correlation between samples of testing and training although there is little difference between the training set and the testing set in the voting process on non-flaw samples.

V. Conclusions

In this research, we developed an approach for fault detection in aluminum castings based on object detection

Classifier	S_n	$1 - S_p$	Correct Rate
LDA	0.8333	0.0000	0.9592
QDA	1.0000	0.4324	0.6735
Mahalanobis	1.0000	0.4324	0.6735
SVM linear	0.9167	0.0541	0.9388
SVM RBF σ = 1	0.9167	0.0541	0.9388
SVM RBF σ = 2	0.8333	0.0000	0.9592
SVM RBF σ = 0.5	0.9167	0.0811	0.9184
KNN	0.9167	0.0541	0.9388
ANN	0.9167	0.1351	0.9167

Table II: Training of $SVM_{Appearance}$.

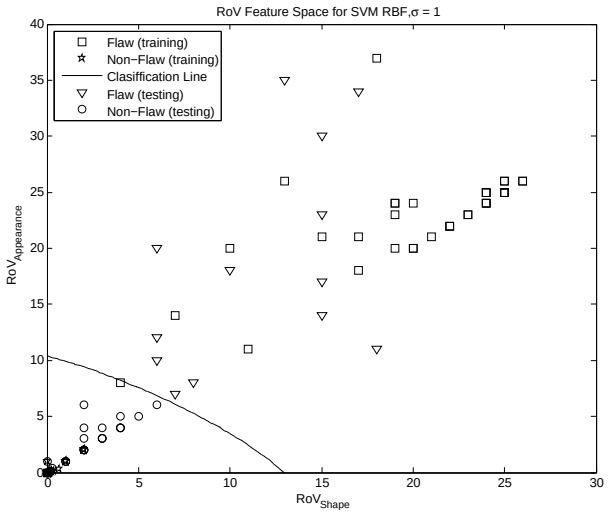


Figure 6: Feature Space for classifier SVM RBF with σ = 1.

methods. Our approach maintains that by using a sliding-box it is possible to integrate information from multiple views and train a classifier with all available information, therefore ruling out a strict segmentation that has been used thus far. This approach allowed us to correctly detect faults without the need for finding correspondences as is classically seen in multiple view schemes.

The results obtained in the preprocessing stage of flaw detection based on saliency encourages the use of this method in non parametrical segmentation of flaws and reinforces our idea of emulating the human visual system of human operators. The detections were successful in the majority of the cases despite variations in the intensity of images, increasing the overall performance of the system.

Correlation between the classification of real samples and simulated ones suggests a novel form of learning which is useful in avoiding the issue of flaw detection when the flaws are uncommon and rare in actuality.

The methodology of machine learning along in addition to other tools gives our system the flexibility to deal with process requirements and therefore the capacity to adapt in the detection of various types of flaws.

We believe the results of our methodology are promising. However, we think that it is very possible to better take advantage of the information available from all views by utilizing models which are more complex. Based on this concept, we are working on a new model of information integration that allows us to define or discover a new structure of the object that we want to describe, that way we will be able to utilize our methodology in other detection problems.

REFERENCES

[1] D. Mery and D. Filbert, "Automated flaw detection in aluminum castings based on the tracking of potential defects in a radioscopic image sequence," *IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION*, vol. 18, no. 6, 2002.

[2] M. Purschke, "IQI-sensitivity and applications of flat panel detectors and X-ray image intensifiers – a comparison," *Insight*, vol. 44, no. 10, pp. 628–630, 2002.

[3] H. Boerner and H. Strecker, "Automated x-ray inspection of aluminum castings," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 1, pp. 79–91, 1988.

[4] D. Mery, D. Filbert, and T. Jaeger, "Image processing for fault detection in aluminum castings," *Analytical Characterization of Aluminum and Its Alloys*, 2002.

[5] D. Mery, T. Jaeger, and D. Filbert, "A review of methods for automated recognition of casting defects," *INSIGHT*, vol. 44, no. 7, pp. 428–436, 2002.

[6] L. Pizarro, D. Mery, R. Delpiano, and M. Carrasco, "Robust automated multiple view inspection," *Pattern Analysis & Applications*, vol. 11, no. 1, pp. 21–32, 2008.

[7] D. Mery, D. Hahn, and N. Hitschfeld, "Simulation of defects in aluminium castings using cad models of flaws and real x-ray images," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 47, no. 10, pp. 618–624, 2005.

[8] M. Carrasco and D. Mery, "Automated visual inspection using trifocal analysis in an uncalibrated sequence of images," *Materials Evaluation*, vol. 64, no. 9, pp. 900–906, 2006.

[9] R. Achanta, F. Estrada, P. Wils, and S. Susstrunk, "Salient region detection and segmentation," *Lecture Notes in Computer Science*, vol. 5008, p. 66, 2008.

[10] A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," pp. 401–408, 2007.

[11] P. Viola and M. Jones, "Robust real-time object detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 37–154, 2004.

[12] N. Dalai, B. Triggs, I. Rhone-Alps, and F. Montbonnot, "Histograms of oriented gradients for human detection," vol. 1, 2005.

[13] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge Univ Pr, 2003.

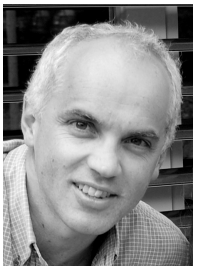
[14] S. Montabone and A. Soto, "Human detection using a mobile platform and novel features derived from a visual saliency mechanism," *Image and Vision Computing*, 2009.

[15] D. Mery, "Crossing line profile: A new approach to detecting defects in aluminium die castings," *Lecture Notes in Computer Science*, pp. 725–732, 2003.

[16] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[17] C. Bishop *et al.*, *Pattern recognition and machine learning*. Springer New York:, 2006.

Christian Pieringer received his B.S degree in Electronic Engineering at the Pontificia Universidad Catï¿½lica de Valparaï¿½so in 2003. At present he is a Ph.D. Candidate in Computer Science at the Pontificia Universidad Catï¿½lica de Chile. He received a scholarship from the Comisiï¿½n Nacional de Investigaciï¿½n Cientï¿½fica y Tecnolï¿½gica (CONICYT). He was an additional reviewer in the 11th IEEE International Conference on High Performance Computing and Communications, 2009.



Dr.-Ing. Domingo Mery Domingo Mery received his M.Sc. degree in Electrical Engineering from the Technical University of Karlsruhe, Germany, in 1992, and his Ph.D. degree with distinction from the Technical University of Berlin, in 2000. He was a Research Scientist at the Institute for Measurement and Automation Technology at the Technical University of Berlin. He was a recipient of a scholarship from the Konrad-Adenauer-Foundation, and of a scholarship from the German Academic Exchange Service (DAAD).