

Document downloaded from:

<http://hdl.handle.net/10251/101898>

This paper must be cited as:



The final publication is available at

<https://doi.org/10.1016/j.rcim.2017.04.009>

Copyright Elsevier

Additional Information

# On the Detection of Defects on Specular Car Body Surfaces

Jaime Molina, J. Ernesto Solanes<sup>1</sup>, Laura Arnal and Josep Tornero

*Universitat Politècnica de València, Institute of design and Manufacturing, Camí de Vera S/N, 46022, València, Spain*

*Email address: {jaimopar,esolanes,lauarbe,jtornero}@idf.upv.es*

---

## Abstract

The automatic detection of small defects (of up to 0.2mm in diameter) on car body surfaces following the painting process is currently one of the greatest issues facing quality control in the automotive industry. Although several systems have been developed during the last decade to provide a solution to this problem, these, to the best of our knowledge, have been focused solely on flat surfaces and have been unable to inspect other parts of the surfaces, namely style lines, edges and corners as well as deep concavities. This paper introduces a novel approach using deflectometry- and vision-based technologies in order to overcome this problem and ensure that the whole area is inspected. Moreover, since our approach, together with the system used, computes defects in less than 15 seconds, it satisfies cycle time production requirements (usually of around 30 seconds per car). Hence, a two-step algorithm is presented here: in the first step, a new *pre-processing* step (image fusion algorithm) is introduced to enhance the contrast between pixels with a low level of intensity (indicating the presence of defects) and those with a high level of intensity (indicating the absence of defects); for the second step, we present a novel *post-processing* step with an image background extraction approach based on a local directional blurring method and a modified image contrast enhancement, which enables detection of defects in the entire illuminated area. In addition, the *post-processing* step is processed several times using a multi-level structure, with computed image backgrounds of different resolution. In doing so, it is possible to detect larger defects, given that each level identifies defects of

---

<sup>1</sup>Corresponding author.

different sizes. Experimental results presented in this paper are obtained from the industrial automatic quality control system *QEyeTunnel* employed in the production line at the Mercedes-Benz factory in Vitoria, Spain. A complete analysis of the algorithm performance will be shown here, together with several tests proving the robustness and reliability of our proposal.

*Keywords:* Automated visual inspection, Image fusion, Specular surfaces, Painted surfaces, Car body inspection.

---

## 1. Introduction

Product quality control is one of the most important processes in a production line, with final products put through several inspections in order to satisfy client expectations. This is especially critical in the automotive industry, where the finish of the car is of the utmost importance, which is why most manufacturers have a dedicated line of quality control in each of its assembly phases: raw plate, painting and final finishing.

Although the majority of production lines today are fully automated, the quality control process is still carried out using manual detection and subjective evaluation by experts, known as *check-men* (shown in Fig. 1). For instance, after a car body has been painted, these check-men inspect the car body based on how the light projected on the body is deflected when a defect is present. In addition, in order to detect defects of very small size, these experts pass their hands over the car body, using special gloves which help them in the detection process.

There are some issues regarding the current manual inspection process: first, the difference in evaluation criteria applied by *check-men* across the board, and second, the difference in the evaluation criteria applied on an individual basis by *check-men* on any one day. These issues make it impossible to guarantee the same minimum quality standard for all vehicles given that the process is affected by human errors. A third issue arising from this is the impossibility of performing a “big-data” analysis that would allow the detection of errors in these previous stages, since any reports compiled are affected by human errors and, moreover, only the most important of these are reported and saved to databases, resulting in the loss of valuable information.

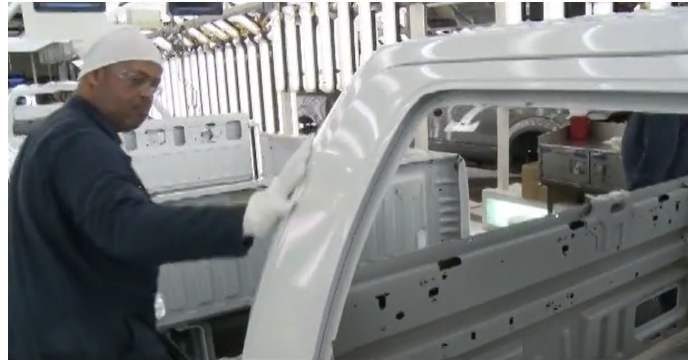


Figure 1: Check-men working on a production line, inspecting car body parts following the painting stage (image courtesy of FORD Dearborn, USA).

As a consequence of limitations in human detection, product quality managers in the automotive industry are still looking for a systematic solution. Although the ideal scenario would be to detect defects as soon as possible and therefore save money and time and reduce the total cost of the final product, it is in the primary stages (raw plate) where small defects (around  $6\mu\text{m}$  of diameter) are more difficult to detect. Note that when a defect of such small size has passed through the painting process, it can grow in size to the order of  $0.2\text{mm}$  in diameter or greater. As a consequence of painting, these defects become more visible if the surface is specular, currently the most common finish in vehicles today and therefore the focus of our research. This effect is not present in matt surfaces.

Let us turn our attention to existing research. Early work in the field of estimating surface deviations using lighting reflection techniques was carried out by [1] using the basic deflectometry principle on metal surfaces. The authors describe a system based on a stripe projection method using a high-resolution line scan camera for the recognition of dents and waviness of only  $10\mu\text{m}$  depth in less than  $10\text{s}$  per  $\text{m}^2$ . The detection is performed in the frequency domain by detecting the differences between an ideal surface and a surface with defects. This approach works on flat surfaces given that an ideal background can be computed a priori but not on surfaces with concavities or where there are positional variations, in which cases it is difficult to compute an ideal background.

A few years later, [2] differentiates between structured and smooth surfaces. In this work, an image fusion technique based on data fusion techniques [3] is presented, expressing the image fusion as an energy term  $E_k(\mathcal{D}, \mathcal{N}, \mathcal{F})$ , which involves raw data  $\mathcal{D}$ , the nuisance parameter  $\mathcal{N}$  [4] and the fusion result  $\mathcal{F}$ . The energy terms are then combined to an energy function  $E$  by means of weighting summation  $E = \sum_k \lambda_k E_k(\mathcal{D}, \mathcal{N}, \mathcal{F})$ , with  $\lambda_k > 0$ . The authors claimed that the advantages of this approach are its generality such as the possibility to add additional information and constraints by simply adding further energy terms. The approach is then tested on several surfaces, one of them the door of a painted car body, where the image fusion is performed by the minimization of the energy function  $E$ . Again,  $E$  can not deal with sudden surface slope changes, such as style lines, corners or edges, so the detection was performed on flat surfaces or those with smooth slope changes.

Despite different laboratory setups, a commercial and very successful system based on the ideas in [2] was developed and installed in Ford factories all around the world. This system, described in [5, 6], uses a moving structure made up of several light bars (fluorescent tubes at high-frequency) and a set of cameras in fixed positions around the stationary car body, and is able to detect defects of up to  $0.2mm$  in diameter, speeding up and improving the quality of the manual inspection carried out until that point. Unfortunately, as with previous research, only flat surfaces and those with smooth slope changes are inspected, meaning again that all style lines, edges and corners are excluded.

In [7], a very similar system installed in the Opel factory in Spain is presented, the main difference of which is that the inspection is carried out whilst the car body is moving. Again, information about the algorithm detection developed, as well as its performance, is limited.

Other commercial systems dealing with defect detection on painted car bodies use robots to cover the entire area, such as in [8]. This approach tries to cover as much area as possible by bringing the vision and illumination systems close to the body surface, but the details of the detection algorithm are not clear or well explained. In this approach, the area that is not expected can be reduced and the system could cover a much larger area compared with [5, 6]. The main problem of this approach is with

regard to cycle time production requirements, since it takes a long time to cover the entire surface of the car body and, as far as we are aware, it is still currently an industrial prototype.

In addition, some patents such as [9, 10, 11, 12, 13] can be found, proposing systems for the inspection and detection of defects on specular surfaces, none of which clarify or indicate how to overcome the problems of detecting defects in concavities, style lines, edges and corners, etc.

Therefore, this paper outlines a novel automated vision inspection approach based on deflectometry and image fusion techniques which is able to detect defects not only in flat areas and those with smooth slope changes, but in car body style lines (sudden surface slope changes), concavities, edges and corners, etc. This means that the approach is able to detect defects in areas that are very difficult to inspect, such as handles for example, assuming of course that the area is well illuminated. Thus, the paper describes a two-step algorithm: a new *pre-processing* step (image fusion algorithm), which is more robust when faced with environmental illumination pollution (or diffuse light) than, for example, and as is proved in this paper, the one used in [5, 6]; and subsequently, we present a new *post-processing* step to extract the image background using a local directional blurring in order to make it possible to detect defects on style lines, edges and corners.

In addition, a multi-level detection approach is presented enabling the detection of larger defects. The input at each additional level is the re-sized image background computed at a higher level. Since one defect can be detected at more than one level, a well-known clustering approach based on distance is used.

Experimental results are obtained from the *QEyeTunnel* industrial automatic control quality system employed in the production line at the Mercedes-Benz factory in Vitoria, Spain. The paper describes briefly both the hardware and software architectures of this system, focusing on the industrial vision system in which the algorithm has been implemented. We provide the computational cost of our algorithm when it is implemented in a dedicated vision system in both *CPU* and *GPU* platforms. We also provide examples of the detection performance and analyze its robustness and reliability by carrying out several tests. In addition, a comparison of the detection using our

approach and the manual inspection carried out by check-men is reported and commented upon.

The paper is organized as follows: in Sec. 2, the preliminaries and objectives of the paper are outlined. Then, in Sec.3 the proposed approach is explained in detail. Afterwards, in Sec. 4, the experimental industrial setup is explained, together with all tests and results, complete with comments. Sec. 5 discusses some details of our proposal where improvement is possible. Finally, conclusions are provided in Sec. 6.

## **2. Preliminaries and Objectives**

*a) Deflectometry-based detection on specular surfaces:* this technique has proven to be a reliable and accurate approach to accomplishing the task of detecting defects on car body surfaces [14]. In fact, this is what check-men do nowadays in manual inspections using their own eyes, analyzing the reflections on the car body reflected in the car body and looking for deformations. In the same way, as in the case with projection techniques, deflectometry is based on the projection of structured light patterns over a surface. When a triangulation method like fringe projection is used, the camera is focused on the surface onto which a light pattern is projected, as shown in Fig. 2. Thus, when a deformation (i.e. ding or dent) appears, the light rays are deviated, producing a sudden change in the shape of the pattern. Throughout the next sections, we will show how we can take advantage of these pattern changes in order to detect very small defects.

*b) Light pattern reflection issues:* since our approach is based on how the light pattern is reflected onto a surface, some issues regarding such reflection need to be considered. Fig. 3 shows a well-known surface reflection model proposed by [15]. In this model, the reflected light has three components: a specular spike, specular lobe, and diffuse lobe, where the diffuse lobe is the result of the internal reflection, while the specular spike and specular lobe are the result of the surface reflection. If the surface is perfectly specular, no diffuse lobe is observed, yet as the surface property changes from specular to non-specular, the specular spike rapidly diminishes and the diffuse lobe becomes stronger. Since the information

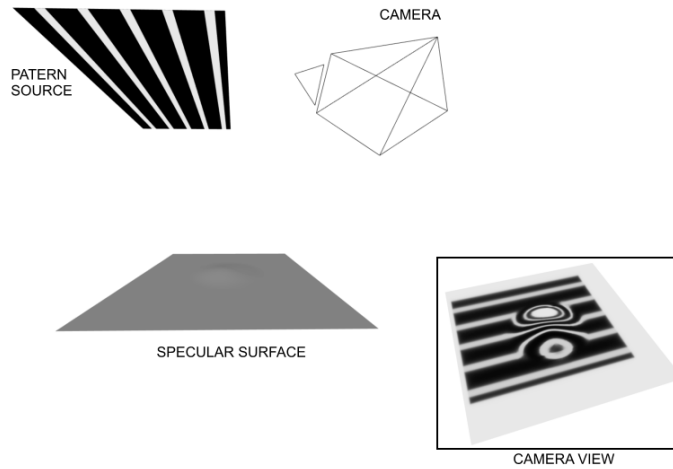


Figure 2: Deformation of the projected pattern due to surface defects.

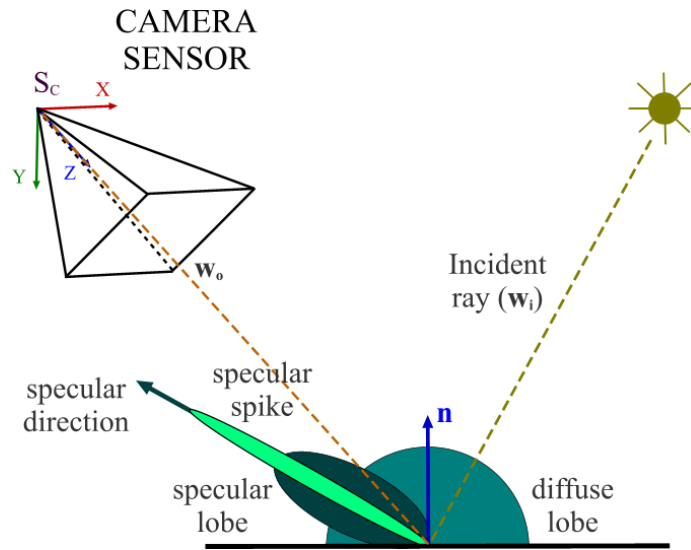


Figure 3: Reflectance model [15]. When light is incident to a surface, the light is reflected as a diffuse lobe, specular lobe and a specular spike, their shape and size dependent on the surface properties.

regarding the surface structure is held mainly by the specular spike, it is very important to minimize the effect of the diffuse lobe. Moreover, the effect of the diffuse lobe is not always due to the surface's direct reflection properties, but can



also be magnified due to environmental light pollution.

- c) *Image fusion technique*: the goal of image fusion is to integrate complementary multi-sensor, multi-temporal and/or multi-view information into one new image containing information of a quality which cannot be achieved otherwise [16]. In this paper, the term *image fusion* refers to the process in which several images acquired by one single camera at different instants of time are merged forming a new image, enhancing information that is of interest in the purposes of defect detection.
- d) *Problem Definition*: the goal of this work is to propose a new image detection algorithm based on deflectometry techniques for detecting small defects on specular surfaces in general, and car body surfaces in particular. We aim to:
- provide an algorithm that is robust in the face of varying surface properties and environmental light pollution
  - offer the possibility of defect detection carried out not only on flat surfaces or those with smooth slope changes, but also on sudden slope changes such as style lines, provided of course the area is well illuminated
  - meet the cycle time production requirements of car body production line manufacturers (less than 30s)

### 3. Proposed Approach

Let us define  $\mathbf{a} \in \mathbb{R}^{\mathbf{X}}$  as the resulting fusion image in which  $\mathbf{X}$  represents the  $n$ -dimensional Euclidean space  $\mathbb{R}^l$  with  $l = 1, 2, \text{ or } 3$  together with the discrete topology, although other topologies such as *von Neuman* or the *odd-even product* are also commonly used in computer vision [17]. So, in this paper  $\mathbf{X} = \mathbb{Z}_m \times \mathbb{Z}_n = \{\mathbf{x} := (x, y) \in \mathbb{Z}^2 : 0 \leq x \leq m - 1, 0 \leq y \leq n - 1\}$ ,  $m$  and  $n$  being the number of columns and rows of the image. Let  $\mathbf{i}(k) := \mathbf{i}_k \in \mathbb{R}^{\mathbf{X}}$  be the acquired image at  $k = 0, 1, 2, \dots, M - 1$  instant of time, with  $M$  being the total number of images acquired.

Therefore, our approach can be divided into two main steps based on the execution of the same:

- *Pre-processing step*, where the combination of  $M$  images acquired during the scanning stage is performed, obtaining as a result the fusion image:

$$\mathbf{a} = f(\mathbf{i}_k, p_k(\boldsymbol{\omega}_i)) \text{ with } k = 0, \dots, M - 1 \quad (1)$$

where  $\mathcal{P} = \{p_k(\boldsymbol{\omega}_i), k = 0, \dots, M - 1, i = 1, \dots, \infty\}$  is the set of patterns reflected on the surface,  $\boldsymbol{\omega}_i$  is each of the infinite incident rays over the surface, and  $p_k$  represents the intensity of the pattern at the location indicated by the incident ray  $\boldsymbol{\omega}_i$  (see Fig. 3).

- *Post-processing step*, where from the resulting fusion image  $\mathbf{a}$ , a discrimination of the part of the image that is background (low-frequency signal) and the part that is defect (high-frequency signal) is performed by a treatment based on the local trend of each pixel followed by a threshold that makes it possible to distinguish between defect (high-frequency signal and high magnitude) and noise (signal of very high frequency and small magnitude). This treatment allows us to detect very small defects in concave areas or sudden changes in the car body surface such as handles, corners or style lines. As far as the authors are aware, this is the first work that proposes a solution which is able to detect very small defects in such areas and that fulfills cycle time production line requirements.

It should be remarked that the *pre-processing step* must be carried out in real-time while the light pattern is sweeping over the body surface, ensuring that no frame is missed, which would produce a poor result. The *post-processing*, meanwhile, starts once the fusion image is ready, allowing us to perform other tasks simultaneously as long as the cycle time requirements are met.

In the following, we detail all the steps of the solution proposed in this paper, putting special emphasis on those aspects which differentiate it from other solutions, using comparative examples where necessary.

### 3.1. *Pre-processing step*

Our image fusion proposal is based on two premises: first, to lose the minimum amount of information possible in each of the mergers; second, to try to minimize the

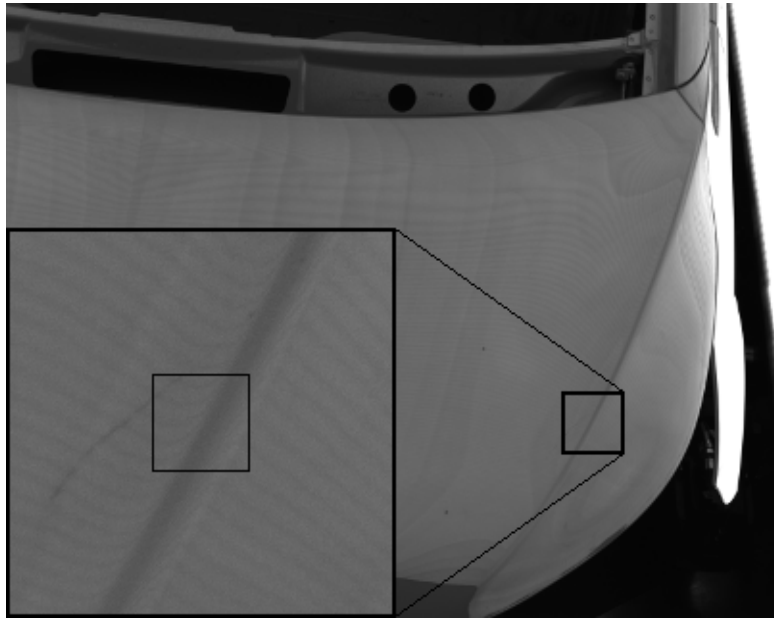
effect of diffuse and ambient parts of the light, since, as we have seen in section 2, it is the specular spike which has the relevant information of the body surface structure. Hence, we propose the following:

$$\mathbf{a} = \sum_{k=1}^{M-1} (\mathbf{i}_k - \mathbf{i}_{k-1})^2 \quad (2)$$

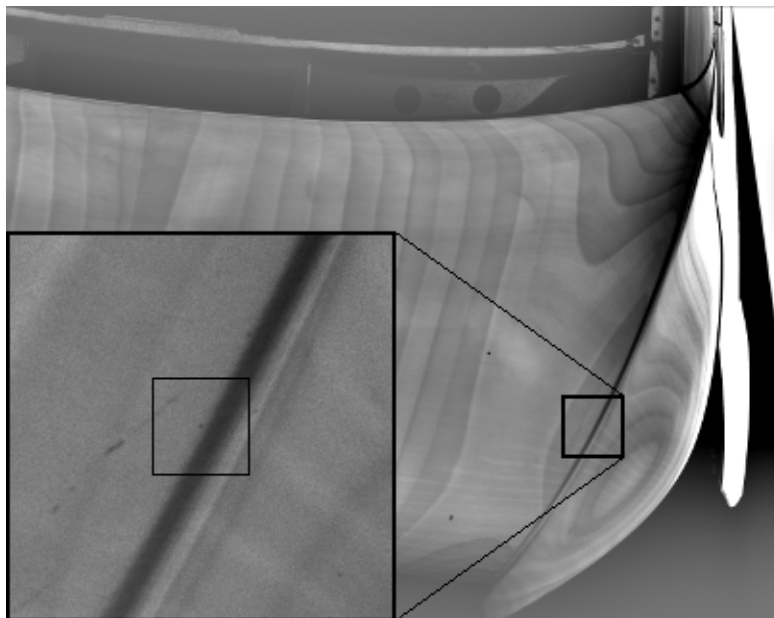
First, note how our proposal uses the difference between two consecutive frames ( $\mathbf{i}_k - \mathbf{i}_{k-1}$ ), thereby reducing the effect of the diffuse light counterpart. In Appendix A we empirically demonstrate this statement, comparing the result obtained with our method with the result described in [5, 6]. On the other hand, with the square of this difference, we try to maximize the effect of the contrast between low and high levels of pixel intensity. Note that a defect has a very low intensity value while a part of the surface without defect has a very high intensity value, so that by enhancing the proposed quadratic before the fusion, it is possible to enhance the effect of the contrast between a defective zone and a non-defective one. Finally, and in order to lose the minimum amount of possible information, the arithmetic sum of the result obtained by the  $M$  quadratic differences is calculated. Thus, as we can see, we move from working on a grayscale value, ranging from 0 to 255, to a half-level precision (float 16bits), ranging from -65504 to 65504, losing the least amount of information as is computationally allowed.

It must be pointed out that, even though we will continue here to refer to images, we are actually working with high-precision matrices. That is why all image shown in this paper are normalized in order to visualize the result of our approach. This normalization will be indicated with  $Trunc(\bullet)$  from here on.

Fig. 4 shows the difference between the fusion image resulting from the proposed fusion strategy Fig. 4(b) and that obtained by one of the most successful systems in the automotive industry [6] to date Fig. 4(a). It is remarkable that the contrast has been enhanced using our approach, making it clearer that a defect is present. Whilst our fusion strategy increases the detection capability, it also increases the contrast of areas which are not defects, such as corners or style lines. The following step of our algorithm deals with this problem, allowing the detection to be carried out in the entire



(a)



(b)

Figure 4: Example of *pre-processing* step result on hood: Fig. 4(a) shows the fusion result of the image fusion strategy developed in [6]; Fig. 4(b) shows the normalized fusion approach,  $Trunc(\mathbf{a}(\mathbf{x}))$ .

area.

### 3.2. Post-processing step

Unlike the majority of methods of defect detection that we can find in the literature [18, 19], which are based on obtaining a background of the analyzed surface offline and then comparing it with the result of each scan, we propose a method that obtains the corresponding image background and meets the cycle time required by the production line. As far as we are aware, only [5] and [6] perform some type of image background extraction in an attempt to homogenize the intensity levels of the fusion image so that, by means of a local thresholding, the detection of the defects is performed. The major drawback of this form of image background extraction is the need to hide parts of the surface such as style lines, corners, edges and concavities (e.g. handles). If such parts are not hidden, they would be detected as defects. Note that in order to exclude parts of a complex surface such as a door or the rear of a car body, a perfect calibration of the vision system must be performed in addition to a  $2D$  or  $3D$  matching of the surface seen by each camera. This usually requires the use of well-known patterns that are not always found in sufficient numbers in order to achieve a correct match. In addition, such techniques significantly increase the computation time of the algorithms, [20].

Therefore, we propose a local directional-based blurring method for obtaining the background corresponding to  $\mathbf{a}$ . This is based on the use of the image gradient [21], which is affected by the image noise. So, in order to reduce such noise, first a linear filter is applied. Let us denote  $\mathbf{b} \in \mathbb{R}^{\mathbf{X}}$  as the resulting smoothed image:

$$\mathbf{b} := \mathbf{a} \circledast_{\Gamma} N \quad (3)$$

where  $\Gamma : \mathbb{R}^{\mathbf{X}}|_N \rightarrow \mathbb{R}$  denotes the averaging function  $\Gamma(\mathbf{a}|_{N(\mathbf{y})}) = \frac{1}{\text{card}(N(\mathbf{y}))} \sum_{\mathbf{x} \in N(\mathbf{y})} \mathbf{a}$ , and  $N(\mathbf{y}) \subset \mathbb{Z}^2$  a predefined neighborhood of  $\mathbf{y} \in \mathbb{Z}^2$ .

Let us denote  $\mathbf{d} \in \mathbb{R}^{\mathbf{X}}$  as the gradient direction image computed from  $\mathbf{b}$  given by:

$$\mathbf{d} := \arctan2((\mathbf{b} \oplus \mathbf{s})|_{\mathbf{m} > 0}, (\mathbf{b} \oplus \mathbf{t})|_{\mathbf{m} > 0}) \quad (4)$$

where  $\mathbf{a1} \oplus \mathbf{a2} = \{(\mathbf{y}, \mathbf{a3}(\mathbf{y})) : \mathbf{y} \in \mathbf{Y} \subset \mathbf{X}, \mathbf{a3} := \sum_{\mathbf{x} \in \mathbf{X} \cap S(\mathbf{a2}_\mathbf{y})} \mathbf{a1} \cdot \mathbf{a2}_\mathbf{y}\}$  is the right linear convolution product,  $S(\mathbf{t}_\mathbf{y})$  is the support of  $\mathbf{t}$  defined by  $S(\mathbf{t}_\mathbf{y}) = \{\mathbf{x} \in \mathbf{X} :$

$\mathbf{t}_y \neq 0\}$  with  $\mathbf{t}_y = \mathbf{t}(\mathbf{y})$ , and  $\mathbf{m}$  is the gradient magnitude image taking the form:

$$\mathbf{m} := \sqrt{(\mathbf{b} \oplus \mathbf{s})^2 + (\mathbf{b} \oplus \mathbf{t})^2} \quad (5)$$

In this paper, templates  $\mathbf{t}$  and  $\mathbf{s}$  are those taken from the well-known *Sobel edge detector*, [21]:

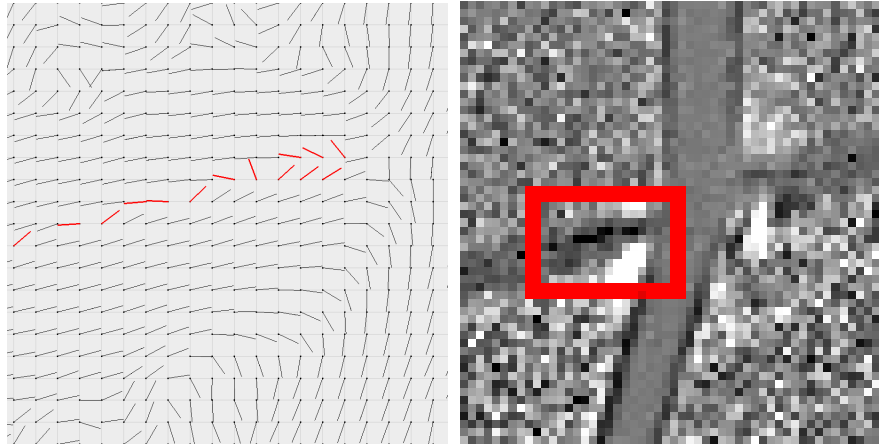
$$\mathbf{s}_{(i,j)}(\mathbf{x}) = \begin{cases} -1 & \text{if } \mathbf{x} = (i-1, j-1) \text{ or } (i-1, j+1) \\ 1 & \text{if } \mathbf{x} = (i+1, j-1) \text{ or } (i+1, j+1) \\ -2 & \text{if } \mathbf{x} = (i-1, j) \\ 2 & \text{if } \mathbf{x} = (i+1, j) \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

$$\mathbf{t}_{(i,j)}(\mathbf{x}) = \begin{cases} -1 & \text{if } \mathbf{x} = (i-1, j-1) \text{ or } (i+1, j-1) \\ 1 & \text{if } \mathbf{x} = (i-1, j+1) \text{ or } (i+1, j+1) \\ -2 & \text{if } \mathbf{x} = (i, j-1) \\ 2 & \text{if } \mathbf{x} = (i, j+1) \\ 0 & \text{otherwise} \end{cases}$$

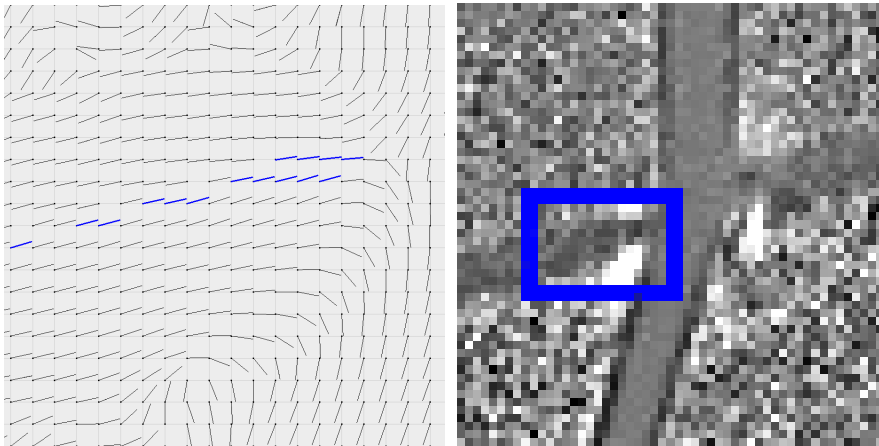
with  $(i, j)$  denoting the evaluated pixel of  $\mathbf{b} \in \mathbb{R}^{\mathbf{X}}$ .

Note that the basis of the proposed image background extraction method is based on obtaining the correct directions (trend) of each pixel. Fig. 5 shows an example of a style line in which its pixel directions are not computed correctly (see Fig. 5(a) and Fig. 5(b)) due to the problem that the gradient algorithm has in computing where image noise and singular points are present. In order to avoid this and to get correct directions, we propose smoothing them in accordance with the trend of their neighborhood. The result of this can be seen in Fig. 5(c) and Fig. 5(d), in which all directions follow the trend of the style line.

Once the gradient direction image  $\mathbf{d}$  is computed correctly, the image background is obtained as follows. Let  $\mathbf{h}$  be a subset of  $\mathbf{d}$  with radius  $R$ . For the sake of formulation simplicity, let us denote  $c = (i, j)$  as the reference pixel and  $p = (-R \leq i \leq R, -R \leq j \leq R)$  as the evaluated pixel. Let  $\mathbf{h}_c$  be the center of the subset, and



(a) Computed directions using the fusion image **a**. (b) Computed image background using Fig. 5(a).



(c) Computed directions using the smoothed fusion image **b**. (d) Computed image background using Fig. 5(c).

Figure 5: Directional blurring approach for image background computation: case of style lines. The image in Fig. 5(a) represents the pixel directions obtained corresponding to the area marked in red in Fig. 5(b). It can be observed that, in the central area of the style line, due to image noise, the computed pixel directions differ from those of the rest of the style line. When applying the proposed directional blurring approach using such directions, defective areas appear (pixels with low intensity) along the style line, as shown in Fig. 5(c). However, when smoothing the fusion image before computing the image gradient, the directions align perfectly with those of the rest of the style line, as shown in Fig. 5(c), thus obtaining a correct image background when applying our directional blurring approach, as shown in Fig. 5(d).

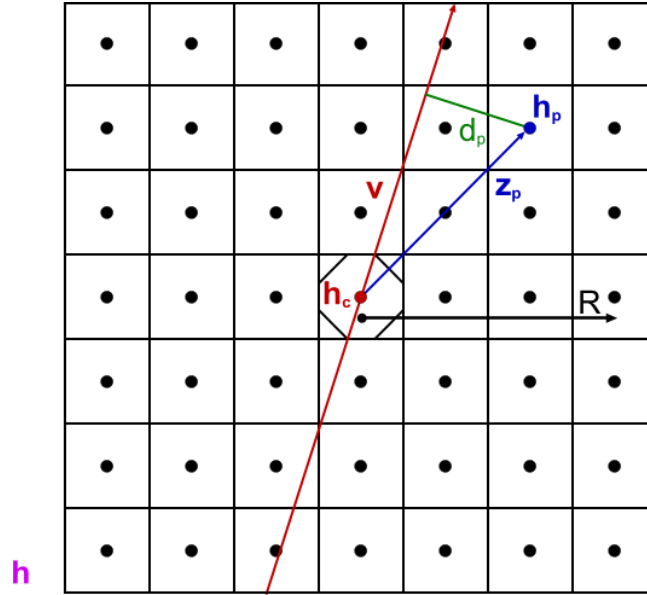


Figure 6: Proposed directional blurring method.

$\mathbf{v} := \mathbf{d}_c$  its vector director. For every pixel  $\mathbf{h}_p$  of the subset, we compute the vector  $\mathbf{z}_p = p - c$ , taking the center cell of the subset as the reference cell (see Fig. 6). Hence, its projection over the direction given by  $\mathbf{h}_c$  is:

$$d_p = |\mathbf{v} \times \mathbf{z}_p| \quad (7)$$

With this, we define a weighting value for each pixel of the subset in the form:

$$w_p = (0 \vee (R - d_p))^3 \quad (8)$$

where  $\vee$  represents the maximum operator between two values. Note that with this weighting, pixels of the subset whose directions are close to the direction of  $\mathbf{h}_c$  (small  $d_p$ ) have greater weight, while pixels whose directions are not related to the direction of  $\mathbf{h}_c$  (large  $d_p$ ) have less weight. Note that for pixels where  $(R - d_p) < 0$ ,  $w_p = 0$ , there is no blurring. Thus, the background image is given by:

$$\mathbf{f} = \frac{\sum (w_p \cdot \mathbf{a}_p)}{\sum w_p} \quad (9)$$



Hence, once the background image is obtained, the next step is to differentiate defects from noise. To do this, a *modified local area contrast enhancement* is proposed. The well-known local area contrast enhancement [21] is based on the computation of the local standard deviation as follows:

$$\sigma := \beta \vee \sqrt{(\mathbf{a} - \boldsymbol{\mu})^2 @ N} \quad (10)$$

with  $\boldsymbol{\mu}$  being the local mean of  $\mathbf{a}$  defined by  $\boldsymbol{\mu} := \mathbf{a} @ N$  and where the image average function is  $\alpha \in \mathbb{R}^{\mathbf{Y}}|_N \rightarrow \mathbb{R}$  of  $\mathbf{g} @ N$ , with  $\mathbf{y} \in \mathbf{Y} \subset \mathbf{X}$ , meaning that the right reduction of  $\mathbf{g} \in \mathbb{R}^{\mathbf{Y}}$ , with  $N(\mathbf{y}) \subset \mathbb{Z}^2$ , yields the neighborhood averaging operation  $(\mathbf{g} @ N)(\mathbf{y}) := \alpha(\mathbf{g}|_{N(\mathbf{y})})$ . Note that  $\beta > 0$  is a lower bound applied to  $\sigma$  in order to avoid problems with division by zero in the next step of the proposed technique. So, the enhancement image  $\mathbf{e} \in \mathbb{R}^{\mathbf{X}}$  takes the form:

$$\mathbf{e} = \frac{1}{\sigma}(\mathbf{a} - \boldsymbol{\mu}) \quad (11)$$

In contrast to this, we propose the following *modified local area contrast enhancement* based on the background obtained in which the weighting of the high-frequency component is a function of the local standard deviation computed only by using the positive values of the image  $\mathbf{a} - \mathbf{f}$ . This means that:

$$\sigma^+ := \beta \vee \sqrt{(0 \vee (\mathbf{a} - \mathbf{f}))^2 @ N} \quad (12)$$

The advantage of our approach is that using  $\mathbf{f}$  instead of  $\boldsymbol{\mu}$  in computing the local high-pass filter of  $\mathbf{a}$  in the spatial domain, is that we remove more efficiently all low frequencies contained in  $\mathbf{a}$  since  $\mathbf{f}$  includes such low signal information. Specifically, in our case,  $\mathbf{f}$  yields all the information regarding the shape of the surface, such as style lines, corners, etc.

It is worth noting that  $\sqrt{(0 \vee (\mathbf{a} - \mathbf{f}))^2 @ N}$  in Eq. 12 actually represents the local standard deviation of the positive part of the signal  $\mathbf{a} - \mathbf{f}$ . The reason for using  $\sigma^+$  instead of  $\sigma$  deals with the effect of both approaches on defects. Fig. 7 shows an example of this behavior when either a small defect (in Fig. 7(a) and Fig. 7(b)) or a large defect (in Fig. 7(c) and Fig. 7(d)) appears. When the defect is small, the result obtained using  $\sigma$  y  $\sigma^+$  is the same except for a signal scaling issue. But when the

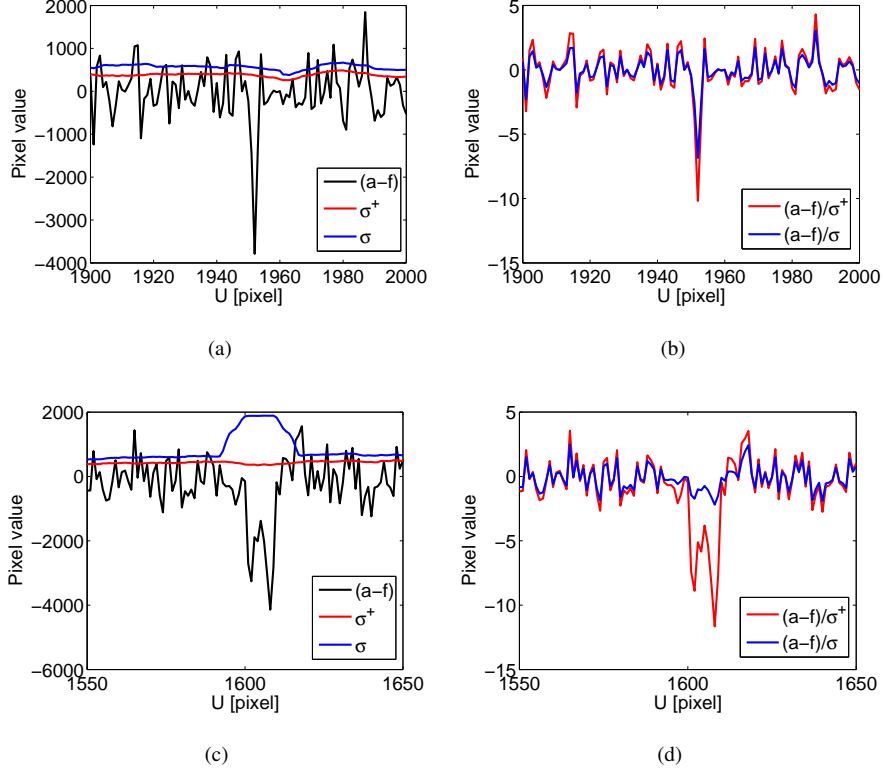


Figure 7: Comparison between the results obtained by using  $\sigma^+$  instead of  $\sigma$ : in Fig. 7(a), the classic and modified standard deviation are compared when a small defect is present; in Fig. 7(b), the comparison of  $\frac{1}{\sigma}(\mathbf{a} - \mathbf{f})$  and  $\frac{1}{\sigma^+}(\mathbf{a} - \mathbf{f})$  indicates when a small defect is present; in Fig. 7(c), the classic and modified standard deviation are compared when a large defect is present; in Fig. 7(d), the comparison of  $\frac{1}{\sigma}(\mathbf{a} - \mathbf{f})$  and  $\frac{1}{\sigma^+}(\mathbf{a} - \mathbf{f})$  indicates when a large defect is present.

defect is larger in size, using  $\sigma$  the defect is lost, whereas by using  $\sigma^+$  the defect is detected.

Hence, we propose the following enhanced image:

$$\mathbf{e} = \frac{1}{\sigma^+}(\mathbf{a} - \mathbf{f}) - th \quad (13)$$

where  $th$  is a constant threshold value.

As in [21], the local gain factor  $\frac{1}{\sigma^+}$  of Eq. 13 is inversely proportional to the local standard deviation. Thus, a larger gain is applied to regions with low contrast.

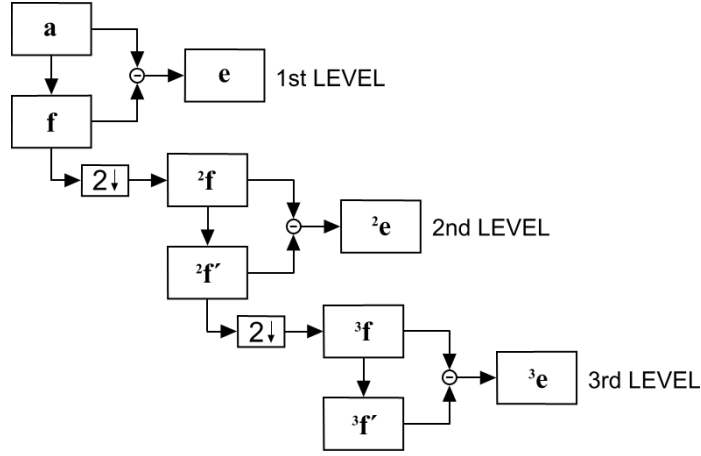


Figure 8: Proposed multi-level detection method: the computed background is half-downsampled and used as input for each detection level.

Note that since  $\mathbf{f}$  is computed using a kernel of constant and determined size, defects with an area larger than such a kernel are lost. However, this kernel cannot be too large since some shapes of the body would be not included in  $\mathbf{f}$ , making the detection in those parts of the surface impossible.

In order to solve this problem, we propose a multi-level detection method as shown in the three-level example in Fig. 8. The first level consists in extracting the image background  $\mathbf{f}$  from the fusion image  $\mathbf{a}$ , as has been explained above. Afterwards, Eqs. 3 to 13 are applied as mentioned previously, obtaining the enhanced image  $\mathbf{e}$ . For the second level,  $\mathbf{f}$  is rescaled in the form  ${}^2\mathbf{f} := \Lambda\mathbf{f}$ , where  $\Lambda\mathbf{f}$  denotes the rescaling of  $\mathbf{f}$  by a generic rescaling operation  $\Lambda : \mathbb{R}^{\mathbf{X}} \rightarrow \mathbb{R}^{\mathbf{Z}}$ , where  $\mathbf{Z} = \mathbb{Z}_{m/q} \times \mathbb{Z}_{n/q} = \{\mathbf{x} := (x, y) \in \mathbb{Z}^2 : 0 \leq x \leq m/q - 1, 0 \leq y \leq n/q - 1\}$ , with  $m$  and  $n$  be the number of columns and rows of the image and  $q > 1$  the scaling factor. In this paper, we propose  ${}^2\mathbf{f}$  to be the half-sized resolution of  $\mathbf{f}$ . Then, we apply Eqs. 3 to 13, thus obtaining the enhanced second-level image  ${}^2\mathbf{e}$ . For the third level, we repeat the same procedure described above for the second level, using as input the background extracted from the second level,  ${}^2\mathbf{f}'$ , half-sizing it using  ${}^3\mathbf{f} := \Lambda^2\mathbf{f}'$  and then applying Eqs. 3 to 13 in order to obtain the enhanced third-level image  ${}^3\mathbf{e}$ . If more levels are allowed, the procedure will be the same.

The last step of the algorithm consists in performing a *global thresholding* of the enhanced images  $\mathbf{e}$ ,  ${}^2\mathbf{e}$  and  ${}^3\mathbf{e}$ , followed by an *up-scaling* operation of  ${}^2\mathbf{e}$  and  ${}^3\mathbf{e}$  to meet the size of  $\mathbf{e}$  and thus provide the result of the detection. Hence, let  $\mathbf{p} \in \{0, 1\}^{\mathbf{X}}$ ,  ${}^2\mathbf{p} \in \{0, 1\}^{\mathbf{Y}}$  and  ${}^3\mathbf{p} \in \{0, 1\}^{\mathbf{Y}}$  be the images obtained following the thresholding of  $\mathbf{e}$ ,  ${}^2\mathbf{e}$  and  ${}^3\mathbf{e}$  respectively:

$$\begin{aligned}\mathbf{p} &= \chi(\mathbf{e}) \\ {}^2\mathbf{p} &= \chi({}^2\mathbf{e}) \\ {}^3\mathbf{p} &= \chi({}^3\mathbf{e})\end{aligned}\tag{14}$$

where  $\chi = \begin{cases} 1 & \text{if } \mathbf{e} < 0 \\ 0 & \text{if } \mathbf{e} \geq 0 \end{cases}$ . The final image result  $\mathbf{r} \in \{0, 1\}^{\mathbf{X}}$  is given by:

$$\mathbf{r} := \mathbf{p} \parallel \nu({}^2\mathbf{p}) \parallel \nu({}^3\mathbf{p})\tag{15}$$

where  $\nu: \mathbb{R}^{\mathbf{Y}} \rightarrow \mathbb{R}^{\mathbf{X}}$  is the up-scaling operator with  $\mathbf{Y} \subset \mathbf{X}$ .

Fig. 9 shows an example of the proposed detection methodology. Fig. 9(a) shows how, at the first detection level, the algorithm has only been able to detect the smallest defects, leaving many others undetected. The larger defects are detected in levels 2 and 3 as shown in Figs. 9(b) and 9(c). The combination of the different levels of detection is shown in Fig. 9(d).

Finally, a masking layer is applied to all those areas seen by the camera that either do not belong to the surface of the body (e.g. mechanical or production line structures), or parts in which detection is not desirable (e.g., interior of the window frame). The method for generating such a mask, which is very specific to the body shape and system structure, is beyond the scope of this paper. Let  $\mathbf{u} \in \{0, 1\}^{\mathbf{X}}$  be the correspondent mask image. Thus, the result  $\mathbf{r}_n \in \{0, 1\}^{\mathbf{X}}$  is:

$$\mathbf{r}_n := \mathbf{r} \& \mathbf{u}\tag{16}$$

where the operator  $\&$  represents the binary “AND” image operation.

### 3.3. Case Study: Defects close to Style Lines

In this section, we will analyze in detail a typical case showing a defect very close to a style line. Let us take the fusion image  $\mathbf{a}$  as shown in Fig. 10(a). Let us focus our

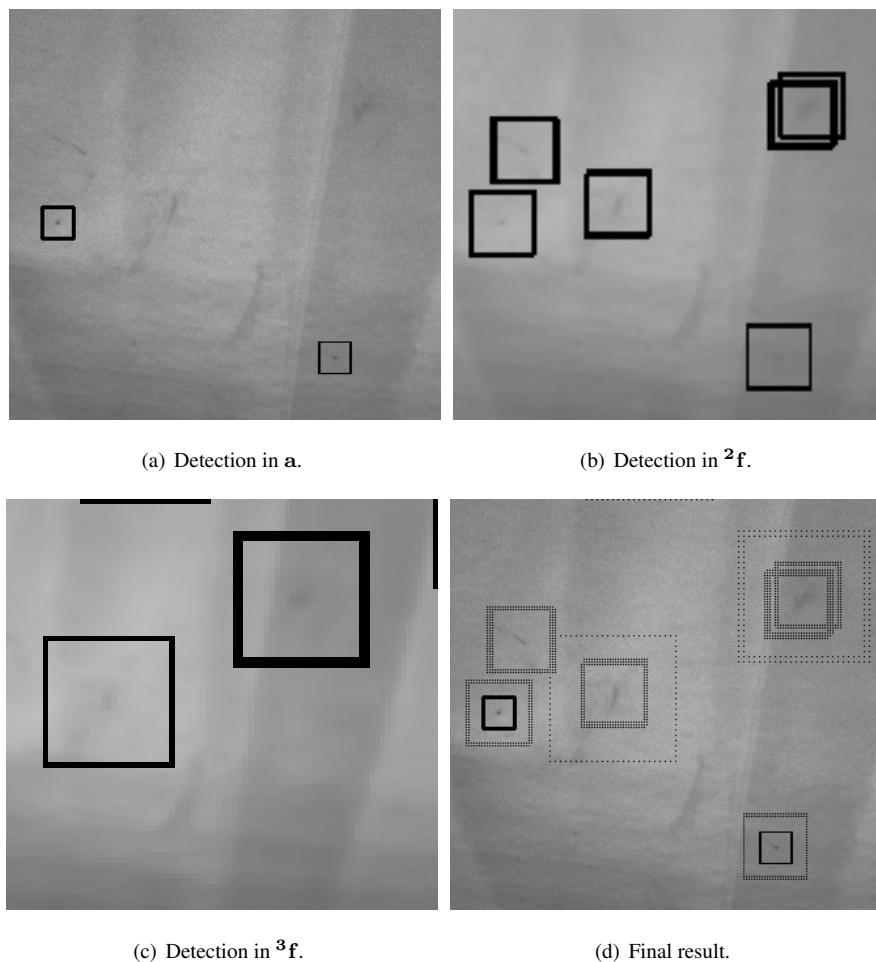


Figure 9: Three-level detection behavior: in Fig. 9(d), first level-detection is represented by a small solid-line box, second-level detection represented by a medium-sized dotted-line box and third-level detection by a large dotted-line box.

attention on the defect highlighted in the magnified area in the same figure. In Figs. 10(b) and 10(c), two different views of the 3D representation of this area are shown. Note in both figures how the defect lies on the edge of the style line.

As we have explained, the first stage of the proposed *post-processing step* deals with the image background extraction. Using the method detailed from Eq. 3 to 9, the background takes the form shown in Fig. 10(d). Note that both image noise and defect

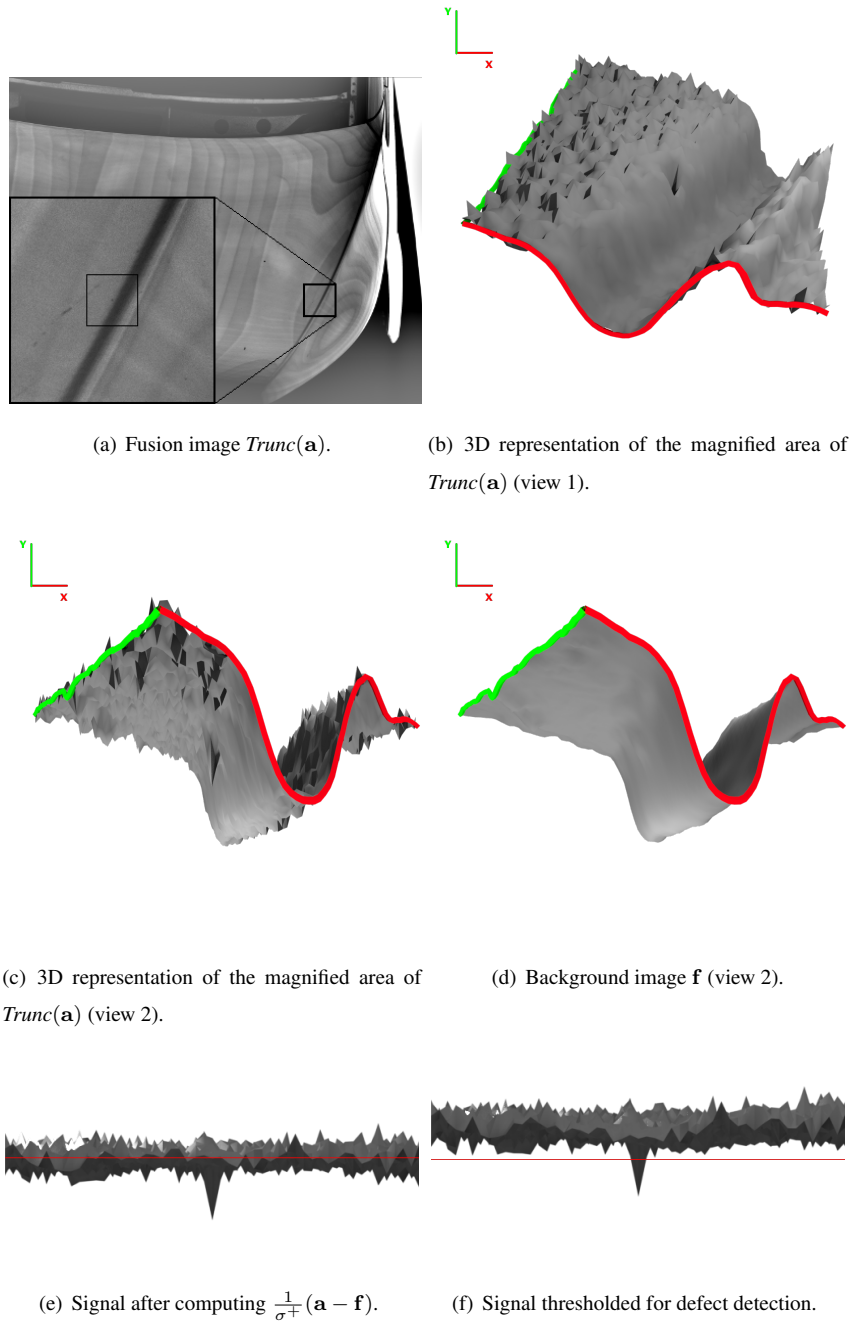


Figure 10: Case study: evaluation of the proposed algorithm dealing with a defect close to or on a style line.



Figure 11: QEyeTunnel: an automated inspection system for car body surface quality control based on artificial vision.

are not present, but the style line remains, which is one of the biggest advantages put forward in our approach.

When we use Eq. 12 and apply  $\frac{1}{\sigma^+}(\mathbf{a} - \mathbf{f})$ , the entire signal is set on zero level, with the defect clearly highlighted, provided that its size is larger than the signal noise as shown in Fig. 10(e).

Finally, using a global threshold,  $th$ , we can distinguish between noise and defect and thus provide the defects on the surface as a result. These last two parts are covered in Eq. 13 as one unique step.

#### 4. Experimental Evaluation and Results

As seen in Sec. 3, the defect detection algorithm proposed in this paper consists of two phases: a *pre-processing* step, in which a light scan of the body to be analyzed is required using light patterns (e.g., straight lines, chess board pattern, etc.), and a *post-processing* step, which uses as input the fusion image resulting from the scan performed in the previous phase. For the first phase, it is necessary to use some hardware that projects light patterns and performs this scanning process. Although a laboratory setup could have been used for validation, we have taken results from the *QEyeTunnel* industrial system recently developed in Mercedes-Benz' factory in Vitoria, Spain (Fig. 11). Following is a brief description of this system avoiding all industrial aspects and

those related to the integration of the system. We have also omitted other aspects such as camera positions, intrinsic parameters, light patterns, etc., which are not deemed important in this paper.

#### 4.1. *QEyeTunnel: an Automated Vision-based Quality Control Inspection System*

Broadly speaking, this inspection system consists of two parts: an external fixed structure where a determined number of cameras are optimally placed in order to see the entire surface of the body to be analyzed, as well as a moving internal structure similar to a scanning machine which houses curved screens known as 'sectors' which act as the light sources that project the illumination patterns over the body surface, as can be seen in Fig. 11. The number of cameras as well as the number and size of sectors depend on the car bodies to be analyzed by the system. In particular, the *QEyeTunnel* inspection system employed at Mercedes-Benz in Vitoria comprises 23 monochrome cameras of 5MP working at 15fps, and two sectors set 3 meters apart, with a resolution of  $192 \times 1344$  each and a total route displacement of 2m moving at 0.267m/s of speed. With this setup, the scanning stage takes around 11s from start to finish.

Fig. 12 shows the elements of the hardware architecture as well as the communication relationships in the *QEyeTunnel* system. The main features are:

- *System Low-level Controller (PLC)*: This controls the mechanical movement of the sectors and also monitors all production line stages, system security as well as visual and sound alarms. The PLC is connected using the *PROFISAFE* communication protocol, and to the *PC PROC* using the *TCP-IP* protocol, through the use of a *PROFISAFE-ETHERNET* converter. The characteristics of each body as well as its position during the inspection process is indicated by this *PLC* to the *PC PROC*. The *PLC* waits for the starting signal from the *PC PROC* before sending back the ending signal when the scanning stage has finished.
- *Vision Industrial System (PC PROC)*: This is the element that controls the program flow. It is an industrial vision system from *Matrox*<sup>®</sup> named *Supersight*



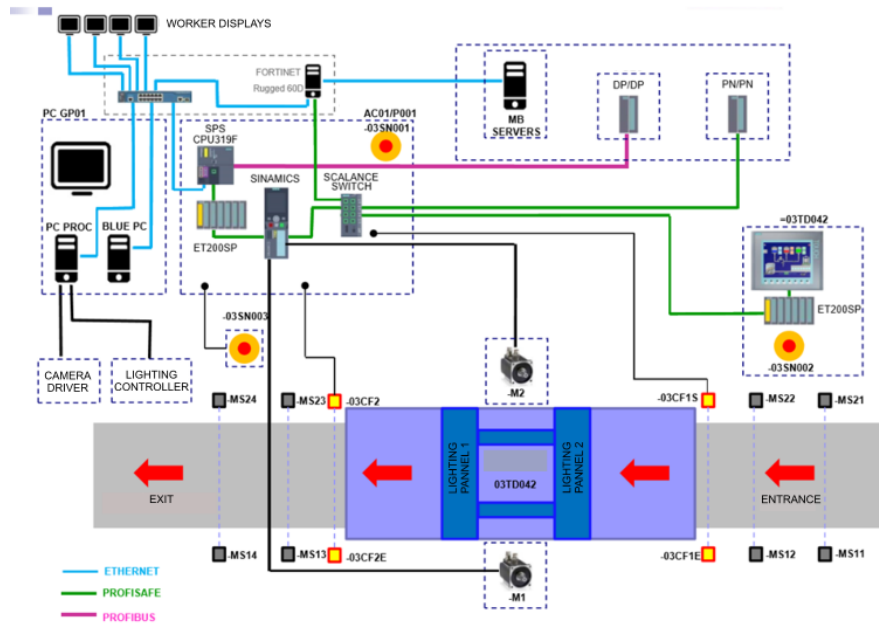


Figure 12: Hardware architecture and communication relationships of the *QEyeTunnel* system.

*Solo*, which is an entry-level configurable single-node high-performance computing (*HPC*) platform supporting two multi-core *Intel*<sup>®</sup> *Xeon* processors, third-party *GPU*s and *Matrox*<sup>®</sup> *FPGA* boards for demanding industrial imaging applications. It is equipped with 6 *GigE Card PCIe AdLink*<sup>®</sup> *GIE64+* with Power over Ethernet technology (*PoE*), and 4 ports for supporting 24 cameras. Moreover, it is equipped with 2 *Gigabyte*<sup>®</sup> *GeForce GTX 1080 8GB GDDR5X Dual Link DVI-D HDMI 3X DisplayPort PCI-E* graphic cards, which were used for the implementation of our proposal and also to run the program which generates the patterns to be projected. This system communicates with the *PLC* and the *BLUE PC* using the *TCP-IP* communication protocol.

- *System High-level Controller (BLUE PC)*: This is responsible for providing the inspection results to the production line workers, as well as acting as an interface between line workers and the *PC PROC* on maintenance issues. It is an industrial

*PC*-based system which communicates with the various *Worker Displays*, *MB Servers* and *PC PROC* systems using the *TPC-IP* communication protocol.

- *Worker displays*: Several screens controlled by small *PCs* are situated throughout the production line to display the results of the inspection. Workers use the information to locate the defects and act accordingly (fixing them when possible or marking them for later repair). Substituting such screens with light-weight augmented reality glasses is currently being considered, which would improve worker comfort and facilitate the identification and repair processes.
- *MB Servers*: The results of each inspection as well as the system backups are saved on these servers. Communication with the *BLUE PC* is through *TCP-IP* but using a *FORTINET* protocol. Apart from serving as a data security system, it is also used to perform big-data analysis in order to identify problems in the painting process, which will ultimately result in significant money savings and an increase in the quality of the final product.

Fig. 13 shows the inspection program flow as each car body enters the system. When a new body is ready to be introduced into the inspection system, the production line control systems inform the *PLC* accordingly and also provide all necessary body-related data (e.g. model, color, etc.). Only the necessary data is transmitted by the *PLC* to the *PC PROC* which, while the body is being positioned correctly for inspection, loads all initialization parameters, such as thresholds, masks, light pattern, etc. Once the car body is in the inspection position (signal transmitted by the *PLC* to the *PC PROC*), the *PC PROC* gives scanning permission to the *PLC* and activates the cameras. The 23 cameras then begin to send images at 15fps to the *PC PROC* and both sectors begin the scanning movement as instructed by the *PLC*. The *PC PROC* computes the previously mentioned *pre-processing* step and at the same time waits for the *PLC* to send the scanning ending signal. With the end of the scanning, the *PLC* warns the *PC PROC* and the latter stops the acquisition of images and proceeds to perform the *post-processing* step for all cameras. Meanwhile, the *PLC* allows the car body to leave the inspection system and, when a new body is ready, repeats the process. When the results

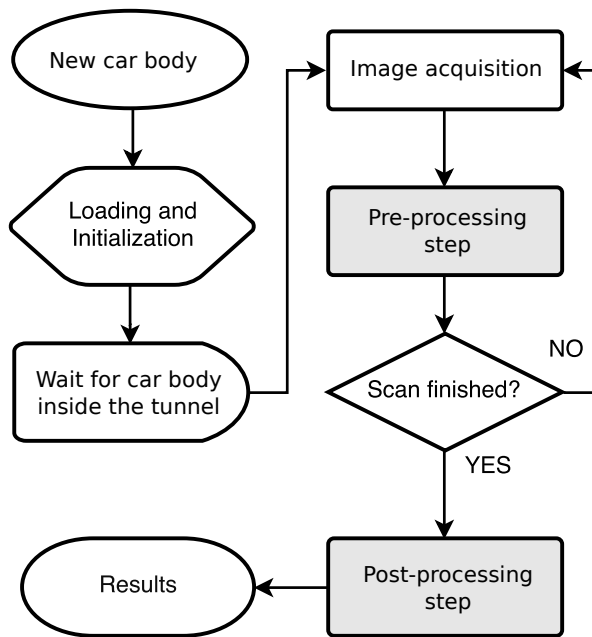
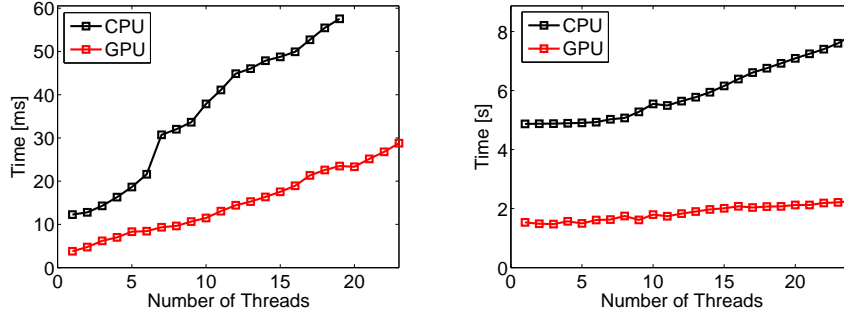


Figure 13: Inspection system program flow.

of the inspection are obtained, the *PC PROC* sends them to the *BLUE PC*, which in turn sends them to the different display monitors so that workers can locate identified defects on the car body and see to their repair.

#### 4.2. Algorithm performance

As mentioned in previous sections, this entire procedure, including the time required to place the car body in the inspection position and to return it to the production line, must not exceed the cycle time requirements of the production line. This time, for small to medium-sized car bodies, is usually 30s. Taking into account that the scanning process takes about 11s and that the positioning of the vehicle and its return to the production line takes about 8s, this usually leaves approximately 10s in which to carry out the entire detection process.



(a) *Pre-processing* step: maximum time per frame. (b) *Post-processing* step: maximum time.

Figure 14: Computational cost comparison between the proposed algorithm implemented in *CPU*<sup>1</sup> and in *GPU*<sup>2</sup>. Performance given 10 iterations in each thread.

In light of the previously mentioned, Fig. 14 shows the results of the computational cost comparison of the algorithm proposed in this paper using the previously mentioned Supersight Solo industrial vision system and the 23 cameras used by Mercedes-Benz’ *QEyeTunnel* inspection system in Vitoria, Spain. It should be mentioned that in this analysis, we have excluded all computational costs and times required by the communication between the *PLC* and *PC PROC*, as well as the time needed to send the inspection results to the *BLUE PC*. Moreover, due to the computational power of the current graphics cards (*GPUs*), it was necessary to perform a comparison of the algorithm implemented in both *CPUs* and *GPUs*.

Thus, Fig. 14(a) shows the *pre-processing* step performance. First, it can be seen that the *CPUs* are unable to handle and correctly perform the online processing required by the 23 cameras, and are only able to support 18 cameras. This is not due to a computational power problem but rather to a RAM bus access problem. When this

<sup>1</sup>The tests were performed using 2 Intel® Xeon® E5645 48GB DDR3-1066 processors. For more information about the main characteristics of the CPU, please visit [http://www.matrox.com/imaging/media/pdf/products/supersight/supersight\\_e2.pdf](http://www.matrox.com/imaging/media/pdf/products/supersight/supersight_e2.pdf)

<sup>2</sup>The tests were performed using 2 Nvidia® GeForce GTX 1080 (GP104-400) GPUs with 8GB GDDR5X of VRAM capacity. For more information about the main characteristics of the GPU, please visit <http://www.nvidia.es/graphics-cards/geforce/pascal/gtx-1080/>

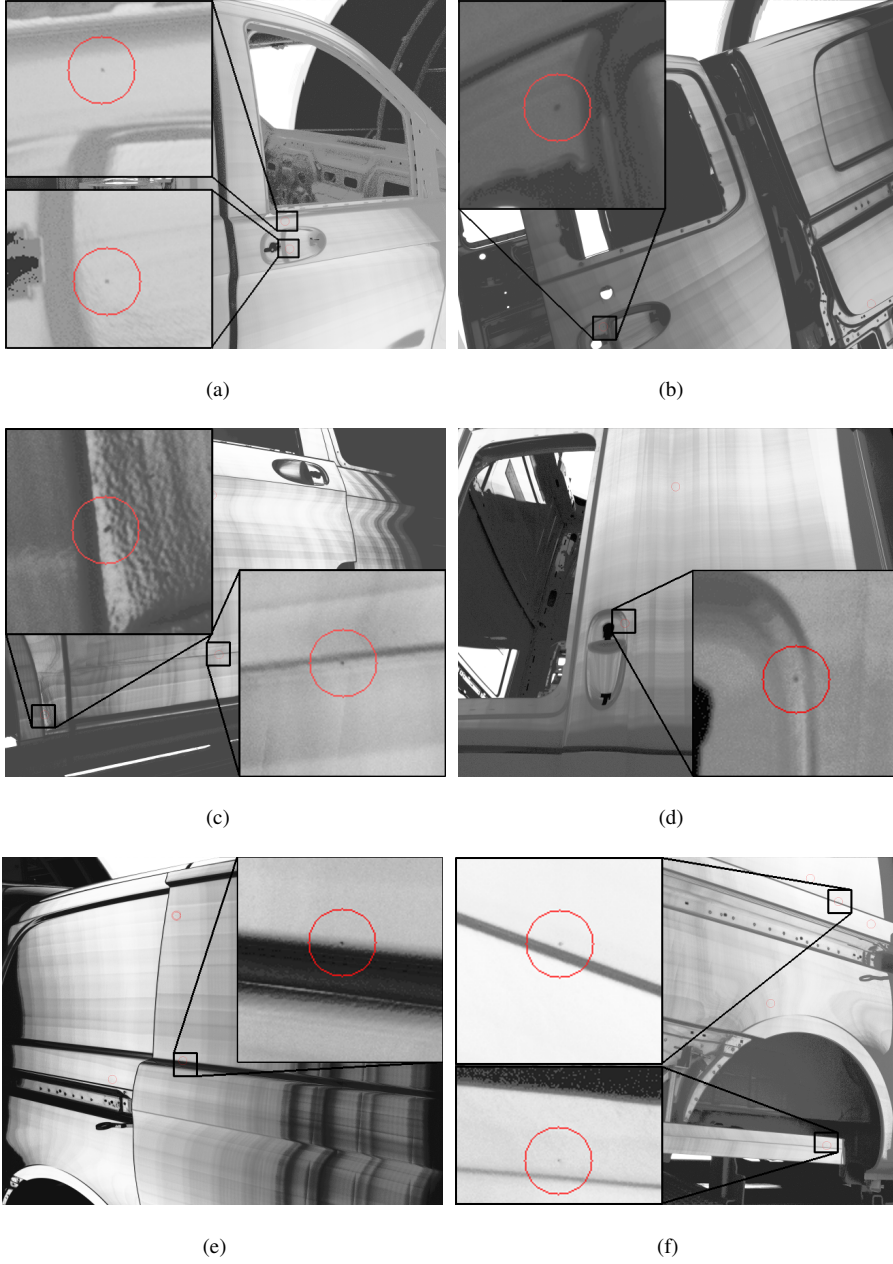


Figure 15: Set of detection examples showing different views and the detection of defects on different car bodies.

happens, the whole system freezes, resulting in a loss of frames, which is of course totally unacceptable.

On the other hand, it is appreciated that the *GPUs* are able to support all cameras without problems, with the maximum time needed to run our algorithm being  $23ms$  per frame. This means that the system is capable of additional computation which could be used to detect other kinds of defects or indeed to improve the current proposal. With regard to the *post-processing* step, Fig. 14(b) shows the results obtained by both implementations, with  $8s$  needed for the *CPU* implementation and  $2s$  for the *GPUs* when dealing with 23 cameras. Both implementations are within the limit of  $10s$  mentioned two paragraphs previously, meaning both satisfy the production line cycle time requirements. However, using the *CPU* implementation, it is not possible to add additional detection algorithms or other kinds of computation such as the communications mentioned previously between the *PLC* and *BLU PC*.

This study shows that the use of the *GPUs* not only allows our approach to be implemented in a system with so many cameras working in parallel but also allows the addition of more algorithms to detect other types of defects not identified by the algorithm proposed in this paper, possibly large dints and dents. That is why Mercedes-Benz' *QEyeTunnel* inspection system in Vitoria uses *GPUs* as the platform for processing artificial vision algorithms and all results referred to here were obtained using implementation on these *GPUs*.

#### 4.3. Defect Detection Analysis and Results

To demonstrate the performance of the defect detection algorithm presented in this paper, a selection of results obtained on the production line with different parts and car bodies is shown in Fig. 15. In some of these, detections are observed close to or on style lines, as in Figs. 15(c), 15(e) and 15(f). In addition, it is possible to see detections carried out in concave areas such as the handle, Figs. 15(a), 15(b) and 15(d), as well as in parts where there is little surface area for detection such as the side door track cover, Fig. 15(f). Note that current industrial solutions, such as [5, 6], have to mask all these areas to give a margin of safety and, as one can see, greatly reduce the area of inspection.

An interesting aspect to be analyzed is how external aspects such as the sectors and movement of illumination patterns, as well as varying car body positions while stationary in the inspection position, affect the proposed detection algorithm. For this, a car body with commonly found defects on the entire surface is used and two kind of experiments are carried out:

- 1) The car body is placed in the inspection position and a total of 15 inspections are carried out without changing its position. The information obtained from this analysis will determine the variability in the detection taking into account only the variability of the movement of the light pattern and its reflection on the body. This analysis is known as a repeatability test.
- 2) The car body is introduced into the system and positioned in the inspection position. Once the inspection is completed, the car body comes out and the cycle is repeated 15 times. This will determine the effect on our proposal of the variability caused by the position of the car body, as well as other factors such as the movement of the same on the production line which can cause movement of some parts (i.e. side doors, rear doors or back door). This analysis is known as a reproducibility test.

Thus, Fig. 16 shows the results obtained in the first test, the repeatability test. It is observed that the detection has a variability of around 12%, which in terms of repeatability means we are at around 90%, a very satisfactory result in terms of industrial vision systems, thus corroborating the robustness of our proposal.

With regard to the reproducibility test, Fig. 17 shows the results obtained. These indicate that our proposal continues to have a variability of around 11%, which in terms of reproducibility means we achieve around 90%, corroborating again the robustness of the proposal.

It should be noted that this study has indicated that most false positives (around 9%, as shown in the figures) are very small defects (less than  $0.1mm$  in diameter), which are at the limit of the system resolution.

In addition, we would like to show an analysis of the detection results using different car bodies with different colors in order to determine the robustness of the proposed

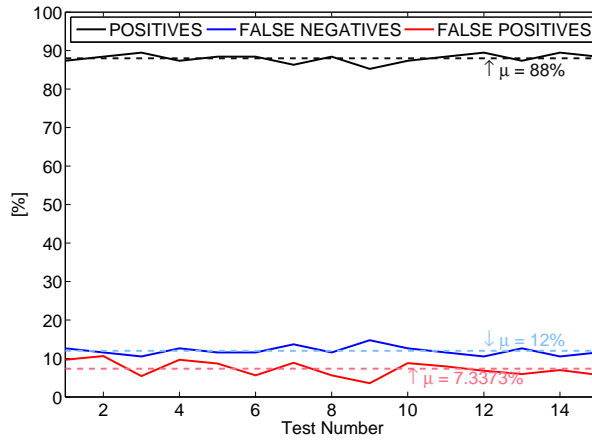


Figure 16: Results of the repeatability test.

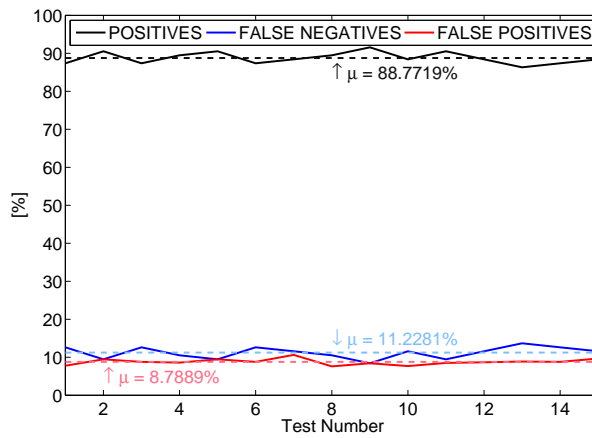


Figure 17: Results of the reproducibility test.

approach in relation to the color variable. For this, Fig. 18 shows the mean of 10 car bodies of several different colors with defects detected, which have been separated into four categories: “C+S”, defects detected by both check-men and the system; “NC”, defects detected by the system but not by the check-men; “NS”, defects detected by the check-men but not by the system; and “FP”, when the system warned of a defect but the check-men did not find it, either because it was not present, or because the size and



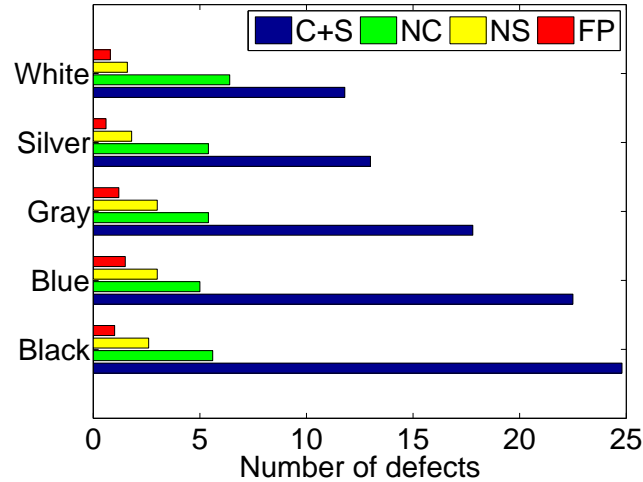


Figure 18: Inspection results comparison between the automatic inspection and the manual visual inspection, classified by car body colors: “C+S”, defects detected by both check-men and the system; “NC”, defects detected by the system but not by the check-men; “NS”, defects detected by the check-men but not by the system; “FP”, when the system warned of a defect but the check-men did not find it, either because it was not present, or because the size and area made it difficult to detect.

area made it difficult to detect.

The figure shows that, regardless of car body color, the number of false positives (around 5% on average), the number of non-detections of the system (around 8% on average), as well as the number of non-detections by check-men (about 13% on average), all remain constant. These results show that reliability is close to 85%, which again is a very strong performance for an industrial system of this complexity.

Finally, to give an indication of its detection capabilities, our algorithm is currently capable of detecting the following types of defects: craters, dirt, stripes, humidity marks, hair, and drops of various types.

## 5. Discussion

Fig. 18 shows an evident improvement in the number of defects detected by our algorithm with respect to the quality control performed by the *check-men*. It also shows a certain dependence on the color of the body analyzed, where for certain colors there

is a clear drop in the number of defects detected not only by the system but by *check-men*. To the best of our knowledge, there are two factors that can explain this result: the first is related to the diffuse light, while the second has to do with paint properties and the painting method used for each color. If there is a direct relation between the diffuse light and color, then, depending on the color, there will be an effect on either the light intensity or the threshold values, or both, and so the setup of the system is made more difficult.

It should be mentioned that, given the supposed robustness of our approach with respect to the quantity of light reflected, as shown in the proof in Appendix A, currently, both the light intensity of the projected patterns and threshold values "*th*" are constant for all bodies regardless of color. However, the results of Fig. 18 led us to query our claims of robustness.

After consulting with car body painting experts, however, they have confirmed to us that the effect is due to the manner in which dark colors are painted, and that it is more common to have a higher number of defects with dark colors than with light colors. Therefore, we are confident that the results in Fig. 18 do not reflect a problem with the robustness of our algorithm but instead can be attributed to the painting process.

Currently, the threshold value "*th*" is obtained following the next process: given a wide set of fusion images from different car bodies, regardless of the color, "*th*" is set so that false positives are minimized and true positives are maximized. The optimal "*th*" is taken as the best for all the samples of the dataset. Since the knowledge of what is a false positive and what is a false negative is unknown a priori, a first adjust is performed, which depends on the subjectivity and knowledge of the person carrying out this setup. For this reason, once the system is working in the production line, a second adjust using the check-men feedback is performed. Due to each camera has a different "*th*", this process is done for every single camera of the system. Given that this manual process is very time consuming in terms of human resources, nowadays we are working in a method for computing threshold values automatically.

Another aspect to be discussed is the proposed multi-level structure method. Images are of  $5MP$ , which is a resolution of  $2592 \times 2048$  pixels, which are repeatedly reduced while moving from one level to another. In particular, in this paper we have used an

optimal third-level structure based on experimental results, with a final image after two reductions of  $648 \times 512$  pixels. Further levels of reduction did not produce any noticeable improvement. However, it is possible to introduce more levels by using higher resolution cameras, therefore increasing the quality of the detection.

Finally, and with respect to the non-detections shown in the experimental results section, it is worth remarking that most of these were found in areas where illumination was not of sufficient quality, a problem sometimes caused by the illumination patterns projected.

In this paper, we are using patterns not related to the 3D model of the body to be inspected. As a consequence, in concavities the light changes its velocity, its width and even its shape. However, were it possible to generate and project patterns that fit the shape of the body at any instant during the scanning stage, the majority of non-detections caused by illumination factors could be eliminated, thus improving the detection capacity of the system. This will be a topic for further research.

## 6. Conclusion

In this paper, a new vision algorithm based on deflectometry techniques for detecting small defects on specular surfaces in general, and car body surfaces in particular, has been presented. Our approach is based on the information obtained by reflecting a light pattern on the specular surface to detect structural deformations thereon.

The algorithm relies on two steps: a *pre-processing* step, performed during the scanning process, and a *post-processing* step, where, using the information from the *pre-processing* step, defects are detected.

With regard to the *pre-processing* step, we have detailed a new image fusion algorithm that allows us to reduce the effect of the diffuse light counterpart, thus leaving us with the interesting information contained in the specular lobe. Moreover, the proposed law enhances the contrast between pixels with low levels of intensity (indicating the presence of defects) and pixels with high levels of intensity (indicating the absence of defects), thus facilitating their detection in the later phase of the presented algorithm. A comparison of the result given by our proposal and that given by other

industrial solutions has been presented, proving that our solution provides better fusion images from a defect detection point of view.

As regards the *post-processing* step, a novel image processing is presented, which allows defect detection in concave areas or those with abrupt changes in the surface such as handles, style lines, edges and corners, provided that the area is well lit. As far as we are aware, this is the first work dealing with this problem. Our approach has as its foundation a new background extraction technique based on a local directional blurring through which low-level information in the fusion image, showing large surface deformations such as concavities or style lines, is differentiated from high-level information, in which small surface defects are found. Once the background is extracted, a new contrast enhancement law, which uses the computed background image as a mean and where only the positive values in standard deviation are taken into account, has been presented. Using this, the entire image is homogenized to zero level and defects are enhanced from noise. Finally, a global threshold is set in order to distinguish defects from noise. We have seen a typical example of a defect very close to a style line, and explained step by step the solutions given by the presented algorithm.

Since the image background extraction solution proposed in this paper is dependent on the kernel size chosen to perform the blurring, defects of a larger size could conceivably be considered part of the surface shape. That is why we have proposed a solution based on a multi-level detection, where each lower level uses as input the image background computed in the higher level but rescaled to a smaller size. Thus, using the same kernel size for the proposed blurring and post-processing algorithm, larger defects are also detected, as illustrated in a practical example presented in the paper.

In addition, the paper shows and analyzes the results of the proposed detection algorithm implemented in the industrial *QEyeTunnel* inspection system developed at the Mercedes-Benz factory in Vitoria, Spain. First, the study of the algorithm's computational cost has shown that, in order to meet cycle time constraints of the production line, our algorithm needed to be implemented using a platform of *GPUs*, with a maximum processing time of *23ms* per frame in the *pre-processing* step and of *2s* in total in the *post-processing* step using 23 cameras working in parallel. Subsequently, several

examples of defects located in or close to style lines, edges, corners and handles have been shown, thus demonstrating the detection power of our algorithm. An analysis of the robustness of the proposed algorithm has also been carried out. The results have shown that the system has a variability of less than 12% if we include false positives and non-detections. Finally, the comparison of the manual detection checks performed by experts at Mercedes-Benz and the detection carried out by our algorithm has shown that we achieved an average improvement in detection of 40% and with a mean of false positives of only 7%.

### AppendixA. Proof of the Ambient and Diffuse Light Reduction

Let us use the *rendering equation* [22, 23] which, in computer graphics, is an integral equation in which the equilibrium radiance leaving a point is given as the sum of emitted plus reflected radiance under a geometric optics approximation:

$$\mathbf{L}_o(\mathbf{p}, \omega_o, \omega_i) = \int_{\Omega} f_r(\mathbf{p}, \omega_i, \omega_o) \cdot \mathbf{L}_i(\mathbf{p}, \omega_i) \cdot (\omega_i \cdot \mathbf{n}) \cdot d\omega_i \quad (\text{A.1})$$

where  $\mathbf{L}_o$  is the total spectral radiance along direction  $\omega_o$ , from a particular position  $\mathbf{p}$ , where  $\mathbf{p}$  is the location in space,  $\omega_o$  is the direction of the outgoing light;  $f_r$  is the bidirectional reflectance distribution function, the proportion of light reflected from  $\omega_i$  to  $\omega_o$  at position  $\mathbf{p}$ ;  $\mathbf{L}_i$  is spectral radiance coming inward toward  $\mathbf{p}$  from direction  $\omega_i$ , and  $\mathbf{n}$  is the surface normal at  $\mathbf{p}$ . Note as  $\omega_i \cdot \mathbf{n}$  is the weakening factor of inward irradiance due to incident angle, as the light flux is smeared across a surface whose area is larger than the projected area perpendicular to the ray.

This equation models the behavior of a light source on a point of a surface with certain physical properties seen from a particular view (see the reflectance model in Fig. 3). Several lighting models have been developed over the years to solve the *rendering equation*. Here we are going to use the *Cook-Torrence model* [24] since it is closer to the physical reality than others like the Phong [25] or Blinn-Phong [26] models. The *Cook-Torrence model* contains both a diffuse and specular part:

$$f_r = k_d \cdot f_{\text{lambert}} + k_s \cdot f_{\text{cook-torrence}} \quad (\text{A.2})$$

Here,  $k_d$  is the ratio of incoming light energy that gets refracted with  $k_s$  being the ratio that gets reflected. The left side of A.2 states the diffuse part and is known as *Lambertian diffuse* which is a constant factor denoted as:

$$f_{lambert} = \frac{c}{\pi} \quad (\text{A.3})$$

with  $c$  being the surface texture. The specular part of the equation is described as:

$$f_{cook-torrance} = \frac{D \cdot F \cdot G}{4 \cdot (\boldsymbol{\omega}_o \cdot \mathbf{n})(\boldsymbol{\omega}_i \cdot \mathbf{n})} \quad (\text{A.4})$$

where  $D$  is the normal distribution function,  $G$  the geometric function and  $F$  the Fresnel equation. Each of these functions is an approximation of their physics equivalents and it is possible to find more than one version of each one that aims to approximate the underlying physics.

In this proof, we use the approximation given by Trowbridge-Reitz [27] for the normal distribution function  $D$ :

$$D_{TR}(\mathbf{n}, \mathbf{c}, \alpha) = \frac{\alpha^2}{\pi \cdot ((\mathbf{n} \cdot \mathbf{c})^2 \cdot (\alpha^2 - 1) + 1)^2} \quad (\text{A.5})$$

where  $\mathbf{c} = \frac{\boldsymbol{\omega}_i + \boldsymbol{\omega}_o}{\|\boldsymbol{\omega}_i + \boldsymbol{\omega}_o\|} \in [0, 1]$  is the halfway vector to measure against the surface's microfacets, with  $\alpha$  being a measure of the surface's roughness.

The geometric function is defined by the Schlick-Beckmann approximation as:

$$G_{Schlick}(\mathbf{n}, \boldsymbol{\omega}_o, k) = \frac{\mathbf{n} \cdot \boldsymbol{\omega}_o}{(\mathbf{n} \cdot \boldsymbol{\omega}_o) \cdot (1 - k) + k} \quad (\text{A.6})$$

$k = \frac{(\alpha+1)^2}{8}$  being a remapping of  $\alpha$  using direct lighting.

The Fresnel equation can be approximated using the Fresnel-Schlick approximation [28]:

$$F_{Schlick}(\mathbf{n}, \boldsymbol{\omega}_o, F_o) = F_o + (1 - F_o) \cdot (1 - (\mathbf{n} \cdot \boldsymbol{\omega}_o))^5 \quad (\text{A.7})$$

where  $F_o$  represents the base reflectivity of the surface, which we calculate using something called the *indices of refraction*.

Without loss of generality, let us suppose that we have an ideal point of light in two different instants of time,  $\boldsymbol{\omega}_i(k)$  y  $\boldsymbol{\omega}_i(k-1)$ , with both the point of view (camera

position)  $\omega_o$  and surface point  $\mathbf{p}$  being constant. Note that by doing this, both  $G_{Schlick}$  and  $F_{Schlick}$  remain constant since they are not affected by  $\omega_i$ . Thus:

$$\mathbf{L}_o(\mathbf{p}, \omega_o, \omega_i, k) - \mathbf{L}_o(\mathbf{p}, \omega_o, \omega_i, k-1) = k_d \cdot \frac{c}{\pi} \cdot \mathbf{n} \cdot (\omega_i(k) - \omega_i(k-1)) + (D_{TR}(k) - D_{TR}(k-1)) \cdot \frac{k_s \cdot G_{Schlick} \cdot F_{Schlick}}{4 \cdot \omega_o \cdot \mathbf{n}} \quad (\text{A.8})$$

In the above equation, it can be observed that the first summing corresponding to both the diffuse and ambient parts of the light is directly affected by the difference of the light source position in two different instants of time. The smaller this difference, the less the contribution of the diffuse and ambient parts of the light. In the same way, the second summing corresponds to the contribution made by the specular part of the light. It is observed in this case that the effect is inversely proportional and raised to the fourth, the reason why small movements of the light source do not “cancel” the specular contribution of the light.

This phenomenon is observed in Fig. A.19 and A.20 in which several consecutive frames and their respective histograms are shown. On one hand, Fig. A.19 shows the result obtained using Eq. A.1 directly. As can be seen, as the light source changes its position ( $\omega_i(k)$ ), the histogram varies considerably due to the contribution of the diffuse and the ambient parts of the light. This effect causes a problem for image fusion techniques because it affects the brightness and contrast of the final result, affecting in our case the contrast between defects and background. On the other hand, Fig. A.20 shows the effect that is obtained when carrying out the proposed operation in this paper:  $\mathbf{L}_o(\mathbf{p}, \omega_o, \omega_i, k) - \mathbf{L}_o(\mathbf{p}, \omega_o, \omega_i, k-1)$ . One can see that histograms at different instants of time remain constant, thus proving that the result seen in Eq. A.8 effectively and considerably reduces the environmental and diffuse effects.

### Acknowledgments

This work is supported by VALi+d (APOSTD/2016/044) and PROMETEO (PROMETEOII/2014/044) Programs, both from Conselleria d'Educaci'ó, Generalitat Valenciana.

The authors would also like to thank the heads of project management and the check-men at the Mercedes-Benz factory in Vitoria, Spain, for all their efforts during

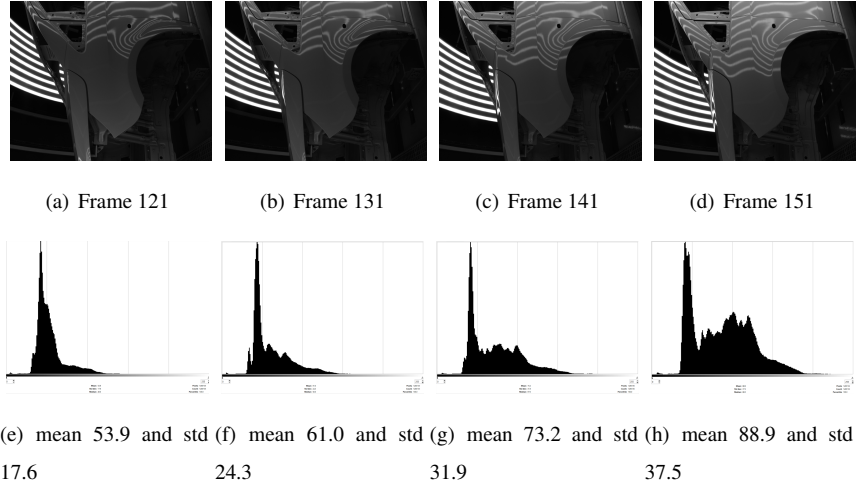


Figure A.19: Effect of the diffuse light at different instants during the scanning stage when using only  $i_k$  for fusing.

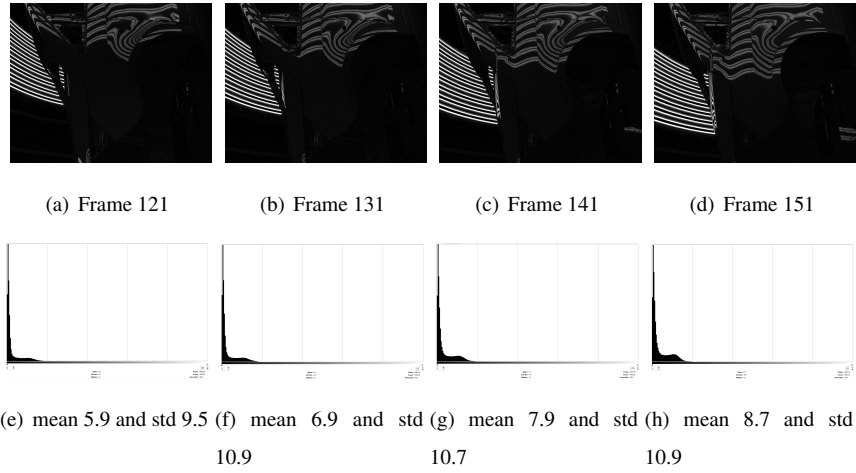


Figure A.20: Effect of the diffuse light at different instants during the scanning stage when using  $i_k - i_{k-1}$  for fusing.

the project, especially Juan Carlos Villanueva, Justo Oria and Juan Carlos Sadaba.



## References

- [1] B. Denkena, H. Ahlers, F. Berg, T. Wolf, H. Tonshoff, Fast inspection of larger sized curved surfaces by stripe projection, *CIRP Annals - Manufacturing Technology* 51 (1) (2002) 499 – 502. doi:[http://dx.doi.org/10.1016/S0007-8506\(07\)61570-4](http://dx.doi.org/10.1016/S0007-8506(07)61570-4).
- [2] F. P. Leon, S. Kammel, Inspection of specular and painted surfaces with centralized fusion techniques, *Measurement* 39 (6) (2006) 536–546.
- [3] X. Gros, *NDT Data Fusion*, Elsevier Science, 1996.  
URL [https://books.google.es/books?id=zDZGZVMG\\_IoC](https://books.google.es/books?id=zDZGZVMG_IoC)
- [4] J. J. Clark, A. L. Yuille, *Data Fusion for Sensory Information Processing Systems*, no. 105 in *The Kluwer International Series in Engineering and Computer Science*, Kluwer, 1990.
- [5] L. Armesto, J. Tornero, A. Herraez, J. Asensio, Inspection system based on artificial vision for paint defects detection on cars bodies, in: *IEEE International Conference on Robotics and Automation, ICRA 2011, Shanghai, China, 9-13 May 2011, 2011*. doi:10.1109/ICRA.2011.5980570.  
URL <http://dx.doi.org/10.1109/ICRA.2011.5980570>
- [6] J. Tornero, L. Armesto, M. C. Mora, N. Montés, A. Herráez, J. Asensio, Detección de defectos en carrocerías de vehículos basado en visión artificial: Diseño e implantación, *Revista Iberoamericana de Automática e Informática Industrial RIAI* 9 (1) (2012) 93–104. doi:<http://dx.doi.org/10.1016/j.riai.2011.11.010>.  
URL <http://www.sciencedirect.com/science/article/pii/S1697791211000793>
- [7] J. Santolaria, J. Velázquez, D. Samper, J. Aguilar, I. Escursell, *Sistema de inspección de defectos opel españa* (2016).  
URL <http://www.opel.es/acerca-de-opel/noticias-opel/>

2016/septiembre/i3a-y-gm-sistema-deteccion-defectos-pintura-coches.  
html

- [8] Micro-Epsilon, Sistema de inspección de defectos opel españa (2015).  
URL <http://www.micro-epsilon.com/measurement-systems/Paint-Inspection/karosserie/index.html>
- [9] M. Imanishi, K. Yoshida, T. Asaeda, Y. Suzuki, S. Chida, M. Watanabe, Surface defect inspection apparatus, uS Patent 5,726,705 (Mar 1998).  
URL <https://www.google.ch/patents/US5726705>
- [10] K. Alders, M. Lehe, G. Wan, Method for the automatic recognition of surface defects in body shells and device for carrying out said method, uS Patent 6,320,654 (Nov 2001).  
URL <https://www.google.ch/patents/US6320654>
- [11] D. Clarke, R. Reynolds, T. Pryor, Panel surface flaw inspection, uS Patent 4,629,319 (Dec 1986).  
URL <https://www.google.com/patents/US4629319>
- [12] M. Tornero, A. Mora, M. Herraéz, S. Montes, System of detection of faults in surfaces by merging of images by means of light sweeping, cA Patent App. CA 2,684,365 (Oct. 23 2008).  
URL <http://www.google.com.ar/patents/CA2684365A1?cl=en>
- [13] C. Prior, P. Simon, M. Herraéz, M. Asensio, M. Tornero, N. Ruescas, A. Armesto, Inspection system and method of defect detection on specular surfaces, wO Patent App. PCT/IB2010/052,193 (Nov. 24 2011).  
URL <https://google.com/patents/WO2011144964A1?cl=en>
- [14] S. Kammel, F. Puente León, Deflectometric measurement of specular surfaces, IEEE Transactions on Instrumentation and Measurement 57 (4) (2008) 763–769.  
doi:10.1109/IMTC.2005.1604173.
- [15] T. Kiuchi, K. Ikeuchi, Roughness and shape of specular lobe surfaces using photometric sampling method, in: Proceedings of IEEE Conference on Computer

- Vision and Pattern Recognition, 1993, pp. 765–766. doi:10.1109/CVPR.1993.341164.
- [16] A. A. Goshtasby, 2DD and 3-D Image Registration: For Medical, Remote Sensing, and Industrial Applications, Wiley-Interscience, 2005.
- [17] S.-s. Chen, Stochastic and neural methods in signal processing, image processing, and computer vision, International Society for Optical Engineering, vol. 1766 of Proceedings of SPIE, (San Diego, CA), July 1992, 1992.
- [18] A. Yogeswaran, P. Payeur, 3D Surface Analysis for Automated Detection of Deformations on Automotive Body Panels, INTECH Open Access Publisher, 2012.  
URL <https://books.google.es/books?id=k4Z9oAEACAAJ>
- [19] V. Borsu, A. Yogeswaran, P. Payeur, Automated surface deformations detection and marking on automotive body panels, in: 2010 IEEE International Conference on Automation Science and Engineering, 2010, pp. 551–556. doi:10.1109/COASE.2010.5584643.
- [20] G. Bergmann, Á. Horváth, I. Ráth, D. Varró, A benchmark evaluation of incremental pattern matching in graph transformation, in: H. Ehrig, R. Heckel, G. Rozenberg, G. Taentzer (Eds.), Proc. 4th International Conference on Graph Transformations, ICGT 2008, Vol. 5214 of Lecture Notes in Computer Science, Springer, Springer, 2008, pp. 396–410, acceptance rate: 40%.
- [21] R. C. Gonzalez, R. E. Woods, Digital Image Processing (3rd Edition), Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.
- [22] D. S. Immel, M. F. Cohen, D. P. Greenberg, A radiosity method for non-diffuse environments, in: Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '86, ACM, New York, NY, USA, 1986, pp. 133–142. doi:10.1145/15922.15901.  
URL <http://doi.acm.org/10.1145/15922.15901>
- [23] J. T. Kajiya, The rendering equation, in: Computer Graphics, 1986, pp. 143–150.

- [24] R. L. Cook, K. E. Torrance, A reflectance model for computer graphics, *ACM Trans. Graph.* 1 (1) (1982) 7–24. doi:10.1145/357290.357293.  
URL <http://doi.acm.org/10.1145/357290.357293>
- [25] B. T. Phong, Illumination for Computer-Generated Pictures, *Comm. of the ACM* 18 (6) (1975) 311–317.
- [26] J. F. Blinn, Models of light reflection for computer synthesized pictures, *SIGGRAPH Comput. Graph.* 11 (2) 192–198. doi:10.1145/965141.563893.
- [27] B. Walter, S. R. Marschner, H. Li, K. E. Torrance, Microfacet models for refraction through rough surfaces, in: *Proceedings of the 18th Eurographics Conference on Rendering Techniques, EGSR'07*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 2007, pp. 195–206. doi:10.2312/EGWR/EGSR07/195-206.
- [28] C. Schlick, An inexpensive brdf model for physically-based rendering, *Computer Graphics Forum* 13 (1994) 233–246.