Article

Machine Vision system for counting small metal parts in electro-deposition industry

Rocco Furferi 1,*, Lapo Governi 1, Luca Puggelli 1, Michaela Servi 1 and Yary Volpe 1

1 Department of Industrial Engineering, University of Florence (Italy); ingind@pec.unifi.it
* Correspondence: rocco.furferi@unifi.it; Tel.: +39-055-2768741

Featured Application: The present work has application in the field of galvanic coating for fashion industry by proposing a method and a machine able to count the number of items attached to a galvanic frame.

Abstract: In the fashion field, the use of electroplated small metal parts such as studs, clips and buckles is widespread. The plate is often made of precious metal, such as gold or platinum. Due to the high cost of these materials, it is strategically relevant and of primary importance for manufacturers to avoid any waste by depositing only the strictly necessary amount of material. To this aim, Companies need to be aware of the overall number of items to be electroplated so that it is possible to properly set the parameters driving the galvanic process. Accordingly, the present paper describes a Machine Vision-based method able to automatically count small metal parts arranged on a galvanic frame. The devised method relies on the definition of a proper acquisition system and on the development of image processing-based routines. Such a system is then implemented on a counting machine is meant to be adopted in the galvanic industrial practice to properly define a suitable set or working parameters (such as current, voltage and deposition time) for the electroplating machine and, thereby, to assure the desired plate thickness from one side and to avoid material waste on the other.

Keywords: Machine Vision; Morphological image filtering; Galvanic Industry; Rear-projection.

1. Introduction

As widely recognized, electroplating (more precisely electrodeposition) is a chemical process that uses electric current to transfer metal from a cation to an electrode (i.e. the object to be treated), to form a coherent thin metal coating [1]. The amount of mass deposit derives from Faraday’s laws on electrolysis [2] and directly depends on current intensity and time:

\[ m = \frac{M \cdot I \cdot t}{Z \cdot F} \]  

Where:
- \( m \) = mass deposited on electrode;
- \( M \) = molar mass of the material to be deposited;
- \( I \) = current intensity (Ampere);
- \( t \) = time;
- \( Z \) = valence of material’s ions;
- \( F \) = Faraday’s constant (96485.33 C mol⁻¹)

The process used by manufacturers working in the electrodeposition of fashion accessories field consists of arranging the (usually) small parts to be plated on a frame by using hooks or, more often, metal wires as shown in Figure 1.
Therefore, electrodeposition simultaneously occurs on a number of parts. Since the material to be deposited on electrode (multiple items to be plated) is required to form a uniform thin layer, the overall mass is given by:

\[ m = n \cdot s \cdot T \cdot \rho \]  

(2)

Where:
- \( n \) = number of items to be electroplated;
- \( s \) = surface of a single item;
- \( T \) = coating thickness;
- \( \rho \) = material mass density.

Hence, in order to obtain a desired coating thickness, both number and surface of items to be plated need to be known. While the items surface is retrievable by means of possibly available items CAD models or by using 3D scanning, the number of items arranged on the galvanic frame is not straightforwardly available in order to compute the overall surface to be electroplated.

To date, the parts attached to the frame are manually counted, however the reliability of the process is limited by ensuing weakness and inattentiveness; in other words, it is inevitably prone to errors due to operators’ tiredness, lack of attention, etc.

In scientific literature, several papers specifically address the topic of designing counting systems with reference to a variety of industrial fields [3], [4]. In addition, many counting machines have been available on the market for years [5], [6]. Unfortunately, regardless of the technology adopted (e.g. weight measurement, free-fall, optical scan lines), almost all the machines available on the market requires items to be physically separated one from each other (i.e. not disposed on package or frames), or to be arranged upon a moving tray. Therefore, such solutions are not suitable or adaptable to count items that are already arranged on a galvanic frame.

In this paper, a Machine Vision-based method to automatically count small metal parts arranged on a galvanic frame is proposed. The devised method, relying on the definition of a proper acquisition system and on the development of image processing-based routines, is implemented on a counting machine to be adopted in the galvanic industrial practice. The correct knowledge of the number of items will allow companies in defining a suitable set or working parameters (such as current, voltage and deposition time) for the electroplating machine and thereby to assure the desired plate thickness from one side and to avoid material waste on the other.

2. Materials and Methods
As shown in Figure 1, the galvanic frame is formed by 4 tubular beams welded to compose a rectangle as shown in Figure 1. In the general configuration, on the shorter sides several hooks are joined. Workers use inert metal wires to knot together a variable number of items. Successively, each wire is linked to a couple of corresponding hooks (i.e., the nth on the upper side with the nth on the lower) so that the wire results to be arranged on the frame along an approximately vertical direction. Once all the couples of hooks are filled, the galvanic frame is sent to the electroplating bath. Considering that items can be very small in size (down to 10 mm on the longer side), the variability in length (thus in mass) of the wire itself precludes the adoption of any weight-based approach for the counting system.

Consequently, the attention has been focused on computer vision-based approaches. The main idea is to properly acquire a 2D digital image of the frame, on which to detect each item by means of computer vision (CV) tools [7].

2.1. Literature methods

According to scientific literature, several different CV approaches can be adopted. Considering the task, three among them seem to deserve further investigation: deterministic template-matching, neural network-based algorithms or brightness-based segmentation. The applicability, effectiveness and robustness of each of them strictly depends on the “typology” of the image to be analysed. It has to be noticed that other approaches such as colour-based ones are not applicable since wire colour can be very close to the items one.

More in detail, deterministic template-matching algorithms are intended to find, into an image, instances of a given template. For example, OCR (Optical Character Recognition) procedures – which recognize text within pictures (e.g., a PDF file) – are usually built based on template matching algorithms. This approach would be optimal to solve our problem, if only items are arranged on a rigid grid, so that each item was oriented in the same way with respect to the camera. Actually, some companies use galvanic frames where items are placed into a fixed position on the frame, as shown in Figure 2.

![Figure 1. Galvanic frame with items placed in a fixed position.](image)

However for such cases, operators usually fill all the available slots with items; therefore the number of items on the galvanic frame is known a priori. As already mentioned, however, in the general configuration described before, items are knotted on wires and, consequently, their orientation in space is far from being equal. This issue inevitably limits the applicability of deterministic template-based algorithms and makes their adoption inconvenient for the specific case analysed in this paper.

With respect to this limitation, an evolution of template-based algorithm, as defined above, can be found on neural network (NN) based approaches [8–11]. Some of them, in fact, are able to detect a specific object independently from its orientation and position in the scene. Among them, YOLO (You Only Look Once) [12] is a state-of-the-art real-time object detection system, targeted for real-time processing. Object detection is a computer technology related to computer vision and image
processing that deals with detecting instances of semantic objects of a certain class (e.g. humans, cars, etc.) in digital images and videos.

Differently from prior approaches which apply the model to an image at multiple locations and scales and then high scoring regions of the image are considered detections, YOLO applies the network to the full image. Specifically, the image is divided into an S x S grid and the algorithm returns bounding boxes and predicted probabilities for each of these regions. The method used to compute these probabilities is logistic regression [13]. This way, other than performing a very fast detection, predictions are informed by global context in the image.

Off-the-shelf YOLO nets with pre-trained weights although cannot be able to predict our subject of interest, as they have not been trained in detecting these particular objects (see Figure 3).

![Figure 2](image)

**Figure 2.** Prediction result obtained with an off-the-shelf pre-trained YOLO net.

On the other hand, to achieve a proper result using these networks it is not sufficient to provide a limited number of training images (e.g. ten to twenty images). In the light of these considerations, and given that the shape of the items can change frequently (every month all lots can be completely new), it is not practical for users to train the algorithm each time.

For this reason, another approach has been explored: the classical brightness-based segmentation [14–16]: assuming that items brightness (or range of brightness) is different from background one, it is possible to isolate the background itself. The resulting image contains only pixels belonging to items and connecting wires. Unfortunately, wire and items colours can be very close, so that colour segmentation cannot be used to separate the ones from the others. Fortunately, since wires are thinner than items, their pixels can be removed by means of CV tools such as pixel erosion/dilation. The number of separate clusters of pixels describing the items can then be easily retrieved by means of labelling tools.

In detail, a threshold value (or at least a range) must be used in order to isolate on the image only pixels relative to the items to be counted. Supposing that thresholding operation works flawlessly, a binary image can be obtained where white pixels represent items to be counted and black pixels are the background and the wires. Afterwards, many well-known algorithms can be used to count the number of isolated regions in binary images.

2.2. Image acquisition requirements

The brightness-based segmentation approach seems the most promising for the specific application but needs to be tailored to the peculiarities entailed by small dimensions and high reflectivity of items to be counted. Consequently, the definition of a proper input image has a primary importance for the success of the method.

Depending on the finishing and on the material of which items are made, their aspect is rarely opaque but rather it is highly reflexive. Obviously, even colour may change, varying from copper-like to silver and gold thus resulting in different brightness. All these characteristics make it very difficult (or even impossible) to obtain a satisfactory thresholding values or ranges, on which set the segmentation.
To make it more complicated, the silhouette of same items knotted to the galvanic frame varies significantly, due to their almost-random orientation. Moreover, the placement is far from being equally spaced.

Therefore, in order to make the segmentation algorithms effective, it is of critical importance to obtain a suitable image where it is possible to separate items from background. To this purpose, three different lighting settings have been considered in order to evaluate their efficacy in favouring image segmentation operations:

1. Frontal lighting with black uniform background;
2. Lighted white background;
3. Light rear-projection.

2.3. Light settings

In the industrial practice, a single galvanic frame is filled with a number of identical items. In order to make the performed tests representative and speed-up the testing process, we chose to use typical galvanic frames filled with a variety of items of different shapes and dimensions arranged on vertical metal wires instead of using multiple frames (each one with a single item typology). Image acquisition is carried out by using a Fujifilm T1 SLR camera (APS-C sensor format) and an 18 mm focal length lens. Acquired images have resolution of 15.8 Mpixel (4826x3264 pixel).

3.1.1. Frontal light with black uniform background

The first tested layout setting is meant to “physically” isolate items from the rest of the scene by putting an opaque black canvas behind the frame. The frame, containing four different item typologies, is positioned approximately perpendicular to the camera optical axis at a distance of 500 mm, so that the frame occupies completely the field of view. The set is illuminated by a frontal lighting source (800lm focusable LED torch). This layout allows obtaining an almost-uniform black background on which items shape appear enhanced.

In the resulting digital image, background pixels are characterized by low brightness. On the contrary, items pixels brightness has, as expected, high brightness thanks to the frontal and strong illumination (see for instance Figure 5, referred to four different items). However, this layout results not optimal for a number of reasons. First, background subtraction (i.e. to subtract from the image to be analysed a reference image of the background canvas acquired prior to positioning the galvanic frame) is not applicable since shadows/reflections projected on the canvas by the items and the wires make the background itself different from the reference.

In addition, brightness-based segmentation leads to two additional main issues as explained below:

- items and background pixels may be incorrectly detected/assigned;
- wires and items brightness are similar, thus difficult to separate.
Starting from Figure 5a, it is not possible to isolate items from wires by thresholding, as shown in Figure 5b. The only filtering operation that could allow wires deletion is image erosion. Unfortunately, since items dimensions and wires thickness are similar, the operation (even if combined with successive dilation filtering thus performing a morphological opening) lead to sub-fragmentation of single items into multiple pixels clusters (Figure 5c), thus invalidating the successive counting operation.

Observing in detail the image of two items (named, respectively, item “A” and “B”) right after thresholding (Figure 6a and Figure 6b), the first issue becomes evident. It can be noticed that, for some items, darkest pixels are mistakenly assigned to background. Consequently, some of them result already fragmented into multiple parts (Figure 6c and Figure 6d). In other cases, effects are less evident but equally dangerous due to the successive (required) erosion operation; as shown in Figure 6c, items may result so thinned that successive operations unavoidably cause fragmentation.

More in detail, sub-fragmentation can occur in two possible scenarios: in presence of bridges (Figure 6a) or in case of inner holes Figure 6c. In the first case, the thickness of the bridge may be similar to the thickness of the wires to be removed. Consequently, it commonly happens that the morphological opening on the binary image (i.e. the erosion followed by dilation) remove both wires and bridges, thus causing undesired fragmentation of the cluster (Figure 6d). Similarly, in case of inner holes – given by the actual shape of the item or caused by thresholding – morphological opening may causes fragmentation.
This issue can be possibly avoided by using a morphological image closure (i.e. the dilation followed by erosion) followed by an additional erosion. Figure 7a demonstrates the result of such an operation applied to Figure 6c.

However, also this solution may lead to some unwanted side effects that make this alternative unsuitable. In fact, in Figure 7b it can be noticed that wires form closed loops in the image. In some cases, loops may result completely closed by a morphological closing filtering. If sufficiently large, they can be easily mistaken for items in the counting phase. In addition, if a couple of items is sufficiently close, filter may cause the fusion of the relative clusters into one (see Figure 7c).

Figure 7. (a) Item “B” after morphological closing + image erosion; (b) Wires loop after morphological closing followed by image erosion; (c) Clusters merged after dilation. The image is intentionally left in grayscale to show actual original items.

Between the two alternatives proposed above, the better performing proved to be the first one i.e. the morphological opening-based solution. Starting from the resulting image (Figure 6), connected regions representing actual items need to be discriminated from the ones representing small wire portions and/or item fragments. To this aim, an elective method could be to perform an area-based discrimination carried out by imposing an appropriate area threshold.

Since item dimension is widely variable and unknown a priori, a fixed area threshold value cannot be based on item dimension itself. On the contrary, wire’s dimension is constant. Accordingly, it is possible to define a fixed area threshold under which clusters are considered too small to be an item, thus must be ignored. Considering that the wire’s thickness is approximately 1 mm – corresponding to 7 pixels in the image (based on the shooting setup described in the previous section), a limit dimension has been set at 4 mm$^2$ – corresponding to 200 pixels. In the example shown in Figure 8a, this method allows to appropriately discard small clusters.

However, in several other situations, such as the one depicted in Figure 8b, this criterion leads to misclassification. This is mainly due to the heavy image cluster fragmentation induced by the acquisition setup and subsequent image filtering. Other than the simple criterion described above, other more complex techniques have been tested in order to cluster pixel regions, namely k-means clustering and SVM [17,18]. The results, not detailed in the present paper, show that this misclassification still occurs.
Figure 8. (a) Discrimination among valid and ignored clusters; (b) Misclassification of small clusters: red coloured clusters are discarded since their area is lower than the selected threshold (i.e. 200 pixel); green coloured clusters are counted thus leading to counting error since both belongs to a single item.

3.1.2. Lighted background

Moving from the issues faced with the first setting, the second layout makes use of backlighting. The galvanic frame, containing a set of identical items, is arranged between the camera and an approximately uniformly illuminated white background (see Figure 9).

![Figure 9. Lighted background setup.](image)

Under the proper camera settings, lights saturate brightness for the background, while pixel belonging to items generally appear darker (Figure 10a).

Awkwardly, many item regions appear bright due to specular reflections/inter-reflections among items themselves. Similarly to the configuration described in the previous Section, over-fragmentation issues arise. In fact, despite with the setup the entire background is better detected and isolated, some items portions, which appear light due to the inter-reflections mentioned above, are mistakenly assigned to the background (see Figure 10b). Even using morphological operators, similar to the ones described in Section 3.1.1 the fragmentation issue persists making it practically unfeasible to correctly classify pixel clusters.

![Figure 10. (a) Original image acquired with backlighting; (b) Image of Figure 10a after thresholding.](image)

3.1.3. Rear Projection
To overcome all the discussed drawbacks related to “direct” backlighting, a third solution has been developed and tested. In detail, a canvas for rear-projection has been placed at a 20 mm distance from the galvanic frame containing seven item typologies while light source (in this case an overhead projector with a 3300 lumen light source) and camera have been arranged as depicted in Figure 11.

![Rear projection setup diagram](image)

Figure 11. Rear projection setup.

Experimental tests have shown that, due to the items thickness and shape, the minimum distance between the projector and the frame needs to be set to 1.8m in order to avoid shadow blurring. As depicted in Figure 12a, items cast a very sharp and uniform dark shadow on the canvas. At the same time, wires appear thinner than (for instance) the ones shown in Figure 5b.

Starting from the image shown in Figure 12a, a straightforward binary threshold and subsequent morphological image opening is applied similarly to the approach described for frontal lighting setup described in Section 1. As shown in Figure 12b, this approach leads to minimally fragmented pixel clusters. Therefore, by using an area threshold equal to 200 pixel it is possible to correctly count the items number.

Summing up, the rear projection setup proved to be the most suitable among the tested ones in order to correctly isolate items to be counted. In fact, the key-point of the procedure resides into the very sharp “native” image, in which shadows are extremely defined. Consequently, required filtering operations are way less aggressive than it was needed in the previous cases.

![Image pair](image)

Figure 12. (a) Original image acquired with the rear projection setup; (b) Binary image after morphological opening process.

Summing up, the rear projection setup proved to be the most suitable among the tested ones in order to correctly isolate items to be counted. In fact, the key-point of the procedure resides into the very sharp “native” image, in which shadows are particularly defined. Consequently, required
filtering operations are way less aggressive than it was needed in the previous cases. For this reason, this method has been selected to design the counting machine, as described in the next Session.

3. Rear projection-based counting machine prototype

Though the preliminary tests performed using the 7 item typologies shown in Figure 12a are deemed representative, a prototypal rear projection-based counting machine has been designed in order to perform extensive testing in an industrial environment. As shown in Figure 13, the system comprises:

- An image acquisition device (industrial monochrome camera IDS UI 3200-SE-M with a 6 mm lens with 12Mpix resolution (4104 x 3006 pixel).
- A CCD light projector to assure uniform lighting.
- A couple of orientable mirrors (used to extend the light path up to the 1.8m mentioned in section 3); such mirrors are used to reduce the overall dimensions of the counting machine, which must not exceed (1.5 x 1.0 x 1.0) m in order not to be excessively cumbersome for an industrial environment.
- An enclosure system to assure environmental light does not affect the scene.

The projector is placed backwards, on the frontal part of the machine. Light is reflected by the first mirror upwards towards the second one; this last reflects it forward to hit the galvanic frame. Its shadow is projected on the rear-projection canvas, which is arranged parallel to the frame.

![Figure 13. Counting system main components.](image-url)

In Figure 14a, a rendering of the designed counting machine architecture shows the arrangement of the above-mentioned components. The final design of the machine is in Figure 14b.
Figure 14. (a) Counting machine architecture; (b) Final design of the counting machine.

A tablet – on which runs the designed application (developed in Matlab®) – command the industrial camera. By means of a dedicated GUI (see Figure 15), the operator can check the position of the frame and can start the acquisition when such position is considered correct.

Figure 15. Dedicated GUI implemented for controlling the counting machine performance.

The procedure is then able, in approximately 0.3s, to print to screen the number of the detected items. Simultaneously, it is showed a control picture on which clusters that have been considered are coloured in red while the ones that have been ignored are white. In this way, the operator can rapidly check the effectiveness of the procedure and make corrections if needed.

5. Discussion and Conclusions

In this paper, a method to devise a method and a machine for counting the number of small metal parts randomly arranged on a galvanic frame is proposed. Knowing a-priori the area of each item that will be treated by galvanic bath, this makes it possible to estimate with a satisfying accuracy the overall area to be treated and, consequently, to optimize the settings of the treatment itself. Especially in high fashion field, in which precious materials are often used to realize plates, this allows to avoid material waste, thus leading to a significant money saving. Considering all the limitations that the application imposes (pieces already mounted on the frame, high reflectivity), many of the approaches usually adopted for counting machines (e.g. free fall, weight analysis) cannot be followed. The procedure is hence based on machine vision and makes use of rear-projection on a canvas to obtain a sufficiently sharp and easy to elaborate image with simple morphological operators. A counting machine, implementing the devised system has been designed.

This prototypical counting machine has been pre-tested with 20 different galvanic frames hosting 20 different kinds of objects with maximum dimension spanning from 10 to 80 mm. For all the test cases, the number of counted objects is exactly equal to the number of actual objects mounted on the frames. However, the system will undergo an extensive test campaign in an Italian Company working in the galvanic coating industry.

Accordingly, future works will be addressed to extensively test the devised procedure both in terms of performance (i.e. counted number of items vs. actual number, verified by visually inspecting the frames) and of usability.

Author Contributions: Conceptualization, L.P. and R.F.; methodology, L.G. and R.F.; software, M.S. and L.P.; validation, Y.V., L.P. and M.S.; data curation, M.S. and Y.V.; writing—original draft preparation, L.P.; writing—review and editing, L.G., R.F and Y.V.; supervision, L.G. and R.F.; project administration, L.G.

Funding: This work has been carried out thanks to the decisive regional contribution from the Regional Implementation Programme co-financed by the FAS (now FSC) and the contribution from the FAR funds made available by the MIUR.
Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript and in the decision to publish the results.

References