

# Towards the use of Augmented Reality techniques for Assisted Acceptance Sampling

Fiorenzo Franceschini<sup>1</sup>, Maurizio Galetto, Domenico Maisano, Luca Mastrogiacomo

<sup>1</sup>[fiorenzo.franceschini@polito.it](mailto:fiorenzo.franceschini@polito.it)

Politecnico di Torino, DIGEP (Department of Management and Production Engineering),  
Corso Duca degli Abruzzi 24, 10129, Torino (Italy)

## Abstract

*Acceptance Sampling* is a statistical procedure for accepting or rejecting production lots according to the result of a sample inspection. Formalizing the concept of *Assisted Acceptance Sampling* (AAS), this paper suggests the use of consolidated tools for reducing the risk of human errors in acceptance sampling activities. To this purpose, the application of Augmented Reality (AR) techniques may represent a profitable and sustainable solution.

An AR-based prototype system is described in detail and tested by an experimental plan. The major original contributions of this work are: (i) introducing the new paradigm of AAS, and (ii) developing a preliminary application in an industrial-like environment. This application is a first step towards the realization of a complete AAS system.

**Keywords:** lot-by-lot acceptance sampling, assisted acceptance sampling, augmented reality, quality control, on-line quality control.

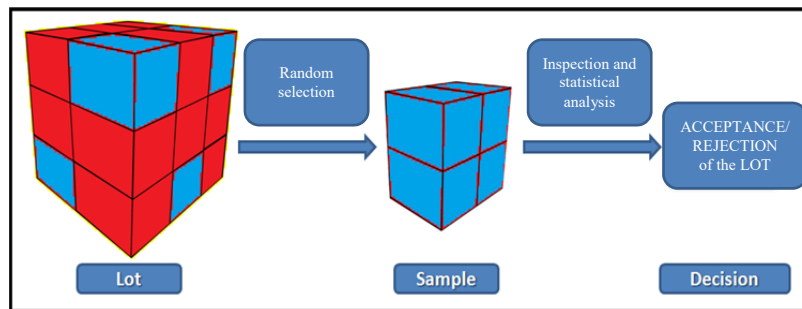
## 1. Introduction and literature review

Quality inspection procedures are generally used for checking the compliance with specification of incoming lots of products. When the lot size is very large and inspections are expensive or destructive, Acceptance Sampling (AS) is generally preferred to 100% inspection. AS uses statistical sampling to decide whether to accept or reject a production lot. This widespread and deep-rooted practice is regulated by several standards [1]. As derivatives of MIL STD 105E [2] and ANSI / ASQ Z1.4 [3], ISO 2859 and ISO 3951 standards [4] address the important role that AS plays when dealing with the product flow, with an emphasis on the producer's process [5].

AS generally requires the random selection of a sample of product units from a lot which should be representative of the lot itself (see Fig. 1). Then the lot is sentenced according to the result of the sample inspection: if the sample defectiveness is tolerable, then the whole lot is accepted, otherwise it is rejected [5].

To date, the human factor still has a crucial role in the AS, which often involves specialized operators dedicated to the inspection activities. In this context, distraction, fatigue, superficiality

and lack of operator's training are among the possible causes of errors that might compromise the effectiveness of AS.



**Fig. 1. Schematic representation of the AS process.**

Being able to superimpose virtual objects and cues upon the real world and in real time, Augmented Reality (AR) potentially constitutes a useful tool to assist operators in the random sampling process. There are many examples of ad-hoc AR systems designed for assisting operators in their specific tasks. Limiting the focus on manufacturing technologies, AR is applied to prototype visualization [6, 7, 8, 9], assistance in maintenance and repairing activities [10, 11, 12], training [13, 14], guidance during product assembly [11], robot path planning [15, 16].

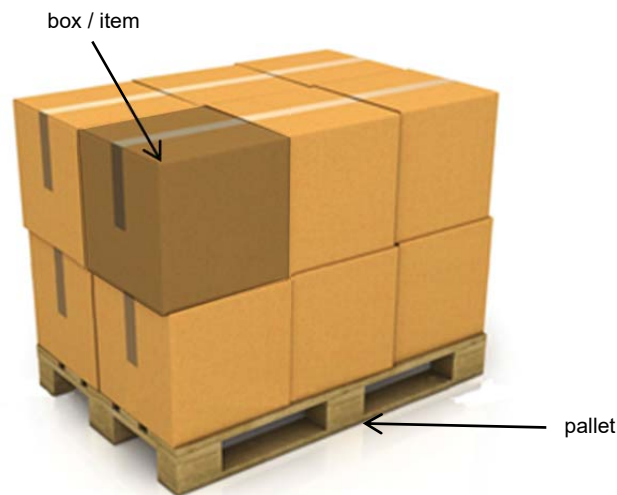
This manuscript introduces the concept of Assisted Acceptance Sampling (AAS), suggesting the use of AR techniques for supporting quality inspectors in the AS activities. This can significantly decrease the time for training operators, also reducing the risk of human error in the selection of the sample while ensuring the compliance with the requirement of sampling randomness. In detail, the paper describes a prototype implementation based on AR techniques which is able to recognize and track production lots in industrial environments. The ability to recognize and track production lots and simultaneously display them on a portable device is a first and fundamental step for any subsequent development of additional support features. As a preliminary implementation, the goal of the prototype is limited to support the user in the univocal identification of the item to sample in order to ensure the hypothesis of random sampling which is a fundamental assumption of AS. The prototype is tested through an experimental screening of the major factors affecting its performance. This paper is not aimed at developing new AR techniques or procedures, but rather extending the use of the available ones for typical AS operations. Moreover, the preliminary prototype application represents a first step towards the realization of a complete AAS system.

The remainder of this paper is organized into five sections. Sect. 2 discusses the problem of implementing an AR system for a lot by lot AS, while Sect. 3 describes the prototype developed in an industrial-like environment, at the industrial metrology and quality laboratories of DIGEP – Politecnico di Torino. Sect. 4 provides an application example of AAS based on the aforementioned prototype. Sect. 5 contains a screening analysis of the major factors affecting the performance of the

AR system. Finally, the concluding section highlights the main implications, limitations and original contributions of this work.

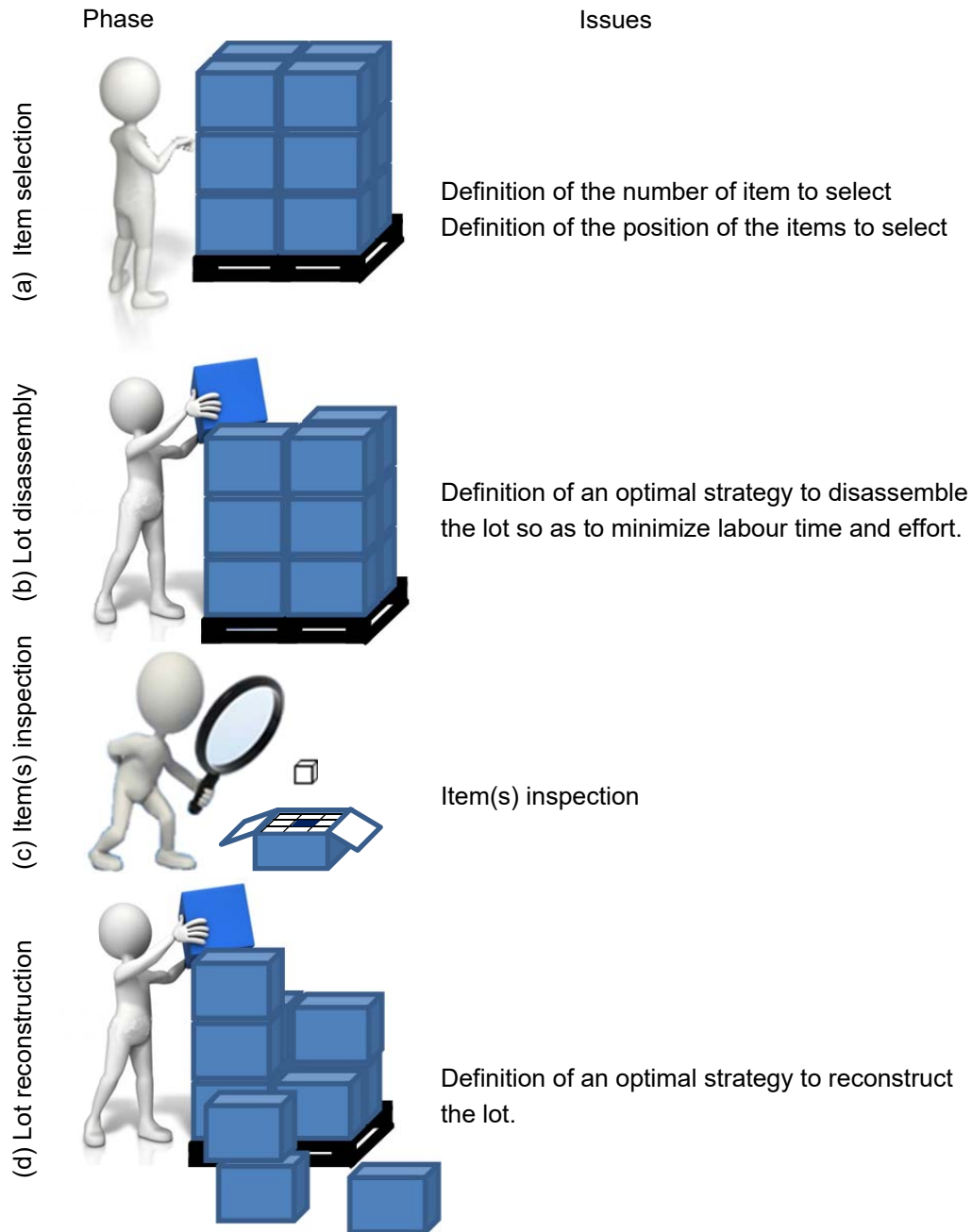
## 2. The concept of AAS: problem introduction

Even for companies with a high level of automation, AS activities are difficult to automate, since they are strongly affected by sensitivity, experience and training of operators. While automatic sampling and inspection systems have been developed and adopted in some specific contexts [17, 18, 19], manual approaches are still very common, with a crucial role played by operators. Unfortunately, the presence of an operator may also cause some errors stemming from distraction, fatigue, superficiality, lack of training, etc.



**Fig. 2. Lot-pallet configuration.**

In AS, a typical situation is that of inspecting a lot of items arranged on a pallet (see Fig. 2). The operator has to disassemble the lot, open the packages to be inspected and eventually reconstruct the lot after the inspection. For large-sized lots or lots with a large number of items, these operations can be quite complex. Fig. 3 depicts the typical activities and issues in AS activities.

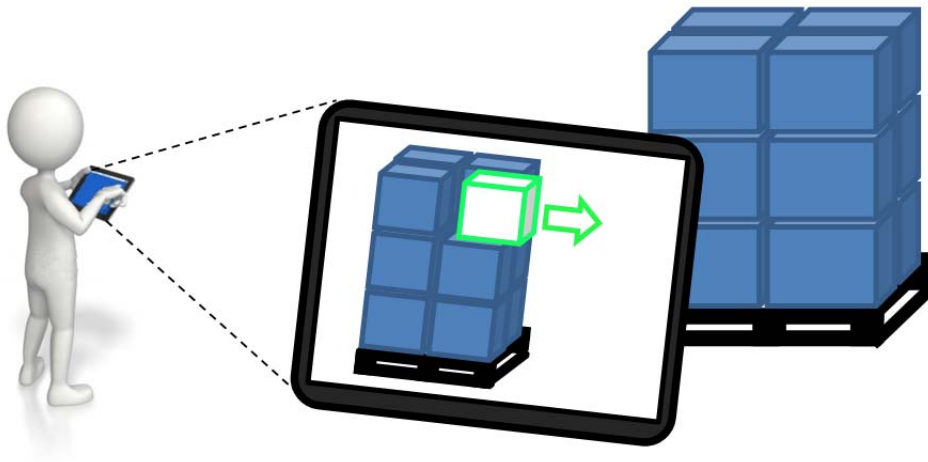


**Fig. 3. Typical phases and issues of non-assisted AS.**

Assisted AS is aimed at overcoming the limitations of the classical AS. In a AAS context, the operator is guided in his specific tasks providing real time information, such as for example the number and the location of the items to inspect and/or the optimal strategy to pick them. To date, specific tags (and readers) – such as Radio-frequency identification (RFID) tags or bar codes – have been used to identify, locate and track lots in warehouses or throughout production lines. Although this technology is mature and widely adopted, alone it is not able to fully assist AS procedures due to its intrinsic limitations [20, 21, 22].

AR techniques can be seen as a complementary tool to enhances AS systems, enriching the

operator's perception of the reality with additional information for guiding the sampling activities (see Fig.4). However, to be effective, they have to be minimally invasive, preferably inexpensive and not requiring particular hardware infrastructures.



**Fig. 4. Schematisation of the paradigm of AAS.**

In the following sections we describe the design and implementation of a first prototype developed according to the requirements of a specific application. Despite this specificity, the following discussion is deliberately general so as to apply to a generic context.

In detail, the prototype tries to solve the problem of tracking a generic lot-pallet, locating and identifying single the items, also guiding the operator in AS process selection.

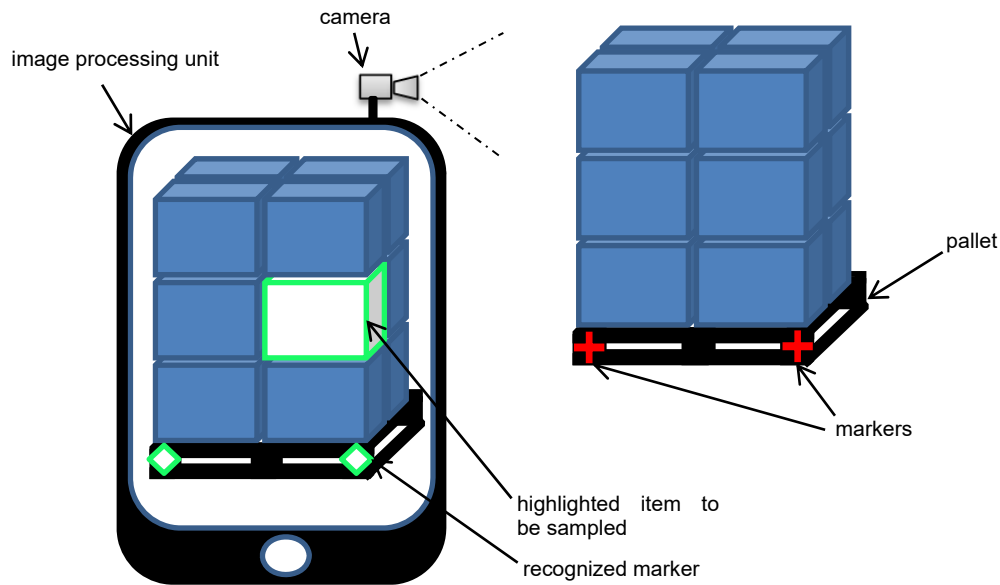
### **3. Prototype implementation**

As schematized in Fig. 5, the prototype was designed as composed by two elements:

- An image processing unit embedding an image acquisition device (i.e. a standard low cost camera);
- A pallet equipped with some markers, to be recognized by the image processing unit.

The idea was to realize a prototype able to manage the camera streaming so as to recognise a generic production lot, highlighting the items to be sampled.

C++ language was used as programming platform due to its large variety of graphical and image processing libraries. In detail, current implementation uses ARToolkit and OpenGL libraries [23, 24]. A tablet with an Intel® i3 processor and 4GB RAM was used as a preliminary processing unit. Two different cameras were used to test prototype performance (see Sect. 5).



**Fig. 5. Scheme of the prototype architecture.**

A simple selection algorithm able to suggest the boxes to sample was implemented. To make the implementation more efficient and robust, a set of markers placed on the base of the pallet was used (see Fig. 6). Being different from each other, markers make it possible to identify each side of the lot-pallet univocally. The choice of using artificial markers is not a relevant constraint in this kind of application: markers are simple labels which do not obstruct the normal functionality of pallets (for instance when handled by a forklift, pallet jack, front loader, etc.). Markers were simply glued onto the pallet in a simple and fast way.



**Fig. 6: Image of a lot with the detail of a marker.**

Fig. 7 synthetically shows the main steps of the implemented algorithm. Their logic is explained in the following sections by means of some graphical examples. Particular attention is paid to the third step (i.e. Transformation matrix estimation), which constitutes the core activity of the whole implementation.

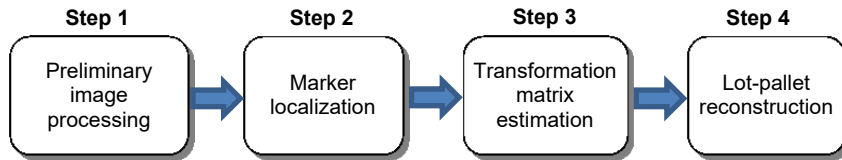


Fig. 7: Schematic flow chart of the preliminary tested procedure.

### 3.1 Step 1 – Preliminary image processing

This phase is aimed at capturing a frame/image from the camera and performing some preliminary image processing, i.e. the image is converted to black and white, improving the contrast for easing the subsequent processing phases. The optimal threshold value for the black and white conversion is automatically estimated by the analysis of the image features [25]. Fig. 8 shows an example of black and white conversion.

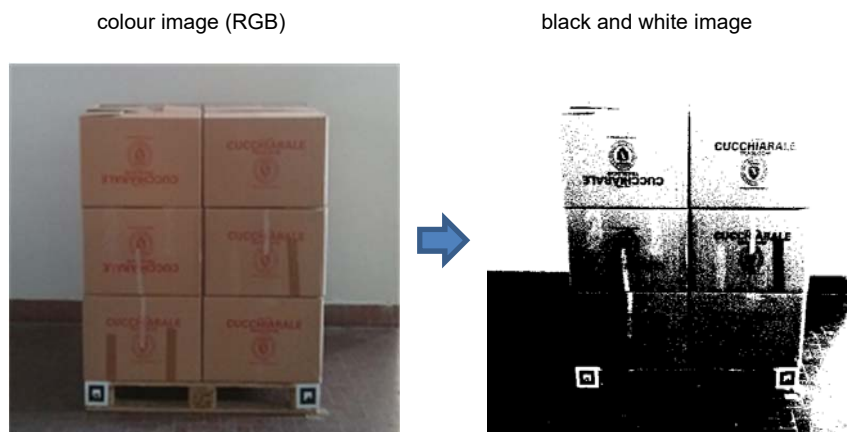


Fig. 8. Step 1 – Example of preliminary image processing.

### 3.2 Step 2 – Marker localization

The purpose of this step is to identify the position of the markers in the captured frame. This phase can be in turn divided into further steps, as described by Fig. 9.

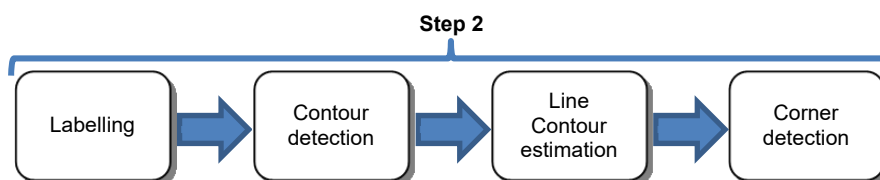
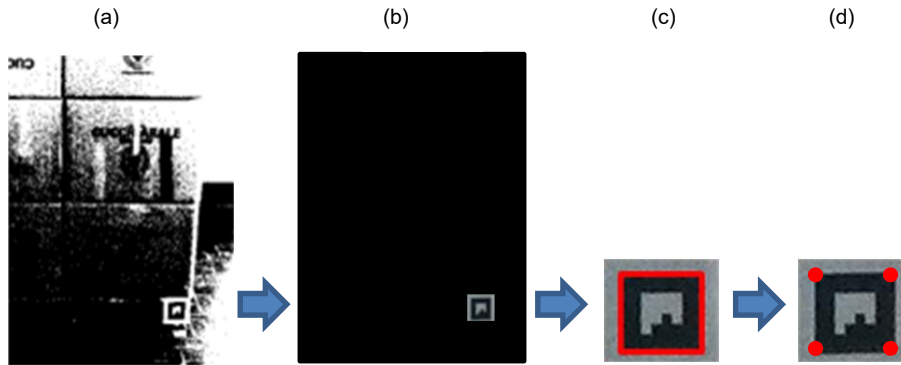


Fig. 9. Step 2 – Flow chart of the marker localization phase.

At this stage, the black and white image is processed for identifying and extracting the so-called “connected-component”, i.e. the regions of the image that are characterised by connected boundaries [26, 27]. The approach here selected is that proposed by Samet et al. [28] which proved

to be robust and efficient.

After identifying the connected boundaries in the analysed frame, the regions for which the outline contour can be fitted by four line segments are extracted. Rectilinear segments are intersected in order to identify the corner of each marker in the frame. Fig. 10 briefly summarizes these steps.



**Fig. 10. Step 2 - Schematic representation of the target segmentation phase. (a) is the original black and white image; (b) is the extracted connected component; (c) is the outline contour of the connected component while (d) is the result of the intersection of rectilinear segments.**

The final output of this phase is the set of 2D homogeneous coordinates  $(X_{k,s}, Y_{k,s})^T$  relating to the corners of each marker in camera screen reference system.

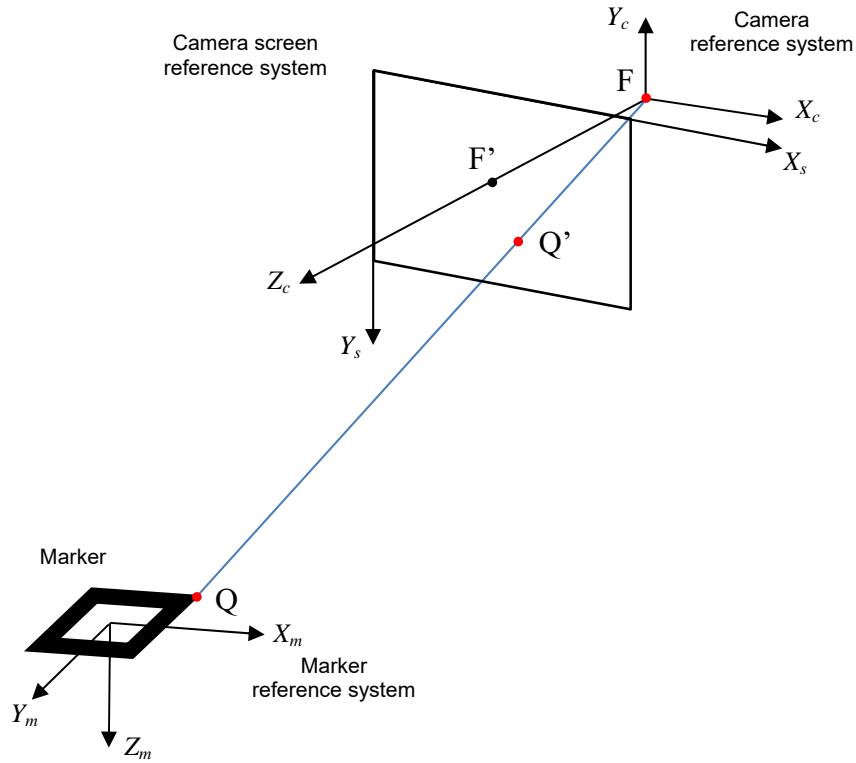
### **3.3 Step 3 – Transformation matrix estimation**

The purpose of this step is to determine the relative position of the image acquisition device with respect to the lot-pallet. This step requires the a-priori knowledge of (a) the relative positions of the markers on the pallet and (b) the technical specifications of the camera.

On the one hand, the position of the marker on the pallet can be defined by a preliminary measurement procedure that, defined a local reference system on the pallet (hereafter referred to as marker reference system), for each marker establishes the coordinates of its four corners.

On the other hand, the technical specifications of the camera are summarized by its internal parameters contained in the projection matrix  $\mathbf{P}$ . In the absence of technical accidents, this matrix can be estimated once in a while through appropriate calibration procedures. Further details concerning these procedures can be found in Appendix B.





**Fig. 11. Step 3 – Schematization of coordinate reference systems. F and Q are respectively the focal point of the camera and a generic corner of the marker. F' and Q' are the projections of F and Q on the camera view-plane.**

From the geometric point of view and focusing only on one marker, the problem is schematized in Fig. 11, where Q is the generic corner of a marker, Q' its projection on the camera screen and F the focal point of the camera. The aim of this phase is to estimate the transformation matrix  $\mathbf{W}$ , which links the coordinates of Q, in the marker reference system  $(X_{Q,m}, Y_{Q,m}, Z_{Q,m})^T$ , to the 2D coordinates of Q in the camera screen reference system  $(X_{Q,s}, Y_{Q,s})^T$ . This transformation can be expressed as a roto-translation in homogeneous coordinates as:

$$\begin{pmatrix} hX_{Q',s} \\ hY_{Q',s} \\ h \\ 1 \end{pmatrix} = \mathbf{P} \cdot \mathbf{W} \cdot \begin{pmatrix} X_{Q,m} \\ Y_{Q,m} \\ Z_{Q,m} \\ 1 \end{pmatrix}, \quad (1)$$

where  $(X_{Q,m}, Y_{Q,m}, Z_{Q,m})^T$  and  $\mathbf{P}$  are known values specified by the calibration procedure and  $(X_{Q,s}, Y_{Q,s})^T$  the result of Step 2 as described in Sect. 3.2. The inversion of Eq. 1, which is possible considering all the corners concurrently recognized by the system, allows the estimation of  $\mathbf{W}$ . Notice that the transformation matrix  $\mathbf{W}$  only depends on the relative position between the camera and the lot-pallet of interest. Since the camera and the lot-pallet may move,  $\mathbf{W}$  has to be estimated

in real time, i.e. for each of the frames captured.

For a more comprehensive description of this step refer to Appendix A.

### 3.4 Step 4 – Lot-pallet reconstruction

In this stage, the shape of the lot-pallet is reconstructed basing on a-priori known information. In a preliminary set-up phase, the coordinates of the lot corners are defined with respect to the markers' reference system. Say  $(X_{k,m}, Y_{k,m}, Z_{k,m})^T$  the coordinate vector of one of the corners of the lot, in the markers' reference system. Having defined  $\mathbf{P}$  and  $\mathbf{W}$ , the position of the corner in the camera screen  $(X_{k,s}, Y_{k,s})^T$  is given by:

$$\begin{pmatrix} hX_{k,s} \\ hY_{k,s} \\ h \\ 1 \end{pmatrix} = \mathbf{W} \cdot \mathbf{P} \cdot \begin{pmatrix} X_{k,m} \\ Y_{k,m} \\ Z_{k,m} \\ 1 \end{pmatrix}. \quad (2)$$

Eq. 2 allows to determine the camera screen position of all the corners. A virtual representation of lot edges can be superimposed over the real image by simply plotting all the lines connecting corners. Fig. 12 shows an example of lot reconstruction.



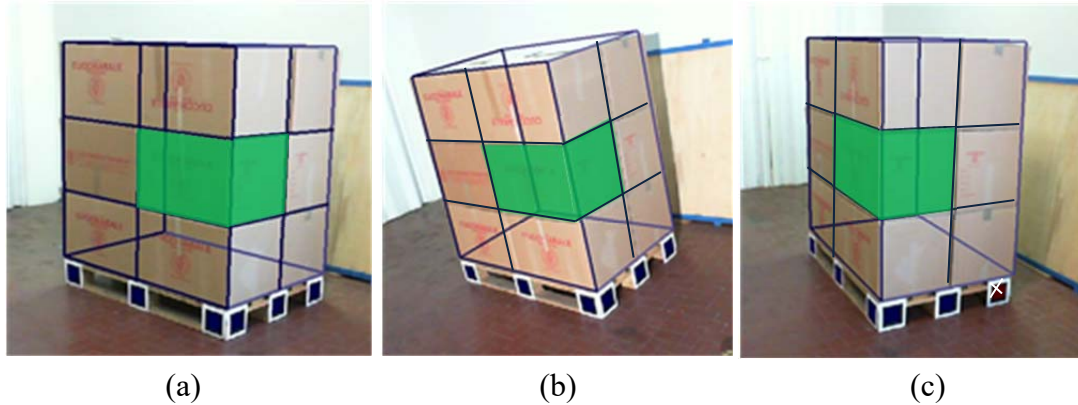
Fig. 12. Step 5 – Example of lot reconstruction.

The items composing the sample are highlighted in the graphical output in order to ease their identification and sampling.

## 4. Application Example

This section illustrate a few screenshots captured by the prototype, moving the camera around the

lot-pallet. In this example, twelve markers (three for each square of the pallet) were used to reconstruct a lot composed by 2 x 2 x 3 boxes. Markers highlighted in blue have been recognized by the system, while white crossed markers not (see Fig. 13c).



**Fig. 13. Qualitative example of system operation. White crossed squares represent markers that are not recognized by the system.**

As Fig. 13 shows, the prototype appears rather robust, even if lot reconstruction is not always perfect (see for instance the upper right corner of the lot in Fig. 13b). The accuracy of lot reconstruction is likely to be affected by several factors relating to the technology used, the environmental conditions as well as the algorithmic approach.

In order to provide a quantitative evidence of the prototype performance, next section provides some preliminary experimental results.

## 5. Factor Screening

Being based on image processing, the prototype performance can be affected by a large number of different factors.

In order to establish which of these factors are particularly significant, this section proposes a screening of five macro factors that were classified as potentially important for the system. The macro factors analysed may be further decomposed into other single factors. In detail, they are:

- *Number of recognized markers.* To reconstruct the lot-pallet, the prototype requires the recognition of at least one marker. However, with the increase of the number of markers it is reasonable to expect a better reconstruction (see Fig. 14).



Fig. 14. Example of images captured by the prototype with 2 and 3 recognized markers.

- *Light.* Lighting conditions can be crucial. An optimal lighting can improve the quality of the image so as to facilitate its processing (see Fig. 15).



Fig. 15. Example of images captured by the prototype in different light conditions.

- *Distance.* The distance between the image acquisition device and the lot compared to the size of the marker may influence the accuracy of the lot reconstruction (see Fig. 16).



Fig. 16. Example of images captured by the prototype at different distances from the lot.

- *Angle.* The angle between the camera and the lot (and thus markers) may complicate the

recognition of the markers (see Fig. 17).



Fig. 17. Example of images captured by the prototype at different angles.

- *Camera type.* Resolution, shutter time, quality of the lens, intrinsic parameters of the camera are all parameters influencing the quality of the image, and as a consequence the performance of the system (see Fig. 18).



Fig. 18. Example of images captured by two different acquisition devices.

### 5.1 Factors Screening

To test the aforementioned factors, a 5 factors complete factorial plan with 2 levels for each factor (i.e.  $2^5$  factorial design) and 3 replications was designed. In the experiments, the prototype was used to reconstruct a lot consisting of  $2 \times 2 \times 3$  boxes as the one exemplified in Fig. 13. The dimensions of each box are 60 x 40 x 40 cm, while that of the pallet are 120 x 80 x 15 cm.

Factors were varied in a total of  $2^5 \cdot 3 = 96$  combinations and the test sequence was randomized using the random number generator provided by Minitab®. For each of the 96 combinations, we recorded the image captured by the camera and the resulting coordinates of the corners ( $c_v$ ) of the reconstructed lot.

The considered response variable was the error in the reconstruction of the lot, intended as the

maximum pixel distance ( $d_{\max}$ ) between the corners of the real and the virtual lot (respectively  $c_r$  and  $c_v$ ), as reconstructed from the procedure:

$$d_{\max} = \max_{i=1,2,\dots,8} (\|c_r - c_v\|). \quad (3)$$

As a first application, the real lot corners were defined manually. Although this operation may introduce an error, its magnitude is considered negligible compared to that in the lot reconstruction. Each of the five factors was changed according to two levels. A detailed description is reported in the following:

1. *Number of recognized markers.* The prototype ideally can work when one marker is recognized. However preliminary tests showed that more robust results in terms of lot reconstruction can be obtained with two or more markers. For this reason 1 and 3 recognized markers were chosen respectively as “-1” and “1” level for this factor.
2. *Light.* The experiments were done in an indoor warehouse illuminated with artificial lights. A 1000-Watt halogen lamp was used. Two different configurations of illumination have been tested, respectively with the halogen lamp turned off and on. These two conditions of illumination correspond to approximately 250 lux and 720 lux in the area in which the pallet is placed.
3. *Distance.* This factor is intended as the distance between the camera and closest marker, the two levels are  $d_1 \approx 130$  cm and  $d_2 \approx 230$  cm. Considering that we used square markers of side  $l_m = 15$  cm, they respectively correspond to a ratio  $d_1 / l_m \approx 8.7$  and  $d_2 / l_m \approx 15.3$ .
4. *Angle.* Two angulations between camera and pallet were analysed: a frontal and a lateral position corresponding respectively to angles of about  $90^\circ$  and  $135^\circ$  between camera axis and the closest pallet side.
5. *Camera type.* Two different cameras were tested. For both of them the resolution was limited to 800x600 pixels in order to facilitate the real-time use of the prototype. The first one is the Microsoft LifeCam NX-6000 model 1082 while the second is the Microsoft LifeCam Cinema MSH5D-00015.

**Tab. 1. Summary of the parameters of the Design of Experiment (DoE).**

	Factors	Levels		Replications
		-1	1	
A	Number of recognized markers	1	3	3
B	Light	Off (250 lx)	On (720 lx)	3

C	Distance	$d_1 = 130$ cm	$d_2 = 230$ cm	3
D	Angle	Frontal ( $\alpha_1 \approx 90^\circ$ )	Lateral ( $\alpha_2 \approx 135^\circ$ )	3
E	Camera Type	LifeCam NX-6000	LifeCam Cinema MSH5D-00015	3

Table 1 summarises the parameters of the configuration of the experimental design.

## 5.2 Results Analysis

The response variable is modelled as:

$$d_{\max} = \beta_0 + \sum_{i=1}^5 \beta_i x_i + \sum_{i=1}^4 \sum_{j>i}^5 \beta_{i,j} x_i x_j + \sum_{i=1}^3 \sum_{j>i}^4 \sum_{k>j}^5 \beta_{i,j,k} x_i x_j x_k + \varepsilon, \quad (4)$$

where  $x_i$  with  $i \in \{1, 2, \dots, 5\}$  are the values of the  $i$ -th factor. For each of the 96 considered combinations of the factors a value of  $d_{\max}$  is produced. Neglecting the residual, the response variable can be modelled as

$$\hat{d}_{\max} = \hat{\beta}_0 + \sum_{i=1}^5 \hat{\beta}_i x_i + \sum_{i=1}^4 \sum_{j>i}^5 \hat{\beta}_{i,j} x_i x_j + \sum_{i=1}^3 \sum_{j>i}^4 \sum_{k>j}^5 \hat{\beta}_{i,j,k} x_i x_j x_k. \quad (5)$$

According to this notation residuals are given by:

$$\varepsilon = d_{\max} - \hat{d}_{\max}. \quad (6)$$

Considering the analysed combinations, a total of 96 residuals are generated. Table 2 details the output of the factorial fit of  $\hat{d}_{\max}$  versus all the factors and the considered interactions. Large values of T and conversely small P-values (say  $<0.05$ ) indicate that the factors (or interactions) have a statistically significant effect on the response.

**Tab. 2. Factorial Fit:  $\hat{d}_{\max}$  versus Markers; Light; Distance; Angle; Camera.**

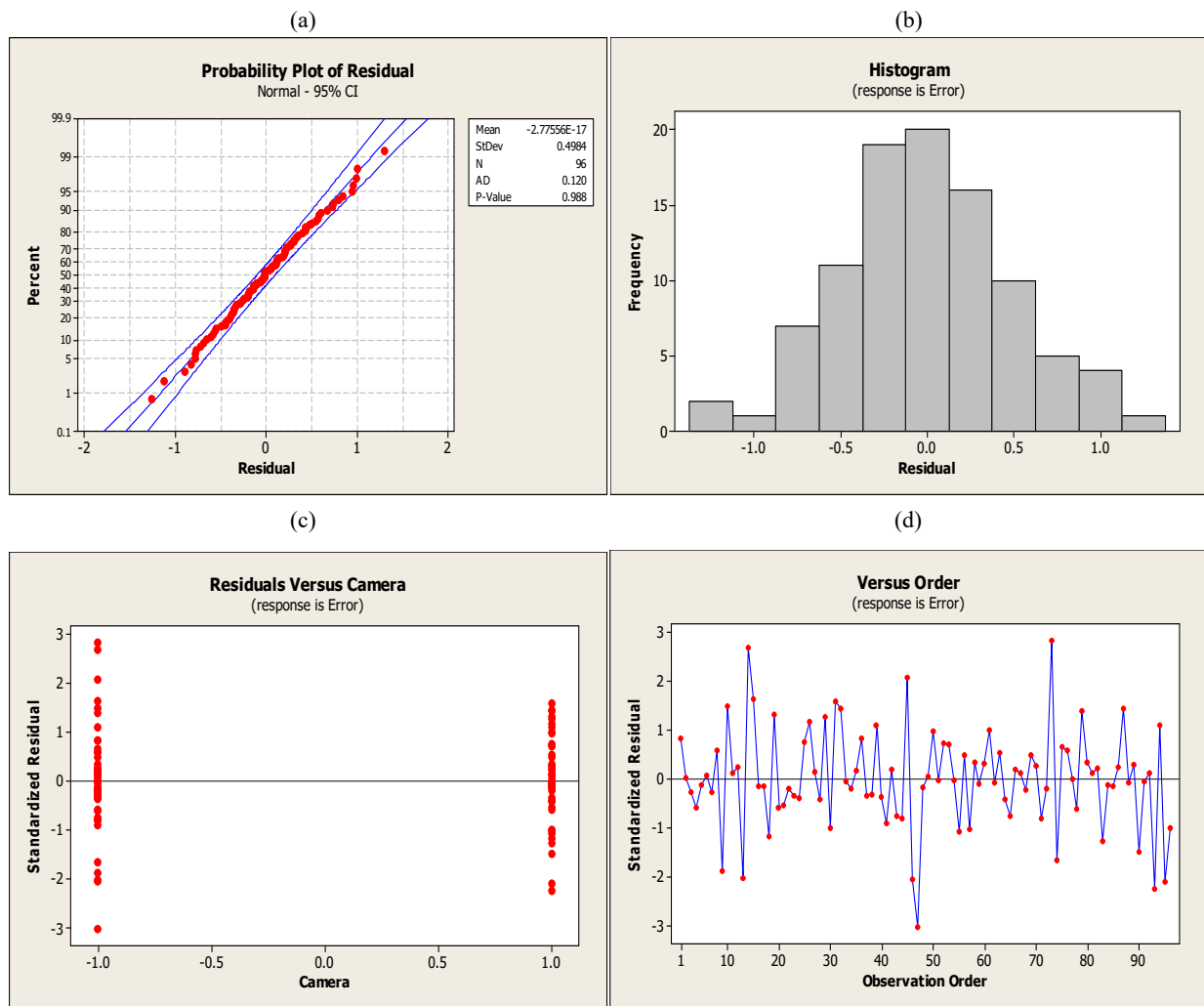
Estimated Effects and Coefficients for Error (coded units)					
Term	Effect	Coef	T	P	
Constant		$\hat{\beta}_0 =$	3.8020	64.16	0.000
Markers ( $x_1$ )	-0.9203	$\hat{\beta}_1 =$	-0.4602	-7.77	0.000
Light ( $x_2$ )	0.9240	$\hat{\beta}_2 =$	0.4620	7.80	0.000
Distance ( $x_3$ )	-0.7643	$\hat{\beta}_3 =$	-0.3822	-6.45	0.000
Angle ( $x_4$ )	0.6595	$\hat{\beta}_4 =$	0.3298	5.56	0.000
Camera ( $x_5$ )	0.6609	$\hat{\beta}_5 =$	0.3305	5.58	0.000
Markers*Light	0.0773	$\hat{\beta}_{1,2} =$	0.0386	0.65	0.517
Markers*Distance	-0.0982	$\hat{\beta}_{1,3} =$	-0.0491	-0.83	0.410
Markers*Angle	0.1169	$\hat{\beta}_{1,4} =$	0.0584	0.99	0.327
Markers*Camera	0.0524	$\hat{\beta}_{1,5} =$	0.0262	0.44	0.660

Light*Distance	0.0338	$\hat{\beta}_{2,3} =$	0.0169	0.29	0.776
Light*Angle	0.1301	$\hat{\beta}_{2,4} =$	0.0650	1.10	0.276
Light*Camera	-0.5626	$\hat{\beta}_{2,5} =$	-0.2813	-4.75	0.000
Distance*Angle	-0.1067	$\hat{\beta}_{3,4} =$	-0.0534	-0.90	0.371
Distance*Camera	-0.2333	$\hat{\beta}_{3,5} =$	-0.1166	-1.97	0.053
Angle*Camera	-0.0828	$\hat{\beta}_{4,5} =$	-0.0414	-0.70	0.487
Markers*Light*Distance	-0.1010	$\hat{\beta}_{1,2,3} =$	-0.0505	-0.85	0.397
Markers*Light*Angle	0.3397	$\hat{\beta}_{1,2,4} =$	0.1699	2.87	0.005
Markers*Light*Camera	0.1270	$\hat{\beta}_{1,2,5} =$	0.0635	1.07	0.288
Markers*Distance*Angle	-0.1191	$\hat{\beta}_{1,3,4} =$	-0.0595	-1.00	0.319
Markers*Distance*Camera	-0.1899	$\hat{\beta}_{1,3,5} =$	-0.0950	-1.60	0.114
Markers*Angle*Camera	0.4502	$\hat{\beta}_{1,4,5} =$	0.2251	3.80	0.000
Light*Distance*Angle	-0.2956	$\hat{\beta}_{2,3,4} =$	-0.1478	-2.49	0.015
Light*Distance*Camera	0.1092	$\hat{\beta}_{2,3,5} =$	0.0546	0.92	0.360
Light*Angle*Camera	-0.0583	$\hat{\beta}_{2,4,5} =$	-0.0292	-0.49	0.624
Distance*Angle*Camera	0.0699	$\hat{\beta}_{3,4,5} =$	0.0349	0.59	0.557
<hr/>					
$R^2 = 80.65\%$		$R\text{-Sq}(\text{adj}) = 73.74\%$			

The high value of  $R^2$  (>80%) indicates the goodness of fit of the model. This consideration is also supported by the value of the adjusted  $R^2$  (~74%) which gives the percentage of variation explained by only those factors (and interactions) that really affect  $\hat{d}_{\max}$  [29]. Hence, the gap between  $R^2$  and the adjusted  $R^2$  (~6%) can be explained by looking the significance of the factors considered for the factorial fit: the value of  $R^2$  is slightly inflated by the presence of few factors – say those with a P-value lower than 0.05 – that are not significant.

Also, the analysis in Fig. 19a shows the normal distribution of the residuals which is further confirmed by an Anderson-Darling test. The homogeneity of residual distribution was tested with respect to all the analyzed factors. As an example, Fig. 19c reports a plot of  $\hat{d}_{\max}$  versus the camera factor. Furthermore, there is no evidence of lack of independence. For instance, Fig. 19d shows the plot of residuals versus order of observation.





**Fig. 19. Residual analysis. (a) Normal Probability Plot, (b) Histogram of Residuals, (c) Residual versus camera and (d) Residual versus order plot. Residual are expressed in pixels.**

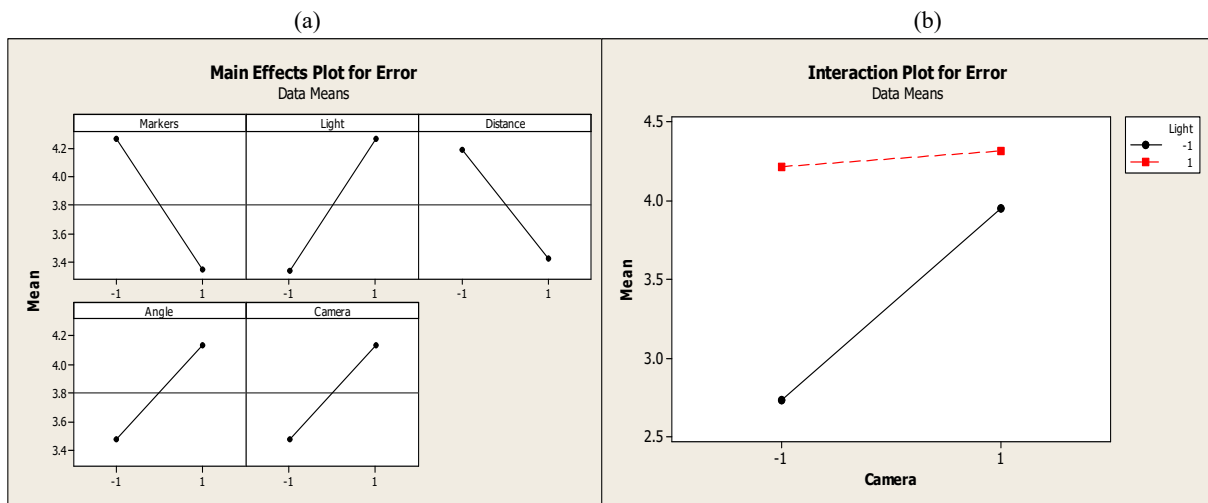
The factorial fit shows which factors and interactions are significant. These results are further confirmed by the main effects plot and the interaction plot for the response variable (see Fig. 20).

In order of importance, light is the most significant factor. Surprisingly, excessive lighting of the scene can significantly worsen the performance of the prototype. This is probably due to the fact that the recognition of a marker is done by isolating a sub-image based on the white frame surrounding each marker (see Sect. 4.2). Increasing the lighting may generate an overexposed image with many “white” parts which may complicate the recognition of the markers. In this case, it would probably be reasonable to expect an optimal value of illumination beyond which the prototype performance worsens.

As expected, the increase in the number of markers improves the performance of the prototype. Obviously, the recognition of a single marker allows the reconstruction of the lot with a lower accuracy with respect to the reconstruction obtained by the recognition of three markers.

Another important factor is distance. In general, the increase of the distance between camera and

pallet results in a reduction of the reconstruction error. This is because the difference between the virtual reconstruction of the lot and the real lot is more evident when the camera is closer to the lot, and the error, in terms of pixels, is larger. Even in this case, the behaviour of the response variable is probably not linear: while increasing the distance may have a positive effect, on the other hand, it is also reasonable to expect a degradation of performance from large distances, when the markers on the pallet are not easily recognizable.



**Fig. 20. (a) Main Effect plot for the mean of the DoE response variable. (b) Interaction plot for the mean of the response variable versus Camera and Light factors**

Another factor that has a significant effect is the camera type. The quality of the components and the different firmware settings of the analysed devices have a significant effect on the prototype. In this case the second camera is the best.

The angle causes an increase of the error for images from angled perspectives. It must be noted that in these tests it was decided to use only markers on one side of the pallet. In normal use, when the camera is angled with respect to the lot-pallet, the prototype is able to see up to six markers. In order to keep under control the number of recognized markers, nine of the markers on the pallet (i.e. all markers apart from those on one side) were hidden.

Fig. 20(b) shows the interaction plot for the response variable versus Camera and Light factors. The interaction between Light and Camera is noteworthy. This can be explained by the camera settings which causing different reactions to light exposure.

Three-way interactions Marker-Angle-Camera and Marker-Light-Angle are less important but significant. It is hard to give a practical explanation of the significance of such interactions: as for the Marker-Angle-Camera interaction, we can think that the different types of camera - which have different lenses and settings - can react differently to shots of different markers more or less angled; considering the Marker-Light-Angle interaction, one can imagine that depending on the angle and the marker, there are lighting conditions that allow a better reconstruction of the lot.

Being a screening, this analysis is far from being exhaustive. In particular, the behaviour of some factors, such as light or camera type, has still to be deepened. However, we remark that the maximum reconstruction error is in the order of a few pixels, i.e. a tolerable value for practical applications.

## 6. Conclusions

This paper introduced the concept of Assisted Acceptance Sampling (AAS), which entails conceiving and developing real-time tools for driving the AS operators in manual sampling operations, while reducing the risk of errors due to distraction, fatigue, lack of training, etc..

Preliminary results, concerning the implementation of a prototype, able to recognize and track a lot arranged on a pallet labelled with special markers, were presented. An experimental screening showed that the most significant factors affecting the performance of the prototype are lighting conditions, the number of markers used, the position with respect to the pallet and the type of camera used.

The good performance of the prototype implementation corroborates the fact that the proposed tool, if properly used by AS operators, may lead to remove human errors concerning the non-random selection of the sample units. A rigorous quantification of this kind of improvement, in real industrial environments, is left to further analysis.

In its current state, the main limitations of the prototype can be summarized as follows:

- The system is able to recognize and track a generic lot, but a procedure for guiding lot disassembly and re-assembly remains yet to be fully developed.
- To date, the prototype is only able to handle images in which at least one marker of the pallet is recognizable.

To be truly applicable, future developments of the prototype need to address and overcome these issues. Efforts in this direction would allow the application of innovative ways of performing lot-by-lot sampling in industrial environments, opening the possibility of reaching new levels of operational efficiency.

## References

1. Franceschini, F., M. Galetto, D. Maisano and L. Mastrogiacomo (2010). "Clustering of European countries based on ISO 9000 certification diffusion." International Journal of Quality & Reliability Management **27**(5): 558-575.
2. US Department of Defence (1989). MIL-STD-105E, Military Standard-Sampling Procedures and Tables for Inspection by Attributes, US Department of Defence, Arlington, VA

3. American Society for Quality (2008). ANSI/ASQ Z1.4-2008: Sampling Procedures and Tables for Inspection by Attributes
4. ISO (2002). ISO 2859 - Sampling procedures for inspection by attributes. International Organisation for Standardisation, Geneva
5. Schilling, E. G. and D. V. Neubauer (2010). Acceptance sampling in quality control, Boca Raton, FL, Chapman and Hall/CRC.
6. Azuma, R., Y. Baillet, R. Behringer, S. Feiner, S. Julier and B. MacIntyre (2001). "Recent advances in augmented reality." Computer Graphics and Applications, IEEE **21**(6): 34-47.
7. Friedrich, W., D. Jahn and L. Schmidt (2002). ARVIKA-augmented reality for development, production and service. Proceedings of the IEEE/ACM International Symposium on Mixed and Augmented Reality ISMAR 2002, 3-4, Citeseer.
8. Regenbrecht, H., G. Baratoff and W. Wilke (2005). "Augmented reality projects in the automotive and aerospace industries." Computer Graphics and Applications, IEEE **25**(6): 48-56.
9. Prieto, P. A., F. D. Soto, M. D. Zúñiga, S. F. Qin and D. K. Wright (2012). "Three-dimensional immersive mixed-reality interface for structural design." Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture **226**(5): 955-958.
10. Schwald, B. and B. De Laval (2003). "An augmented reality system for training and assistance to maintenance in the industrial context." Journal of WSCG **11**(1): 1-8.
11. Webel, S., U. Bockholt, T. Engelke, N. Gavish, M. Olbrich and C. Preusche (2012). "Augmented reality training platform for assembly and maintenance skills." Robotics and Autonomous Systems **61**(4): 398-403.
12. Zhu, J., S. Ong and A. Y. Nee (2012). "An authorable context-aware augmented reality system to assist the maintenance technicians." The International Journal of Advanced Manufacturing Technology **66**(9-12): 1699-1714.
13. De Crescenzo, F., M. Fantini, F. Persiani, L. Di Stefano, P. Azzari and S. Salti (2011). "Augmented reality for aircraft maintenance training and operations support." Computer Graphics and Applications, IEEE **31**(1): 96-101.
14. Gavish, N., T. Gutierrez, S. Webel, J. Rodriguez and F. Tecchia (2011). Design Guidelines for the Development of Virtual Reality and Augmented Reality Training Systems for Maintenance and Assembly Tasks. The International Conference SKILLS 2011.
15. Abbas, S. M., S. Hassan and J. Yun (2012). Augmented reality based teaching pendant for industrial robot. 12th International Conference on Control, Automation and Systems (ICCAS) 2210-2213, IEEE.
16. Fang, H., S. Ong and A. Nee (2012). "Robot Path and End-Effector Orientation Planning Using Augmented Reality." Procedia CIRP **3**: 191-196.
17. James, R. E. and D. Karner (1988). "Pallet inspection and repair system." U. S. Patents 4,743,154, 10 May 1988.
18. Ouellette, J. F. (1992). "Pallet inspection and stacking apparatus." U. S. Patent 5,096,369, Mar 17 1992.
19. Townsend, S. and M. D. Lucas (2007). "Automated digital inspection and associated methods." U. S. Patent US 2007/0163099 A1, Jul 19 2007.
20. Seidel, T. and R. Donner (2010). "Potentials and limitations of RFID applications in packaging industry: A case study." International Journal of Manufacturing Technology and Management **21**(3-4): 225-238.
21. Friedemann, S. and M. Schumann (2011). Potentials and limitations of RFID to reduce uncertainty in production planning with renewable resources, 17-26.
22. Expósito, I. and I. Cuiñas (2013). "Exploring the limitations on RFID technology in traceability systems at beverage factories." International Journal of Antennas and Propagation **2013**.

23. Kato, H. (2013). "ARToolkit." Retrieved 02/07/2013, from <http://www.hitl.washington.edu/artoolkit/>.
24. Silicon Graphics. (2013). "OpenGL." Retrieved 02/07/2013, from <http://www.opengl.org/>.
25. Otsu, N. (1975). "A threshold selection method from gray-level histograms." *Automatica* **11**(285-296): 23-27.
26. Bailey, D. G., C. T. Johnston and M. Ni (2008). Connected components analysis of streamed images. International Conference on Field Programmable Logic and Applications, 2008. FPL 2008., 679-682.
27. Klaiber, M., L. Rockstroh, W. Zhe, Y. Baroud and S. Simon (2012). A memory-efficient parallel single pass architecture for connected component labeling of streamed images. International Conference on Field-Programmable Technology (FPT), 159-165.
28. Samet, H. and M. Tamminen (1988). "Efficient Component labeling of images of arbitrary dimension represented by linear bintrees." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **10**(4): 579-586.
29. Draper, N. R. and H. Smith (2014). Applied regression analysis, John Wiley & Sons.
30. Montgomery, D. C. (2008). Design and analysis of experiments, Wiley, New York.
31. Luhmann, T., S. Robson, S. Kyle and I. Harley (2006). Close Range Photogrammetry Principles, Methods and Applications, Whittles, Scotland.
32. Franceschini, F., M. Galetto, D. Maisano, L. Mastrogiacomo and B. Pralio (2011). Distributed Large Scale Dimensional Metrology: New Insights, Springer.

## Appendix A – Transformation Matrix

This appendix provides further details about Step 3 of the procedure schematized in Fig. 9. The goal of this step is to determine the relative position of the image acquisition device with respect to the lot-pallet. With reference to the problem described by Fig. 11 and the notation introduced in Sect. 3.3, the aim of this phase is to estimate the transformation matrix  $\mathbf{W}$ , which links the coordinates of  $Q$ , in the local reference system of the markers  $(X_{Q,m}, Y_{Q,m}, Z_{Q,m})^T$ , to the coordinates of  $Q$  in the camera reference system  $(X_{Q,c}, Y_{Q,c}, Z_{Q,c})^T$ . This transformation can be expressed as a roto-translation in homogeneous coordinates as:

$$\begin{pmatrix} X_{Q,c} \\ Y_{Q,c} \\ Z_{Q,c} \\ 1 \end{pmatrix} = \begin{pmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_1 \\ r_{2,1} & r_{2,2} & r_{2,3} & t_2 \\ r_{3,1} & r_{3,2} & r_{3,3} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} X_{Q,m} \\ Y_{Q,m} \\ Z_{Q,m} \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0} & 1 \end{pmatrix} \cdot \begin{pmatrix} X_{Q,m} \\ Y_{Q,m} \\ Z_{Q,m} \\ 1 \end{pmatrix} = \mathbf{W} \cdot \begin{pmatrix} X_{Q,m} \\ Y_{Q,m} \\ Z_{Q,m} \\ 1 \end{pmatrix}. \quad (\text{A1})$$

The transformation matrix  $\mathbf{W}$ , which is composed of a rotation ( $\mathbf{R}$ ) and a translation ( $\mathbf{T}$ ) component, only depends on the relative position between the camera and the marker of interest. Since the camera and the lot-pallet may move,  $\mathbf{W}$  has to be estimated in real time, i.e. for each of the frames captured.

The projection matrix  $\mathbf{P}$ , i.e. the function linking the coordinates of  $Q$  in the camera reference system  $(X_{Q,c}, Y_{Q,c}, Z_{Q,c})^T$  to the coordinates of the same point (or its projection  $Q'$ ) in the camera screen reference system  $(X_{Q,s}, Y_{Q,s})^T$ , is assumed to be known:

$$\begin{pmatrix} hX_{Q',s} \\ hY_{Q',s} \\ h \\ 1 \end{pmatrix} = \begin{pmatrix} P_{1,1} & P_{1,2} & P_{1,3} & 0 \\ 0 & P_{2,2} & P_{2,3} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} X_{Q,c} \\ Y_{Q,c} \\ Z_{Q,c} \\ 1 \end{pmatrix} = \mathbf{P} \cdot \begin{pmatrix} X_{Q,c} \\ Y_{Q,c} \\ Z_{Q,c} \\ 1 \end{pmatrix}. \quad (\text{A2})$$

This matrix is a function only of the intrinsic camera parameters (focal length, principal point, scale factors), so it does not depend on the position of the image acquisition device.  $\mathbf{P}$  can be determined through appropriate calibration procedures [31]; Appendix B will give more details about them.

Combining Eqs. A2 and A1 we obtain:

$$\begin{pmatrix} hX_{Q',s} \\ hY_{Q',s} \\ h \\ 1 \end{pmatrix} = \mathbf{P} \cdot \mathbf{W} \cdot \begin{pmatrix} X_{Q,m} \\ Y_{Q,m} \\ Z_{Q,m} \\ 1 \end{pmatrix}. \quad (\text{A3})$$

Knowing the four pairs of coordinates of the marker corners in the camera screen reference system (i.e. the output of step 2), Eq. A3 can be reversed to find  $\mathbf{W}$ . Focusing on  $\mathbf{W}$ ,  $\mathbf{R}$  and  $\mathbf{T}$  are

estimated separately. In particular,  $\mathbf{R}$  is estimated basing on the following geometric considerations. When two parallel sides of a square marker are projected on the image, the equations of the line segments in the camera screen reference system are:

$$\begin{aligned} a_1 X_s + b_1 Y_s + c_1 &= 0 \\ a_2 X_s + b_2 Y_s + c_2 &= 0 \end{aligned} \Rightarrow \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} X_s \\ Y_s \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow \mathbf{A} \cdot \begin{pmatrix} X_s \\ Y_s \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (\text{A4})$$

being  $(X_s, Y_s)^T$  generic coordinates in camera screen reference system. By using Eq. A2, Eq. A4 can be reformulated as a function of generic coordinates in camera reference system (i.e.  $(X_c, Y_c, Z_c)^T$ ):

$$\mathbf{A} \cdot \begin{pmatrix} X_s \\ Y_s \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow \mathbf{A} \cdot \mathbf{P} \cdot \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (\text{A5})$$

In matrix form, Eq. A5 expresses the equation of two bundles of plans including the two parallel sides of the marker. Normal vectors of these planes  $\mathbf{n}_1$  and  $\mathbf{n}_2$  are:

$$\begin{aligned} \mathbf{n}_1 &= (a_1 P_{11} \quad a_1 P_{12} + b_1 P_{22} \quad a_1 P_{13} + b_1 P_{23} + c_1)^T \\ \mathbf{n}_2 &= (a_2 P_{11} \quad a_2 P_{12} + b_2 P_{22} \quad a_2 P_{13} + b_2 P_{23} + c_2)^T. \end{aligned} \quad (\text{A6})$$

Thus, the direction vector of two parallel sides of the marker can be obtained as the cross product  $\mathbf{n}_1 \times \mathbf{n}_2$ . Given the two sets of parallel sides of the marker, two nominally perpendicular unit direction vectors ( $\mathbf{u}_1$  and  $\mathbf{u}_2$ ) can be obtained. The cross product among  $\mathbf{u}_1$  and  $\mathbf{u}_2$  gives a third unit vector ( $\mathbf{u}_3$ ) so that the rotation component  $\mathbf{R}$  in the transformation matrix  $\mathbf{W}$  can be expressed as:

$$\mathbf{R} = (\mathbf{u}_1^T \quad \mathbf{u}_2^T \quad \mathbf{u}_3^T). \quad (\text{A7})$$

Given  $\mathbf{R}$ , the translation component of  $\mathbf{W}$  can be estimated by Eq. A1. In each frame, the coordinates of the four marker corners are known both in the marker and the camera screen reference system. Thus, a system of eight equations in three unknown parameters ( $\mathbf{T} = (t_1 \quad t_2 \quad t_3)^T$ ) is defined for every frame.

Ideally, the recognition of a single marker is sufficient for the recognition and reconstruction of the entire lot-pallet. However, for a robust recognition, up to twelve markers were used (i.e. three in

each side, see Fig. 12). Finally,  $\mathbf{W} = \begin{pmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0} & 1 \end{pmatrix}$  is estimated.

## Appendix B – Projection Matrix

The goal of camera calibration is finding the projection matrix ( $\mathbf{P}$ ) introduced in Eq. A2 [32]. The camera calibration is performed using a simple cardboard frame with a ruled grid of lines. The coordinates of all intersection points are a-priori known in the cardboard local 3D coordinates. The cardboard frame is captured by the camera from different points of view. Also the coordinates of intersection points in the camera screen reference system are identified by image processing.

Considering the generic intersection point  $I_i$  on the cardboard frame, the relationship among its camera screen coordinates  $(X_{I_i,s}, Y_{I_i,s})^T$ , the camera coordinates  $(X_{I_i,c}, Y_{I_i,c}, Z_{I_i,c})^T$  and cardboard coordinates  $(X_{I_i,m}, Y_{I_i,m}, Z_{I_i,m})^T$  can be modelled as:

$$\begin{pmatrix} hX_{I_i,s} \\ hY_{I_i,s} \\ h \\ 1 \end{pmatrix} = \mathbf{P} \cdot \begin{pmatrix} X_{I_i,c} \\ Y_{I_i,c} \\ Z_{I_i,c} \\ 1 \end{pmatrix} = \mathbf{P} \cdot \mathbf{W} \cdot \begin{pmatrix} X_{I_i,m} \\ Y_{I_i,m} \\ Z_{I_i,m} \\ 1 \end{pmatrix} = \mathbf{C} \cdot \begin{pmatrix} X_{I_i,m} \\ Y_{I_i,m} \\ Z_{I_i,m} \\ 1 \end{pmatrix} = \begin{pmatrix} C_{1,1} & C_{1,2} & C_{1,3} & C_{1,4} \\ C_{2,1} & C_{2,2} & C_{2,3} & C_{2,4} \\ C_{3,1} & C_{3,2} & C_{3,3} & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} X_{I_i,m} \\ Y_{I_i,m} \\ Z_{I_i,m} \\ 1 \end{pmatrix} \quad (\text{B1})$$

where  $\mathbf{P}$  is the projection matrix to be estimated. Since many pairs of  $(X_{I_i,s}, Y_{I_i,s})^T$  and  $(X_{I_i,m}, Y_{I_i,m}, Z_{I_i,m})^T$  are generally obtained, matrix  $\mathbf{C}$  can be estimated by a least square approach.

Combining Eqs. A2 and B1, it is obtained:

$$\begin{pmatrix} P_{1,1} & P_{1,2} & P_{1,3} & 0 \\ 0 & P_{2,2} & P_{2,3} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_1 \\ r_{2,1} & r_{2,2} & r_{2,3} & t_2 \\ r_{3,1} & r_{3,2} & r_{3,3} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} C_{1,1} & C_{1,2} & C_{1,3} & C_{1,4} \\ C_{2,1} & C_{2,2} & C_{2,3} & C_{2,4} \\ C_{3,1} & C_{3,2} & C_{3,3} & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{B2})$$

Notice that matrix  $\mathbf{C}$  has 11 independent variables, while  $\mathbf{P}$  and  $\mathbf{W}$  5 and 6 (three rotation angles and three translation components) respectively. As a result, the matrix  $\mathbf{C}$  can be easily decomposed into  $\mathbf{P}$  and  $\mathbf{W}$ , considering the upper triangular form of  $\mathbf{P}$ .