

# Active X-ray Testing of Complex Objects

*Vladimir Riffo and Domingo Mery*

Department of Computer Science – Pontificia Universidad Catolica de Chile

vriffo1@uc.cl, dmery@ing.puc.cl

## Abstract

X-ray testing of complex objects –such as luggage screening at airports– is usually performed manually. This is not always effective, since it depends strongly on the pose of the objects of interest (target objects), occlusion and human capabilities as well. Additionally, certain target objects are difficult to be detected using only one viewpoint. For this reason, we developed an active X-ray testing framework that is able to adequate the viewpoint of the target object in order to obtain better X-ray images to analyze. The key idea of our method is to adapt automatically the viewpoint of the X-ray images in order to project the target object in poses where the detection performance should be higher. Thus, the detection inside of complex objects can be performed in a more effective way. Using a robotic arm and a semi-automatic manipulator system, the robustness and reliability of the method have been verified in the automated detection of razor blades located inside of nine different objects showing promising preliminary results: in 130 experiments we were able to detect 115 times the razor blade with 10 false alarms, achieving recall of 89% and precision of 92%.

**Keywords:** X-ray inspection, baggage screening, framework, SIFT, active vision, target object detection.

## 1 Introduction

It is well known that X-ray inspection can be performed by human inspectors or automatic systems. Although humans in many cases can do the job better than machines, they are slower

and tire quickly. In addition, the human inspectors are not always consistent and effective to evaluate objects, because the inspection tasks are monotonous and exhausting, even for experts in the field. Moreover, human experts are difficult to find or maintain in an industry, require training and learning can take time.

The main areas of application of X-ray testing are:

- **Baggage screening:** Since 9/11 X-ray imaging has become an important issue. Threat items are more difficult to recognize when placed in close packed bags, when superimposed by other objects, and when rotated [1].
- **Foods:** In order to ensure food safety inspection, the following interesting applications have been developed: detection of foreign objects in packaged foods, detection of fishbones in fishes, identification of insect infestation, and fruit and grain quality inspection [2, 3].
- **Cargos:** With the ongoing development of international trade, cargo inspection becomes more important. X-ray testing has been used for the evaluation of the contents of cargo, trucks, containers, and passenger vehicles to detect the possible presence of many types of contraband [4].
- **Castings:** In order to ensure the safety of construction of certain automotive parts, it is necessary to check every part thoroughly using X-ray testing. Within these parts –considered important components for overall roadworthiness–, non-homogeneous regions like bubble-shaped voids or fractures, can be formed in the production process [5].
- **Weldings:** In welding process, a mandatory inspection using X-ray testing is required in order to detect defects (porosity, inclusion, lack of fusion or penetration and cracks). X-ray images of welds are widely used for the detecting those defects in the petroleum, chemical, nuclear, naval, aeronautics and civil construction industries [6].

We observe, there are some application areas, like castings inspection, where automated systems are very effective. Additionally, there are certain application areas, like weldings and cargos, where the inspection is semi-automatic; and there is some research in food science where food quality is beginning to be characterized using X-ray imaging. Finally, there are application areas, like baggage screening, where human inspection is still used, although the effectiveness is around 90% (in best-case) [7–9].

In order to aid human inspection tasks, baggage screening algorithms have been mainly developed to solve image segmentation problems using single views. Multiple view analysis, however, can be exploited in the inspection of complex objects where the detection is extremely difficult with only one viewpoint, because the target object may be occluded or its projection is intricate as shown in Fig. 1.

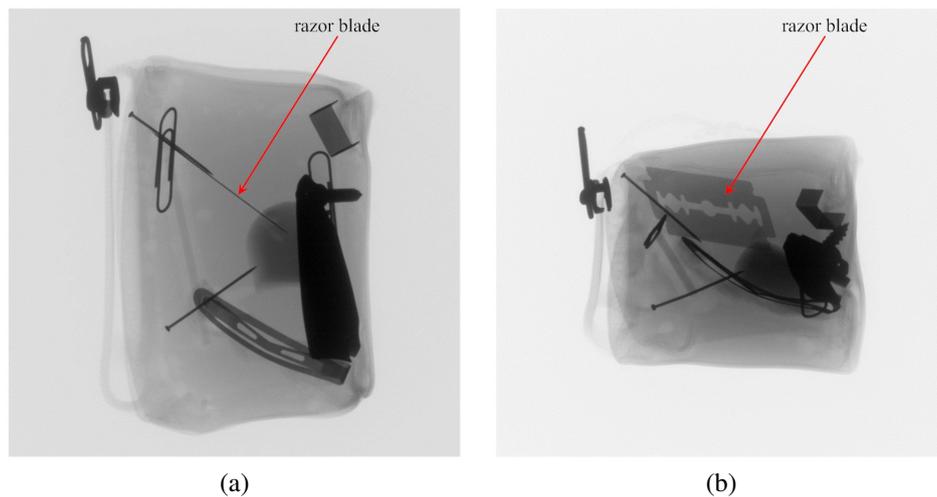


Figure 1: Two radioscopic views of the same inspection object (a pencil case), where the detection of the target object (a razor blade) is; a) impossible, and b) possible.

In this sense, multiple view analysis can be a powerful option for examining complex objects where uncertainty can lead to misinterpretation. Multiple view analysis offers advantages not only in 3D interpretation. Two or more views of the same object taken from different viewpoints can be used to confirm and improve the diagnostic done by analyzing only one image. In the last years, there are two main strategies developed for X-ray testing with multiple views: The first strategy is based on the tracking of potential objects of interest across an X-ray image sequence. The key idea can be summarized as follows: *i*) potential objects of interest are segmented in each view of the sequence using an application dependent method that analyzes 2D features in each single view ensuring the detection of the target object (not necessarily in all views) and allowing false detections, *ii*) the potential regions are tracked across the sequence based on similarity and geometrical constraints (according to multiple view geometry [10]) eliminating those potential regions that cannot be tracked, and *iii*) the tracked regions are analyzed including those views where the segmentation fails (the positions can be predicted by reprojection [10]). This strategy was developed for aluminum castings in [11] using an off-line calibration approach [12] for

the geometrical model, and an uncalibrated model [13]. Additionally, this strategy was used for other complex objects using a structure model obtained from the views [14]. The second strategy is based on a virtual 3D scanning of the target object. Despite existence techniques for linking information between different views, there are still relevant challenges to solve problems such as data association, correlation establishment, and occlusions. For this reason, in [15], the classical 2D sliding-window scheme (see for example an application in inspection of welds [16]) was changed to a *3D sliding-box* approach. The key idea of this approach is to analyze –for each 3D position of a sliding-box– only the portion of the images where the sliding-box is projected. Again, the strategy was tested in aluminum castings.

We observe in these works that the automated analysis in multiple views can achieve high detection rates because there is no substantial overlap effect, however, there is an initial requirement of a considerable high number of images, many of them are certainly unnecessary. Additionally, in all cases the estimation of a complex geometrical model is required.

For these reasons, we developed an X-ray testing framework for the active multiple view inspection of complex objects that avoids the mentioned disadvantages. The key idea of our method is to adapt automatically the viewpoint of the X-ray images in order to project the target object in poses where the detection performance should be higher. Thus, analyzing a first view, we can automatically search a second, third and even fourth view, to confirm and improve the analysis of only one image. In our proposal, we incorporate some ideas of *best view* and *view planning* where there are a variety of published works in the areas of recognition and 3D reconstruction (see for example [17, 18]), however, to best our knowledge there are no publications on X-ray inspection using these ideas.

In this paper we present: the proposed approach (Section 2), the results obtained in several experiments (Section 3), and some concluding remarks and suggestions for future research (Section 4).

## 2 Proposed Approach

In our approach, a *target object*, *i.e.*, the object of interest to be detected, may be located inside of a *container object*. Usually, there are many objects inside of a container object, and only

a few of them –if any– are target objects. For this reason, we say that the whole object to be inspected is a *complex inspection object*. In order to test our framework, we used –without loss of generality– *razor blades* (target objects) inside of different cases (container objects), as illustrated in Fig. 1. This inspection object is complex enough to perform our experiments. Additionally, a razor blade has the advantage of being an object with interesting characteristics to address: thinness, low X-ray absorption coefficient, symmetry in all quadrants, easy portability and dangerous (in baggage screening at airports for example).

## 2.1 General framework

The general framework attempts to find a *good view* of the inspection object, *i.e.*, an image in which a target object should be viewed from a *good pose* that ensures its detection. The good poses of the target object correspond to those ones from them the acquired view should have a high probability of detection. In our experiments, the good poses of a razor blade correspond to the frontal views. Thus, the key idea is to rotate and/or translate the inspection object from an initial position to a new one in which the detection probability of the target object should be higher. Clearly, if the initial position corresponds to a good view, no more positions will be required, in these cases the inspection is performed with only one X-ray image avoiding the analysis of more images.

The proposed algorithm consists of two parts (A and B), as is illustrated in Fig. 2: In part A, an arbitrary initial position of the inspection object is chosen and an X-ray image is acquired. In the detection step, a target object is searched, and if a potential target object is detected, its pose is estimated. If the estimated pose corresponds to one of the good poses then the inspection is finished detecting a target object, else the inspection object should be moved –using the estimated pose– so that one of the good poses of the target object is expected. This process is interrupted after  $C_{max}$  times in order to avoid looping. On the other hand, in part B, if no potential target object is found –in the detection step– then the object is arbitrarily moved to a new position very different from the first one, repeating the detection step. The last situation can repeated until  $D_{max}$  times in order to ensure the inspection from all relevant viewpoints. In our experiments,  $C_{max} = 4$  and  $D_{max} = 3$ .

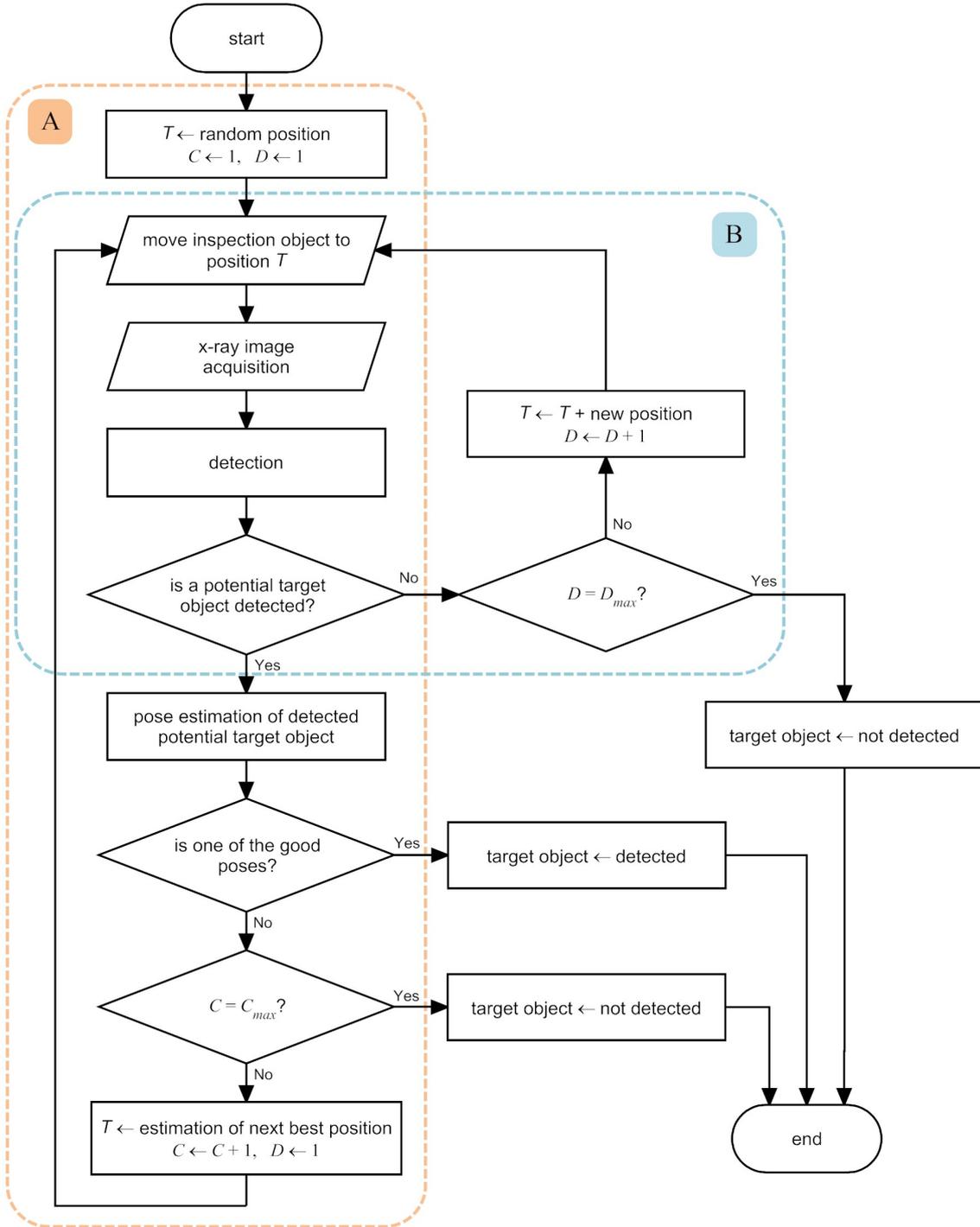


Figure 2: Framework for active inspection with X-rays.

This proposed framework is general and easily adaptable, however, the designed algorithms for detection and pose estimation are application dependent. As mentioned above, in our experiments we use a razor blade as target object. For this reason, the details of our method explained below are for this particular application.

## 2.2 Detection of potential target objects

An application dependent algorithm must be defined to detect automatically potential target objects in a single test image. The objects detected in this step are called *potential target objects*. As mentioned above, in order to test our framework, we developed an algorithm that is able to detect razor blades. In this section, the algorithm will be explained in further details.

The algorithm is based on a well known technique used by the computer vision community called SIFT (*Scale Invariant Feature Transform*) [19]. SIFT is able to detect and extract –in certain *keypoints* of an image– local feature descriptors that are very robust against noise and changes in scale, rotation, viewpoint and contrast. A SIFT descriptor is typically a vector with 128 elements computed from 16 histograms of gradients in 8 directions in a neighborhood of the keypoint (see more details in [19]). Thus, the matching between keypoints of two different images of the same object is efficiently performed by comparing their SIFT descriptors, *i.e.*, by finding the minimal Euclidean distance between the descriptors.

In our approach, we use a SIFT description of the target object in all feasible poses by rotating two axes in nine steps as shown in Fig. 3. All extracted descriptors are stored in an arrange  $\mathbf{P}$ , where  $\mathbf{p}_j$  means the  $j$ -th descriptor, for  $j = 1 \dots m$ . Each descriptor  $\mathbf{p}_j$  has a corresponding pose  $r_j$ . In our example,  $r_j \in [1, 81]$  for  $9 \times 9$  poses.

In the detection step, all SIFT descriptors of the test image of the inspection object are extracted and stored in an arrange  $\mathbf{Q}$ , where  $\mathbf{q}_i$  means the  $i$ -th descriptor of the test image for  $i = 1 \dots n$ . Now, all duplets  $(\mathbf{q}_i, \mathbf{p}_j)$  that fulfill the condition  $\|\mathbf{q}_i - \mathbf{p}_j\| < \theta_E$  for  $i = 1 \dots n$  and  $j = 1 \dots m$  are selected, where  $\theta_E$  is a minimum distance threshold, and  $\|\mathbf{q}_i - \mathbf{p}_j\|$  means the Euclidean distance between both vectors. Afterwards, for each selected descriptors the corresponding pose  $r_j$  is obtained. The selected descriptors and their corresponding poses will be stored in  $\mathcal{Q}$  and  $\mathcal{R}$  respectively. Thus, we have *i)*  $\mathcal{Q}$ : all keypoints of the test image that have been matched with keypoints of the target object, and *ii)*  $\mathcal{R}$ : the corresponding poses for the selected keypoints  $\mathcal{Q}$ . The detection is performed in the following two steps: *i)* Clustering: in  $\mathcal{Q}$ , we find all keypoints of the same pose that are close to each other in the test image. Thus, we define subwindows  $\mathcal{W}_B$  that have at least  $\theta_B$  keypoints of the same pose (see blue squares in 4b). In our experiments, we set the size of  $\mathcal{W}_B$  equal to  $80 \times 80$  pixels, and  $\theta_B = 3$ . *ii)* Merging: all subwindows  $\mathcal{W}_B$  that

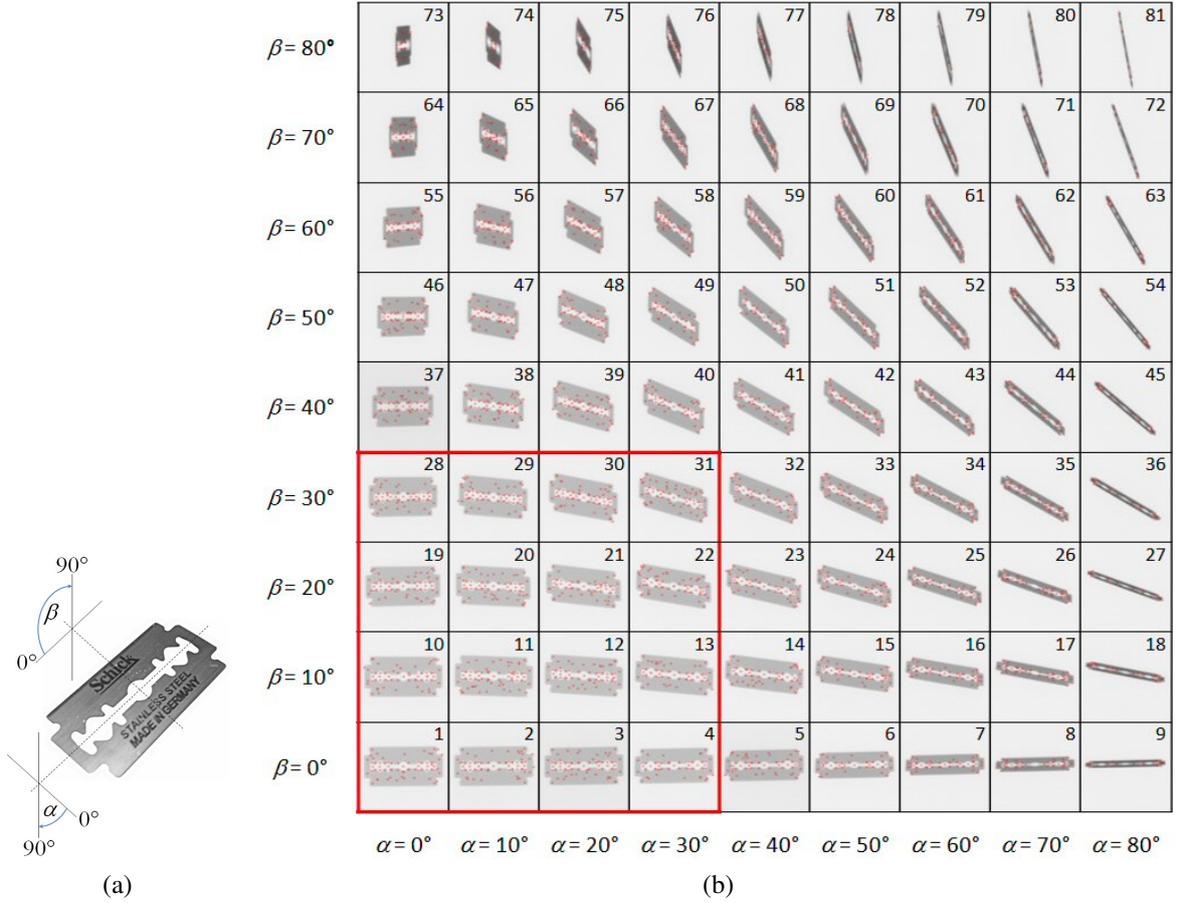


Figure 3: Description of the target object: the razor blade is located and irradiated in 81 different poses ( $\alpha = 0, 10, \dots, 80^\circ$ , and  $\beta = 0, 10, \dots, 80^\circ$ ), from their X-ray images SIFT descriptors are extracted (see red keypoints). The good poses are the frontal views defined by  $0 \leq \alpha \leq 30$  and  $0 \leq \beta \leq 30$ , *i.e.*, the poses 1, 2, 3, 4, 10, 11, 12, 13, 19, 20, 21, 22, 28, 29, 30 and 31 (see large red square at the bottom left).

are connected or overlapped, will be merged in a new larger subwindow  $\mathcal{W}_G$  (see green squares in Fig. 4b and 4c). The subwindow that encloses the highest number of keypoints of the same pose will be selected if this number is equal or greater than  $\theta_G$ , ensuring at least  $\theta_G$  descriptors of the same pose in the selected window (see red square in Fig. 4d). In our experiments, we set  $\theta_G = 8$ . The selected subwindow will be called  $\mathcal{W}_S$  and it corresponds to a potential target object. If no subwindow fulfills this condition, then no potential target object is detected.

### 2.3 Pose estimation

Clearly, the estimated pose of the detected potential target object will be the corresponding pose of the selected subwindow  $\mathcal{W}_S$  (see Section 2.2 and label in Fig. 4d). If this pose is one of the *good poses* then the algorithm will be finished detecting a target object. In our example,

the set of good poses corresponds to frontal views of the razor blade as shown in Fig. 3. If the estimated pose does not belong to the set of good poses, or no potential target object is detected, then the inspection object will be moved to a new position attempting to get a good pose of the target object. The estimation of the *next best position*, called *next best view* in active vision, will be explained in following.

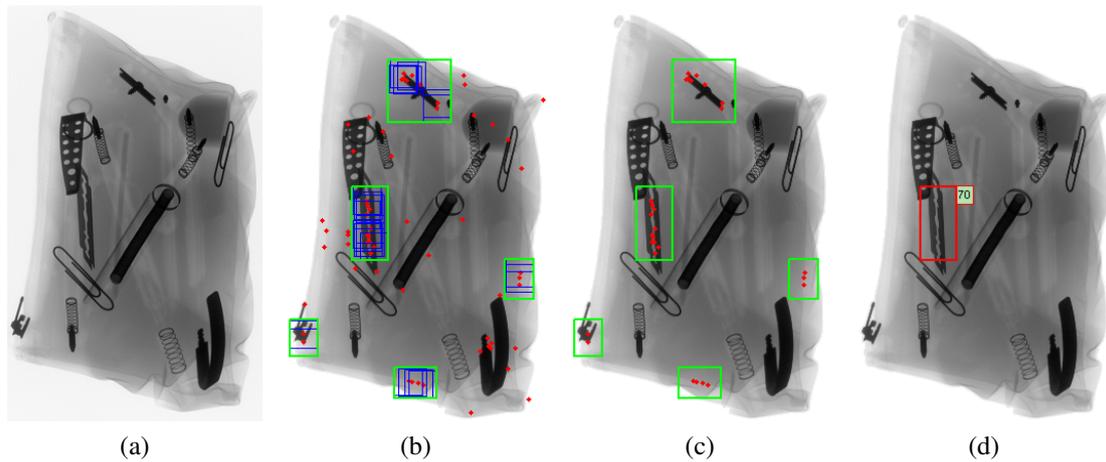


Figure 4: Detection and pose estimation: a) Test image. b) Extracted SIFT keypoints (red points) and clusters of keypoints with the same pose (blue subwindows  $\mathcal{W}_B$ ). b) and c) Merged clusters (green subwindows  $\mathcal{W}_G$ ). d) Selected potential target object (red subwindow  $\mathcal{W}_S$ ) and its corresponding pose '70' (compare to pose 70 in Fig. 3).

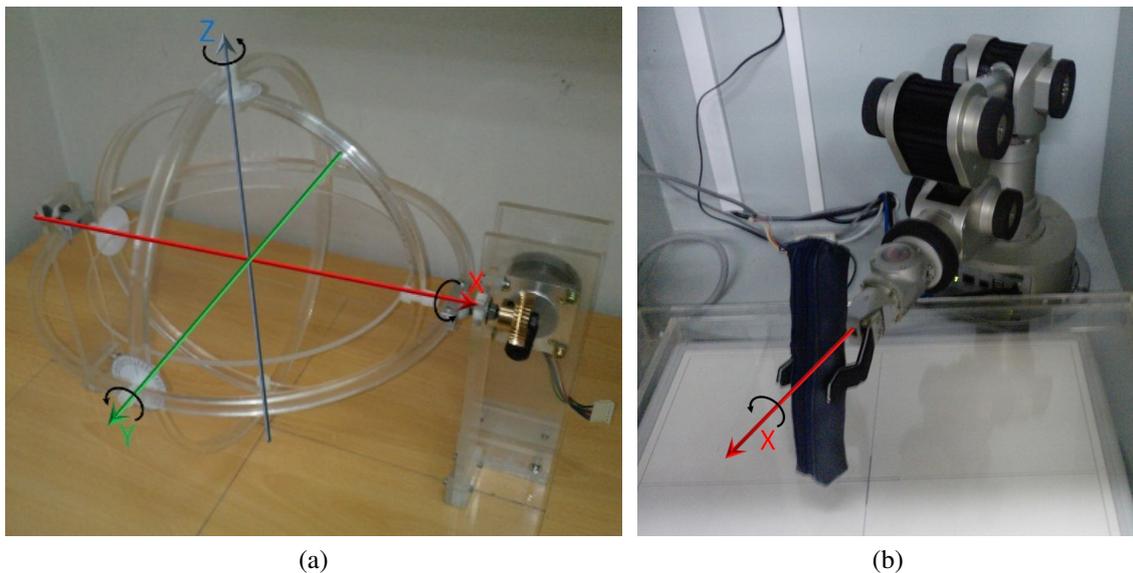


Figure 5: Manipulation systems: a) Semi-automatic system: axis  $X$  can be controlled using a stepper motor, whereas axes  $Y$  and  $Z$  are manually rotated. b) Robotic arm: only axis  $X$  can be rotated by the manipulator.

## 2.4 Estimation of next best position

As outlined in previous subsections, our framework attempts to move the inspection object so that a *good pose* of the target object should be viewed. In this Section, we explain in further details, how this *next best position* is estimated.

As shown in Fig. 2, there are two different scenarios by estimating the next best position. Scenario A: when a potential target object was detected in previous step. It means that the pose of the detected potential target object does not correspond to a *good pose*. Scenario B: when a potential target object was not detected in previous step. It happens when there is no target object inside of the inspection object, or when the pose of a existing target object is so intricate that the detection may fail, as shown in Fig. 1a. Scenario B is easily dealt with by moving the inspection object to a very different position. In our case, we rotate the inspection object around an axis perpendicular to the optical axis of the X-ray projection in  $\omega_X = -40^\circ$ . This value was heuristically determined by studying Fig. 3. Thus, the target object -if any- could be quickly aligned to the region of the *good poses*, as shown in Fig. 1b. On the other hand, for scenario A, the next best position will depend on the estimated (actual) pose of the potential detected target object. Generally, the problem here is that the manipulation system can rotate only the axes of the inspection object and these axes do not correspond to the axes of the target object. It would be very simple to obtain a good pose of our target object if the manipulation system could rotate the target object around  $\alpha$  or  $\beta$  axes of the razor blade (see Fig. 3). For this reason, we present in this paper two different manipulation systems as shown in Fig. 5: a) a semi-automatic system (with three rotation axes ( $X, Y, Z$ )), and b) a robotic arm with only one rotation axis ( $X$ ). In both cases, axis  $Z$  corresponds to the optical axis of the X-ray projection. In following we explain how this problem can be solved using these constraints.

In order to move the manipulation system to a position that corresponds to a *good pose* of the target object, first, the  $Z$  axis is rotated  $\omega_Z = \gamma$  -if it is possible- so that the major axis of the razor blade is parallel to the  $X$  axis, and afterwards the  $X$  axis is rotated  $\omega_X = -\alpha$ , where  $\alpha$  is the angle of the estimated pose as shown in Fig. 3, *e.g.*, for pose labeled as ‘70’ the rotation will be  $\omega_X = -60^\circ$ . In the semi-automatic manipulation system both rotations are possible, whereas in the robotic arm only the last rotation is possible.

In order to estimate the orientation of the razor blade, an ellipse will be fitted to the keypoints of the estimated pose contained in the selected subwindow  $\mathcal{W}_S$  as shown in Fig. 6. The angle between the major axis of the ellipse and the horizontal  $X$  axis will be denoted by  $\phi$ . Thus, if the angle  $\phi \geq 90^\circ$  then the rotation of  $Z$  axis will be  $\gamma = 180^\circ - \phi$ , and if the angle  $\phi < 90^\circ$  then  $\gamma = -\phi$  (see Fig. 6).

Using this approach, the attempt is made to obtain a new position of the inspection object where the pose of the target object should be a *good pose* or close to a *good pose*. It is worth noting that due to perspective, razor blade X-ray images acquired from four different poses  $(\pm\alpha, \pm\beta)$  are very similar. For this reason the pose estimation, and therefore the estimation of the next best position, may fail. In these cases, the next view may not be a good view and thus, no target object will be detected or no good pose will be obtained. There are two alternatives to deal with this erroneously estimated position: *i*) to go back and then to correct the pose estimation, and *ii*) to try to obtain a new next best position from this erroneous position. From a practical standpoint, since option *i*) can be repeated until three times, we decided on option *ii*). Our experiments validated this option.

Theoretically, in case of inspecting with a manipulation system with only one rotation axis (case of the robotic arm), the detection will fail if the razor blade is placed perpendicular to rotation axis, because by rotating the target object it will never achieve a good pose. In order to overcome this problem, the object must be inspected again from another viewpoint. This can be performed by leaving the object on a table and taking it again from another pose.

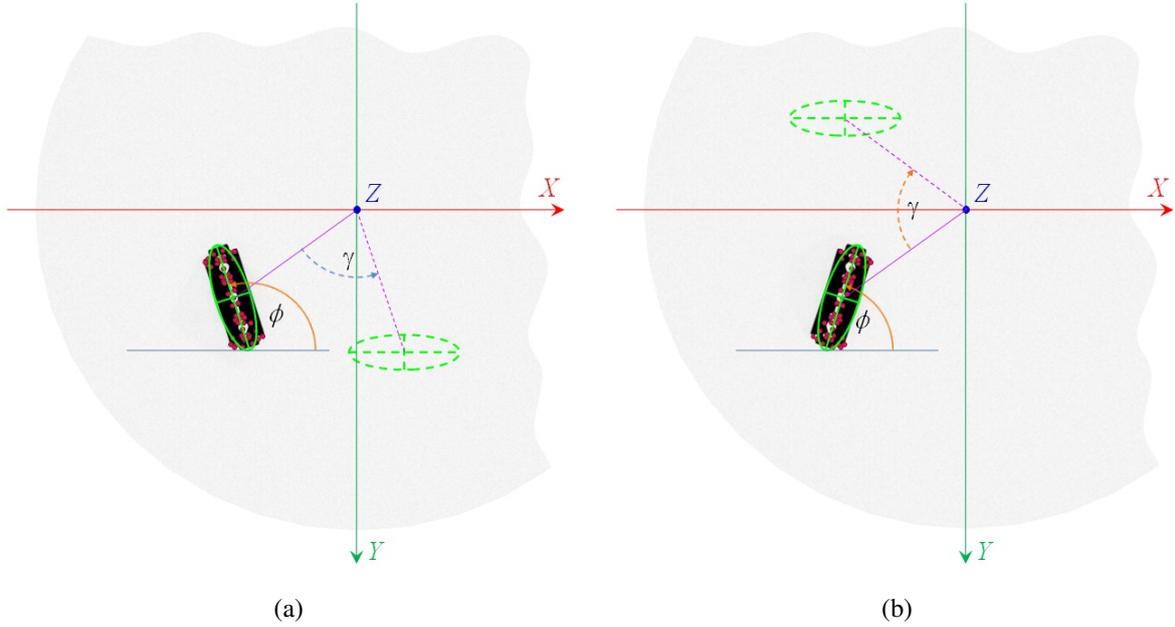


Figure 6: rotation around the axis  $Z$ : a) when  $\phi \geq 90^\circ$ , b) when  $\phi < 90^\circ$ .

### 3 Experimental Results

In the experiments, we used a digital X-rays detector (Canon, model CXDI-50G), X-rays emitter tube (Poskom, model PXM-20BT ) and lead security cabinet to isolate inspection environment (see Fig. 7). The manipulation systems were a clamping mechanism semi-automatic (Gyroscope), and a robotic arm (Neuronics, model Katana 6M), as shown in Fig. 5. The nine inspection objects (a,b,c,d,e,f,g,h,i) are illustrated in Fig. 8. Inside of all inspection objects was a razor blade. Seven objects (a,b,c,d,e,f,g) were inspected using the semi-automatic system, whereas six objects (a,c,d,g,h,i) were inspected using the robotic arm. Each object was tested starting from 10 arbitrary initial positions, yielding  $(7 + 6) \times 10 = 130$  experiments.

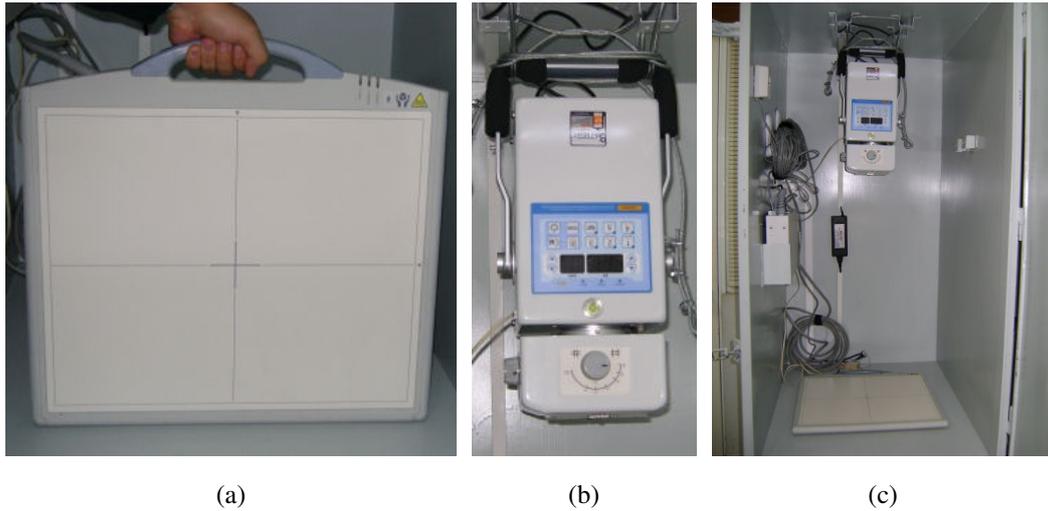


Figure 7: X-ray imaging system: a) X-rays detector (flat panel), b) X-rays source, and c) lead cabinet with source and flat panel.

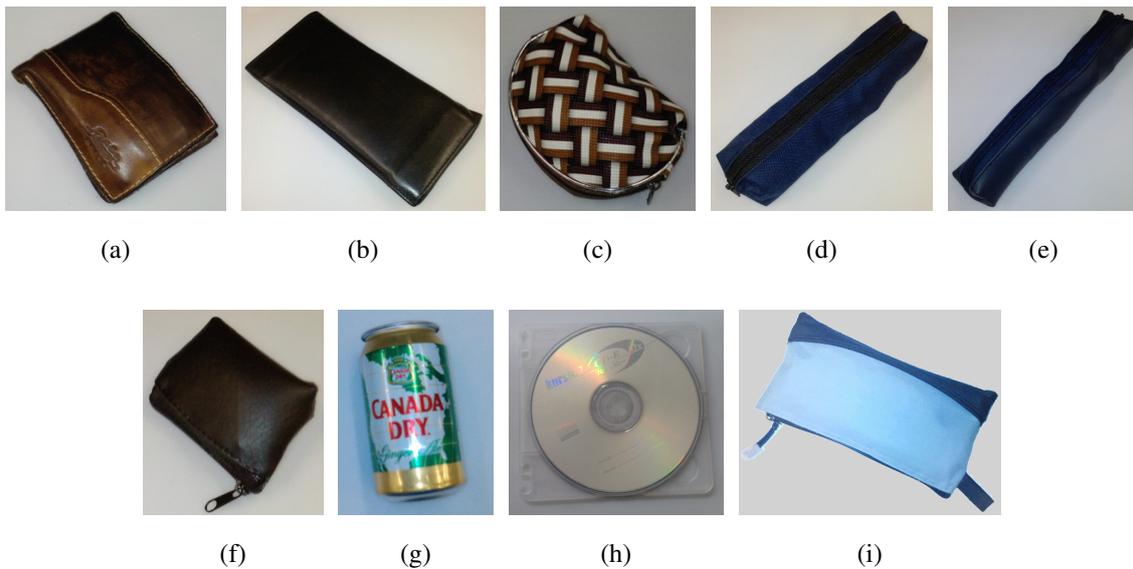


Figure 8: Inspection objects.

### 3.1 Description of the target object

As mentioned in Section 2.2, the target object –a razor blade– was described from 81 different poses by rotating  $\alpha$  and  $\beta$  axes (Fig. 3). The razor blade was located in the center of a sphere of EPS (*Expanded PolyStyrene*) as shown in Fig. 9a. We used a sphere of EPS, due to its low X-ray absorption coefficient. The Fig. 9b illustrates the description of a razor blade using SIFT descriptors for pose ‘1’ of Fig. 3. In the experiments, we used the SIFT implementation of [20].



Figure 9: Characterization of a razor blade, a) sphere of EPS (*Expanded PolyStyrene*), and b) razor blade with two SIFT representations: magnitude-orientation, and keypoints.

### 3.2 Inspection experiments

Every object was inspected 10 times where the first image of each inspection test was acquired in a random position. The obtained results are summarized in Table 1 for the semi-automatic system and Table 2 for the robotic arm. As shown in the tables, sometimes only the first image was enough to detect the target object, however, in other cases 2, 3 or 4 images were necessary. The algorithm for detection of potential target objects failed in some cases (see *ND* for no-detection), however, in many of these cases the target object could be detected in the next view(s). Thus, in these cases, our approach was able to overcome intricate views and some occlusions. Nevertheless, in certain cases our approach could not detect any target object due to severe occlusion and unfavorable initial positions.

Figures 10, 11, 12 and 13 illustrate the inspection process for four different inspection objects. We can see the ability of our approach to find the target object looking for good views even with partial occlusions. Since the clamping system of the semi-automatic system has a high X-ray absorption coefficient, some regions of the test image were so dark that caused false detections (*FP*). On the other hand, since the robotic arm could move the inspection object using only one degree of freedom (rotation of  $X$  axis), the whole inspection in certain cases was performed in more views than using the semi-automatic system that was able to rotate the  $Z$  axis as well.

In order to evaluate the performance of our approach, we calculated *precision* and *recall* rates from Tables 1 and 2. They are defined as  $precision = TP/(TP + FP)$  and  $recall = TP/P$ , where  $TP$  is the total number of true positives (detected razor blades),  $FP$  is the total number of false positives (false detections) and  $P$  is the total number of razor blades (one per experiment).

Thus, the semi-automatic system yielded a rate of  $recall = 88.6\%$  and  $precision = 92.5\%$ , and the robotic arm achieved rates of  $recall = 88.3\%$  and  $precision = 91.4\%$ .

The framework was implemented using MATLAB. Each single view was processed in average in 10 seconds on an iMac OS X 10.6.8, processor 3.06GHz Intel Core 2 Duo, 4GB RAM memory.

Table 1: Inspection sequences and performance indicators with semi-automatic system.

Object	Sequence	Pose	Rotation		Pose	Rotation		Pose	Rotation		Pose	Performance	
			X	Z		X	Z		X	Z		TP	FP
a	1	2	-	-	-	-	-	-	-	-	-	1	0
	2	2	-	-	-	-	-	-	-	-	-	1	0
	3	ND	-40°	-	2	-	-	-	-	-	-	1	0
	4	ND	-40°	-	28	-	-	-	-	-	-	1	0
	5	2	-	-	-	-	-	-	-	-	-	1	0
	6	ND	-40°	-	ND	-40°	-	28	-	-	-	1	0
	7	2	-	-	-	-	-	-	-	-	-	1	0
	8	ND	-40°	-	2	-	-	-	-	-	-	1	0
	9	2FP	-	-	-	-	-	-	-	-	-	0	1
	10	ND	-40°	-	ND	-40°	-	2	-	-	-	1	0
b	1	1	-	-	-	-	-	-	-	-	-	1	0
	2	3	-	-	-	-	-	-	-	-	-	1	0
	3	ND	-40°	-	10	-	-	-	-	-	-	1	0
	4	65	-10°	-35°	56	-10°	-17°	55	-	-	-	1	0
	5	59	-40°	-31°	2FP	-	-	-	-	-	-	0	1
	6	60	-50°	-69°	2	-	-	-	-	-	-	1	0
	7	16	-60°	0°	10	-	-	-	-	-	-	1	0
	8	7	-60°	20°	19	-	-	-	-	-	-	1	0
	9	8	-70°	23°	10	-	-	-	-	-	-	1	0
	10	ND	-40°	-	23	-40°	24°	1	-	-	-	1	0
c	1	19	-	-	-	-	-	-	-	-	-	1	0
	2	ND	-40°	-	28	-	-	-	-	-	-	1	0
	3	ND	-40°	-	ND	-40°	-	64	-	-	-	1	0
	4	ND	-40°	-	ND	-40°	-	ND	-	-	-	0	0
	5	ND	-40°	-	ND	-40°	-	3	-	-	-	1	0
	6	ND	-40°	-	2FP	-	-	-	-	-	-	0	1
	7	10	-	-	-	-	-	-	-	-	-	1	0
	8	ND	-40°	-	ND	-40°	-	37	-	-	-	1	0
	9	ND	-40°	-	ND	-40°	-	19	-	-	-	1	0
	10	ND	-40°	-	ND	-40°	-	5	40°	-85°	21	1	0
d	1	50	-40°	9°	28	-	-	-	-	-	-	1	0
	2	39	-20°	-77°	51	-50°	43°	ND	-40°	-	3	1	0
	3	2FP	-	-	-	-	-	-	-	-	-	0	1
	4	28	-	-	-	-	-	-	-	-	-	1	0
	5	69	-50°	-83°	38	-10°	-77°	2	-	-	-	1	0
	6	4	-	-	-	-	-	-	-	-	-	1	0
	7	6	-50°	27°	ND	-40°	-	32	-40°	15°	20	1	0
	8	30	-	-	-	-	-	-	-	-	-	1	0
	9	48	-20°	-58°	37	-	-	-	-	-	-	1	0
	10	50	-40°	-44°	2	-	-	-	-	-	-	1	0
e	1	31	-	-	-	-	-	-	-	-	-	1	0
	2	ND	-40°	-	ND	-40°	-	ND	-	-	-	0	0
	3	77	-40°	-41°	30	-	-	-	-	-	-	1	0
	4	ND	-40°	-	ND	-40°	-	46	-	-	-	1	0
	5	13	-	-	-	-	-	-	-	-	-	1	0
	6	ND	-40°	-	2	-	-	-	-	-	-	1	0
	7	ND	-40°	-	46	-	-	-	-	-	-	1	0
	8	ND	-40°	-	ND	-40°	-	2	-	-	-	1	0
	6	ND	-40°	-	2	-	-	-	-	-	-	1	0
	10	59	-40°	83°	ND	-40°	-	2	-	-	-	1	0
f	1	ND	-40°	-	2	-	-	-	-	-	-	1	0
	2	7	-60°	-4°	7	-60°	-8°	ND	-40°	-	2	1	0
	3	ND	-40°	-	ND	-40°	-	ND	-	-	-	0	0
	4	2	-	-	-	-	-	-	-	-	-	1	0
	5	ND	-40°	-	2	-	-	-	-	-	-	1	0
	6	ND	-40°	-	ND	-40°	-	2	-	-	-	1	0
	7	ND	-40°	-	37	-	-	-	-	-	-	1	0
	8	16	-60°	-1°	2	-	-	-	-	-	-	1	0
	9	2FP	-	-	-	-	-	-	-	-	-	0	1
	10	ND	-40°	-	2	-	-	-	-	-	-	1	0
g	1	2	-	-	-	-	-	-	-	-	-	1	0
	2	41	-40°	17°	19	-	-	-	-	-	-	1	0
	3	ND	-40°	-	66	-20°	68°	20	-	-	-	1	0
	4	42	-50°	3°	19	-	-	-	-	-	-	1	0
	5	5	-40°	-64°	19	-	-	-	-	-	-	1	0
	6	ND	-40°	-	64	-	-	-	-	-	-	1	0
	7	41	-40°	4°	20	-	-	-	-	-	-	1	0
	8	22	-	-	-	-	-	-	-	-	-	1	0
	9	24	-50°	10°	19	-	-	-	-	-	-	1	0
	10	48	-20°	-4°	38	-10°	-19°	46	-	-	-	1	0
TOTAL:												62	5
$recall = TP/P:$												88,6%	
$precision = TP/(TP + FP):$												92,5%	

ND: No-Detection, TP: True Positive, FP: False Positive, P: total of experiments.

Table 2: Inspection sequences and performance indicators with robotic arm.

Object	Sequence	Pose	Rotation	Pose	Rotation	Pose	Rotation	Pose	Performance	
			X		X		X		TP	FP
a	1	2	-	-	-	-	-	-	1	0
	2	ND	-40°	ND	-40°	4	-	-	1	0
	3	ND	-40°	5	-40°	2	-	-	1	0
	4	ND	-40°	2	-	-	-	-	1	0
	5	ND	-40°	2	-	-	-	-	1	0
	6	11	-	-	-	-	-	-	1	0
	7	ND	-40°	14	-40°	2	-	-	1	0
	8	ND	-40°	19	-	-	-	-	1	0
	9	ND	-40°	ND	-40°	1	-	-	1	0
	10	24	-50°	2FP	-	-	-	-	0	1
c	1	ND	-40°	4	-	-	-	-	1	0
	2	5	-40°	1	-	-	-	-	1	0
	3	ND	-40°	ND	-40°	4	-	-	1	0
	4	ND	-40°	4	-	-	-	-	1	0
	5	ND	-40°	10	-	-	-	-	1	0
	6	46	-	-	-	-	-	-	1	0
	7	ND	-40°	47	-10°	38	-10°	30	1	0
	8	55	-	-	-	-	-	-	1	0
	9	52FP	-	-	-	-	-	-	0	1
	10	48	-20°	31	-	-	-	-	1	0
d	1	28	-	-	-	-	-	-	1	0
	2	ND	-40°	ND	-40°	29	-	-	1	0
	3	3	-	-	-	-	-	-	1	0
	4	ND	-40°	37	-	-	-	-	1	0
	5	ND	-40°	64	-	-	-	-	1	0
	6	55	-	-	-	-	-	-	1	0
	7	ND	-40°	67	-30°	61	-60°	ND	0	0
	8	37	-	-	-	-	-	-	1	0
	9	68	-40°	ND	-40°	2FP	-	-	0	1
	10	ND	-40°	55	-	-	-	-	1	0
g	1	28	-	-	-	-	-	-	1	0
	2	20	-	-	-	-	-	-	1	0
	3	55	-	-	-	-	-	-	1	0
	4	46	-	-	-	-	-	-	1	0
	5	65	-10°	56	-10°	46	-	-	1	0
	6	65	-10°	57	-20°	39	-20°	20	1	0
	7	ND	-40°	46	-	-	-	-	1	0
	8	ND	-40°	ND	-40°	10	-	-	1	0
	9	5	-40°	1	-	-	-	-	1	0
	10	ND	-40°	52	-60°	3	-	-	1	0
h	1	ND	-40°	6	-50°	1	-	-	1	0
	2	48	-20°	29	-	-	-	-	1	0
	3	5	-40°	2	-	-	-	-	1	0
	4	ND	-40°	3	-	-	-	-	1	0
	5	55	-	-	-	-	-	-	1	0
	6	ND	-40°	19	-	-	-	-	1	0
	7	2FP	-	-	-	-	-	-	0	1
	8	ND	-40°	6	-50°	11	-	-	1	0
	9	59	-40°	21	-	-	-	-	1	0
	10	6	-50°	10	-	-	-	-	1	0
i	1	19	-	-	-	-	-	-	1	0
	2	ND	-40°	48	-20°	29	-	-	1	0
	3	ND	-40°	ND	-40°	37	-	-	1	0
	4	55	-	-	-	-	-	-	1	0
	5	ND	-40°	46	-	-	-	-	1	0
	6	ND	-40°	2FP	-	-	-	-	0	1
	7	ND	-40°	ND	-40°	ND	-	-	0	0
	8	ND	-40°	ND	-40°	2	-	-	1	0
	9	55	-	-	-	-	-	-	1	0
	10	41	-40°	56	-10°	55	-	-	1	0
<b>TOTAL:</b>									53	5
recall = TP/P:									88,3%	
precision = TP/(TP + FP):									91,4%	

ND: No-Detection, TP: True Positive, FP: False Positive, P: total of experiments.

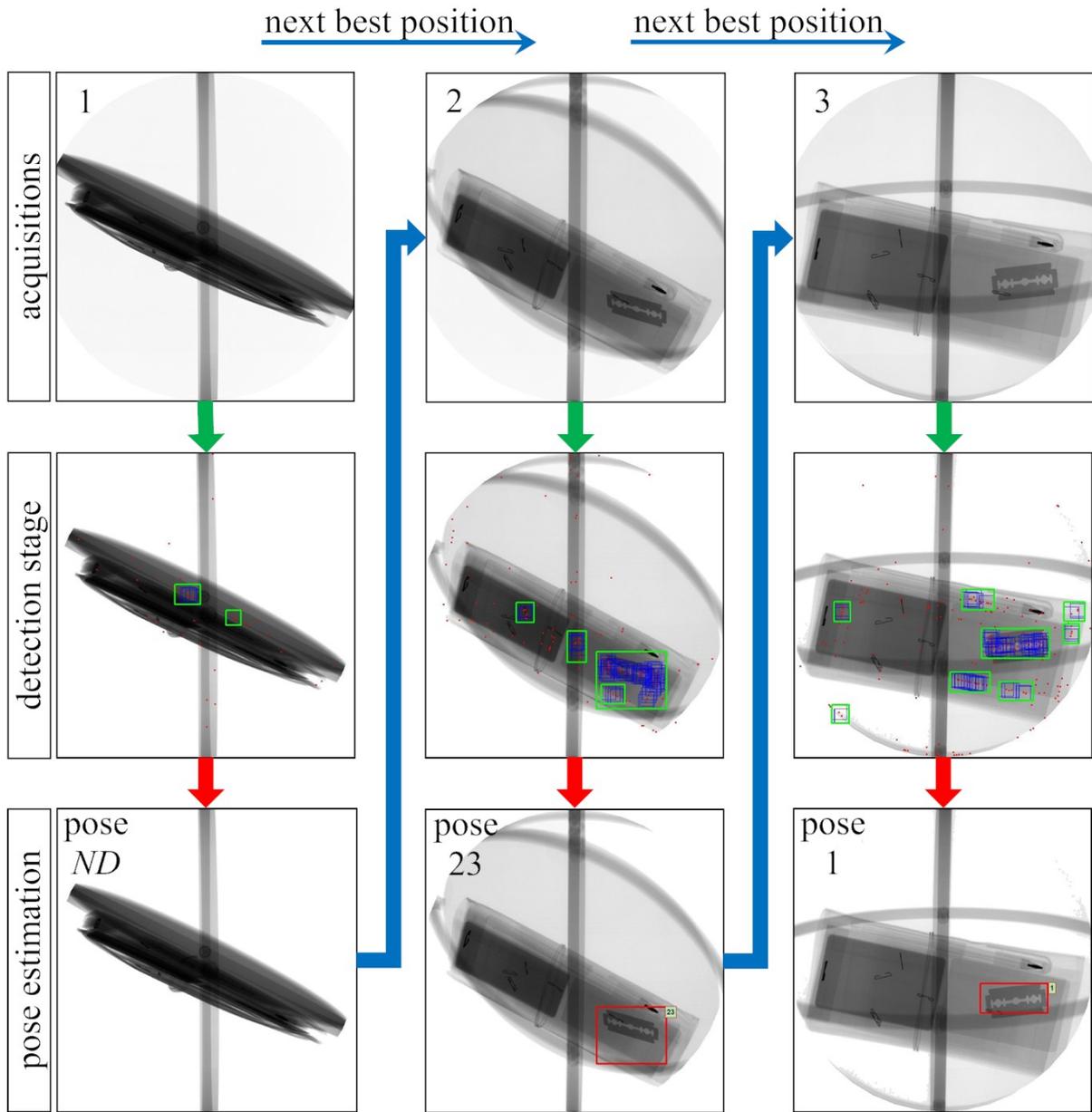


Figure 10: Inspection of object *b*, sequence 10 with semi-automatic system (see Table 1).

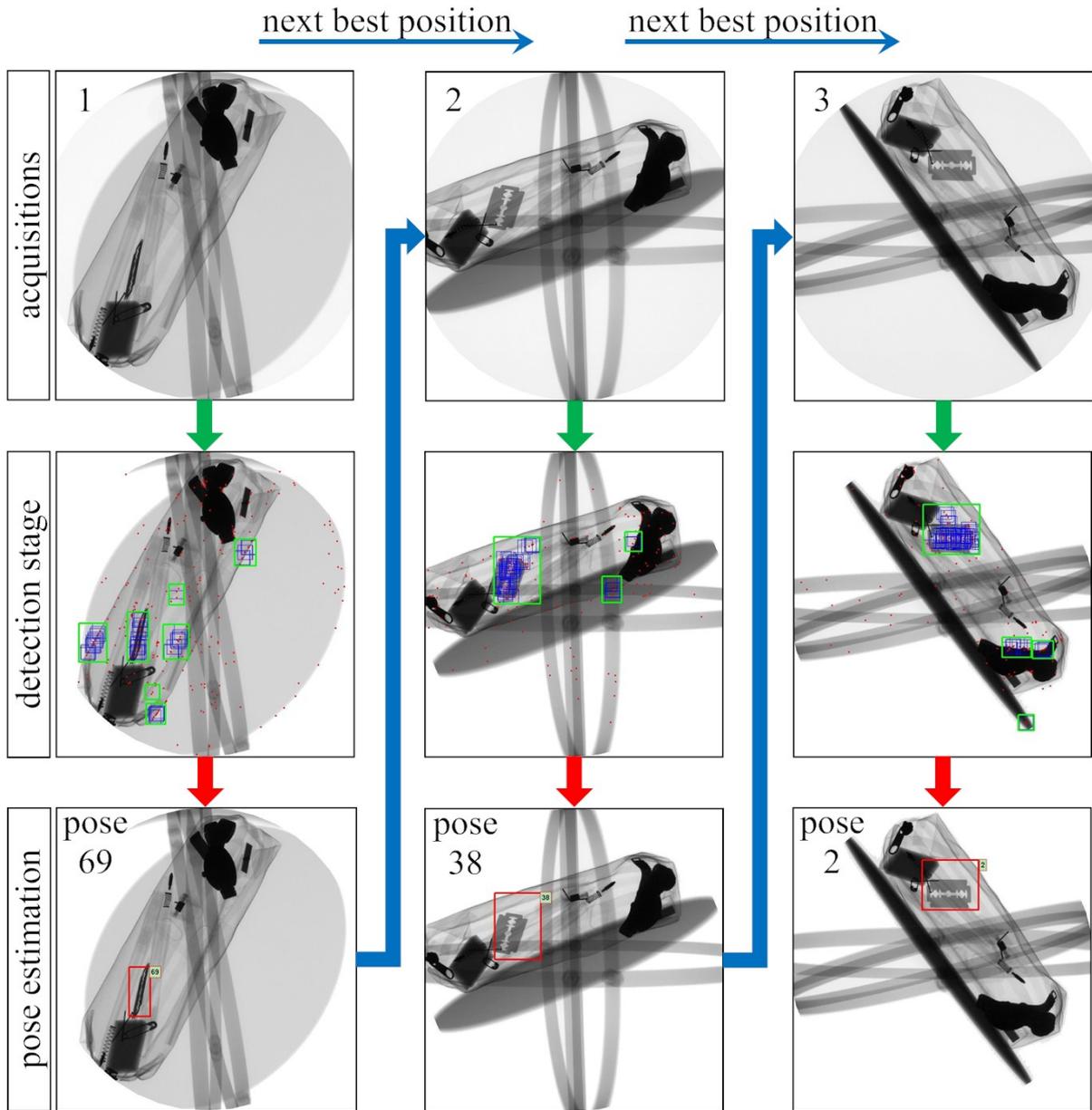


Figure 11: Inspection of object *d*, sequence 5 with semi-automatic system (see Table 1).

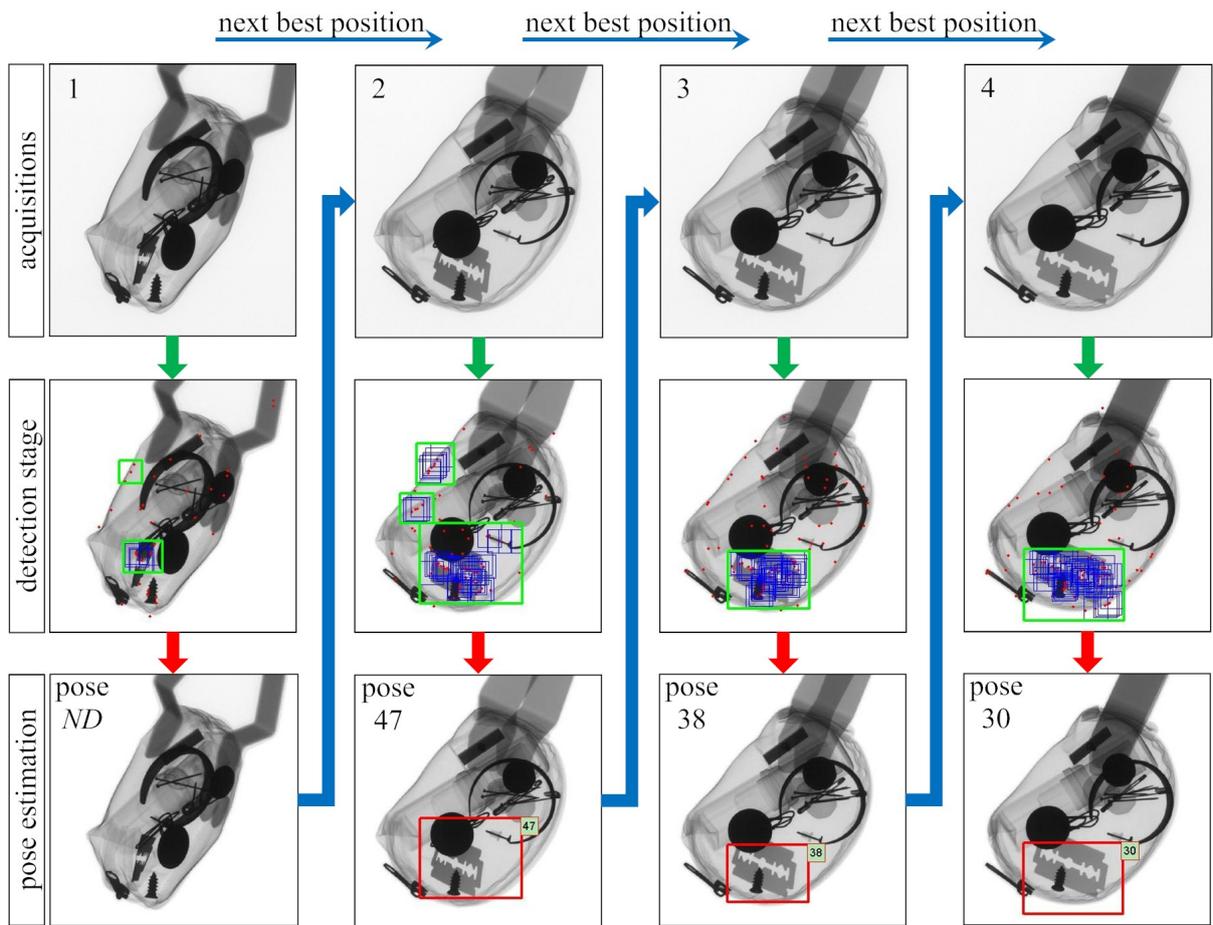


Figure 12: Inspection of object *c*, sequence 7 with robotic arm (see Table 2).

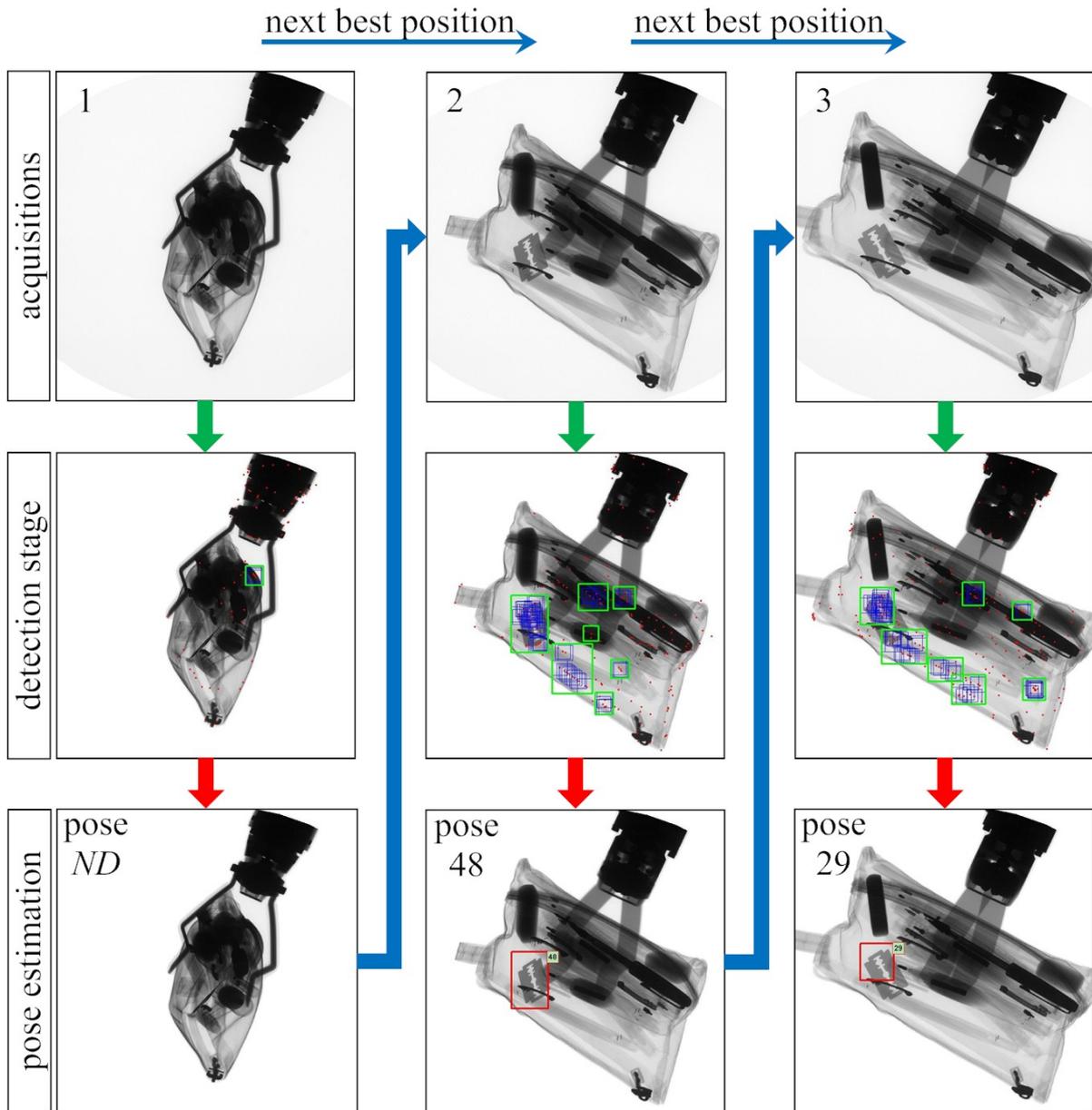


Figure 13: Inspection of object  $i$ , sequence 9 with robotic arm (see Table 2).

## 4 Conclusions

In this paper, we proposed a radiosopic multiple view framework which is able to inspect complex objects using *active vision*. The framework we proposed was exemplified by detection of a razor blade, which in places such as airports and customs centers is considered a dangerous object. Although the detection of a razor blade may seem simple, conceptual and experimental evidence showed us the complexity of solve this problem due to the symmetry in all quadrants,

thinness and smallness, and low X-ray absorption coefficient.

Using a robotic arm and a semi-automatic manipulator system, the robustness and reliability of the method have been verified in the automated detection of razor blades located inside of nine different objects showing promising preliminary results: in 130 experiments we were able to detect 115 times the razor blade with 10 false alarms, achieving recall of 89% and precision of 92%. Comparing our results with human inspection processes that can achieve a around than 90% in baggage screening at airports [7–9], we found that our approach is promising.

As feature work, we will implement this method in more complex scenarios with more occlusion and other target objects. Additionally, we will incorporate a calibrated model of the X-ray projection in order to manipulate the inspection object with more accuracy.

## Acknowledgments

This work was supported by grant Fondecyt No. 1100830 from CONICYT–Chile.

## References

- [1] G. Zentai, “X-ray imaging for homeland security,” *Imaging Systems and Techniques, 2008. IST 2008. IEEE International Workshop on*, pp. 1–6, Sept. 2008.
- [2] R. P. Haff and N. Toyofuku, “X-ray detection of defects and contaminants in the food industry,” *Sensing and Instrumentation for Food Quality and Safety*, vol. 2, no. 4, pp. 262–273, 2008.
- [3] D. Mery, I. Lillo, H. Loebel, V. Riffo, A. Soto, A. Cipriano, and J. Aguilera, “Automated fish bone detection using x-ray testing,” *Journal of Food Engineering*, vol. 105, no. 3, pp. 485–492, 2011.
- [4] X. Duan, J. Cheng, L. Zhang, Y. Xing, Z. Chen, and Z. Zhao, “X-ray cargo container inspection system with few-view projection imaging,” *Nuclear Instruments and Methods in Physics Research A*, vol. 598, pp. 439–444, 2009.
- [5] D. Mery, “Automated radiosopic testing of aluminum die castings,” *Materials Evaluation*, vol. 64, no. 2, pp. 135–143, 2006.
- [6] T. W. Liao, “Improving the accuracy of computer-aided radiographic weld inspection by feature selection,” *NDT & E International*, vol. 42, no. 4, pp. 229–239, 2009.
- [7] A. Schwaninger, D. Hardmeler, and F. Hofer, “Aviation security screeners visual abilities visual knowledge measurement,” *Aerospace and Electronic Systems Magazine, IEEE*, vol. 20, pp. 29 – 35, june 2005.
- [8] D. Hardmeier, F. Hofer, and A. Schwaninger, “The role of recurrent cbt for increasing aviation security screeners’ visual knowledge and abilities needed in x-ray screening,” pp. 338–342, 2006.

- [9] A. Wales, C. Anderson, K. Jones, A. Schwaninger, and J. Horne, “Evaluating the two-component inspection model in a simplified luggage search task,” *Behavior research methods*, vol. 41, no. 3, p. 937, 2009.
- [10] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge Univ Pr, 2003.
- [11] D. Mery and D. Filbert, “Automated flow detection in aluminum castings based on the tracking of potential defects in a radiosopic image sequence,” *IEEE Transactions on Robotics and Automation*, vol. 18, no. 6, pp. 890–901, 2002.
- [12] D. Mery, “Explicit geometric model of a radiosopic imaging system,” *NDT & E International*, vol. 36, no. 8, pp. 587–599, 2003.
- [13] M. Carrasco and D. Mery, “Automated visual inspection using trifocal analysis in an uncalibrated sequence of images,” *Materials Evaluation*, vol. 64, pp. 900–906, 2006.
- [14] D. Mery, “Automated detection in complex objects using a tracking algorithm in multiple x-ray views,” in *Proceedings of the 8th IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum (OTCBVS 2011), Colorado*, pp. 41–48, 2011.
- [15] C. Pieringer and D. Mery, “Flaw detection in aluminum die castings using simultaneous combination of multiple views,” *Insight*, vol. 23, no. 10, pp. 548–552, 2010.
- [16] D. Mery, “Automated detection of welding discontinuities without segmentation,” *Materials Evaluation*, no. June, pp. 657–663, 2011.
- [17] F. Farshidi, S. Sirouspour, and T. Kirubarajan, “Robust sequential view planning for object recognition using multiple cameras,” *Image and Vision Computing*, vol. 27, no. 8, pp. 1072 – 1082, 2009.
- [18] H. Dutagaci, C. P. Cheung, and A. Godil, “A benchmark for best view selection of 3d objects,” in *Proceedings of the ACM workshop on 3D object retrieval, 3DOR ’10, (New York, NY, USA)*, pp. 45–50, ACM, 2010.
- [19] D. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [20] A. Vedaldi and B. Fulkerson, “VLFeat: An open and portable library of computer vision algorithms.” <http://www.vlfeat.org/>, 2008.