

# Inspección Radioscópica Activa de Objetos Complejos Usando Múltiples Vistas

Vladimir Riffo and Domingo Mery

Departamento de Ciencia de la Computación - Pontificia Universidad Católica de Chile  
Av. Vicuña Mackenna 4860(143), Santiago de Chile.  
vriffo1@uc.cl, dmery@ing.puc.cl

**Resumen** La inspección radioscópica de objetos complejos, como por ejemplo la revisión de maletas en los aeropuertos, por lo general es realizada manualmente. Esto no siempre es efectivo, ya que depende en gran medida de la pose del objeto de interés, la oclusión y las capacidades de los inspectores humanos. Por esta razón hemos desarrollado un *framework* para la inspección radioscópica de objetos complejos, que mediante visión activa basada en la estimación de pose, permite detectar objetos de interés tales como una hoja de afeitar. Creemos que este trabajo constituye el primer paso para la automatización de los sistemas de seguridad, ya que los resultados obtenidos han sido alentadores en la detección de una hoja de afeitar oculta en seis objetos, y en sesenta secuencias de inspección (diez por cada objeto) logrando índices de *recall* = 86,7 % y *precision* = 91,2 %.

## 1. Introducción

Los rayos X no sólo se han utilizado en medicina sino también en *inspección no destructiva* (NDT: *Non-Destructive Testing*) de materiales u objetos [9], donde el propósito es analizar las partes internas que son imperceptibles a simple vista. Con el fin de lograr inspecciones con rayos X eficaces y eficientes, los sistemas automáticos están siendo desarrollados para ejecutar tareas difíciles, tediosas y a veces peligrosas con el fin de ofrecer una inspección objetiva y reproducible en cada prueba. Algunas aplicaciones son análisis de productos alimenticios [3], detección de objetos en equipajes [11], inspección de piezas de automóviles [5], control de calidad de las soldaduras [10], entre otras. En la inspección de equipaje se han desarrollado algoritmos orientados principalmente a la segmentación de imágenes, para ayudar visualmente a inspectores humanos, y sólo se hace inspección de una imagen radioscópica, lo cual no siempre es adecuado ya que el elemento de interés que se encuentra dentro de un objeto, puede estar en oclusión parcial o total, o en una pose que impida su detección. Hoy en día sin embargo está cobrando fuerza la inspección de múltiples vistas que logra reducir el efecto de la oclusión, y aumentar así los índices de desempeño [5, 6].

Nuestro trabajo presenta un método de inspección activa de múltiples vistas radioscópicas, es decir que mediante el análisis de una primera vista, se pueda buscar de manera automática una segunda, tercera y hasta cuarta vista que ayude a realizar un análisis más confiable. En este sentido nuestra propuesta incorpora algunas ideas de *mejor vista* y *planeamiento de vistas* en donde existe una gran variedad de trabajos

publicados en las áreas de reconocimiento y reconstrucción 3D [1, 2], pero no se han encontrado publicaciones enfocadas a la inspección radioscópica.

Este artículo presenta en la Sección 2 un resumen del método propuesto, en la Sección 3 los experimentos y resultados y en la Sección 4 las conclusiones.

## 2. Método Propuesto

El método propuesto consta de dos etapas como se ilustra en la Fig. 1: A) Se analiza una primera vista en la que se busca el objeto de interés, si éste es encontrado se estima su pose (localización y orientación) y se mueve el objeto a una siguiente posición en la que se espera una mejor vista para corroborar la detección. Este proceso se puede repetir varias veces. B) En caso de no encontrar el objeto de interés en la primera vista, el objeto es movido arbitrariamente a una nueva posición –muy distinta a la primera– con el fin de obtener una vista muy diferente en la que, si hay un objeto a detectar, lo podamos apreciar mejor. Este *framework* propuesto es de carácter general y fácilmente adaptable, sin embargo, los algoritmos implicados deben ser *ad-hoc* al objeto a detectar. En nuestros experimentos hemos ejemplificado el proceso de detección de una hoja de afeitar, por esta razón los detalles de nuestro método entregados a continuación son para esta detección en particular.

Para la detección es necesario contar con una caracterización del objeto de interés, con este fin usamos descriptores SIFT [4] que son invariantes a la posición, rotación, escalamiento e intensidad, condiciones requeridas en el análisis de múltiples vistas de rayos X, donde las estructuras se ven representadas como imágenes sombra, aumentadas de tamaño, según sea la posición del centro óptico o punto de emisión de rayos x [8]. Así, se construye una base de datos (*BD*) con los descriptores SIFT del objeto de interés en todas las poses factibles, tal como se muestra en la Fig.2.

La detección se realiza siguiendo la idea sugerida en [4], que consiste en realizar el *matching* entre los descriptores  $q_i$  obtenidos de la imagen de rayos X de inspección y los descriptores  $p_j$  almacenados en la base de datos *BD* del objeto de interés. Se

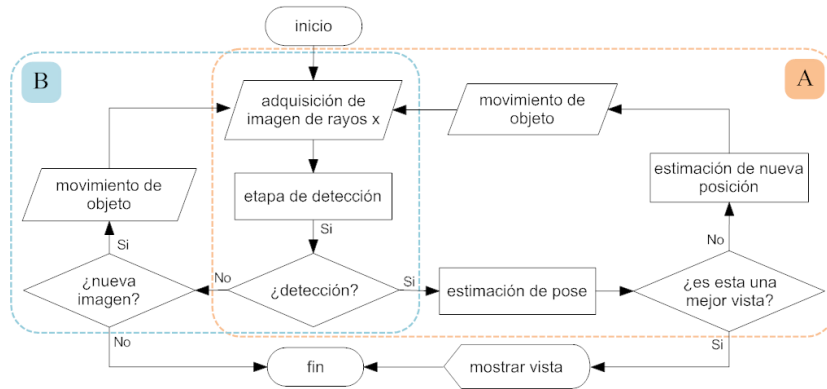


Figura 1: Framework para inspección activa con rayos X.

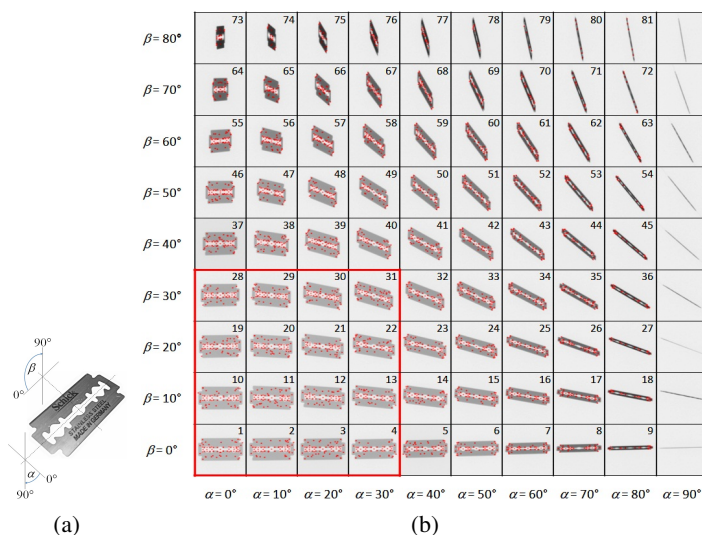


Figura 2: Caracterización con SIFT del objeto de interés: a) rotaciones aplicadas, y b) plantilla con descriptores SIFT (puntos rojos) e indicación de pose. El recuadro rojo ubicado abajo a la izquierda contiene las mejores vistas de la hoja de afeitar.

encuentran los mejores descriptores  $q_i$  que hagan *matching* con algún  $p_j$  que cumplan con  $d_\epsilon = \|q_i - p_j\| < \theta$ , donde  $\theta$  es un umbral de distancia mínima. Además, cada descriptor  $q_i$  encontrado, tendrá asociada la pose  $r_j$  del descriptor  $p_j$  que ha hecho el mejor *matching*. Cada imagen sometida a inspección tendrá ahora sólo los descriptores  $\mathcal{Q}$  y las poses  $\mathcal{R}$  que tienen la menor distancia a algún descriptores en  $P$  con su pose en  $R$  de la base de datos, respectivamente. A partir de esto, la detección se realiza en dos etapas: *localización* y *encuadramiento*.

**Localización:** En esta etapa se seleccionan algunos de los descriptores de  $\mathcal{Q}$  que están próximos entre sí y que tienen una pose común. Para esto, se define una pequeña ventana de tamaño  $w \times h$  en píxeles,  $\mathcal{W}_{w \times h}$ , cuyo centro corresponde a la posición  $(x, y)$  de cada descriptor. Así, sólo permanecerán en  $\mathcal{Q}$  los descriptores en la posición  $(x, y)$  que tengan un número  $\theta_2$  de descriptores situados dentro de  $\mathcal{W}_{w \times h}$  y que tengan la misma pose (ver ejemplo en Fig. 3b).

**Encuadramiento:** Todas las ventanas  $\mathcal{W}_{w \times h}$  que se encuentren unidas o traslapadas, formarán una nueva ventana  $\mathcal{W}_G$  de tamaño variable, mayor o igual que el tamaño de  $\mathcal{W}_{w \times h}$  (ver Fig. 3b, 3c).

Para estimar la pose, se exige a cada ventana detectada ( $\mathcal{W}_G$ ) que a lo menos contenga  $\theta_3$  descriptores con la misma pose (en nuestros experimentos  $\theta_3 = 8$ ), la pose que tiene mayor frecuencia, será la pose asignada a  $\mathcal{W}_G$ . Si lo anterior no ocurre, se descarta la  $\mathcal{W}_G$ , y por el contrario, si hay más de una  $\mathcal{W}_G$  con pose asignada, se selecciona como detección válida a la  $\mathcal{W}_G$  que tenga mayor número de descriptores con igual pose (Fig. 3d).

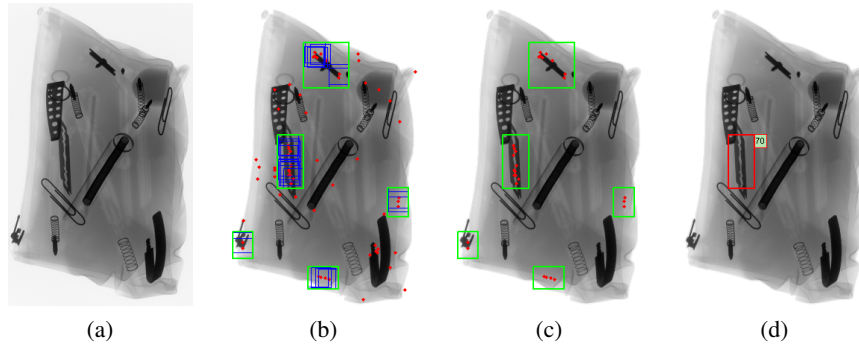


Figura 3: Etapas de detección y estimación de pose: a) imagen inicial, b) localización de ventanas  $\mathcal{W}_{w \times h}$  (ventanas azules) y encuadramiento de ventanas  $\mathcal{W}_G$  (ventanas verdes), c) ventanas  $\mathcal{W}_G$  encuadradas, y d) pose detectada = 70.

Un objeto sometido a inspección radioscópica, puede estar en distintas posiciones, sin embargo, no todas serán adecuadas para la detección de algún objeto de interés. Sumado a este inconveniente, está el hecho de que un objeto de interés tiene vistas que dificultan su reconocimiento visual. Es necesario entonces, establecer una estrategia de rotación del objeto y que permita detectar al objeto de interés en un reducido número de imágenes (*visión activa*). La rotación del objeto se realiza considerando dos situaciones: *no detección* y *detección*, y en ambas se realizan rotaciones, de tal forma de colocar al objeto de interés en una mejor vista (Fig.2). De forma general, se podrían realizar rotaciones en torno a los tres ejes  $X, Y, Z$  del sistema cartesiano del objeto. A modo de restricción, en nuestra propuesta hemos definido sólo dos rotaciones, una en torno al eje  $Z$  y otra en torno al eje  $X$ , considerando que cualquier mecanismo de inspección (manual, semi-automático y automático) tendrá limitados grados de libertad sin causar oclusión en las imágenes de rayos X. Cuando el objeto de interés no es detectado, ya sea por la no existencia al interior del objeto, o por la posición en que se encuentra, se realizará sólo una rotación de  $\alpha = -40^\circ$  en torno al eje  $X$ . Esto se ha determinado de manera heurística, ya que al realizar una rotación de  $40^\circ$  (positiva o negativa) en torno al eje  $X$ , se podría aproximar rápidamente a una zona de mejor vista (Fig.2).

Para detectar el objeto de interés y estimar su pose, se han requerido que a lo menos  $\theta_3$  descriptores hayan hecho *matching* con los descriptores en la  $BD$ . Estos descriptores son útiles para calcular la rotación  $\gamma$  en torno al eje  $Z$ . Esta rotación permite dejar al objeto de interés de forma horizontal o lo más próximo a esa posición, es decir con  $\beta = 0^\circ$  (Fig.2). Para esto es necesario ajustar a una elipse la posición  $(x, y)$  de los descriptores que hicieron *matching*, y determinar el ángulo  $\phi_1$  del eje mayor de la elipse con la horizontal. Si el ángulo  $\phi_1 \geq 90^\circ \Rightarrow \gamma = 180^\circ - \phi_1$ , y si el ángulo  $\phi_1 < 90^\circ \Rightarrow \gamma = -\phi_1$ , ver (Fig.4).

Cuando el objeto de interés es detectado y se ha estimado su pose, se puede buscar en la base de datos  $BD$  (Fig.2) la rotación en torno al eje  $X$  que dio origen a dicha pose, entonces la rotación será la misma pero en sentido contrario, es decir  $-\alpha$ . Las sucesivas

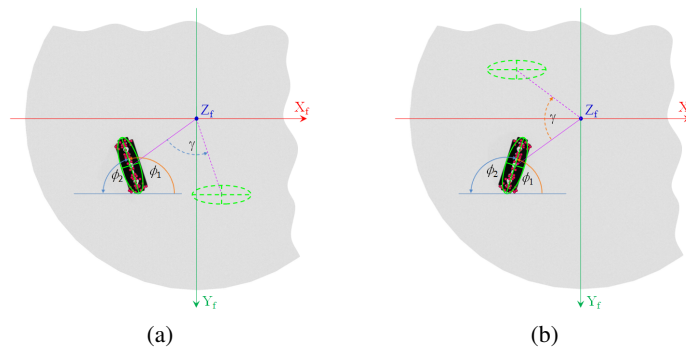


Figura 4: Rotación en  $Z$ : a) cuando ángulo  $\phi_1$  del eje mayor de la elipse con la horizontal es  $\geq 90^\circ$ , b) cuando ángulo  $\phi_1$  del eje mayor de la elipse con la horizontal es  $< 90^\circ$ .

rotaciones terminan cuando; en tres ocasiones sucesivas no se ha logrado la detección, cuando se alcanza una mejor vista y excepcionalmente cuando se ha alcanzado una pose con  $\alpha = 0^\circ$  (debido a la restricción de sólo rotaciones en torno a los ejes  $X$  y  $Z$ ).

### 3. Resultados Experimentales

En los experimentos hemos usado un detector digital de rayos X (marca Canon, modelo CXDI-50G), tubo emisor de rayos X (marca Poskom, Modelo PXM-20BT), cabina de seguridad (de plomo) para aislar el entorno de inspección y un mecanismo de posicionamiento de objetos semi-automático (giroscopio). Se inspeccionaron 6 objetos (ver  $Obj_1, \dots, Obj_6$  en Fig. 5), cada uno de ellos con una hoja de afeitar.

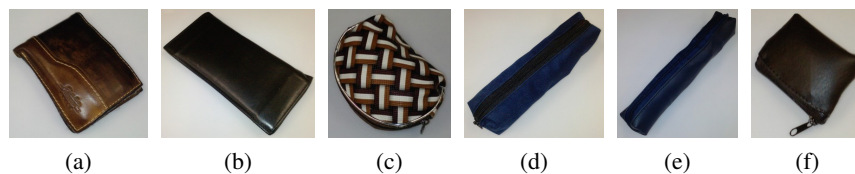


Figura 5: Objetos sometidos a inspección radioscópica: a)  $Obj_1$  billetera, b)  $Obj_2$  chequera, c)  $Obj_3$  monedero 1, d)  $Obj_4$  estuche, e)  $Obj_5$  estuche delgado, y d)  $Obj_6$  monedero 2.

Cada objeto fue inspeccionado 10 veces, en donde para cada vez la primera imagen fue tomada en una posición aleatoria. Un resumen de los resultados se encuentra en la Tabla 1. Como se aprecia, se lograron secuencias de hasta 4 imágenes de rayos X como máximo, sin embargo, algunas veces bastó la primera imagen para la detección. En algunos casos se pasó de la no-detección ( $ND$ ) –en una vista–, a la detección –en la siguiente–, lo cual es muy meritorio, ya que demuestra que el método es capaz

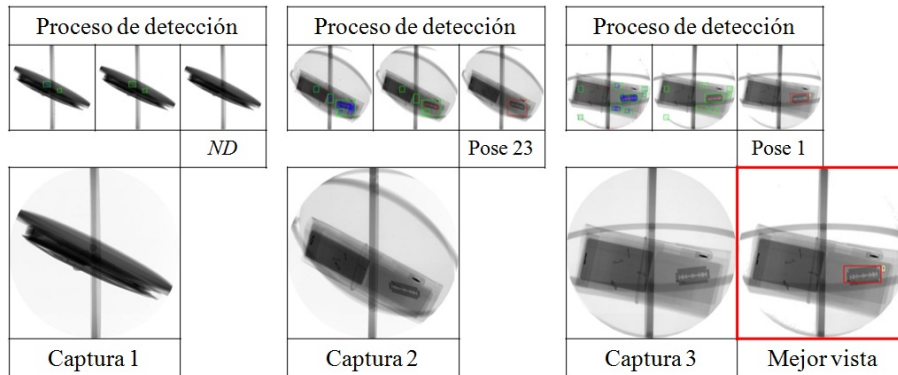


Figura 6: Objeto 2 sometido a inspección, ( $Obj_2$  en secuencia 10, ver Tabla 1).

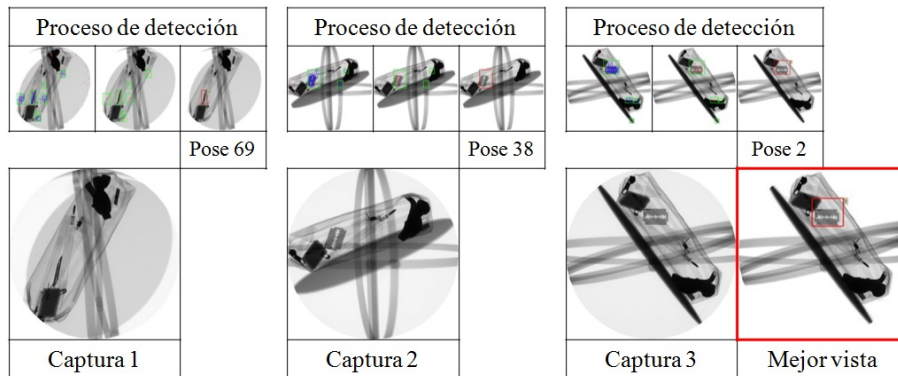


Figura 7: Objeto 4 sometido a inspección, ( $Obj_4$  en secuencia 5, ver Tabla 1).

de soslayar vistas iniciales inadecuadas y oclusiones. Sin embargo, debido a fuertes oclusiones y posiciones iniciales desfavorables, algunas veces la detección falló.

La Fig. 6 y la Fig. 7 muestran 2 objetos distintos que han sido inspeccionados, en ellas se aprecia las capacidades de nuestra propuesta para encontrar el objeto de interés con una mejor vista, se aprecia también que en algunos casos existen oclusiones parciales y que de igual manera es posible avanzar en la detección. A partir de estos resultados y los que están representados en forma numérica en la Tabla 1, podemos decir que el sistema de sujeción (giroscopio) tiene un alto grado de absorción a los rayos X, oscureciendo fuertemente las imágenes, esto en ocasiones causó algunas zonas de falsas alarmas,  $FP$ , asociadas comúnmente a la pose 2 y debido a la aparición de descriptores SIFT en dichas zonas. Para estimar el desempeño de nuestra propuesta, hemos calculado a partir de los datos registrados en la Tabla 1 los índices *precision* y *recall*, definidos como;  $precision = TP/(TP + FP)$  y  $recall = TP/P$ , donde,  $TP$  es la suma total de verdaderos positivos,  $FP$  es la suma total de falsos positivos y  $P$  corresponde al total de secuencias realizadas. Aquí se obtuvo  $TP = 52$ ,  $FP = 5$  con  $P = 60$ , y a partir de estos

valores se obtuvo:  $precision = 91,2\%$  y  $recall = 86,7\%$ , valores alentadores que nos indican cuánto realmente es capaz de detectar nuestro sistema, lo cual es comparativamente superior a la eficiencia de los inspectores humanos en procesos de inspección visual, que pueden alcanzar como máximo el  $80\%$  [7].

Tabla 1: Secuencias de inspección para distintos objetos e indicadores de desempeño.

Objeto	Secuencia	Pose		Rotación		Pose		Rotación		Pose		Desempeño		
		X	Z	X	Z	X	Z	X	Z	X	Z	TP	FP	
Obj1	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
Obj2	1	3	-	-	-	-	-	-	-	-	-	-	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
Obj3	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
Obj4	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
Obj5	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
Obj6	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
	6	2FP	-	-	-	-	-	-	-	-	-	-	0	1
	10	ND	-40°	-	2	-	-	-	-	-	-	-	1	0
TOTAL:												52	5	
recall = TP/P:												86,7%		
precision = TP/(TP + FP):												91,2%		

ND: Ninguna Detección, TP: Verdadero Positivo, FP: Falso Positivo, P: suma total de secuencias.

## 4. Conclusiones

En este trabajo nosotros hemos propuesto un framework de inspección visual radioscópica que permite hacer inspección de objetos que contengan algún objeto de interés. Los resultados de  $precision = 91,2\%$  y  $recall = 86,7\%$  nos indican que estos son

los primeros pasos para lograr sistemas de inspección con rayos X, totalmente automatizados. El *framework* que hemos propuesto lo hemos ejemplificado con la detección de una hoja de afeitar, que en lugares como aeropuertos y centros aduaneros es considerado peligroso, si bien la detección de una hoja de afeitar pudiese parecer simple, la evidencia conceptual y experimental nos indicó que la simetría en todos los cuadrantes, la delgadez y pequeñez resultaron ser problemas difíciles de solventar, debido a la naturaleza de absorción de los rayos X y a la invarianza a la rotación que presentan los descriptores SIFT. Finalmente todas las soluciones de caracterización, rotación y estrategias de inspección han sido desarrolladas pensando en una solución genérica y adaptable a cualquier objeto de interés que requiera ser detectado. Como idea de trabajo futuro pensamos que un sistema calibrado y con dos o tres vistas iniciales podría resolver el tema de oclusiones indeseadas y con esto robustecer el sistema de inspección.

## Referencias

1. Dutagaci, H., Cheung, C.P., Godil, A.: A benchmark for best view selection of 3d objects. In: Proceedings of the ACM workshop on 3D object retrieval. pp. 45–50. 3DOR '10, ACM, New York, NY, USA (2010)
2. Farshidi, F., Sirouspour, S., Kirubarajan, T.: Robust sequential view planning for object recognition using multiple cameras. *Image and Vision Computing* 27(8), 1072 – 1082 (2009)
3. Haff, R.P., Toyofuku, N.: X-ray detection of defects and contaminants in the food industry. *Sensing and Instrumentation for Food Quality and Safety* 2(4), 262–273 (2008)
4. Lowe, D.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2), 91–110 (2004)
5. Mery, D., Filbert, D.: Automated flow detection in aluminum castings based on the tracking of potential defects in a radioscopic image sequence. *IEEE Transactions on Robotics and Automation* 18(6), 890–901 (2002)
6. Mery, D.: 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)
7. Newman, T., Jain, A.: A survey of automated visual inspection. *Computer Vision and Image Understanding* 61(2), 231–262 (1995)
8. Quinn, R., Sigl, C., Company, E.K.: Radiography in modern industry, vol. 4. Eastman Kodak (Rochester, NY) (1980)
9. Richter, H.U.: Chronik der Zerstörungsfreien Materialprüfung. Deutsche Gesellschaft für Zerstörungsfreie Prüfung, DGZfP, Verlag für Schweißen und verwandete Verfahren, DVS-Verlag GmbH, Berlin, 1 edn. (1999)
10. Silva, R., Mery, D.: State-of-the-art of weld seam inspection using X-ray testing: Part I – image processing. *Materials Evaluation* 65(6), 643–647 (2007)
11. Zentai, G.: X-ray imaging for homeland security. *Imaging Systems and Techniques*, 2008. IST 2008. IEEE International Workshop on pp. 1–6 (Sept 2008)