

INSPECCIÓN VISUAL AUTOMÁTICA USANDO ANÁLISIS TRIFOCAL EN UNA SECUENCIA DE IMÁGENES NO CALIBRADAS

Miguel A. Carrasco & Domingo Mery

Departamento de Ciencia de la Computación
Pontificia Universidad Católica de Chile
Av. Vicuña Mackenna 4860(143), Santiago de Chile
e-mail: {mlcarras@puc.cl; dmery@ing.puc.cl}

Resumen: La Inspección Automática usando Múltiples Vistas (AMVI) ha sido desarrollada recientemente para detectar fallas automáticamente en objetos manufacturados. AMVI ha sido implementada con éxito en secuencias de imágenes calibradas donde se conoce con precisión la función de transferencia 3D \rightarrow 2D de la proyección de las vistas, sin embargo, su aplicación en ambientes industriales es una tarea compleja debido a las inestabilidades propias del sistema. Esta investigación propone una nueva estrategia, basada en la detección de fallas en una secuencia de imágenes no calibradas. La metodología diseñada consiste en construir un modelo y realizar un análisis trifocal que determine la posición real de una falla, utilizando puntos de control correspondientes en la secuencia. Los resultados experimentales demuestran la potencialidad en la detección de defectos en imágenes no calibradas, detectando todos los defectos que permanecen en la relación trifocal.

Keywords: Computer vision, multiple view geometry, automated visual inspection, defect detection.

1 Introducción

La calidad de los productos manufacturados es uno de los principales objetivos en el proceso productivo. Uno de los factores más relevantes respecto de la calidad, tiene relación con el cumplimiento de estándares de seguridad impuestos por el fabricante, o bien por alguna norma reguladora. Si bien, algunos fabricantes toleran productos con fallas, existen otros, donde la seguridad tiene un rol crítico. Tales como equipos de alta presión, contenedores químicos, llantas de aluminio, etc. La existencia de fallas en estos, pueden causar graves accidentes. Una de las formas convencionales para detectarlas, utiliza rayos X o rayos Gamma, los cuales penetran el material generando una imagen radiológica al ser expuesta sobre una placa fotográfica o sensor fotoeléctrico. Las fallas quedan reveladas debido a la variación de intensidad respecto al material analizado. El análisis e interpretación de estas imágenes es realizado por un inspector en forma manual, sin embargo, el proceso es deficiente, ya que toma una gran cantidad de tiempo, y es inconsistente, ya que depende del inspector para su análisis [1].

La Inspección Visual Automática (AVI) ha solucionado en gran parte el control de calidad, estableciendo políticas de control precisas y objetivas [2]. Mediante técnicas de procesamiento de imágenes y reconocimiento de patrones, permite detectar fallas en los

productos asegurando dos condiciones básicas del proceso productivo; eficiencia y rapidez. Sin embargo, la mayoría de estas técnicas utilizan el procesamiento individual sobre cada falla de la imagen. Implicando un posterior proceso de análisis sobre las características individuales de cada falla. Luego, a través del reconocimiento de patrones, determinamos si la falla es efectivamente una falsa alarma o una falla real.

Recientemente, se ha desarrollado una nueva metodología para detectar fallas automáticamente denominada Inspección Automática usando Múltiples Vistas (AMVI) [3]. En forma equivalente al proceso de detección de fallas que realiza un inspector, la metodología AMVI detecta las fallas utilizando los siguientes dos pasos. El primer paso, denominado *identificación*, consiste en detectar todas las regiones anómalas o fallas hipotéticas en cada imagen sobre una secuencia de movimientos del objeto, sin conocimiento a priori sobre la estructura de éste; El segundo paso, denominado *seguimiento o tracking*, consiste en seguir en cada imagen de la secuencia las fallas hipotéticas detectadas en el primer paso. Si las fallas hipotéticas continúan a lo largo de la secuencia de imágenes, se identifica la falla hipotética como una falla real, y el objeto es clasificado como defectuoso. En cambio, si las fallas hipotéticas no tienen correspondencias en la secuencia serán consideradas como falsas alarmas. La

metodología AMVI tiene como principal fundamento el hecho que sólo las fallas reales (y no las falsas alarmas) pueden ser observadas a lo largo de la secuencia de imágenes, ya que éstas permanecen estables en la posición relativa al movimiento del objeto. Por lo tanto, teniendo dos o más vistas de un mismo objeto, sobre diferentes puntos de vista, es posible mejorar el rendimiento en la detección de fallas reales.

Esta estrategia original presentada en [3], requiere una calibración previa del sistema de adquisición de secuencias de imágenes. En la calibración se pretende establecer la función de transferencia que proyecta un punto 3D del objeto en un punto 2D de la imagen. Lamentablemente, el proceso de calibración es difícil de realizar en ambientes industriales ya que existen vibraciones y movimientos aleatorios que varían con el tiempo y no son considerados en la función de transferencia original estimada. Un primer método alternativo para la realización de la estrategia AMVI en secuencias no calibradas se presentó en [4] para secuencias de dos imágenes. Sin embargo, debido a que la robustez de la metodología AMVI aumenta con el número de imágenes analizadas en la secuencia, en el presente trabajo se propone una modificación al sistema robusto diseñado en [4] procesando tres imágenes en vez de dos. Adicionalmente, en este trabajo se ha aumentado la cantidad de puntos correspondientes de control que relacionan las imágenes utilizando curvas B-Spline, lo que mejora significativamente la estimación del modelo multi-focal necesario para hacer el *tracking*.

El resto de este documento se divide en las siguientes secciones. La sección 2 incluye los trabajos relacionados sobre la metodología AMVI. La sección 3, dedicada al método propuesto, incluye una descripción de la metodología empleada para segmentar fallas hipotéticas, estimar la matriz fundamental en forma robusta, generar puntos de control artificiales y estimar los tensores trifocales. La sección 4 incluye los resultados experimentales. Finalmente, la sección 5 presenta las conclusiones y líneas de trabajo futuro.

2 Antecedentes

El principal objetivo de AMVI es seguir sólo las fallas hipotéticas, y no estimar la estructura del objeto. Inicialmente, la metodología fue implementada para automatizar la inspección de llantas de aluminio utilizando una secuencia de imágenes en un sistema calibrado [3]. En este caso, la fase de calibración del objeto se generó fuera de línea. Luego se elaboró un modelo de proyección para rastrear las fallas a lo largo

de la secuencia de imágenes, utilizando los principios de la geometría de múltiples vistas [5]. Los resultados obtenidos demostraron la factibilidad técnica para detectar la totalidad de fallas reales, junto con una alta tasa en la discriminación de falsas alarmas. Sin embargo, en ambientes industriales, la calibración es un proceso complejo, producto de las vibraciones en la adquisición de las imágenes del objeto, implicando una falta de precisión para estimar los parámetros del modelo geométrico de múltiples vistas.

La investigación desarrollada en [4], presentó un método robusto alternativo que no requiere calibrar el sistema de adquisición de imágenes. Uno de los principales factores para estimar el movimiento del objeto, se debe a que éste es un cuerpo rígido que tiene un movimiento de rotación y/o traslación con velocidad constante y una trayectoria suave. El método presentado en [4] busca regiones significativas de la estructura del objeto a analizar que estén presentes a lo largo de la secuencia de imágenes. Una vez establecida la correspondencia entre puntos de estas regiones se procede a construir un modelo de dos vistas que sirva para establecer la correspondencia posterior entre las fallas hipotéticas.

El modelo propuesto utilizó la siguiente metodología: primero, identificar los puntos estructurales en cada imagen de la secuencia; segundo, encontrar la correspondencia de los puntos del primer paso entre imágenes consecutivas; tercero, estimar en forma robusta la matriz fundamental [5] de los posibles puntos correspondientes. Una vez finalizado este proceso, se dispone de un modelo matemático que relaciona el movimiento entre pares de imágenes. La siguiente fase, relaciona las fallas hipotéticas de ambas imágenes. Utilizando la proyección de línea epipolar [5], buscamos en ésta las posibles fallas que puedan concordar con las propiedades de la falla hipotética de la imagen anterior. Esta evaluación se realiza a través de la distancia Euclidiana sobre el espacio de propiedades de área e intensidad. En caso que la falla no tenga una correspondencia en la siguiente imagen, se rechaza esta última, siendo considerada como falsa alarma ya que no cumple con la restricción epipolar.

Existen dos factores relevantes sobre el proceso desarrollado en [4]. Primero, la fase de segmentación está diseñada para detectar la mayoría de los defectos sin un conocimiento a priori del material y/o posición del objeto. Dentro de las fases del procesamiento de imágenes, la segmentación se encuentra en una etapa inicial, no obstante, tiene un papel preponderante en todo el proceso, ya que un rendimiento deficiente en esta fase puede causar una pobre predicción en la

detección de fallas reales [6]. Segundo, el sistema de calibración debe evaluar sólo los puntos de control que sean correspondientes. Sin embargo, no existe una correspondencia perfecta entre ambas imágenes, debido a distorsiones geométricas u otras anomalías en el proceso de captura. Por este motivo, se ha empleado un algoritmo robusto que descarta las posiciones que contengan un error en la proyección. Empleando sólo un conjunto de pares relacionados que minimice el error entre la posición real y la proyectada, según una métrica de distancia Euclidiana. El método anterior es conocido como aproximación de RANSAC [7]. El término robusto, tiene relación con la flexibilidad para determinar un mínimo conjunto de coordenadas relacionadas, y descartar aquellas que no cumplen con un error mínimo permitido entre la proyección y la posición real.

3 Método propuesto

Los resultados en laboratorio han demostrado un muy buen rendimiento de AMVI en la detección de fallas en estructuras de aluminio en ambientes calibrados [3]. Sin embargo, en ambientes industriales la calibración es un proceso complejo y de alto costo para el fabricante. Esta sección, presenta un nuevo método AMVI propuesto para la detección automática de fallas usando un análisis trifocal en secuencias de imágenes no calibradas, perfeccionando así, el método diseñado por los mismos autores en [4]. La mejora se debe a que hemos extendido el análisis de dos a tres imágenes, estimando los tensores trifocales en forma robusta, e incrementado los puntos de control en forma artificial para establecer las correspondencias entre las vistas. A continuación presentamos los pasos de nuestro nuevo método:

1. Para cada una de las tres imágenes de la secuencia (A, B y C), buscar las fallas hipotéticas utilizando el algoritmo de segmentación *crossing line profile* [8].
2. Para todas las estructuras del objeto de análisis en las imágenes A, B, y C, buscar las relaciones entre las estructuras, y generar puntos de control artificiales con curvas B-Spline.
3. Estimar la matriz fundamental entre las imágenes A y B utilizando el método RANSAC [7] con los puntos estructurales y artificiales encontrados en el paso 2.
4. Para todas las fallas hipotéticas de A encontradas en el paso 1, generar la proyección epipolar con la matriz fundamental en la imagen B, y recorrer la

línea epipolar en busca de fallas hipotéticas encontradas en el paso 1 en la imagen B:

- a. Si existe más de una falla hipotética en la línea epipolar, buscar la mejor relación a base de las propiedades de área e intensidad relacionada a través de la menor distancia Euclidiana, y almacenar dicha relación.
 - b. En caso de que la línea epipolar no pase por una falla hipotética en la imagen B, quiere decir que no existe una falla proyectada en la imagen B y se descarta la falla hipotética de la imagen A.
 - c. Si existe sólo una falla contenida en la línea epipolar que cumpla con la distancia Euclidiana mencionada, entonces se almacena la relación de la falla entre A y B.
5. Estimar los tensores trifocales entre las imágenes A, B y C utilizando RANSAC con los puntos estructurales y artificiales encontrados en el paso 2.
 6. Para todas las relaciones de falla entre A y B encontradas en el paso 4, buscar la posición del centro de masa de las fallas hipotéticas y re-proyectar estas posiciones utilizando el tensor trifocal:
 - a. Si existe una falla proyectada que esté a una distancia mínima a la falla hipotética en la imagen C del paso 1, asignar dicha posición como matching en tres vistas y se determina así la detección de una falla en la secuencia.
 - b. Si no existe una falla en la imagen C que esté relacionada con la proyección, entonces eliminar dicha falla hipotética y considerarla como falsa alarma.

A continuación, presentamos una explicación de estos pasos.

3.1 Identificación de fallas hipotéticas

La segmentación de fallas hipotéticas permite identificar regiones en cada imagen de la secuencia que pueden corresponder a fallas reales (Fig. 1). Existen dos características generales que son utilizadas para identificarlas: i) una falla se considera como un subconjunto conectado dentro de la imagen, ii) la diferencia entre los niveles de gris de la falla y sus vecinos es considerable. Inicialmente, la identificación es llevada a cabo a través de un proceso sin conocimiento a priori utilizando la convolución de la imagen con un núcleo *Laplacian-of-Gaussian* (LoG) y posteriormente analizando los cruces por cero [6]. El operador LoG intrínsecamente utiliza un filtro pasa

bajos gaussiano para disminuir los niveles de ruido en la imagen. El resultado del operador y el cruce por cero, genera una imagen binaria que contiene las fallas reales con contornos conectados. Sin embargo, no siempre los contornos de las fallas quedan totalmente cerrados. Esto sucede cuando se encuentran cercanos al borde de una estructura regular (Fig.1c). La solución consiste en aumentar los bordes de las estructuras regulares. Este procedimiento consiste en calcular el gradiente de la imagen para identificar dichas posiciones (Fig.1d) y luego, generar una imagen binaria que emplee sólo los niveles de mayor energía del gradiente (Fig.1e). Una vez que cada región cerrada es segmentada, son extraídas características a través del perfil de tonos de

gris de una línea recta que pase por el centro de la región segmentada, siendo identificadas como fallas hipotéticas aquellas que presenten un perfil de gran varianza (ver detalles en [8]). Este detector de fallas hipotéticas contiene un alto nivel de falsas alarmas, sin embargo, tiene las siguientes ventajas: i) el mismo detector se aplica para todas las imágenes, ii) permite identificar las fallas hipotéticas, independiente de la posición o de la estructura del objeto estudiado, es decir, sin conocimiento a priori del diseño de la estructura, iii) la detección de las fallas reales es muy alta (mayor al 90%).

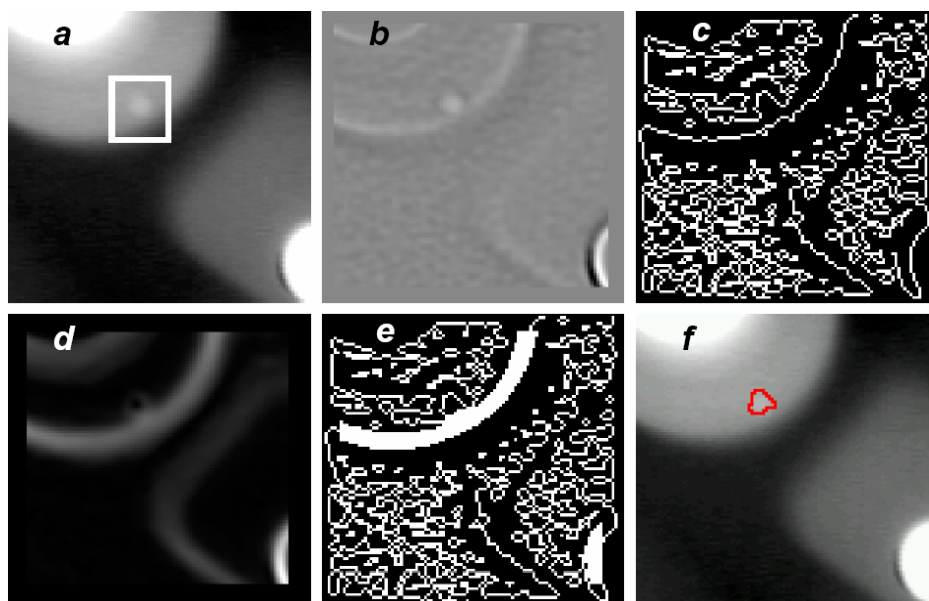


Fig.1. Detección de fallas: a) sección de imagen radiológica con una falla inscrita en el borde de una estructura regular, b) aplicación del filtro laplaciano sobre la imagen con $\sigma = 1.25$ píxeles ($kernel = 11 \times 11$), c) imagen binaria con cruces por cero, d) gradiente de la imagen, e) detección de bordes, después de aumentar los bordes los niveles máximos del gradiente, y f) detección de fallas utilizando la varianza del cruce de línea del contorno.

3.2 Identificación de puntos de control

El seguimiento de regiones es parte primordial del proceso AMVI ya que permite modelar las trayectorias de todas las posiciones en una secuencia de imágenes. Dicho modelo utiliza la posición de cada defecto en cada imagen y la compara con la posición modelada, reduciendo así el número de falsas alarmas. En esta fase del proceso, determinamos la matriz fundamental en forma automática, utilizando los cambios de posición de cada región para cada par de imágenes.

Debido a que el sistema no está calibrado, es necesario que el seguimiento utilice las características morfológicas de cada imagen. Este proceso realiza los siguientes pasos: i) se construye una imagen binaria

utilizando el método de Otsu [7], de esta forma se aíslan las regiones del fondo de la imagen; ii) se eliminan todas aquellas regiones que estén conectadas con el borde de la máscara, ya que éstas tienen una variación en el área que las contiene. El objetivo consiste en obtener la información de posición de sólo aquellas regiones que tienen una mínima variación de área y desplazamiento, iii) para cada región, calcular la posición del centro de masa, los puntos externos, verificando que los desplazamientos y el área concuerde con la región en la siguiente imagen. Además, introducir nuevos puntos de posición generados artificialmente a través de una curva B-Spline cuadrática, la cual utilice los centros de masa como puntos de control.

A diferencia de la investigación realizada en [4], disponemos de un reducido grupo de estructuras relacionadas en las tres vistas. Por este motivo, se han introducido puntos artificiales (con resolución subpíxel), generados a través de las curvas B-Spline. Una de las mayores ventajas de este tipo de curvas, es que está contenida en el casquete convexo formado por los puntos de control. Estos últimos corresponden a los centros de masa de las estructuras válidas en las tres vistas. El resultado de la interpolación, genera una curva suave inscrita en la región formada por los centros de masa (Fig. 3).

La construcción de la curva B-Spline tiene como supuesto, que la posición de los centros de masa para cada estructura tiene un alto grado de precisión en cada vista. En caso de existir un error en dicha posición, todos los puntos artificiales incluirán este error en la estimación de la curva. Bajo esta condición, se penaliza los cambios considerables en el área de las estructuras correspondientes, ya que esto implica un cálculo incorrecto en la posición del centro de masa.

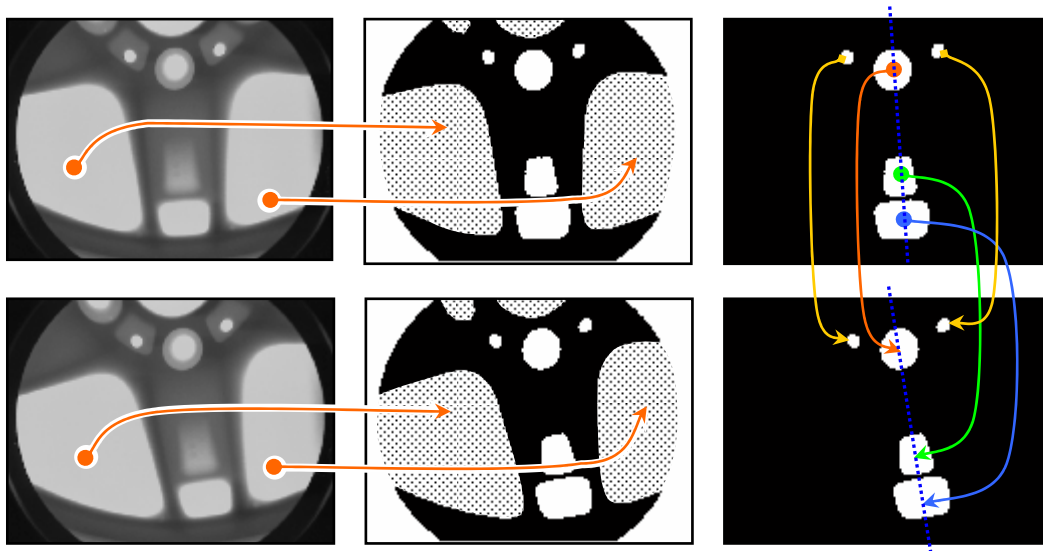


Fig. 2. Selección de posiciones ideales: izquierda) imágenes originales, centro) segmentación después de aplicar el método de Otsu, derecha) determinación de regiones internas sin variación de área.

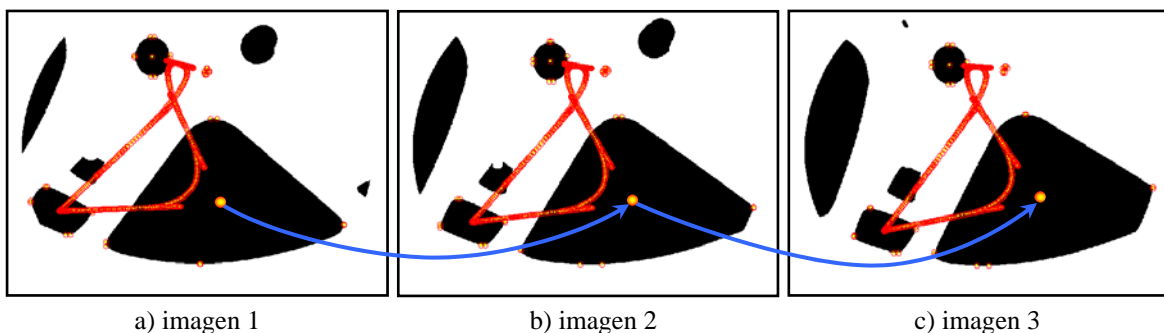


Fig. 3. Matching de las posiciones relativas de cada estructura. Líneas azules) matching de las posiciones de los centros de masa. Líneas rojas) curva B-Spline cuadrática utilizando centros de masa como punto de control de las estructuras válidas en la largo de la secuencia.

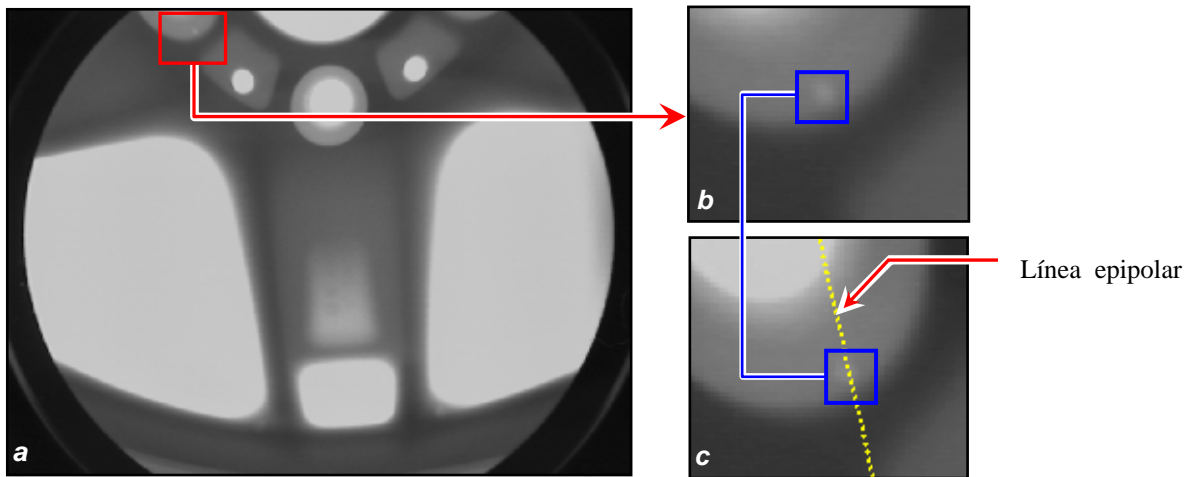


Fig. 4. Línea epipolar generada automáticamente a partir de la matriz fundamental: a) imagen original, b) sección de la imagen original con falla, c) línea epipolar sobrepuesta sobre la imagen par.

3.3 Estimación de la matriz fundamental

Utilizando las posiciones intrínsecas de cada región en ambas imágenes y los puntos generados artificialmente, emplear el algoritmo robusto RANSAC [7] para estimar la matriz fundamental. Debemos recordar que existe la probabilidad de error entre la posición de una región y su respectivo par correspondiente, sin embargo, este error es mínimo debido a que RANSAC utiliza el conjunto de puntos que obtenga la mejor estimación de la matriz fundamental.

La matriz fundamental es vital para el proceso AMVI ya que relaciona cualquier posición en pares de imágenes. La precisión de ésta permite determinar correctamente fallas hipotéticas a lo largo de la línea epipolar, no obstante, si la matriz fundamental no es robusta, la línea epipolar será incorrecta y el posterior proceso de tensores trifocales queda errado. Establecer coordenadas correspondientes en forma precisa, permite aumentar la precisión del cálculo.

Existen dos factores que perturban el rendimiento de la matriz fundamental en forma automática. Primero, las regiones de interés para el cálculo de posición, son aquellas que permanecen invariantes ante el movimiento del objeto. Todas aquellas regiones que tienen una variación en el área, no son utilizadas ya que la posición del centro de masa y los extremos varían de acuerdo al movimiento generado por la rotación y/o traslación (Fig. 2). Segundo, las posiciones de centro de masa y extremas son calculadas para cada región tal que cumpla con el paso anterior, sin embargo, existe la probabilidad de error en el emparejamiento de las posiciones relacionada entre sí.

Por este motivo, RANSAC considera las mejores n posiciones del total de puntos ingresados. En caso que el centro de masa tenga una variación importante en la posición relativa de la siguiente imagen, es posible estimar erróneamente la matriz fundamental, y generar un error en el posterior proceso de cálculo del tensor trifocal. Afortunadamente, no es relevante para el proceso que todos los puntos concuerden en forma exacta. Siendo ésta la mayor ventaja del algoritmo robusto para estimar la matriz fundamental.

3.4 Evaluación en dos imágenes

Para determinar si una falla es correspondiente en la siguiente imagen, se utiliza una máscara que recorra la línea epipolar comparando ésta con la región segmentada de la imagen original. Una vez determinada una falla hipotética (Fig. 4a), ésta debe encontrarse a lo largo de la línea epipolar (Fig. 4b). Si la región se encuentra en ambas imágenes, entonces se considera como falla hipotética, en caso contrario, la región es descartada. Experimentalmente, una misma línea epipolar puede contener más de una falla hipotética. En este caso, establecemos un grado de similaridad para cada falla hipotética, bajo el supuesto que todas éstas cumplen con la restricción epipolar. El criterio compara las características de área e intensidad para cada falla (Fig 4.b). La similaridad se establece cuando dos fallas (una en la imagen A, y la otra contenida en la línea epipolar de la imagen B) están a una distancia mínima en el espacio de las propiedades, siendo utilizada la distancia Euclidiana como métrica de similitud. Si aún permanecen falsas alarmas luego del proceso de análisis de dos vistas, es posible utilizar el tensor trifocal para eliminar las falsas alarmas restantes.

3.5 Estimación de los tensores trifocales

El análisis trifocal permite modelar todas las relaciones geométricas en tres vistas siendo independiente de la estructura contenida en cada imagen [5]. El tensor depende sólo del movimiento entre las vistas y de los parámetros internos de las cámaras, y es definido únicamente por las vistas de las matrices. Sin embargo, puede ser calculado a través de las correspondencias de las imágenes, sin conocimiento a priori del movimiento o calibración del objeto. Esta característica justifica su utilización, ya la estimación de la matriz fundamental no siempre elimina la totalidad de las falsas alarmas. Un análisis en tres vistas aumenta la probabilidad de formar tríos que cumplen la condición trifocal [9].

La reproyección del tensor permite estimar la posición de un conjunto de puntos en la tercera vista utilizando las dos primeras vistas. Este proceso tiene un buen rendimiento en sistemas calibrados, sin embargo, dado que nuestro objetivo es modelar la proyección en sistemas no calibrados, utilizamos el mismo criterio aplicado en la estimación de la matriz fundamental. Siendo calculado el tensor a través del algoritmo RANSAC [7]. Los principios de este algoritmo son similares para el cálculo de la matriz fundamental, modificando en forma particular la métrica de distancia

aplicada en el análisis trifocal. A diferencia en el cálculo de la línea epipolar, el tensor permite determinar una reproyección en la tercera vista, utilizando las posiciones correspondientes en las dos primeras vistas (Fig. 3a y 3b).

3.6 Evaluación en tres vistas

El criterio para aceptar una falla hipotética como falla real, sólo debe verificar que la posición entre la posición reproyectada en la tercera vista se encuentre contenida en un rango de distancia mínima. En caso que las fallas segmentadas en la tercera vista no concuerden con la proyección del tensor, se descarta la falla hipotética.

4 Resultados experimentales

Esta sección, presenta los resultados de los experimentos realizados en una secuencia de imágenes radiológicas de llantas de aluminio (Fig. 5). Separamos el análisis en dos pasos; primero, detectar pares de fallas utilizando la estimación de la matriz fundamental en dos vistas; segundo, utilizando el resultado anterior, reproyectamos los pares de fallas hipotéticas en la tercera vista utilizando la estimación del tensor trifocal. Ambos pasos se detallan a continuación.

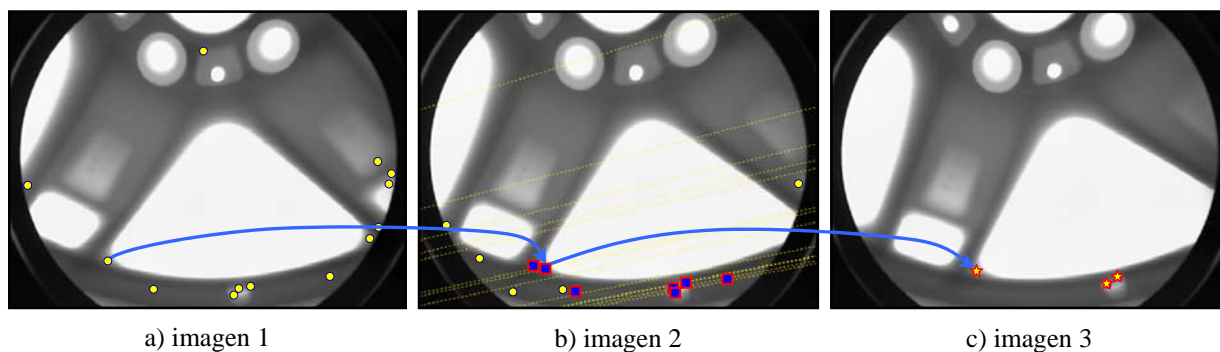


Fig. 5. Proceso general de estimación de las fallas, a) segmentación de fallas hipotéticas, b) estimación de la línea epipolar utilizando la matriz fundamental, c) proyección de las coordenadas de la imagen 1 y 2, utilizando los tensores trifocales estimados con RANSAC, sobre la imagen 3

Tabla 1. Rendimiento de la detección de fallas en dos vistas a través de la Matriz Fundamental en 24 imágenes (12 imágenes estéreo).

| Imagen | Fallas Hipotéticas | Fallas Reales | Fallas S/Corr. | Matching | Falsas Alarmas | Detección | Detección S/Corr. |
|--------------|--------------------|---------------|----------------|----------|----------------|-----------|-------------------|
| Izquierda | 14.46 | 4.92 | 0.62 | 4.31 | 20.2% | 100% | 87.5% |
| Derecha | 13.23 | 4.92 | 0.62 | 4.31 | 22% | 100% | 87.5% |
| Promedio | 13.85 | 4.92 | 0.62 | 4.31 | 21.1% | 100% | 87.5% |
| Total fallas | 180 | 64 | 8 | 56 | 35 | | |

(S/Corr: Sin Correspondencias)

Tabla2. Rendimiento de la detección de fallas en tres vistas utilizando el Tensor Trifocal en 66 imágenes (22 tríos).

| Imagen | Fallas Hipotéticas | Fallas Reales | Fallas S/Corr. | Matching | Falsas Alarmas | Detección | Detección S/Corr. |
|--------------|--------------------|---------------|----------------|----------|----------------|-----------|-------------------|
| Izquierda | 7.14 | 2.82 | 1.27 | 1.55 | 4.5% | 100% | 54.8% |
| Centro | 6.82 | 2.68 | 1.14 | 1.55 | 4.7% | 100% | 57.6% |
| Derecha | 6.68 | 2.82 | 1.27 | 1.55 | 4.8% | 100% | 54.8% |
| Promedio | 6.88 | 2.77 | 1.23 | 1.55 | 4.6% | 100% | 55.8% |
| Total fallas | 454 | 183 | 81 | 102 | 21 | | |

(S/Corr: Sin Correspondencias)

4.1 Rendimiento en dos vistas

El primer paso, es evaluar el rendimiento del algoritmo en dos vistas, utilizando el método propuesto en la sección 3.4. Este consiste en determinar las fallas correspondientes entre dos imágenes de una secuencia, a través de la búsqueda de fallas en la línea epipolar. El estudio se aplicó a doce pares de imágenes, es decir, a 24 imágenes radiológicas de 578×768 píxeles.

Los resultados indican que la detección de fallas reales en secuencia de dos vistas es perfecta. Es decir, el modelo detecta todas las fallas reales que tienen una correspondencia en la siguiente imagen, contenidas en la línea epipolar. No obstante, debido a que el proceso de segmentación no ha detectado el 100% de las fallas reales, existe un conjunto de fallas sin correspondencia entre los pares de imágenes. En promedio, existen 0.62 fallas reales por imagen sin correspondencia. De estos, sólo 56 pares formaron un matching, de un total de 64 fallas reales, con un 87.5% de rendimiento en promedio. Además, el estudio reveló una tasa de 21.1% de falsas alarmas, es decir, sólo 2.92 falsas alarmas por imagen (Tabla 1). Aunque el porcentaje sea elevado, no penalizamos las falsas alarmas, ya que éstas pueden ser reducidas utilizando una tercera imagen.

Por lo tanto, la estimación del modelo de movimiento en dos vistas confirma los resultados generados en [4]. En general, es posible determinar con precisión la matriz fundamental para cada par de imágenes debido a: i) la mayoría de las posiciones de cada región concuerda con la rotación y/o traslación de la siguiente imagen; ii) se consideran sólo las regiones que sean similares tanto en área como en posición dentro de la imagen, eliminando posibles errores de *matching*; iii) RANSAC utiliza sólo los mejores siete pares de puntos dentro del total de posiciones en todas las regiones; iv) se asegura que siempre el centro de masa de cada región corresponde con su par en la siguiente imagen, aumentando la precisión en la estimación.

4.2 Rendimiento en tres vistas

El segundo paso, utiliza el algoritmo propuesto en la sección 3.6. Luego de completar la detección de

posibles pares de fallas en dos imágenes, extendemos la detección de las fallas en la tercera imagen de la secuencia. En este caso, el estudio se aplicó a 22 tríos de imágenes, es decir, 66 imágenes radiológicas.

Al igual que el paso anterior, ocurrieron los mismos problemas de falta de correlación entre las fallas reales en secuencia. En efecto, de las 181 fallas reales, sólo 102 fallas tienen una correspondencia en las tres imágenes, esto debido a problemas en la segmentación inicial, o bien porque una falla es vista sólo en una o dos imágenes (y no en las tres). Utilizando el conjunto de 102 fallas en correspondencia, el tensor trifocal tuvo un rendimiento perfecto. Es decir, la proyección del tensor tiene una correspondencia en la tercera imagen, muy cercana a la posición real al emplear las coordenadas correspondientes en la primera y segunda imagen. En cambio, si consideramos el total de fallas reales, (con y sin correspondencias), el rendimiento bajó a un 55.8%. El estudio reveló una considerable baja en la detección de falsas alarmas. Los resultados indican que sólo existen 0.32 falsas alarmas por imagen, y un total de 21/458 fallas hipotéticas, lo cual representa un 4.6% de falsas alarmas.

Los experimentos demuestran que es muy baja la probabilidad que una falsa alarma se encuentre en tres vistas. Sólo las fallas reales que tienen una correspondencia perfecta en las tres imágenes, pueden ser encontradas con este tipo de técnica.

Para aumentar el rendimiento general, tanto en dos y tres vistas, debemos incrementar el número de fallas hipotéticas en la fase de segmentación. En algunos casos, comprobamos que la estimación de la proyección epipolar fue correcta, sin embargo, en la siguiente imagen, el proceso de segmentación no generó fallas correspondientes con las regiones anteriores. De esta forma, no fue posible generar un par de fallas en correspondencia. Recordemos que nuestro análisis trifocal, sólo admite pares de fallas contenidas en las dos primeras imágenes de cada secuencia, para estimar la proyección en la tercera imagen.

Un resumen de los contenidos de las tablas 1 y 2, es dado a continuación: i) *Fallas Hipotéticas*, es el promedio de regiones detectadas en la fase de segmentación por cada imagen (más detalles en [8]); ii) *Fallas Reales*, es el promedio de regiones detectadas como fallas reales, utilizando un proceso de detección manual; iii) *Fallas S/Corr.*, es el promedio de fallas reales que no tiene correspondencias entre pares o tríos de imágenes, es decir, fallas reales que no estén contenidas en la siguiente imagen de la secuencia; iv) *Matching*, es el promedio de fallas reales que tiene una correspondencia total entre pares o tríos de imágenes, utilizando el método propuesto en la sección 3; v) *Falsas Alarmas*, es el porcentaje de regiones que cumpliendo la condición epipolar y/o trifocal, no son fallas reales; vi) *Detección*, es el porcentaje de detección de fallas reales que cumplan la condición epipolar y/o trifocal. Se expresa como la relación porcentual entre: $\text{Matching} / (\text{Fallas Reales} - \text{Fallas S/Corr.})$; vii) *Detección S/Corr.*, es el porcentaje que relaciona el total de matching / Fallas Reales.

5 Conclusiones

Esta investigación presenta el desarrollo de un nuevo algoritmo de detección de fallas en productos manufacturados, en una secuencia de imágenes no calibradas. Utilizando la nueva metodología AMVI [3], hemos diseñado un nuevo sistema de calibración automático basado sólo en las posiciones espaciales de las estructuras. El enfoque propuesto, utiliza la proyección epipolar, generada por la matriz fundamental y los tensores trifocales en forma robusta, con el objetivo de construir un modelo de movimiento sin conocimiento a priori de su estructura. Basamos nuestra investigación en el supuesto que las fallas hipotéticas son fallas reales, si son correspondientes en posición en una secuencia de imágenes, ya que éstas permanecen estables en su posición relativa al movimiento del objeto.

Respecto la investigación desarrollada en [4], hemos extendido el análisis de dos a tres imágenes por secuencia, a través de la estimación de los tensores trifocales. Además, hemos introducido nuevos puntos de control generados artificialmente a través de las curvas B-Spline, debido a la baja cantidad de estructuras que permanecen estables en tres imágenes de una secuencia.

Los resultados indican que es posible generar un modelo automático para una secuencia de imágenes que represente el movimiento entre los puntos y regiones contenidos en éstas. La posibilidad de introducir puntos de control no correspondientes en

tríos de imágenes, es la principal ventaja del algoritmo RANSAC para estimar en forma robusta la matriz fundamental y los tensores trifocales. De esta forma, podemos utilizar como puntos referencia, los extremos de las estructuras, los centros de masa, y los puntos artificiales, suponiendo que siempre existe un mínimo conjunto de puntos que genera el menor error en la proyección.

La principal ventaja de nuestro modelo, es la estimación automática del movimiento. Sin embargo, en algunos casos experimentamos errores en la proyección debido a distorsiones geométricas en la adquisición de las imágenes. Nuestro trabajo futuro es estimar fallas en secuencias de cuatro imágenes, e introducir al modelo la corrección de las distorsiones ópticas.

6 Referencias

- [1]. Liao, T. W.; Li, Y.: “*An Automated radiographic NDT system for weld inspection: Part II--Flaw detection*”, NDT&E International, Vol. 31, No. 3, pp.183-192, 1998.
- [2]. Newman, T.S.; Jain, A.K.: “*A survey of automated visual inspection*”. Computer Vision and Image Understanding, 61(2): 231-262, 1995.
- [3]. Mery, D.; Filbert, D.: “*Automated Flaw 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.
- [4]. Mery, D.; Carrasco, M.A.: “*Automated multiple view inspection based on uncalibrated image sequences*”. Lecture Notes in Computer Science, LNCS 3540: 1238-1247, 2005.
- [5]. Hartley R.; Zisserman A.: “*Multiple View Geometry in Computer Vision*”. Cambridge University Press, Cambridge, UK, 2000.
- [6]. Castleman, K. “*Digital image processing*”. Prentice-Hall. Englewood Cliffs, New Jersey, 1996.
- [7]. Haralick, R.; Shapiro, L.: “*Computer and robot vision*”. Addison-Wesley Publishing Co., New York, 1992.
- [8]. Mery, D.: “*Crossing line profile: a new approach to detecting defects in aluminium castings*”. Lecture Notes in Computer Science 2749. 725-732, 2003.
- [9]. Faugeras, O. et. al: “*The Geometry of Multiple Images: The Laws That Govern the Formation of Multiple Images of a Scene and Some of Their Applications*”, The MIT Press, Cambridge MA, London, 2001.