

EST 2015, 18º Concurso de Trabajos Estudiantiles.

Construcción de un escáner 3D y procesamiento de imágenes con Python y OpenCV

Alejandro Dario Villalba

Universidad Gastón Dachary. Posadas, Misiones
villalba.alejandro@gmail.com

Resumen. La técnica de escaneo 3D es capaz de potenciar la detección y descripción de fallas en la superficie de listones de madera destinados a la fabricación de pisos. El sistema de escaneo 3D captura una señal de video compuesta por la proyección de una línea láser sobre la superficie de un listón de madera. Sobre la señal de video se aplican diferentes técnicas de procesamiento de imágenes con la finalidad de reducir el ruido y aplicar mejoras. Se procede con el análisis y la manipulación de la imagen a fin de poder obtener un perfil de escaneo que representa a la sección de la pieza de madera escaneada. Para lograr esto, se requirió de la construcción de un prototipo de escaneo 3D más la implementación del sistema de procesamiento de imágenes con Python y OpenCV. Los resultados de los procesos descriptos en este documento serán utilizados en un Trabajo Final de Carrera.

Keywords: Escaneo 3D, procesamiento de imágenes, binarización, umbral, suavizado, esqueletizado.

1 Introducción

En el ámbito de la Fabricación industrial de pisos de madera sólidos, se debe prestar especial atención a las fallas existentes en la superficie de la materia prima, a fin de poder ofrecer al mercado un determinado nivel de calidad del producto terminado. La regulación de calidad utilizada se encuentra definida en las Normas IRAM 9552 Pisos de madera, en sus partes 2 (1) y 3 (2).

La industria maderera moderna tiene la posibilidad de implementar la informatización del proceso productivo como un recurso estratégico permitiendo reducir costos, lograr ventajas competitivas y la generación de puestos de trabajo calificado. Una de las ramas de la informática es la Visión por Computadora y su aporte a la industria maderera (según (3) y (4)) consiste en el aumento de la interpretación humana sobre el producto o dotar a una máquina de percepción de visión para aprovechar su capacidad de automatización.

En este trabajo se plantearán un conjunto de técnicas de procesamiento de imágenes a implementar, para lograr obtener un perfil de escaneo resultante del funcionamiento de un escáner 3D y del procesamiento de la señal de video. El perfil de escaneo proveerá los datos necesarios para la posterior detección y descripción de irregularidades en la superficie de la madera.

2 Alcance

Este trabajo es parte de una investigación realizada en el contexto del Trabajo Final de Carrera de Ingeniería en Informática cursada en la Universidad Gastón Dachary. Dicho trabajo consiste en el desarrollo de un sistema de escaneo 3D para la detección y descripción de fallas en la superficie de listones de madera, destinados a la fabricación de pisos machihembrados sólidos. Para dicho proyecto se llevó a cabo la construcción de un prototipo de escáner 3D motorizado, controlado por una placa Arduino. El desarrollo de software es realizado con Python y OpenCV. El documento presente aborda el análisis e implementación de las técnicas de procesamiento de imágenes necesarias para la obtención del perfil de escaneo.

3 Marco Teórico

La técnica de medición 3D por triangulación láser requiere la disposición triangular entre 3 componentes: El proyector Láser, el dispositivo de captura de video y la plataforma de escaneo (ver Fig. 1)

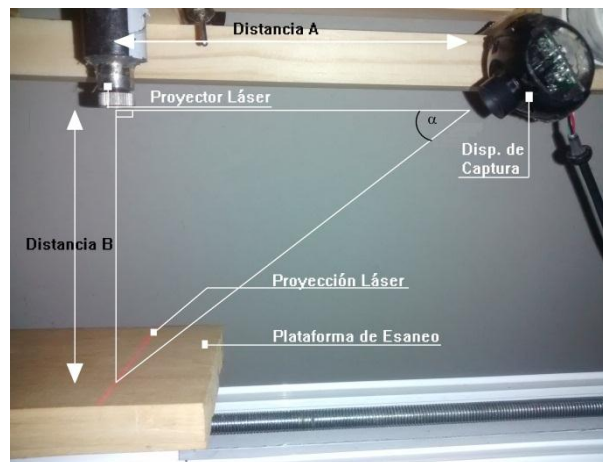


Fig. 1. Elementos de un sistema de Triangulación Láser

El dispositivo de captura de video debe estar correctamente enfocado de forma que la proyección láser se encuentre en su campo de visión. En la Fig. 2 se puede apreciar

la proyección láser dentro del campo de visión del dispositivo de captura. Dicho dispositivo realiza capturas a una tasa de 30 fps con una resolución de 640 x 480 px.

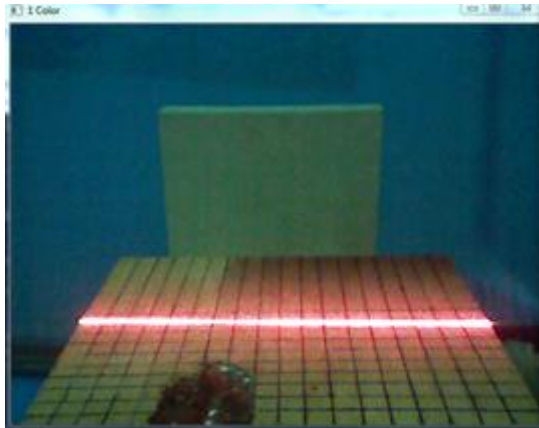


Fig. 2. Captura de proyección láser a color

Las particularidades de funcionamiento de un dispositivo de medición por triangulación láser pueden ser apreciados en (5) para la medición de alturas de cajas de cartón. Por otra parte, en el trabajo (6) es implementado un dispositivo de medición láser más acorde a las necesidades de este proyecto al verificar la calidad de soldadura.

En el modelo de color RGB, cada color aparece representado por una combinación de sus componentes espectrales primarios: rojo, verde y azul. Este modelo está basado en el sistema de coordenadas cartesianas donde cada color se corresponde con 1 eje (ver (3) pg. 116). Cada píxel es una combinación de 3 tonos de color. Si se utiliza 8 bits para representar cada componente de color, es posible obtener una paleta de 256 tonos (donde el 0 es el más oscuro y el 255 el más claro). La señal de video capturada, con 3 canales de colores, deberá ser transformada a tonos de grises. Este proceso transforma cada píxel de color a su correspondiente tono de gris, respetando la escala de tonalidades anteriormente expuesta.

El histograma de una imagen es una representación gráfica de la cantidad de ocurrencias de los distintos tonos de la imagen. En la Fig. 3 se puede apreciar un ejemplo.

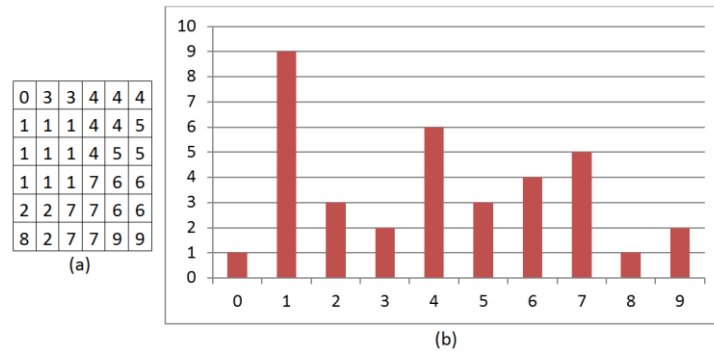


Fig. 3. Ejemplo de Histograma: (a): Imagen Original. (b) Gráfico de Histograma

Una de las técnicas más utilizadas para la extracción de regiones consiste en la binarización mediante detección de umbral. Para esta técnica se requiere realizar un estudio del histograma de la imagen en tonos de grises a fin de poder determinar el valor de umbral a aplicar. El resultado de este proceso se corresponde con:

$$g(x, y) \begin{cases} 0 & \text{si } f(x, y) > T \\ 1 & \text{si } f(x, y) \leq T \end{cases} \quad (1)$$

Donde $g(x, y)$ es la imagen binarizada, $f(x, y)$ es la imagen en tonos de grises y T es el umbral. Se puede apreciar un ejemplo en la Fig. 4.

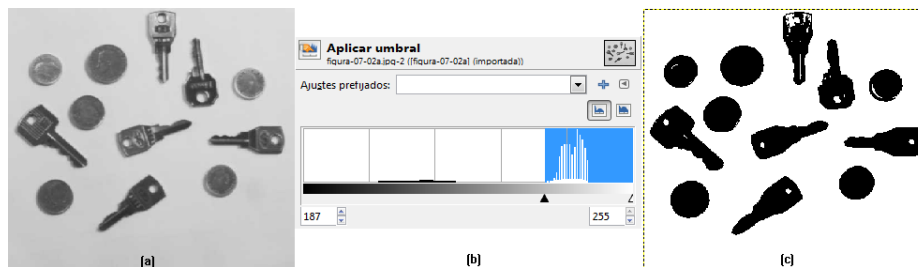


Fig. 4. Ejemplo de Binarización: (a): Imagen Original. (b) Gráfico de Histograma. (c) Imagen Binarizada con el umbral de (b) (187)

El ruido en las imágenes binarias produce efectos no deseados. Dentro de estos defectos, se pueden listar: contornos irregulares, esquinas perdidas, superficies poco definidas y los llamados "sal y pimienta" (ver Fig. 5) Por esta razón, una de las primeras fases del procesamiento de imágenes conlleva la reducción del ruido. Una técnica consiste en realizar la transformación del tono de un píxel dependiendo de un conjunto de píxeles vecinos acorde a un tamaño de vecindad (usualmente de 3, 5 o 7).

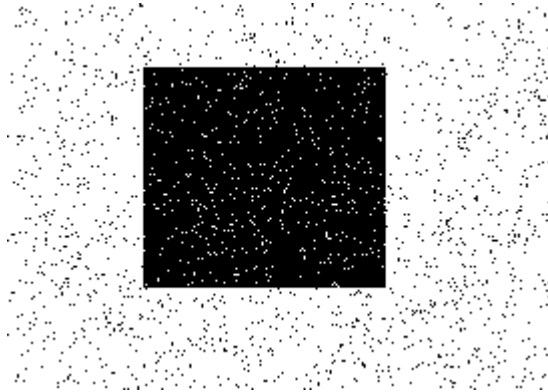


Fig. 5. Ejemplo de Ruido denominado Sal (puntos blancos) y Pimienta (puntos Negros)

Luego de poder reducir o eliminar el ruido, es posible aplicar técnicas de esqueletización para obtener el soporte de la figura (el esqueleto). La figura resultante será una un conjunto de líneas que correspondan con la región sin pérdida de la forma esencial. Para lograr esto se realiza un proceso de adelgazamiento de la región hasta obtener el resultado deseado. En la Fig. 6 se puede apreciar una imagen binaria y el resultado su esqueletización.

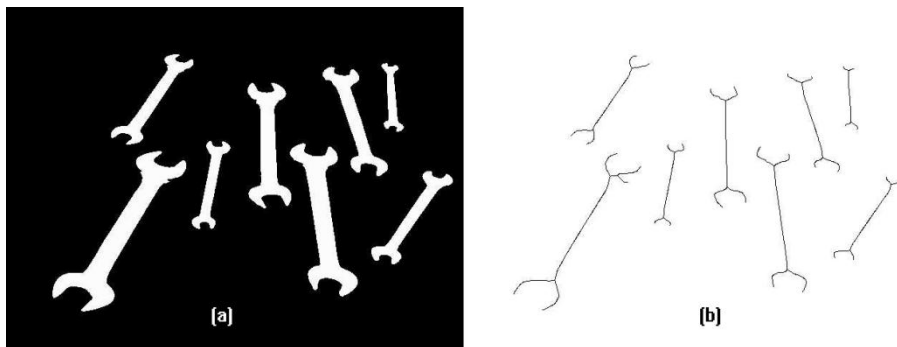


Fig. 6. Ejemplo de Esqueletización: (a) Imagen original Binarizada. (b) Resultado de la esqueletización de (a). Ver (3) pg. 230

El resultado final del procesamiento de la señal de video consiste en la obtención de una línea continua a partir de la captura de la proyección láser. La forma que adquiera dicha línea será una representación topográfica de la superficie de la pieza escaneada en la sección afectada por la proyección láser.

4. Implementación

Se describirán a continuación las técnicas y los procedimientos implementados para la obtención del perfil de escaneo a partir de una señal de video.

4.1. Construcción del Hardware

El principio de funcionamiento de un escáner 3D requiere que el cuerpo a escanear se desplace debajo de una proyección láser. El desplazamiento es realizado por un motor Paso a Paso controlado por una placa Arduino UNO y un panel LCD. El panel LCD cuenta con una pantalla y una serie de botones, permitiendo visualizar y cambiar el comportamiento del motor.

El sistema requiere de un alto grado de precisión para obtener los resultados deseados. La estructura de sostén está armada con columnas para estanterías metálicas pre-perforada y ajustadas con bulones, de forma que sea posible la realización de ajustes. La forma de la estructura se asemeja a la forma de una L. El lado más corto (destinado a apoyar la plataforma de escaneo) se encuentra horizontal y apoyado en la mesa. El lado más largo se encuentra vertical y es utilizado para sostener el proyector láser y el dispositivo de captura de video.

Los componentes del sistema de escaneo están montados de tal forma que es posible realizar modificaciones tanto en la altura de los dispositivos respecto de la plataforma (Distancia B en la Fig. 1), la separación entre el proyector láser y el dispositivo de captura (Distancia A en la Fig. 1) y el ángulo de incidencia del dispositivo de captura (Angulo α en la Fig. 1).

4.2. Calibración

Se deberá realizar la calibración del proyector láser y del dispositivo de captura a fin de determinar las distancias mínimas a las que se deberá ajustar el sistema.

4.2.1. Calibración del Proyector Láser

Se deberá conocer la distancia mínima que debe existir entre el proyector láser y el objeto a fin de poder asegurar que la longitud de la proyección sobre la superficie del objeto alcance a cubrir todo su ancho. Para lograr esto, es necesario calcular el ángulo de proyección en el lente del proyector láser (ver “alfa” en la Fig. 7).

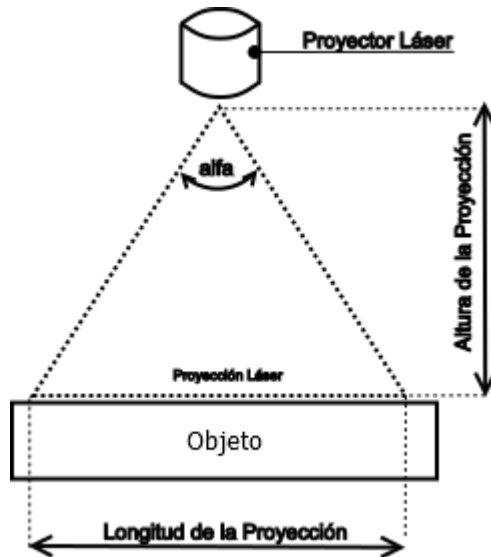


Fig. 7. Diagrama de Calibración láser

El proceso de cálculo inicia situando al proyector láser a una distancia conocida respecto del objeto y la posterior medición de la longitud de la proyección láser sobre la superficie del objeto. Luego, es posible calcular el ángulo utilizando las relaciones trigonométricas. Para el cálculo se dispuso el proyector láser a una distancia de 26.35 mm del objeto, dando una proyección de 50 mm sobre la superficie del mismo. Los resultados obtenidos indican que el ángulo es de 86.98° y que se deberá respetar una distancia mínima de aprox. 60 mm entre el proyector láser y el objeto.

4.2.2. Calibración del Dispositivo de Captura

El dispositivo de captura es una cámara web comercial con una resolución de 640 x 480 px. Al poseer diferente resolución entre el horizontal y el vertical, los ángulos de apertura (horizontal y vertical) del campo de visión son diferentes y tienen utilidades diferentes. En el caso de la horizontal, será utilizado para determinar la distancia mínima que debe existir entre la cámara y el objeto, a fin de poder asegurar que todo el ancho del objeto se encuentre en el campo de visión de la cámara. En el caso de la vertical, será utilizada para calcular la equivalencia entre 1 píxel y los ángulos que se corresponde. Esta relación será explotada en el Trabajo Final de Carrera en la etapa de medición.

Para llevar a cabo las mediciones pertinentes, se requirió la impresión de un patrón cuadrulado de 10 mm x 10 mm dispuesto a una distancia de 154 mm de la cámara. Luego de realizar la captura de la señal de video, se realizó un trabajo manual sobre la captura a fin de poder identificar los píxeles que conforman el patrón cuadrulado, tal como se aprecia en la Fig. 8}.

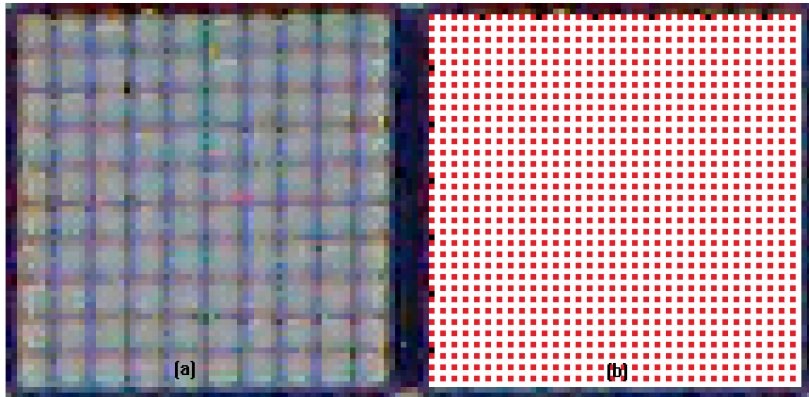


Fig. 8. Análisis de Patrón: (a) Captura Original con zoom. (b) Resultado del reconocimiento de la superficie

El resultado obtenido indica que por cada cuadro de 10 x 10 mm en el patrón existen 60 x 60 px. Luego, es posible calcular el largo y ancho del campo de visión teniendo en cuenta la existencia de una resolución de 640 x 480 px dando como resultado 96.96 mm x 72.72 mm respectivamente. A partir de esto, se pudo determinar que, en el sentido vertical, cada píxel se corresponde con 0.0553° y que en el sentido horizontal se requiere de aprox. 151 mm para asegurar que toda la proyección láser se encuentre dentro del campo de visión de la cámara. Estos valores serán corroborados en la etapa de prueba del sistema y dependiendo de los resultados, se procederá a realizar correcciones.

4.3. Implementación del Software

4.3.1. Estudio del Histograma y Binarización

La finalidad del estudio del histograma es la determinación del umbral óptimo a utilizar en la etapa de binarización. Previo al estudio del histograma se aplicó sobre la señal de video la transformación de los colores a sus correspondientes tonos de grises. En esta transformación se realizaron 2 experimentos, variando la iluminación del ambiente y se obtuvieron los siguientes resultados:

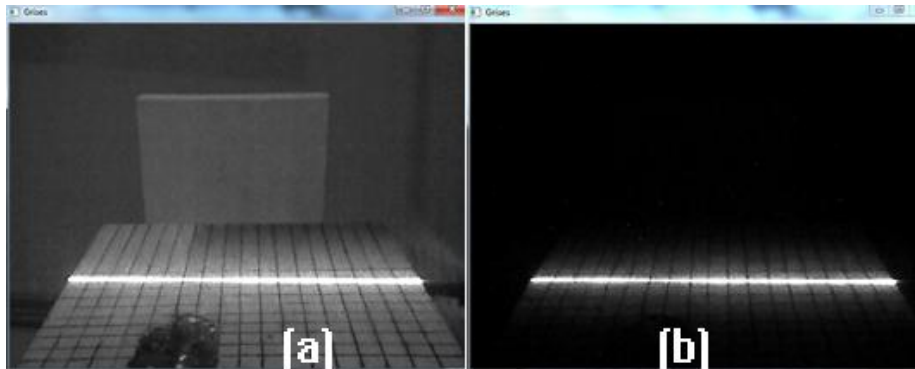


Fig. 9. Tonos de Grises: (a) Con Iluminación Ambiente. (b) Sin Iluminación Ambiente

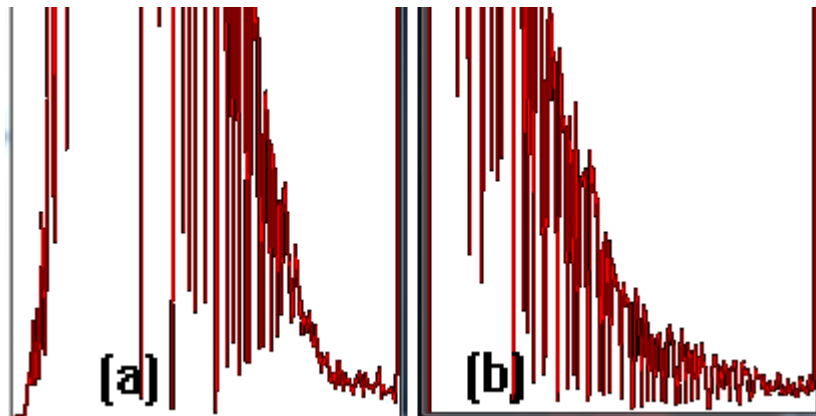


Fig. 10. Histogramas Resultantes: (a) Con Iluminación Ambiente. (b) Sin Iluminación Ambiente

En los histogramas de la Fig. 10 es posible notar la mayor carga de datos del lado izquierdo (tonos oscuros) pero, casi imperceptiblemente, se grafican valores del lado derecho (más iluminado) que son los que se corresponden con la línea láser.

Es necesario entonces realizar un estudio cuya finalidad es poder definir cuál es el valor más adecuado para el umbral. Cuando la proyección láser incide sobre la superficie de una muestra se produce un efecto de reflexión. El efecto de esta reflexión sobre la señal digital produce una línea notablemente más gruesa que la propia proyección, con un centro más brillante degradándose hacia sus bordes (en la imagen b de la Fig. 9 se puede apreciar como la proyección láser es capaz de iluminar incluso parte de la superficie de la muestra).

Se realizó un estudio estadístico sobre los resultados obtenidos al aplicar la binarización sobre una señal de video para un conjunto de umbrales que fueron desde el 249 al 255. Por cada valor de umbral se tomaron 50 muestras de la señal de video y

se calcularon una serie de indicadores (Máximo, Mínimo, Media, Varianza y Desviación Estándar). Los resultados fueron los siguientes:

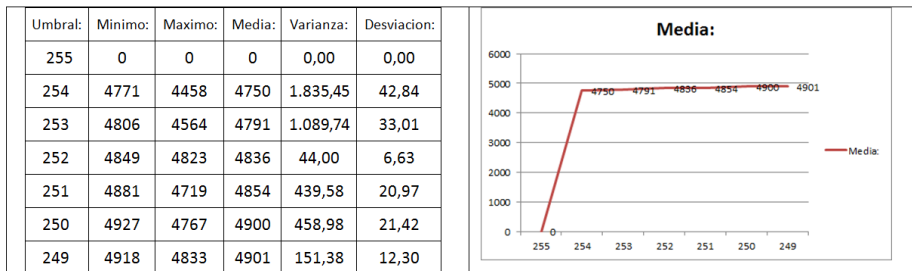


Fig. 11. Estudio estadístico del umbral ideal

Los resultados obtenidos indican que el valor de 254 como umbral produce una menor cantidad de píxeles haciendo que se descarten a aquellos que no son lo suficientemente brillantes. En la Fig. 12 se puede apreciar una comparación entre la imagen original y el resultado de aplicar la siguiente función:

```
ret,binario=cv2.threshold (
    gray,
    valorUmbral,
    255,
    cv2.THRESH_BINARY_INV
)
```

Dónde los parámetros son:

- **gray:** La imagen original en tonos de grises
- **valorUmbral:** El número del umbral a aplicar
- **255:** El máximo valor admitido
- **cv2.THRESH_BINARY_INV:** Tipo de Umbralización

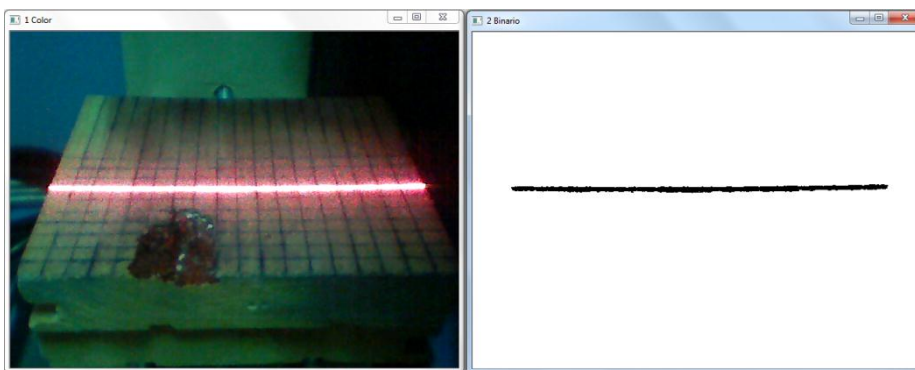


Fig. 12. Resultado de aplicar un valor umbral de 254

4.3.2. Suavizado Binario

Si se realiza una ampliación a la imagen binarizada de la Fig. 12 se podrá notar la existencia de píxeles irregulares en el contorno de la línea (ver Fig. 13)



Fig. 13. Ampliación de la Fig. 12 binarizada

OpenCV ofrece 3 funciones para poder aplicar el suavizado sobre imágenes, las cuales están listadas a continuación y ofreciendo diferentes enfoques:

```

1 ladoKernel=5
2 GBlur=cv2.GaussianBlur(binario, (ladoKernel, ladoKernel)
3   , 0)
4 Blur = cv2.blur(binario, (ladoKernel, ladoKernel))
5 MBlur = cv2.medianBlur(binario, ladoKernel)
    
```

Los experimentos realizados arrojan los siguientes resultados:

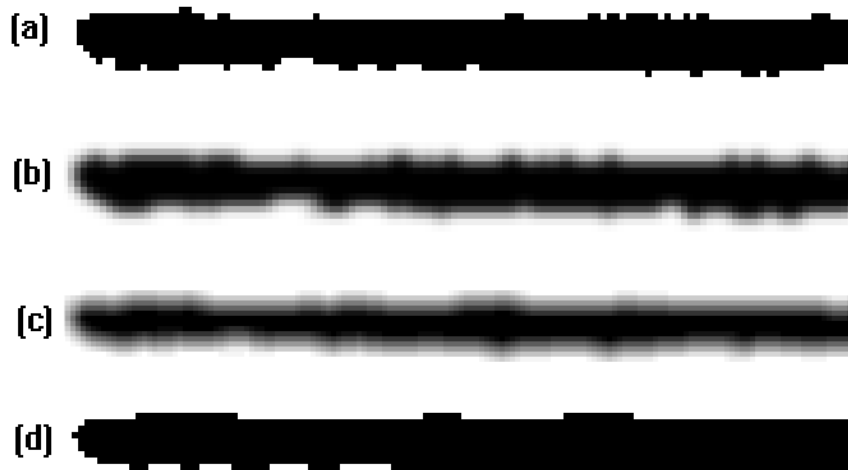


Fig. 14. Resultado de aplicar 3 técnicas de suavizado: a) Original, b) GaussianBlur, c) blur y d) medianBlur. Por razones de espacio, se presentan los extremos izquierdos de las imágenes

Como se puede apreciar en los resultados de la Fig. 14 (con un tamaño de vecindad de 5), la función medianBlur (d) es la que aporta el mejor resultado. Las otras 2 técnicas no son adecuadas al incluir píxeles con otros tonos diferentes al binario y cuya posterior binarización produciría el ensanchamiento de la línea alejándose así del objetivo buscado.

El estudio a realizar a continuación consiste en determinar el tamaño de la vecindad a utilizar (kernel). En la Fig. 15 es posible apreciar los resultados obtenidos con tamaños de kernel correspondientes a 3, 5, 7 y 9.

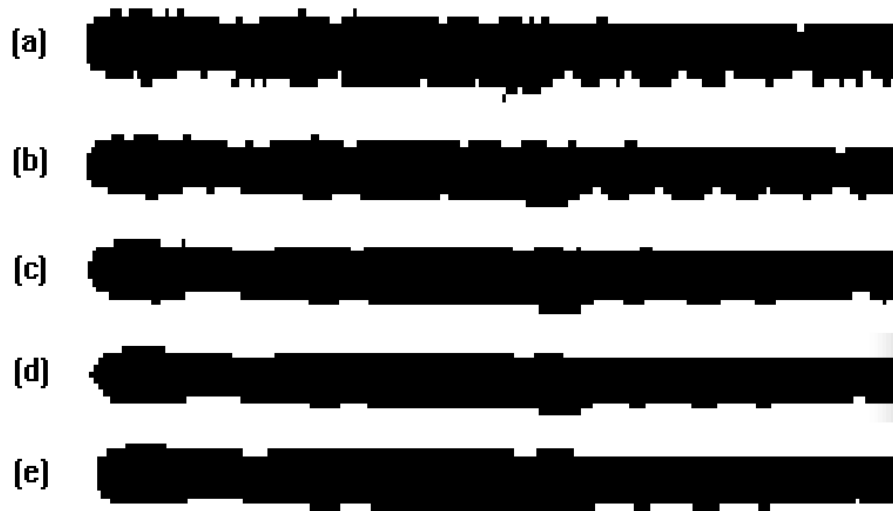


Fig. 15. Resultado de aplicar 4 tamaños de kernel: a) Original, b) kernel=3, c) kernel=5, d) kernel=7 y e) kernel=9. Por razones de espacio, se presentan los extremos izquierdos de las imágenes.

Se realizó un estudio para determinar el espesor de las líneas obtenidas luego del proceso de binarizado y el resultado indica que como máximo hay 7 píxeles. Si se observan los resultados obtenidos en la Fig. 15, el kernel=5 (c) resulta el más adecuado por el momento. Estudios futuros podrán modificar la esta selección.

4.3.3. Esqueletización

La esqueletización tiene por objetivo obtener una línea representativa del objeto inicial. Una de las necesidades de este trabajo implica la no modificación del largo de la muestra por lo que se requiere reducir solo el espesor. Para la esqueletización se aplicó un proceso de análisis y filtrado de forma que solo sean mostrados los píxeles de la parte superior de la figura binarizada. En el caso de que existan discontinuidades en el trazo de la muestra, se procedió a crear puntos en la parte inferior del resultado a fin de indicar la existencia de una falla y lograr la continuidad del resultado. En la Fig. 16 se muestran los resultados obtenidos sin suavizar (b) y la aplicación del suavizado previo (c) utilizando el método explicado en el apartado anterior.



Fig. 16. Resultado de aplicar la esqueletización: a) Imagen Original, b) Esqueletización sin suavizado, c) Esqueletización con suavizado (kernel=5)

5. Resultados

En la Fig. 17 y Fig. 18 y se pueden apreciar 2 casos de procesos completos:



Fig. 17. Resultado de una muestra sin falla

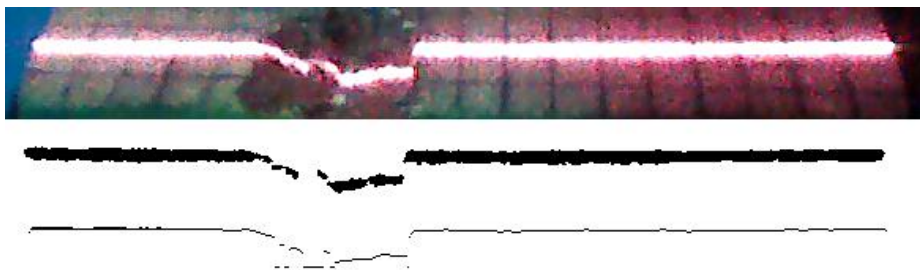


Fig. 18. Resultado de una muestra con falla

6. Trabajos Futuros

Las técnicas y resultados detallados en este documento serán utilizados para lograr la implementación de un sistema de escaneo 3D. Los trabajos a futuro implican lograr que el sistema realice el procesamiento de la señal de video en tiempo real y sea

posible realizar las mediciones necesarias para obtener la vista tridimensional de la pieza escaneada. El trabajo finalizará con la implementación de un sistema de Software que analice las figuras tridimensionales y logre la detección y descripción las irregularidades.

7. Conclusiones

Las técnicas aplicadas al estudio del Histograma y la posterior binarización dieron resultados satisfactorios. En cuanto a la selección del método de suavizado, la función medianBlur es la más adecuada de entre las opciones nombradas. La aplicación del método de esqueletizado deberá ser puesto a prueba junto con el método de suavizado en una instancia futura a este trabajo.

Referencias

1. Argentina, IRAM. IRAM_9552-2. *Pisos de Madera. Parte 2 - Requisitos Generales, marcado y evaluación de conformidad*. 2010.
2. —. IRAM_9552-3. *Pisos de Madera. Parte 3 - Macizos (sólidos). Requisitos*. 2010.
3. Pajares, G y De La Cruz, J. M. *Visión por Computador. Imágenes digitales y aplicaciones*. Madrid : RA-MA, 2008.
4. Sossa, J. H. y Rodriguez, M. R. *Procesamiento y Analisis de Imagenes Digitales*. México : RA-MA, 2015.
5. Battiston, Federico. *Medición con láser por Triangulación*. San Nicolás de los Arroyos : Universidad Tecnológica Nacional, 2011.
6. *A Laser-Based Vision System for Weld Quality Inspection*. Huang, Radovan y Huang, Wei. s.l. : Sensors, 2011, Sensors.