



Casa abierta al tiempo

UNIVERSIDAD AUTONOMA METROPOLITANA

UNIDAD IZTAPALAPA

DIVISIÓN DE CIENCIAS BÁSICAS E INGENIERÍA
Posgrado en Ciencias y Tecnologías de la Información

ANÁLISIS DE LA MORFOLOGÍA VEGETAL UTILIZANDO UNA CÁMARA 3D

TESIS

que presenta

SERGIO CONTRERAS MURILLO

para obtener el grado de

DOCTOR EN CIENCIAS

(Ciencias y Tecnologías de la Información)

Asesores:

Dr. Sergio Gerardo de los Cobos Silva

Dr. Román Anselmo Mora Gutiérrez

Ciudad de México, agosto de 2020

RESUMEN

La **morfología** vegetal es la rama de la biología que estudia la forma externa de las plantas (Müller 2000). El análisis de la forma y tamaño de las hojas de las plantas permite determinar su estado de salud y su capacidad para absorber la luz solar.

En esta investigación se aborda el uso novedoso de una **cámara 3D** para analizar la forma de las plantas, en especial aquellas cuya forma imposibilita la aplicación de las técnicas de medición previamente reportadas en la literatura.

Se presenta entonces la medición del área de cladodios de *Opuntia Spp.* pues sus espinas dificultan su manejo y sus dimensiones imposibilitan la utilización de aparatos diseñados para medir hojas.

Las especies del género *Opuntia* (nopales) son cactáceas comestibles, nativas de América; En ellas, las hojas se encuentran transformadas en espinas y son los segmentos aplanados de su tronco (**cladodios**) los que presentan capacidades fotosintéticas.

El método propuesto no es “**destructivo**”, el resultado se obtiene de inmediato, puede utilizarse *in situ*, es de bajo costo, emplea tecnología de punta y resuelve los problemas que presentarían los métodos clásicos de medición del área foliar al aplicarse a dichas plantas espinosas.

La medición de los cladodios permite evaluar el rendimiento de un cultivo, decidir el momento para la cosecha o censar y caracterizar especímenes silvestres.

En contraste con las imágenes 2D, con las imágenes 3D se considera la **inclinación** y **curvatura** de la muestra para determinar su superficie. Se propone calcular la inclinación vía la **regresión lineal múltiple** de una muestra al centro de la imagen.

Para evaluar la estrategia se obtuvo el área de 12 cladodios con el método propuesto y otras cinco estrategias contra las que se comparó. El algoritmo presentó una exactitud del **95%**, al nivel de las mejores estrategias destructivas, pero sin dañar la planta y es al menos 20 veces más rápido en términos del tiempo necesario para implementarlo.

Para esta investigación se utilizó una cámara 3D Intel© RealSense© R200 que, si bien funciona con **luz infrarroja** (le afecta la luz solar), permite por primera vez capturas en campo.

El algoritmo codificado en el lenguaje Java, tarda menos de 2 segundos en ejecutarse en un equipo con procesador Intel core I7 con 8 GB de RAM a 2.5 GHz.

La base de datos con las imágenes 3D generadas, así como el código fuente pueden descargarse gratuitamente desde el sitio <http://www.serconmur.com/area-foliar-3d.html>



Figura 1. Nopal (*Opuntia Spp*).

Dedicatoria

Este trabajo está dedicado a mis papás, a mis hermanos, a mis familiares, a mis amigos y a mis profesores, que me inspiran a seguir esforzándome día a día con alegría y dedicación.

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología por su apoyo y patrocinio durante mis estudios de doctorado a través de su programa de becas nacionales para estudios de posgrado.

Le agradezco al personal del vivero forestal de la Universidad Autónoma Chapingo por las facilidades que me prestaron en las visitas a sus instalaciones.

Le agradezco con mucho cariño a mis asesores, los Dres. Sergio Gerardo de los Cobos Silva y Román Anselmo Mora Gutiérrez por su ayuda y sus consejos.

Finalmente, le agradezco al Dr. Enrique Rodríguez de la Colina, coordinador del posgrado; al personal de la división de Ciencias Básicas e ingeniería y a todos los que laboran en la Universidad Autónoma Metropolitana por su dedicación y por la ardua labor que realizan. En estos momentos difíciles en los que enfrentamos la pandemia del COVID-19 su entereza se hace patente.

ÍNDICE

Introducción.....	17
1.1 Hipótesis	19
1.2 Objetivos	19
1.2.1 Objetivo general	19
1.2.2 Objetivos particulares.....	19
1.3 Organización del documento.....	20
1.4 Definición del problema	21
1.5 Complejidad del problema	22
La triangulación	22
La segmentación.....	23
Conceptos básicos	26
2.1 La morfología vegetal	26
2.2 Las partes de la hoja	27
2.3 Los tipos de hojas	28
2.4 Las técnicas de medición	30
Estado del arte	32
3.1 Estado del Arte	32
3.2 Las imágenes 3D	36
3.2.1 los formatos de imagen 3D.....	36
3.2.2 las tecnologías 3D	39
3.3 El tratamiento digital de imágenes 3D	42
3.3.1 Filtrado	42
3.3.2 Segmentación.....	45
3.3.3 Reconstrucción	46
3.3.4 Rellenado.....	46
3.3.5 Otras operaciones de morfología matemática	49
3.4 Los dispositivos 3D	49

3.4.1 El sensor 3D Kinect de Microsoft	50
Medición de cladodios de <i>Opuntia SPP</i>	56
4.1 Materiales y metodología	59
4.2 Regresión Lineal Múltiple.....	61
4.3 Crecimiento de región con semilla	62
4.4 Rellenado.....	62
4.5 Erosión.....	63
4.6 Dilatación	63
4.7 Triangulación.....	64
4.8 La aplicación.....	65
Resultados y discusión.....	67
Conclusiones.....	74
Apéndice I: Especificaciones técnicas del dispositivo 3D Kinect de Microsoft	76
Apéndice II: Especificaciones técnicas de la cámara 3D Intel R200	78
Apéndice III: Especificaciones técnicas de la cámara 3D Intel SR300	80
Apéndice IV: Definición xml.....	82
software.....	86
Curriculum vitae del autor.....	88
glosario	90
Bibliografía.....	92

ÍNDICE DE FIGURAS

Figura 1. Nopal (<i>Opuntia Spp</i>).....	5
Figura 2. Reconexión de una malla existente alrededor de un nuevo punto insertado con el algoritmo de Bowyer-Watson (Rebay, 1993).....	22
Figura 3. En el octree cada nodo puede estar dividido en 8 partes.	24
Figura 4. Octree (PCL. 2019).....	24
Figura 5. Hoja compuesta de <i>Jacaranda Caerulea</i> con: (1) base foliar, (2) peciolo, (3) raquis, (4) foliolos.....	28
Figura 6. Clasificación de las hojas (Con imágenes de Kraus et al, 2004; e imágenes propias).....	29
Figura 7. Diagrama de la forma de las variables Largo (L) y ancho (A) en hojas de a) Brevo; b) Curuba; c) Feijoa; d) Fresa; e) Mora y f) Uchuva (Casierra-posada et al, 2007) 30	
Figura 8. Parámetros biométricos medidos en los cladodios de <i>O. Ficus-Indica</i> . L: Largo; W: Ancho; T: Espesor medio (obtenido de la media aritmética de los puntos T1, T2 y T3); D: Diámetro del cuello (Gaspar et al. 2018).....	31
Figura 9. Integradores portátiles CID Bio Science (CID Bio-Science Inc., 2018)	32
Figura 10. Determinación del área foliar en hojas de <i>Manihot Esculenta</i> (Cock et al, 2014)	33
Figura 11. Escaneo 3D de una planta de coliflor (<i>Brassica oleracea</i>) (Andújar et al, 2016)	34
Figura 12. monitoreo del crecimiento de una planta de frijol (<i>Phaseolus vulgaris</i>) (Okayamaa et al, 2015).	34
Figura 13. Caracterización rápida de estructuras vegetales con un sensor Kinect de Microsoft (Sarbolandi et al, 2015)	35
Figura 14. Resultado del Escaneo de una planta (Sarbolandi et al, 2015).....	35
Figura 15. Imagen 3D de arbusto de adelfa (<i>Nerium Oleander</i>) y su flor en detalle	36

Figura 16. Posición X e Y de los píxeles en la imagen.	37
Figura 17. Posición X, Y y Z de los vóxeles en el espacio.	37
Figura 18. Imagen a color y mapa de profundidades en escala de grises de dos plantas de lechuga (<i>Lactuca Sativa</i>).....	38
Figura 19. Escáner 3D de luz estructurada HP Sprout© (HP Development Company, L.P, 2016).....	39
Figura 20. Proyección de puntos de luz infrarroja en una planta de Níspero (<i>Eriobotrya japonica</i>) utilizando un sensor Kinect TOF de microsoft	40
Figura 21. Mapa de profundidades de una planta de Níspero (<i>Eriobotrya japonica</i>) obtenida de un sensor Kinect TOF de Microsoft.....	40
Figura 22. ARCore, el SDK de realidad aumentada de Google funciona en dispositivos móviles con 2 cámaras (Google, 2018).....	41
Figura 23. Convolución.....	42
Figura 24. Aplicación del operador de Robert Cross para detectar bordes	43
Figura 25. Bordes obtenidos mediante la convolución en una imagen de una hoja a blanco y negro	44
Figura 26. Bordes obtenidos mediante la convolución en una imagen a blanco y negro de una hoja, tonalidad invertida.....	44
Figura 27. Imagen de una hoja en blanco y negro	45
Figura 28. Píxeles de una imagen etiquetados como pertenecientes a dos regiones, a y b	46
Figura 29. Triangulación de una nube de puntos	47
Figura 30. Dos ejemplos de triangulación para el mismo conjunto de datos. El segundo ejemplo cumple la condición de Delaunay	47
Figura 31. Pertenencia de un punto a un triángulo.....	48
Figura 32. Vectores u y v de un triángulo	48
Figura 33. Erosión y dilatación con un elemento estructurante circular (UNS, 2019).....	49

Figura 34. Imagen obtenida con el sensor Kinect. La escena muestra una planta de chile <i>Capsicum annuum</i> en una maceta, iluminada por miles de puntos de luz infrarroja.	50
Figura 35. Mapa de profundidades	51
Figura 36. La luz ambiente afecta el rendimiento del dispositivo Kinect (Sarbolandi et al, 2015)	52
Figura 37. Cámara 3D Intel RealSense R200.....	53
Figura 38. Mapa de profundidades de una planta de nopal en escala de grises obtenida con la cámara Intel RealSense R200.....	53
Figura 39. Intel Realsense SR300 3d camera (Intel corporation, 2015, 3)	55
Figura 40. Luz estructurada en cámara 3D Realsense SR300 (Intel corporation, 2015)	55
Figura 41. Opuntia Ficus Indica (Pixabay)	57
Figura 42. Camello comiendo el cactus espinoso en Etiopia (FAO, 2017)	58
Figura 43. Corte manual de cladodios del cactus sin espinas en Túnez. Ali Nefzaoui (FAO, 2017).....	58
Figura 44. Mapa de vértices de una planta de nopal obtenida con la cámara Intel RealSense R200	60
Figura 45. Pasos del método propuesto	60
Figura 46. Representación del plano obtenido con la regresión lineal múltiple.....	61
Figura 47. Región segmentada, vóxeles vecinos a la semilla y que están a menos de 4 cm. de la ecuación del plano.	62
Figura 48. Región con sus orificios rellenados con valores generados de la ecuación del plano.....	62
Figura 49. Región erosionada, la erosión separa un cladodio de los demás	63
Figura 50. Dilatación del cladodio segmentado para recobrar su dimensión original	63
Figura 51. Pasos del algoritmo de segmentación	64

Figura 52. Ejercicio con dos círculos de nopal unidos entre sí, A) mapa de vértices, B) segmentación inicial, C) segmentación final.....	65
Figura 53. La imagen se divide en 4 cuadrantes con 3 hilos cada uno.	66
Figura 54. Gráfica de cajas y bigotes de las mediciones realizadas mediante 1) planimetría, 2) cámara 3D, 3) estimación matemática de elipse, 4) estimación matemática de triángulo y semicírculo, 5) volumetría	69
Figura 55 componentes del sensor Kinect (Microsoft, 2018, 1)	76
Figura 56. Conector Kinect – PC (Microsoft, 2018)	77
Figura 57. Ejercicio con un sensor Kinect de Microsoft ©.....	77
Figura 58. Cámara 3D Intel RealSense R200 (diagrama) (Intel corporation, 2015).	78
Figura 59. Diagrama de la cámara Intel RealSense R200 (Intel corporation, 2015).....	79
Figura 60. Aplicación para la toma de fotografías en 3D.....	86
Figura 61. Aplicación para el análisis de las imágenes 3D	87

ÍNDICE DE CUADROS

Cuadro 1: Comparación de tecnologías 3D que pueden utilizarse en campo	41
Cuadro 2: Clasificación científica de las plantas del género <i>Opuntia</i>	56
Cuadro 3: Áreas en cm ² obtenidas mediante Escaneo 2D (medición de referencia), el método propuesto y otras 4 metodologías	68
Cuadro 4: Margen de error de las áreas obtenidas en cm ² de las 5 técnicas de medición respecto del escaneo 2D en términos de la distancia Euclidiana	69
Cuadro 5: Margen de error de las áreas obtenidas en cm ² por toma de fotografías 3D respecto del escaneo 2D en términos de la distancia Euclidiana	70
Cuadro 6: Comparación entre el método propuesto y otras técnicas de medición del área foliar al aplicarse a cladodios espinosos de <i>Opuntia Spp</i>	72
Cuadro 7: Propiedades de las cámaras a color RealSense R200	79
Cuadro 8: Propiedades de las cámaras infrarrojas RealSense R200	79
Cuadro 9: Parámetros de operación de la cámara RealSense SR300	80

ÍNDICE DE FÓRMULAS

Fórmula 1: Área de las hojas de <i>Solanum Lycopersicum</i> (Jitomate)	33
Fórmula 2: Operador de Robert-Cross	43
Fórmula 3: Detección de bordes	43
Fórmula 4: Matrices que representan una zona plana, un borde y una esquina	44
Fórmula 5: Kernels para la detección de esquinas	45
Fórmula 6: Área de un triángulo dados 3 vértices	48
Fórmula 7: Regresión lineal	61

INTRODUCCIÓN

El tamaño de las plantas es monitoreado por investigadores y agrónomos para evaluar el estado de salud de los cultivos. Así como las plantas tienen formas muy variadas, también son muchas y muy variadas las métricas y las técnicas de medición.

En este trabajo se presenta el uso de una **cámara 3D** para analizar la forma de las plantas y se escogió trabajar con cladodios de *Opuntia Spp.* (nopales) porque la presencia de espinas en la planta y sus dimensiones dificultan la aplicación de las estrategias clásicas de medición del **área foliar**.

En las plantas del género *Opuntia* las hojas se encuentran transformadas en espinas y sus troncos, transformados en segmentos circulares y aplanados (cladodios), los que presentan capacidades fotosintéticas.

Obtener el **área** de un cladodio de nopal mediante **estimación matemática** o **planimetría** depende de mediciones directas que se dificultan por la presencia de espinas mientras que los equipos especializados (**integradores**) no están diseñados para piezas gruesas ni grandes.

Los métodos que se basan en la toma de **fotografías 2D**, tienen problemas para segmentar especímenes que se encuentran rodeados de otros del mismo color, forma, textura y/o tamaño.

En las imágenes 3D los **pixeles** tienen información espacial y se convierten en **vóxeles**.

El tratamiento digital de imágenes 3D empieza con su **adquisición** mediante tecnologías como: **Estereoscópica**, **Tiempo de Vuelo** o **Luz estructurada**. La imagen 3D se presenta en formato de **nubes de puntos**, **mapas de profundidades** o **mapas de vértices**.

En la etapa de **filtrado** se elimina el ruido de la imagen. También se acentúan o atenúan el brillo, contraste, etc.

En la **segmentación**, se separan del fondo las figuras presentes en la imagen. En los sistemas de visión artificial se necesita determinar a priori el modelo descriptor de la estructura a reconocer (La ecuación que más se parece a los objetos). En este caso, la superficie de los cladodios es descrita por la ecuación del plano. Es por ello por lo que la segmentación inicia con la determinación del plano que caracteriza la inclinación y posición del nopal mediante la aplicación de la regresión lineal múltiple de una muestra de la figura a segmentar.

Posteriormente, durante la fase de **reconstrucción**, los vértices señalados por los vóxeles segmentados se unen mediante aristas para formar una figura cerrada.

Finalmente, se selecciona una opción de **visualización** para mostrar el resultado.

La medición con cámaras 3D es una opción que nos permite no destruir ni manipular la planta espinosa y que resuelve las deficiencias de las estrategias 2D (tales como la necesidad de un fondo de un color o patrón específicos). A diferencia de las estrategias que utilizan fotografías 2D, las imágenes 3D si compensan el efecto de la inclinación y de la curvatura de la muestra en la medición del área foliar.

El resultado se muestra inmediatamente y permite medir con exactitud, *in situ*, muestras con un área foliar grande o que se encuentren en la parte alta de la planta.

1.1 Hipótesis

El análisis de fotografías en tercera dimensión tiene un mejor desempeño para caracterizar cladodios de *Opuntia Spp.* en campo que las estrategias clásicas de medición del área foliar.

1.2 Objetivos

1.2.1 Objetivo general

Desarrollar una metodología para el estudio morfológico, no destructivo, en campo, de cladodios de *Opuntia Spp.* basado en el análisis de imágenes en tercera dimensión.

1.2.2 Objetivos particulares

Determinar el dispositivo 3D y formato de imagen adecuados para desarrollar la investigación.

Desarrollar un algoritmo para la segmentación de los cladodios de *Opuntia Spp.*

Desarrollar un algoritmo para determinar la inclinación de los segmentos vía la regresión lineal múltiple.

Especificar la estructura de datos más adecuada para la persistencia y representación de las características morfológicas de los cladodios.

1.3 Organización del documento

En el **capítulo 2** se abarcan conceptos preliminares de la morfología vegetal.

En el **capítulo 3** se describe el estado del arte. Se mencionan las estrategias clásicas de medición del área foliar y se revisan investigaciones recientes en el campo de la biología descriptiva. También se abarcan los sistemas de visión artificial 3D. En "*Las imágenes 3D*" se exponen las diferentes maneras de obtener y representar las imágenes con información espacial. En "*El tratamiento digital de imágenes 3D*" se describen los algoritmos del área de la morfología matemática que se requieren para el análisis de las imágenes 3D.

En el **capítulo 4** se presenta la metodología propuesta para el análisis de las imágenes 3D de cladodios de *Opuntia*. El método consta de 5 fases: obtención, filtrado, segmentación, reconstrucción y presentación. Se utiliza la regresión lineal múltiple para determinar la inclinación y posición de los nopales en el espacio.

En el **capítulo 5** se presentan y analizan los resultados. La metodología propuesta se compara contra otros 5 métodos.

Finalmente, en el **capítulo 6** se presentan las conclusiones.

En el **apéndice 1, 2 y 3** se incluye la información técnica de las cámaras 3D evaluadas.

En el **apéndice 4** se ofrecen los esquemas XSD de los archivos XML diseñados para almacenar la información 3D y sus metadatos.

En el **apéndice 5** se incluye el manual de usuario de la aplicación desarrollada.

1.4 Definición del problema

El monitoreo del desarrollo de las plantas es de vital importancia en la producción e investigaciones agrícolas, pero las plantas del género *Opuntia* presentan espinas y otras características morfológicas que dificultan su estudio mediante métodos clásicos.

Los cladodios de *Opuntia* tienen **espinas**, son **gruesos**, altamente **reflectivos** y sus plantas pueden llegar a medir varios metros de altura. Los equipos especializados (integradores) tienen restricciones en cuanto al tamaño de las muestras que pueden medir y el manejo de las espinas.

Los procesos **destruictivos** ofrecen mayor exactitud, pero a costa de dañar el espécimen e imposibilitan monitorear el desarrollo de una misma planta a lo largo del tiempo.

Las estrategias de medición basadas en fotografías 2D encuentran difícil segmentar piezas que se encuentran rodeadas de otras del mismo color y textura, por ejemplo, en un cultivo.

Hay que tener en cuenta que los sistemas de visión artificial por computadora tienen un alto costo de **memoria** y **procesamiento**. En algunos casos, la luz solar afecta su rendimiento y dificulta su uso en campo.

Otras estrategias de medición basadas en la toma de fotografías presentan algunos de los siguientes inconvenientes que dificultan su implementación en campo:

Requieren **manipular** la planta y/o

Requieren que el **fondo** de la imagen sea de un color en particular y/o

Requieren que el **fondo** de la imagen tenga marcas de referencia y/o.

Requieren que la fotografía se tome a una **distancia** conocida y/o.

No compensan la **inclinación** de la cámara al momento de tomar la fotografía o requieren colocarla de manera paralela a la muestra y/o

No compensan la **curvatura** de la muestra o requieren que esté plana.

En los desarrollos previos que han utilizado cámaras 3D que funcionan con **luz infrarroja**, la luz solar interfiere con su funcionamiento y que requieren una conexión eléctrica y computadora, lo que dificulta su uso en campo.

En algunos casos se requiere construir estructuras de soporte para la cámara (Azzari *et al*, 2013), lo cual aumenta el costo y disminuye la usabilidad de una plataforma. En algunos casos es necesario montar la cámara en un automóvil, transportar las plantas enteras a un área de medición (Okayamaa *et al*, 2015), restringir su uso a invernaderos (Nissimov *et al*, 2015), etc.

1.5 Complejidad del problema

El tratamiento digital de imágenes 2D y 3D es computacionalmente costoso; requiere de mucha memoria y poder de procesamiento. Las estrategias seleccionadas para esta investigación buscaron en todo momento ahorrar recursos. Las dos etapas a implementar en esta investigación con el mayor costo computacional son la **triangulación** y la **segmentación**.

Con esto en mente, la investigación comienza evaluando distintos dispositivos 3D para seleccionar aquel que disminuya la complejidad de los pasos posteriores.

La triangulación

La triangulación se refiere a la unión de los vóxeles mediante aristas que conviertan la nube de puntos en una figura cerrada. En este proceso hay que preferir la configuración en la que la superficie se maximiza, es decir, en la que se cumpla la condición de **Delaunay**. La triangulación de Delaunay *maximiza el ángulo mínimo de cada triángulo de la triangulación* (Cerrolaza & Flórez López, 2000).

Un primer método de triangulación, el método incremental, fue propuesto por Bowyer y Watson en 1981. Este “es esencialmente un método de reconexión, ya que calcula cómo se modificará una triangulación de Delaunay existente debido a la inserción de un nuevo punto. (Rebay, 1993) “

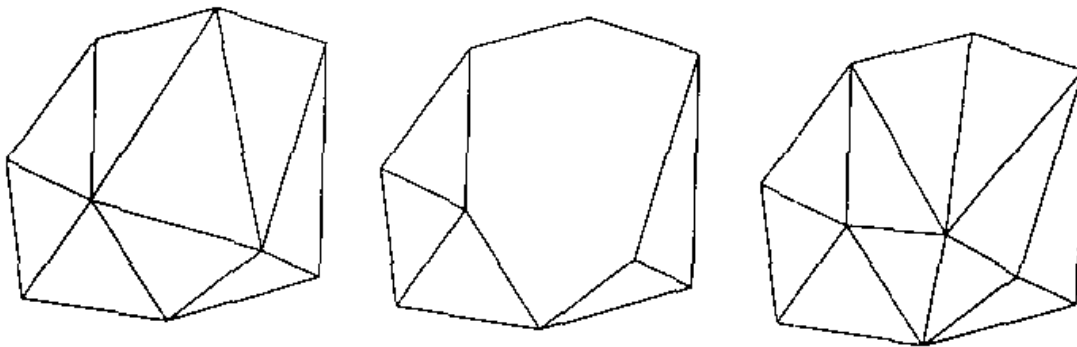


Figura 2. Reconexión de una malla existente alrededor de un nuevo punto insertado con el algoritmo de Bowyer-Watson (Rebay, 1993).

La eficiencia del algoritmo depende de la rapidez con la que se realice la búsqueda de los triángulos que se eliminarán en cada inserción de puntos y esto se hace mucho más fácil gracias al conocimiento de los triángulos adyacentes a cada triángulo. Dado que todos los triángulos que se eliminarán son siempre contiguos, se puede usar una búsqueda de árbol entre los triángulos vecinos para encontrar todos los demás triángulos que se eliminarán después del primero. En el caso típico, el número de triángulos que se eliminarán en cada inserción de punto no depende del número total de

triángulos existentes. Si la información correspondiente a los triángulos vecinos está disponible y se emplea la búsqueda multidimensional del primer triángulo que se eliminará, el algoritmo puede calcular la triangulación de Delaunay de un conjunto de n puntos en $O(n \log n)$ Operaciones. (Rebay, 1993)

Divide y vencerás

La técnica de divide y vencerás consiste en dividir el conjunto de puntos a triangular en 2 partes, resolver recursivamente cada mitad y juntar las soluciones de los subproblemas para obtener una solución general. Los puntos se presentan ordenados según su primera coordenada. Al tratarse nuevamente de un método en el que la complejidad recae en la rapidez con la que se buscan los puntos en la estructura de datos, el tiempo en el que se triangulan n puntos es $O(n \log n)$. (Martínez-Rodríguez, 2015)

La segmentación

El algoritmo de crecimiento de región con semilla es recursivo y su complejidad varía dependiendo de la estructura de información seleccionada y del tamaño mínimo de las regiones.

Otro aspecto que considerar es que el algoritmo de crecimiento de región puede implementarse de maneras diferentes. Por ejemplo, si se necesitara segmentar todas las figuras pertenecientes a la escena se necesitan varias semillas. Una vez que una semilla terminó de formar una región, la siguiente semilla encontrara el espacio de búsqueda reducido, pues no necesita recorrer los vóxeles que ya fueron asignados, resultando en una complejidad logarítmica.

Para esta investigación se considera que solo habrá **una semilla**, por lo que su complejidad es lineal, es decir que se calcula respecto del número de vóxeles de la imagen y no respecto del número de semillas-regiones. Para una imagen con n vóxeles se realizan n operaciones para determinar si se agregan o no a la región. Cuanto más grande sea la imagen, más operaciones se realizarán.

A esta complejidad se le suma el tiempo que le toma a la máquina leer y actualizar la estructura de datos en la que se presenten los vóxeles.

En este caso, los vóxeles están ordenados en una **matriz**, pero no siempre es así. Por ejemplo, la información 3D puede almacenarse en un **octree**, que es árbol que divide el espacio 3D en 8 cuadrantes, cada cuadrante es dividido a su vez en 8 nodos que son divididos sucesivamente.

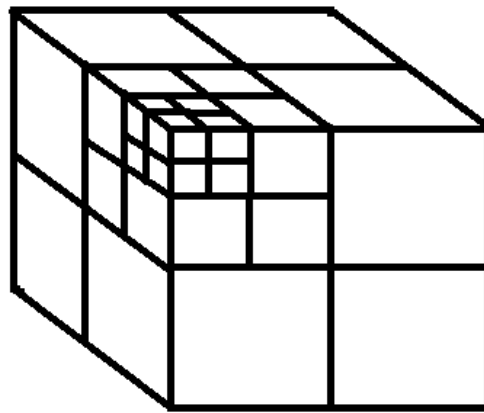


Figura 3. En el *octree* cada nodo puede estar dividido en 8 partes.



Figura 4. Octree (PCL. 2019)

El octree está diseñado para mostrar la figura con diferentes niveles de resolución (Rajasekaran & Reif, 2007) y por ende con diferentes velocidades de renderizado. Cuanto más detalle se muestre, más memoria se utilizará y más lento será su procesamiento.

En estructuras de datos tipo árbol y lista el tiempo de exploración es de orden logarítmico; pero en este caso, el uso de una matriz añade una complejidad constante pues el acceso a la memoria es aleatorio.

Además, hay que considerar si las regiones pueden ser de tamaño 1, o si los vóxeles sueltos son considerados ruido. Ese aspecto genera diferencias significativas en el número de figuras segmentadas.

En este caso se considera que el usuario solo mediría un clado a la vez, y que por tanto el algoritmo de segmentación parte de una sola semilla previamente determinada. También se decidió ignorar el ruido, es decir que las regiones tengan un tamaño mayor a 1.

El costo computacional del tratamiento digital de imágenes 3D es tan alto que el público en general ha tardado en adoptarlos a su vida diaria. Si bien existían los fundamentos teóricos, es en la última década que los procesadores gráficos y los elementos de hardware de los equipos especializados pudieron miniaturizarse al grado de constituir aparatos de consumo.

CONCEPTOS BÁSICOS

La **morfología** vegetal es la rama de la biología que estudia la forma externa de las plantas mientras que la **anatomía**¹ su estructura interna. El conocimiento de la forma y estructura de la planta es esencial para poder entender su funcionamiento (Müller 2000).

Los agricultores monitorean el tamaño de las plantas para evaluar su estado y detectar deficiencias en su manejo o predecir el momento ideal de la cosecha. De las plantas puede medirse, por ejemplo, el largo, ancho y área de sus hojas, el peso de las hojas frescas, el peso de las hojas secas o el número de hojas por rama.

El **área foliar** se refiere a la superficie que abarca un sólo lado de una hoja. Esta variable representa la cantidad de material de hojas en los ecosistemas y controla los vínculos entre la biosfera y la atmósfera a través de diversos procesos como la fotosíntesis, la respiración, la transpiración y la intercepción de la lluvia (López-Tolentino *et al*, 2016).

Los sistemas de visión artificial son utilizados en la agricultura para resolver problemas de trabajo estacional de alto volumen. Se utilizan para detectar y localizar los frutos en las ramas, determinar la madurez de los frutos, determinar la geometría de la copa de los árboles y localizar un árbol específico en la huerta (Zujevs *et al*, 2015).

Para esta investigación se trabajó con una cámara 3D **Intel RealSense R200** que fusiona la información de dos imágenes infrarrojas para formar un modelo estereoscópico. La estrategia desarrollada permite medir las plantas sin interferir con su crecimiento y muestra el resultado inmediatamente. El uso de equipo especializado reduce la posibilidad de cometer errores propios de un proceso manual.

2.1 La morfología vegetal

La morfología vegetal es la rama de la biología que estudia la forma externa de las plantas (Müller 2000). “La constitución genética de las plantas y el ambiente son dos grandes fuerzas entre las cuales ha intervenido la selección natural para generar diferentes tipos de estructuras vegetales. Desde el punto de vista evolutivo, las hojas han sido las más plásticas, por lo que numerosas formas se han desarrollado en respuesta a los variados tipos de presión de selección” (Kraus *et al*, 2004).

¹ La anatomía incluye a la citología que es el estudio de las células y la histología que es el estudio de los tejidos

2.2 Las partes de la hoja

Las hojas son los órganos laterales del tallo que poseen capacidades fotosintéticas (Kraus et al, 2004).

Partes de las hojas (Kraus et al, 2004):

Lámina o limbo: es la parte más visible de la hoja, es generalmente delgada, aplanada y expandida con una superficie superior (**haz**) y una inferior (**envés**). La lámina puede encontrarse entera en las hojas simples o dividida en **foliolos** en las hojas compuestas.

Base foliar: es la parte proximal de una hoja. Cuando se desarrolla tanto que abraza al **tallo** se denomina **vaina**.

Pecíolo: es un eje que une la base foliar con la lámina. En las hojas compuestas el pecíolo se divide en varios **raquis**.

Estípulas: son dos pequeños apéndices en el inicio de la base foliar a veces transformados en **espinas**.

Estípelas: son dos pequeños apéndices en la base de los foliolos a veces transformados en espinas.

Ócrea: es una formación membranosa que rodea al tallo por encima de la inserción de cada hoja. Se considera que se trata de 2 estípulas soldadas.

Lígula: es un apéndice membranoso o piloso en la región que une la lámina y la vaina. Evita la penetración de agua entre la vaina y la caña.

Aurículas: son 2 apéndices en la base de la lámina y a los costados de la lígula que abrazan al tallo.

Nervaduras: son los haces vasculares que recorren la hoja en toda su longitud.



Figura 5. Hoja compuesta de *Jacaranda Caerulea* con: (1) base foliar, (2) peciolo, (3) raquis, (4) foliolos.

2.3 Los tipos de hojas

Las hojas de las plantas se dividen en simples y compuestas dependiendo de si tienen uno o más foliolos. La **figura 6** muestra los distintos tipos de hojas simples.

Clasificación de las hojas: 1.-Flabelada, 2.- Rómbica, 3.-Asimétrica, 4.-Cletrada, 5.-Falcada, 6.-Linear, 7.-Espatulada, 8.- Lanceolada, 9.-Oblongo, 10.-Oval, 11.-Obovado, 12.-Runcinada, 13.-Escamiforme, 14.-Ovado, 15.-Reniforme, 16.-Acorazonada, 17.-Lirado, 18.-Sagitada, 19.-Acicular, 20.-Orbicular, 21.-Deltoide.



Figura 6. Clasificación de las hojas (Con imágenes de Kraus et al, 2004; e imágenes propias)

2.4 Las técnicas de medición

Los agricultores monitorean el tamaño de las plantas para evaluar su estado y detectar deficiencias en su manejo o predecir el momento ideal de la cosecha.

De las plantas puede medirse, por ejemplo, el largo, ancho y área de sus hojas, el peso de las hojas frescas, el peso de las hojas secas o el número de hojas por rama. Las longitudes máximas son diferentes para cada tipo de hoja (Figura 7).

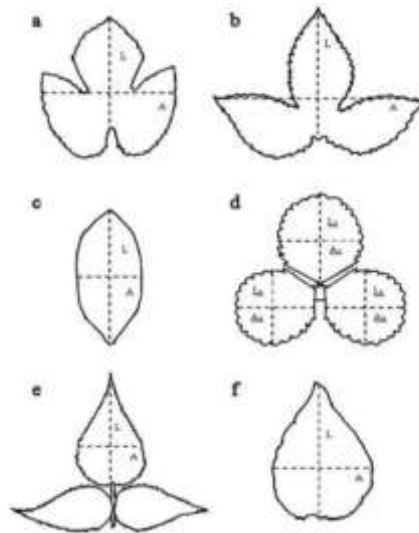


Figura 7. Diagrama de la forma de las variables Largo (L) y ancho (A) en hojas de a) Brevo²; b) Curuba³; c) Feijoa⁴; d) Fresa⁵; e) Mora⁶ y f) Uchuva⁷ (Casierra-posada et al, 2007)

En los nopales, las hojas se encuentran transformadas en espinas y son los cladodios los que pueden medirse (Figura 8).

Para evaluar el éxito de los injertos⁸ puede analizarse el largo del corte, el diámetro del corte, el número de hojas en el esqueje, el área promedio de las hojas, el peso de las hojas frescas, el peso de las hojas secas, el porcentaje de éxito y el número de brotes nuevos (Ramhari et al, 2015).

² Brevo: Higuera; *Ficus carica*.

³ Curuba: *Passiflora mollissima*.

⁴ Feijoa: Guayaba de Brasil; *Acca sellowiana*

⁵ Fresa: *Fragaria vesca*.

⁶ Mora: *Rubus glaucus*

⁷ Uchuva: *Physalis peruviana*.

⁸ Injerto: Método que consiste en disponer sobre un vegetal, una parte viva de otro, a fin de que se identifiquen y formen un solo individuo. (Barioglio, 2006)

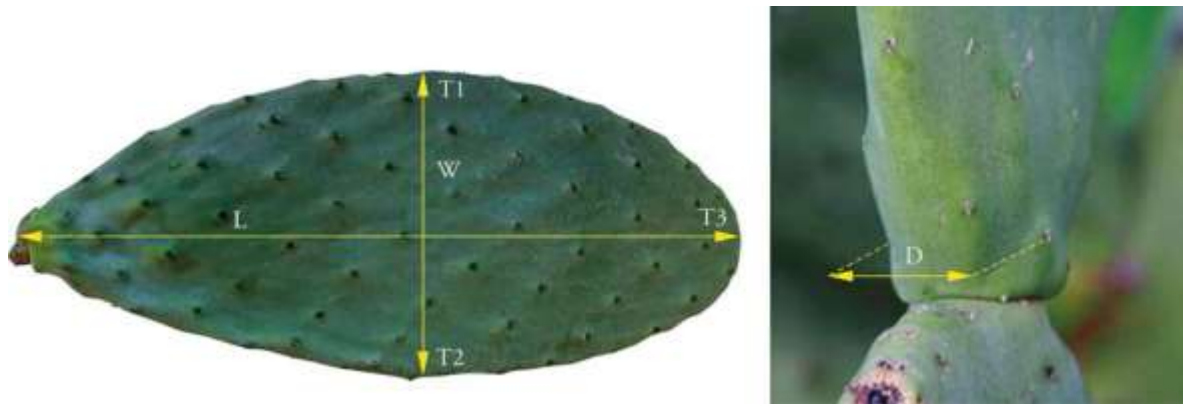


Figura 8. Parámetros biométricos medidos en los cladodios de O. Ficus-Indica. L: Largo; W: Ancho; T: Espesor medio (obtenido de la media aritmética de los puntos T1, T2 y T3); D: Diámetro del cuello (Gaspar et al. 2018).

Las técnicas de medición de plantas se dividen en **destructivas** y **no destructivas** dependiendo de si la muestra es separada o no de la planta. Los métodos de medición de frutas u hojas que requieren cortar muestras interfieren con el desarrollo de la planta e imposibilitan medir más de una vez el mismo espécimen.

Técnicas de medición destructivas: La muestra es separada de la planta

1. Integrador de laboratorio
2. Fotografías a una distancia conocida
3. Fotografías con un fondo de color o patrón específicos
4. Escaneo 3D en laboratorio

Técnicas de medición no destructivas: La muestra no es separada de la planta

1. Integrador portátil
2. Estimación indirecta
3. Fotografía con fondo de color particular y sensor ultrasónico

ESTADO DEL ARTE

3.1 Estado del Arte

Los escáneres especializados en la medición de hojas son llamados **integradores**. Hay integradores portátiles y de laboratorio. El precio de estos equipos ronda los 10,000 USD (Kosmos Scientific de México, 2018). Cada modelo de integrador mide tamaños y formas específicas de hojas (**figura 9**).



Figura 9. Integradores portátiles CID Bio Science (CID Bio-Science Inc., 2018)

El área foliar también puede ser estimada mediante **modelos matemáticos**. En ellos, el área es estimada a partir de la multiplicación del largo y ancho de la hoja por un coeficiente de proporcionalidad. También puede estimarse a partir del **peso** y/o **volumen** de la pieza o de una sección de ella, como en la técnica de **corte de discos de hoja** (Leaf disc) (Mendes *et al.*, 2015).

Por ejemplo, en su trabajo, *Astegiano et al* (2001) establecen que el Área (A) de las hojas de jitomate (*Solanum lycopersicum*) está dada por la fórmula en la que l_m representa el largo máximo y am el ancho máximo (Astegiano et al, 2001):

$$A = 0.364 \times (lm \times am) - 9.31 \quad (1)$$

Sin embargo, el modelo matemático solo sirve para una especie en particular y la operación no considera la presencia de deformaciones o mutilaciones en las hojas. La estimación matemática es una alternativa al uso de equipo especializado de medición, pero depende a su vez de otras mediciones que, de realizarse manualmente, pueden propagar errores. La exactitud de un modelo estadístico depende de la técnica empleada y de la calidad de las muestras utilizadas para su desarrollo.

La **planimetría** es otro método de medición. Para ello, la silueta de la muestra se marca en un papel para poder medirla sin tener que cortar el espécimen.

En los cladodios también puede medirse su **inclinación**. *Sortibrán et al.* (2005) estudiaron la relación entre la inclinación inusual de los cladodios de *Opuntia Puberula Pfeiffer* (que crecen horizontalmente) y su efecto en la captación de luz.

Puede medirse la cantidad de **radiación solar** que es interceptada por la planta y no llega al suelo mediante equipo especializado (Mendes, 2015).

Finalmente, la determinación del área foliar a través de la toma de **fotografías 2D** es un método no destructivo. Para esta técnica la muestra se aplana entre 2 hojas de acrílico, una transparente y otra con un patrón que permita determinar la escala en la imagen. Utilizando algún software de edición de imágenes se cuentan los píxeles y se convierten a unidades cuadradas. Esta técnica requiere que la muestra esté paralela a la cámara al momento de tomar la fotografía (**figura 10**).



Figura 10. Determinación del área foliar en hojas de *Manihot Esculenta* (Cock et al, 2014)

Los sistemas de visión artificial son utilizados en la agricultura para resolver problemas de trabajo estacional de alto volumen. Se utilizan para detectar y localizar los frutos en las ramas, determinar la madurez de los frutos, determinar la geometría de la copa del árbol y localizar el árbol en la huerta (Zujevs et al, 2015).

Por ejemplo, en *Andújar et al* (2016) se utiliza un sensor 3D kinect© para escanear una coliflor (*Brassica Oleracea*) y posteriormente tratar la nube de puntos con un software para obtener un modelo completo (**Figura 11**).



Figura 11. Escaneo 3D de una planta de coliflor (*Brassica oleracea*) (Andújar et al, 2016)

Esto, porque debido a que las coliflores están cubiertas de hojas, es difícil determinar su nivel de maduración y el momento ideal para cosechar. Ahora, el mejor momento para el corte de frutas individuales puede detectarse utilizando estos modelos, lo que permitiría la optimización de la cosecha e incrementaría el rendimiento.

Esta investigación muestra el potencial de las cámaras de profundidad para ser utilizadas como una herramienta precisa para estimar el grado de maduración de un cultivo y así optimizar su rentabilidad (Andújar et al, 2016).

En algunos casos la necesidad de una conexión eléctrica, equipo de cómputo y evitar la luz solar directa dificultan su uso en el campo. Por ejemplo, en *Okayamaa et al (2015)* se menciona la necesidad de la construcción de un armazón que mantenga la cámara en una posición específica. En este caso se utiliza una cámara 3D de la compañía Asus©.

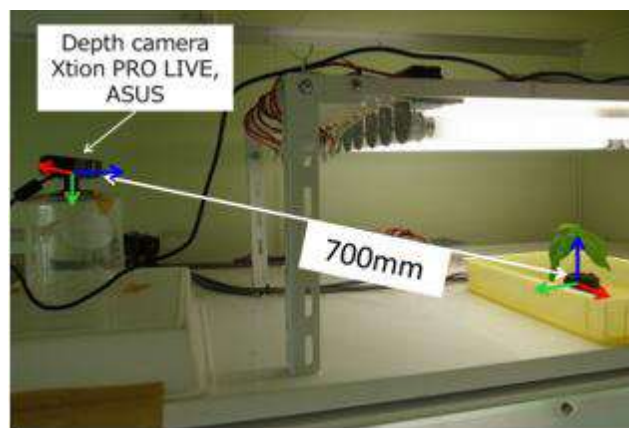


Figura 12. monitoreo del crecimiento de una planta de frijol (*Phaseolus vulgaris*) (Okayamaa et al, 2015).

En otros, la luz solar interfiere con su funcionamiento pues utilizan equipos basados en luz infrarroja. En *Sarbolandi et al (2015)* vemos otro ejemplo de escaneo de plantas. En este caso la nube de puntos se procesó con un API llamada *point cloud library* PCL (Cousins, 2013) que se abordará más adelante. En la **figura 13** se observa que fue

necesaria una estructura metálica para que el sensor se colocara en una posición específica. En la **figura 14** vemos el resultado del escaneo.



Figura 13. Caracterización rápida de estructuras vegetales con un sensor Kinect de Microsoft (Sarbolandi et al, 2015)

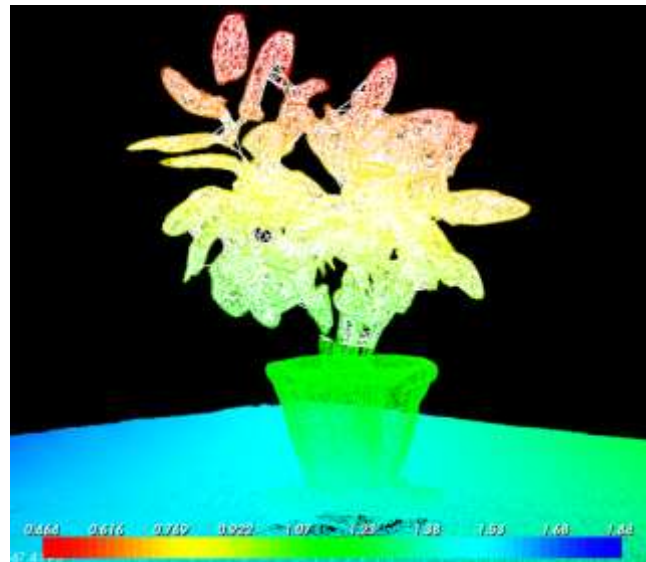


Figura 14. Resultado del Escaneo de una planta (Sarbolandi et al, 2015)

El diseño de estrategias y equipos para la medición de especies vegetales se enfrenta a la vastedad de especies y formas de las plantas y a los retos que plantea el trabajo en campo. Conforme la tecnología mejora, aumenta la presencia de sistemas de visión artificial aplicados a la agricultura.

3.2 Las imágenes 3D

Las cámaras de tercera dimensión se utilizan en los sistemas de visión artificial de automóviles autónomos y robots, en el escaneo de terrenos agrícolas, para inventariar obras de arte y restos arqueológicos, en el reconocimiento biométrico e incluso en videojuegos y realidad aumentada.

3.2.1 Formatos de imagen 3D

La unidad mínima de información de las imágenes 2D, es el **pixel**. En las imágenes 3D, los pixeles incluyen información de su posición en el espacio. Los pixeles que poseen información sobre su posición en el espacio se convierten en **vóxeles**. Cuando los vóxeles registran su posición respecto de un plano cartesiano, a la imagen se le llama **nube de puntos**. Cuando los vóxeles contienen la distancia a la que se encuentran de la cámara entonces el archivo es un **mapa de profundidades**.

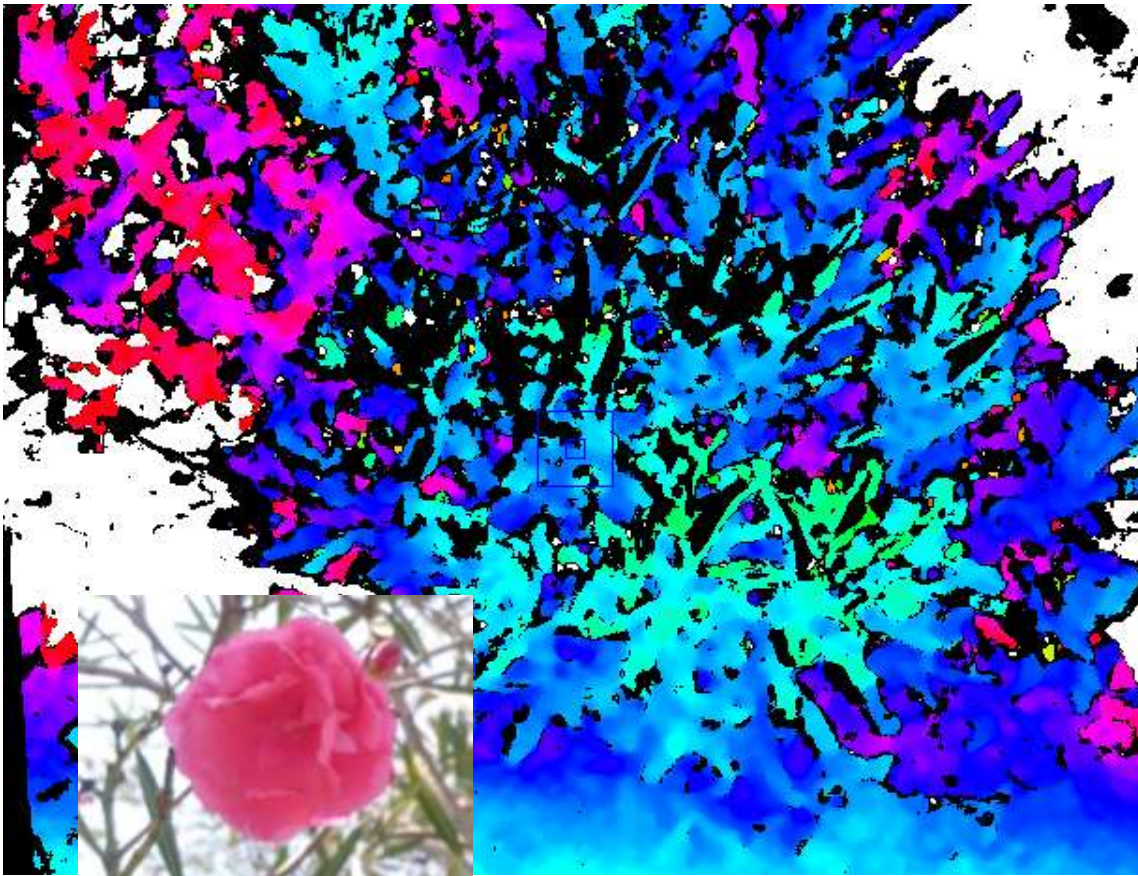


Figura 15. Imagen 3D de arbusto de adelfa (*Nerium Oleander*) y su flor en detalle

Las imágenes en 2D tienen uno o varios **canales de color**. El mapa de profundidades tiene además un **canal de profundidad**. Los pixeles de una imagen 2D están ordenados en una matriz de pixeles determinada por los ejes X e Y. En las imágenes 3D los vóxeles

tienen además la posición X, Y, Z del punto en el espacio. Los valores de X e Y de la posición de un punto en la matriz, no son los mismos que del punto en el espacio.

La posición de un vóxel en una imagen 2D (generalmente) se contabiliza en números naturales a partir de la esquina superior izquierda. En dichas imágenes el eje Y va de arriba para abajo (**figura 16**).

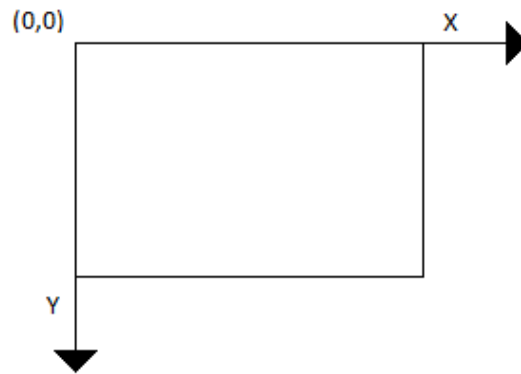


Figura 16. Posición X e Y de los píxeles en la imagen.

En las imágenes 3D la posición en el espacio se marca en centímetros o pulgadas desde un plano cartesiano colocado al centro de la imagen y a la altura de la cámara (**figura 17**). El eje Y va de abajo hacia arriba.

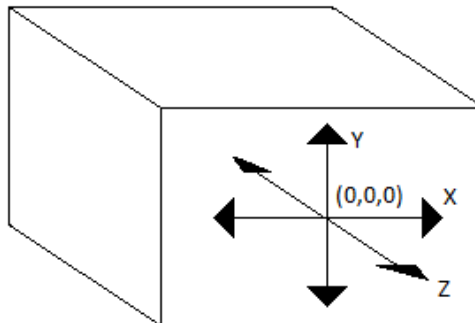


Figura 17. Posición X, Y y Z de los vóxeles en el espacio.

En las fotografías 3D el valor 0 del eje Z señala la cámara, así que si se utiliza la cámara frontal de los dispositivos el valor de Z será negativo. Al momento de representar la imagen 3D debe saberse si la fotografía fue tomada con la cámara trasera o delantera.

Las imágenes 3D pueden representarse como una imagen en **escala de grises**, en la que los tonos más claros están más cerca de la cámara, porque son directamente proporcionales a la intensidad de la luz infrarroja reflejada sobre los objetos (**figura 18**).



Figura 18. Imagen a color y mapa de profundidades en escala de grises de dos plantas de lechuga (*Lactuca Sativa*)

Existen varios formatos en los que puede almacenarse la información de las imágenes 3D, tanto abiertos como propietarios.

Cuando la imagen 3D está en escala de grises puede guardarse como tal y su factor de conversión almacenarse en sus metadatos. También pueden agregarse los bytes de la imagen 3D a los metadatos de la imagen 2D. Esto permite que el usuario que recibe la imagen pueda revisar su contenido sin necesidad de un software especial, aunque sí lo necesitaría si quiere visualizar el contenido en 3D.

Por otro lado, para las nubes de puntos hay muchos formatos propietarios por que las aplicaciones con las que se visualizan tienen diferentes motivaciones, desde la animación 3D hasta la arquitectura y manufactura.

Algunos formatos 3D son (PCL, 2019):

- **PLY**: un formato de archivo de polígonos, desarrollado en la Universidad de Stanford (Burkardt, 2012).
- **STL**: un formato de archivo nativo del software CAD de estereolitografía⁹ creado por 3D Systems ©.
- **OBJ**: un formato de archivo de definición de geometría desarrollado por *Wavefront Technologies* ©.
- **X3D**: el formato de archivo basado en XML de la norma ISO para representar datos de gráficos de computadora en 3D

En el campo de la visión artificial destaca el formato de código abierto PCD (Point Cloud Data) establecido por la **Open Perception Foundation** para aplicaciones que utilizan su biblioteca de algoritmos PCL (Point Cloud Library) (Rusu y Cousins, 2011). El formato PCD almacena metadatos tales como el histograma (utilizado en la descripción de características).

⁹ Impresión 3D en la que un láser UV solidifica una resina de fotopolímero (3D Systems, 2019)

3.2.2 Las tecnologías 3D

Existen 3 tecnologías de visión artificial: la de proyección de luz estructurada, la de visión estereoscópica y el mapeo del tiempo de vuelo.

En la estrategia de proyección de **Luz estructurada** se toman varias fotografías desde diferentes ángulos a una muestra sobre la que se proyecta un conjunto sucesivo de patrones de sombras. Posteriormente se analiza la forma en que las sombras se deforman sobre la superficie para recrear la figura en tercera dimensión. Esta estrategia está diseñada para ser utilizada en interiores y necesita que el objetivo permanezca inmóvil dado que requiere tomar varias fotografías (**figura 19**).



Figura 19. Escáner 3D de luz estructurada HP Sprout© (HP Development Company, L.P, 2016)

Por su parte, las cámaras **tiempo de vuelo (ToF)** crean un mapa de las distancias a cada uno de los píxeles de una imagen registrando el tiempo que le toma a un haz de luz infrarroja rebotar en los objetos y regresar a la cámara.



Figura 20. Proyección de puntos de luz infrarroja en una planta de Níspero (*Eriobotrya japonica*) utilizando un sensor Kinect TOF de microsoft

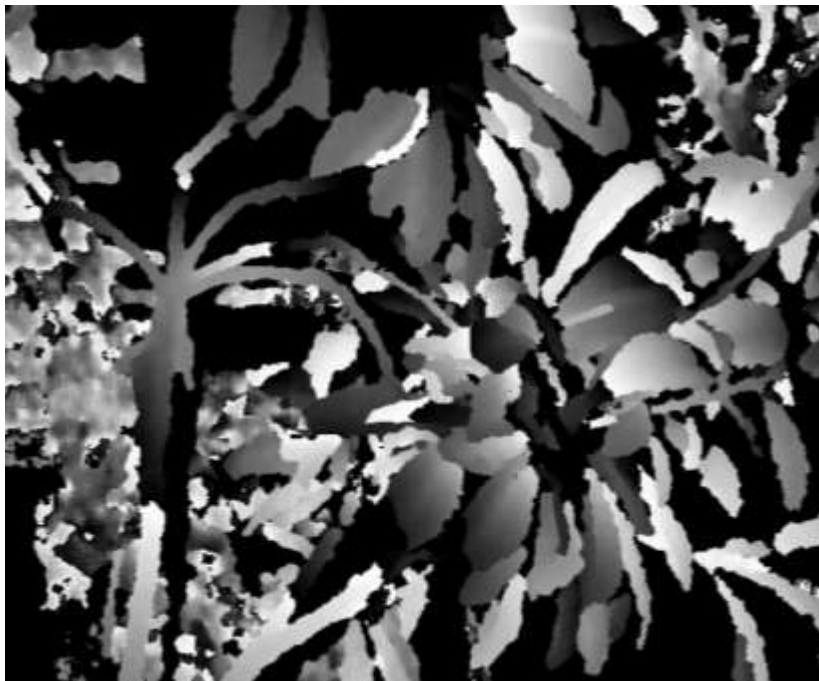


Figura 21. Mapa de profundidades de una planta de Níspero (*Eriobotrya japonica*) obtenida de un sensor Kinect TOF de Microsoft

En la **visión estereoscópica** dos o más cámaras captan la misma escena y por triangulación [de los puntos de interés] se determina la distancia a los vóxeles (Pajares, 2008). La visión estereoscópica tiene mayor complejidad computacional que la TOF, pero las imágenes tienen mayor resolución porque se utilizan cámaras normales de cualquier capacidad.



Figura 22. ARCore, el SDK de realidad aumentada de Google funciona en dispositivos móviles con 2 cámaras (Google, 2018)

Cuadro 1. Comparación de tecnologías 3D que pueden utilizarse en campo

Técnica	Equipo	Resolución	Complejidad computacional
Visión binocular	2 cámaras	Mayor. La resolución puede ser tan alta como la cámara que se utilice, pero la exactitud depende del software.	Mayor
Tiempo de vuelo	Proyector laser y cámara infrarroja especializados	Menor. Es difícil miniaturizar los sistemas laser, pero las mediciones son más precisas.	Menor

3.3 El tratamiento digital de imágenes 3D

Los píxeles de una imagen pueden conceptualizarse como una matriz de valores numéricos. En el sistema RGB los colores se obtienen de la combinación de 3 tonos de rojo, verde y azul cuyo rango va desde 0 (que representa la ausencia de color) hasta 255.

El mapa de profundidades, la profundidad puede considerarse como un canal extra para las imágenes 2D en el que los valores representan la distancia a la cámara en centímetros o pulgadas. Así que pueden analizarse como si se tratara de una imagen en escala de grises.

En las nubes de puntos los vóxeles poseen un componente X, Y y Z que establecen su posición en el espacio y no necesariamente están ordenados en forma de matriz. El análisis en las nubes de puntos se centra en su unión y segmentación, así como en su representación.

Una vez que la imagen 3D es adquirida, las siguientes etapas de su análisis son:

- Filtrado.
- Segmentación.
- Reconstrucción.
- Rellenado.
- Visualización.

3.3.1 Filtrado

En la etapa de filtrado, la matriz que representa una imagen es multiplicada mediante **convolución** por otra matriz (**kernel**) para aumentar o disminuir algunos de sus valores. Como resultado, algunos atributos se resaltan o desaparecen.

En la convolución, se seleccionan porciones de la matriz del mismo tamaño que el kernel y sus elementos se multiplican entre sí. Los resultados se suman y se escriben en la matriz del resultado.

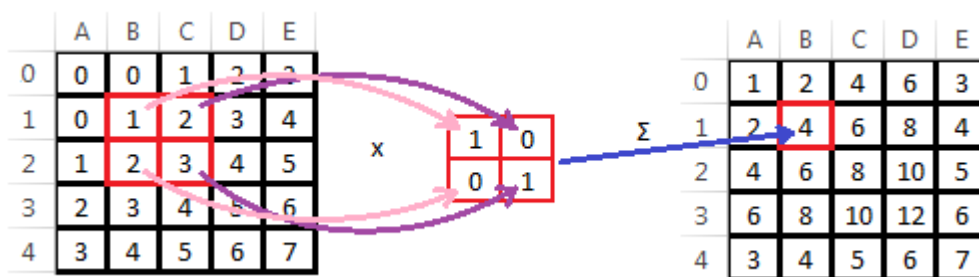


Figura 23. Convolución

El filtrado se utiliza para eliminar el ruido de una imagen, aumentar el contraste, resaltar los bordes, segmentar una figura, etc.

3.3.1.1 Detección de bordes.

Un borde es el límite entre zonas con distintas características. Los bordes se detectan con la aplicación mediante convolución del operador de **Roberts Cross**. El operador de Roberts Cross utiliza dos kernels que diferencian los valores de los pixeles de forma diagonal, los kernels se denominan M- y M+.

Kernels M- y M+ para el operador de Roberts Cross:

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (2)$$

Al implementar la convolución, el valor máximo arrojado por la aplicación de dichos patrones es guardado como el valor del borde en ese punto. El punto $E_{x,y}$ en el borde es entonces el máximo de los dos valores absolutos derivados de la convolución de los dos kernels en un punto $P_{x,y}$ de la imagen (Nixon & Aguado, 2002)

Detección de bordes:

$$E_{x,y} = \max\{|M^+ * P_{x,y}|, |M^- * P_{x,y}|\} \quad \forall x, y \in 1, \dots, N - 1 \quad (3)$$

Con la aplicación del operador de Robert Cross, desaparecen de la matriz los conjuntos de 4 pixeles contiguos con valores similares y persisten las zonas con al menos un pixel diferente. Antes de aplicar el kernel hay que binarizar la matriz, es decir, utilizar solo dos valores (0 y 1) para representarla (**figura 24**).

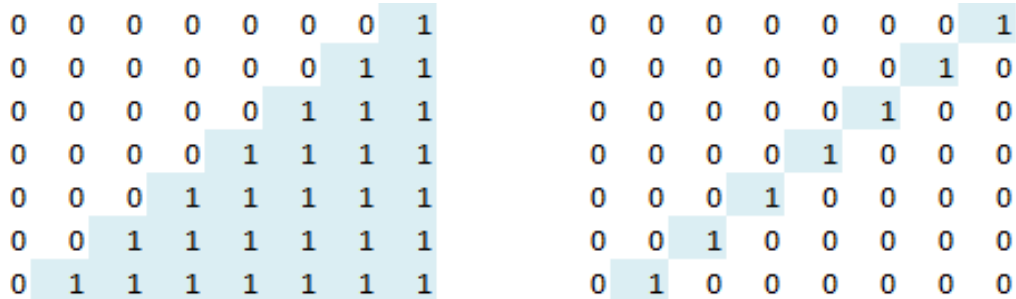


Figura 24. Aplicación del operador de Robert Cross para detectar bordes

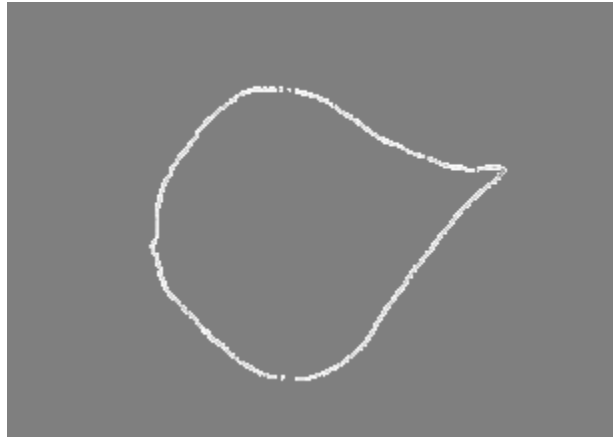


Figura 25. Bordes obtenidos mediante la convolución en una imagen de una hoja a blanco y negro

Originalmente el fondo es obscuro y el borde claro pero los valores se invierten para presentarlos al usuario (**figura 26**).

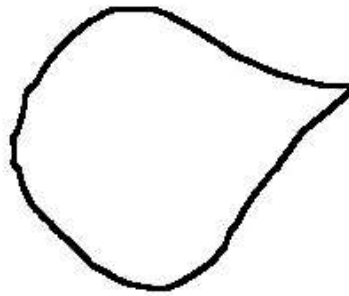


Figura 26. Bordes obtenidos mediante la convolución en una imagen a blanco y negro de una hoja, tonalidad invertida

3.3.1.2 Detección de esquinas

Una esquina es el punto en el que la dirección del borde cambia más allá de cierto número de grados. La detección de esquinas puede ayudarnos a reconocer el ápice y las puntas de las hojas.

Las matrices mostradas a continuación son un ejemplo de secciones que pueden encontrarse en la matriz numérica que representa una imagen, los valores reales van del 0 al 256 pero se presenta binarizada para su análisis.

Matrices que representan una zona plana, un borde y una esquina:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & \mathbf{1} & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 \\ 0 & \mathbf{1} & 0 & 0 \\ \mathbf{1} & 0 & 0 & 0 \end{bmatrix} \quad (4)$$

Para detectar esquinas se aplica la convolución con un kernel que represente la esquina, tal como los presentados a continuación:

Kernels para la detección de esquinas:

$$\begin{bmatrix} -1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad (5)$$

3.3.2 Segmentación

La segmentación se refiere a la separación de la escena en componentes. Cuando la información está en escala de grises la imagen es tratada como matriz numérica donde cada pixel es representado con un número que va del 0 al 256 y que representa tonos de oscuros a claros respectivamente.



Figura 27. Imagen de una hoja en blanco y negro

3.3.2.1 Etiquetado de regiones

Los pixeles que quedan encerrados entre bordes constituyen **regiones**. La separación de una imagen en regiones puede lograrse etiquetando los pixeles. El etiquetado de regiones es un proceso en el que la imagen se recorre y sus pixeles se marcan con la etiqueta de sus vecinos a menos que se tope con un borde o haya que fusionarse con otra región (**figura 28**).

a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	
a	a	a	a	a	0	0	0	0	0	a	a	a	a	a	a	a	a	
a	a	a	a	0	0	b	b	b	0	0	a	a	a	a	a	a	a	
a	a	a	0	0	b	b	b	b	b	0	0	0	a	a	a	a	a	
a	a	a	0	b	b	b	b	b	b	b	b	0	0	0	0	0	a	a
a	a	a	0	b	b	b	b	b	b	b	b	b	b	0	0	a	a	a
a	a	0	0	b	b	b	b	b	b	b	b	b	b	0	0	a	a	a
a	a	0	b	b	b	b	b	b	b	b	b	b	0	0	a	a	a	a
a	a	0	b	b	b	b	b	b	b	b	b	0	a	a	a	a	a	a
a	a	0	b	b	b	b	b	b	b	b	0	0	a	a	a	a	a	a
a	a	0	0	b	b	b	b	b	b	b	0	a	a	a	a	a	a	a
a	a	a	0	b	b	b	b	b	b	b	0	0	a	a	a	a	a	a
a	a	a	0	0	b	b	b	b	b	b	0	a	a	a	a	a	a	a
a	a	a	a	a	0	0	b	b	0	0	a	a	a	a	a	a	a	a
a	a	a	a	a	a	0	0	0	0	a	a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a

Figura 28. Píxeles de una imagen etiquetados como pertenecientes a dos regiones, a y b

3.3.2.2 Crecimiento de región con semilla

Otra opción de segmentación es utilizar un algoritmo de crecimiento de región, puede utilizarse cuando se conoce el lugar en el que un segmento está situado.

El algoritmo de crecimiento de región comienza con un conjunto de puntos *semilla* y permite que las *regiones* crezcan al agregar los píxeles vecinos que tienen propiedades similares; por ejemplo: intensidad, textura o color. Dos problemas serios en esta técnica son: la selección de las semillas iniciales que representan adecuadamente las regiones de interés, y la selección de propiedades adecuadas para asignar los puntos a las diversas regiones durante el proceso de crecimiento. Se puede mencionar un tercer problema: el criterio para detener el algoritmo. En general, impediremos que una región crezca cuando no haya más píxeles que satisfagan los criterios de inclusión en esa región. El uso de este descriptor implica que hay un conocimiento a priori de al menos un *modelo espacial* de los resultados esperados (Pajares et al, 2008).

3.3.3 Reconstrucción

La reconstrucción es la unión de varias nubes de puntos capturadas desde distintas perspectivas para formar una sola figura.

3.3.4 Rellenado

En la fase de relleno, se exploran estrategias que permitan que la nube de puntos forme figuras cerradas, por ejemplo, mediante la definición de aristas.

3.3.4.1 Triangulación

Triangular un conjunto S de n puntos en el plano se refiere a unir los puntos mediante segmentos de líneas rectas que no se intersecan de modo que cada región al interior de la estructura convexa de S sea un triángulo (Chen, 1996).

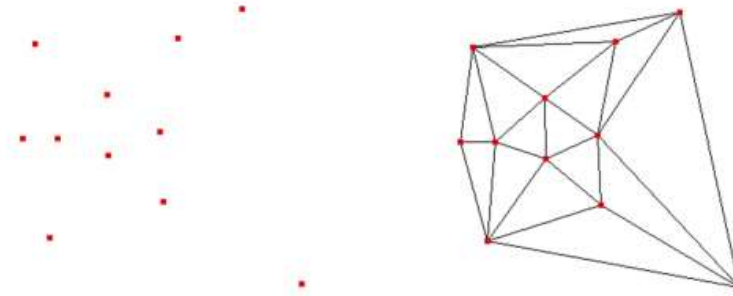


Figura 29. Triangulación de una nube de puntos

3.3.4.2 Triangulación de Delaunay

El mismo conjunto de puntos puede triangularse de maneras distintas. La condición de **Delaunay** garantiza la construcción de una triangulación lo más cercana posible a la equilátera (Maximiza el ángulo mínimo de cada triángulo de la triangulación).

3.3.4.3 Condición de Delaunay

Dado un conjunto de puntos en el plano, tres puntos definen un triángulo de Delaunay si el círculo que circunscribe a dicho triángulo no contiene ningún otro punto del conjunto de datos. En la **figura 30** se observan dos triangulaciones diferentes para los mismos 4 puntos "a", "b", "c" y "d". La primera triangulación no cumple la condición de Delaunay porque el punto "d" queda encerrado en el círculo formado por los otros tres puntos. La segunda triangulación si cumple la condición de Delaunay porque el punto c queda fuera del círculo formado por los puntos "a", "b" y "d".

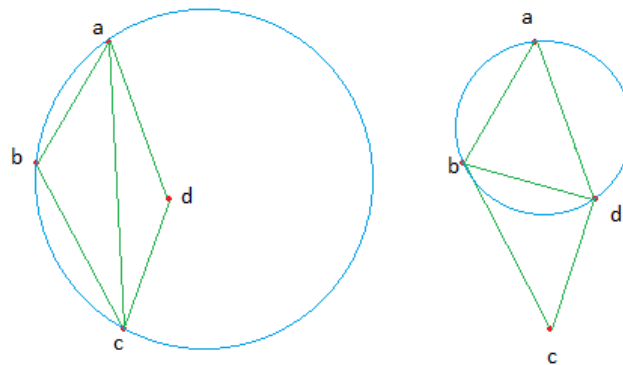


Figura 30. Dos ejemplos de triangulación para el mismo conjunto de datos. El segundo ejemplo cumple la condición de Delaunay

En espacios tridimensionales se considera una esfera circunscrita en lugar de un círculo.

3.3.4.4 Pertenencia de un punto a un triángulo

Un punto pertenece a un triángulo si recorriendo sus vértices en el sentido de las manecillas del reloj, el punto se encuentra del mismo lado para los tres. En la siguiente imagen, para el primer triángulo (A) el punto evaluado pertenece al triángulo por que se encuentra a la derecha de los tres vértices. en el segundo triángulo (B) el punto se encuentra a la izquierda de uno de sus vértices.

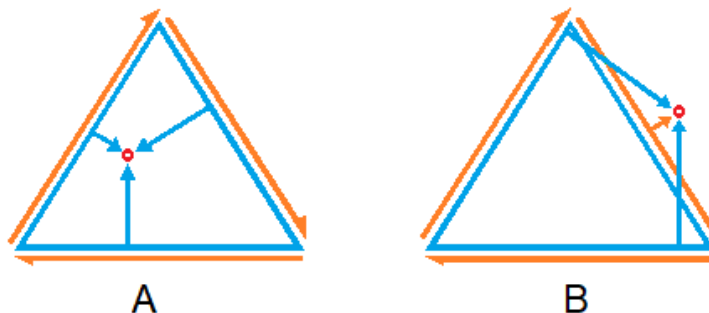


Figura 31. Pertenencia de un punto a un triángulo.

Para el caso de figuras en tercera dimensión, la triangulación evalúa si los puntos pertenecen a una esfera, en lugar de a un círculo.

3.3.4.5 Área de un triángulo dados 3 vértices

Puede medirse el área de los triángulos resultantes para obtener el área de la superficie completa.

El área de un triángulo dados 3 vértices P_1, P_2, P_3 es igual a la mitad del producto escalar de los vectores u y v (Kolman, 2006):

$$A_T = \frac{1}{2} \|u \times v\| \quad (6)$$

Donde u señala el vector que va de P_1 a P_2 (base del triángulo) y v el vector que va de P_1 a P_3 .

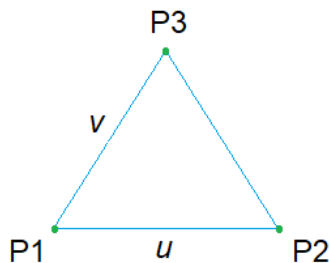


Figura 32. Vectores u y v de un triángulo

3.3.5 Otras operaciones de morfología matemática

3.3.5.1 Erosión

La erosión es una operación que pertenece al área de la morfología matemática en la que se evalúa un patrón en cada uno de los vóxeles de la imagen original, que se mantienen si el patrón está completamente contenido. Esta operación reduce el tamaño de las figuras.

3.3.5.2 Dilatación

La operación de dilatación agranda las regiones incluyendo en los segmentos los píxeles (o vóxeles) que se encuentren a una distancia a la redonda. La dilatación se utiliza en las imágenes para rellenar los agujeros.

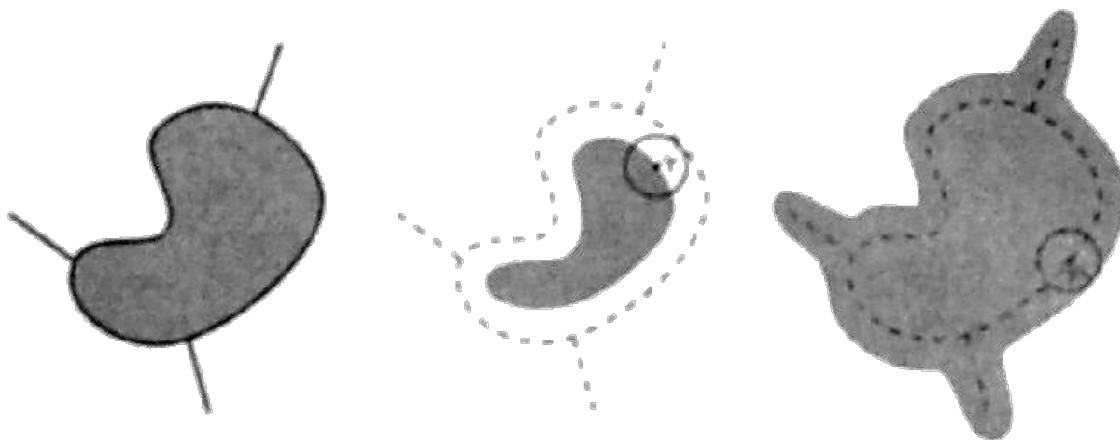


Figura 33. Erosión y dilatación con un elemento estructurante circular (UNS, 2019)

3.4 Los dispositivos 3D

De las cámaras 3D diseñadas para ser utilizadas en equipos de consumo, la más conocida es la cámara **Kinect** de Microsoft. Sin embargo, después de evaluar su rendimiento se determinó buscar otra alternativa pues no puede utilizarse en exteriores y la nube de puntos que arroja no está ordenada lo que aumenta la complejidad de la triangulación.

La segunda y tercera cámaras evaluadas son de la familia **RealSense**© de la compañía **Intel**©. La versión de escritorio utiliza tecnología **ToF** y la versión para dispositivos móviles **Visión estéreo**, seleccionándose la última por la posibilidad de utilizarla en campo.

A continuación, se presentan las características de las cámaras y sus ventajas y desventajas.

3.4.1 El sensor 3D Kinect de Microsoft

El dispositivo Kinect© de la compañía Microsoft© permite capturar un mapa 3D de una escena y con ello reconocer los movimientos de los usuarios que utilizan su consola de videojuegos Xbox© (Microsoft, 2018, 1).



Figura 34. Imagen obtenida con el sensor Kinect. La escena muestra una planta de chile *Capsicum annuum* en una maceta, iluminada por miles de puntos de luz infrarroja.

La **figura 35** muestra una imagen obtenida con la cámara infrarroja del sensor Kinect. En la escena, puntos de luz infrarroja iluminan una maceta sobre una mesa con una puerta abierta al fondo.

La **figura 36** es el mapa de profundidades. Las zonas en color verde están demasiado cerca para ser captadas por el sensor y las azules demasiado alejadas. Los píxeles en color rojo están ocluidos, y los grises muestran distintos tonos de acuerdo con su lejanía.

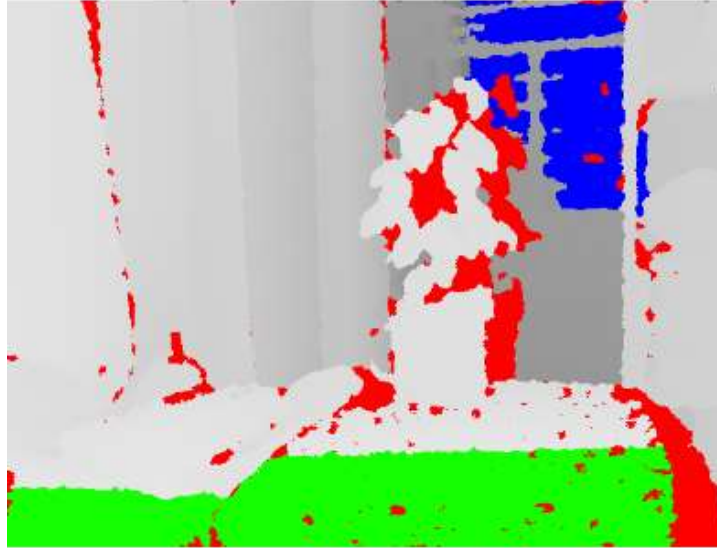


Figura 35. Mapa de profundidades

EL sensor Kinect no puede utilizarse en exteriores y por su diseño presenta varios inconvenientes al momento de utilizarlo para realizar mediciones. En el artículo *Rango de sensibilidad Kinect: Luz estructurada vs Kinect Tiempo de vuelo* se compilan algunos (Sarbolandi *et al*, 2015):

Errores que pueden presentarse al utilizarse un dispositivo Kinect para medir:

1. Iluminación ambiente: aun cuando Kinect sólo toma en cuenta cierto ancho de banda las fuentes de luz lo afectan.
2. Interferencia multi - aparato: Cuando se usan dos o más aparatos, la señal de uno interfiera con la señal de los otros
3. Desvío por temperatura: cuando la temperatura del aparato aumenta las distancias obtenidas cambian
4. Error sistemático en la distancia: cuando no hay calibración previa
5. Falta de homogeneidad de profundidad: la geometría de los objetos puede presentar zonas ocluidas que presentan pixeles inválidos
6. Efecto multi - trayectoria: la luz no sólo viaja directamente al objeto y de regreso, puede dispersarse en superficies altamente reflejantes
7. Error de distancia asociado a la intensidad: cuando la distancia es menor el láser tiene más intensidad y su medición es más exacta. Cuanto mayor sea la distancia menor será la intensidad del láser y menos exacta la medición.
8. Medios difusores y semitransparentes
9. Escenario dinámico: los objetos en movimiento generan corrimiento en las mediciones.

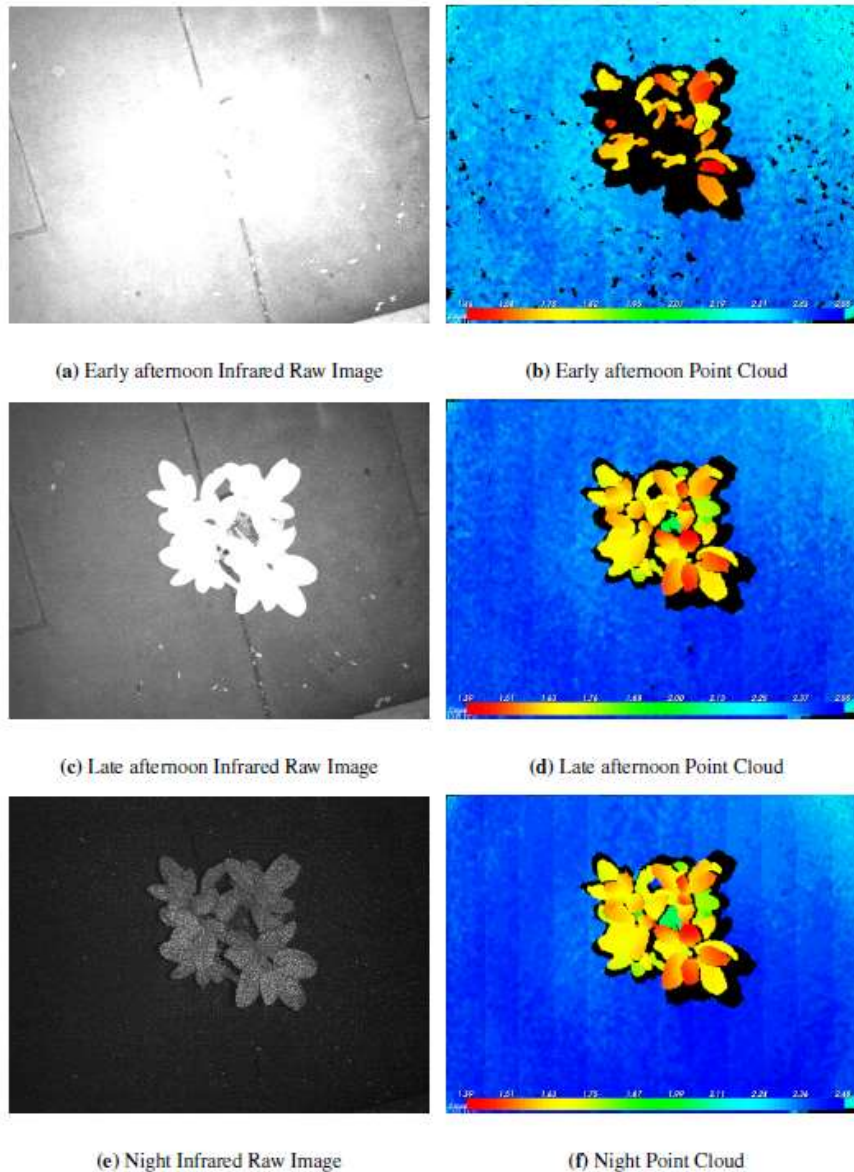


Figura 36. La luz ambiente afecta el rendimiento del dispositivo Kinect (Sarbolandi et al, 2015)

Algunos de los problemas operacionales del sensor Kinect se presentan en equipos de otras empresas. Si bien, los desarrollos previos realizados con el dispositivo Kinect muestran el camino a seguir, las desventajas del aparato plantearon la necesidad de considerar otras cámaras que permitan operar en campo.

3.4.2 Cámara Intel RealSense R200

RealSense© es la línea de cámaras 3D de la compañía Intel©. Algunas utilizan tecnología ToF y otras, visión estereoscópica. Funcionan para dispositivos con sistema operativo Windows y procesadores Intel.

La cámara Intel Real Sense R200 toma fotos en tercera dimensión de escenas entre los 0.5 y 3.5m de distancia. Captura imágenes de 480 x 360 píxeles a 60 cuadros por segundo. Tiene un campo de visión de 59° horizontales y 46° verticales. La cámara se conecta al puerto USB de dispositivos portátiles (Intel Corporation, 2015). Puede utilizarse para realizar mediciones, escanear en 3D y tomar fotografías multi - foco. (**figura 37**)



Figura 37. Cámara 3D Intel RealSense R200

El dispositivo proyecta un haz de luz infrarroja en el rango de los 850 nanómetros que es luego captado por dos cámaras infrarrojas para formar una imagen estereoscópica. En el mapa de profundidades se expresa la distancia de cada píxel a la cámara.

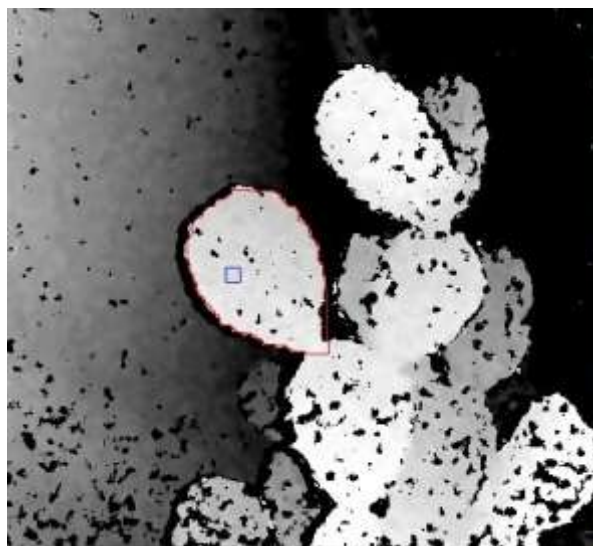


Figura 38. Mapa de profundidades de una planta de nopal en escala de grises obtenida con la cámara Intel RealSense R200

La cámara de profundidad funciona mejor en entornos que tienen detalles y textura. Los datos de profundidad serán menos precisos para superficies negras o altamente reflectantes.

Con base en los ejercicios realizados (véase capítulo 3) se encontró que al fotografiar plantas la cámara funciona mejor si recibe menos de 500 **lux**, es decir, que puede usarse al amanecer, al atardecer o dentro de invernaderos. En otras circunstancias el fotógrafo debe controlar las condiciones de iluminación.

No significa que la cámara no funcione en exteriores bajo la luz solar, sino que su funcionamiento depende de la textura de la superficie a fotografiar. Las hojas de las

plantas son lisas y absorben la luz infrarroja. Algunas plantas desérticas reflejan parte de la luz solar para evitar deshidratarse.

La cámara permite, mediante software, configurar los valores de exposición y ganancia para utilizar la cámara en exteriores. El tiempo de **exposición** que puede estar entre 1 y 33 milisegundos debe disminuirse. Así mismo, la **ganancia**, que es la medida de la apertura del obturador, debe ser baja.

La cámara incluye su propio procesador, que convierte el mapa de profundidades en un mapa de vértices. En el mapa de vértices, cada vóxel contiene la coordenada espacial x , y , z con respecto a un plano cartesiano en el que la coordenada central $(0,0,0)$ está en el centro de la imagen

El flujo de video en profundidad se genera con la tecnología de visión estéreo asistida por un proyector láser infrarrojo y dos sensores de imagen infrarroja.

La cámara de color en el módulo de cámara R200 proporciona información de textura. Los usos de la información de textura incluyen la superposición en una imagen de profundidad para crear una nube de puntos de color y la superposición para la reconstrucción del modelo 3D.

El módulo R200 usa visión estéreo para calcular la profundidad. La implementación de la visión estéreo consiste en una cámara de luz infrarroja izquierda, una cámara de infrarrojos derecha y un proyector láser infrarrojo. Los datos de la cámara izquierda y derecha se envían al módulo ASIC R200. El chip ASIC calcula los valores de profundidad para cada píxel en la imagen. El proyector de infrarrojos se utiliza para mejorar la capacidad del sistema para calcular la profundidad en escenas con poca textura. Tradicionalmente, las escenas con poca textura, como las paredes, representaban un desafío para los sistemas de visión estéreo al calcular la profundidad.

(Intel corporation, 2015).

3.4.3 Cámara Intel RealSense SR300

El modelo **Intel RealSense SR300 3D** es la versión para equipos de escritorio. Esta versión toma videos de 720p a 60 cuadros por segundo o de 1080p a 30 cuadros por segundo. Opera en un rango de profundidades de 20 a 150 cm y se conecta al puerto USB 3.0. (**figura 39**)



Figura 39. Intel Realsense SR300 3d camera (Intel corporation, 2015, 3)

La cámara Intel® RealSense™ SR300 implementa un sistema de luz estructurada de corto alcance.

Para generar un cuadro de profundidad, el proyector de infrarrojos ilumina la escena con un patrón de barras verticales de luz infrarroja codificado en una frecuencia espacial predefinida. El patrón se deforma sobre las figuras, se refleja y es capturado por la cámara IR. Los valores de píxeles de la cámara IR son procesados por el chip ASIC de imágenes para generar un cuadro de profundidad. Los cuadros de profundidad subsiguientes crean un flujo de video que se transmite al sistema cliente. (Intel corporation, 2015).

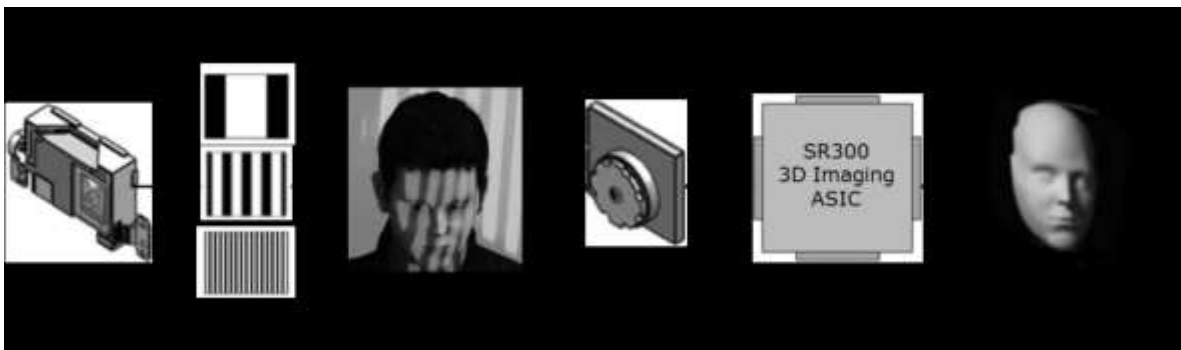


Figura 40. Luz estructurada en cámara 3D Realsense SR300 (Intel corporation, 2015)

MEDICIÓN DE CLADODIOS DE *OPUNTIA SPP.*

Las espinas presentes en las plantas del género *Opuntia* dificultan su manejo y estudio. De ahí la importancia de diseñar un método indirecto de medición. La toma de fotografías 3D in situ es un método de bajo costo que permite medir la planta espinosa sin tocarla, sin dañarla, y el resultado es visualizado de inmediato. Bajo el supuesto de que los cladodios son planos, se propuso utilizar la regresión lineal múltiple al aplicar las diversas técnicas de tratamiento digital de imágenes.

Cuadro 2: Clasificación científica de las plantas del género *Opuntia*

Reino	Plantas
Fílum	Espermatofitas
Clase	Dicotiledóneas
Orden	Opuntiales
Familia	<i>Cactaceae</i>
Tribu	<i>Opuntiae</i>
Subfamilia	<i>Opuntioidea</i>
Género	<i>Opuntia spp.</i>

Las plantas del género *Opuntia* pertenecen a la familia de las **cactáceas** (Cuadro 2) y son nativas del continente americano. Son plantas perennes, por lo general arbustiformes o rastreras, y espinosas con tallos carnosos (suculentas). Su tronco es definido o ramificado desde la base con ramas erguidas, extendidas o postradas formadas por artículos o pencas, llamados “**cladodios**”. Estos pueden ser globosos, cilíndricos o aplanados, muy carnosos o leñosos, de color verde, con funciones fotosintéticas. En ellos se observan hojas transformadas en espinas y flores de colores vivos y brillantes que dan origen a frutos conocidos como “**tunas**”, que son bayas ovoides cilíndricas de diversos colores. Los nopales poseen características morfológicas y fisiológicas que les permiten almacenar y conservar el agua en sus tejidos, con lo cual sobreviven a la escasez de agua y a las marcadas variaciones de temperatura características de las zonas áridas y semiáridas (Cervantes-Ramírez, 2008) (ver **figura 41**).



Figura 41. *Opuntia Ficus Indica* (Pixabay)

México es el principal productor de nopales frescos para consumo humano en el mundo, con más de 12 000 hectáreas cultivadas y una cosecha que ronda las 825 000 toneladas anuales. **China** produce nopal a gran escala para el área cosmética. **Brasil** es el principal productor y exportador de la cactácea como forraje para ganado. La nación sudamericana tiene 600 mil hectáreas para cultivar nopal de forraje, que junto con otros elementos soportan una producción lechera diaria de un millón 200 mil litros (La razón de México, 2016).

La FAO destaca que el cultivo del nopal será clave en el futuro, ante un aumento de las sequías y la desertificación, especialmente en África y Asia. Es por ello que países como Marruecos, Etiopía, Sudáfrica, Kenia, India y Pakistán están incrementando su producción y uso (FAO, 2017).



Figura 42. Camello comiendo el cactus espinoso en Etiopía (FAO, 2017)

Figura 43. Corte manual de cladodios del cactus sin espinas en Túnez. Ali Nefzaoui (FAO, 2017)

La presencia de espinas y la anchura de los segmentos en los especímenes de *Opuntia* imposibilitan el uso de otras técnicas y aparatos normalmente utilizados para medir el área foliar. Además, el nopal y su fruto tienen el mismo color, textura y silueta, lo único que los diferencia es que el segmento es plano y la fruta es esférica. Esta situación certifica la necesidad y pertinencia de la tercera dimensión para la segmentación de las pencas y frutos.

En las cactáceas, el agua disponible durante el desarrollo de la planta genera diferencias marcadas de grosor y tamaño (Scalisi et al, 2016). La medición de los segmentos ayudaría a diagnosticar la salud de las plantas, determinar la madurez de un cultivo para decidir su fecha de cosecha, personalizar las dosis de fertilizantes e insecticidas, describir e inventariar especímenes silvestres y/o calificar la reacción de las plantas a diferentes configuraciones de suelo, agua y temperatura.

Mendes (2015) reportó que no había antecedentes en la literatura de la medición del área de cladodios por métodos indirectos e intentó aplicar cuatro estrategias clásicas:

estimación matemática, planimetría, medición de discos foliares e intercepción de radiación solar. Encontró que “los métodos de intercepción de la luz y sus derivaciones, medición de discos foliares y planimetría no son recomendables, debido a su baja precisión y exactitud” y que si bien “el método de modelado matemático presentó coeficientes de precisión y exactitud por encima de 0.90, su aplicabilidad depende del uso de parámetros estimados y ecuaciones de valores ajustados.”

La posibilidad de medir los segmentos de la planta espinosa de manera no destructiva y sin manipularla, permitiría monitorear el desarrollo de un cladodio en particular durante un período de tiempo.

El desarrollo de equipo especializado reduciría los errores cometidos en una medición manual, mientras que el resultado se visualizaría inmediatamente sin interferir con el desarrollo de las plantas.

El método propuesto utiliza tecnología de última generación que por primera vez se utiliza para medir cladodios espinosos in situ. El resultado de la medición se muestra de inmediato, el método es “no destructivo”, y el tiempo de implementación es mucho menor que el de otras técnicas reportadas en la literatura.

4.1 Materiales y metodología

La segmentación de los cladodios de nopal inicia con la captura de la imagen con la cámara 3D Intel RealSense R200. En la imagen, además de tener diferentes profundidades, los segmentos tienen diferentes inclinaciones que ayudarían a segmentarlos. La imagen se almacena como un mapa de vértices (**ver figura 42**).

La plataforma planteada en este trabajo utiliza una cámara de última generación que, si bien funciona con luz infrarroja, permite por primera vez capturas **en campo** ya que los pulsos de luz con los que trabaja se generan con un patrón de anchos de banda característicos y aleatorios susceptibles a ser separados del ruido generado por la luz solar. Además de que puede conectarse al dispositivo móvil del usuario lo que ahorra costos.

Bajo el supuesto de que los cladodios son planos, se propuso el uso de la **regresión lineal múltiple** para caracterizarlos. De la regresión lineal múltiple se obtiene un plano que se utiliza para evaluar los vóxeles al momento de aplicar el algoritmo de crecimiento de región.

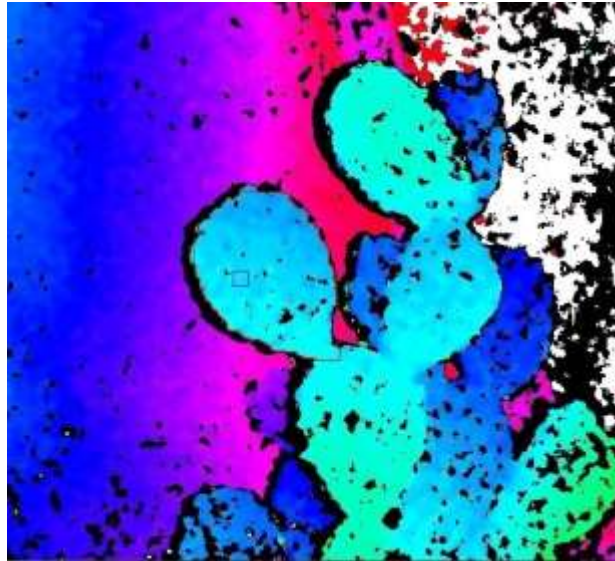


Figura 44. Mapa de vértices de una planta de nopal obtenida con la cámara Intel RealSense R200

El primer paso para refinar el resultado de la segmentación es rellenar la pieza tomando en cuenta la información de la ecuación del plano. Posteriormente, la silueta del cladodio es separada del tronco a la que está unida y de sus vástagos con la aplicación sucesiva de etapas de erosión y dilatación.

Finalmente, se triangula la nube de puntos y se suman las áreas de los triángulos para obtener el resultado final.



Figura 45. Pasos del método propuesto

Para la investigación se trabajó con un equipo con procesador Intel© Core I7 con 8 GB de RAM a 2.5 GHz. El algoritmo¹⁰ fue codificado en lenguaje Java© para el sistema operativo Windows 10©.

4.2 Regresión Lineal Múltiple

La regresión es el método mediante el cual se descubren las relaciones entre una o más variables de respuesta (también llamadas variables **dependientes**) y los predictores (también llamadas variables **independientes**).

La regresión lineal múltiple supone que la variable de respuesta es una función lineal de los parámetros del modelo y que hay más de una variable independiente en el modelo. La forma general del modelo de regresión lineal múltiple es la siguiente:

Ecuación 7: Forma normal de la regresión lineal múltiple

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p + \varepsilon \quad (7)$$

Donde "y" es la variable independiente, $\beta_0, \beta_1, \dots, \beta_p$ son los coeficientes de la regresión y x_1, x_2, \dots, x_p son las variables independientes del modelo; ε representa el error (Hernández et al, 2004).

La operación de regresión lineal múltiple se aplica a un conjunto semilla de 12 por 12 vóxeles al centro del cladodio a segmentar. Como resultado se obtiene la ecuación de un plano (**ver figura 43**).



Figura 46. Representación del plano obtenido con la regresión lineal múltiple

¹⁰ El código fuente puede descargarse gratuitamente desde www.serconmur.com/area-foliar-opuntia.html

4.3 Crecimiento de región con semilla

Bajo la suposición de que los cladodios son mayormente planos, los píxeles agregados a la región que crece son los que se ubican sobre el plano descrito por la operación de regresión. Dependiendo del espesor y curvatura de la planta y de la distancia a la cámara se considera una tolerancia de entre 1.5 y 4 centímetros (**figura 44**).

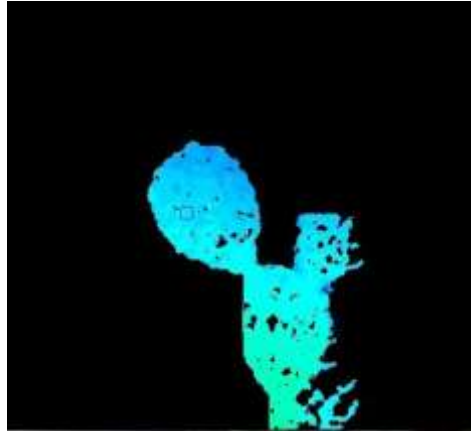


Figura 47. Región segmentada, vóxeles vecinos a la semilla y que están a menos de 4 cm. de la ecuación del plano.

4.4 Rellenado

Los orificios en la pieza segmentada deberán rellenarse con **puntos artificiales** antes de continuar con las otras etapas. Puede haber vóxeles sin información en las espinas o en las imperfecciones.

El relleno generalmente se realiza mediante una operación de dilatación, pero en este caso los vértices sin información de profundidad no tienen la coordenada X ni Y.

Para rellenar cualquier punto sin información que no pertenezca al fondo se utilizan los valores de X e Y de los vóxeles vecinos. El punto artificial se evalúa en la ecuación del plano para encontrar el valor de z (**figura 45**).



Figura 48. Región con sus orificios rellenados con valores generados de la ecuación del plano

4.5 Erosión

Generalmente, parte del tronco que sostiene al cladodio o sus vástagos serán incluidos en la segmentación porque podrían tener la misma inclinación. Para separarlos se realiza una operación de erosión con un elemento estructurante circular.

Cuando se aplica la operación de erosión con un elemento estructurante circular (de entre 5 y 10 cm de diámetro, correspondientes al ancho de la unión a romper) la región se separa en varias (ver figura 46).



Figura 49. Región erosionada, la erosión separa un cladodio de los demás

4.6 Dilatación

Para recuperar las dimensiones de la región de interés, se realiza una operación de dilatación con un elemento estructurante circular de radio similar al utilizado en la erosión. En este caso, los vóxeles con información de profundidad de la imagen original que se encuentran dentro del radio del elemento estructurante se agregan a la región.

El proceso de erosión y dilatación separa un cladodio en particular de otras ramas a las que pudiera estar unido y es inocuo para las piezas que ya se hubieran segmentado correctamente (ver figura 47).



Figura 50. Dilatación del cladodio segmentado para recobrar su dimensión original

Pasos de la Segmentación



Figura 51. Pasos del algoritmo de segmentación

4.7 Triangulación

Finalmente, para obtener el área de la pieza, se representan los vóxeles como una superficie formada por triángulos en tercera dimensión y luego se suman sus áreas.

El mapa de vectores generado por la cámara posee una textura de **crestas** y **valles** debido a la presencia de espinas y al proceso de relleno. El resultado de la triangulación de una superficie así es mayor al valor esperado para una figura idealmente lisa. Si se considera que el área que queremos medir es la de una figura plana, entonces los valores de la distancia "Z" pueden generarse a partir de la evaluación de "X" e "Y" en la ecuación del plano obtenida en la regresión.

Otra manera de no tomar en cuenta el ruido generado por las espinas es triangular solo los bordes de la figura. Cuando los triángulos se forman horizontalmente, se considera la curvatura vertical y cuando los triángulos se forman verticalmente, la horizontal, así que puede tomarse el promedio del resultado de ambos ejercicios.

Finalmente puede considerarse que, en la mayoría de los casos, los cladodios son planos y que en un cultivo interesaría medir la parte aprovechable de la planta, aun en las piezas curvadas.

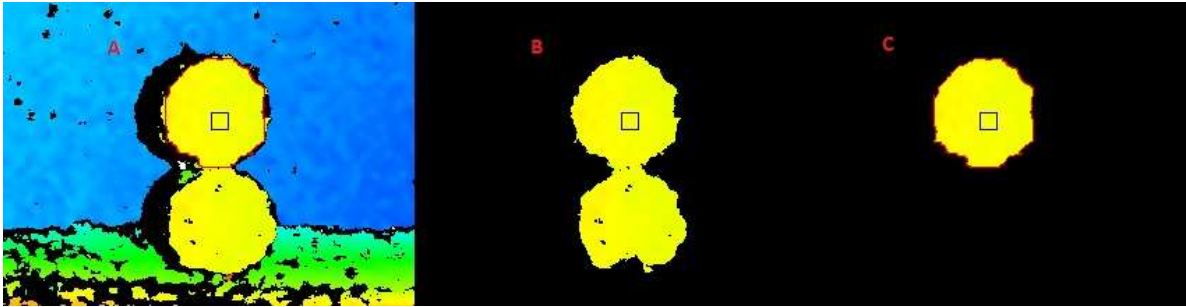


Figura 52. Ejercicio con dos círculos de nopal unidos entre sí, A) mapa de vértices, B) segmentación inicial, C) segmentación final.

4.8 La aplicación

El algoritmo fue implementado en lenguaje Java. Los programas codificados en lenguaje java son interpretados por una máquina virtual. Internamente, la máquina virtual de java utiliza dos grandes estructuras de memoria: **heap** (montón) y **stack** (pila).

En la memoria *stack*, java almacena las variables primitivas y los apuntadores a los objetos, en la memoria *heap* almacena los objetos. El *garbage collector*¹¹ recorre permanentemente la memoria *heap* y cuando encuentra un objeto que no es referenciado en *stack*, lo elimina para que el espacio pueda reutilizarse.

La máquina virtual de java le asigna una pila a cada **hilo**. Por defecto, la pila inicia con un tamaño de 64MB y crece hasta un máximo de 128MB. El tamaño de la pila es configurable, pero en los sistemas operativos de 32 bits, no puede pasar de 2GB.

La cámara 3D funciona en sistemas de 32 y 64 bits, así que el software debe utilizar menos de 2GB de memoria en *stack* o restringir su uso a solo sistemas de 64 bits.

Para esta investigación se tomaron diversas medidas para ahorrar memoria y recursos computacionales que permitan ejecutar la aplicación en el menor tiempo posible y brindarle al usuario una mejor experiencia.

La cámara seleccionada arroja la información en forma de matriz, por lo que para la triangulación se sabe de antemano cuales son los vecinos de cada vóxel disminuyendo su complejidad computacional.

Con la triangulación resuelta, el siguiente paso más demandante es la segmentación, específicamente el algoritmo de crecimiento de región con semilla, ya que es recursivo. Para acelerarlo, se determinó dividir la imagen en 4 **cuadrantes** a partir de la semilla. En cada cuadrante el trabajo se divide entre 3 hilos (12 hilos en total) que operan en direcciones contrarias.

¹¹ El *recolector de basura* es una aplicación de la máquina virtual que se encarga de eliminar automáticamente los objetos en desuso

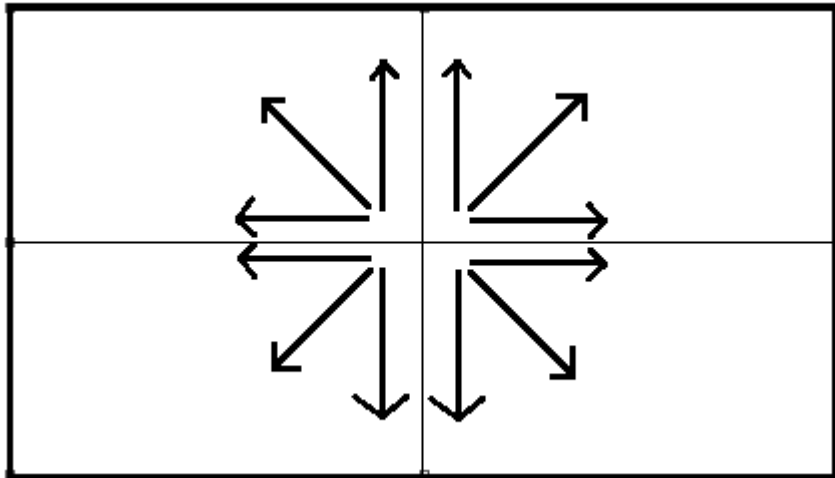


Figura 53. La imagen se divide en 4 cuadrantes con 3 hilos cada uno.

Como ya se había mencionado antes, cada hilo tiene su propia pila y eso reparte el uso de la memoria y de los núcleos del procesador.

Las imágenes se almacenan en formato XML y PCD.

RESULTADOS Y DISCUSIÓN

El rendimiento del algoritmo se comparó contra otras 5 estrategias clásicas. Se calculó el área de 13 cladodios de nopal por el método propuesto, por **planimetría**, **volumetría**, **escaneo 2D** y 2 estrategias de **estimación matemática**. Se consideró a la medición mediante el escaneo 2D y uso de software especializado como la medida de referencia contra la cual se evaluaron los otros métodos.

Primero se fotografiaron las plantas con la cámara 3D, el usuario selecciona los mejores parámetros de tolerancia, y radios de erosión – dilatación. La tolerancia (distancia permitida de los vóxeles al plano) aumenta cuando el cladodio es muy grueso y disminuye cuando la pieza está pegada a otra. El radio de erosión – dilatación depende del ancho de las uniones entre los cladodios.

Posteriormente se cortaron los cladodios y se digitalizaron con un **escáner** con el objetivo de obtener su área mediante el uso de un software especializado¹² (Contreras *et al.* 2015). También se marcó su silueta en papel milimétrico para estimar su área mediante planimetría.

Se calculó cuál sería el área del cladodio si se considerara que tienen forma de elipse, para ello se midió el largo y ancho con cinta métrica y se sustituyeron los valores en la fórmula del área de la elipse. Del mismo modo se calculó cual sería el área si se considerara que los cladodios tienen la forma de un semicírculo unido a un triángulo.

Finalmente se calculó el volumen de un cilindro que representara la pieza (y el área de su base). Fue necesario entonces medir el ancho de una muestra del centro del cladodio con un vernier, su volumen sumergiéndola en agua y su peso (**ver Cuadro 3**).

¹² Disponible gratuitamente en <http://www.serconmur.com/area-foliar-escaner.html>

Cuadro 3. Áreas en cm² obtenidas mediante Escaneo 2D (medición de referencia), el método propuesto y otros 4 métodos.

muestra	Escaneo 2D	Cámara 3D	volumetría	Planimetría	Aproximación elipse	Aproximación triángulo semicírculo
1	172.70	172.00	154.92	180.24	187.86	107.99
2	50.70	50.80	52.31	58.37	64.08	36.93
3	259.70	267.70	271.28	304.30	306.42	169.95
4	271.60	295.60	272.89	313.77	318.55	172.15
5	491.84	530.10	426.66	629.48	546.16	291.92
6	77.37	86.70	81.37	92.69	96.89	55.96
7	97.00	113.80	110.01	113.24	108.85	57.47
8	239.41	253.40	234.44	278.23	274.88	153.96
9	296.57	319.60	287.51	348.32	345.57	192.53
10	279.12	311.60	261.56	331.58	338.66	190.15
11	259.26	325.20	273.73	301.58	311.01	168.78
12	278.30	340.60	288.32	311.27	313.64	171.16
13	712.30	767.70	679.00	785.20	790.89	428.14

De las estrategias evaluadas, la volumetría y la toma de fotografías 3D tuvieron el menor error promedio (**Cuadro 3**), con la salvedad de que la volumetría fue la estrategia más difícil de implementar y requirió varias horas para aplicarse. Para el escaneo 2D fue necesario cortar los cladodios en partes para poder meterlos en el dispositivo. La toma de fotografías 3D se implementó al menos 20 veces más rápido que las otras técnicas.

Cuadro 4. Margen de error de las áreas obtenidas en cm² de las 5 técnicas de medición respecto del escaneo 2D en términos de la distancia Euclidiana.

Método	promedio de error	mediana de error	error máximo	error mínimo	varianza	desviación estándar
volumetría	556.60	134.25	4247.58	1.68	1412787.01	1188.60
cámara 3D	1196.40	530.38	4348.08	0.01	2398380.42	1548.67
triángulo semicírculo	14673.29	8053.81	80742.39	189.41	494375375.40	22234.55
elipse	1967.83	2183.04	6177.32	140.56	2922166.40	1709.43
planimetría	2958.14	1778.30	18944.76	56.85	25170478.55	5017.01

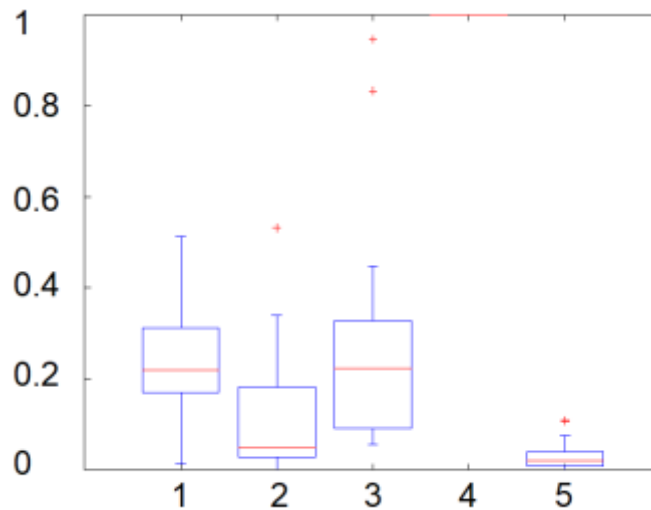


Figura 54. Gráfica de cajas y bigotes de las mediciones realizadas mediante 1) planimetría, 2) cámara 3D, 3) estimación matemática de elipse, 4) estimación matemática de triángulo y semicírculo, 5) volumetría

Se detectó que los resultados de la medición 3D suelen estar por encima del resultado esperado, así que se calculó un factor de conversión que al multiplicarlo por los resultados compensase el error, para ello se realizó un remuestreo.

La técnica de remuestreo mas comúnmente utilizada es el remuestreo con remplazo, o Bootstrap¹³ que consiste en tomar de forma repetida muestras de la población, seleccionadas al azar (Drennan, 2019).

Se utilizó una estrategia *bootstrap* de 100,000 remuestreos con 2 colas y 95 % de confianza, donde se determina la relación existente entre el área obtenida y el área esperada. Se obtuvo un factor de conversión promedio de 0.90918, un factor máximo de 0.93981 y mínimo de 0.87708, siendo la varianza máxima 0.00594 y la mínima 0.00141.

Los resultados iniciales se multiplicaron por el factor de conversión promedio y se recalculó el error promedio. De los nuevos datos se desprende que la multiplicación del resultado de la medición 3D por su razón de conversión (0.90918) mejora los resultados al punto de estar al nivel de las mejores estrategias destructivas (**Cuadro 5**).

Cuadro 5. Margen de error de las áreas obtenidas en cm² por toma de fotografías 3D respecto del escaneo 2D en términos de la distancia Euclidiana.

método	promedio de error	mediana de error	error máximo	error mínimo	varianza	desviación estándar
cámara 3D	589.28	70.8	3407.37	0.3341	1052472.87	1025.90

A través de una prueba no paramétrica de medianas se determinó que la toma de fotografías 3D y la volumetría son estadísticamente similares pues ambas muestran una **exactitud del 95%**. Con base en ello se puede decir que el procedimiento propuesto en este trabajo es tan bueno como las mejores técnicas destructivas, con las ventajas señaladas anteriormente.

En una cámara 2D conforme aumenta la inclinación la figura es mostrada como más delgada. Al utilizar una cámara 3D también hay menos información disponible cuando la inclinación aumenta, pero el uso del plano determinado por la regresión lineal múltiple disminuyó la desviación estándar de los resultados.

El promedio del resultado de ambas estrategias representaría el área de una superficie curva y minimizaría el efecto de la textura espinosa.

¹³ En 1979, Bradley Efron publica el análisis formal de Bootstrap, término que procede de la expresión inglesa *to pull oneself up by one's Bootstrap* (que podría traducirse como: *levantarse mediante el propio esfuerzo*), tomada de *las aventuras del barón Munchausen*, personaje del siglo XVIII creado por el escritor Rudolph Erich Raspe, en la cual el barón había caído al fondo de un lago profundo y, cuando creía que todo estaba perdido, tuvo la idea de ir subiendo tirando hacia arriba de los cordones (*bootstrap*) de sus propias botas (Miranda Moles, 2003).

Los mejores resultados se obtienen cuando las muestras están lo más paralelas posible a la cámara. La forma plana y ovalada de los segmentos de *Opuntia* favorecen su medición.

Las imágenes se tomaron con la **resolución** más baja que ofrece la cámara para favorecer la experiencia del usuario en términos de rapidez, en este caso de 640 * 480 pixeles a un metro de distancia evitando la luz solar directa (menos de 300 lux).

El éxito de las etapas de erosión y dilatación depende de que la semilla se encuentre lo más cerca posible al centro del cladodio, de otro modo existe el riesgo de que la semilla se elimine al erosionar la pieza; Para evitarlo, la interface gráfica permite reposicionar la semilla y configurar el tamaño del elemento estructurante cuyo radio no puede ser mayor a la distancia de la semilla al borde de la pieza.

El diámetro de la erosión y de la dilatación dependen del ancho de las uniones a separar. Es decir, que dependen de la especie de *Opuntia* y de su disposición en la planta (las uniones son más anchas en los segmentos más bajos). Además, es inversamente proporcional de su distancia a la cámara ya que un vóxel representa un área más grande cuanto más lejos esté de la cámara.

Otra variable susceptible de ser configurable o dinámica es la tolerancia del algoritmo de crecimiento de región. Hay que recordar que la regresión lineal múltiple determina un plano respecto del cual se toman los vóxeles más cercanos (cuya distancia al plano sea menor a un valor de tolerancia). Esta medida que ronda los 2 cm. depende del ancho y curvatura de los nopales (que a su vez depende de la especie). La tolerancia del algoritmo de crecimiento de región depende del ancho y la curvatura de los cladodios, de la densidad del cultivo y es inversamente proporcional de su distancia a la cámara.

Finalmente, se menciona la posibilidad de configurar valores del hardware. El tiempo de exposición y la ganancia (apertura del diafragma) de la cámara tienen efectos significativos sobre la calidad de las imágenes ante variaciones de luz.

Los mejores resultados se obtienen cuando las muestras están lo más paralelas posible a la cámara. La forma plana y ovalada de los segmentos de *Opuntia* favorecen su medición.

El método propuesto soluciona algunos de los inconvenientes que presentan las metodologías clásicas al aplicarse a los cladodios espinosos, y sin dañar la planta genera un resultado confiable que es presentado inmediatamente (**Cuadro 6**).

El algoritmo codificado en el lenguaje Java tarda menos de 2 segundos en ejecutarse en un equipo con procesador *Intel core I7* con 8 GB de RAM a 2.5 GHz.

Cuadro 6. Comparación entre el método propuesto y otras técnicas de medición del área foliar al aplicarse a cladodios espinosos de *Opuntia Spp.*

Método	Problemática en los otros métodos	Método propuesto: Fotografía 3D
Métodos destructivos	Cortar la muestra daña la planta y no permite medir el desarrollo de una misma pieza a lo largo del tiempo.	Este algoritmo no requiere dañar la planta
Métodos destructivos, Integrador portátil, Estimación matemática, Fotografía 2D	Manipular plantas espinosas puede causarle heridas al usuario	No requiere manipular la planta
Integrador portátil, Escaneo 2D	El cladodio de nopal es demasiado grande y grueso para entrar en el dispositivo	Puede medir piezas de gran tamaño y no importa el grosor
Integrador	El dispositivo tiene un costo elevado	Este dispositivo tiene bajo costo, La cámara se conecta al dispositivo móvil del usuario
Medición con cámara 3D tradicional (Kinect)	Requiere conexión eléctrica, por lo que disminuye su usabilidad en campo	Primera cámara 3D diseñada para usarse en campo
Medición con cámara 3D tradicional (Kinect)	El sol afecta su funcionamiento, No puede usarse en exteriores	Primera cámara 3D infrarroja diseñada para usarse en exteriores y filtrar el ruido de la luz solar pero solo en horarios de baja radiación solar
Fotografía 2D	Necesita un fondo con un patrón para convertir de pixeles a centímetros cuadrados	No necesita un fondo con un patrón determinado.

Fotografía 2D	Necesita tomar la fotografía a una distancia conocida para convertir de pixeles a centímetros cuadrados	La distancia la arroja la cámara.
Fotografía 2D	La inclinación de la muestra respecto de la cámara afecta la medición	La regresión lineal múltiple compensa la inclinación de la muestra respecto de la cámara
Fotografía 2D	La curvatura de la muestra afecta la medición	La triangulación puede tomar en cuenta la curvatura de la muestra
Estimación matemática	La fórmula que determina el área de los cladodios solo funciona para una especie en particular.	Los parámetros de operación se ajustan de acuerdo a la especie y variables ambientales.
Estimación matemática	No considera las imperfecciones de los especímenes.	Las imágenes 3D pueden detectar los agujeros en la muestra producto de una enfermedad o lesión.

CONCLUSIONES

La segmentación de las partes de las plantas utilizando imágenes 3D resuelve diversos inconvenientes de las metodologías clásicas de medición del área foliar al momento de medir cladodios espinosos de *Opuntia*. Algunas metodologías y aparatos ni siquiera pueden aplicarse en estas plantas.

El método desarrollado:

- No es destructivo,
- No requiere manipular la planta,
- Puede medir piezas de gran tamaño o que se encuentre a gran altura

Respecto de las técnicas que utilizan fotografías en 2D:

- Permite medir una pieza que está rodeada de otras del mismo color y textura
- No requiere que la cámara esté paralela a la muestra
- No requiere que la cámara este a una distancia determinada (aunque si presenta un rango de operación)
- Compensa el efecto de la inclinación y curvatura de las muestras

El algoritmo presentó una exactitud del 95%, al nivel de las mejores estrategias destructivas, pero sin dañar la planta y es al menos 20 veces más rápido en términos del tiempo necesario para implementarlo. El uso de una cámara 3D permite que el resultado sea mostrado inmediatamente.

Aunque la foto se tome desde diferentes perspectivas los resultados son consistentes, gracias al uso de la regresión lineal múltiple en las distintas etapas del algoritmo.

De las cámaras evaluadas se seleccionó el dispositivo RealSense R200 de Intel porque está diseñada para ser utilizada en exteriores, obtener la electricidad del puerto USB y arrojar la nube de puntos ordenada en una matriz (lo que disminuye la complejidad de la triangulación).

La cámara 3D seleccionada se ve afectada por la luz solar (como cualquier dispositivo que usa luz infrarroja). La cámara funciona mejor cuando la escena recibe menos de 300 lux, es decir, que puede usarse al amanecer, al atardecer o dentro de invernaderos. En otras circunstancias el fotógrafo debe controlar las condiciones de iluminación.

El algoritmo entero puede volver a aplicarse para refinar el resultado, en la primera ejecución el plano se calcula de una semilla, en las ejecuciones sucesivas el plano se calcularía a partir de la figura segmentada.

La forma y tamaño de los cladodios de *Opuntia* favorece su análisis en tercera dimensión, pero el que sean lisos y reflectivos aumenta el efecto adverso de la luz solar directa.

La medición de cladodios de *opuntia* mediante la toma de fotografías 3D no interfiere con el desarrollo de la planta y permite monitorear el desarrollo de un mismo espécimen a lo largo del tiempo.

La aplicación desarrollada ayuda a agrónomos e investigadores a:

- Determinar el estado de salud de las plantas.
- Medir el daño producido por una plaga.
- Determinar el estado de madurez de un cultivo para decidir su fecha de cosecha.
- Predecir el volumen de la cosecha.
- Personalizar por planta las dosis de fertilizantes e insecticidas.
- Describir e inventariar especímenes silvestres.
- Calificar la respuesta de las plantas a diferentes configuraciones de suelo, agua y temperatura.

La combinación de este tipo de cámara con un dispositivo portátil abre nuevas posibilidades para el almacenamiento y análisis de los datos que ahora pueden ser enviados a un servidor automáticamente añadiendo posicionamiento geográfico y generando estadísticas en tiempo real.

Producto de la investigación se desarrolló una aplicación que permite la toma y análisis de las fotografías 3D. El algoritmo codificado en el lenguaje Java, tarda menos de 2 segundos en ejecutarse en un equipo con procesador Intel core I7 con 8 GB de RAM a 2.5 GHz. Se desarrolló un esquema XML propio para almacenar la información 3D.

La base de datos con las imágenes 3D generadas, así como el código fuente pueden descargarse gratuitamente desde el sitio <http://www.serconmur.com/area-foliar-3d.html>.

Apéndice I

ESPECIFICACIONES TÉCNICAS DEL DISPOSITIVO 3D KINECT DE MICROSOFT

En 2010, la empresa Microsoft lanzó la cámara 3D Kinect para su consola de videojuegos Xbox 360. El dispositivo proyecta una nube de puntos infrarrojos invisibles para las personas. Existen 2 versiones de Kinect, la primera se utilizaba en la consola de videojuegos XBOX 360 y la segunda para la consola de videojuegos XBOX One. La primera versión de Kinect fotografía objetivos situados a una distancia de entre 80 cm y hasta 4 metros de distancia.

La segunda versión puede configurarse además para capturar objetivos que estén entre 40 cm y 3 metros. El mapa de profundidades tiene una resolución de 640 * 480 píxeles.

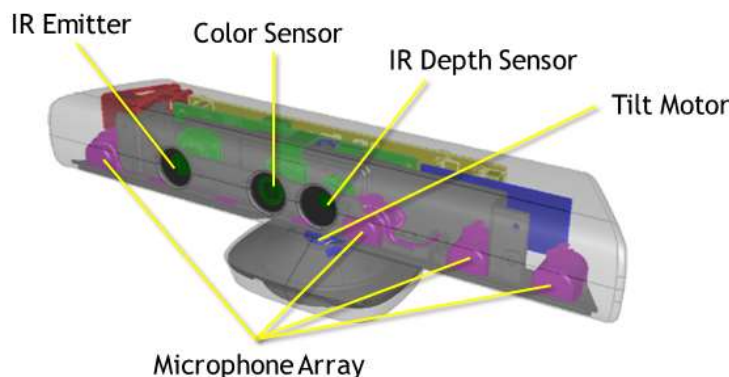


Figura 55 componentes del sensor Kinect (Microsoft, 2018, 1)

En su interior, el sensor contiene:

- Una cámara RGB que almacena tres canales de información en una resolución de 1280*960 píxeles.
- Un emisor infrarrojo (IR) y un sensor de profundidad infrarrojo. El emisor emite rayos de luz y el sensor de profundidad lee los haces de luz reflejados de regreso hacia el sensor. Los rayos reflejados son convertidos en información de profundidad midiendo la distancia entre un objeto y el sensor.
- Un arreglo de 4 micrófonos. Los 4 micrófonos permiten encontrar la fuente del sonido y la dirección de las ondas de audio.

- Un acelerómetro de 3 ejes que determina la orientación del aparato (Microsoft, 2018, 2).

El dispositivo tiene un campo de visión de 43° verticales por 57° horizontales. Su motor puede inclinar la cámara +/- 27° y captura 30 cuadros por segundo.

El sensor incluye su propio eliminador para conectarlo a una fuente de electricidad. Con la ayuda de un adaptador puede conectarse a una computadora.

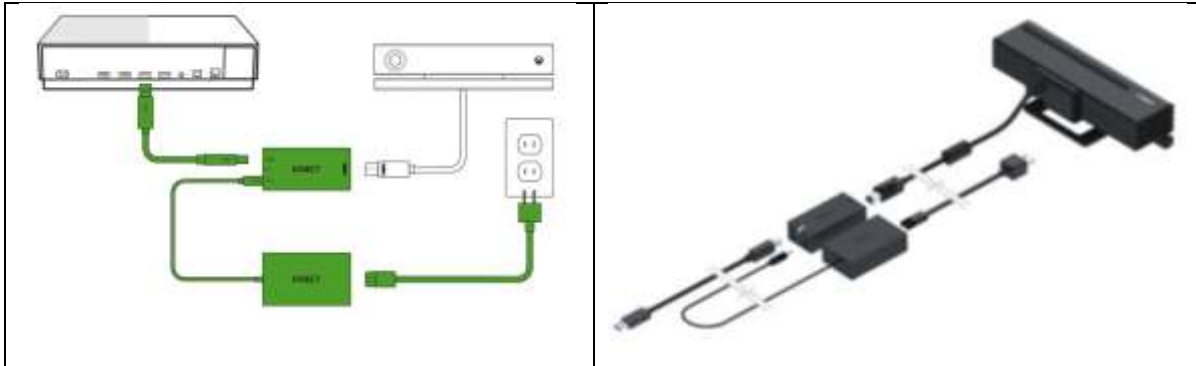


Figura 56. Conector Kinect – PC (Microsoft, 2018)

Las aplicaciones para el sensor Kinect se realizan en el entorno de desarrollo visual estudio, con el lenguaje de programación C# (figura 51).

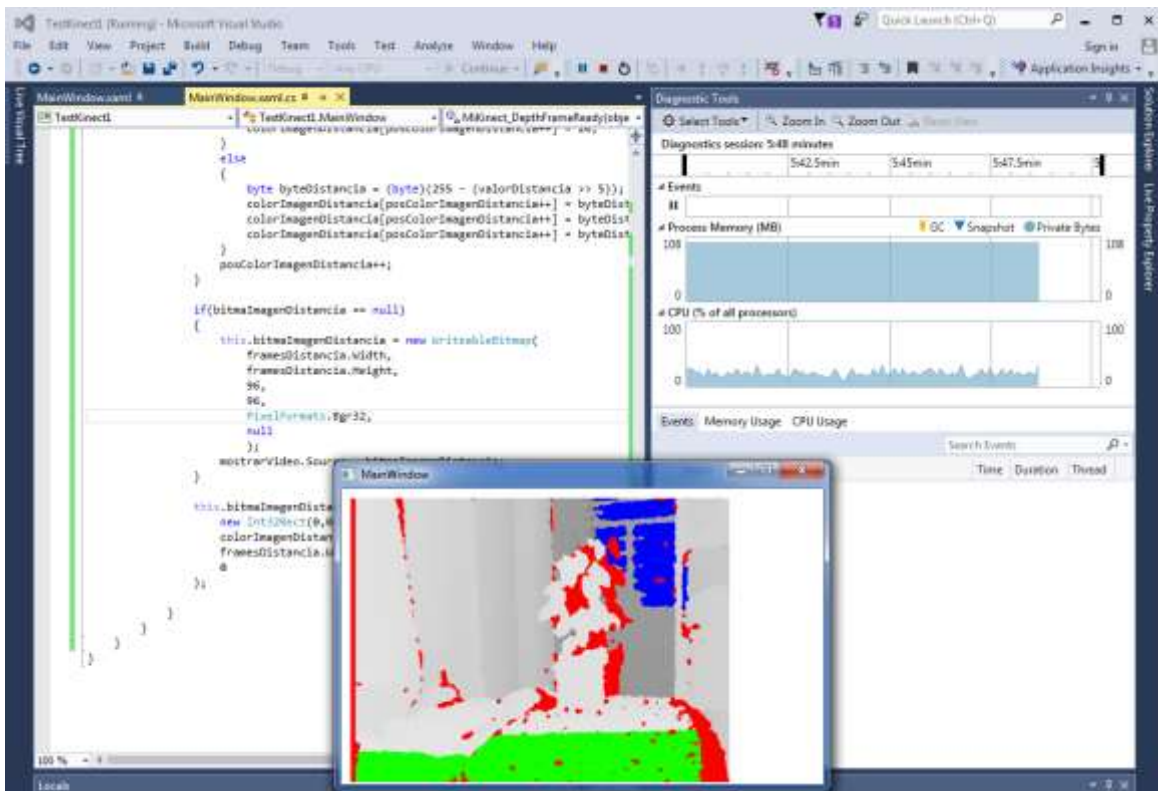


Figura 57. Ejercicio con un sensor Kinect de Microsoft ©

ESPECIFICACIONES TÉCNICAS DE LA CÁMARA 3D INTEL R200

En el año 2015 la compañía Intel presentó la familia de cámaras 3D Realsense, compuesta de un modelo para equipos de escritorio y otro para dispositivos móviles. La cámara modelo R200 está diseñada para utilizarse en exteriores conectada a una tableta electrónica que ejecute el sistema operativo Windows de Microsoft.

La cámara R200 es un dispositivo USB 3.0 que puede proporcionar secuencias de video en color, profundidad e infrarrojos. Su tamaño pequeño proporciona a los programadores flexibilidad para diseñar una amplia gama de productos.

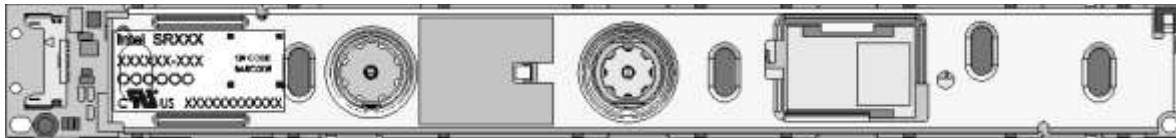


Figura 58. Cámara 3D Intel RealSense R200 (diagrama) (Intel corporation, 2015).

Características

- ASIC (chip) de imágenes embebido.
- Resolución VGA desde 0.4m hasta 2.8m.
- Producto laser infrarrojo clase 1
- Dimensiones de 101.56 mm x 9.55mm x 3.8mm.
- RGB Full HD.

Requerimientos mínimos

- 1GB de espacio en disco duro
- 2GB de memoria RAM
- USB3
- Cable R200

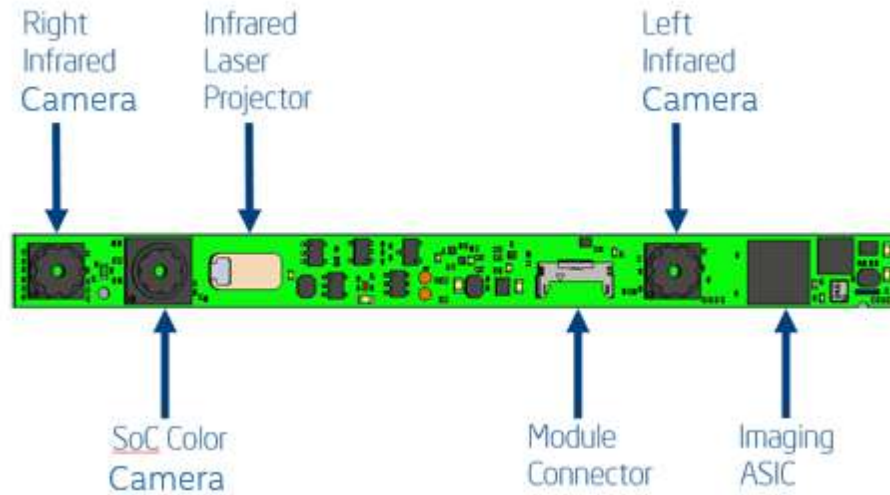


Figura 59. Diagrama de la cámara Intel RealSense R200 (Intel corporation, 2015)

Cuadro 7: Propiedades de las cámaras a color RealSense R200

Píxeles activos	1920x1080
Relación de aspecto	16:9
Tipo de Filtro	Corte Infrarrojo
Enfoque	Automático
Tipo de disparador	Rolling Shutter
Campo de visión Vertical	43° +/-2°
Campo de visión Horizontal	70° +/-2°
Campo de visión Diagonal	77° +/-4°

El módulo tiene 2 cámaras infrarrojas idénticas

Cuadro 8: Propiedades de las cámaras infrarrojas RealSense R200

Píxeles	640x480
Relación de aspecto	4:3
Tipo de filtro	Pase de banda infrarroja
Enfoque	Automático
Tipo de disparo	Global Shutter
Campo de visión Vertical	46° +/-5°
Campo de visión Horizontal	59° +/-5°
Campo de visión Diagonal	70° +/-4.5°

ESPECIFICACIONES TÉCNICAS DE LA CÁMARA 3D INTEL SR300

El modelo SR300 es el dispositivo para equipos de escritorio de la familia de cámaras Intel Realsense. Este aparato utiliza tecnología tiempo de vuelo para capturar la imagen en tercera dimensión. Captura videos de 720 pixeles a 60 cuadros por segundo o de 1080 pixeles a 30 cuadros por segundo. Opera en un rango de profundidades de 20 a 150 cm y se conecta a un puerto USB 3.0.

Aplicaciones

- Reconocimiento facial y seguimiento
- Escaneo y Mapeo
- Segmentación de escena
- Seguimiento de manos y dedos
- Realidad aumentada

Cuadro 9: Parámetros de operación de la cámara Realsense SR300

Parámetro	Cámara	
	Infrarroja	Color
Pixeles activos	640 x 480	1920 x 1080
Relación de aspecto	4:3	16:9
Campo de visión Vertical	55° +/-2°	41.5° +/-2°
Campo de visión Horizontal	71.5° +/-2°	68° +/-2°
Campo de visión Diagonal	88° +/-3°	75.2° +/-4°
Inclinación	+/-1°	+/-1°

DEFINICIÓN XML

Los archivos XML (*Extensible Markup Language*) permiten especificar los metadatos de un objeto. Los archivos XML se diseñan para que su contenido pueda ser procesado automáticamente por una computadora. Los archivos de metadatos pueden crearse, compartirse y consumirse por cualquier software que siga sus directrices.

Los archivos XML implementan una estructura de datos que se define en un archivo XSD (*XML Schema Definition*).

Para esta investigación, se diseñaron archivos de metadatos para almacenar las imágenes 3D, las circunstancias en las que las fotos fueron tomadas y el resultado de su análisis. Además de contener a la imagen en sí, guardan las configuraciones del software al momento de su análisis, ya que configuraciones diferentes generan resultados diferentes.

Almacenar las imágenes en un archivo XML permite distribuirlas en internet y que otras aplicaciones desarrolladas por terceras personas puedan analizar e interpretar la información. También permite la compatibilidad entre varias versiones de un mismo software.

Esquema XSD de las imágenes a Color.

Las imágenes a color están compuestas por un vector de píxeles.

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  elementFormDefault="qualified"
  attributeFormDefault="unqualified">
  <xs:element name="imagenColor">
    <xs:complexType>
      <xs:attribute name="height" type="xs:int"/></xs:attribute>
      <xs:attribute name="pitches" type="xs:int"/></xs:attribute>
      <xs:attribute name="width" type="xs:int"/></xs:attribute>
      <xs:sequence>
        <xs:element name="pixel" maxOccurs="unbounded" type="xs:int">
          </xs:element>
        </xs:sequence>
      </xs:complexType>
    </xs:element>
  </xs:schema>
```

Extracto de un archivo XML de una imagen a color.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<imagenColor height="480" pitches="640" width="640">
  <pixel>-4738635</pixel>
  <pixel>-4343877</pixel>
  <pixel>-4342342</pixel>
  <pixel>-4407879</pixel>
  <pixel>-4473412</pixel>
  <pixel>-4473155</pixel>
  <pixel>-4866371</pixel>
  ...
</ imagenColor >
```

Esquema XSD del mapa de profundidades. Los mapas de profundidades pueden interpretarse como imágenes en escala de grises en los que la intensidad de cada pixel es directamente proporcional a la distancia a la que se encuentra el objetivo.

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  elementFormDefault="qualified"
  attributeFormDefault="unqualified">
  <xs:element name="imagenProfundidad">
    <xs:complexType>
      <xs:attribute name="height" type="xs:int"/></xs:attribute>
      <xs:attribute name="pitches" type="xs:int"/></xs:attribute>
      <xs:attribute name="semillaX" type="xs:int"/></xs:attribute>
      <xs:attribute name="semillaY" type="xs:int"/></xs:attribute>
      <xs:attribute name="width" type="xs:int"/></xs:attribute>
      <xs:sequence>
        <xs:element name="voxel" maxOccurs="unbounded" type="xs:int">
          </xs:element>
        </xs:sequence>
      </xs:complexType>
    </xs:element>
  </xs:schema>
```

Extracto de un archivo XML de un mapa de profundidades.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<imagenProfundidad
  height="468" pitches="640" semillaX="320" semillaY="240" width="628">
  <voxel>-16777216</voxel>
  <voxel>-16777216</voxel>
  <voxel>-16777216</voxel>
  ...
</ imagenprofundidad >
```

Esquema XSD del mapa de vértices.

A diferencia del mapa de profundidades, la posición X, e Y del voxel hacen referencia a su posición en el espacio, y no a su posición en la matriz.

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified"
attributeFormDefault="unqualified">
  <xs:element name="imagenVertices">
    <xs:complexType>
      <xs:attribute name="height" type="xs:int"/></xs:attribute>
      <xs:attribute name="width" type="xs:int"/></xs:attribute>
      <xs:sequence>
        <xs:element name="vertice" maxOccurs="unbounded">
          <xs:complexType>
            <xs:attribute name="x" type="xs:int"/>
            </xs:attribute>
            <xs:attribute name="y" type="xs:int"/>
            </xs:attribute>
            <xs:attribute name="z" type="xs:int"/>
            </xs:attribute>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

Extracto de un archivo XML de un mapa de vértices

```
<?Xml version="1.0" encoding="utf-8" standalone="yes"?>
<imagenvertices height="468" width="628">
  <vertice x="0.0" y="0.0" z="0.0"/>
  <vertice x="0.0" y="0.0" z="0.0"/>
  <vertice x="0.0" y="0.0" z="0.0"/>
  <vertice x="0.0" y="0.0" z="0.0"/>
  <vertice x="0.0" y="0.0" z="0.0"/>
  <vertice x="0.0" y="0.0" z="0.0"/>
  ...
</ imagenvertices >
```

SOFTWARE

La cámara necesita forzosamente:

- Sistema operativo: Windows
- Procesador: Intel
- Puerto: USB 3.0

La aplicación se separó en dos, una para tomar las fotografías y otra para visualizar los resultados, con el objetivo de que incluso las personas que no posean una cámara puedan ejecutar los algoritmos de segmentación utilizando el banco de imágenes generado en la investigación.

La aplicación, su código fuente y el banco de imágenes se encuentra disponible para su descarga gratuita desde www.serconmur.com/opuntia.html

Aplicación para tomar las fotografías

La aplicación permite encender el emisor de luz infrarroja cuando la foto se toma en interiores. Del mismo modo puede disminuirse el tiempo de exposición (tiempo que el diafragma dura abierto) y ganancia (nivel de apertura del obturador) cuando la foto se toma en exteriores para disminuir el efecto adverso del brillo del sol.



Figura 60. Aplicación para la toma de fotografías en 3D

La aplicación arroja las dimensiones del cladodio de nopal señalado por una semilla en el centro de la imagen. Las coordenadas X e Y de la semilla son configurables.

A partir del conjunto de 36 voxeles contenidos en la semilla, se calcula la ecuación de un plano que corta al cladodio. No todos los voxeles que conforman la figura del cladodio se encuentran exactamente en el plano, sino que quedan algunos milímetros por encima o por debajo. Ese valor se configura con el botón "tolerancia".

Para separar un cladodio del tronco que lo sostiene o de sus vástagos se realiza una operación de erosión en la que se borra de la figura un margen especificado por el botón "radio erosión" este valor es directamente proporcional al ancho de las uniones entre cladodios que deben romperse.

No hay problema si el radio de dilatación es mayor al de erosión para no perder definición en los bordes, pues, aunque el valor sea mayor, no crece más allá de la señal original.

Aplicación para explorar las fotografías

Si no se cuenta con la cámara, de todos modos, la imagen puede ser analizada con esta versión del software que no incluye los controles de la cámara.

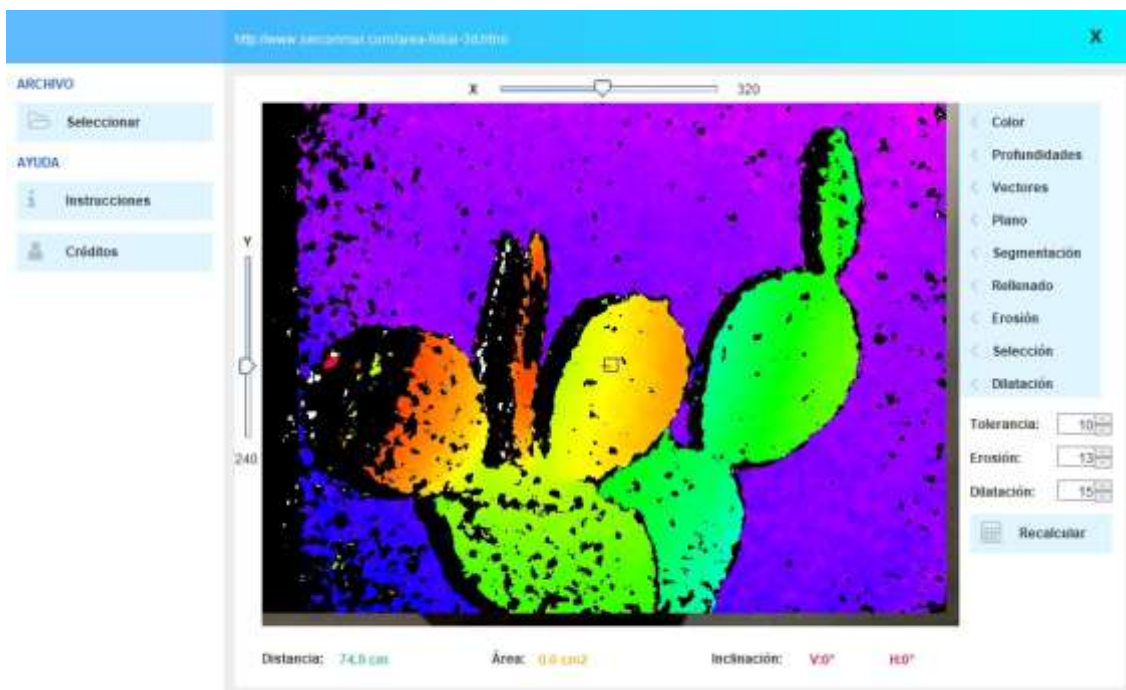


Figura 61. Aplicación para el análisis de las imágenes 3D

CURRICULUM VITAE DEL AUTOR

2014 | *Maestría en ciencias de la computación* |

Universidad Autónoma del Estado de México | Mención Honorífica.

2010 | *Ingeniería en Computación* |

Universidad Autónoma del Estado de México | EGEL – testimonio de desempeño sobresaliente.

2013 | *Integración del lenguaje estadístico R en aplicaciones para el sistema operativo Android* | 4to Congreso nacional de tecnología computacional e informática | UAEM Texcoco.

2014 | *Aplicación para obtener el área foliar desde imágenes obtenidas de un escáner* | 3er. Congreso Internacional de Investigación en Ciencias Básicas y Agronómicas | Universidad autónoma Chapingo.

2016 | *Desarrollo de una aplicación móvil para obtener el área foliar de hojas simples* | Agroproductividad | Colegio de Posgraduados | Vol. 9, Núm. 10, octubre 2016. pp.: 75-78.

2019 | *Medición de cladodios de Opuntia Spp. en fotografías en tercera dimensión* | Agrociencia | Colegio de posgraduados.

GLOSARIO

ASIC: *Application Specific Integrated Circuit*; Circuito Integrado para aplicaciones específicas.

CAD: *Computer Aided Design*; Diseño Asistido por Computadora.

DTO: Data Transfer Object; Objeto de transferencia de datos.

Ganancia: Apertura del obturador.

Exposición: Tiempo que el obturador permanece abierto.

FPS: *Frames per second*; cuadros por segundo.

FHD: *Full High Definition*; Alta definición completa. Resolución de 1920×1080 píxeles.

FOV: *Field of view*; Campo de visión.

GHz: Gigahertz.

GB: Gigabyte.

Heap: Montículo, memoria de la máquina virtual de Java.

IR: *Infrared*; Infrarrojo.

ISO: *International Organization for Standardization*. Organización Internacional de Normalización.

LUX: lumen /m².

MB: Megabyte.

Píxel: unidad mínima de información de las imágenes digitales.

PCL: *Point Cloud Library*. Biblioteca para el tratamiento de nubes de puntos.

RAM: *Random Access Memory*; Memoria de acceso aleatorio.

RGB: *Red, Green, Blue*; Rojo, Verde, Azul.

Spp: especies.

Stack: Pila, Memoria de la máquina virtual Java.

ToF: *Time of Flight*; Tiempo de vuelo.

UML: *Unified Modeling Language*; Lenguaje Unificado de Modelado.

VGA: *Video Graphics Array*; Matriz de gráficos de vídeo. Resolución de 640 × 480 píxeles.

Vóxel: píxel con información espacial.

XML: *Extensible Markup Language*; Lenguaje de Marcado Extensible.

XSD: *XML schema definition*.

BIBLIOGRAFÍA

- Andújar, D., Ribeiro, A., Fernández-Quintanilla, C., & Dorado, J. 2016. Using depth cameras to extract structural parameters to assess the growth state and yield of cauliflower crops. *Computers and Electronics in Agriculture*, 122, 67-73.
- Astegiano, E., J. Favaro, y C. Bouzo. 2001. Estimación del área foliar en distintos cultivares de tomate (*Lycopersicon esculentum* Mill.) utilizando medidas foliares lineales. *Inves. Agrar. Prod. Protecc. Veg.* 16: 2.
- Azzari, G., M. L. Goulden, y R. B. Rosu. 2013. Rapid characterization of vegetation structure with a Microsoft Kinect Sensor *Sensors* 13: 2384-2398.
- Baraona Cockrell, M., & Sancho Barrantes, E; 2000. *Fruticultura general, fruticultura I.* Costa Rica: EUNED, 2000.
- Barioglio, C. F., 2006. *Diccionario de las ciencias agropecuarias - 1ra ed.* Encuentro grupo editor. Argentina. pp 242.
- Burkardt J, 2012, PLY a polygon file format, Stanford University. [online] consultado el 1 de febrero de 2019 en <https://people.sc.fsu.edu/~jburkardt/data/ply/ply.html>
- Casierra-Posada, F.; Peña Z, G. R. y Peña-Olmos J. E. 2007. Estimación indirecta del área foliar en *Fragaria vesca* L., *Physalis peruviana* L., *Acca sellowiana* (Berg.) Burret, *Rubus glaucus* L., *Passiflora Mollissima* (Kunth) L. H. Bailey Y *Ficus Carica* L. *Revista U.D.C.A Actualidad & Divulgación Científica* 11 (1): 95-102, 2007.
- Cerrolaza M. y Flórez López J., 2000, *Modelos matemáticos en ingeniería moderna.* Universidad Central de Venezuela, Consejo de desarrollo científico y humanístico.
- Cervantes R, M. C. 2002. *Plantas de Importancia Económica en las Zonas Áridas y Semiáridas de México.* Instituto de Geografía UNAM. México. pp: 103-104.
- Chen, J. 1996. *Computational geometry: Methods and applications.* Computer science department, Texas A&M University. EEUU.
- CID Bio-Science Inc. 2018. CI-202 Portable laser leaf area meter. <https://cid-inc.com/plant-science-tools/leaf-area-measurement/ci-202-portable-laser-leaf-area-meter/> (Consulta: julio 2018).
- Cock, J., P. Moreno C., M. A. Toro, y J. Arango. 2014. Sistema de Monitoreo No Destructivo en Yuca. *Centro Int. Agric. Trop.* pp: 9-10.

- Contreras M, S., O. Lugo E., y J. Ayala. V. 2015. Aplicación para obtener el área foliar desde imágenes obtenidas de un escáner. Ciencias Químicas y Matemáticas-©ECORFAN. Universidad Autónoma Chapingo. Handbook T-I: 37-46.
- Cousins, R. B. 2013. 3D is here: Point Cloud Library (PCL). IEEE International Conference on Robotics and Automation (ICRA). [online]. <http://www.pointclouds.org/>
- Drennan, R. D. y Gonzales V., 2019, Estadística para arqueólogos: un enfoque de sentido común. Primera edición en español. Universidad de los Andes, Facultad de Ciencias Sociales, Departamento de Antropología, Ediciones Uniandes, Bogotá.
- FAO, 2017. Crop ecology, cultivation and uses of cactus pear. The food and Agriculture Organization of the United Nations and The International Center for Agricultural Research in the Dry Areas Rome.
- Faúndez Zanuy, Marcos. 2000. Tratamiento Digital de Voz e Imagen y Aplicación a la Multimedia. Marcombo. España. pp: 172.
- Forestry suppliers, 2015. CID Bio-Science Portable Laser Leaf Area Meter. Recuperado el 6 de septiembre de 2015, de <https://www.cid-inc.com/plant-science-tools/leaf-area-measurement/ci-202-portable-laser-leaf-area-meter/>
- Gaspar R, C. M., L. C. Gazarini, T. F. Fonseca, and M. M. Ribeiro. 2018. Above-ground biomass estimation of *Opuntia Ficus-indica* (L.) Mill. for forage crop in a Mediterranean environment by using non-destructive methods. *Experimental Agriculture*, Cambridge University Press. April 2018: 1-16.
- Hernández O., J., M. J. Ramírez Q., y C. Ferri R. 2004. Introducción a la Minería de Datos. Pearson Educación. 680 p.
- HP Development Company, L.P, 2016. Sprout Pro by HP | HP® Official Site 2019, HP Development Company, L.P. (s.f.). [online]. Consultado el 30 de abril de 2016 en <http://www8.hp.com/us/en/sprout/home.html>
- Intel corporation. 2015. SDK Design Guidelines, Intel RealSense camera R200, versión 1.1.
- Intel corporation, 2016. Tecnología Intel Real Sense. [online]. Consultado el 30 de abril de 2016, de <http://www.intel.la/content/www/xl/es/architecture-and-technology/realsense-overview.html>
- Kosmos Scientific de México. 2018. Medidor portátil de área foliar ci-203. <https://www.kosmos.com.mx/tienda/catalogo/medidor-portatil-area-foliar-p-2607.html> (Consulta: julio 2018).

- Kolman, B., y D. R. Hill. 2006. Aplicaciones de vectores en R2 y R3. In: Álgebra Lineal. 8va ed. Pearson Educación. pp: 261.
- Kraus, T. A., Blanco, C. A. y Vegetti, A. C. 2004. La hoja: Morfología externa y anatomía. Universidad nacional de rio cuarto, Universidad nacional del litoral, Argentina.
- La razón de México. 2016. China y Brasil compiten con México en producción de nopal. <https://www.razon.com.mx/negocios/china-y-brasil-compiten-con-mexico-en-produccion-de-nopal/> (Consulta: julio 2018).
- López-Tolentino, G., Lira-Saldivar R. H. y Méndez-Argüello B., 2016. Medición de Intercambio Gaseoso, Área Foliar e Índice de Clorofila en Plantas Elicitadas con Nanopartículas. 2do mini simposio – taller de agro-nano tecnología. CoNaCyT.
- Martínez-Rodríguez, Ulises, 2015. Aplicación de la geometría computacional en la reconstrucción 3D basada en diagramas de Voronoi. Tesis para la obtención de grado de licenciatura. Benemérita Universidad Autónoma de Puebla. Puebla, México.
- Mendes P, K., T. G. Freire S., W. J. da Silva D., H. F. de Sousa C., y M. S. Beserra M. 2015. Indirect methods for determining the area index of forage cactus cladodes. *Pesq. Agropec. Trop.* 45: 63-171.
- Microsoft 2018 1, Kinect for Windows, Kinect - Windows app development, [online]. <https://developer.microsoft.com/en-us/windows/kinect>
- Microsoft 2018 2 Uso de un sensor Kinect para Xbox One con una consola Xbox One S o Xbox One X; [online]. <https://support.xbox.com/es-ES/xbox-one/accessories/kinect-adapter>
- Miranda Moles, A., 2003, Tesis: El método de remuestreo y su aplicación en la investigación biométrica. Escuela nacional de salud pública, Cuba.
- Müller L. 2000, Manual de laboratorio de morfología vegetal, Centro agronómico tropical de investigación y enseñanza (CATIE) Turrialba Costa Rica, prefacio p. iv.
- Nissimov S., Goldberger J. y Alchanatis V., 2015. Obstacle detection in a greenhouse environment using the Kinect sensor, *Computers and Electronics in Agriculture* 113 (2015) 104–115.
- Nixon, M., & Aguado, A. 2002. Feature Extraction and image processing. Newnes, Elsevier. RU.
- Nobel, P., y R. W. Meyer. 1991. Biomechanics of cladodes and cladode-cladode junctions for *Opuntia ficus-indica* (Cactaceae). *Am. J. Bot.* - AMER J BOT. 78. 10.2307/2444929.

- Okayamaa, T.; Gotoa, T.; Suzukid, T.; Toyodaa, A. 2015. Monitoring Nest Building in Mice and Leaf Movement of Kidney Beans using an Infrared Range Camera, *Agriculture and Agricultural Science Procedia*, Elsevier, Número 3, 40 – 44.
- Pajares M, G., y J. de la Cruz G. 2008. *Visión por Computador: Imágenes Digitales y Aplicaciones*. 1ra ed. Alfaomega Grupo Editor, México. 764 p.
- PCL. 2019, The PCD (Point Cloud Data) file format — PCL 0.0 documentation - Point Cloud Library, point cloud library http://pointclouds.org/documentation/tutorials/pcd_file_format.php
- Pérez-Perales G, 2016, localización de puntos en subdivisiones planas. Algoritmo de Kirkpatrick. Universidad politécnica de Madrid, facultad de informática. TRABAJO de FIN DE CARRERA, [Online]. Consultado el 1 de febrero de 2016 en <http://www.dma.fi.upm.es/personal/mabellanas/tfcs/kirkpatrick/Aplicacion/algoritmos.htm>
- Ramhari, G., Pravin, B., Mahadev, A., A. Bhange, S., y Ishrat, K. 2015. Influence of Different Rootstocks on Growth, Photosynthesis, Biochemical Composition, and Nutrient Contents in 'Fantasy Seedless' Grapes. *International Journal of fruit science*, 15:3, 251-266.
- Rajasekaran, S. y Reif J. 2007. *Handbook of Parallel Computing. Models, Algorithms and Applications*. CRC Press. U.S. pp. 16.
- Rebay, S. 1993. Efficient Unstructured mesh generation by means of Delaunay triangulation and Bowyer-Watson Algorithm. *Journal of computational physics* 106, 125 – 138.
- Rincón G., N., M. Olarte Q, y J. Pérez N. 2012. Determinación del Área Foliar en Fotografías Tomadas con una Cámara Web, un Teléfono Celular o una Cámara semiprofesional. *Revista Facultad Nacional de Agronomía Medellín*. Universidad Nacional de Colombia. 65: 6399 -6405.
- Rusu R. B. and S. Cousins, 2011. "3D is here: Point Cloud Library (PCL)," 2011 IEEE International Conference on Robotics and Automation, Shanghai, pp. 1-4.
- Sarbolandi H., D. Lefloch, y A. Kolb. 2015. Kinect range sensing: Structures-light versus time-of-flight kinect. *Comp. Vision and Image Understand*. 139: 1-20.
- Scalisi, A., B. Morandi, P. Inglese, and R. Lo-Bianco. 2016. Cladode growth dynamics in *Opuntia Ficus-indica* under drought. *Environ. Exp. Bot*. 122: 158-167.
- Sortibrán, L., C. Tinoco-Ojanguren, T. Terrazas, and A. Valiente-Banuet. 2005. Does cladode inclination restrict microhabitat distribution for *Opuntia Puberula* (cactaceae)? *Am. J. Bot*. 92: 700-708.

Universidad nacional del sur, 2019. Procesamiento Digital de Imágenes - Idea subyacente. Argentina [Online] Consultado el 1 de febrero de 2019 en <http://www.lip.uns.edu.ar/pdi/index42.htm>

Zujevs A., V. Osadcuck, and P. Ahrendt. 2015. Trends in robotic sensor.

3D systems, 2019, Estereolitografía, [online]. Consultado el 1 de febrero de 2019 en <https://es.3dsystems.com/on-demand-manufacturing/stereolithography-sla#>