

Inspección automática de jitomate usando digitalización 3D

Hector Josue de la Riva Martinez.^a, Delgadillo Wong Graciela^b, Espejel Zarate Ricardo^c

^aInstituto Politécnico Nacional, CECYT 9 "Juan de Dios Bátiz",

^bTecnológico Nacional de México campus Culiacán,

^cInstituto Tecnológico Superior del Oriente del Estado de Hidalgo,

Abstract

Industry 4.0 is constantly changing traditional methods. It is important to implement solutions in the food sector today. 3D digitization technology with fringe projection and structured light is a broad field of exploration. The use of optical methods for 3D reconstruction of fruits is innovative because it uses minimal computational and technological resources, making it a viable alternative for fruit classification. In this project, we studied the application of 3D reconstruction technology by fringe projection for tomato classification. The results obtained indicated error margins of 1 to 4 millimeters in the detection of the actual size of Saladettes and smooth round tomatoes. It is interesting to explore the use of mathematics for the estimation of these dimensions. Finally, we provide a retrospective on the existing foreign technology today, which is expensive for small farmers who do not have access to it. The implementation of this local and accessible technology provides a valuable opportunity to improve classification processes in the agricultural and food sector, especially benefiting Mexican farmers.

Keywords: Digitalización 3D, proyección de luz estructurada, agricultura, clasificación.

Índice		5. Discusión y Resultados	11
1. Resumen	1	6. Conclusiones	11
2. Introducción	2	7. Fuentes Bibliográficas	12
2.1. Planteamiento del problema	2	1. Resumen	
2.2. Antecedentes	2	La industria 4.0 produce cambios constantes en los métodos tradicionales. Implementar soluciones en el sector alimentario resulta importante en la actualidad. La tecnología de digitalización en 3D con proyección de franjas y luz estructurada es un campo amplio de exploración. El uso de métodos ópticos para reconstrucción 3D de frutos resulta innovadora, ya que emplea recursos computacionales y tecnológicos mínimos, siendo una alternativa viable para la clasificación de frutos. En este proyecto, se estudió la aplicación de la tecnología de reconstrucción 3D por proyección de franjas para la clasificación de jitomate. Los resultados obtenidos indicaron márgenes de error de 1 a 4 milímetros en la detección del tamaño real de los jitomates Saladettes y redondos lisos. Es interesante explorar el uso de las matemáticas para la estimación de estas dimensiones. Finalmente, se realiza una retrospectiva sobre la tecnología extranjera existente en la actualidad, que resulta costosa para los pequeños granjeros que no tienen acceso a ella. La implementación de esta tecnología local y accesible brinda una oportunidad valiosa para mejorar los procesos de clasificación en	
2.3. Objetivos	2		
3. Marco teórico	3		
3.1. Metodología	3		
3.1.1. Modelo pinhole de formación de imagen	3		
3.1.2. Homografías: relación proyectiva entre planos	3		
3.2. Calibración: estimación de parámetros usando homografías	4		
3.3. Demodulación de fase	4		
3.4. Triangulación	5		
3.4.1. Intersección línea-línea	5		
3.4.2. Intersección línea-plano	5		
3.5. Reconstrucción de una esfera	6		
4. Evaluación Experimental	7		
4.1. Descripción del sistema	7		
4.1.1. Materiales	7		
4.2. Calibración	7		
4.3. Proyección y procesamiento	8		
4.4. Implementación real	9		

el sector agrícola y alimentario, beneficiando especialmente a los agricultores mexicanos.

2. Introducción

El constante diagnóstico de la calidad de los alimentos ha sido un tema de interés durante mucho tiempo, ya que su calidad depende de varios factores que pueden afectar su idoneidad para el consumo. Sin embargo, los procesos de clasificación pueden resultar costosos, lo que limita el acceso a estas tecnologías a empresas de gran tamaño, dejando en desventaja al sector micro-empresarial con respecto a las ventajas tecnológicas y biológicas.

Los sistemas de clasificación de alimentos son esenciales para ofrecer productos de alta calidad al mercado (Babio, N, Casas-Agustech, P., Salas, J., 2020). Por lo tanto, es crucial poder detectar características como textura, forma, color, tamaño, entre otras, para garantizar la excelencia de los productos, especialmente aquellos destinados a la exportación.

En este contexto, la visión por computadora se presenta como una alternativa viable para la detección de características, en lugar de depender únicamente del recurso humano, que puede ser susceptible a errores. Esto permite redirigir recursos hacia un mejor cultivo o conservación de los alimentos. La reconstrucción de objetos mediante tecnología de industria 4.0 se convierte en un camino prometedor tanto para el sector alimentario como para el campo computacional, ya que en un mundo globalizado es esencial interactuar eficientemente con los recursos disponibles. La digitalización 3D, utilizando cámaras y unidades ópticas, resulta más accesible y fácil de implementar en comparación con métodos tradicionales, aunque el tiempo de captura puede variar.

En este proyecto, abordaremos los principios teóricos que sustentan la digitalización de objetos en el espacio, así como su aplicación en la reconstrucción de alimentos redondos, y su manipulación computacional.

2.1. Planteamiento del problema

En la actualidad, el mercado del sector alimentario en México, especialmente para huertos pequeños o empresas de invernadero, presenta una oportunidad de venta durante los meses de julio, agosto y noviembre. Dentro de este mercado, el jitomate ha sido uno de los principales cultivos en el país, con más de 60 especies distintas y variadas formas. Sin embargo, el proceso de clasificación de este fruto en empresas de gran envergadura resulta costoso, lo que genera una desventaja significativa para los agricultores de menor tamaño. Estos agricultores dependen de la visión y experiencia humana para la clasificación, lo que no garantiza la

calidad del producto y consume tiempo significativo, afectando la eficiencia y calidad del proceso.

Las limitaciones debidas a la inspección manual hace evidente la necesidad de una solución que permita a los pequeños agricultores acceder a un método de clasificación más eficiente y preciso para el jitomate, que garantice la calidad del producto y reduzca el tiempo y esfuerzo invertidos. La aplicación de tecnologías de visión por computadora y reconstrucción 3D se presenta como una posible alternativa para mejorar el proceso de clasificación del jitomate, ofreciendo una mayor precisión y eficiencia en la selección del producto final, con el objetivo de impulsar su competitividad en el mercado y contribuir al crecimiento sostenible del sector alimentario en México.

2.2. Antecedentes

Multiscan S90 Tomato(2023)

Sistema compacto especialmente diseñado para la selección de diversas variedades de tomate con un tamaño superior a 40 mm, basándose en criterios de calidad, color, tamaño y forma. Este sistema permite el transporte del fruto mientras realiza un giro completo de 360°, logrando una inspección integral que detecta problemas de calidad empleando sistemas de visión artificial.

CUPSIZER Pentacup-M(2022)

Sistema de clasificación para tomates usando visión por cámaras y bandas expansivas en V que aseguran una transferencia delicada.

TOMRA Sorting NFM(2016)

Maquinaria diseñada para detectar y rechazar frutos descoloridos, con moho, podredumbre, manchas solares, quemados, verdes, y aquellos que no cumplan con el tamaño especificado. Además, tiene la capacidad de identificar y eliminar todo tipo de materiales extraños, como materia animal, tallos de algodón, metal, plástico, piedras, cristal, madera, entre otros. Esta máquina robusta se monta sobre una cosechadora y realiza una eficaz clasificación de alimentos, utilizando tecnología de clasificación por color y detección de materiales extraños.

2.3. Objetivos

Objetivo General

Desarrollar un prototipo basado en visión por computadora para la clasificación automatizada de jitomates, permitiendo obtener información detallada y lograr una clasificación precisa.

Objetivos Específicos

1. Diseñar un prototipo fácil de transportar, con un ajuste sencillo para una cinta transportadora de jitomates.
2. Reemplazar el uso de algoritmos de aprendizaje automático por el cálculo matemático de centroides y dimensiones de los jitomates, con el fin de automatizar el proceso de clasificación en la distribución.
3. Reducir el tiempo y la experiencia requerida por los agricultores en la clasificación de los productos, mediante el empleo de técnicas de digitalización óptica, con el propósito de disminuir los costos de operación.
4. Contribuir al estudio de clasificación de verduras y frutas mediante el procesamiento de datos y técnicas matemáticas, utilizando visión por computadora, con el fin de construir bases de datos adecuadas para su posterior análisis.

3. Marco teórico

3.1. Metodología

3.1.1. Modelo pinhole de formación de imagen

Para el proceso de formación de imagen es necesario conocer el concepto del plano imagen y del sistema de formación de imágenes. El plano imagen es el lugar en donde la cámara detecta la luz que recibe de las escenas tridimensionales del mundo real. En el caso de la cámara pinhole, el plano imagen es la pared posterior de la cámara. Mientras que el sistema óptico es el conjunto de lentes y otros elementos que capturan la luz que viene de la escena, dirigiendola al plano imagen, en el caso del modelo pinhole es posible estudiar el proceso de formación de imagen considerando que todo el sistema óptico es un único agujero pequeño conocido como el como *pinhole*.

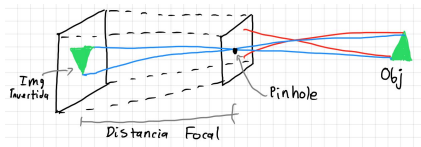


Figura 1: Modelo del sistema pinhole.

Para el proceso de formación de imagen en el modelo del pinhole, todos los puntos de la imagen se comportan de tal forma que emiten rayos de luz en todas direcciones, en el caso más ideal, la pared del sistema bloqueará todos los rayos que esta emite a excepción de uno, el cual cruza por el pinhole, siendo el único que llegue a la pared del fondo de la cámara, grabandose como un punto de la imagen, este procedimiento se repite para cada uno de los puntos de la escena, por lo que cada punto esta relacionado mediante una linea recta que es

la distancia entre la pared del pinhole y la pared de la imagen (distancia focal).

Este proceso tiene como principal problemática la inversión tanto del eje X como del eje Y. Para corregir este problema tenemos que ayudarnos matematicamente del concepto de las coordenadas homogéneas.

3.1.2. Homografías: relación proyectiva entre planos

Las homografías son transformaciones proyectivas que establecen un correspondencia entre elementos. Esto conlleva a que a un punto le corresponde otro punto y a una recta le corresponde otra recta.

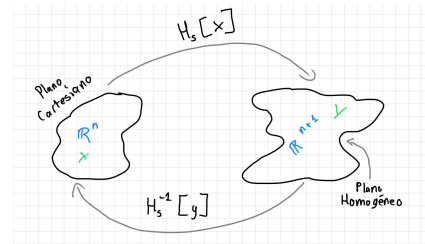


Figura 2: Relación entre coordenadas cartesianas y coordenadas homogéneas a través del operador \mathcal{H} .

Para poder obtener dos puntos ρ y μ , debemos auxiliarnos del concepto de coordenadas homogéneas, para el cual es necesario realizar homografías. Si queremos obtener la coordenada correspondiente en el plano homogéneo teniendo la coordenada del plano real tenemos que utilizar el siguiente modelo matemático:

$$\mu = G\mathcal{H}_s(\rho), \quad (1)$$

donde G es una matriz 3×3 , no singular y con determinante distinto de cero.

A su vez el operador \mathcal{H}_s nos permite ir del espacio de las coordenadas cartesianas (n dimensiones) al espacio de coordenadas homogéneas ($n + 1$ dimensiones). Este operador esta definido de la siguiente forma:

$$\mathcal{H}_s[v] = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \\ s \end{pmatrix} \quad (2)$$

Podemos usar este proceso de forma inversa, para pasar de coordenadas homogéneas a coordenadas cartesianas utilizamos el siguiente modelo matemático:

$$\rho = \mathcal{H}_s^{-1}(G\mathcal{H}_s(\rho)), \quad (3)$$

donde G es una matriz 3×3 , no singular y con determinante distinto de cero.

Y el operador \mathcal{H}_s inverso nos permite ir del espacio de las coordenadas homogéneas ($n + 1$ dimensiones) al espacio de coordenadas cartesianas (n dimensiones). Este operador está definido de la siguiente forma:

$$\mathcal{H}_s^{-1}[h] = \frac{s}{s[h]} \mathcal{H}_s[h] = v, \quad (4)$$

donde definimos h como:

$$h = \begin{pmatrix} \lambda h_1 \\ \lambda h_2 \\ \vdots \\ \lambda h_n \\ s \end{pmatrix} \quad (5)$$

Es útil obtener las homografías porque necesitamos obtener los parámetros de la cámara y el proyector con los que trabajaremos, por lo que es necesario al menos 3 homografías para obtener dichos parámetros.

3.2. Calibración: estimación de parámetros usando homografías

La calibración en la estimación de parámetros utilizando homografías se refiere a un proceso para determinar parámetros intrínsecos y extrínsecos de una cámara a partir de por lo menos tres homografías.

El proceso de calibración de cámaras, es necesario para poder extraer información métrica a partir de imágenes 2D del mundo 3D.

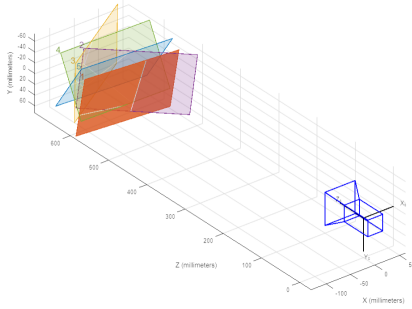


Figura 3: Calibración con 5 homografías.

Este método se basa en el posicionamiento fijo de la cámara observando una escena que a partir de sus desplazamientos son tomadas cinco imágenes por una misma cámara con parámetros intrínsecos fijos que serán suficientes para obtener tanto los parámetros extrínsecos como intrínsecos.

Los parámetros intrínsecos son aquellos que caracterizan la geometría interna de la cámara, como la distancia

focal, la ubicación del punto principal, la distorsión radial y tangencial, entre otros.

Los parámetros extrínsecos son aquellos que describen la posición y orientación de la cámara con respecto a un sistema de coordenadas externo, como la posición del centro óptico y los ángulos de rotación en los ejes del mundo real.

Para obtener dichos parámetros utilizamos los siguientes modelos matemáticos:

$$K = k \Xi_f^{-1}, \quad (6)$$

donde K es una matriz triangular superior de tamaño 3×3

$$R = [r_1 r_2 r_3], \quad (7)$$

donde R es la matriz de rotación de tamaño 3×3 y r_1, r_2, r_3 son vectores. La posición de la cámara está dada por el vector.

$$t = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad (8)$$

donde t es el vector de translación de la cámara de tamaño 3×1 , con coordenadas x, y, z .

$$G = K[\bar{r}_1 \bar{r}_2 - R^T t], \quad (9)$$

donde G es una matriz de tamaño 3×3 , que no debe de ser singular y su determinante debe ser distinto de cero.

Una vez obteniendo las imágenes se detectan las correspondencias entre puntos del objeto o patrón en diferentes imágenes. Estas correspondencias se utilizan para calcular la matriz de homografía, que es una matriz 3×3 que representa la transformación proyectiva entre las dos imágenes.

Con la información de múltiples homografías y un objeto de calibración con propiedades conocidas (por ejemplo, una cuadrícula con puntos de coordenadas bien definidas), es posible estimar los parámetros intrínsecos y extrínsecos de la cámara.

3.3. Demodulación de fase

En el contexto de este trabajo, es esencial codificar las coordenadas proyectadas en la diapositiva (patrón de calibración) para capturar la forma del objeto. Para lograr esto, utilizamos modelos matemáticos que definen la proyección de franjas en forma de coseno:

$$Gx(x, y) = \frac{1}{2} + \frac{1}{2} \cos(2\pi Fx + \delta_k), \quad (10)$$

$$Gy(x, y) = \frac{1}{2} + \frac{1}{2} \cos(2\pi Fy + \delta_k). \quad (11)$$

Cuando un punto p en el espacio 3D es iluminado por las rejillas Gx y Gy , el proceso de demodulación de fase devuelve las coordenadas v_x y v_y que representan la ubicación donde el proyector "ve" el punto p .

La demodulación de fase generalmente se lleva a cabo en dos pasos: extracción de la fase envuelta y desenvolvimiento de fase. La fase codificada varía entre $-\pi$ y π radianes.

Una ventaja de la demodulación de fase absoluta es que se logran mediciones precisas al proyectar múltiples rejillas con diferentes frecuencias. Específicamente, las rejillas de baja frecuencia proporcionan una fase absoluta y simplifican el proceso de desenvolvimiento de fase, mientras que las rejillas de alta frecuencia aumentan la precisión de la medición.

Para obtener las coordenadas proyectadas, utilizamos un conjunto de patrones de franjas con desplazamientos de fase uniformemente distribuidos en el intervalo $[0, \frac{2\pi}{p}]$. Este conjunto está definido por:

$$d_p = \left(\frac{2\pi}{p} \right) (1, 2, \dots, n), \quad (12)$$

donde p es el número de desplazamientos de fase (también conocido como el número de pasos).

3.4. Triangulación

La triangulación es una herramienta matemática que nos sirve para poder encontrar la posición de un objeto dado dos puntos de referencia que estén observando dicho objeto, en este caso esos dos puntos serán la cámara y el proyector respectivamente de nuestro sistema previamente calibrado, ya que para realizar la triangulación es necesario contar con los parámetros intrínsecos y extrínsecos tanto de la cámara como del proyector.

3.4.1. Intersección línea-línea

Para la triangulación línea-línea es necesario utilizar la demodulación de fase tanto en el eje X, como en el eje Y, esto quiere decir que tendremos que realizar más capturas de franjas para poder realizarla, el beneficio que este método tiene es que la calidad y precisión en la triangulación de puntos va a ser mejor, lo que dará una nube de puntos de la reconstrucción con mejor calidad, sin embargo tiene como principal desventaja el consumo de más almacenamiento. como se ve en la siguiente imagen.

Para poder realizar la triangulación por intersección línea-línea es necesario tener la dirección del proyector y de la cámara, las cuales serán respectivamente una

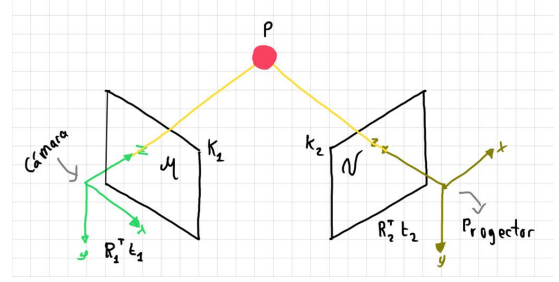


Figura 4: Determinación de puntos en el espacio 3D por línea-línea.

matriz de $3 \times N$, en donde cada columna será una dirección. Además se necesitarán contar con los vectores de posición correspondientes a cada uno, que serán de tamaño 3×1 .

Como resultados obtendremos una matriz de tamaño $3 \times N$ donde cada columna será la posición del punto observado, con lo cual podremos construir la nube de puntos del objeto. Para esto utilizamos el siguiente modelo matemático.

$$P = \frac{1}{2}(t_1 + t_2 + \lambda_1 d_1 + \lambda_2 d_2), \quad (13)$$

donde d_1 y d_2 son respectivamente:

$$d_1 = R_1 K_1^{-1} H[\mu], \quad (14)$$

$$d_2 = R_2 K_2^{-1} H[\nu]. \quad (15)$$

3.4.2. Intersección línea-plano

La intersección línea-plano se refiere al punto en el espacio tridimensional donde una línea y un plano se cruzan o se encuentran. En matemáticas y geometría, este punto de intersección se determina al encontrar las coordenadas que satisfacen tanto la ecuación de la línea como la ecuación del plano, tal como se muestra en la siguiente figura 5.

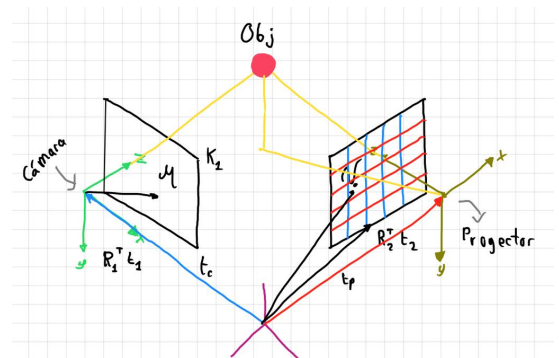


Figura 5: Determinación de puntos en el espacio 3D por línea-plano.

Es importante destacar la ventaja de este método en comparación con el método de intersección línea-línea, ya que aunque es más robusto, su procesamiento es más lento, lo que requiere la mitad de las imágenes y solo proporciona $(2m + n)$ restricciones para la estimación tridimensional de puntos, donde 'm' es el número de cámaras y 'n' es el número de proyectores.

Para encontrar la intersección entre la línea y el plano, debemos resolver un sistema de ecuaciones que incluya tanto la ecuación de la línea como la ecuación del plano. La solución de este sistema proporcionará las coordenadas del punto de intersección.

3.5. Reconstrucción de una esfera

Para la reconstrucción de la nube de puntos, se emplea como referencia la ecuación de la esfera con centro fuera del origen, ya que se llegó a la conclusión de que la esfera se aproxima mucho a la forma del objeto estudiado y no es tan difícil de trabajar como un ovaloide.

$$(x - Cx)^2 + (y - Cy)^2 + (z - Cz)^2 = r^2. \quad (16)$$

Ecuación de la esfera con centro fuera del origen.

Lo primero que necesitamos para aproximar la nube de puntos, es conocer su centroide, lo cual conseguimos utilizando la siguiente ecuación:

$$C = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \frac{1}{n} \sum_{k=1}^n P. \quad (17)$$

Ecuación para calcular el centroide de una esfera.

Donde definimos P como:

$$P_i = \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix}. \quad (18)$$

Una vez que se obtiene el valor aproximado del centroide, se procede a calcular el valor del radio, que es otro dato que se necesita para poder reconstruir la nube de puntos, para lograr esto utilizamos la siguiente ecuación:

$$r = \frac{1}{n} \sum_{k=1}^n \|P_k - C\|. \quad (19)$$

Esta forma de aproximación que hicimos hasta este punto no es la forma más precisa que se puede utilizar para poder reconstruir la nube de puntos, por lo que tenemos que proceder con otro método matemático para poder lograr una mayor exactitud en la aproximación, por lo que primero se debe continuar con el desarrollo de la ecuación de la esfera con centro fuera del origen.

$$\begin{aligned} x^2 + Cx^2 - 2xCx + \dots \\ y^2 + Cy^2 - 2yCy + \dots \\ z^2 + Cz^2 - 2zCz - r^2 = 0, \end{aligned} \quad (20)$$

donde despejamos la ecuación para dejar de un lado los siguientes términos que renombraremos por facilidad con otra variable:

para el lado derecho de la ecuación:

$$x^2 + y^2 + z^2 = y, \quad (21)$$

para el lado izquierdo de la ecuación:

$$\begin{aligned} f(\Theta) = Cx^2 + Cy^2 + Cz^2 \dots \\ -2xCx - 2yCy - 2zCz + r^2, \end{aligned} \quad (22)$$

donde Θ es:

$$\Theta = (Cx, Cy, Cz, r^2), \quad (23)$$

por lo que ahora la ecuación reescrita es :

$$f(\Theta) = y. \quad (24)$$

Como siguiente paso, necesitamos linealizar la ecuación a la que llegamos, por lo que se procede a utilizar mínimos cuadrados:

$$x = (A^T A)^{-1} A^T y. \quad (25)$$

Dada la estructura a la que queremos llegar, se debe de utilizar series de Taylor para poder linealizar la función:

$$\sum_{k=0}^{\infty} \frac{f^k(x_0)(x - x_0)^k}{k!} = f(x)$$

$$f(\Theta) = f(x_0) + f'(x_0)x + \frac{1}{2}f''(x_0)(x - x_0)^2. \quad (26)$$

Por lo que tomando como referencia de nuevo la ecuación de la esfera con centro fuera del origen, para poder definir la siguiente función:

$$f(\Theta) = (x - Cx)^2 + (y - Cy)^2 + (z - Cz)^2 - r^2. \quad (27)$$

Una vez definida la función se debe obtener las derivadas parciales utilizando cada uno de los valores que adquiero Θ :

$$\Theta = (Cx, Cy, Cz, r^2), \quad (28)$$

respecto al radio cuadrado

$$\frac{\partial f}{\partial r^2} = -1, \quad (29)$$

respecto a C_x

$$\frac{\partial f}{\partial C_x} = -2(x - C_x), \quad (30)$$

respecto a C_y

$$\frac{\partial f}{\partial C_y} = -2(y - C_y), \quad (31)$$

respecto a C_z

$$\frac{\partial f}{\partial C_z} = -2(z - C_z). \quad (32)$$

De lo que se concluye:

$$J = \begin{bmatrix} -1 & -2(x-x_0) & -2(y-y_0) & -2(z-z_0) \end{bmatrix}, \quad (33)$$

$$\Theta_{k+1} = \Theta_k - J^+(Y_{mod}). \quad (34)$$

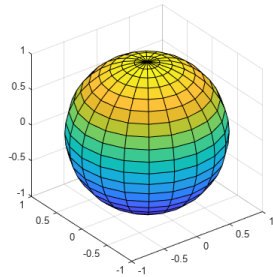


Figura 6: Reconstrucción aproximada de la esfera en Matlab.

El último paso, se realiza con la finalidad de obtener una aproximación con un margen mínimo de error en la reconstrucción de la nube de puntos, cabe mencionar que para la serie de Taylor se realizaron 25 iteraciones en los cálculos para maximizar los resultados.

4. Evaluación Experimental

4.1. Descripción del sistema

4.1.1. Materiales

1. Cámara Webcam Logitech C270 con resolución 1280x720px.
2. Proyector Epson PowerLite W39 con resolución 800x1280px.
3. Patrón de ajedrez de 7x10 cuadrados de 17x17m.
4. Computadoras gama media con software Matlab, Meshlab y Blender.
5. Esferas de unicel de 8cm, 7.2cm, 6.5cm, 5.5cm, 4.1cm, 3.5cm y 2.5cm, en color blanco y rojo.
6. Muestras de jitomates Saladette, Cherry y Redondo liso.7



Figura 7: Materiales.

4.2. Calibración

Para llevar a cabo la calibración del sistema propuesto, es necesario emplear un patrón de calibración que tenga colores amarillo y cian, ya que estos colores son sensibles a la luz y son adecuados para la detección de los colores azul y rojo por parte de la cámara y el proyector, respectivamente. Para este propósito, se utiliza comúnmente un tablero de ajedrez como patrón de calibración.

El tablero físico debe contener el color amarillo, y el proyector debe desplegar el color cian de forma superpuesta para capturar simultáneamente el patrón desde diferentes posiciones como se muestra en la fig 8



Figura 8: Patrón amarillo-cyan.

Durante el proceso de calibración de la cámara y el proyector, se requiere al menos 5 homografías capturadas y los canales rojo y azul de ambos patrones de ajedrez superpuestos. Esto se debe a que los colores blanco y amarillo contienen el componente rojo para la cámara, mientras que los colores cian y verde son generados por el tablero de ajedrez del proyector como se puede observar en la fig 9.

El canal azul de la cámara captura exclusivamente el tablero de ajedrez de referencia. Esto se debe a que

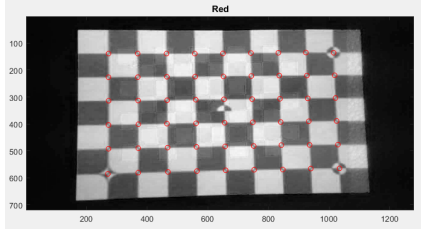


Figura 9: Patrón Red.

los colores blanco y cian tienen el componente azul, mientras que el amarillo y el verde del tablero de ajedrez de referencia iluminado no contienen este componente en la misma proporción, como se muestra en la fig 10.

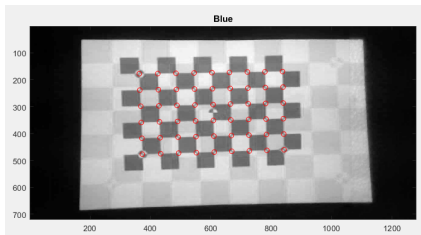


Figura 10: Patrón Blue.

En un escenario ideal, el canal verde debería mostrar un brillo uniforme sin detectar ninguno de los tableros de ajedrez, lo que garantizaría una detección nula de los mismos con un bajo contraste. Sin embargo, debido a la presencia de iluminación ambiental y desequilibrios en el brillo y colores tanto en la cámara como en el proyector, se puede producir una detección no nula, lo que puede afectar la precisión del proceso de calibración.

En el sistema descrito, se emplean 6 homografías con ángulos de rotación de izquierda a derecha y de arriba a abajo, tomando la primera y última captura como referencias para el sistema sin otras posiciones. De esta manera, se obtienen los puntos de las esquinas de cada cuadro respectivamente para el proyector y la cámara, lo que permite obtener los parámetros intrínsecos y extrínsecos del sistema.

Con este enfoque, se pueden obtener un total de 54 puntos que deben ser detectados en todos los canales, es decir, en los canales rojo y azul. Estos puntos son fundamentales para la calibración precisa del sistema y aseguran una correcta correspondencia entre el proyector y la cámara durante la proyección y captura de las imágenes.

El uso de las homografías y la detección de los puntos en ambos canales son pasos cruciales para obtener la información necesaria que permitirá una calibración

adecuada del sistema, garantizando así una proyección y captura precisa y consistente en todo el proceso.

4.3. Proyección y procesamiento

Para realizar la proyección de franjas en el sistema, es necesario tener en cuenta varios factores para generar las fases que luego serán demoduladas. Previamente se deben emplear las formulas $G_x(x,y)$ y $G_y(x,y)$ para producir las franjas por función cosenoidales.

En el presente sistema, se realizó previamente un procesamiento con 7 esferas de color blanco y rojo. Esto permitió ajustar el sistema para realizar una captura posterior del fruto en cuestión. Para lograrlo, se realizaron proyecciones de franjas tanto en el eje X como en el eje Y, obteniendo ambas fases. Aunque este proceso es más lento, es más completo y preciso, como se puede observar en la figura 11:

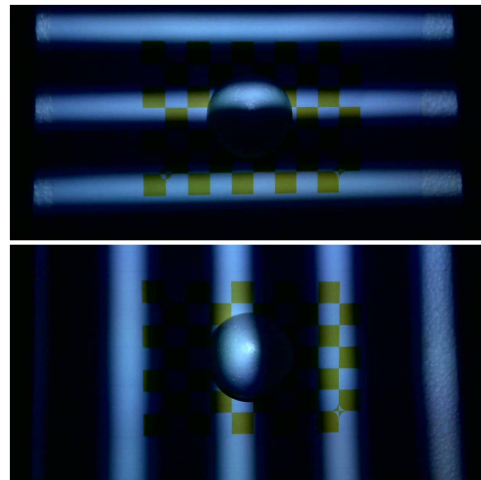


Figura 11: Proyección de franjas con phase X y Y para esfera de 8cm

Por otro lado, se utilizan proyecciones de franjas únicamente para el eje X empleando para capturar el jitomate, lo que proporciona un único parámetro de fase. Luego, se calcula la distancia d_2 utilizando la fase en el eje Y en función de la fase en el eje X mediante la siguiente fórmula:

$$d_2 = R_2 K_2^T \begin{pmatrix} ones \\ zeros \\ -\phi X^T \end{pmatrix}. \quad (35)$$

Con esto, se logra capturar más fotografías del jitomate, de la cual se tomaron al menos 4 perspectivas del mismo utilizando la proyección solo en eje X. Para este propósito, se realizaron 16 desplazamientos de fase con 4 rejillas, cada una con una frecuencia de píxeles de 20 y una variación de 1.0, lo que permite capturar un total de 64 imágenes del objeto, como se observa en la figura 12.

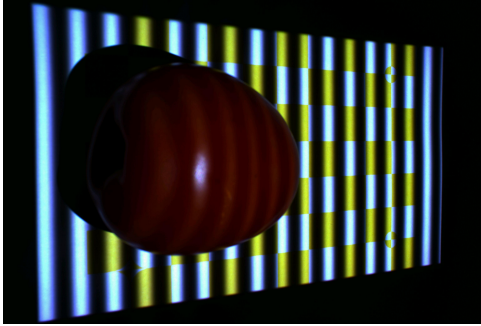


Figura 12: Proyección de franjas sobre jitomate

Posteriormente, se procede a realizar el envolviendo de fase, el cual debe tener en cuenta los parámetros de fase en los ejes X e Y para las 14 esferas procesadas, como se muestra en la Figura 13.

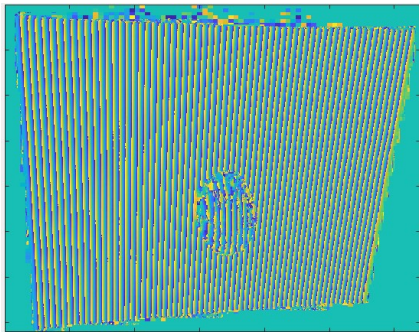


Figura 13: Envolvimiento de fase en Y para esfera roja

Luego, en continuidad con lo anterior, se procede a realizar el desenvolvimiento de fase para los ejes X e Y, obteniendo así los parámetros de fase en X y Y, como se muestra en la Figura 14.

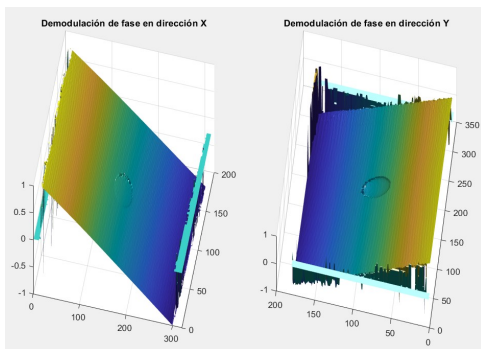


Figura 14: Demodulación en ejes X y Y de esferas blancas y rojas

Consecutivamente, se debe realizar la triangulación de las coordenadas para así reconstruir el objeto, que es la esfera en estudio. Esto se realizó para las 14 muestras

tanto en blanco como en rojo, como se muestra en la Figura 15.

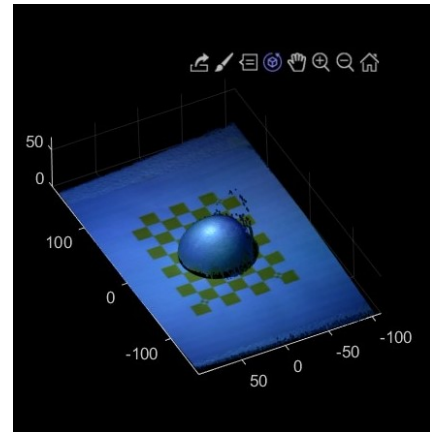


Figura 15: Triangulación de esfera blanca, mismo para rojas

Después de obtener la nube de puntos correspondientes al objeto, se procede a estimar su radio. Para ello, se exporta el archivo resultante y se eliminan todos los puntos no correspondientes al objeto, marcados como "nan.º no definidos". Luego, se utiliza una función en Python o Matlab para seleccionar el punto mínimo y máximo, lo que permite realizar una estimación de su estructura utilizando una serie de Taylor. De esta manera, se calcula iterativamente el promedio del radio hasta obtener un valor estable que representa el radio de la esfera, multiplicándolo por 25 veces para mayor precisión. Figura 16.

Una vez obtenida la estimación del radio de la esfera, se puede representar gráficamente el resultado. Se superpone la esfera calculada con la nube de puntos original para visualizar la efectividad de este método empleado. Esta comparación permitirá observar si el modelo aproximado se ajusta bien a la forma real del objeto, es decir, si la nube de puntos se asemeja a una esfera. La rapidez de este proceso es otra ventaja destacada, ya que se logra un procesamiento rápido y eficiente de los datos, lo que resulta beneficioso para aplicaciones en tiempo real o análisis de grandes conjuntos de datos. Se puede destacar que los resultados calculados por software y funciones programadas en contraste con las mediciones reales de las esferas resultaron con un margen de error de 1mm o menos, por lo que se puede decir que es efectivo.

4.4. Implementación real

De acuerdo a las complicaciones y facilidades de la aplicación experimental en las esferas rojas y blancas, se procede a realizar la aplicación a frutos de forma

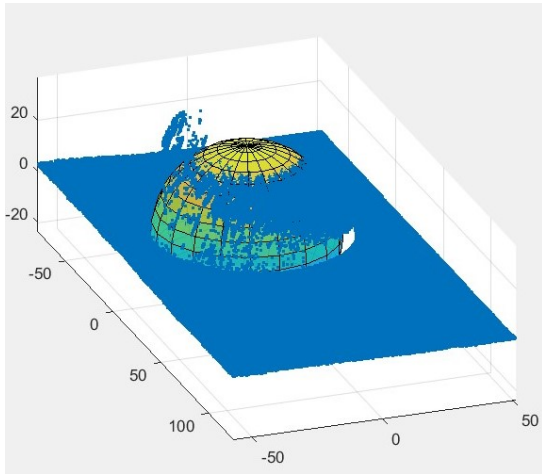


Figura 16: Grafico de esfera estimada con superposicion de nube de puntos en triangulación

ovalada y con superficies distintas, es decir a los jitmates saladette y redondo liso. Para digitalizarlo, se capturan datos de al menos 4 lados diferentes del jitmate, y para cada lado se toman 64 fotografías, lo que resulta en un total de 256 imágenes.

Para reducir la cantidad de capturas requeridas, se ha empleado el método de proyección de franjas en el eje X, como se explicó anteriormente. Esto nos permite reducir el número de capturas a 64 por lado, en lugar de tener que tomar el doble de imágenes para el eje Y, de acuerdo a la figura 17.

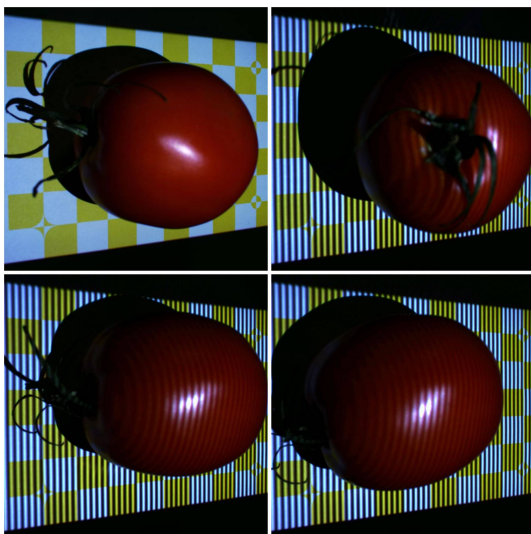


Figura 17: Captura de cada lado del jitmate en solo utilizando franjas en eje X

Cada una de las 64 capturas tomadas por lado debe ser procesada mediante el envolvimiento o modulación de

fase. Este proceso nos permite obtener la información de la forma tridimensional del jitmate, lo que es esencial para el análisis y estudio posterior como se muestra en la figura 18.

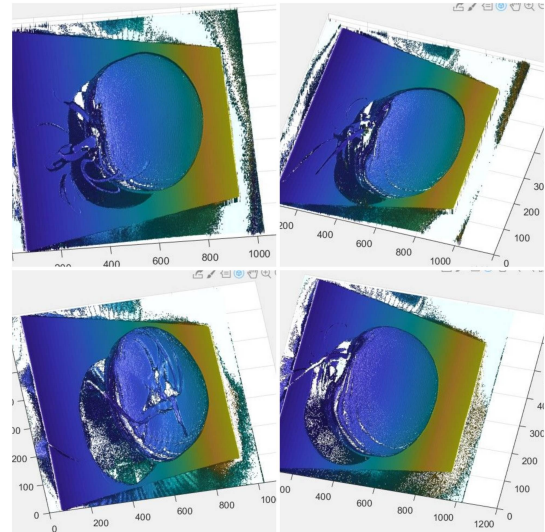


Figura 18: Demodulacion o desenvolvimiento de fase solo en eje X para los 4 lados del jitmate

Continuando con el procesamiento, se procede a realizar la triangulación para cada uno de los lados del jitmate, lo que nos permite obtener la digitalización tridimensional del objeto y de cada uno de sus puntos. El resultado de este proceso se muestra en la siguiente figura 20.

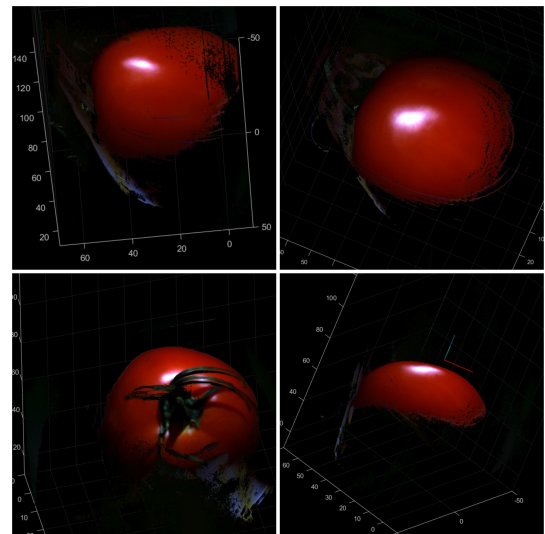


Figura 19: triangulación y vista en 3D para los 4 lados del jitmate.

En contraste con la experimentación de las esferas, ahora se procede a unir cada una de las nubes de puntos obtenidas para el jitmate. Cada nube de puntos se

guarda en formato '.PLY' para poder procesarlas en el software Meshlab y lograr reducir el ruido y pixeles adicoooneles que no oportan una observación para el fruto, por lo que despues de ello se exportan todas las nubes de puntos para seguir con su renderizado en el software Blender. De esta manera, se unen cada cara correspondiente en su posición definitiva, logrando así la reconstrucción tridimensional completa del jitomate. Luego, se exporta el modelo a Meshlab y se guarda la nube de puntos resultante, como se muestra en la figura20.

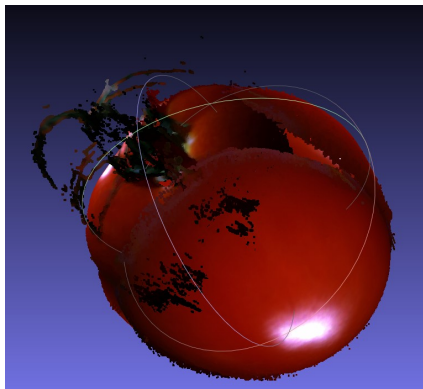


Figura 20: Reconstrucción del jitomate con todas sus posiciones capturadas para su exportación en nube de puntos.”

Para finalizar, debe calcular el radio del fruto para determinar su tamaño del mismo, por lo cual se procesa en la función previamente empleada para las esferas rojas y identificar el resultado, como se muestra en la figura 21.

5. Discusión y Resultados

Los resultados que se obtuvieron mediante el uso de la proyección de luz estructurada mediante franjas, se logró obtener una estimación apropiada y precisa respecto a las dimensiones de las esferas y de los jitomates. Los datos recopilados revelaron un margen de error mínimo, que oscila entre uno y dos milímetros en el caso de las esferas, y entre dos y cuatro milímetros en el caso de los jitomates. Estos resultados respaldan con firmeza la viabilidad y eficacia de este método en particular.

En cuanto al resultado obtenido al implementar el método en el jitomate redondo liso, se obtuvo un radio estimado de 36.31 mm, lo que indica que el diámetro del objeto es de aproximadamente 72.62 mm.

Sin embargo, tras comparar los resultados, se observó que el tamaño real del jitomate es de 68.9 mm, mientras que el obtenido fue bastante cercano, como se muestra

Esfera	Color	Tamaño Real (Radio)	Tamaño Matlab (Radio)
1	Blanco	39.0 mm	41.0 mm
2	Blanco	36.0 mm	37.0 mm
3	Blanco	32.5 mm	34.5 mm
4	Blanco	25.2 mm	27.2 mm
5	Blanco	20.5 mm	22.1 mm
6	Blanco	15.5 mm	15.1 mm
7	Blanco	12.7 mm	13.0 mm

Cuadro 1: Resultados obtenidos y contrastados de las esferas blancas.

Esfera	Color	Tamaño Real (Radio)	Tamaño Matlab (Radio)
1	Rojo	39.0 mm	44.6 mm
2	Rojo	36.0 mm	40.4 mm
3	Rojo	32.5 mm	38.3 mm
4	Rojo	25.2 mm	32.8 mm
5	Rojo	20.5 mm	28.9 mm
6	Rojo	15.5 mm	16.7 mm
7	Rojo	12.7 mm	13.4 mm

Cuadro 2: Resultados obtenidos y contrastados de las esferas rojas.

en la figura21.

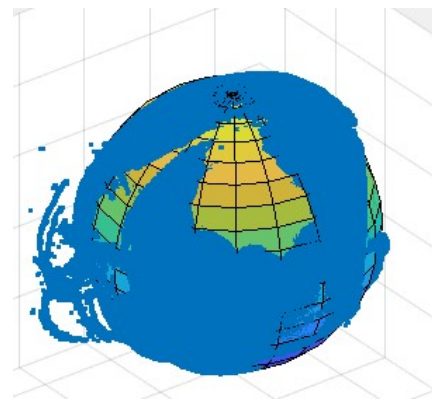


Figura 21: Ajuste y superposición de la nube de puntos del jitomate para estimación del radio.”

6. Conclusiones

En conclusión, la técnica de proyección de luz estructurada a través de franjas ha demostrado ser altamente efectiva para la estimación precisa de las dimensiones de esferas y jitomates. Nuestros resultados mostraron un margen de error mínimo, con fluctuaciones de solo uno a dos milímetros en el caso de las esferas y entre dos y cuatro milímetros en el caso de los jitomates.

En comparación con otras alternativas existentes en el mercado para la clasificación del jitomate, nuestro

enfoque se destaca por su simplicidad y economía. Mientras que otras soluciones requieren recursos más amplios, como redes neuronales convolucionales y bases de datos extensas, nuestra técnica basada en proyección de franjas ofrece una solución más accesible.

La relevancia del jitomate en la dieta básica del mexicano promedio y su uso en diversas recetas y platillos tradicionales hacen que contar con una solución local y accesible para su clasificación sea de gran importancia para el sector agrícola y alimentario de nuestro país.

Nuestros resultados validan la viabilidad y eficacia de este enfoque particular. El siguiente paso del proyecto consistirá en la construcción de un prototipo que albergue un mini-proyector y una cámara Raspberry Pi 4, ofreciendo una solución prometedora para ser implementada en invernaderos de agricultores con recursos limitados, lo que representa una alternativa práctica a los sistemas más costosos empleados por grandes empresas.

En resumen, este estudio demuestra que la proyección de franjas es una técnica viable y efectiva para la estimación de dimensiones de esferas y jitomates, con potencial para su aplicación en el sector agrícola y alimentario. La investigación futura y el desarrollo del prototipo permitirán seguir avanzando en esta prometedora dirección y brindar una herramienta valiosa para los agricultores mexicanos.

7. Fuentes Bibliográficas

1. Máquina para clasificar tomate — Aweta. (s. f.). <https://www.aweta.com/es/productos/tomate>.
2. Asa, T. S. (s. f.). Máquina clasificadora de alimentos por colores. Tomra. <https://language.sites.tomra.com/es/es/sorting/alimentacion/equipamiento-de-clasificacion/nfm>.
3. Multiscan Technologies en el sector del tomate — Multiscan. (2022, 23 junio). Multiscan Technologies. <https://multiscan.eu/aplicacion-tomate/:.text=Multiscan>
4. R. Juarez-Salazar, “How do phase-shifting algorithms work?,” *Eur. J. Phys*, vol. 39, p. 065302, 2018.
5. R. Juarez-Salazar, C. Robledo-Sanchez, F. Guerrero-Sanchez, and A. Rangel-Huerta, “Generalized phase-shifting algorithm for inhomogeneous phase shift and spatiotemporal fringe visibility variation,” *Opt. Express*, vol. 22, pp. 4738–4750, Feb 2014.
6. Juarez-Salazar, R., Robledo-Sanchez, C., Guerrero-Sanchez, F., Barcelata-Pinzón, A., González-García, J., Santiago-Alvarado, A. (2016). Intensity normalization of additive and multiplicative spatially multiplexed patterns with N encoded phases. *Optics and Lasers in Engineering*. <https://doi.org/10.1016/j.optlaseng.2015.09.001>
7. Rigoberto, J. S. (2021, 24 junio). Perfilometría óptica de superficies discontinuas. <https://tesis.ipn.mx/handle/123456789/28991>
8. A. Zhang. “Flexible New Technique for Camera Calibration”; 1999 <http://research.microsoft.com/zhang>.