

# Cuadernos de Educación y Desarrollo

Vol 1, Nº 2 (abril 2009)

<http://www.eumed.net/rev/ced/index.htm>

---

## VISIÓN TRIDIMENSIONAL BASADA EN METROLOGÍA LÁSER Y ALGORITMOS DE INTELIGENCIA ARTIFICIAL

J. Apolinar Muñoz Rodríguez  
Centro de Investigaciones en Óptica A.C.  
México

Edgar Fernando Velázquez Pedroza.  
Universidad de Guadalajara, Centro universitario De Los Lagos  
México

Diego Alonso Castañeda Aguilar  
Instituto Tecnológico de Celaya  
México.

### Resumen

Se presenta una técnica de visión tridimensional basada en metrología láser y algoritmos de inteligencia artificial. En esta técnica, la visión se efectúa mediante proyección de una línea láser y procesamiento de imágenes. Mediante un arreglo óptico móvil, la línea láser barre la superficie bajo estudio. En base al comportamiento de la línea láser, una red de aproximación produce un modelo matemático para calcular la topografía del objeto. La arquitectura de la red se construye con imágenes de línea láser proyectada sobre objetos con dimensiones conocidas. El enfoque de esta red es evitar mediciones sobre el arreglo óptico. Además, esta red calcula los parámetros de visión en forma automática. Así, se mejora la precisión de los resultados y el manejo del sistema de visión. Esto se debe a que errores de medición no se pasan al sistema de reconstrucción tridimensional. Para describir la precisión de los resultados, se calcula el error cuadrático medio usando un método de contacto como referencia. Esta técnica es probada con objetos reales y se presentan resultados experimentales.

### 1. Introducción

En metrología óptica y visión por computadora se han desarrollado diversas técnicas para digitalización de objetos. La metrología es una herramienta muy valiosa en visión artificial, particularmente en reconstrucción de objetos. El uso de luz estructurada, permite que el sistema de visión sea confiable y los datos adquiridos son muy fáciles de interpretar. Así mismo, la incorporación de redes neuronales a los métodos ópticos permite la automatización del proceso de reconstrucción. Una técnica en particular para reconstrucción de objetos es la de proyección de una línea láser [1-3].

En esta técnica, se proyecta una línea de luz sobre la superficie del objeto. La línea se deforma en el plano de la imagen debido a la variación de la superficie del objeto. Con la posición de la línea y los parámetros de la geometría del arreglo óptico, se deduce la información de la topografía del objeto. En esta técnica, las distancias de la geometría del arreglo óptico se miden mediante un proceso externo al sistema de reconstrucción. Posteriormente estos parámetros son usados por el sistema computacional para calcular las dimensiones de la superficie del objeto.

En la técnica propuesta, la forma del objeto se obtiene detectando la línea láser mediante procesamiento de imágenes y redes de aproximación. En este caso, la posición de la línea láser es procesada por la red para obtener la superficie del objeto. Con la aplicación de la red se evitan mediciones de los parámetros de la geometría del arreglo óptico. La arquitectura de esta red neuronal se construye con imágenes de la línea láser proyectada sobre patrones con dimensiones conocidas. En el arreglo experimental, el objeto se desplaza en el eje- $x$  con lo que se barre la superficie de este. En cada paso del movimiento, se captura la línea láser mediante una cámara CCD. Mediante aproximación Gaussiana, se calcula la posición de la línea láser. Posteriormente, la red neuronal calcula la dimensión del objeto en base a la posición de la línea láser. La información producida por la red se guarda en un arreglo de memoria para obtener la forma del objeto. La viabilidad del método es examinado en base a el error cuadrático medio. Este análisis se efectúa tomando como referencia un método de contacto.

## 2. Descripción del arreglo óptico

La configuración del arreglo óptico de ésta técnica incluye: un dispositivo electromecánico, un proyector de línea láser, una cámara CCD y una computadora. El arreglo experimental se muestra en la Figura 1. En este arreglo, el objeto se fija sobre una plataforma, la cual se mueve en el eje  $x$ . Este dispositivo está controlado por computadora y su movimiento está basado en un algoritmo computacional. Sobre el objeto se proyecta una línea vertical, la cual barre la superficie de éste. En cada paso del movimiento, la línea cambia de posición en el plano de la imagen de acuerdo a la topografía del objeto. Cada línea se captura mediante la cámara CCD. La topografía del objeto se obtiene mediante el procesamiento de la línea y la red de aproximación. La información producida por cada línea corresponde a una sección transversal del objeto. La forma completa del objeto se genera con la información de todas las líneas procesadas.

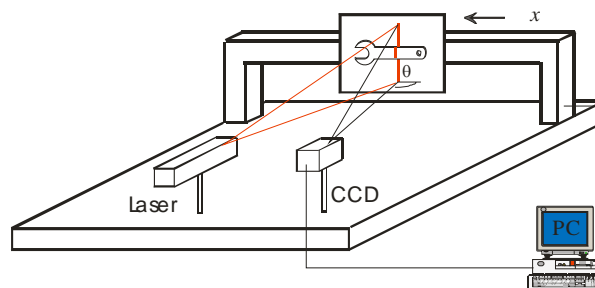


Figura 1. Arreglo experimental.

Para describir la relación entre la posición de la línea láser y la superficie del objeto, se usa la geometría del arreglo óptico mostrado en la Figura 2. En el plano de referencia están localizados el eje  $y$ , el eje  $x$  y la superficie del objeto se indica como  $h(x, y)$ . Los puntos  $A$  y  $B$  corresponden a la línea láser proyectada sobre en el plano de referencia y sobre el objeto, respectivamente. Cuando una línea láser se proyecta sobre el objeto, esta se mueve en el plano de la imagen del punto  $x_A$  al  $x_B$ . Este desplazamiento se representado como:

$$s(x,y) = x_A - x_B \quad (1)$$

El desplazamiento  $s(x,y)$  es directamente proporcional a la superficie del objeto  $h(x,y)$  [4]. Usando este desplazamiento, la superficie del objeto se determina mediante la red de aproximación. De esta manera, se obtiene la información de una sección transversal del objeto. El desplazamiento  $s(x,y)$  se obtiene midiendo la posición de la línea de luz en cada renglón de la imagen. Esta posición es obtenida con una resolución de fracción de píxel mediante el método de Aproximación Gaussiana.

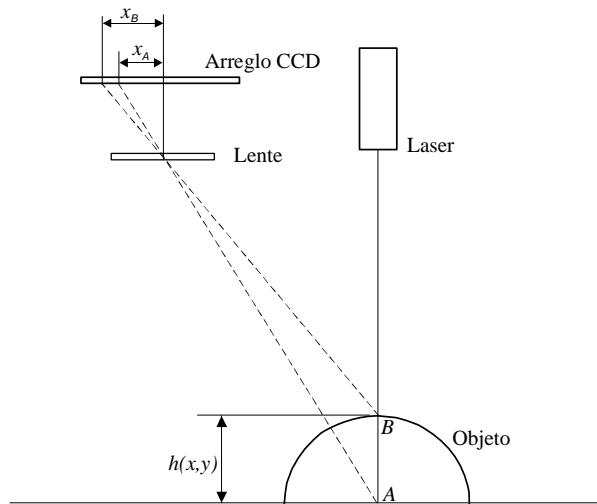


Figura 2. Geometría del arreglo experimental.

La intensidad proyectada por un diodo láser es una distribución *Gaussiana* en dirección del eje  $x$  [5]. Los valores de intensidad de la línea láser son representados como  $(x_0, z_0), (x_1, z_1), (x_2, z_2), \dots, (x_n, z_n)$ , donde  $x_i$  es la posición del píxel y  $z_i$  es la intensidad del píxel. Una manera de representar estos valores es una función Gaussiana, la cual se define como:

$$f(x) = \frac{Ni}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (2)$$

Donde  $Ni$  es el área bajo la curva,  $\mu$  es la media de la función, y  $\sigma$  es la desviación estándar. Para determinar la posición de la línea láser se calcula la media  $\mu$ , la cual representa el centro de la función Gaussiana. La expresión para calcular la media puede describirse como:

$$\mu = \frac{\sum_{i=1}^n z_i x_i}{\sum_{i=1}^n z_i} \quad (3)$$

La Figura 3 muestra un conjunto de píxeles de un renglón de una línea láser. De estos píxeles se calcula la posición central  $\mu$  para determinar la posición de la línea. Para efectuar este procedimiento, se sustituye la posición  $x_i$  y la intensidad del píxel  $z_i$  en la Eq. (3). En este caso, el resultado es  $\mu=14.664$ . Por consiguiente, la posición de la línea de luz es  $x_B=14.664$  píxeles. También, el factor  $Ni$  y la desviación estándar son calculados para ajustar la función Gaussiana mostrada en la Figura 3.

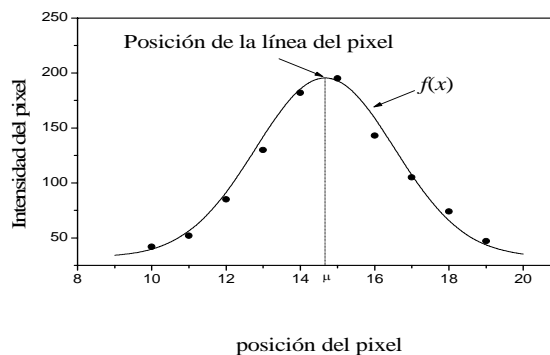


Figura 3. Píxeles aproximados a una función de Gaussiana.

El valor de la media se usa en la Ec.(1) para obtener el desplazamiento  $s(x,y)$ . El procedimiento de calcular la posición de la línea, se aplica a todos los renglones de la imagen. Con esto, se obtiene el

perfil del objeto de la región donde se proyectó la línea láser. Este perfil corresponde a los desplazamientos  $s_1, s_2, s_3, \dots, s_n$ , que se procesarán por la red para obtener una sección transversal del objeto. En la Figura 4, se muestra una línea láser.

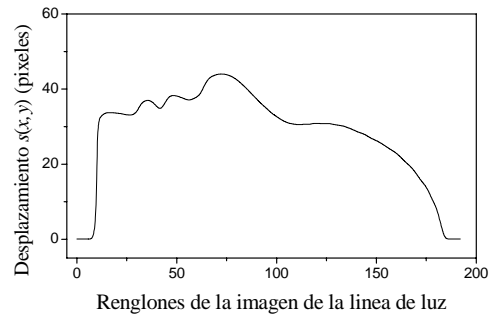


Figura 4. Línea láser proyectada sobre un objeto.

Figura 5. Perfil del objeto de la línea láser de la Fig.4

De cada renglón de esta imagen se calcula la media  $\mu$  mediante la Ec.(3) para obtener la posición de la línea láser a lo largo de esta. Usando la posición de la línea se calcula el desplazamiento  $s(x,y)$  mediante la Ec.(1). Este desplazamiento representa el perfil del objeto, el cual se muestra en la Figura 5. Este perfil es extraído de cada una de las imágenes capturadas en el movimiento del objeto.

### 3. Estructura de la red de aproximación

Las funciones de Bezier se han usado para representar curvas y superficies [6]. La red Bezier propuesta consta de un vector de entrada, una entrada paramétrica, una capa oculta y una capa de salida. La estructura de la red se muestra en la Figura 6. Los datos de entrada  $h_1, h_2, h_3, \dots, h_n$  son las dimensiones de objetos conocidos y los valores de entrada  $s_1, s_2, s_3, \dots, s_n$  son los desplazamiento de la línea. La entrada paramétrica  $u$  es un valor en proporcional al desplazamiento  $s_i$ . La capa oculta se construye con funciones de base Bezier. La capa de salida se forma con la sumatoria de las neuronas de la capa oculta, las cuales son multiplicadas por un peso. Cada capa de la estructura de la red se obtiene de la siguiente forma. El desplazamiento de la línea  $s_1, s_2, s_3, \dots, s_n$ , se obtiene mediante procesamiento de imágenes. La relación entre el desplazamiento y el valor  $u$  es representado por

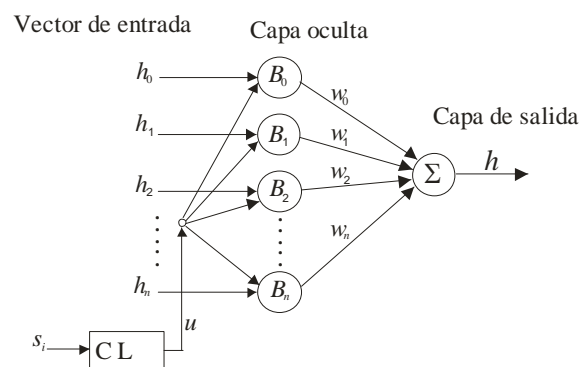


Figura 6. Estructura de la red Bezier

La relación entre el desplazamiento y el valor paramétrico se describe mediante la siguiente expresión:

$$u = a_0 + a_1 s, \tag{4}$$

donde  $a_i$  son constantes a determinar y  $s$  es la posición del desplazamiento de línea. Usando  $s$  y su posición  $u$ , la ecuación (4) se puede determinar. Las funciones Bezier están definidas en el intervalo  $0 \leq u \leq 1$  [7]. Por lo tanto  $u = 0$  para el primer desplazamiento  $s_0$  y  $u=1$  para el último desplazamiento  $s_n$ . Sustituyendo los valores  $(s_0, u=0)$  y  $(s_n, u=1)$  en la ecuación (4), se obtienen dos ecuaciones con

dos incógnitas. Resolviendo estas ecuaciones, se obtiene  $a_0$  y  $a_1$ . Los datos  $h_1, h_2, h_3, \dots, h_n$ , se obtienen mediante un método de contacto. Las neuronas de la capa oculta se construyen mediante funciones de base Bezier, las cuales se describen como:

$$B_i(u) = \binom{n}{i} u^i (1-u)^{n-i}, \quad 0 \leq u \leq 1, \quad \binom{n}{i} = \frac{n!}{i!(n-i)!}, \quad (5)$$

La capa de salida se determina mediante la sumatoria de las neuronas, las cuales son multiplicadas por un peso. La respuesta de salida de la red es representada por

$$h(u) = \sum_{i=0}^n w_i B_i(u) h_i, \quad 0 \leq u \leq 1, \quad (6)$$

donde  $w_i$  son los pesos,  $h_i$  es la dimensión del objeto y  $B_i$  es la función de base Bezier Ec.(6). Para calcular los pesos  $w_i$  se implementa un mecanismo de ajuste. Para efectuar este proceso,  $u_i$  y  $h_i$  se sustituyen en la Ec.(6). En esta forma se obtiene el sistema de ecuaciones:

$$\begin{aligned} h_0 &= w_0 \binom{n}{0} (1-u)^n u^0 h_0 + w_1 \binom{n}{1} (1-u)^{n-1} u h_1 + \dots + w_n \binom{n}{n} (1-u)^0 u^n h_n, \quad 0 \leq u \leq 1. \\ h_1 &= w_0 \binom{n}{0} (1-u)^n u^0 h_0 + w_1 \binom{n}{1} (1-u)^{n-1} u h_1 + \dots + w_n \binom{n}{n} (1-u)^0 u^n h_n, \quad 0 \leq u \leq 1. \\ &\vdots \\ h_n &= w_0 \binom{n}{0} (1-u)^n u^0 h_0 + w_1 \binom{n}{1} (1-u)^{n-1} u h_1 + \dots + w_n \binom{n}{n} (1-u)^0 u^n h_n, \quad 0 \leq u \leq 1. \end{aligned} \quad (7)$$

Este sistema de ecuaciones se puede representar como

$$\begin{aligned} h_1 &= w_1 \beta_{1,1} + w_2 \beta_{1,2} + \dots + w_n \beta_{1,n} \\ h_2 &= w_1 \beta_{2,1} + w_2 \beta_{2,2} + \dots + w_n \beta_{2,n} \\ &\vdots \\ h_n &= w_1 \beta_{n,1} + w_2 \beta_{n,2} + \dots + w_n \beta_{n,n} \end{aligned} \quad (8)$$

Este sistema de ecuaciones a su vez se puede escribir en forma de matriz como  $\beta W = H$ :

$$\begin{bmatrix} \beta_{1,1} & \beta_{1,2} & \beta_{1,3} & \dots & \beta_{1,n} \\ \beta_{2,1} & \beta_{2,2} & \beta_{2,3} & \dots & \beta_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \beta_{n,1} & \beta_{n,2} & \beta_{n,3} & \dots & \beta_{n,n} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} \quad (9)$$

Resolviendo el sistema de ecuaciones (9), se obtienen los pesos  $w_i$ . En esta forma la red Bezier ha sido completada. De esta manera para un  $s_i$  la red calcula una dimensión  $h_i$ . El resultado de esta red es una función que calcula la dimensión del objeto en base al desplazamiento de línea láser.

#### 4. Parámetros de visión

En visión tridimensional, las dimensiones de un objeto se determinan en base a los parámetros de la cámara y la geometría del arreglo experimental. Los parámetros de la cámara incluyen: distancia focal, orientación, centro de la imagen, factor de escala y distorsión. En la técnica propuesta, los

parámetros de la cámara se calculan mediante la red y procesamiento de imágenes. La orientación de la cámara se determina en base al eje óptico. En este caso, el eje óptico es perpendicular al plano de referencia como se muestra en la Figura 7. La distancia entre el centro de la imagen y la línea láser se denota mediante  $\ell_a$ . La dimensión del objeto se indica mediante  $h_i$  y  $D = z_i + h_i$ . De acuerdo a al eje perpendicular, la dimensión  $h_i$  del objeto tiene una proyección  $k_i$  en el plano de referencia. De la Figura 7, el desplazamiento se define como  $s_{ij} = (x_c - x_B) - (x_c - x_A)$ . Así, la proyección  $k_i$  en el plano de referencia se calcula mediante la siguiente expresión.

$$k_i = \frac{F h_i}{s_{ij} + x_c - x_A} \quad (10)$$

De la ecuación 10,  $F$ ,  $x_c$ ,  $x_A$  son constantes y  $h_i$  es calculada por la red. En este caso,  $k_i$  es una función lineal y la derivada  $dk/ds$  es una constante. Debido a la distorsión la derivada  $dk/ds$  no exactamente una contante. Pero la derivada es lo más aproximado a una constante.

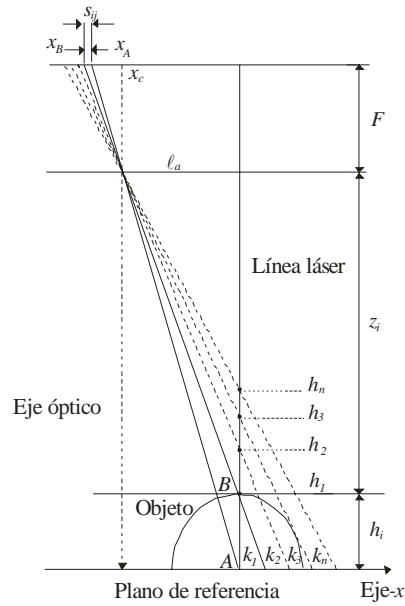


Figura 7. Geometría de un eje óptico perpendicular al plano de referencia.

En base a un eje perpendicular, se deducen los parámetros de la cámara. Para efectuar el proceso, la red produce la dimensión del objeto  $h_i$ . La geometría de la Figura 7, se describe mediante la siguiente expresión

$$\frac{z_i}{\ell_a} = \frac{z_i + F}{\eta(x_c - x_B) + \ell_a}, \quad (11)$$

De esta ecuación,  $\eta$  es el factor de escala que convierte los pixeles en milímetros. Usando  $D = z_i + h_i$  y  $\eta s_{ij} = \eta(x_c - x_B) - \eta(x_c - x_A)$ , la ecuación (11) se reescribe como

$$\frac{D - h_i}{\ell_a} = \frac{D - h_i + F}{\eta(s_{ij} + x_c - x_A) + \ell_a} \quad (12)$$

Donde  $D$  es la distancia de la lente al plano de referencia. De la ecuación (12), las constantes  $D$ ,  $\ell_a$ ,  $F$ ,  $\eta$  y  $x_c$  deben ser determinadas. Esto se realiza reescribiendo la ecuación (12) como el siguiente sistema de ecuaciones

$$\begin{aligned}
h_0 &= D - \frac{F \ell_a}{\eta(s_{00} + x_c - x_A)} \\
h_1 &= D - \frac{F \ell_a}{\eta(s_{01} + x_c - x_A)} \\
h_2 &= D - \frac{F \ell_a}{\eta(s_{02} + x_c - x_A)} \\
h_3 &= D - \frac{F \ell_a}{\eta(s_{03} + x_c - x_A)} \\
h_4 &= D - \frac{F \ell_a}{\eta(s_{04} + x_c - x_A)} \\
h_5 &= D - \frac{F \ell_a}{\eta(s_{05} + x_c - x_A)}
\end{aligned} \tag{13}$$

Los valores  $h_1, h_2, \dots, h_5$ , son calculados mediante la red y se sustituyen en la ecuación (13). Resolviendo este sistema de ecuaciones, se obtienen las constantes  $D, \ell_a, F, \eta$ , and  $x_c$ . De esta manera, se obtienen los parámetros del sistema de visión en base a la red y procesamiento de imágenes. La distorsión del sistema se puede observar mediante la posición de la línea de luz en el plano de la imagen, la cual se describe mediante la siguiente expresión

$$x_B = \frac{F \ell_a}{D - h_i} + x_c \tag{14}$$

En base a esta ecuación, el comportamiento de  $x_B$  con respecto a  $h_i$  debería ser una función lineal. Sin embargo, debido a la distorsión  $x_B$  no es una función lineal. La red se construye mediante datos que se obtienen del desplazamiento  $s_i = (x_c - x_B) - (x_c - x_A)$ . Por lo tanto, la red produce datos  $h_i$  no lineales. Así, la distorsión es incluida en la red, la cual calcula la dimensión del objeto. Por medio de la información producida por la red de todas las imágenes, se reconstruye la forma del objeto.

## 5. Resultados experimentales

La Figura 1 muestra el arreglo experimental. El objeto se mueve en pasos de  $1.27 \text{ mm}$  a lo largo del eje  $x$ . Una línea láser vertical se proyecta sobre la superficie del objeto por medio de un diodo de láser de  $5 \text{ mW}$ . La línea láser se deforma de acuerdo a la topografía del objeto en el proceso de barrido. Estas líneas deformadas son capturadas por medio de una cámara CCD y se digitalizan con una resolución de  $320 \times 200$  píxeles y 256 niveles del gris. El experimento es probado con dos objetos. El primer objeto es un maniquí, el cual se muestra en la Figura 8. Este objeto es barrido por la línea láser para obtener la información de la superficie. El resultado de este paso son los datos  $s_1, s_2, \dots, s_m$ , los cuales contienen los desplazamientos de cada renglón de la imagen de la línea. Estos datos son procesados por la red vía Ec.(8) para obtener los datos de la altura  $h(s)$  de la superficie.



Figura 8. Cara de maniquí.

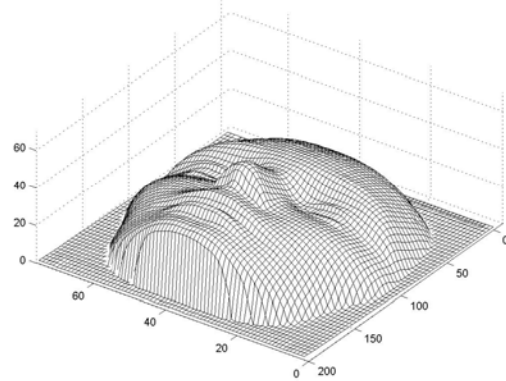


Figura 9. Forma 3D de la cara de maniquí.

Con el procesamiento de cada línea, se obtiene una sección transversal del objeto. Los datos de cada sección transversal se guardan en un arreglo de la memoria para construir la forma 3D completa del objeto. Para conocer la exactitud de los resultados obtenidos, se calcula *la raíz media cuadrática (rms)* del error [8]. El valor *rms* se describe como

$$rms = \sqrt{\frac{1}{m} \sum_{i=1}^m (h_{oi} - hc_i)^2} , \quad (15)$$

donde  $h_{oi}$  son los datos obtenidos por un método de contacto,  $hc_i$  son datos calculados por la red Bezier y  $m$  es el número de datos analizados. El método de contacto consiste en una máquina de medición por coordenadas (MMC). El error obtenido para el maniquí mediante la red es un  $rms = 0.161 \text{ mm}$ . Se procesaron setenta y dos líneas mediante la red Bezier para obtener la forma del maniquí, la cual muestra en la Figura 9. En esta Figura, la escala de los ejes está en  $mm$ .

El segundo objeto utilizado es una horma, la cual se muestra en la Figura 10. El procedimiento aplicado al maniquí se aplica a la horma. Con esto se obtiene la vista frontal de la horma, la cual se muestra en la Figura 11. En esta Figura la escala de los ejes esta en  $mm$ . En este caso, el error es  $rms = 0.173 \text{ mm}$ .

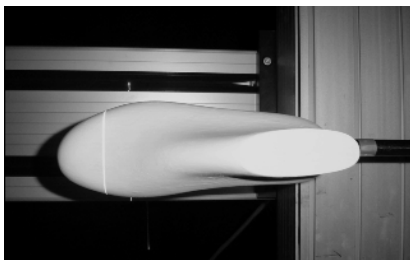


Figura 10. Superficie de la horma.

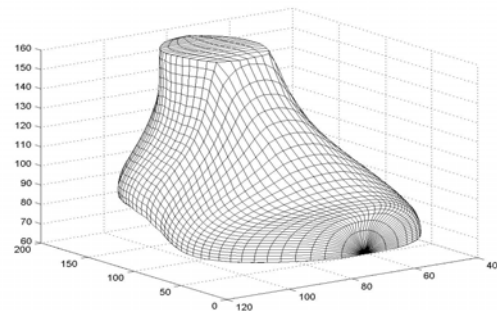


Figura 11. Reconstrucción del la horma.

La computadora empleada en este experimento es una PC a 1 Ghz. Cada imagen se procesa en un tiempo de 0.023 seg. Este tiempo de proceso es mínimo ya que los datos de la imagen se extraen con pocas operaciones via Eq. (3). Usando la red Bezier en esta técnica, la forma del objeto es determinada sin el cálculo de las distancias de la geometría del arreglo óptico. Por consiguiente, el procedimiento es más fácil que las técnicas que usan las distancias de los componentes del arreglo óptico. En esta técnica todos los pasos se efectúan por proceso computacional y se evitando así mediciones sobre el arreglo experimental. De esta manera se logra buena repetitibilidad en cada medición.



## 6. Conclusiones

Se ha presentado una técnica para detección de forma mediante una línea láser y una red Bezier. Esta técnica representa una valiosa herramienta para la detección de forma de objetos en inspección industrial. El método perimétrico de la técnica es automático y evita mediciones de los componentes del arreglo experimental, como es común en los métodos de proyección de una línea láser. Estos parámetros son obtenidos por proceso computacional. Esto mejora la exactitud de los resultados obtenidos por la técnica, ya que los errores de mediciones físicas no se introducen en el sistema. En este trabajo, la habilidad de medir la posición de la línea de luz con resolución de subpixel se ha logrado de una manera muy rápida mediante aproximación de Gaussiana. Esto se logra ya que la posición de la línea de luz se determina con pocas operaciones. Usando este arreglo experimental se obtiene una buena repetibilidad en cada medición, por consiguiente esta técnica es efectuada de buena manera.

## Referencias

- [1] L. Zagorchev and A. Goshtasby, "A paintbrush laser range scanner", *Computer vision and image understanding*, Vol. 10, p. 65-86 (2006).
- [2] Z. Wei, G. Zhang and Y. Xu, "Calibration approach for structured-light-stripe vision sensor based on invariance of double cross-ratio", *Opt. Eng.* Vol. 42 No. 10, p. 2956-2966 (2003).
- [3] A. M. McIvor, "Nonlinear calibration of a laser profiler", *Opt. Eng.* Vol. 42 No.1, p. 205-212, (2002).
- [4] J. A. Muñoz Rodríguez and R. Rodríguez-Vera, "Evaluation of the light line displacement location for object shape detection", *Journal of Modern Optics*, Vol. 50 No.1, p. 137-154 (2003).
- [5] F. Causa and J. Sarma, "*Realistic model for the output beam profile of stripe and tapered superluminescent light-emitting diodes,*" *Appl. Optics*, Vol. 42 No.21, p. 4341-4348, (2003).
- [6] Y. J. Ahn, Y. S. Kim and Y. Shin, "Approximation of circular arcs and other curves by Bezier curves of high degree", *Journal of Computational and Applied Mathematics*, Vol. 167, p. 405 – 416, (2004).
- [7] M. E. Mortenson, *Geometric Modeling*, Wiley, Second edition, U.S.A. (1997),
- [8] T. Masters, *Practical Neural Networks Recipes in C++*, Academic Press, U.S.A 1993.