

Aplicación de redes neuronales de aproximación a una línea de luz para reconstrucción 3D de objetos

J.A. Muñoz Rodríguez^a, A. Asundi^b y R. Rodríguez Vera^a

^aCentro de Investigaciones en Óptica A.C.

Apartado Postal 1- 948, León, GTO, 37000, México

^bNanyang Technological University, School of Mechanical & Production engineering

Nanyang avenue, Singapore 639798, Tel: (65) 6790-5936

E-mail: munoz@foton.cio.mx

Recibido el 1 de julio de 2003; aceptado el 10 de diciembre de 2003

Se presenta una técnica para extraer la forma 3D de objetos mediante el procesamiento de imágenes de una línea de luz. En este proceso, una red neuronal de aproximación reconstruye la forma 3D del objeto. Esta red neuronal, se genera con una estructura en forma de capas, neuronas, conexiones y reciben información para ser procesada, con lo que se obtiene una respuesta de salida. En este caso, la información que recibe la estructura de la red neuronal, corresponde a imágenes de una línea de luz proyectada sobre objetos con dimensiones conocidas. Estas imágenes se obtienen durante el barrido de la línea de luz sobre el objetos patrón. El método perfilométrico que usa la red neuronal de aproximación, está basado en las deformaciones que sufre una línea de luz cuando ésta se proyecta sobre un objeto. Estas deformaciones son medidas mediante aproximación gaussiana. En esta técnica se reconstruye la forma 3D sin utilizar los parámetros del arreglo experimental. Esto constituye una ventaja sobre los métodos comunes de proyección de línea. En esta forma la precisión de los resultados se mejora, ya que no se introducen errores de medición al sistema. La precisión de esta técnica se obtiene mediante el valor *rms*. Esta técnica es probada con simulaciones y con objetos reales. También se presenta el tiempo de procesamiento y la precisión de los resultados.

Descriptor: Reconstrucción 3D; proyección de línea de luz; red neuronal de aproximación; aproximación gaussiana

A technique for 3D object shape detection based on light line image processing is presented. In this process, an approximation neural network is used to reconstruct the 3D object shape. This neural network is generated using images of a light line projected onto the objects, whose dimensions are known. These images are obtained in the scanning step of the light line onto the objects. The profilometric method used by the neural network is based on the light line deformations. These deformations are measured by the Gaussian approximation method. In this technique, the 3D shape is obtained without use the parameters of the experimental set-up. It is an advantage over conventional methods of the light line projection. In this manner, the accuracy is improved due to the errors are not introduces in the system. The accuracy in this technique is deduced by the rms value. This technique is tested with simulations and real objects. Also, the time processing and accuracy results are presented.

Keywords: 3D reconstruction; light line projection; approximation neural network; gaussian approximation.

PACS: 84.35; 42.62E; 42.79

1. Introducción

En metrología óptica y visión robótica se han desarrollado diversas técnicas para la reconstrucción 3D de objetos, usando métodos ópticos. El uso de luz estructurada permite que el sistema sea más confiable y los datos adquiridos se interpretan en forma muy simple para obtener el resultado final. Una técnica en particular es la de proyección de una línea de luz [1-7]. Cuando una línea de luz es proyectada sobre la superficie de un objeto, ésta se deforma en el plano de la imagen. Esta deformación se debe a la variación de la superficie del objeto. El objetivo de esta técnica es la detección de la posición de estas deformaciones en la imagen. Con la medición de estas deformaciones y los parámetros de la geometría del arreglo óptico se deduce la información de la topografía del objeto. En algunos casos, estas deformaciones se miden usando el método de detección de pico [1]. En este método, el pico es asignado al *pixel* con mayor intensidad. Este pico y su centro se obtienen usando una ventana sobre la región de la línea y aplicando un valor de umbral. Otro método consiste en medir el grueso de una línea de luz como una función

de la superficie del objeto [2]. Este grueso es medido directamente sobre la pantalla usando un ocular. También se ha utilizado la máquina de medición de coordenadas (MMC) [3]. Esta técnica usa bloques patrón como calibrador. Las coordenadas espaciales de la deformación de la línea se detectan determinando la posición de la línea sobre la gradación de los bloques patrón. Otro método de detección de pico usa un sensor electrónico para la detección del pico, por medio de un integrador numérico [4]. En el método de integración de líneas (*TDI: Time Delay Integration*), una cámara graba línea por línea un conjunto de imágenes para generar un patrón de franjas [5-7]. En este método, la forma del objeto se obtiene calculando la fase del patrón de franjas. En todos estos métodos las distancias de la geometría del arreglo óptico son medidos mediante un proceso externo al sistema computacional. Posteriormente estos parámetros son introducidos al sistema computacional para poder obtener las dimensiones de la superficie 3D del objeto.

En la técnica propuesta en este trabajo, la forma 3D del objeto se obtiene usando únicamente el proceso de detección

de la deformación de la línea de luz. Esta deformación tiene una relación geométrica bien definida con la superficie del objeto [8]. Para evitar la medición de los parámetros de la geometría del arreglo óptico, se construye una red neuronal de aproximación (RNA) cuya estructura es alimentada con imágenes de líneas de luz proyectadas sobre objetos patrones con dimensiones conocidas. La estructura de esta RNA, consta de una capa de entrada, una capa oculta de neuronas y una capa de salida [9]. En la capa de entrada no se realiza ningún proceso, simplemente se transfiere un vector de datos de entrada a la capa oculta. Este vector de datos contiene la información de líneas de luz proyectadas sobre los objetos patrón. La conexión de los datos hacia la capa oculta no se asocia ningún peso, ya que no se realiza ningún proceso en la capa de entrada. En la capa oculta, cada neurona se determina mediante una función radial, la cual está centrada a la distancia del dato trasferido de la capa de entrada hacia la capa oculta. La capa de salida es formada con la suma de la neuronas, las cuales están multiplicadas por un peso. Esta función es la respuesta de salida, la cual se obtiene mediante un proceso de entrenamiento en donde se determinan los pesos asociados a cada neurona de la capa oculta. En este trabajo también se explica la forma de seleccionar los parámetros de la arquitectura de la RNA. El vector de los datos de entrada se obtiene detectando la posición en la imagen de la línea de luz proyectada sobre los objetos patrón. Para medir la posición de la línea de luz con resolución de fracción de *pixel* se usa el método de aproximación gaussiana. Este método es usado, ya que la distribución de intensidad de la línea de luz también es gaussiana. Con este método la posición de la línea de luz se obtiene calculando la posición central de la función gaussiana. Esta posición es calculada con una simple sumatoria. Esto representa una ventaja en tiempo de procesamiento, ya que normalmente la posición se obtiene calculando el máximo de intensidad mediante la aplicación de derivadas.

El arreglo experimental usado en esta técnica es muy simple, el objeto es desplazado en forma horizontal y a su vez es barrido por una línea de luz. En cada paso del movimiento se captura una línea de luz mediante una cámara CCD. De estas imágenes se extrae la posición de la línea de luz mediante aproximación gaussiana. Los datos de la posición de la línea son introducidos en la RNA, la cual genera una sección transversal del objeto. La información producida por el conjunto de imágenes procesadas durante el barrido de la línea se guarda en un arreglo de memoria para obtener la forma completa del objeto.

2. Descripción del arreglo óptico

En metrología óptica existen tres tipos de configuraciones geométricas para las técnicas de proyección de iluminación. En la primera configuración, la cámara y el proyector son alineados a un ángulo respecto a la superficie del objeto. En la segunda configuración, el proyector se alinea perpendicularmente a la superficie del objeto y la cámara se alinea a un ángulo. En la tercer configuración, la cámara está alineada

perpendicularmente a la superficie del objeto y el proyector está alineado a un ángulo. El método propuesto en este trabajo puede ser usado por estas tres configuraciones. En este trabajo, en particular, se usó la configuración donde el proyector de línea de luz está alineado perpendicularmente a la superficie del objeto y la cámara alineada a un ángulo. El arreglo experimental es muy simple como se muestra en la Fig. 1. En este arreglo, el objeto se fija sobre una plataforma, la cual se mueve en el eje x mediante un dispositivo electromecánico. Este dispositivo está controlado por computadora, y su movimiento está basado en un programa computacional. Sobre el objeto se proyecta una línea de luz vertical por medio de un diodo láser. En cada paso del movimiento, la línea de luz se deforma de acuerdo a la topografía del objeto. Cada una de estas líneas, es capturada por una cámara CCD y una tarjeta digitalizadora. La deformación de la línea de luz tiene una relación geométrica bien definida con la superficie del objeto. La topografía correspondiente a una región del objeto donde se proyecta la línea de luz es obtenida mediante el procesamiento de la RNA. 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.

Para describir la relación geométrica entre la deformación de la línea de luz y la superficie del objeto, se usa la geometría del arreglo experimental. En la Fig. 2 se muestra la geometría del arreglo óptico usado en este experimento. En el plano de referencia están localizados el eje x y el eje y , y la superficie del objeto se indica como $h(x, y)$. Los puntos A y B corresponden a la línea de luz proyectada sobre el plano de referencia y sobre la superficie del objeto, respectivamente. Cuando una línea de luz se proyecta sobre la superficie del objeto, ésta se mueve en el plano de la imagen del punto x_A al x_B . Este desplazamiento puede ser representado como

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

Este desplazamiento $s(x, y)$ es directamente proporcional a la superficie del objeto $h(x, y)$. Usando este desplazamiento, la superficie del objeto se determina mediante la RNA. Esta RNA contiene los datos de la superficie del objeto para un desplazamiento $s(x, y)$ de línea de luz dado. De esta manera,

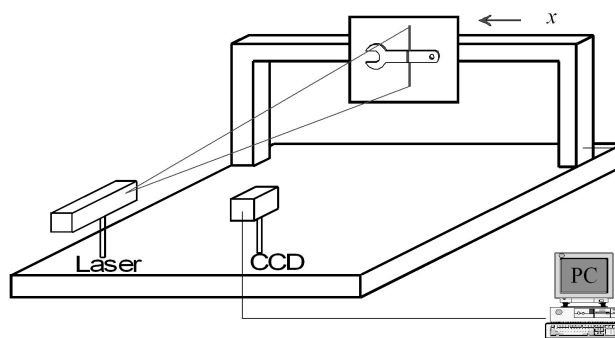


FIGURA 1. Arreglo experimental.

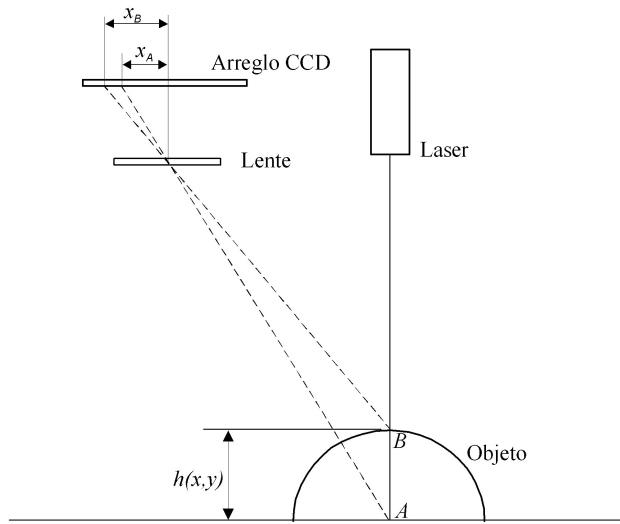


FIGURA 2. Geometría del arreglo experimental.

se obtiene la información de la superficie del objeto correspondiente a 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 *pixel* mediante el uso del método de aproximación gaussiana, el cual se explica en la Sec. 3.

3. Aproximación gaussiana de una línea de luz

La intensidad proyectada por un diodo láser es una distribución gaussiana en dirección lateral (eje \$x\$) [10]. Los valores de intensidad de cada renglón de una imagen de línea de luz 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 *pixel* y \$z_i\$ es la intensidad del *pixel*, para \$i = 1, 2, 3, 4, \dots, n\$. Una manera natural de representar estos valores de intensidad, es una función gaussiana. La ecuación de una función gaussiana 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 [11]. Para determinar la posición de la línea de luz, el parámetro más importante de esta función gaussiana es el valor de la media. Esto es porque \$\mu\$ representa el centro de la función gaussiana. Por lo tanto, la posición de la línea de luz se obtiene calculando la media \$\mu\$ en la imagen. 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 Fig. 3 muestra un conjunto de *pixeles* de un renglón de una imagen de línea de luz. De este conjunto de *pixeles*,

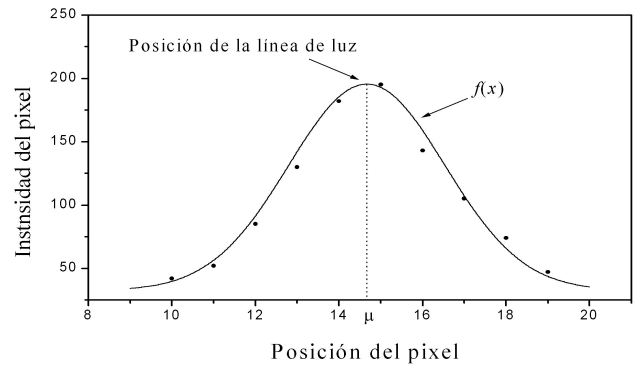


FIGURA 3. Función Gaussiana correspondiente a un conjunto de *pixeles*.

se calcula la correspondiente media \$\mu\$ que representa la posición central. Para realizar este procedimiento, se sustituye la posición \$x_i\$ y la intensidad \$z_i\$ de estos 10 *pixeles* en la Ec. (3), para \$i = 1, 2, 3, \dots, n\$. En este caso el resultado de la posición central es \$\mu = 14.664\$. Este resultado corresponde a la posición de la línea de luz \$x_B = 14,664\$ *pixeles*.

También se calcula el factor de normalización y la desviación estándar para graficar la función gaussiana mostrada en la Fig. 3, la cual representa a estos *pixeles*. Generalmente para una función gaussiana, el centro de esta se calcula mediante la media aritmética [12]. Esta media es definida como:

$$\mu_A = \frac{\sum_{i=1}^n x_i}{n} \quad (4)$$

El resultado de esta media es siempre el punto medio de la posición de los *pixeles* en el eje \$x\$. La media aritmética asume que todos los *pixeles* tienen la misma importancia. Por lo tanto, la posición central de la función gaussiana es el punto medio. Sin embargo, el máximo de intensidad de los *pixeles* no siempre está en el punto medio. Esto se debe a que la cámara o el diodo láser están alineados a un ángulo con respecto a la superficie del objeto. Por lo tanto, el máximo de intensidad se encuentra al lado izquierdo o al lado derecho del punto medio del conjunto de *pixeles*. El resultado de la media aritmética para los *pixeles* de la Fig. 3 es \$\mu = 14.5\$ *pixeles*. Como se puede ver, el resultado es diferente al resultado obtenido por la Ec. (3), la cual es llamada media ponderada. Para la media ponderada, se asigna un peso a cada punto. En este caso, el peso es el valor de la intensidad del *pixel*. Esto significa que los *pixeles* con mayor intensidad son más importantes que los *pixeles* con menor intensidad. Por lo tanto, la media ponderada Ec.(3) es usada para calcular la posición de la línea de luz. El procedimiento de calcular la posición de la línea de luz mediante la media ponderada, 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 de luz. Este perfil corresponde al vector de desplazamiento \$\mathbf{S} = (s_1, s_2, s_3, \dots, s_m)\$, el cual se introducirá en la RNA para obtener una sección transversal del objeto. En la Fig. 4 se muestra una línea de luz. Para cada renglón de esta imagen



FIGURA 4. Imagen de una línea de luz.

se calcula la media ponderada [Ec. (3)] para obtener la posición de la línea de luz a lo largo de ésta. Usando esta posición a lo largo de la imagen (eje y), se calcula el desplazamiento $s(x,y)$. Este desplazamiento representa el perfil del objeto, el cual es mostrado en la Fig. 5. Como se puede observar el perfil se determina con un mínimo de operaciones, vía Ec. (3). Este perfil es extraído de cada una de las capturadas en el movimiento del objeto.

4. Construcción de la arquitectura de la RNA

Las redes neuronales son estructuras compuestas de un número de neuronas conectadas entre sí. Cada neurona tiene una entrada, una salida y una función que procesa la información de entrada para obtener la respuesta de salida. A la salida de cada neurona se le asocia un peso, con lo cual se conecta a otras neuronas. Cada una de estas neuronas es activada cuando recibe el nivel de excitación adecuado, el cual proviene de las influencias separadas que recibe de sus enlaces de entrada. La función de activación corresponde a una función de umbral, la cual determina la activación de la salida de la neurona. De esta estructura se obtiene una función, la cual corresponde al modelo de la información con la cual se construye la red neuronal. Las redes neuronales, son aplicadas a problemas de clasificación, aprendizaje supervisado, regresión, identificación de formas y detección de objetos [13–19]. Una RNA es la representación de una red de interpolación, la cual se puede determinar por medio de funciones radiales [20]. La estructura de la red neuronal usada en este trabajo consta de tres capas: una capa de entrada, una capa oculta y una capa de salida. Esta estructura, la cual consta de tres capas, es equivalente a una estructura de dos capas [21–23], ya que la capa de entrada no realiza ningún proceso y únicamente se encarga de conectar los datos de entrada a cada una de las neuronas de la capa oculta. Esta capa de entrada es simplemente un

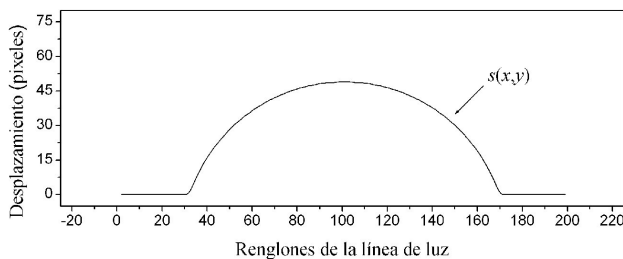


FIGURA 5. Perfil de un objeto obtenido de la línea de luz de la Fig. 4.

vector de datos muestreados y cada dato se transfiere a cada neurona de la capa oculta. La capa oculta se genera mediante funciones radiales, las cuales son centradas a la distancia de los datos muestreados de entrada. La capa de salida se obtiene con la sumatoria de las neuronas de la capa oculta multiplicadas por su peso w_i . Los pesos se obtienen después de un proceso de entrenamiento, con lo que se determina la respuesta deseada. La estructura de esta RNA se muestra en la Fig. 6. En esta RNA, la capa de entrada es simplemente un vector de datos: $\mathbf{S} = (s_1, s_2, s_3, \dots, s_m)$. La información de este vector de entrada se obtiene del procesamiento de imágenes de una línea de luz proyectada sobre objetos con dimensiones conocidas: $\mathbf{H} = (h_1, h_2, h_3, \dots, h_m)$, usando el método de aproximación gaussiana [Ecs. (3) y (1)]. Estos desplazamientos son mostrados en la Fig.7. La capa oculta recibe estos valores del vector de entrada y se genera una neurona en cada nodo. Estas neuronas son determinadas por una función radial, la cual opera sobre una área localizada a el valor del dato de entrada. La expresión de cada neurona se describe como $f_i(\|s - s_i\|^2)$, donde s_i es el centro de la función. El rango efectivo de la función es determinado por el centro y el ancho de la función. La respuesta para la capa de salida es la suma de todas las neuronas, las cuales son multiplicadas por un peso. La respuesta deseada para la capa de salida de esta RNA es

$$h(s) = \sum_{i=1}^n w_i f_i(\|s - s_i\|^2), \tag{5}$$

donde n es el número de neuronas. Para completar la arquitectura de esta RNA es necesario determinar sus parámetros propios. Estos parámetros son: umbral de activación, forma de la neurona, número de neuronas, centro de la neurona, ancho de la neurona y los pesos. El primer parámetro a determinar es la función de umbral. Para seleccionar la función de umbral de activación, se analiza la salida de la RNA para determinar la influencia del umbral en el comportamiento de la RNA. Existen diversas funciones de umbral como: función sigmoide, umbral lógico, y umbral lineal, las cuales han sido usadas ampliamente como funciones de activación de una red neuronal [24–25]. La selección de una función de umbral debe ser de acuerdo a la respuesta de salida deseada. De acuerdo a la arquitectura de esta RNA, la respuesta de salida esta dada por la Ec.(5). De esta manera, la respuesta de salida deseada es $h(s)$. Por lo tanto, se desea que la RNA se active cuando se tenga en la salida $h(s)$. La función umbral que cuenta con estas características es la función Umbral Lineal, la cual se describe como:

$$\text{salida} = \sum_{i=1}^n w_i f_i. \tag{6}$$

En esta forma, la salida se activa cuando la respuesta es la RNA. En el caso de la función sigmoide, la activación corresponde a la sigmoide de la RNA. Debido a esto no es posible emplearla, ya que no es la respuesta de salida deseada. En el caso de la función umbral lógico, la salida se activa en cero

si la red está abajo de un umbral o en uno si la red esta arriba del umbral. En este caso tampoco se puede aplicar esta función, ya que no es la respuesta de salida deseada. El segundo parámetro a determinar es la forma de la función de cada neurona de la capa oculta. Existen diversas funciones que pueden ser usadas para construir cada neurona, tales como función cúbica, función inversa cuadrática, función multicuadrática y función gaussiana [19]. Las definiciones de cada una de estas funciones se muestran en la Tabla I. En esta tabla r representa la distancia radial $\|s - s_i\|$. El análisis del comportamiento de la RNA, de acuerdo a la forma de las neuronas de la capa oculta, está basado en el error de la RNA con respecto a la respuesta de salida deseada. El criterio MSE (MSE: *Mean Squared Error*, error cuadrático medio), provee información del error cuadrático medio de la respuesta de salida generada por la RNA con respecto a la respuesta de salida deseada [18]. El uso de este criterio se debe a que lo más importante de esta técnica de detección de forma es la precisión de los resultados. El criterio MSE es la media de la diferencia al cuadrado entre la respuesta de salida obtenida por la RNA y la respuesta de salida deseada. El criterio MSE se define por

$$MSE = \frac{1}{m} \sum_{i=1}^m (hc_i - h_i)^2, \tag{7}$$

donde h_i es la respuesta de salida de la RNA, hc_i es la respuesta de salida deseada y m es el número de muestras a analizar. Para definir cuál función se debe de utilizar como base radial en la capa oculta, se entrenó la red usando 10 neuronas para cada una de las funciones propuestas. Posteriormente, se determinó una respuesta de salida de la RNA vía Ec.(5) con cada una de las funciones radiales. Para cada una de las respuestas de salida obtenidas se calcula el valor MSE usando valores de superficie de objetos con dimensiones conocidas. El criterio MSE, indica cuál RNA generada por estas funciones ajusta mejor a la respuesta de salida deseada. De acuerdo a la Tabla I, la función Gaussiana es la función con la que la RNA ajusta mejor a la respuesta de salida deseada. Por lo tanto, la función gaussiana es usada para construir la capa oculta de la estructura de la RNA. De esta manera las neuronas de la capa oculta se expresan como

$$f_i = \exp\left(-\frac{\|s - s_i\|^2}{2\sigma^2}\right), \tag{8}$$

de esta expresión, s_i y σ , son el centro y el ancho de la función gaussiana, respectivamente. Por lo tanto, la expresión de la RNA Ec.(5) se expresa como

$$h(s) = \sum_{i=1}^n w_i \exp\left(-\frac{\|s - s_i\|^2}{2\sigma^2}\right). \tag{9}$$

TABLE I. Comparación de la forma de las funciones radiales usadas para generar una RNA. En esta tabla, se muestra la definición de funciones radiales y número de neuronas que se usan para determinar el criterio MSE. En la primer columna, se muestra la definición de la funciones radiales. En la segunda columna, se muestra el nombre de cada una de estas funciones. En la tercer columna se muestra el numero de neuronas usadas para entrenar la RNA. En la cuarta columna, se muestra el valor del MSE para cada una de las funciones radiales.

Función radial	Nombre	Neuronas	MSE
r^3	Función radial cubica	10	0.00374
$(r^2 + \sigma^2)^{0.5}$	Función multi-cuadrática	10	0.01721
$1/(1 + r^2)$	Función Inversa Cuadrática	10	0.00302
$\exp(r^2/2\sigma^2)$	Función gaussian	10	0.00246

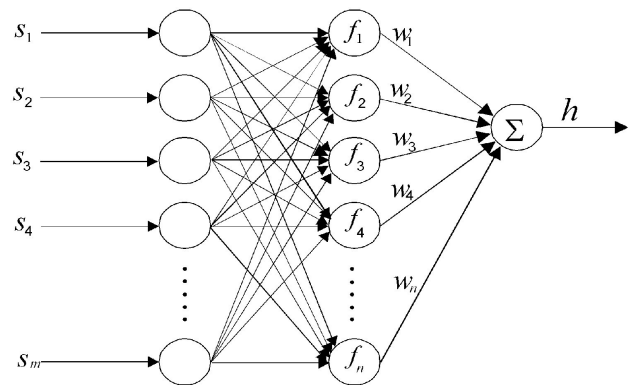


FIGURA 6. Estructura de la Red Neuronal de Aproximación.

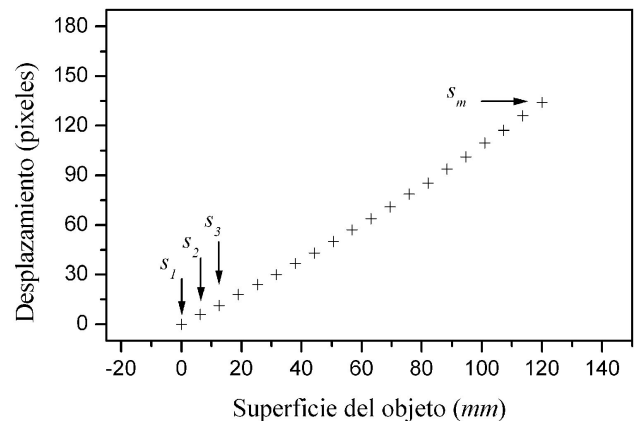


FIGURA 7. Desplazamientos de una línea de luz proyecta sobre objetos patrón.

El tercer parámetro es el centro s_i , el cual proviene del vector de entrada de datos muestreados con lo que se efectuara el entrenamiento. El cuarto parámetro el ancho de las

funciones gaussianas. En este caso el ancho de las funciones gaussianas se denota por σ y se define como

$$\sigma = \frac{d}{\sqrt{2n}}, \tag{10}$$

donde d es la distancia entre el centro de cada par de neuronas y n es el número de neuronas. El quinto parámetro a determinar es el número de neuronas que se deben utilizar en el proceso de entrenamiento de la RNA. Para la capa de entrada el número de neuronas es igual al número de datos m del vector de entrada \mathbf{S} . Esto se debe a que en la capa de entrada no se realiza ningún proceso y los datos son transferidos directamente a la capa oculta. Pero en la capa oculta se usa el criterio MSE para determinar el número de neuronas de esta capa. El criterio MSE también nos proporciona el error de la respuesta de salida de la RNA de acuerdo al número de neuronas en la capa oculta usadas en el entrenamiento. La Fig. 8, muestra el error en la salida la RNA de acuerdo la número de neuronas usadas en el entrenamiento. Para la estructura de esta RNA, el criterio MSE indica que el error es mínimo y estable con más de diez neuronas. Por lo tanto, el número de neuronas para construir esta RNA es a partir de diez. El último parámetro a determinar son los pesos w_i de la RNA.

La estructura de esta RNA de tres capas trabaja igual que una estructura de dos capas, calculan pesos, ya que en la capa de entrada no se realiza ningún proceso. Por lo tanto, no se asocia ningún peso a la salida de la capa de entrada y los datos son transferidos directamente a la capa oculta. Para determinar los pesos de la capa oculta se realiza un proceso de entrenamiento, en el cual éstos son ajustados. Este tipo de entrenamiento es supervisado, ya que la respuesta de salida de la red es conocida. Por lo tanto, los pesos se deben ajustar para obtener la respuesta de salida deseada $h(s)$. En este proceso de entrenamiento, la RNA recibe un conjunto de datos que consiste de duplas (s_i, h_i) , para $i = 1, 2, 3, \dots, m$. De estas duplas, (s_i, h_i) son la entrada y la respuesta de salida, respectivamente, y puede ser representadas como un vector o una matriz. En este caso los datos de entrada provienen del vector $\mathbf{S} = (s_1, s_2, s_3, \dots, s_m)$, el cual proviene de imágenes de línea de luz procesadas y la repuesta de salida proviene del vector $\mathbf{H} = (h_1, h_2, h_3, \dots, h_m)$, el cual corresponde a las dimensiones de los objetos de las imágenes procesadas. La respuesta de salida de la RNA es definida por el cálculo de la suma de productos de acuerdo a la Ec. (9). De acuerdo a la estructura de esta red [Fig. 6 y la Ec. (9)] para cada entrada s_i se genera un sistema de ecuaciones que implican a los pesos no conocidos:

$$\begin{aligned} h_1 &= w_1 \exp\left(-\frac{\|s_1 - s_1\|^2}{2\sigma^2}\right) + w_2 \exp\left(-\frac{\|s_1 - s_2\|^2}{2\sigma^2}\right) + \dots + w_n \exp\left(-\frac{\|s_1 - s_n\|^2}{2\sigma^2}\right) \\ h_2 &= w_1 \exp\left(-\frac{\|s_2 - s_1\|^2}{2\sigma^2}\right) + w_2 \exp\left(-\frac{\|s_2 - s_2\|^2}{2\sigma^2}\right) + \dots + w_n \exp\left(-\frac{\|s_2 - s_n\|^2}{2\sigma^2}\right) \\ &\vdots \\ h_m &= w_1 \exp\left(-\frac{\|s_m - s_1\|^2}{2\sigma^2}\right) + w_2 \exp\left(-\frac{\|s_m - s_2\|^2}{2\sigma^2}\right) + \dots + w_n \exp\left(-\frac{\|s_m - s_n\|^2}{2\sigma^2}\right) \end{aligned} \tag{11}$$

Usando la notación de la Ec. (8), este sistema de ecuaciones lineales Ec. (11), se puede describir como:

$$\begin{aligned} h_1 &= w_1 f_{1,1} + w_2 f_{1,2} + \dots + w_n f_{1,n}, \\ h_2 &= w_1 f_{2,1} + w_2 f_{2,2} + \dots + w_n f_{2,n}, \\ &\vdots \\ h_m &= w_1 f_{m,1} + w_2 f_{m,2} + \dots + w_n f_{m,n}. \end{aligned} \tag{12}$$

en la Ec. (14), m es el índice del dato que llega a cada neurona para ser procesado y también al índice de la respuesta de salida, de acuerdo a la Ec. (11). El índice n corresponde al dato de cada neurona y a cada peso de la neurona con respecto a la Ec. (11). Este sistema de ecuaciones lineales

[Ec. (14)], también se puede representar como producto de matrices: $\mathbf{FW} = \mathbf{H}$. De esta manera se obtiene un sistema lineal, del cual son calculados los pesos. Este sistema lineal se define como:

$$\begin{bmatrix} f_{1,1} & f_{1,2} & f_{1,3} & \dots & f_{1,n} \\ f_{2,1} & f_{2,2} & f_{2,3} & \dots & f_{2,n} \\ \vdots & \vdots & \vdots & & \vdots \\ f_{m,1} & f_{m,2} & f_{m,3} & \dots & f_{m,n} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} \tag{13}$$

Cuando el número de datos del vector de la capa de entrada es igual al numero de neuronas de la capa oculta, se obtiene una matriz cuadrada \mathbf{F} . Resolviendo el sistema lineal de la Ec. (13), se obtienen los pesos w_i . Este sistema lineal se resuelve por el método de chelosky [26]. Este método usa

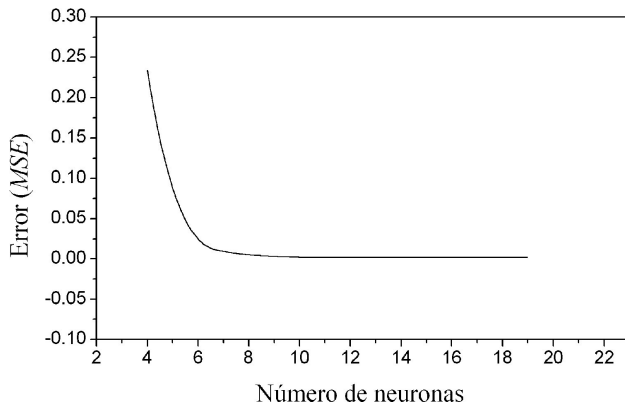


FIGURA 8. Error MSE de salida de la RNA de acuerdo al número de neuronas.

la descomposición LU para transformar el sistema de matrices. En la descomposición LU, una matriz simétrica y positiva definida se descompone en una matriz triangular inferior y en una matriz triangular superior. Usando la matriz triangular superior e inferior, el sistema lineal de la Ec. (13) se transforma en $LUW = H$. Entonces el sistema se resuelve como $LY = H$ y $UW = Y$. Con este procedimiento, los pesos w_i son calculados y la estructura de la RNA ha sido completada. Esta RNA es una función continua, la cual describe las dimensiones de la superficie de un objeto usado en el desplazamiento de una línea de luz. La Fig. 9 muestra la RNA obtenida en este trabajo, construida con diez neuronas Gaussianas.

5. Resultados experimentales

En la Fig. 1 se muestra el arreglo experimental usado en este trabajo. En este arreglo, el objeto se mueve en pasos de fracción de milímetro en forma horizontal, mediante un dispositivo electromecánico. Este dispositivo está conectado a una computadora, la cual manipula el movimiento mediante un programa. Sobre la superficie de este objeto se proyecta una línea de luz, por medio de un diodo láser de 5 mW. En

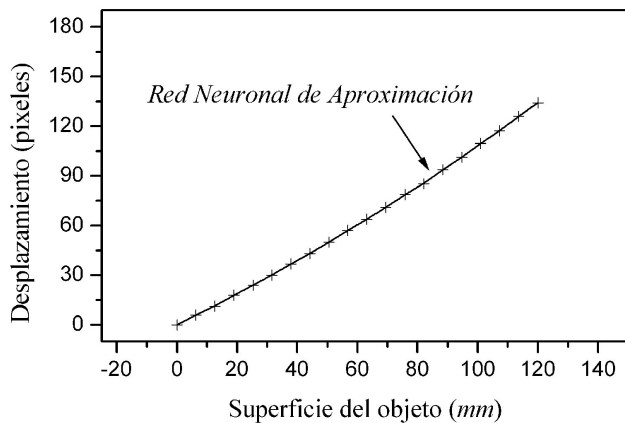


FIGURA 9. RNA construida con diez neuronas.

cada paso del movimiento, la línea de luz se deforma de acuerdo a la topografía de la superficie del objeto. Cada una de estas líneas es capturada por una cámara CCD y una tarjeta digitalizadora, las cuales están conectadas a computadora. La resolución de estas imágenes es de 320×200 *pixeles* y 256 niveles de gris. Los valores de la superficie $h(s)$ se obtienen sustituyendo el desplazamiento $s(x, y)$ de las imágenes de línea de luz en la RNA [Ec. (9)]. Con la información obtenida de todas las líneas procesadas se obtiene la forma 3D completa del objeto.

En este trabajo se presentan dos objetos, de los cuales se extrajo su forma 3D. El primer objeto es la la superficie de un vaso, el cual se muestra en la Fig. 10. De este objeto se captura un conjunto de imágenes, con las cuales se reconstruyó su topografía. De cada una de estas imágenes se extrae un perfil del objeto, mediante el método de aproximación gaussiana como se describe en la Sec. 3. Esto se realiza calculando la posición central μ de la función gaussiana en cada renglón de la imagen, usando la Ec.(3). Usando el resultado de la posición central, se calcula la deformación de la línea de luz usando la Ec.(1). El resultado de este paso es un vector $S(s_1, \dots, s_m)$. Los valores de este vector se sustituyen en la RNA, con lo que se obtiene una sección transversal del objeto. Cada una de las secciones transversales obtenidas del conjunto de imágenes se guardan en un arreglo de memoria para obtener la forma 3D completa.

Para conocer la precisión de los resultados obtenidos se calcula el valor *rms* (*root mean squared*) [27]. Este valor *rms* se determina usando datos calculados mediante la RNA y datos obtenidos de una máquina de medición de coordenadas (MMC). El valor *rms* es definido por

$$rms = \sqrt{\frac{1}{m} \sum_{i=1}^m (ho_i - hc_i)^2}, \tag{14}$$

donde ho_i son los datos medidos con la MMC, hc_i son los datos $h(s)$ calculados por la RNA y m es el número de muestras. El resultado del valor *rms* obtenido para la superficie del

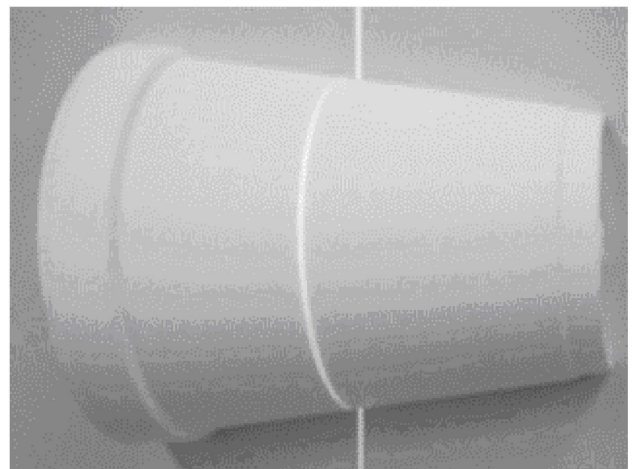


FIGURA 10. Superficie de un vaso usado para extraer su forma 3D.

vaso es $rms = 0.1721$ mm. La forma 3D completa de la superficie del vaso se obtuvo mediante el procesamiento de 84 líneas de luz. La topografía de este objeto se muestra en la Fig. 11. La escala de los ejes de esta figura esta en mm.

El segundo objeto usado para reconstruir su forma 3D es un arnés, el cual se muestra en la Fig. 12. Al igual que para la superficie del vaso, se aplica aproximación gaussiana y RNA para reconstruir la forma 3D del arnés. De esta manera, se extrae la información del desplazamiento de las imágenes y posteriormente se envía a la RNA para extraer secciones transversales del objeto. El resultado del valor rms obtenido para el arnés es $rms = 0.1672$ mm. En este caso se procesaron 96 líneas para reconstruir la forma 3D completa, como se muestra en la Fig. 13. En esta figura, la escala de los ejes está en mm. La computadora que se usó en este proceso es una PC 5×86 a 166 MHz. Cada línea de luz es procesada en 0.062 segundos. Este tiempo de procesamiento es muy rápido, ya que la información se extrae de la imagen con un mínimo de operaciones vía la Ec.(3). En este proceso se puede observar que los parámetros de la geometría del arreglo

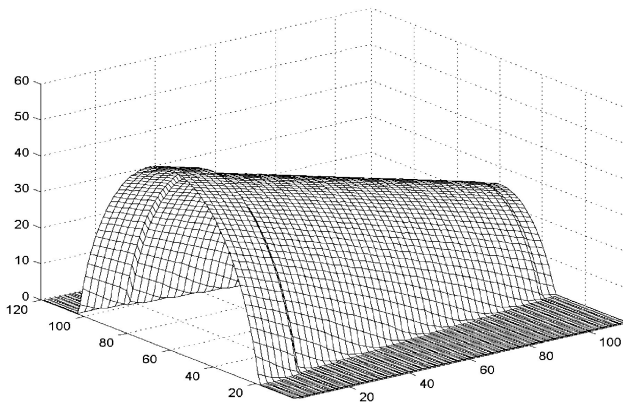


FIGURA 11. Forma 3D de la superficie del vaso correspondiente a la Figura 10.

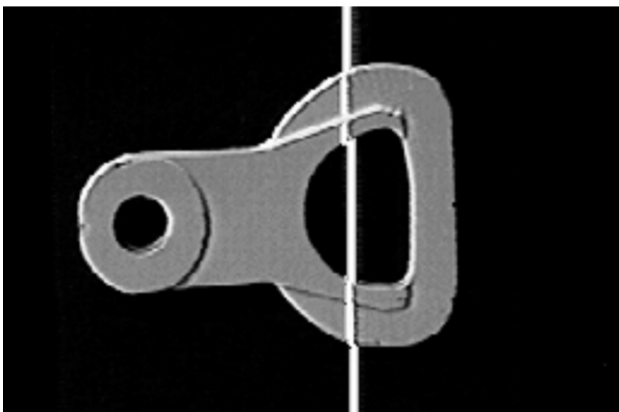


FIGURA 12. Arnés usado para extraer su forma 3D.

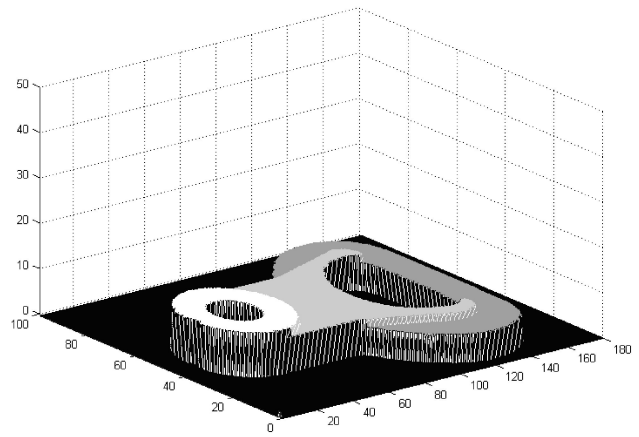


FIGURA 13. Forma 3D del arnés correspondiente a la Figura 12.

experimental no se incluyen en el algoritmo computacional para determinar la forma 3D del objeto. De esta manera, todo el proceso se facilita, ya que no se introducen parámetros medidos en forma externa. Esto a diferencia de cómo se hace en los métodos de proyección de línea de luz. En esta técnica es posible modificar la posición de los componentes del arreglo experimental y la RNA construye otra función $h(s)$, mediante el ajuste de pesos. Por lo tanto, todos los pasos de esta técnica se realizan por medio de un proceso computacional y se evitan mediciones sobre el arreglo experimental. De esta manera se logra una buena repetibilidad en cada medición.

6. Conclusiones

En este trabajo se ha presentado una técnica para reconstrucción 3D de objetos basada en una RNA de una línea de luz. Esta técnica provee una herramienta valiosa para detección 3D de forma en diversas áreas de inspección industrial. Esta técnica permite que se eviten mediciones sobre el arreglo experimental, como es común en los métodos de proyección de línea. Esto mejora la precisión de los resultados, ya que no se introducen al sistema computacional errores de mediciones externas. De esta manera, todo el procedimiento se realiza en forma computacional mediante la RNA y aproximación gaussiana. En este trabajo también se muestra que el procesamiento de imágenes se realiza con una resolución de fracción de *pixel*, mediante aproximación gaussiana. Este procesamiento se efectúa de manera muy rápida, ya que la posición de la línea de luz se determina con muy pocas operaciones. Por otro lado, las mediciones que se efectúan con esta técnica, se realizan con una buena repetibilidad.

Agradecimientos

J. Apolinar Muñoz Rodríguez, agradece al CONACYT, México y CONCYTEG Guanajuato, el apoyo brindado a este trabajo.

1. X.X. Cheng, X.Y. Su y L.R. Guo, *Appl. Optics* **30** (1991) 1274.
2. L.P. Thomas, R. Gratton, B.M. Marino y J.M. Simon, *Appl. Optics* **33** (1994) 2455.
3. W.Ch. Tai y M. Chang, *Opt. Eng.* **35** (1996) 2730.
4. M. Baba, T. Konishi y N. Kobayashi, *Journal of Optics* **29** (1998) 241.
5. S.R. Marokkey, Ch.J. Tay, H.M. Shang y A. K. Asundi, *Opt. Eng.* **36** (1997) 2573.
6. M.R. Sajan, C.J. Tay, H.M. Shang y A. Asundi, *Optics Comm.* **150** (1998) 66.
7. A. Asundi y W. Zhou, *Opt. Eng.* **38** (1999) 339.
8. J.A. Muñoz-Rodríguez, R. Rodríguez-Vera y M. Servin, *Opt. Eng.* **39** (2000) 2463.
9. B. Muller, J. Reinhardt y M. T. Strickland, *Neural networks An Introduction*, (Springer, U.S.A 1995).
10. W.D. Herzog, M.S. Unlu, B.B. Goldberg y G.H. Rhodes, *Appl. Phys. Lett.* **70** (1997) 688.
11. W.J. Dixon y F.J. Massey Jr., *Introduction to statistical analysis*, (Mc Graw Hill, U.S.A 1969).
12. J.P. Holman, *Experimental methods for Engineers*, (Mc Graw Hill, U.S.A 1984).
13. A.M. Neifeld y D. Psaltis, *Appl. Optics* **32** (1993) 1370.
14. E.W. Foor and A.M. Neifeld, *Appl. Optics* **34** (1995) 7545.
15. F.J. Cuevas, M. Servin y R. Rodríguez-Vera, *Optics Comm.* **163** (1999) 270.
16. Y. Fravel, G. Pauliat, Avilling y G. Roosen, *Appl. Optics* **40** (2001) 5162.
17. T. Sabisch, A. Ferguson y H. Bolouri, *IEEE Transactions on Neural Networks* **11** (2000) 921.
18. H. Leung, N. Dubash y N. Xie, *IEEE Transactions on Aerospace and Electronic Systems* **38** (2002) 98.
19. I. Rojas *et al.*, *Neural Processing Letters* **12** (2000) 1.
20. P.H. Winston, *Artificial Intelligence*, (Addison-Wesley, USA, 1992).
21. P. Picton, *Neural networks*, (Polgrave, U.S.A 1998).
22. N. Mark y P. Demetri, *Appl. Optics* **32** (1993) 1370.
23. F. Wesley y N. Mark, *Appl. Optics* **34** (1995) 7545.
24. R.J. Schalkoff, *Artificial Neural Networks*, (Mc Graw Hill, USA, 1997).
25. J.A. Freeman y D.M. Skapura, *Neural Networks algorithms, Applications and Programming techniques*, (Addison Wesley, USA 1991).
26. W.H. Press, B.P.Flannery, S.A.Teukolsky y W.T.Vetterling, *Numerical Recipes in C*, (Cambridge Press, 1993).
27. T. Masters, *Practical Neural Networks Recipes in C++*, (Academic Press, U.S.A 1993).