

Proyecto Final:

Percepción de profundidad

Bernardo Quintino Guzmán, b.quintinoguzman@ugto.mx,
Departamento de Estudios Multidisciplinarios, Universidad de Guanajuato.

I. INTRODUCCIÓN

Junto con la revolución tecnológica que se ha presentado en las últimas décadas, se tendió a simular las destrezas que los humanos poseen, como lo es la visión, los seres vivos son poseedores de dos ojos que permiten ver imágenes y procesarlas con el cerebro, la visión por computadora es un campo de la ciencia y tecnología que se enfoca en contener esas destrezas y transformar en algo que una maquina pueda realizar, pues dada algunas imágenes una computadora puede procesarlas, analizarlas y generar cualquier explicación de cualquier tipo de información.

El ser humano y otros seres vivos están dotados de una visión estero, la cual permite percibir la realidad en tres dimensiones, se percibe la profundidad del entorno gracias al fenómeno visual conocido como estereopsis. Este fenómeno permite saber qué cosas están más cercas o que cosas están más lejanas. Todo este proceso es muy complejo, pero se ha logrado llegar a emular por medido de algoritmos de computadora.

En este documento se presenta la realización de un programa que puede obtener mediante dos imágenes la profundidad de un escenario, dado que, al procesar la información obtenida en un par de imágenes estereoscópicas, se obtiene una gran cantidad de información geométrica de la escena capturada, y con ella generar información útil para sistemas de navegación, robots, mapas topográficos, modelos tridimensionales etc. ([1]).

El principal factor para el algoritmo realizado es la obtención de un mapa de disparidad denso, la *disparidad* en los humanos es la diferencia entre dos imágenes obtenida con los dos ojos, debido a su separación, mientras que la *estereopsis* interpreta la disparidad que permite el fenómeno la sensación de la percepción de la profundidad y para la programación se usa la geometría epipolar, que es la geometría proyectiva intrínseca entre dos imágenes, la cual es independiente de la estructura de la escena y solo depende de los parámetros internos de la cámara y su posición relativa, ya cuando obtenemos el resultado del programa se presentan algunos errores que son parte de las oclusiones, las oclusiones son los detalles que se encuentran en una imagen, pero no en la otra, pues se encuentran ocluidos por objetos cercanos debido a la diferencia de perspectiva entre ellas y a los cambios en profundidad ([2]).

El programa contiene diversos algoritmos, el primero que se ejecuta es cuando se tiene las dos imágenes estereoscópicas, primero se les aplica el algoritmo de ecualización del histograma a cada una para que mejoren su contraste y al final el resultado sea mejor, luego ya se realiza el algoritmo principal para generar un mapa de disparidad denso que se denomina SSD (Sum of Squared Dierences) por pixel sobre una ventana de un tamaño fijo, del cual se consigue el valor mínimo de disparidad para cada pixel, y la posición que ocupe se definirá como el nuevo pixel de la imagen, y por ultimo se le aplica a la imagen resultado el método de expansión del histograma para lograr un mejor y más claro resultado.

II. METODOLOGÍA

A. Expansión del histograma

El método de expansión del histograma o procesamiento puntual, se utilizó una fórmula para transformar uno por uno los pixeles de la imagen.

$$O[x, y] = 255 * (I[x, y] - h_{min}) / (h_{max} - h_{min})$$

La figura 1 se muestra el diagrama que representa a grandes rasgos el algoritmo de expansión del histograma.

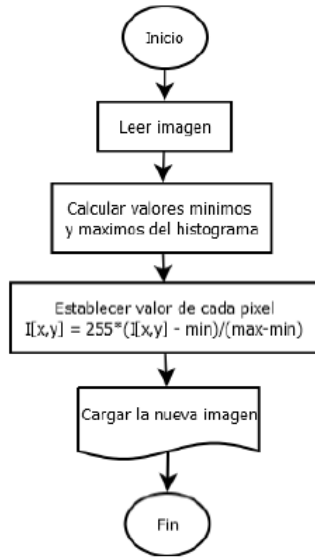


Figura 1: Diagrama de expansión del histograma.

B. Ecualización del histograma

La ecualización del histograma, se realiza más trabajo en cuanto a calcular el valor para cada pixel, primero se consiguió el histograma de la imagen transformada a escala de grises, en las siguientes formulas se expresan las utilizadas en el programa:

$$O[x, y] = 255 * acumulado(I[x, y]) / (I_{width} * I_{height})$$

$$acumulado(k) = \sum_{i=0}^k H[i]$$

Donde H[i] es un valor del histograma de escala de grises.

La figura 2 se muestra el diagrama que representa a grandes rasgos el algoritmo de expansión del histograma.

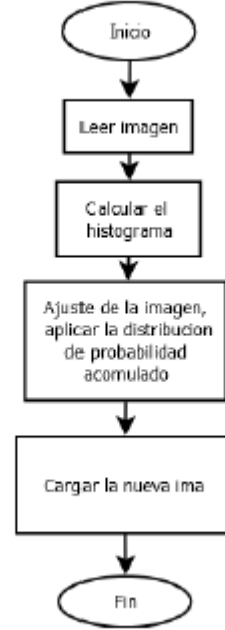


Figura 2: Diagrama de expansión del histograma.

C. Sum of Squared Dierences

La técnica tradicional básica para generar un mapa de disparidad denso se denomina SSD (Sum of Squared Dierences). Se trata de una técnica de correspondencia basada en la correlación entre los pixeles de la imagen izquierda y derecha dentro de un vecindario. Normalmente se escoge una ventana de un cierto tamaño NxM alrededor de un pixel en la imagen izquierda 'Iizq' y se desplaza una ventana del mismo tamaño a lo largo de la misma línea en la imagen derecha 'Ide', considerando el pixel homologo (correspondiente) aquel que tenga un menor valor de correlación. La correlación sobre cada punto de la imagen izquierda 'Iizq (xl; yl)' y cada valor de disparidad d de la siguiente manera ([3]):

$$C_{SSD}(x^l, y^l, d) = \frac{1}{nm} \sum_{i=x^l-n/2}^{x^l+n/2} \sum_{j=y^l-m/2}^{y^l+m/2} (I_{izq}(i, j) - I_{der}(i - d, j))^2$$

De esta manera se obtiene el mapa de disparidad denso $D_{calculada}(u; v)$ dado un par de imágenes estéreo:

$$D_{calculada}(x^l, y^l) = \arg \min_{d \in [0, D]} C_{SSD}(x^l, y^l, d)$$

Donde D es el valor de disparidad máximo esperado en la escena. se puede establecer el valor mínimo de N y M a 7 unidades, y a d se puede establecer a 60.

D. Procedimiento general

En el siguiente diagrama se describe a grandes rasgos el algoritmo de percepción de profundidad, ver figura 3.

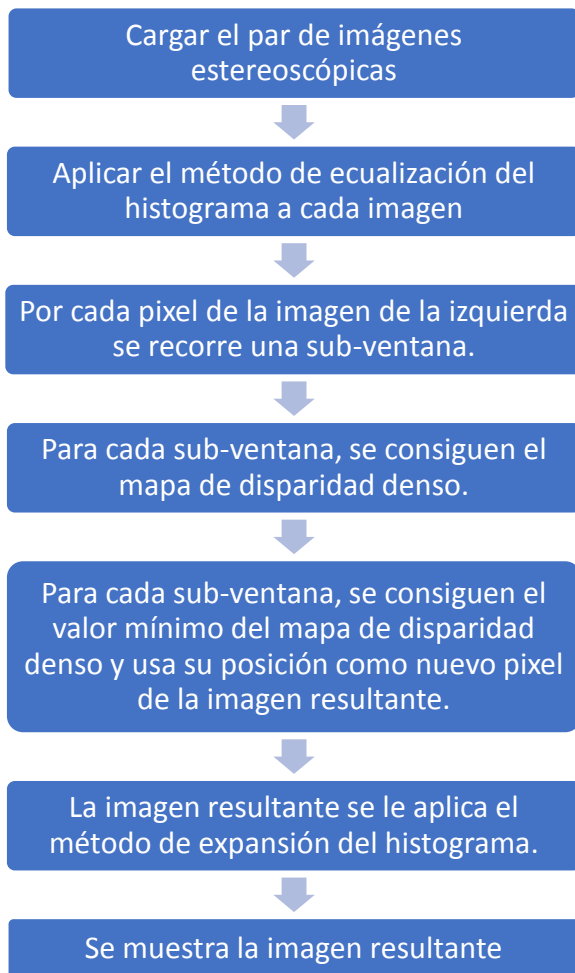


Figura 3: Diagrama del algoritmo

III. RESULTADO

Para el programa se implementó un interfaz que contiene dos elementos tipo imagen, tres botones y un elemento de texto, los dos primeros botones permiten seleccionar cualquier imagen dentro del almacenamiento de la computadora, el primer botón establece a la imagen de la izquierda y el segundo botón sirve para establecer la imagen de la derecha, el tercer y último botón sirve para ejecutar todo el algoritmo, el elemento de texto sirve para mostrar el progreso en porcentaje del algoritmo ejecutado, dado que dependiendo del tamaño de las imágenes durará mucho tiempo y así poder cuantificar el tiempo de duración de la ejecución del algoritmo, por default al iniciar la aplicación se carga dos imágenes dentro del proyecto y siempre que se inicie este programa se establecerá esas imágenes.

En la figura 4 se muestra la interfaz obtenida, por default el número de la dimensión de la sub-ventana es 7.



Figura 4: Interfaz

Para poder obtener un mejor resultado algunos autores agregan un algoritmo dinámico para modificar el tamaño de la sub-ventana, sin embargo esto complica la situación y agrega mucho más tiempo para poder obtener algún resultado.

En la figura 5 se muestra el primer resultado obtenido, pero como fue la primera prueba, este resultado es únicamente aplicando el algoritmo del SSD, en este resultado ya se muestran la percepción de profundidad indicada por el tono de los objetos, pues mientras más cercano es más claro.

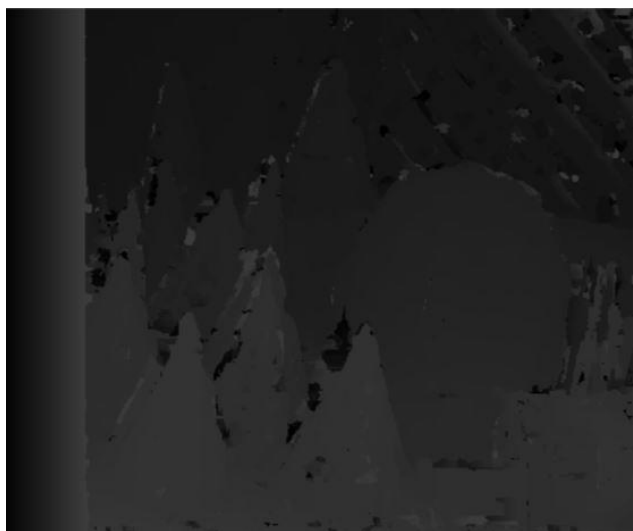


Figura 5: Resultado sin contraste

La imagen no es perfecta puesto a que se notan algunos errores, debido a las ocluciones, se puede implementar algoritmos para establecer el tamaño de la sub-ventana de forma dinamica y mejorar el resultado. Deduciendo que se deberian mejorar el resultado se implementaron otros algoritmos de mejoramiento de imagen.

En la figura 6 se muestra el resultado pero en este caso, al par de imágenes estereoscópicas se les aplico el metodo de ecualizacion del histograma que permitió obtener una imagen resultante mas clara.



Figura 6: Resultado con contraste

En la figura 7 se muestra resultados final, en este caso ademas del resultado obtenido en la figura 6, se le aplico el metodo de expansión del histograma el cual permite obtener una imagen mas clar y resalta mas los detalles.

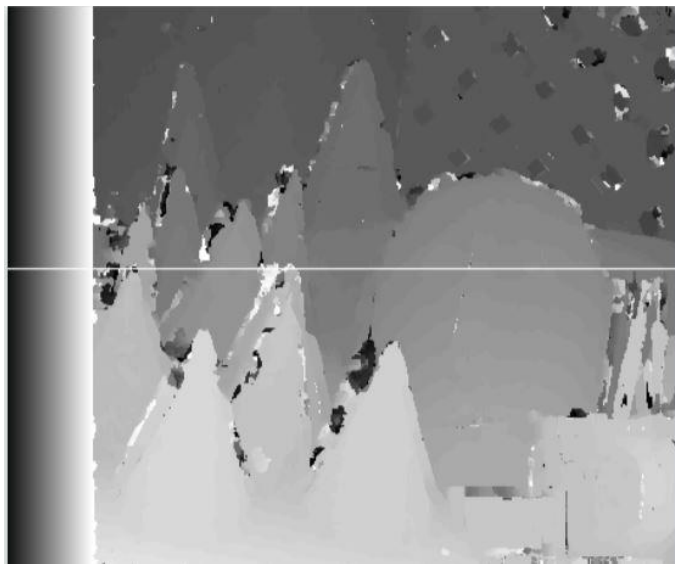


Figura 7: Resultado final

En la figura 8 se muestra el resultado obtenido de otra persona, en comparacion con la figura 7 se nota que se parece mucho, sin embargo el resultado es mejor en la implementacion de contrastes usado en esta proyecto.

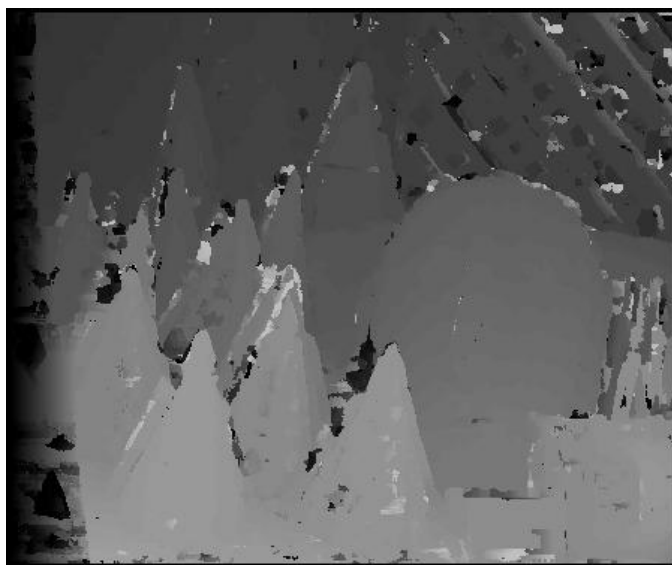


Figura 8: Resultado de otra persona

Se hizo otra prueba con otras dos imágenes estereoscópicas, en este caso se usó las imágenes de Tsukuba. Ver figura 9.



Figura 9: Segunda prueba

El resultado de la segunda prueba se muestra en la figura 10.

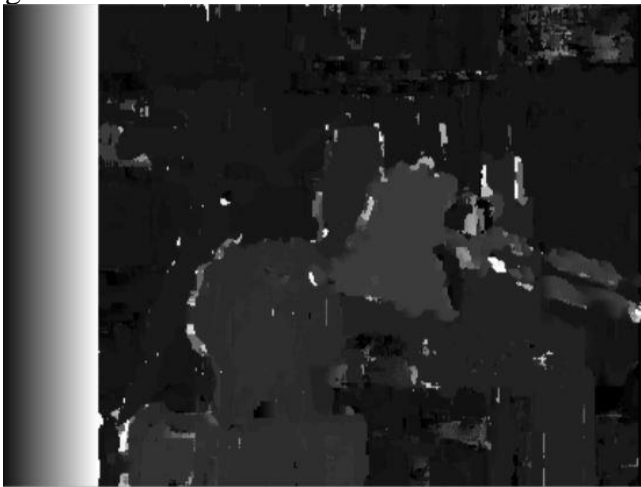


Figura 10: Resultado de la segunda prueba.

Para la tercera prueba se usaron las imágenes “Venus” en la figura 11.



Figura 11: Tercera prueba.

El resultado de la tercera prueba se muestra en la figura 12.

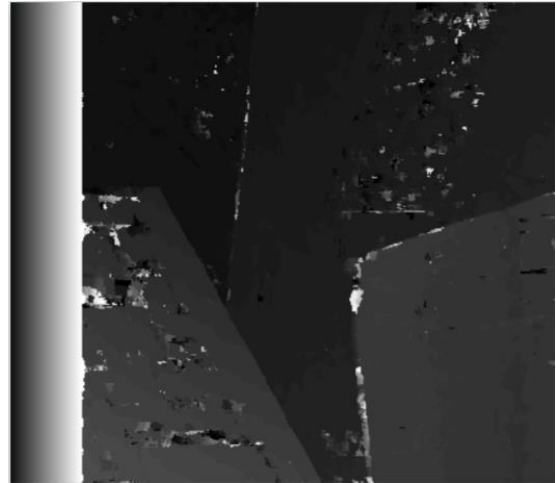


Figura 12: Resultado de la tercera prueba.

Para la cuarta prueba se usaron las imágenes “Teddy” en la figura 13.



Figura 13: Cuarta prueba.

El resultado de la cuarta prueba se muestra en la figura 14.

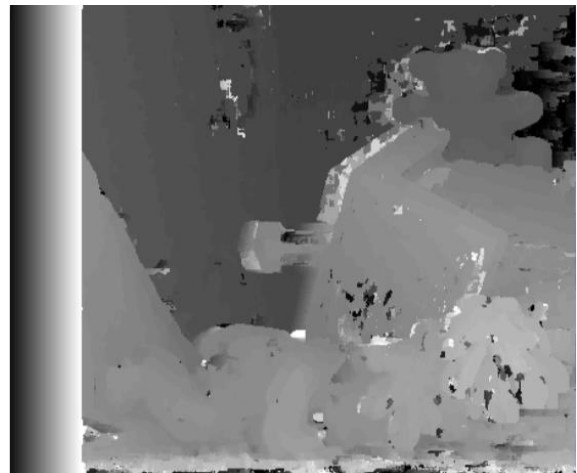


Figura 14: Resultado de la cuarta prueba.

IV. CONCLUSIÓN

Es fascinante lograr que una maquina simule las habilidades de los humanos, ya sea contar, manipular, analizar y procesar cualquier tipo de información, como los robots que pueden interactuar ya con el mundo real, en la visión por computadora, se enfoca en transformar las imágenes en información, lo que permite manipularlas, en este documentos se expresó la percepción de profundidad procesada por una computadora usando dos imágenes estereoscópicas, para un ser humano es fácil distinguir la cercanía de los objetos, esto por que nuestro cerebro esta acostumbrado a realizar esa tarea, aquí el principal algoritmo que se sigue es el SSD, que permite obtener un buen resultado de la posición de los objetos respecto a la cámaras, sin embargo las imágenes no son perfectas, se notan algunos erros, para evitar dichos errores, se usan otros algoritmos que establecen dinámicamente el tamaño de la sub-ventana y el rango de disparidad, los resultados obtenidos por las pruebas son buenos, ya se percibe la diferencia entre los objetos, este programa puede ser utilizado junto con otros que permite obtener resultados aún más interesantes.

V. REFERENCIAS

- [1] Stefano Mattoccia, Stereo Vision: Algorithms and Applications, University of Bologna, January 12, 2013.
- [2] Mario Gonzalo Chirinos Colunga, “Visión Artificial: Percepción de Profundidad”, Instituto Tecnológico de Mérida., Noviembre 2004.
- [3] Martin Peris Martorell, Aproximación del Mapa de Disparidad Estéreo Mediante Técnicas de Aprendizaje Automático, pág. 12, septiembre 2011.