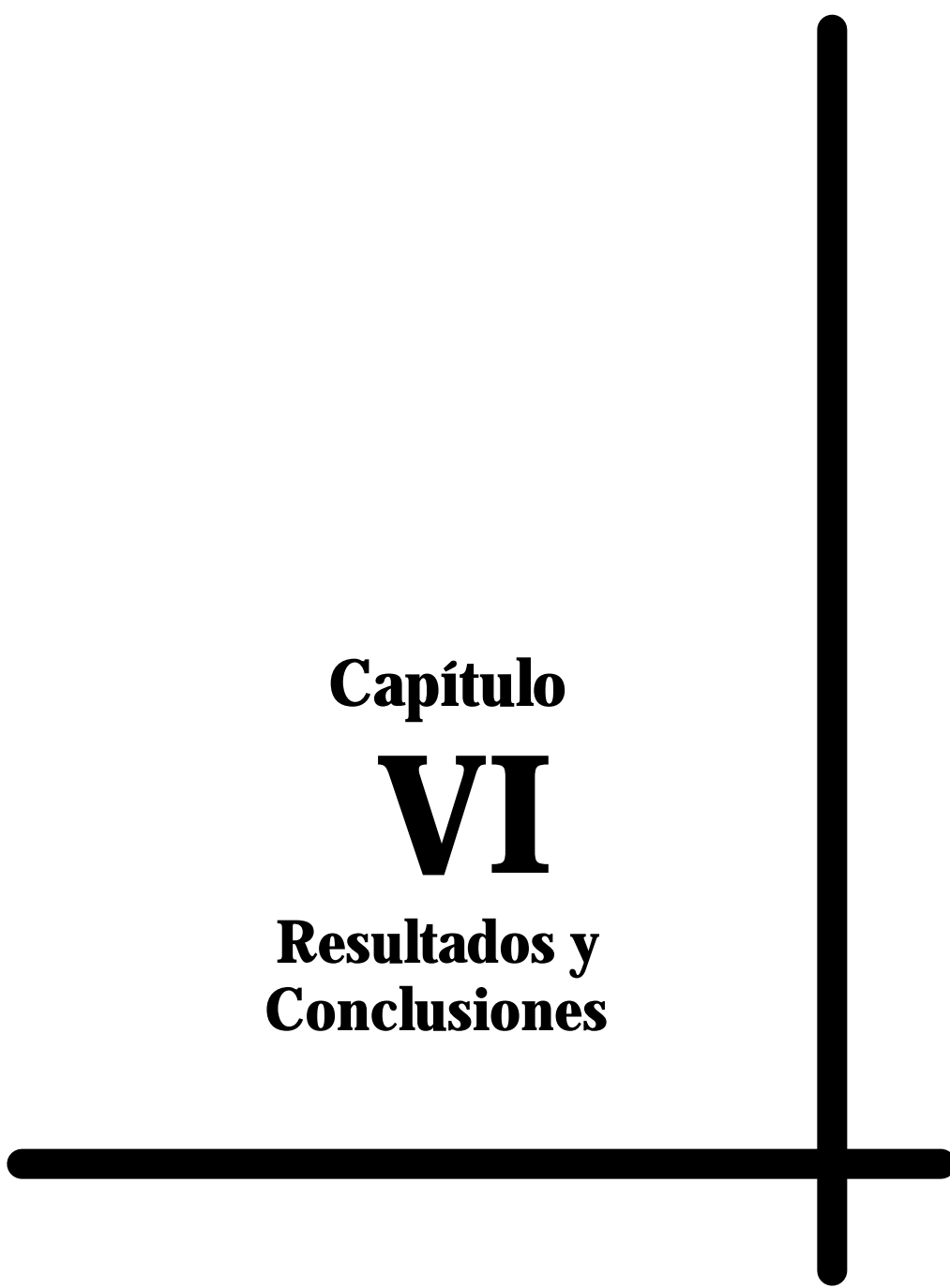


Capítulo
VI
Resultados y
Conclusiones



VI

6.1 Resultados

Aunque al principio la adquisición de imágenes se presentó de manera muy problemática, se logró adquirir manipular y procesar una secuencia de imágenes provenientes de un par de cámaras para realizar un procesamiento digital de imágenes.

El tiempo de encendido y apagado de las cámaras es de 1250 milisegundos, este tiempo fue asignado debido a que aproximadamente una cámara del modelo usado, tarda casi 1100 ms para estabilizar el sensor de la imagen y trabajar en su región lineal; por lo que se le da un margen de aproximadamente 100 ms. para capturar la imagen. Sin embargo, se debe evitar el uso de otros programas que consuman gran cantidad de recursos del sistema, de lo contrario; se puede originar un impacto en el sistema al no encenderse una o ambas cámaras en el tiempo establecido, lo que originaría un error en la salida de nuestros datos.

Al inicio de la programación del sistema, no se tenía contemplado el uso de algún tipo de realce de la imagen, pero se pudo observar que en ocasiones las características eran atenuadas en gran medida, por lo que se optó por añadir un realce de tipo contraste; el realce implantado fue de tipo exponencial sigmoïdal.

Además de la aplicación de un contraste, es posible el mejoramiento de la detección de bordes si es aplicado un filtrado de suavizamiento. El filtro es especialmente útil en situaciones cuando la imagen presenta ruido causado por efecto de iluminación deficiente por ejemplo.

Al utilizar el algoritmo de detección de bordes Sobel, el matching se hace más difícil porque las características son poco emparejables, es decir: es difícil encontrar el par conjugado de la imagen izquierda con el de la derecha y viceversa, esto es debido a que con Sobel los bordes obtenidos son gruesos; por el contrario al utilizar Canny y computar bordes de un pixel de ancho se obtienen de una manera más fácil los pares izquierdo-derecho ya que existen pocos pixeles para emparejar debido a que la búsqueda se concentra en la frontera del borde, esto resulta en un menor y más efectivo emparejamiento de pixeles.

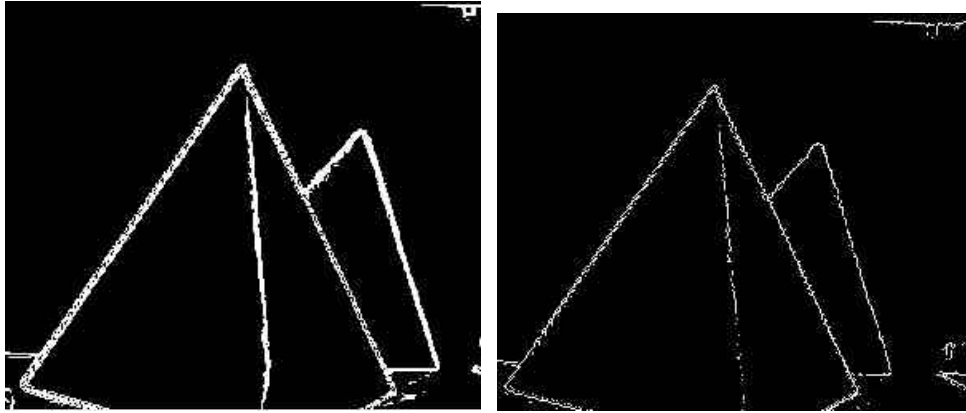


Fig. 6.1 Borde mediante Sobel y Canny

La Figura 6.1 presenta una misma figura procesada mediante algoritmos de detección Sobel y Canny, como se puede apreciar al utilizar Canny, los bordes obtenidos son delgados.

Cuando se aplica la transformada Hough y se vectorizan las características, estos vectores pueden ser graficados en la aplicación sobre las imágenes obtenidas para referencia, sin embargo; como es conocido, las rutinas de pintado y refresco de la pantalla son de mayor consumo de tiempo computacional, por lo que normalmente se prescindirá de su visualización, sin embargo mediante un parámetro en el programa, se pueden graficar estas líneas vectores para su visualización. Cabe mencionar que aunque no se dibujen estas líneas, el algoritmo realiza la operación de estimación de los vectores. De manera experimental se determinó que parámetros como el determinar ángulos cada 11.25 grados resulta bien para la mayoría de las aplicaciones y con un tiempo de corrida aceptable, se encontró que en aplicaciones donde existe una gran cantidad de características inclinadas es necesario indicar al sistema realizar el barrido cada 5.6 grados con lo cual el sistema tiene una mejor detección con un tiempo de respuesta mayor. En otros trabajos que funcionan con un esquema similar, se observó que éstos trabajan encontrando vectores aproximadamente cada 22.5 y 30 grados, nuestro sistema puede determinar ángulos de hasta 2 grados lo que lo hace mas preciso.

Otro parámetro importante con el que trabaja el sistema al vectorizar las características es, qué tan grande es necesario que sea una característica para ser considerada como vector representativo; de manera experimental se determinó que un vector de la quinta parte de la longitud total de la imagen resulta bien para determinar los vectores en la mayoría de las aplicaciones; éste parámetro también puede ser modificado llegando a necesitar en algunas ocasiones la representación de vectores de la sexta parte de la imagen.

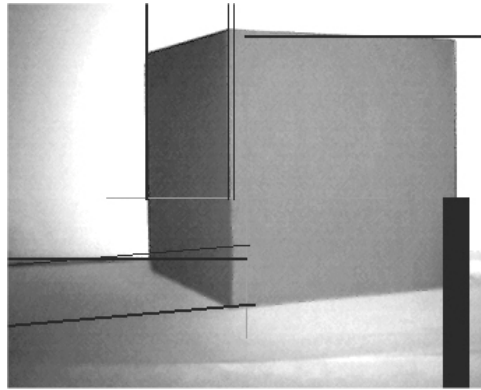


Fig. 6.2 Objeto vectorizado

En la figura 6.2 es mostrada la imagen de un objeto cuyos bordes han sido vectorizados.

El resultado del uso o no de la inclinación es como a continuación se indica: sin inclinación se obtiene aproximadamente el 80% de exactitud, mientras que al usar la inclinación la exactitud, se eleva a cerca del 98% para el algoritmo Canny.

Como se puede apreciar en la figura 5.3 con las cámaras montadas a una distancia de 7 cms. entre sí, después de 1.20 metros de distancia, la diferencia en pixeles tiende a ser casi constante, por lo que si se quiere trabajar con distancias mayores a 1.5 metros, se puede empezar a tener problemas para determinar su distancia, este problema con las cámaras montadas a una distancia de 10 o 14 cm. ocurre a una distancia mayor, por lo que al tener una mayor distancia entre las cámaras se puede manipular objetos que se encuentren a una mucho mayor distancia, en contra se tiene que objetos menores a 20 o 27 cm. Respectivamente, causan problemas al momento de determinar su disparidad, puesto que la distancia de las cámaras y su ángulo de visión no abarcan lo suficiente para que aparezcan en las dos escenas y por lo tanto hay oclusión en una de las imágenes.

La fórmula contiene un parámetro f que fue obtenido de forma experimental y al analizar los datos reales obtenidos; esta f es como a continuación se indica: una f de 630 cuando las cámaras están colocadas a 7 cm entre sí; 700 y 620 cuando se tiene una distancia entre las cámaras de 10 y 14 centímetros respectivamente.

La coordenada Z finalmente fue obtenida por medio de la fórmula planteada y se compararon otros métodos como interpolación y polinomio, la fórmula presentó algunas diferencias respecto a su comportamiento teórico, debido en parte a imperfecciones en la cámara y en la distancia del punto focal de las mismas aunque la obtención de datos por medio de fórmula resultó la forma más efectiva y es la que se recomienda para su uso.

La calidad y la cantidad de luz que ilumina la escena sobre la que se trabaja es fundamental para obtener los resultados satisfactorios con el sistema, la fuente de luz debe estar colocada frontalmente con un desplazamiento máximo de 12 grados de la normal de la vista cíclope de las cámaras; la luz natural es buena, pero también se han encontrado resultados satisfactorios usando luz artificial, la luz natural presenta mucho más problemas cuando su incidencia es mayor a los 12 grados y la luz artificial de tubos fluorescentes usada en conjunto con luz dicróica es la que presenta mayores problemas.

La distancia mínima de sensado de las cámaras con una separación de 7 centímetros entre sí es de 14 centímetros, es decir; todos los objetos que se encuentren a una distancia menor tendrán algún tipo de oclusión en alguna de sus partes y por lo tanto no podrá ser procesado de forma exitosa; la distancia mínima de sensado respecto a la separación entre sí está dada en la tabla 6.1.

Distancia camaras (cm.)	minima distancia sensado (cm.)
7	14
10	20
14	27.8
20	19.5

Tabla 6.1 Distancia cámara vs. distancia mínima de sensado

De lo anterior se puede concluir que la distancia mínima de sensado para las cámaras es aproximadamente del doble de la separación entre las mismas.

Para aplicaciones grandes en robótica, la distancia de separación entre las cámaras debe ser de al menos 10 centímetros para poder procesar de manera eficiente obstáculos de aproximadamente un metro de distancia, aunque se recomienda una separación de 14 centímetros o mayor, si los obstáculos se encuentran a una distancia mayor.

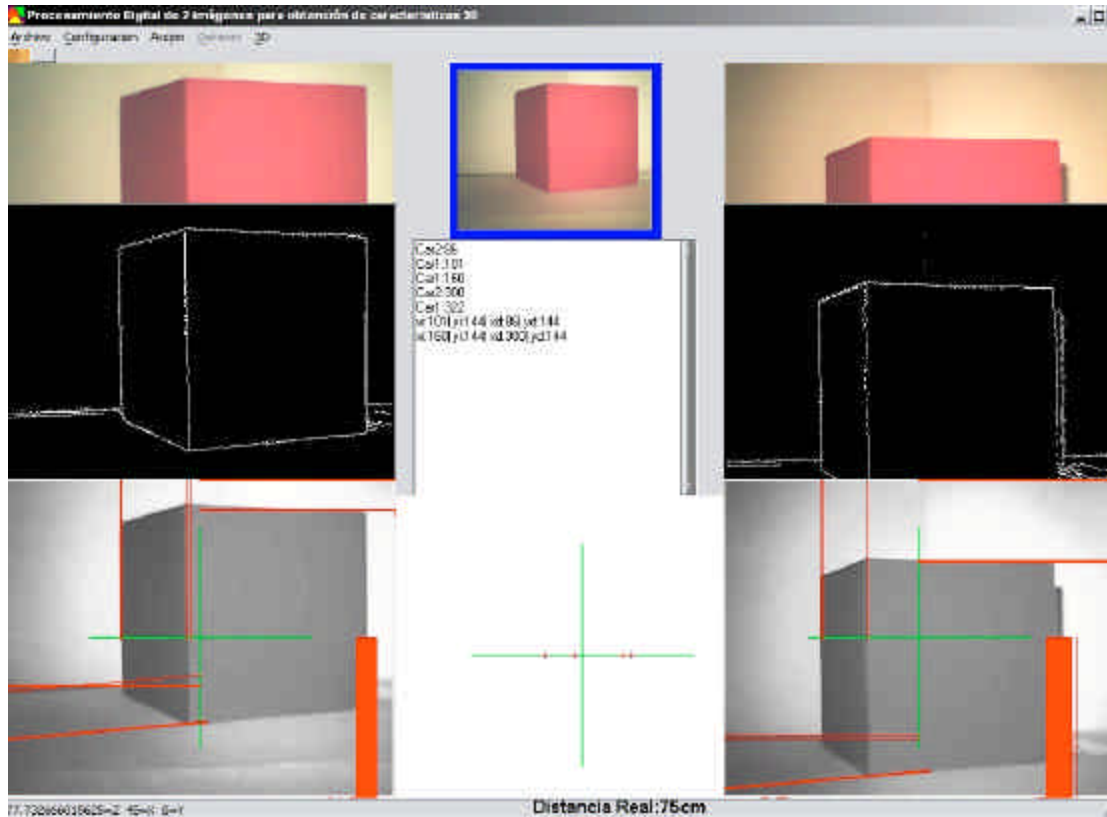


Fig. 6.3 Imágenes de corrida del sistema: Cubo

En la figura 6.3 y 6.4 se muestran algunos ejemplos de corrida del sistema y su funcionamiento, como puede apreciarse en el cubo (fig. 6.3), se detectan 2 características en la imagen izquierda y 2 características en la imagen derecha, correspondientes a las 2 líneas verticales que delimitan al cubo, estas 2 características son emparejadas y dan lugar a pares conjugados, que al aplicar las ecuaciones de transformación 2D a 3D originan las coordenadas tridimensionales que delimitan a la figura; Las coordenadas 2D aparecen en la parte central de nuestro programa y la coordenada 3D en la parte inferior, es decir se está utilizando escaneo de líneas con una separación de 144.

En la figura 6.4 se muestra una corrida del programa en donde, además se obtienen más coordenadas 3D del objeto y se hace una representación de las mismas en 2D en la parte central inferior; en este caso se realiza un análisis con una separación de 25 entre línea y línea.

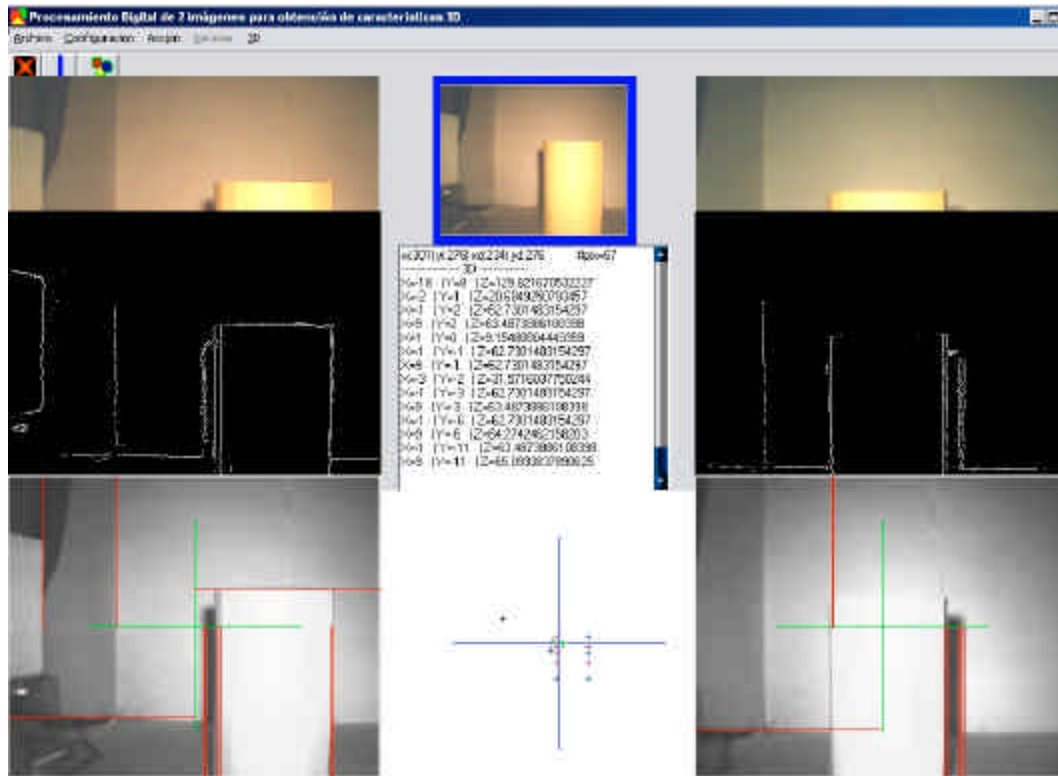


Fig. 6.4 Imágenes de corrida del sistema: Cilindro

A continuación se muestra el resultado de la corrida de nuestro sistema, en primer lugar en la Tabla 6.2 se puede observar el listado de las características encontradas tanto en la imagen izquierda (car1) como en la imagen derecha (car2) después de encontrar estas características, se procede a realizar el matching, que son los datos encontrados, para cada par encontrado se despliega la diferencia estéreo; este proceso se realiza de acuerdo a el incremento previamente establecido, en este caso se trabaja con una línea cada 25 renglones, aunque esto puede ser modificado, aquí únicamente se presenta con propósitos de demostración; finalmente se presenta el resultado de la transformación de 2D a 3D a los pares encontrados.

Car1:3	Car1:169	
Car1:70	Car2:202	
Car1:69	Car2:205	
Car2:97	Car1:271	
Car1:152	Car1:272	
Car1:154	xi:169 yi:126 xd:97 yd:126	#pix=72
Car1:157	xi:272 yi:126 xd:202 yd:126	#pix=70
Car1:158	xi:164 yi:126 xd:205 yd:126	#pix=-41
Car1:159	Car2:34	
Car1:163	Car1:70	
Car1:164	Car2:98	

Car1:151		xi:163 yi:201 xd:216 yd:201	#pix=-53
Car1:152		xi:168 yi:201 xd:98 yd:201	#pix=70
Car1:154		xi:270 yi:201 xd:96 yd:201	#pix=174
Car1:163		Car2:35	
Car1:168		Car2:96	
Car2:202		Car2:98	
Car2:204		Car1:153	
Car2:217		Car1:161	
Car2:220		Car1:163	
Car1:272		Car1:168	
xi:70 yi:151 xd:34 yd:151	#pix=36	Car2:202	
xi:168 yi:151 xd:98 yd:151	#pix=70	Car2:216	
xi:272 yi:151 xd:202 yd:151	#pix=70	Car2:220	
xi:154 yi:151 xd:204 yd:151	#pix=-50	Car1:270	
xi:163 yi:151 xd:217 yd:151	#pix=-54	xi:153 yi:226 xd:35 yd:226	#pix=118
xi:70 yi:151 xd:220 yd:151	#pix=-150	xi:161 yi:226 xd:216 yd:226	#pix=-55
Car2:35		xi:163 yi:226 xd:216 yd:226	#pix=-53
Car1:70		xi:168 yi:226 xd:220 yd:226	#pix=-52
Car2:96		xi:270 yi:226 xd:96 yd:226	#pix=174
Car2:97		Car2:97	
Car1:151		Car2:100	
Car1:154		Car1:162	
Car1:163		Car1:168	
Car1:168		Car2:202	
Car2:201		Car1:270	
Car2:202		xi:162 yi:251 xd:97 yd:251	#pix=65
Car2:204		xi:168 yi:251 xd:100 yd:251	#pix=68
Car2:218		xi:270 yi:251 xd:202 yd:251	#pix=68
Car2:221		Car2:99	
Car1:271		Car1:161	
Car1:273		Car1:167	
xi:70 yi:176 xd:35 yd:176	#pix=35	Car2:201	
xi:151 yi:176 xd:96 yd:176	#pix=55	Car2:204	
xi:154 yi:176 xd:204 yd:176	#pix=-50	Car1:268	
xi:163 yi:176 xd:204 yd:176	#pix=-41	xi:161 yi:276 xd:99 yd:276	#pix=62
xi:168 yi:176 xd:221 yd:176	#pix=-53	xi:167 yi:276 xd:201 yd:276	#pix=-34
xi:271 yi:176 xd:201 yd:176	#pix=70	xi:268 yi:276 xd:204 yd:276	#pix=64
xi:273 yi:176 xd:97 yd:176	#pix=176	----- 3D -----	
Car1:70		X=-4 Y=2 Z=62.7301483154297	
Car2:96		X=6 Y=2 Z=64.2742462158203	
Car2:98		X=-24 Y=-1 Z=133.908996582031	
Car1:152		X=-4 Y=-1 Z=64.2742462158203	
Car1:163		X=6 Y=-1 Z=64.2742462158203	
Car1:168		X=-25 Y=-6 Z=138.186019897461	
Car2:202		X=-7 Y=-4 Z=81.6794128417969	
Car2:204		X=6 Y=-3 Z=64.2742462158203	
Car2:216		X=0 Y=-1 Z=5.85598564147949	
Car2:219		X=-4 Y=-6 Z=64.2742462158203	
Car2:221		X=0 Y=-2 Z=4.91839694976807	
Car1:270		X=-5 Y=-5 Z=32.3906745910645	
xi:70 yi:201 xd:219 yd:201	#pix=-149	X=0 Y=-3 Z=4.91839694976807	
xi:152 yi:201 xd:219 yd:201	#pix=-67		

Tabla 6.2 Resultados corrida del sistema

A continuación se muestran las pantallas de corrida de nuestro sistema, en la figura 6.5 aparece la pantalla de configuración de los umbrales del sistema y sus valores típicos; en la figura 6.6 se presenta

la pantalla con las imágenes procesadas y en la Figura 6.7 se muestra la reconstrucción 3D obtenida al graficar las coordenadas 3D del objeto que en este caso corresponde a una pluma. Como se puede apreciar en la pantalla inferior central de la aplicación aparecen unas cruces donde de una forma rápida se puede observar la posición de nuestro objeto de estudio que se ha encontrado, en este caso las coordenadas finales estarán posicionadas cerca del centro de nuestro mapa. Aunque no es posible apreciarlo, la profundidad es dada mediante la coloración de las cruces. En la parte central aparecen todas las coordenadas 3D calculadas.

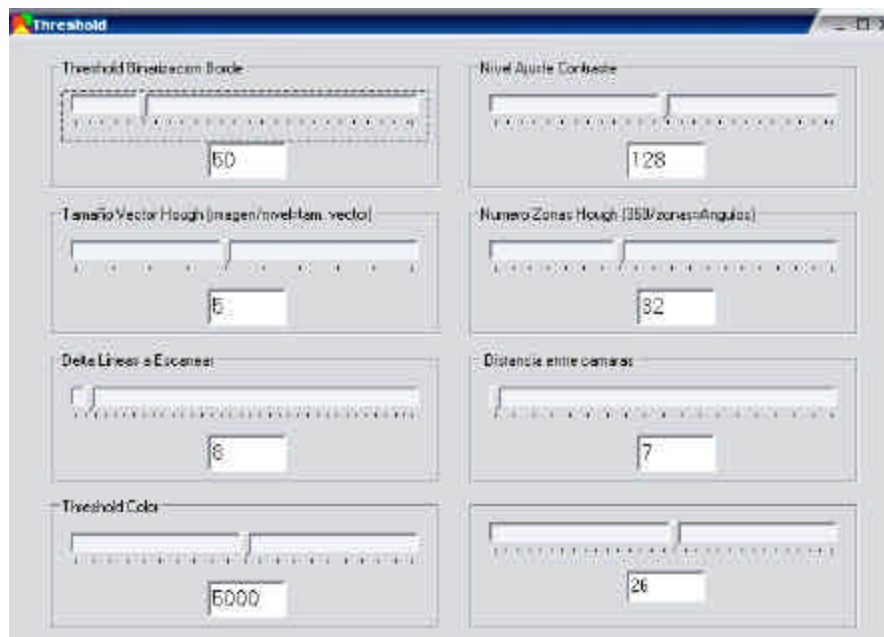


Figura 6.5 Pantalla Umbrales

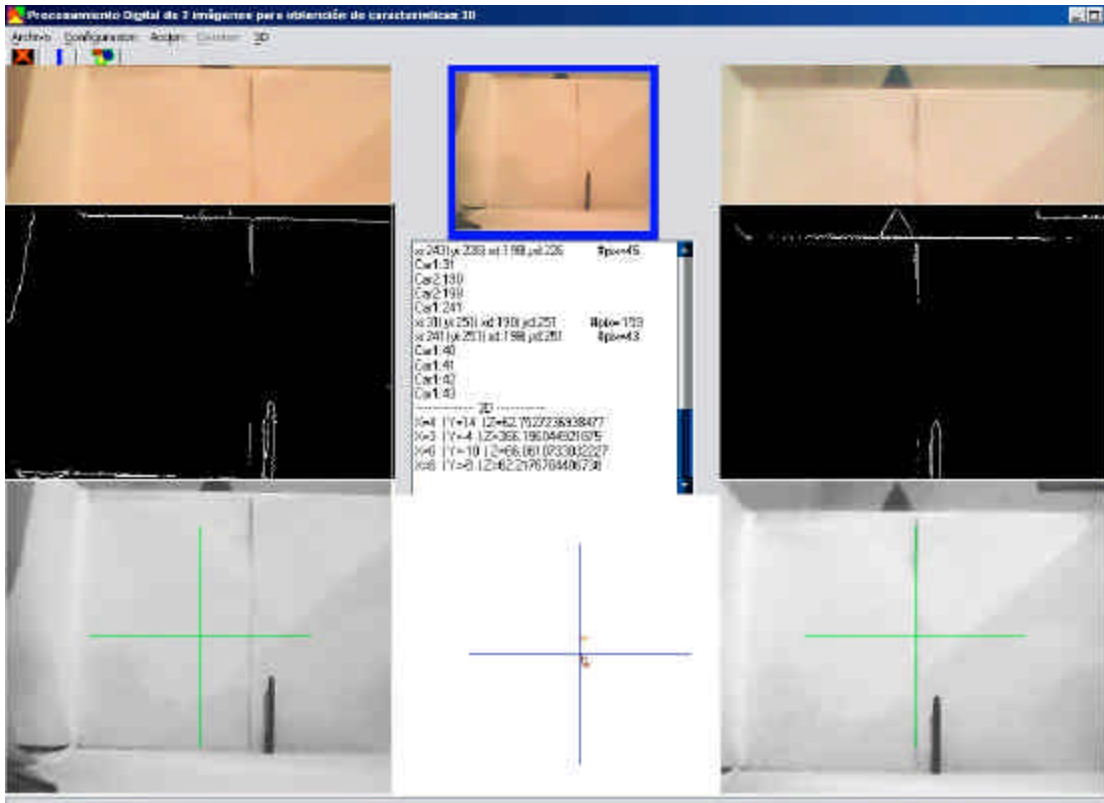


Fig. 6.6 Corrida del sistema

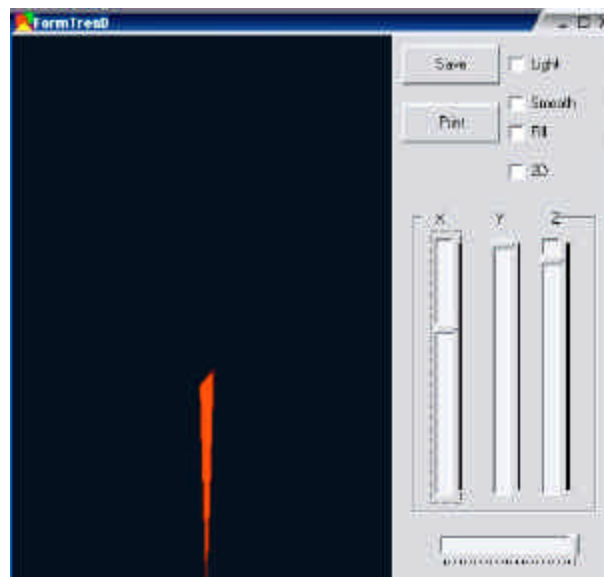


Fig. 6.7 Pantalla representación 3D

En la figura 6.8 es posible apreciar la corrida del sistema al procesar una pirámide triangular y es mostrada su correspondiente aproximación 3D con los datos obtenidos (ver fig. 6.9).

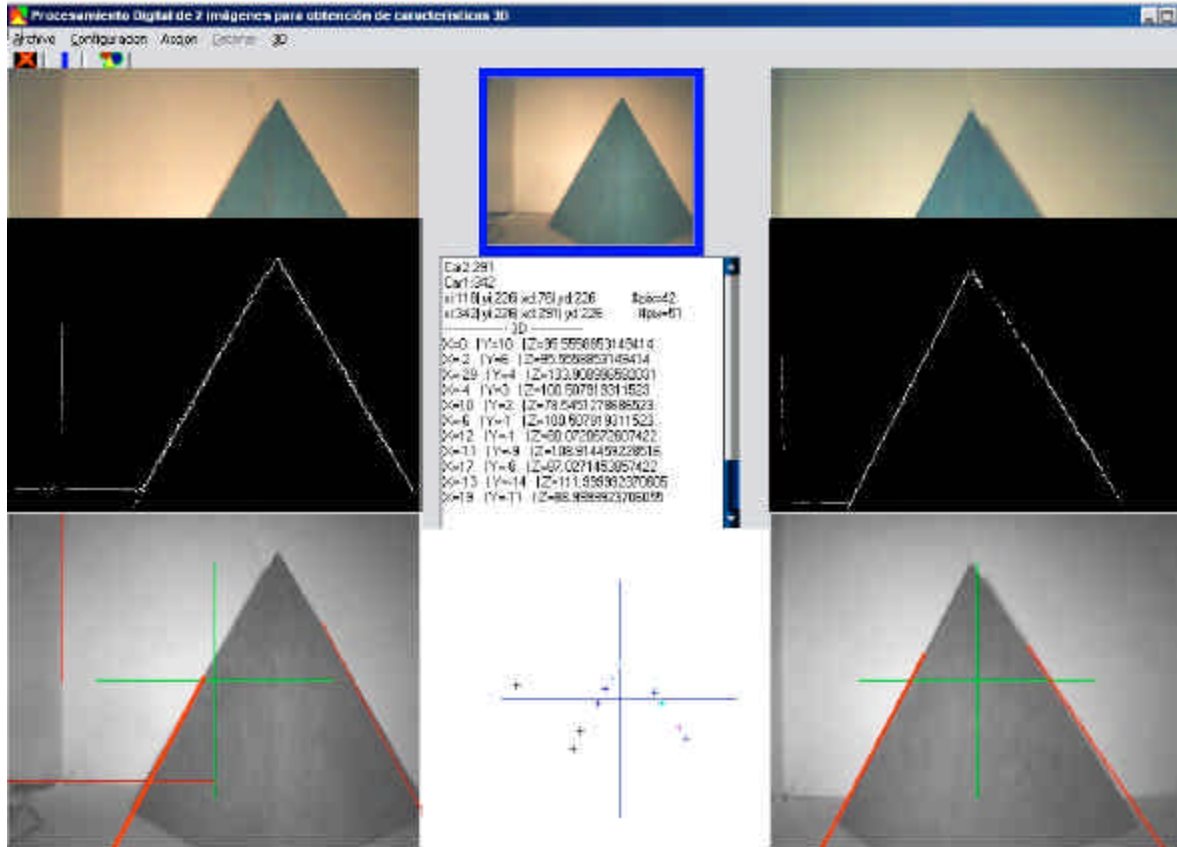


Fig. 6.8 Corrida del sistema: Pirámide

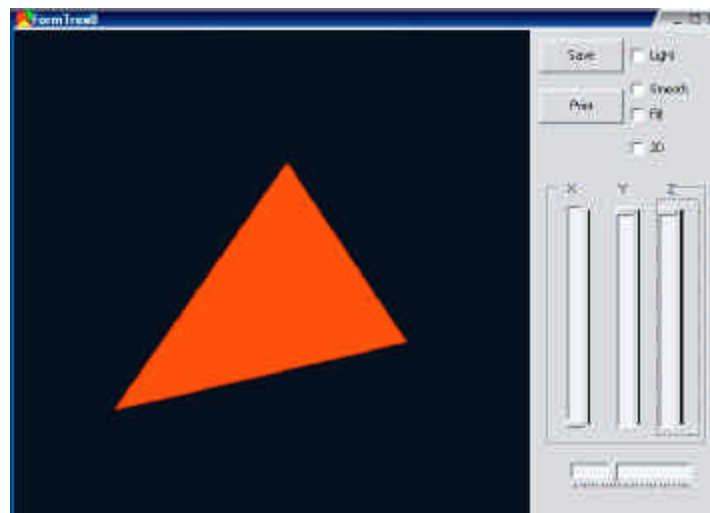


Fig. 6.9 Representación 3D: Pirámide

Se estudiaron varias formas para la representación gráfica de los datos obtenidos, entre las herramientas consideradas estuvo el uso de Java3D, las rutinas de graficado normal de C++ y OpenGL,

finalmente OpenGL fue usada para la representación por su facilidad de uso, estandarización y calidad de los resultados.

6.2 Conclusiones

Los objetivos fueron cumplidos, ya que a través de la manipulación y adquisición de secuencias de imágenes se implantaron métodos para obtener datos del ambiente que se explora y representa y por medio de la información así obtenida se elabora su representación en un mundo virtual 3D y estos datos pueden ser utilizados para el desplazamiento de autómatas.

Como primer paso se llevó a cabo una investigación y valoración sobre métodos y técnicas empleados para la percepción del entorno de autómatas, llegándose a la conclusión de que aunque los sistemas de infrarrojo y de ultrasonido son muy buenos puesto que no requieren de contacto físico y su precio es reducido, el futuro está en sistemas como los de visión con los cuales se puede obtener una mucho mayor cantidad de información de manera más rápida.

Se realizó una valuación de la selección de herramientas a utilizar en el sistema encontrándose a C++ Builder una herramienta muy efectiva para el desarrollo de este tipo de sistemas; su desventaja: la portabilidad, aunque existe la posibilidad de ser emigrado a Kylix (compilador de C++ Builder para S.O. basados en Unix); los dispositivos de entrada resultaron suficientes para propósitos de demostración aunque para aplicaciones críticas es necesario contar con mejores dispositivos de entrada.

Se hizo una investigación, implantación y evaluación exhaustivo de los distintos algoritmos de procesamiento digital de imágenes necesarios y se escogieron los que presentaron mejores resultados para obtener características tridimensionales de los objetos del ambiente; llegándose a la conclusión de que: en cuanto a el proceso de pre-procesamiento, un bloque de contraste es sumamente útil, y es mucho mas provechoso entre mas similares son los objetos que se analizan, el filtrado resulta bueno aunque podría prescindirse de su uso; en el caso de los algoritmos detectores de bordes, el tipo de algoritmo seleccionado es de vital importancia, encontrándose que aquellos que tengan contornos gruesos pueden resultar problemáticos en el momento de realizar el matching por tener mas características que comparar, es por eso que es recomendado el uso del algoritmo Canny.

Otro tipo de procesamiento que fue analizado para su implantación: La Transformada Hough, que resultó útil para propósitos de determinación de la inclinación de las características, lo cual resulta de peso para determinar el emparejamiento, y para las pruebas realizadas resulta en una mejora del sistema, aunque los resultados obtenidos no distan mucho del emparejamiento sin utilizar la inclinación.

Se realizó investigación en cuanto a visión estéreo y algoritmos de emparejamiento, y se realizó la implantación de un sistema estéreo lo más robusto posible considerando una variedad de técnicas de emparejamiento, aunque no se llega a evaluar la totalidad de los mismos y se realiza la evaluación tomando en cuenta cámaras paralelas; No es posible llegar a una conclusión única porque todavía es necesario terminar de evaluar el sistema en una mayor cantidad de pruebas y ambientes, aunque para las pruebas realizadas el emparejamiento considerando pixeles vecinos y con inclinación resultó el mas satisfactorio.

Para una mejor representación de los resultados obtenidos se consideró a OpenGL como una forma sencilla y práctica además de ser una tecnología que actualmente se encuentra ampliamente en uso, así que una vez obtenidas las coordenadas 3D por medio de Open GL se realiza el modelado del ambiente reconocido; cabe aclarar que este modelado es únicamente con fines demostrativos más que de investigación y el uso de OpenGL es más que recomendado.

Ahora bien, comparando el desarrollo de este sistema con trabajos previos y trabajos que se están desarrollando, a nivel nacional poco se ha desarrollado en el campo del análisis volumétrico de los objetos mediante visión estéreo, sin embargo existen importantes trabajos en cuanto a procesamiento digital de imágenes y su uso en autómatas, y uno de estos trabajos lo tenemos en esta misma universidad, J. A. Chavez estudia técnicas para la exploración de ambientes de un robot móvil valiéndose del procesamiento de imágenes se lleva a cabo una implementación en donde se realiza el control de un robot Khepera para evadir obstáculos y además se dibuja en forma de plano 2D, pero debido a que el sistema utiliza una orientación totalmente distinta a la nuestra no es posible analizarlo desde un punto de vista cuantitativo.

Este sistema presenta una mayor velocidad de detección de obstáculos contra otros tipos de formas de construcción de mapas y detección de obstáculos como es el caso de métodos que involucran el uso de sensores de contacto o de borde de objetos con ayuda de infrarrojo por ejemplo.

Ahora bien, existen trabajos en universidades como Rochester, Purdue, y Carnegie Mellon, por citar algunas, donde en general su investigación está orientada a dotar a autómatas de información de su entorno por medio de un análisis digital de imágenes provenientes de un par de cámaras generalmente a color y, además se ayudan de equipo adicional como es el caso de sensores de ultrasonido y sensores de rango láser. (Ver Cap. 1). En esta misma dirección se encuentra trabajando Honda para dotar a sus robots de visión estéreo. Es decir, el análisis de la percepción del robot es realizado mediante visión; Por lo anterior se concluye que esta orientación en la investigación es el camino correcto a seguir en cuanto a visión por computadora estéreo. Sin embargo, aun la información existente apenas empieza a darse a conocer por ser un área de reciente estudio.

Existen varias restricciones en el sistema hasta el momento, una de ellas es que si el movimiento es rápido, debido a la poca velocidad de cambio de las cámaras y a que su tiempo de encendido es relativamente lento conduce a que las cámaras no pueden llevar el ritmo de cambio si este sucede rápidamente.

Como se había planteado desde el inicio de la tesis, los objetos que aquí se analizan y reconocen son muy simples, cubos, pirámides, etc. por ejemplo y presenta algunos problemas con ambientes complejos.

Este sistema cumple con gran parte de las expectativas iniciales, aunque gran parte de su mejora radica en la calidad de los sistemas de adquisición de datos, otra mejora en el sistema se podría dar si se aplicara un algoritmo específico para el reconocimiento, aunque hay que señalar el tiempo extra de procesamiento que implica el implantar un algoritmo de este estilo.

Se llegó a la conclusión de que este sistema puede ser utilizado para procesar objetos con características simples y detectar sus coordenadas tridimensionales reales, esta información de esta forma puede ser utilizada para la predicción de trayectorias en autómatas que requieran el uso de este tipo de información tridimensional, esto sin hacer uso de sensores de tipo de presencia o infrarrojos y puede significar un ahorro en el tiempo de desplazamiento y control.

Si es requerido trabajar en campos donde el área de trabajo se encuentra en un área de 30 cm. a 3 metros aproximadamente, es decir en aplicaciones generales de robótica, es recomendable el uso de las

cámaras montadas a una distancia entre sí de 14 cm. o mayor si es que el área de trabajo se sitúa a una mayor distancia.

Como resultado obtenido aunque no planteado, se presenta la posibilidad de ser una punta de lanzamiento para investigación en el procesado estéreo de imágenes a nivel local y es un punto de partida para el establecimiento de un departamento de visión aquí en la Universidad de las Américas.

6.3 Mejoras y Desarrollos posteriores

Cabe mencionar que se investigó el uso de dispositivos de entrada de captura de video, así como la programación en otros sistemas operativos que no fueran Microsoft; el principal sistema operativo alternativo fue un Sistema Operativo Linux, pero en realidad nunca se llegó a la implantación; por lo que queda como desarrollo la implantación y la comparación de las rutinas corriendo en un sistema operativo basado en Unix.

El algoritmo Canny fue implantado utilizando un solo nivel de umbral y como mejora del mismo se encuentra el implementar el doble threshold y su encadenación entre las dos imágenes, el resultado obtenido con este tipo de detector de borde es muy bueno y los resultados han sido expuestos. Sin embargo, la implantación conducirá a una mejor detección de los bordes en las imágenes.

Asimismo se pueden investigar e implementar otros algoritmos u otras variaciones de los algoritmos matching utilizados, esto se debe a que es una gran área que ha sido poco estudiada y explotada, que repercutiría en el desempeño del sistema y sobre lo cual aun existe escasa información. También queda como investigación futura el uso de otro tipo de contrastes y filtros que resulten en una mejora al sistema.

Como mejora a esta investigación se propone el uso de tarjetas digitalizadoras de entrada de video, que aunque son de un costo relativamente elevado, en caso necesario podrían ser utilizadas con el efecto de tener mejores resultados en cuanto a calidad y tiempo de captura de las imágenes.

Puesto que el tiempo de procesado de una imagen es mucho menor a los 1250 ms. empleados para encender una u otra cámara, como trabajo futuro se plantea el estudio del tiempo de encendido de

cámaras de otras compañías y modelos, esto con el fin de el ahorro de costosas tarjetas digitalizadoras de video.

En cuanto a dispositivos de entrada, se propone la mejora del sistema al utilizar no cámaras paralelas sino cámaras que puedan ser rotadas de acuerdo a la posición del objeto de interés. Además, se propone el uso de un sistema de estas características junto con sistemas auxiliares que pueden aportar una mayor cantidad de información como es el caso del láser y el ultrasonido.

Un trabajo actualmente en desarrollo es el que además de determinar la forma del objeto se determine el color del mismo.

Un objetivo secundario planteado fue explorar nuevas formas para el reconocimiento que no precisaran del uso de redes neuronales, aunque sería de interés el adaptar algún tipo de algoritmo que permita el mejor y más variado reconocimiento de objetos y se considera el estudiar e investigar el uso y el impacto de una red neuronal para este objetivo.

Lo que vale la pena hacer, vale la pena hacerlo bien. Y
Siempre hay una manera de hacer mejor las cosas