

Documento Expression of Interest (EOI) de la generación de trayectoria de un objeto mediante visión artificial.

Sergio Uribe Sáenz de Cámara

Email: suribe003@ikasle.ehu.eus

Alumno de Grado de Ingeniería en Tecnología Industrial
Escuela de Ingeniería de Bilbao

Abstract:

The following project consist on the development of a software tool with the aim of making the tracking 3D of a moving object based on the video images obtained from two cameras (stereo vision). This project was born with the purpose of monitoring aquatic organisms in the Marine Station of Plentzia.

1. Introducción

La visión artificial es la disciplina científica que trata de adquirir, procesar y analizar imágenes del mundo real con el objetivo de que éstas puedan ser tratadas por un computador. El fin último de esta disciplina es recrear el mundo real, ya sea mediante reconocimiento y detección de objetos, reconstrucción de la escena 3D y restauración de imágenes.

El problema reside en que la escena real es tridimensional, mientras que las imágenes captadas son bidimensionales. Es decir, en el paso de la realidad a la imagen hemos perdido una dimensión. Se emplea la visión estereoscópica, que es un procedimiento para la obtención de esa tercera dimensión, para dar solución a dicho problema.

2. Contexto

Estando tanto las cámaras como los computadores totalmente implantados en nuestra sociedad, la visión artificial y estereoscópica es un campo de investigación que cada día va a más.

Una analogía entre sistema artificial y humano sería el siguiente: el sistema visual humano percibe la escena 3D y el cerebro procesa esa información. Los ojos son un sistema estereoscópico o de ejes ópticos paralelos. Por tanto, mediante la visión estereoscópica artificial se intenta imitar esta capacidad natural del ser humano.

Esto es posible gracias al avance en las tecnologías. Así, es posible encontrar diversas bibliotecas de Visión Artificial, destacando sobre todo OpenCV. Además años atrás el lenguaje utilizado para la programación era el lenguaje C, mientras que en la actualidad hemos dado paso a lenguajes de más alto nivel como Matlab, Python o Java, siendo el primero el entorno favorito para aplicaciones de Visión Artificial.

Respecto a las cámaras, también se ha producido un gran avance. Han aparecido cámaras de grabaciones y toma de fotografías de deportes extremos, como las GoPro. Son cámaras de acción, ligeras, resistentes y pueden colocarse en vehículos, manillares, cascos y otros soportes, lo que facilita su uso, incluso dentro del agua con su debida protección.

3. Objetivos y Alcance

Siendo la orientación de la visión estereoscópica hacia la navegación autónoma de vehículos, entre otros, se observa fácilmente la repercusión que va a tener esta disciplina en los próximos años.

Por tanto, este proyecto va a consistir en realizar una programación capaz de detectar un objeto en movimiento, calcular la distancia a la que se encuentra dicho objeto de las cámaras mediante relaciones trigonométricas y posteriormente, reconstruir la escena para poder realizar el seguimiento del objeto.

Finalmente, se valorará mediante un estudio la viabilidad de ampliar el sistema a 4 cámaras. Por otro lado, este proyecto puede servir como base en investigaciones en el campo de la Visión Artificial.

4. Propuesta Técnica

Mediante dos cámaras GoPro Hero 3 se grabará un video con un único objeto en movimiento. Mediante el programa MPEG Streamclip se sacarán las imágenes o frames con los que se trabajará posteriormente en Matlab, el cual es la herramienta matemática adecuada para el estudio de este ámbito, debido a su facilidad para la manipulación de matrices, que como veremos posteriormente, una imagen se define como una matriz. Además, esta herramienta de software resulta idónea por incorporar diversas Toolbox o herramientas para procesado de imágenes y visión por computador.

5. Plan de Trabajo

El primer paso para comenzar el proyecto será, como ya se ha mencionado con anterioridad, grabar un vídeo en el que se observe sin confusión alguna un único objeto en movimiento. A partir del video obtendremos los frames, y con ellos ya podremos trabajar en Matlab.

Con la ayuda de la "Stereo Camera Calibrator App", se realizará una minuciosa calibración, de forma que se obtenga los parámetros estéreo de las cámaras. Una vez hecho esto, se rectificarán todas las imágenes para eliminar la diferencia en el eje y de cada par de imágenes (de las cámaras derecha e

izquierda) -se explicará con detalle este problema a lo largo de este informe-. Posteriormente, después de restar cada imagen a su anterior, se convertirán a blanco y negro, de tal forma que nos quedemos únicamente con el movimiento en color blanco (imagen binaria). Tendremos que aplicar diversos filtros morfológicos para quedarnos únicamente con nuestro objeto de interés.

El siguiente paso es obtener el centroide de nuestro objeto y almacenar sus coordenadas para posteriormente poder calcular a partir de relaciones trigonométricas la distancia del objeto a las cámaras. Para este último paso, habrá que implementar diversas fórmulas que relacionan el mismo punto (centroide del objeto) pero que en las imágenes (izquierda y derecha) del mismo instante, aparecen las coordenadas del centroide con un desplazamiento relativo en el eje x.

Obtenida la distancia del objeto a las cámaras, procederemos a reconstruir la escena.

Por último, cabría valorar las mejoras que se le podría añadir a este proyecto, como puede ser la incorporación de dos cámaras más para tener una reconstrucción de la escena más definida.

6. Conclusiones

Se espera que este proyecto sirva de base para futuras investigaciones en el área de la visión artificial, en especial de la detección de objetos y seguimiento de ellos.

Es un proceso un tanto innovador, que puede que se imponga a otros procesos para realizar el mismo fin, pero consideramos que este tiene ciertas ventajas respecto a los otros, sobre todo en coste que es lo primordial en el mundo empresarial.

Referencias

- [1] W. E. L. Grimson. «Computational experiments with a feature based stereo algorithm».
- [2] Ingemar J. Cox and Sunita L. Hingorani and Satish B. Rao. «Computer Vision and Image Understanding».
- [3] E. Trucco and A. Verri. «Introductory Techniques for 3-D Computer Vision».



Figura 1: Imagen del objeto de interés. Cámara izquierda.



Figura 2: Imagen del objeto de interés. Cámara derecha.



Figura 3: Anáglifo de las imágenes anteriores