

Seguimiento de Objetos en Secuencias de Imágenes RGB-D

Propuesta de Tesis

18 de octubre de 2014

Alumno: Mariano Bianchi

marianobianchi08@gmail.com

Director: Francisco Gómez Fernández

fgomezf@gmail.com

1. Motivación

En la actualidad, las posibles aplicaciones de métodos de seguimiento o tracking son muchas y van desde el uso en la industria hasta juegos de consola. Un ejemplo de ello es la fabricación de barcos y autos mediante el uso de robots. Estas tareas se caracterizan por la necesidad de posicionar de manera precisa una herramienta sobre una pieza de trabajo. A través del uso de métodos de tracking se puede conocer la posición y pose de la pieza que se desea utilizar con respecto a la pose de la cámara y de esta forma saber cómo ubicar la herramienta necesaria para trabajar sobre la pieza en cuestión.

Otra área en donde se utiliza tracking de objetos es para la generación de estadísticas durante un partido de fútbol, tanto de jugadores como de un equipo, aunque las posibles aplicaciones en este contexto son mucho más amplias, como por ejemplo análisis de tácticas, verificación de las decisiones del árbitro, resúmenes automáticos de un partido, etc.

Actualmente existen sensores de profundidad que en conjunto con una cámara RGB pueden ser utilizados para detectar y seguir a una o más personas en tiempo real. De esta manera, mediante un sistema que procese las imágenes RGB-D de estos sensores, las personas puedan utilizar su cuerpo y sus movimientos para interactuar naturalmente con un dispositivo.

La utilización de sensores RGB-D se ha popularizado en los últimos años, cobrando un gran interés científico el estudio de aplicaciones y métodos capaces de procesar y entender la información que los mismos proveen.

La información de profundidad que nos provee un sensor RGB-D es un dato fundamental que nos posibilita encontrar la distancia de un objeto al sensor pudiendo recuperar su información 3D (tridimensional) junto a su textura RGB en tiempo real: 30 cuadros por segundo. El video RGB-D que se obtiene provee una gran ayuda al mejoramiento y desarrollo de nuevas técnicas de procesamiento de imágenes y video ya conocidas. En particular, es de interés en esta tesis, el seguimiento de objetos en secuencias de imágenes RGB-D.

Un sistema de seguimiento se puede dividir en tres etapas bien definidas:

1. Entrenamiento
2. Detección
3. Seguimiento cuadro a cuadro

La etapa de entrenamiento consiste en obtener una representación del objeto al cuál se pretende seguir. Para llevarla a cabo se puede utilizar un patrón (template) ya conocido o aprenderlo de imágenes capturadas del mismo objeto. Este template luego se utiliza en la detección para ubicar la representación del objeto dentro de una imagen cualquiera. Una vez conocido el template no se requiere de una nueva ejecución del entrenamiento.

La segunda etapa, la de detección, radica en encontrar dentro de un frame del video al objeto en cuestión utilizando el método de detección deseado, valiéndose de la información registrada en la etapa de entrenamiento. Esta etapa se ejecuta, con el propósito de encontrar en la imagen el objeto a seguir, al comienzo del sistema de seguimiento y cuando el seguimiento cuadro a cuadro falla. Dado que la etapa de detección suele ser la más costosa en términos de desempeño computacional es deseable que se ejecute la menor cantidad de veces

Finalmente, la tercera etapa consiste en seguir cuadro a cuadro el objeto detectado en la etapa anterior. Es decir, teniendo la ubicación del objeto en un cuadro de video se desea identificar la posición del mismo objeto en el siguiente frame. Esta etapa es la más importante ya que es la que se ejecuta en cada frame del video. La eficiencia del método de seguimiento es lo que determinará que todo el sistema de seguimiento se consiga realizar eficientemente. Si la técnica de seguimiento tiene una efectividad baja, es decir, no logra identificar la nueva posición del objeto en el siguiente cuadro, se debe volver a la etapa de detección cuyo desempeño computacional es mayor.

2. Objetivos

El objetivo principal de esta tesis es la implementación, estudio y evaluación de un sistema de seguimiento de objetos en secuencias de imágenes RGB-D, con las siguientes características:

- Performance Real-time: procesamiento de imágenes mayor a 10 cuadros por segundo
- Seguimiento de objetos tridimensionales con forma conocida previamente y de objetos aprendidos mediante una fase de entrenamiento previa
- Funcionamiento en sensores de profundidad de bajo costo (Kinect, XTion, etc.)

3. Metodología y Antecedentes

Para poder cumplir con los objetivos de esta tesis se comenzará con el desarrollo e implementación del artículo [PLW11] el cual implementa las tres etapas de un sistema de seguimiento de objetos 3D de la siguiente manera:

1. Entrenamiento: obtención del modelo 3D a seguir y calibración
2. Detección: usando DOT [HLI⁺10] adaptado a 3D
3. Seguimiento 3D cuadro a cuadro: usando ICP [Zha94, BM92] alineando las nubes de puntos del modelo y la detección

La etapa de entrenamiento consiste en obtener una representación tridimensional del objeto al cuál se pretende seguir. En el artículo [DC99] se utiliza un entrenamiento off-line que consiste en obtener un modelo CAD (computer-aided design) del objeto que se desea seguir. Luego, en el artículo [PLW11] se presenta una etapa de entrenamiento novedosa que se realiza de manera on-line, en donde utiliza un marcador conocido para definir las coordenadas de los objetos y calibrar la cámara.

La primera etapa del sistema puede ser prescindible si contamos con el modelo 3D del objeto a seguir y una cámara calibrada. Este es el caso de estudio de esta tesis, ya que, con el propósito de poder evaluar cuantitativamente el seguimiento de objetos en secuencias de imágenes RGB-D, utilizaremos la base de datos [LBRF11] la cual nos provee de información de ground truth sobre el posicionamiento de los objetos cuadro por cuadro en video RGB-D. Por otro

lado, se evaluará la posibilidad de implementar la fase de entrenamiento on-line presentada en el artículo [PLW11] y la evaluación cualitativa del seguimiento de objetos 3D.

Sobre la implementación desarrollada se propone modificar la etapa de detección utilizando otros métodos conocidos en la literatura [Bru09, KRTA13] con el fin evaluar el desempeño del algoritmo DOT y sus oportunidades de mejora.

La etapa de seguimiento 3D cuadro a cuadro es la más importante y de la que depende el éxito o fracaso de todo el sistema de seguimiento. La utilización del algoritmo ICP [Zha94, BM92] para esta tarea resulta natural e intuitiva. Por ello, es que en esta tesis se estudiará el algoritmo ICP y sus variantes [EBW04, SHT09], con el fin de evaluar cómo sus parámetros afectan cuantitativamente al sistema de seguimiento y la performance computacional del mismo. Asimismo, se evaluará la adaptabilidad del filtro de Kalman [WB95] para seguimiento de objetos 3D en imágenes RGB-D con posibilidad de desempeño en tiempo real. El filtro de Kalman es un filtro muy popular y estudiado extensivamente en la literatura [JU97, WVDM00] debido a su gran desempeño para realizar seguimiento en imágenes 2D. Por lo tanto, su aplicación en seguimiento de objetos 3D resulta de especial interés.

Referencias

- [BM92] P.J. Besl and Neil D. McKay. A method for registration of 3-d shapes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(2):239–256, 1992.
- [Bru09] Roberto Brunelli. *Template matching techniques in computer vision: theory and practice*. Wiley. com, 2009.
- [DC99] Tom Drummond and Roberto Cipolla. Real-time tracking of complex structures with on-line camera calibration. In *BMVC*, pages 1–10. Citeseer, 1999.
- [EBW04] Raúl San José Estépar, Anders Brun, and Carl-Fredrik Westin. Robust generalized total least squares iterative closest point registration. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004*, pages 234–241. Springer, 2004.
- [HLI⁺10] Stefan Hinterstoisser, Vincent Lepetit, Slobodan Ilic, Pascal Fua, and Nassir Navab. Dominant orientation templates for real-time detection of texture-less objects. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2257–2264. IEEE, 2010.
- [JU97] Simon J Julier and Jeffrey K Uhlmann. New extension of the kalman filter to nonlinear systems. In *AeroSense’97*, pages 182–193. International Society for Optics and Photonics, 1997.
- [KRTA13] S. Korman, D. Reichman, G. Tsur, and S. Avidan. Fast-match: Fast affine template matching. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 2331–2338, 2013.

- [LBRF11] Kevin Lai, Liefeng Bo, Xiaofeng Ren, and Dieter Fox. A large-scale hierarchical multi-view rgb-d object dataset. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 1817–1824. IEEE, 2011.
- [PLW11] Youngmin Park, Vincent Lepetit, and Woontack Woo. Textureless object tracking with online training using an rgb-d camera. In *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*, pages 121–126. IEEE, 2011.
- [SHT09] Aleksandr Segal, Dirk Haehnel, and Sebastian Thrun. Generalized-icp. In *Robotics: Science and Systems*, volume 2, page 4, 2009.
- [WB95] Greg Welch and Gary Bishop. An introduction to the kalman filter, 1995.
- [WVDM00] Eric A Wan and Rudolph Van Der Merwe. The unscented kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. ASSPCC. The IEEE 2000*, pages 153–158. IEEE, 2000.
- [Zha94] Zhengyou Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994.