

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería



**Algoritmos de Visión por computadora para el reconocimiento
de la pose de agentes empleando programación orientada a
objetos y multi-hilos**

Protocolo de trabajo de graduación presentado por José Pablo Guerra
Jordán, estudiante de Ingeniería Electrónica

Guatemala,

2020

Resumen

Este proyecto busca implementar un algoritmo de visión por computadora de una manera eficiente utilizando la programación multi-hilos. Esto se busca mediante el análisis y mejora de los algoritmos ya planteados en el lenguaje C++ y hacer sus respectivas comparaciones con el lenguaje de programación Python, mediante pruebas como la medición de tiempos de ejecución, uso del CPU, entre otras.

El objetivo principal de esto es poder reconocer la pose de agentes (denominados así normalmente en el área de robótica de enjambre) para poder unificarlo con una mesa de pruebas para futuros proyectos de investigación. Esta mesa será armada para poder realizar pruebas en la calibración de la cámara y que permita identificar los puntos de mejora que necesiten una aplicación de una programación multi-hilos.

Antecedentes

Visión por Computadora Aplicado a Robótica de Enjambre

En el documento de tesis de doctorado escrita por Luis Antón Canalís[1], titulada "Swarm Intelligence in Computer Vision: an Application to ObjectTracking", describe como la visión por computadora se puede aplicar a la robótica de enjambre y a la orientación de sus agentes.

La robótica de enjambre va aplicada a la emulación del comportamiento de las colonias (por ejemplo de hormigas o abejas) con algoritmos por computadora. La mayoría de estos algoritmos han utilizado como herramientas diferentes algoritmos de visión por computadora. Esto significa que, por ejemplo, las imágenes emulan terrenos o espacios donde las colonias virtuales las recorren buscando información significativa basada en el procesamiento obtenido por visión por computadora.

Además de eso, en el documento "Visión artificial y comunicación en robots cooperativos omnidireccionales"[2] se detalla también la importancia del uso de (a lo que los autores se refieren como visión artificial) visión por computadora también para el reconocimiento de pose de agentes. Se denota el uso de Python como un lenguaje útil y simple para la programación, así como el uso de la librería OpenCV para la aplicación de algoritmos para visión por computadora.

Continuación Fase II

Este trabajo es la fase 2 del Proyecto de Graduación propuesto por André Rodas titulado "Desarrollo e implementación de algoritmo de visión por computador en una mesa de pruebas para la experimentación con micro-robots móviles en robótica de enjambre"[3]

La fase anterior consistió en la realización de un algoritmo que permitiera reconocer objetos dentro de una cama de pruebas para ser aplicada en robótica de enjambre. Dicha implementación se realizó utilizando OpenCV y el lenguaje C++, con el objetivo de obtener el mayor rendimiento de procesamiento, aunque también se realizó un análisis de otros

posibles candidatos de lenguajes para su uso.

El objetivo es poder identificar objetos dentro de las mesas de pruebas. Para esto, lo que se realizó fue una serie de pruebas para determinar el mejor identificador para el reconocimiento de objetos en dichas mesas. Agregado a esto, se realizaron pruebas con diferentes métodos de procesamiento de la librería OpenCV para obtener los mejores resultados y poder comparar con cuales de estos se obtenía el mejor margen de reconocimiento y procesamiento de la imagen en la mesa de pruebas.

Justificación

Hoy en día, las herramientas computacionales son de gran ayuda para las distintas actividades que se realizan en diferentes áreas, tanto de la industria como en las ramas científicas.

Una de estas herramientas es la visión por computadora, que está enfocada en el uso de algoritmos para el procesamiento de imágenes, dándole la capacidad a la computadora de reconocer datos significativos que pueden ser orientados a diferentes aplicaciones. Su principal uso esta basado en la resolución de problemas o toma de decisiones que requieran gran poder de computo.

Sin embargo, estas tareas pueden representar un costo alto en recursos computacionales. Para esto, la programación multi-hilos puede ser una herramienta muy útil. Al tener varios hilos de procesamiento, es posible capturar y procesar datos de forma más eficiente. Esto permite que las computadoras (y las diferentes aplicaciones que se puedan realizar en estas) puedan llegar a resultados de manera más óptima y rápida.

Visto de otra forma, la programación multi-hilos es una herramienta que permite una mejora en los escenarios en donde se implemente. Ahora bien, logrando la combinación de este tipo de programación y la visión por computadora, es posible obtener reconocimientos de imágenes o entornos (como mesas de pruebas, mapas, entre otros) de manera mucho más adecuada, permitiendo utilizar estos datos en posteriores proyectos, como en la robótica de enjambre, por ejemplo, siendo esta una de sus principales escenarios de aplicación.

Ahora bien, la Programación Orientada a Objetos agrega muchas otras ventajas a los procesos de computo. Primero, ofrece una reutilización del código, es decir, permite utilizar distintos métodos en diversas partes de un código y en variedad de proyectos y aplicaciones. Segundo, permite modificaciones de manera sencilla y práctica ya que se puede añadir o eliminar objetos según sea la necesidad de la aplicación, así como también fiabilidad, ya que es posible reducir códigos grandes en partes más pequeñas, permitiendo encontrar errores de manera más rápida y precisa.

Objetivos

Objetivo General

Mejorar el algoritmo de visión por computadora desarrollado para la mesa de pruebas Robotat para experimentación de robótica de enjambre, usando programación orientada a

objetos, la librería OpenCV, y programación multi-hilos.

Objetivos Específicos

- Seleccionar el lenguaje de programación orientado a objetos más adecuado para la implementación de algoritmos de visión por computadora para la mesa de pruebas Robotat.
- Desarrollar algoritmos computacionalmente eficientes por medio de programación multi-hilos.
- Diseñar e implementar una herramienta de software para aplicaciones de robótica de enjambre, usando los algoritmos desarrollados.
- Validar la herramienta de software en la mesa de pruebas Robotat.

Marco teórico

Visión por Computadora

Visión por computadora se refiere al uso de cámaras o cualquier otro dispositivo de toma de fotografías o vídeos, para recolectar información para su posterior análisis, desarrollando algoritmos para hacer entender a la computadora que es lo que hay (en cuanto a datos o información significativa) en este tipo de archivos. [4]

En otras palabras, consiste en obtener la información relevante, realizando un procesamiento a imágenes y vídeos, para que los seres humanos puedan entender de mejor manera que es lo que hay en ellos. Es decir, poder visualizar lo que una computadora hace en este tipo de procesamientos. Normalmente, este tipo de procesamiento de datos es utilizado para obtener información del medio o entorno (mapas, carreteras, imágenes de todo tipo, etc.) y poder ser utilizado en resolución de problemas o toma de decisiones por parte de una computadora, basado en su entorno o aplicación [4]

OpenCV

Esta es la librería de **Open Source Computer Vision Library** y está disponible para Windows, MacOS y Linux. Es una librería para visión por computadora y machine learning. Incluye mas de 2500 algoritmos que permiten ser utilizados para reconocimiento de rostros, identificación de objetos, clasificación del comportamiento humano, rastreo del movimiento de los ojos entre otros.

Su implementación puede realizarse en C++, Python, Java y Matlab para las diferentes aplicaciones mencionadas. [5]

Programación Multi-hilos

Los procesadores orientados a multiprocesos, permiten realizar diferentes tareas al mismo tiempo. Estos a su vez, son responsables de manejar los recursos que se le asignen a cada uno de estos.

Más específicamente, cuando un programa es ejecutado, la computadora crea algo llamado **proceso** que contiene toda la información relevante del programa, como su identificador por ejemplo, hasta que dichos los programas terminan su ejecución.

Los sistemas operativos multiprocesos, permiten la ejecución de varios programas de manera simultanea y es la computadora la encargada de asignar los recursos a los diferentes programas que los necesiten. El objetivo principal es obtener un uso adecuado de los recursos del CPU. [6]

Un hilo, por tanto, es una unidad de control dentro de un proceso. El programa corre un hilo principal o **main thread** encargado de crear otros hilos según sea su programación, siendo este el principio básico del multi-hilos. Básicamente, orientar los programas ejecutados para realizar varias tareas y en muchos casos realizar procesos más eficientes. [6]

Existen ventajas en el uso de multi-hilos. El primero, es la ventaja de tener un programa realizando captura o procesamiento de información, mientras otro está esperando estas salidas. Es decir, en lugar de tener un programa que espera una imagen o dato, para luego procesarla y dar un resultado, se pueden tener dos procesos corriendo, uno capturando datos y el otro procesando. El resultado de esto será mucho más rápido.

Además, mencionar que la comunicación multi-hilos es mucho más eficiente que la comunicación entre procesos [6]

Programación Orientada a Objetos

Es un enfoque para la organización y el desarrollo de programas que intenta incorporar características más potentes y estructuradas para la programación. Es una nueva forma de organizar y desarrollar programas y no tiene nada que ver con ningún lenguaje de programación específico. Sin embargo, no todos los lenguajes son adecuados para implementar fácilmente los conceptos de la Programación orientada a objetos (POO). [7]

Ofrece grandes ventajas, ya que permite hacer modular un código, es decir, tener partes dedicadas a ciertas funciones y hacerlas hasta cierto punto independientes de las demás.

Además de eso, escribir códigos resulta sencillo ya que, por el concepto de objeto dentro de este tipo de programación, se permite tratarlo como un objeto en la vida real y se puede obtener sus atributos y funciones de manera más práctica que en una programación secuencial.

Objetos

Un objeto en POO cumple básicamente las mismas funciones que un objeto de la vida real. Un objeto puede guardar atributos o características y ser utilizadas mediante métodos. Esto es útil porque permite de alguna forma esconder esta parte interna para solo enfocarse en su función principal (aquella para la cual fueron diseñados). A esto se le conoce como *encapsulación*. [8]

Clase

Una clase es el plano a partir del cual se crean métodos o atributos para los diferentes objetos individuales que pertenecen a esta clase. Es, lo que se podría llamar, como un molde o modelo para la creación de objetos. [9] [8]

Metodología

Investigar

Implica buscar información referente a los lenguajes de programación para alcanzar los objetivos. El primer punto es investigar el funcionamiento de la programación multi-hilos y la sintaxis para los lenguajes seleccionados (funciones, restricciones de uso, diferencias entre ellos, etc.), así como todo lo referente a la librería de OpenCV para comprender su uso y realizar las pruebas que se requieren con la cámara.

Entre este paso también se puede incluir investigar sobre la programación orientada a objetos como una herramienta para el desarrollo de un algoritmo eficiente para aplicaciones de robótica de enjambre.

Implementar

Una vez investigado lo referente a la programación multi-hilos y OpenCV, se procederá a realizar códigos de prueba para validar el funcionamiento de las diferentes aplicaciones que se necesiten implementar. Esto implica programas ejemplo de aplicación multi-hilos, así como el uso de una cámara y OpenCV. Para esto, se tiene estipulado el uso de los lenguajes de C++ y Python

Luego de haber realizado y comprendido el funcionamiento, se procederá a analizar los programas existentes (implementados en el lenguaje C++) para intentar encontrar puntos de mejora y poder aplicar lo investigado. Finalmente, se realizará una migración hacia el lenguaje Python para tener puntos de comparación.

Además de esto, se requiere implementar una mesa de pruebas que busca ser una herramienta para validar el uso de OpenCV y sus aplicaciones en el procesamiento de imagen. Por lo que, con la finalidad de reconocer el entorno para la pose de agentes y realizar calibraciones a la cámara, se utilizará esta mesa para realizar las pruebas necesarias.

Medir y Validar

Finalmente, luego de tener dos elementos de comparación, se procederá a realizar pruebas para obtener parámetros que permitan medir y validar cual de las dos opciones es la mejor, o si ambas aportan de igual forma.

Uno de las primeras pruebas es realizar comparaciones entre la programación multi-hilos de ambos lenguajes definidos. Esto con el fin de medir su rendimiento y comparar el resultado de las tareas que se dispongan en esta primera prueba.

La segunda prueba consiste en la calibración de la cámara y el procesamiento de imágenes. Utilizando OpenCV y la mesa de pruebas diseñada, se pretende que de manera autónoma la computadora pueda reconocer la mesa y tomar acciones de calibración con la cámara. Para esto se tiene ya implementado un código en `C++` (fase anterior de esta tesis) el cual se comparará con los resultados en `Python` realizando el mismo procedimiento, comparando además la sintaxis y facilidad de programación como primer punto comparativo entre estos dos lenguajes.

Al final, las métricas de validación pueden ir desde el rendimiento de la computadora, tiempos de ejecución, entre otros, así como que los resultados para los cuales el programa fue diseñado sean los adecuados (correcta calibración de la cámara, obtención de resultados del procesamiento de imagenes adecuado, etc.)

Cronograma de actividades

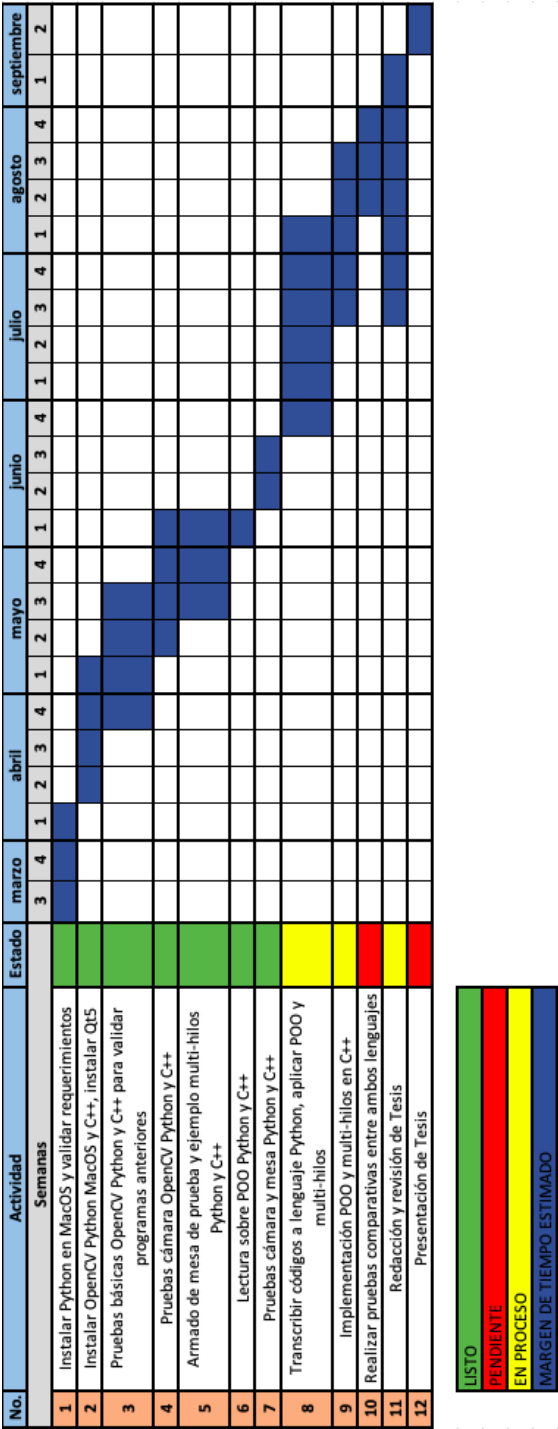


Figura 1: Cronograma de actividades

Índice preliminar

- Resumen
- Introducción
- Objetivos
 - Objetivo general
 - Objetivos específicos
- Justificación
- Marco teórico
 - Visión por computadora
 - OpenCV
 - Programación multi-hilos
 - Programación orientada a objetos.
- Antecedentes
- Alcance
- Metodología
 - Calibración de la cámara mediante OpenCV
 - Implementación de programación multi-hilos y programación orientada a objetos.
- Implementación
- Resultados
- Conclusiones
- Recomendaciones

Referencias

- [1] Luis Antón Canalís, “Swarm Intelligence in Computer Vision: an Application to Object Tracking”, Univesidad de las Palmas de Gran Canaria, mar. de 2010.
- [2] Javier Andrés Lizarazo Zambrano y Mario Alberto Ramos Velandia, “Visión artificial y comunicación en robots cooperativos omnidireccionales”, Universidad Central, Facultad de Ingeniería y Ciencias Básicas, ene. de 2016.
- [3] André Josué Rodas Hernández, “Desarrollo e implementación de algoritmo de visión por computador en una mesa de pruebas para la experimentación con micro-robots móviles en robótica de enjambre”, Departamento de ingeniería electrónica, mecatrónica y biomédica, Universidad del Valle de Guatemala, ene. de 2019.
- [4] Raúl E. López Briega, *Visión por computadora*, <https://iaarbook.github.io/vision-por-computadora/>, visitado el 10/05/2020.

- [5] OpenCV, *About*, <https://opencv.org/about/>, visitado el 05/04/2020.
- [6] R. H. Carver y K.-C. Tai, *Modern multithreading: implementing, testing, and debugging multithreaded Java and C++/Pthreads/Win32 programs*. John Wiley & Sons, 2005.
- [7] P. Chawla, *OOP with C++*, <http://www.ddegjust.ac.in/studymaterial/mca-3/ms-17.pdf>, visitado el 05/06/2020.
- [8] Oracle, *Lesson: Object-Oriented Programming Concepts*, <https://docs.oracle.com/javase/tutorial/java/concepts/index.html>, visitado el 05/06/2020.
- [9] E. Doherty, *What is Object Oriented Programming? OOP Explained in Depth*, https://www.educative.io/blog/object-oriented-programming?aid=5082902844932096&utm_source=google&utm_medium=cpc&utm_campaign=blog-dynamic&gclid=CjOKCQjwoPL2BRDxARIsAEMm9y_jZ7Mu3oH1wRJc5uHdWIeMdhGHbLeR22kkNxokV1-YZS-isYFjZKIaArzfEALw_wcB, visitado el 05/06/2020.