Jornadas de Automática

Monitorización mediante vehículos aéreos multicámara. Caso de uso de Unreal Engine

López Ruiz, J.F.[a], Rico, J.[a], Alamo, T.[a], Ortega, M.G.[a], Vargas, M.[a,*]

[a] *Dpto. Ing. de Sistemas y Automática. Universidad de Sevilla. Avda. de los Descubrimientos s/n. 41092, Sevilla, España.*

To cite this article: López Ruiz, J.F., Rico, J., Alamo, T., Ortega, M.G., Vargas, M. 2024. Monitoring through multi-camera aerial vehicles: A case study using Unreal Engine.
Jornadas de Automática, 45. https://doi.org/10.17979/ja-cea.2024.45.10800

Resumen

El trabajo propuesto en este artículo parte de la consideración de un vehículo aéreo autónomo, dotado de múltiples cámaras, como agente capaz de proporcionar mayor versatilidad en aplicaciones de monitorización y seguimiento de múltiples objetivos móviles. Dicho agente tendría la capacidad de orientar sus cámaras de forma completamente independiente y éstas estarían dotadas con capacidad de *zoom*, lo que permitiría ajustar la distancia focal de cada una de ellas a conveniencia. Partiendo de este concepto y de una estrategia de optimización desarrollada en trabajos anteriores, se propone una generalización de la misma que permita la colaboración de varios agentes en una misma misión. De forma complementaria, parte del trabajo se centra en explorar las posibilidades que ofrece *Unreal Engine 5* como herramienta de simulación gráfica para la implementación de la propuesta.

Palabras clave: Robots aéreos, Trabajo en entornos reales y virtuales, Soporte al operador humano, Percepción y detección, Sistemas multi-vehículo, Redes de robots y sensores inteligentes.

Monitoring through multi-camera aerial vehicles: A case study using Unreal Engine.

Abstract

The research work presented in this article is based on the concept of an autonomous aerial vehicle equipped with multiple cameras, serving as an agent capable of providing enhanced versatility in monitoring and tracking multiple mobile targets. The onboard cameras can be independently oriented and feature zoom functionality. This feature would enable adjusting the focal length of each camera as needed, optimizing the visual coverage of the targets. Building upon this concept and leveraging an optimization strategy developed in previous works, this study proposes a generalization of the approach to facilitate collaboration among multiple agents in a single mission scenario. Additionally, a significant component of the research focuses on exploring the potential of *Unreal Engine 5* as a graphical simulation tool for implementing the proposal and validating its effectiveness.

Keywords: Flying robots, Work in real and virtual environments, Human operator support, Perception and sensing, Multi-vehicle systems, Networks of robots and intelligent sensors.

1. Introducción

Hoy en día resulta bastante familiar la imagen de un vehículo aéreo no tripulado (UAV) dotado de una cámara orientable mediante un balancín motorizado (*gimbal*), que permita la monitorización visual del entorno de dicho vehículo. Si bien las primeras aplicaciones se limitaban esencialmente al ámbito de la inteligencia militar, en la última década se ha ampliado el espectro de uso a muchos otros dominios, incluyendo inspección de infraestructuras (Jordan et al. (2018); Salahat et al. (2019)), misiones de búsqueda y rescate (Sun et al. (2016); Burke et al. (2019)), vigilancia del tráfico (Khan et al. (2017); Kumar et al. (2020)), seguimiento de objetivos en general (Ahmed et al. (2021); Sun et al. (2023)), sin olvidar los usos recreativos, cada vez más extendidos.

Cuando el propósito es realizar el seguimiento simultáneo de varios objetivos móviles que puedan estar distantes entre

*Autor para correspondencia: mvargas@us.es.

sí, la solución convencional pasa bien por desplegar varios vehículos con capacidades similares, cada uno dedicado a un objetivo específico o grupo de objetivos cercanos, como en Robin and Lacroix (2016); Senanayake et al. (2016), bien por recurrir al concepto de atención compartida. En este último caso, un único vehículo puede ajustar periódicamente la dirección de apuntamiento de su cámara entre diferentes objetivos, intentando no perder la pista de ninguno de ellos (Sharma and Pack (2013); Zhao et al. (2019); Baek and York (2020)).

En un trabajo reciente (Vargas et al. (2022)), se explora la idea de un único agente aéreo equipado con múltiples cámaras orientables de forma independiente, con objeto de rastrear simultáneamente varios objetivos con mayor efectividad, al tiempo que se minimice el despliegue de recursos. En el citado trabajo, se asume como premisa que las cámaras destinadas a las tareas de seguimiento poseen una focal fija. En Vargas et al. (2024), por su parte, se generaliza la propuesta anterior. En primer lugar, se introduce la capacidad de agrupación de los objetivos (*clustering*). Esto permite abordar situaciones en las que el número de objetivos sea sustancialmente superior al número de cámaras a bordo, de forma que cada una de ellas se centre en un grupo de objetivos, en lugar de en uno individual. Por otra parte y más relevante, se introduce la capacidad de *zoom* ajustable en las cámaras. Esto proporciona una flexibilidad muy superior al agente, al permitirle encontrar automáticamente el equilibrio óptimo entre la distancia focal a fijar en cada una de las cámaras y la distancia real al objetivo o *cluster* en cuestión.

El presente trabajo plantea, como primera contribución destacada, una generalización del concepto original, considerando que se dispone de varios agentes multicámara, para desempeñar una misma misión de supervisión y seguimiento.

Por otra parte, cabe destacar que las herramientas de simulación gráfica y física desempeñan un papel fundamental en la investigación y desarrollo de sistemas robóticos aéreos, entre otros. Ello se debe a que proporcionan un entorno ideal para llevar a cabo pruebas de concepto y ensayos preliminares con seguridad, comodidad y bajo coste. En este sentido, la segunda contribución relevante del presente trabajo es la propuesta de una implementación del concepto referido en un entorno de simulación que ofrezca, entre otros aspectos, cualidades de animación foto-realistas.

La estructura del artículo es la siguiente. En primer lugar, en la Sección 2, se describe el caso de uso tratado, detallando los conceptos teóricos necesarios e incluyendo la generalización a múltiples agentes de los trabajos previos. En la Sección 3, se hace un rápido repaso a algunos de los entornos de simulación más aptos para la aplicación de interés, justificando la elección realizada, siendo ésta *Unreal Engine*. La Sección 4, por su parte, describe los aspectos más relevantes de la implementación llevada a cabo en *Unreal Engine* y los detalles de la misión recreada. Finalmente, en la Sección 5, se presentan las conclusiones del trabajo.

2. Caso de uso: Monitorización aérea mediante equipo de agentes multi-cámara

2.1. *Concepto de agente aéreo autónomo multi-cámara*

Tal y como se describe en Vargas et al. (2022, 2024), cada uno de los agentes considerados sería un vehículo aéreo con una cámara central, instalada sobre un *gimbal* que puede ser de dos grados de libertad y un número variable de cámaras de seguimiento, colocadas sobre respectivos *gimbals* de tres grados de libertad. En la Figura 1, se muestra una imagen esquemática de un agente de tipo multirrotor con una cámara central flanqueada por dos cámaras de seguimiento. La cámara central mantendría en todo momento un punto de vista cenital, porporcionando una vista general de la escena. Por su parte, las cámaras de seguimiento se centrarán en un objetivo o *cluster* particular, intentando mantenerlo centrado y encuadrado en la imagen.

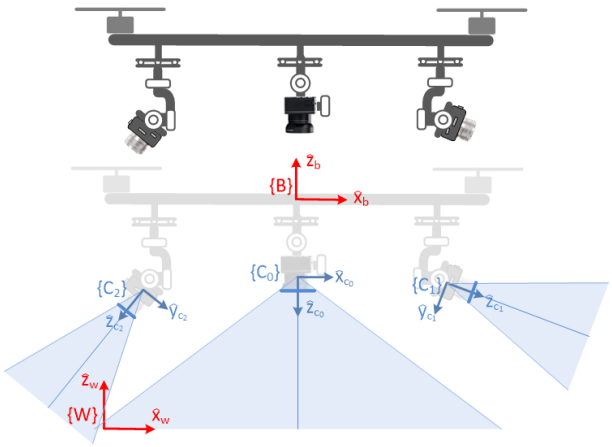

Figura 1: Concepto de agente aéreo multi-cámara empleado.

En este trabajo se hará la suposición de mundo plano, de forma que, partiendo del conocimiento de la altitud del vehículo, y de la posición proyectada en la imagen de un cierto objetivo, pueda deducirse la posición tridimensional del mismo. En la descripción que sigue, se denotará por x_k la posición del agente k-esimo (mediante $\mathrm{x}_k(t)$, en caso de que se requiera explicitar la dependencia con el tiempo, donde t representaría el número de ciclo de ejecución), dotado con n_k cámaras de seguimiento. Por su parte, z_i representa la posición del objetivo i-ésimo (o centro del *cluster* i-èsimo). La variable $d_{(k,i)}$ representa la distancia entre ambos: $d_{(k,i)} = \|\mathrm{x}_k - \mathrm{z}_i\|$.

Basándose en la idea anterior, en Vargas et al. (2024) se plantea un índice de optimización que proporcione referencias de posición en tiempo real, que permitan posicionar el vehículo de forma óptima para la observación del conjunto de objetivos. Para el objetivo o *cluster* i-ésimo, se determina, en primer lugar, la distancia ideal para la observación del mismo, $d_{(k,i),r}$, en función del tamaño equivalente que se estima que tiene en la realidad dicho objetivo o *cluster*, y del tamaño deseado del mismo en la imagen. A partir de esta distancia, se determina que, si efectivamente no hubiera más objetivos, la posición deseada a la que se debería dirigir el vehículo sería una posición en la vertical del objetivo, a una distancia igual a $d_{(k,i),r}$. A esta posición se le denota por $\mathrm{p}_{(k,i),r}$.

Cuando hay más de un objetivo o grupo, salvo en casos límite, no será posible satisfacer la condición anterior para todos los objetivos, de modo que se plantea una segunda condición deseable, en defecto del cumplimiento de la primera. Esta nueva condición pretende que se respete la distancia de observación deseada, aunque no se tenga un punto de vista vertical del objetivo en cuestión.

Cuando hay más de tres objetivos o grupos, en general, no será posible satisfacer la condición anterior para todos ellos. La siguiente condición pretenderá que la distancia al objetivo sea tal que se garantice el tamaño deseado en a imagen, fijando un valor de distancia focal en la cámara correspondiente que esté dentro de su rango operativo: $f_{(k,i),min} \leq f_{(k,i)} \leq f_{(k,i),max}$.

Si la condición anterior tampoco es posible, el siguiente paso es intentar que tanto la distancia focal como el tamaño aparente del objeto en la imagen estén dentro del rango aceptable.

Finalmente, en caso de que lo anterior no sea viable, tenemos la situación en la que no se podrá ajustar un valor de distancia focal que permita que el tamaño del objeto en la imagen esté dentro del rango deseado, penalizándose la desviación respecto a los límites de dicho rango.

Como resumen de lo anterior, las condiciones de observación pueden enumerarse, comenzando por la situación ideal hasta llegar a la más desfavorable, del siguiente modo:

1. $\mathrm{x}_k = \mathrm{p}_{(k,i),r}$: Alineamiento vertical óptimo respecto a un objetivo individual, consiguiéndose al mismo tiempo la distancia focal ideal y el tamaño deseado en la imagen.
2. $d_{(k,i)} = d_{(k,i),r}$: Se viola el alineamiento vertical óptimo, pero la distancia focal y el tamaño del objeto en la imagen son los ideales.
3. $d_{(k,i)} \in [L_{(k,i),1}, U_{(k,i),1}]$: La distancia focal no coincide con el valor óptimo, pero permanece dentro del rango admisible, mientras que se garantiza que el tamaño del objeto en la imagen es el deseado.
4. $d_{(k,i)} \in [L_{(k,i),2}, U_{(k,i),2}]$: La distancia focal y el tamaño aparente del objeto en la imagen no son los ideales, pero sí están dentro del rango deseado.
5. $d_{(k,i)} \notin [L_{(k,i),2}, U_{(k,i),2}]$: No puede conseguirse mantener el tamaño aparente del objetivo en la imagen dentro del rango deseado, dadas las limitaciónes de la distancia focal de la cámara.

A partir de todo lo anterior, el índice se formula como una suma de términos, cada uno correspondiente a una cámara de seguimiento. A su vez, cada uno de los términos se construye mediante la suma de tres sub-términos que pretenden penalizar que la solución viole la secuencia de condiciones expuestas. En particular, el problema de optimización propuesto determina la referencia óptima para la posición del agente resolviendo:

$$\mathrm{x}_k^* = \arg\min_{\mathrm{x}_\in \mathcal{X}_k} J_{k,o}(\mathrm{x}_k), \tag{1}$$

siendo $\mathcal{X}_k$ la caja que define las restricciones a imponer sobre la posición del agente k-ésimo, de forma que, por ejemplo, pueda obligarse a que dicha posición esté restringida a una cierta franja de altitud. Por su parte, el índice de optimización asociado a la "calidad" en la observación del conjunto de objetivo o *clusters* vinculados a dicho agente, $J_{k,o}$, se define como:

$$J_{k,o}(\mathrm{x}_k) = \sum_{i=1}^{n_k} \Gamma_{(k,i)}(\mathrm{x}_k, \mathrm{z}_i) + \Psi_{(k,i)}(\mathrm{x}_k, \mathrm{z}_i) + \Lambda_{(k,i)}(\mathrm{x}_k, \mathrm{z}_i) \tag{2}$$

Los términos $\Gamma_{(k,i)}(\mathrm{x}_k, \mathrm{z}_i)$ penalizan el incumplimiento de las condiciones 1 y 2. Los términos $\Psi_{(k,i)}(\mathrm{x}_k, \mathrm{z}_i)$ penalizan el hecho de que se rebasen los umbrales superiores, $U_{(k,i),j}$; mientras que los términos $\Lambda_{(k,i)}(\mathrm{x}_k, \mathrm{z}_i)$ penalizan el incumplimiento de los umbrales inferiores, $L_{(k,i),j}$. Estos dos últimos, por tanto, cuantifican las condiciones 3 a 5. Para una descripción más detallada del desarrollo expuesto, se remite al lector a Vargas et al. (2024).

2.2. *Gestión de múltiples agentes*

En este trabajo, se propone la generalización de la idea anterior para la inclusión de varios agentes, cuyo cometido sea la monitorización conjunta de un grupo de objetivos. Como se verá, se trata de una solución subóptima, donde la coordinación no se plantea mediante la reformulación integral del problema de optimización ampliado. Por contra, se adopta la solución óptima individual existente, y se combina con una estrategia de optimización combinatorial aplicada al mecanismo de *clustering*.

Como punto de partida, se asumirá que se dispone de m agentes para una misma misión, pudiendo ser éstos de diversa naturaleza y estar dotados con diferente número de cámaras de seguimiento. Se denotará por $\mathfrak{a}_k$ al identificador del k-ésimo agente, mientras que $\mathcal{A} := \{\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_m\}$ constituiría el conjunto de los identificadres de todos los agentes disponibles. De modo similar, t_i representará el identificador del *cluster* (u objetivo, *target*) i-ésimo, siendo el conjunto ordenado de los n *clusters* $\mathcal{T} := \{\mathrm{t}_1, \mathrm{t}_2, \ldots, \mathrm{t}_n\}$. Por otra parte, X representará el conjunto de las posiciones adjudicadas a los agentes en un cierto momento: $\mathrm{X} := \{\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_m\}$, mientras que Z hará referencia al conjunto de las posiciones de los *clusters* en un cierto instante: $\mathrm{Z} := \{\mathrm{z}_1, \mathrm{z}_2, \ldots, \mathrm{z}_n\}$. Para cada agente, se definirá un coste compuesto por tres términos: en primer lugar, el asociado a la observación de los objetivos, acorde a la definición (2); en segundo lugar, el coste del desplazamiento del agente desde su posición actual y, por último, el coste asociado a eventuales cambios en las asignaciones cámara-*cluster* (*swaps*):

$$J_k(\mathrm{x}_k) = \alpha_o\, J_{k,o}(\mathrm{x}_k) + \alpha_d\, \|\mathrm{x}_k(t) - \mathrm{x}_k(t-1)\| + \alpha_{sw}\, S_k. \tag{3}$$

donde se han introducido como factores de ponderación las constantes α_o, α_d, α_{sw} y donde S_k representa el número de cambios de asignación cámara-*cluster* que ha experimentado el agente k-ésimo respecto al ciclo previo. En relación a esto último, cabe destacar el interés de introducir dicha penalización, con idea de reducir la frecuencia de estos intercambios, favoreciendo con ello la tarea de supervisión por parte de operadores humanos.

El primer paso del algoritmo consistirá en agrupar los objetivos en un número n de *clusters*, coincidente con el número de cámaras de seguimiento disponibles, $n = \sum_{k=1}^{m} n_k$. Para tener en cuenta todas las posibles asignaciones de n_k *clusters* que se puedan hacer al agente $\mathfrak{a}_k$, se obtendrán las permutaciones de los n *clusters*, tomados de n_k en n_k, $\mathcal{P}(\mathcal{T}, n_k)$. Las permutaciones resultantes se denotarán por $\mathcal{T}_{n_k,j}$, $j = 1, \ldots, n_k!/(n-n_k)!$. Dado un subconjunto ordenado de identificadores de *cluster*, $\mathcal{T}_{n_k,j}$, tentativamente asignados al agente $\mathfrak{a}_k$, se asume que el primero se está asignando a la cámara 1 del agente, el segundo a la cámara 2, etc. El coste de cada asignación completa, considerando todos los agentes, se obtendría, a partir de (3), del siguiente modo:

$$J(\mathrm{X}) = \sum_{k=1}^{m} J_k(\mathrm{x}_k).$$

En un principio, cabría pensar en calcular, para cada una de las posibles asignaciones de *clusters* a cámaras, el correspondiente coste y simplemente elegir aquélla que proporcione el valor mínimo del mismo. Esto, sin embargo, supondría resolver, en cada ciclo de ejecución, un problema de optimización relativamente costoso un número de veces igual a $n!$. Para evitar esta gravosa estrategia exhaustiva, se propone una alternativa de *branch & bound*, que evite el cálculo del índice para aquellas permutaciones parciales que superen o igualen el coste mínimo, obtenido hasta el momento a partir de las permutaciones que ya han sido completamente desarrolladas. El procedimiento propuesto se detalla mediante los Algoritmos 1 y 2.

Algorithm 1: Cálculo del coste para un agente

Function $\{J_k,\ \mathrm{x}_k\} = \texttt{CosteAgente}(\mathfrak{a}_k, \mathcal{T}_{n_k,j})$**:**
  $\mathrm{x}_k(t-1) = \mathfrak{a}_k.\texttt{PosActual}();$
  $\mathcal{T}_{n_k,j}(t-1) = \mathfrak{a}_k.\texttt{AsignacionActual}();$
  if $\mathfrak{a}_k.\texttt{CosteYaCalculado}(\mathcal{T}_{n_k,j})$ **then**
    $J_k = \mathfrak{a}_k.\texttt{LeeCosteCalculado}(\mathcal{T}_{n_k,j})$;
  else
    $\mathrm{Z} = \{\mathrm{z} : \mathrm{z} = \mathrm{t}_i.\texttt{PosActual}(),\ \forall \mathrm{t}_i \in \mathcal{T}_{n_k,j}\};$
    Resolver $\mathrm{x}_k(t) = \arg\min_{\mathrm{x}\in\mathcal{X}_k} J_{k,o}(\mathrm{x},\ \mathrm{Z});$
    $S_k = \texttt{NumSwaps}(\mathcal{T}_{n_k,j}(t-1),\ \mathcal{T}_{n_k,j});$
    $J_k = \alpha_o J_{k,o}(\mathrm{x}_k(t), \mathrm{Z}_j) + \alpha_d \|\mathrm{x}_k(t) - \mathrm{x}_k(t-1)\| + \alpha_{sw} S_k;$
    $\mathfrak{a}_k.\texttt{RegistraCosteCalculado}(\mathcal{T}_{n_k,j},\ J_k);$
  end if

Algorithm 2: Cálculo del coste global

Function $\{J^*,\ \mathrm{X}^*\} = \texttt{CosteGlobal}(\mathcal{A},\ \mathcal{T},\ J,\ \mathrm{X},\ J^*,\ \mathrm{X}^*)$**:**
  if $\mathcal{A} == \emptyset$ **then**
    if $J < J^*$ **then**
      $J^* = J;\quad \mathrm{X}^* = \mathrm{X};$
    end if
    return;
  end if
  for *cada* $\mathfrak{a}_k \in \mathcal{A}$ **do**
    $n_k = \mathfrak{a}_k.\texttt{NumCamaras}();$
    Obtener $\mathcal{P}(\mathcal{T},\ n_k);$
    for *cada* $\mathcal{T}_{n_k,j} \in \mathcal{P}(\mathcal{T},\ n_k)$ **do**
      $\{J_k,\ \mathrm{x}_k\} = \texttt{CosteAgente}(\mathfrak{a}_k, \mathcal{T}_{n_k,j});$
      if $J + J_k < J^*$ **then**
        $\{J^*,\ \mathrm{X}^*\} = \texttt{CosteGlobal}(\mathcal{A}\setminus\{\mathfrak{a}_k\},\ \mathcal{T}\setminus\mathcal{T}_{n_k,j},\ J+J_k, \{\mathrm{X},\ x_k\}, J^*, \mathrm{X}^*);$
      end if
    end for
  end for

El Algoritmo 1 calcula, de acuerdo con (3), el coste para un agente individual, supuesta una asignación de *clusters* $\mathcal{T}_{n_k,j}$, proporcionando tanto el valor óptimo del índice, como la posición correspondiente. El segundo algoritmo, por su parte, realiza el cálculo del coste global óptimo de forma recursiva, desarrollando el árbol de combinaciones agente-*clusters* necesarias, mediante la estrategia de *branch & bound* citada. Este segundo algoritmo recibe un conjunto o subconjunto de identificadores de agentes y *clusters*, $\{\mathcal{A}, \mathcal{T}\}$, el coste acumulado hasta el momento y el conjunto de posiciones calculadas hasta el momento al ir descendiendo por las ramas del árbol, $\{J, \mathrm{X}\}$, y, finalmente, el coste mínimo encontrado hasta el momento y la solución correspondiente, $\{J^*, \mathrm{X}^*\}$. Por otra parte, devuelve como salida la actualización de este último par de parámetros. El caso base de la estrategia recursiva en el Algoritmo 2 se alcanza cuando el conjunto $\mathcal{A}$ recibido es el conjunto vacío, en cuyo caso, se habrá alcanzado el nivel de mayor profundidad del árbol, habiéndose hecho, por tanto, una asignación completa de todos los *clusters*; procede entonces comprobar si el coste correspondiente al camino desarrollado es inferior al mínimo que se tenía hasta el momento, en cuyo caso, este mínimo es actualizado. El operador \ usado en la llamada recursiva del algoritmo representa sustracción de conjuntos. La llamada principal al Algoritmo 2 sería del siguiente modo: `CosteGlobal` $(\mathcal{A}, \mathcal{T}, 0, \emptyset, \infty, \emptyset)$, donde $\mathcal{A}$ y $\mathcal{T}$ representan el conjunto completo de agentes y *clusters*.

3. Entornos de simulación para aplicaciones de robótica aérea

Esta sección ofrece una breve enumeración de algunas de las herramientas de simulación gráfica más utilizadas en el contexto de vehículos aéreos no tripulados.Una interesante revisión reciente a este respecto puede encontrarse en Collins et al. (2021), donde se dedica un apartado específico a los sistemas de simulación para robótica aérea.

En primer lugar, cabe citar a *Gazebo* (Koenig and Howard (2004)), como el simulador multi-robot de código abierto más ampliamente utilizado en combinación con ROS. Posee una capacidad de renderizado relativamente limitada, pero ofrece, por el contrario, un amplio soporte para modelos UAV y controladores hardware. Como segunda alternativa, se puede hacer referencia a la ofrecida por el ecosistema *Matlab*, incluyendo *Simulink* y el *UAV Toolbox*, que posibilitan la simulación de vehículos aéreos basadas en *Unreal Engine* (The MathWorks (2024)). Sin embargo, al sustentarse esta altarnativa en *Matlab*, podría ser menos deseable a la hora de trasladar de forma ágil el desarrollo a un entorno real.

Por otra parte, podemos mencionar a *Webots* (Michel (2004)). Se trata de un software de simulación de robots móviles 3D de código abierto que ofrece un conjunto diverso de robots y sensores predefinidos, pero carece de renderización foto-realista del entorno. *AirSim* (Shah et al. (2018)) es otro simulador de código abierto, desarrollado por *Microsoft* y basado en *Unreal Engine 4*, que sí está más específicamente orientado a vehículos aéreos. Si bien *Airsim* ofrece renderizado realista y detección de colisiones, se trata de una herramienta muy exigente en términos computacionales.

Dentro de la categoría de motores de videojuegos empleados para simulación en aplicaciones robóticas, destacan *Unity* y *Unreal Engine*. *Unity* ofrece renderizado realista y permite simulación en tiempo real mediante la integración de ROS. Basado en *Unity*, puede mencionarse también a *Flightmare* (Song et al. (2021)), que es un simulador de vehículos aéreos modular de código abierto que combina el motor de renderizado *Unity* con un motor de física flexible, adecuado para simulaciones de alta fidelidad.

En este trabajo se ha optado, sin embargo, por *Unreal Engine* (Epic Games (2024)). La elección se fundamenta en que se trata una herramienta permanentemente en la vanguardia,

gratuita para desarrolladores, respaldada por una fuerte comunidad y que ofrece por sí misma las capacidades de modelado y renderizado realista deseadas, con una carga computacional configurable según las limitaciones del hardware. En particular, se ha optado por *Unreal Engine 5* (*UE5*), la cual introdujo importantes mejoras de rendimiento y un mayor foto-realismo de los diseños.

4. Implementación en Unreal Engine 5

Como se ha mencionado anteriormente, el motor de simulación empleado es *UE5*, concretamente su versión 5.2.1. *UE5* ofrece numerosas herramientas para la programación, como *Python Scripting*, *Blueprint Visual Scripting*[1] y *C++ Coding*. Es esta última herramienta en la que se basa gran parte de la implementación llevada a cabo. En esta primera aproximación a *UE5*, sin embargo, no se simulan de forma totalmente realista la dinámica de objetos ni las colisiones, sino que el desarrollo se ha centrado mayoritariamente en los elementos cinemáticos y en los aspectos gráficos. Hay diversos conceptos, propios de la terminología *UE5*, que resulta relevante referir. En primer lugar el concepto de *actor*, referido a cualquier objeto existente en la escena. En segundo lugar, un *static mesh* es cualquier actor compuesto por un grupo estático de polígonos, mientras que un *scene component* es un actor no visible que proporciona un marco de referencia.

4.1. Creación del escenario

Se ha optado por recrear, en el contexto de una misión de salvamento y rescate marítimo tras accidente aéreo, el despliegue de un equipo de 4 agentes multi-cámara. La Figura 4 permite apreciar como fondo de la misma el escenario simulado, cuyos elementos más destacados se pasan a comentar a continuación. En primer lugar, aparecen objetos estáticos creados a través de modelos 3D de libre uso e introducidos como *static meshes*. Otro elemento relevante es el agua con física realista (olas, refracción de la luz, etc.), creada con el complemento *Water* de *UE5*. El paisaje ha sido implementado con la herramienta *Landscape*. Finalmente, el sistema de iluminación, compuesto por varios actores, entre los cuales destaca *SunSky*, que recrea la luz solar junto a la atmósfera terrestre.

4.2. Simulación de objetivos de interés

Para la recreación de las víctimas del incidente, que serán los objetivos de interés para la aplicación, se puede usar la tecnología *Meta Human*, que permite la creación de réplicas de humanos de calidad foto-realista, de forma simple y rápida. Esta será una opción seleccionable, de forma opcional, en la aplicación, dado que supone un aumento significativo del coste computacional de la simulación. Se simularán trayectorias elípticas de distinta amplitud, excentricidad y velocidad para estos objetivos, desarrollándose todas ellas a nivel ligeramente inferior al de la superficie del agua. En la simulación realizada, se ha simplificado la inicialización del problema, asumiendo que, a la hora de poner en marcha la estrategia de monitorización y seguimiento de los objetivos, los agentes están situados de forma que aquéllos queden encuadrados por las cámaras centrales de éstos.

4.3. Creación de los agentes

Como se ha comentado con anterioridad, se ha optado por un equipo de cuatro agentes, $\mathcal{A} = \{\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3, \mathfrak{a}_4\}$. El número de cámaras de los agentes será diferente, poseyendo 4, 3, 2 y 1 cámaras de seguimiento, respectivamente, además de las respectivas cámaras centrales. La implementación de los agentes se ha llevado a cabo principalmente mediante funciones en C++ diseñadas para integrarse con *Blueprint*. De este modo, se configura el flujo de información entre dichas funciones, dando lugar a la lógica completa de cada agente. Además, se ha modelado cada componente del agente (vehículo, *gimbals*, cámaras, etc) mediante *CAD* y se han creado los *static meshes* correspondientes, lo que permite una visualización precisa del movimiento de los mismos. La Figura 2 muestra un agente con cuatro cámaras de seguimiento.

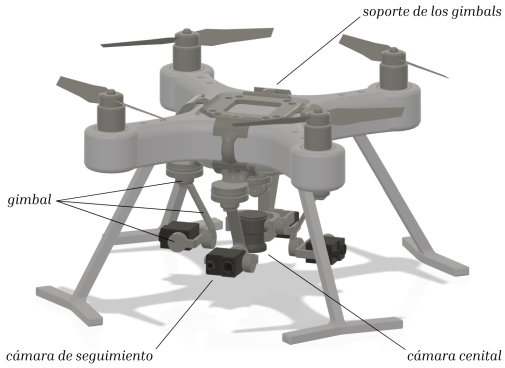

Figura 2: Modelo usado como agente en la simulación. El caso particular de cuatro cámaras de seguimiento corresponde al agente $\mathfrak{a}_1$.

En la Figura 3, se muestran las vistas proporcionadas por los cuatro agentes en un momento dado, durante el desarrollo de la misión. A cada agente se le ha asignado automáticamente un número de *clusters* igual al número de cámaras de seguimiento que posee. El agente debe ajustar su posición y focales de forma óptima, acorde al índice de minimización descrito, para el adecuado encuadre de los *clusters* asignados.

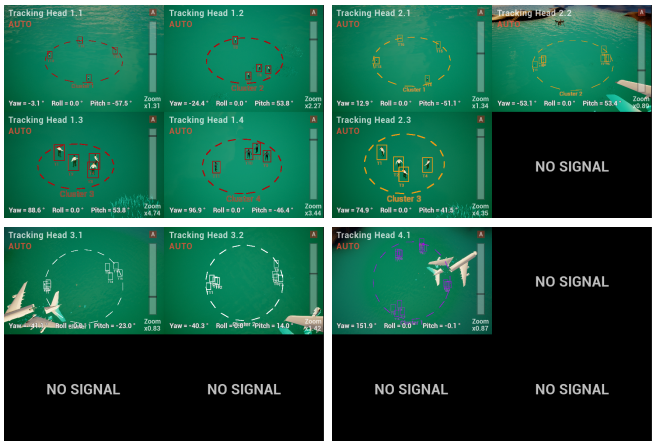

Figura 3: Vistas capturadas por las cámaras de cada uno de los cuatro agentes.

[1] *Blueprint* es un sistema de *scripting* visual nativo de *UE*, que está basado en la interconexión de nodos.

4.4. *Interfaz de usuario*

Se ha diseñado la interfaz de usuario presentada en la Figura 4. Se ha priorizado una disposición eficiente de las distintas vistas proporcionadas por la simulación, las cuales incluyen: la cámara cenital (esquina superior izquierda), las cámaras de seguimiento (esquina superior derecha) y la vista del agente seleccionado (esquina inferior izquierda). Además de esta disposición, la interfaz introduce controles básicos, tales como la pausa o reanudación de la simulación y la selección de un agente específico. También se han implementado controles avanzados que permiten ajustar parámetros específicos del agente seleccionado, incluyendo las focales de las cámaras de seguimiento y central, así como el nivel de *zoom* aplicado sobre la vista del agente previamente comentada. Adicionalmente, la interfaz ofrece la capacidad de modificar la orientación y el *zoom* de la vista principal del escenario. Por último, proporciona una ayuda visual al usuario al resaltar los distintos objetivos y *clusters* en las vistas de las cámaras del agente. En López Ruiz et al. (2024) puede encontrarse un vídeo demostrativo del experimento descrito.

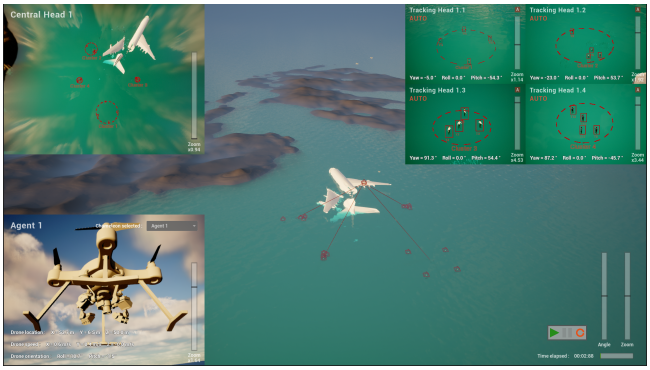

Figura 4: Interfaz mostrando el escenario de la misión y la información e imágenes relativas al agente a_1.

5. Conclusiones

El presente trabajo se centra en la implementación en *UE5* de un sistema de monitorización y seguimiento de un número arbitrario de objetivos móviles, mediante un equipo de vehículos aéreos no tripulados o agentes. Cada agente estaría dotado de múltiples cámaras orientables de forma independiente y habilitadas con *zoom* ajustable. Partiendo de la solución teórica del problema propuesta en trabajos anteriores, en los que se gestionaba únicamente un agente, se ha realizado una adaptación de la misma, para permitir la colaboración de varios agentes en una misión común. La estrategia propuesta combina el algoritmo de *clustering* con una estrategia de optimización combinatorial basada en *branch & bound*. Mediante el simulador, se ha recreado la monitorización y seguimiento de objetivos en alta mar, como asistencia en una hipotética misión de salvamento tras una catástrofe.

Agradecimientos

Ayudas PID2020-115561RB-C32, PID2022-142946NA-I00 y PID2022-141159OB-I00, financiadas por MICIU/AEI/10.13039/501100011033 y FEDER/UE.

Referencias

Ahmed, I., Din, S., Jeon, G., Piccialli, F., Fortino, G., 2021. Towards collaborative robotics in top view surveillance: A framework for multiple object tracking by detection using deep learning. IEEE/CAA Journal of Automatica Sinica 8 (7), 1253–1270.

Baek, S., York, G., 2020. Optimal sensor management for multiple target tracking using cooperative unmanned aerial vehicles. In: 2020 Int. Conference on Unmanned Aircraft Systems (ICUAS). IEEE, pp. 1294–1300.

Burke, C., McWhirter, P. R., Veitch-Michaelis, J., McAree, O., Pointon, H. A., Wich, S., Longmore, S., 2019. Requirements and limitations of thermal drones for effective search and rescue in marine and coastal areas. Drones 3 (4), 78.

Collins, J., Chand, S., Vanderkop, A., Howard, D., 2021. A review of physics simulators for robotic applications. IEEE Access 9, 51416–51431.

Epic Games, 2024. The most powerful real-time 3d creation tool - unreal engine. `https://www.unrealengine.com/en-us`, accessed: 2023-10-17.

Jordan, S., Moore, J., Hovet, S., Box, J., Perry, J., Kirsche, K., Lewis, D., Tse, Z. T. H., 2018. State-of-the-art technologies for UAV inspections. IET Radar, Sonar & Navigation 12 (2), 151–164.

Khan, M. A., Ectors, W., Bellemans, T., Janssens, D., Wets, G., 2017. UAV-based traffic analysis: A universal guiding framework based on literature survey. Transportation Research Procedia 22, 541–550.

Koenig, N., Howard, A., 2004. Design and use paradigms for gazebo, an open-source multi-robot simulator. In: 2004 IEEE/RSJ international conference on intelligent robots and systems (IROS)(IEEE Cat. No. 04CH37566). Vol. 3. Ieee, pp. 2149–2154.

Kumar, P., Sonkar, S., Ghosh, A. K., Philip, D., 2020. Real-time vision-based tracking of a moving terrain target from light weight fixed wing UAV using gimbal control. In: 2020 7th International Conference on Control, Decision and Information Technologies (CoDIT). Vol. 1. IEEE, pp. 154–159.

López Ruiz, J. F., Rico, J., Vargas, M., 2024. Enlace a vídeo demostrativo. `https://uses0-my.sharepoint.com/:f:/g/personal/mvargas_us_es/EoNpD79xbc9FpxZKeUnZwxMB34Q5rtbs6xgIAoqGl4s8NQ?e=fxHfUb`, Accessed: 2024-06-01.

Michel, O., 2004. Cyberbotics ltd. webots™: professional mobile robot simulation. International Journal of Advanced Robotic Systems 1 (1), 5.

Robin, C., Lacroix, S., 2016. Multi-robot target detection and tracking: taxonomy and survey. Autonomous Robots 40 (4), 729–760.

Salahat, E., Asselineau, C.-A., Coventry, J., Mahony, R., 2019. Waypoint planning for autonomous aerial inspection of large-scale solar farms. In: IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society. Vol. 1. IEEE, pp. 763–769.

Senanayake, M., Senthooran, I., Barca, J. C., Chung, H., Kamruzzaman, J., Murshed, M., 2016. Search and tracking algorithms for swarms of robots: A survey. Robotics and Autonomous Systems 75, 422–434.

Shah, S., Dey, D., Lovett, C., Kapoor, A., 2018. Airsim: High-fidelity visual and physical simulation for autonomous vehicles. In: Field and Service Robotics: Results of the 11th International Conference. Springer, pp. 621–635.

Sharma, R., Pack, D., 2013. Cooperative sensor resource management to aid multi target geolocalization using a team of small fixed-wing unmanned aerial vehicles. In: AIAA Guidance, Navigation, and Control (GNC) Conference. p. 4706.

Song, Y., Naji, S., Kaufmann, E., Loquercio, A., Scaramuzza, D., 2021. Flightmare: A flexible quadrotor simulator. In: Conference on Robot Learning. PMLR, pp. 1147–1157.

Sun, J., Li, B., Jiang, Y., Wen, C., 2016. A camera-based target detection and positioning UAV system for search and rescue (SAR) purposes. Sensors 16 (11), 1778.

Sun, P., Li, S., Zhu, B., Zuo, Z., Xia, X., 2023. Vision-based fixed-time uncooperative aerial target tracking for UAV. IEEE/CAA Journal of Automatica Sinica 10 (5), 1322–1324.

The MathWorks, I., 2024. Unreal engine simulation for unmanned aerial vehicles - matlab and simulink. `https://www.mathworks.com/help/uav/ug/3d-simulation-for-unmanned-aerial-vehicles.html`, accessed: 2024-03-28.

Vargas, M., Vivas, C., Alamo, T., 2024. Optimal positioning strategy for multi-camera, zooming drones. IEEE/CAA J. Autom. Sinica 11 (8), In press.

Vargas, M., Vivas, C., Rubio, F. R., Ortega, M. G., 2022. Flying chameleons: A new concept for minimum-deployment, multiple-target tracking drones. Sensors 22 (6), 2359.

Zhao, Y., Wang, X., Wang, C., Cong, Y., Shen, L., 2019. Systemic design of distributed multi-UAV cooperative decision-making for multi-target tracking. Autonomous Agents and Multi-Agent Systems 33 (1), 132–158. DOI: `10.1007/s10458-019-09401-5`