

Universidad de Alcalá  
Escuela Politécnica Superior

Grado de Ingeniería Informática



**Trabajo Fin de Grado**

Soporte Remoto Mediante Realidad Aumentada

**Autor:** Héctor Campos Romero

**Tutor/es:** José María Gutiérrez Martínez

2018

UNIVERSIDAD DE ALCALÁ

Escuela Técnica Superior de Ingeniería  
Informática

INGENIERÍA SUPERIOR EN INFORMÁTICA

---

# Trabajo Fin de Carrera

“SOPORTE REMOTO MEDIANTE REALIDAD  
AUMENTADA”

**Autor:** Héctor Campos Romero

**Director:** José María Gutiérrez Martínez

Tribunal:

Presidente: .....

Vocal 1º: .....

Vocal 2º: .....

Fecha: .... de ..... de .....

## **Palabras clave**

Realidad aumentada, soporte remoto, vuforia, visión artificial, unity, voip.

## **Key words**

Augmented reality, remote support, vuforia, artificial vision, unity, voip.

## **Resumen corto**

Este proyecto da una solución al problema de dar soporte a alguien en la distancia, para ello, se emplea la realidad aumentada y otras características, como chat de voz por IP, rejillas de coordenadas, etc. Todo esto se lleva a cabo empleando Unity como plataforma y Vuforia como framework de realidad aumentada.

## **Short summary**

This project gives a solution to the problema of giving support remotely. To achieve this task, augmented reality and other features are used, such as a voice chat over IP, coordinate grid and so on. The project is made using Unity as platform and Vuforia as augmented reality framework.

# Índice resumido

<b>Introducción</b>	9
<b>Objetivo</b>	11
<b>Estado del Arte</b>	13
<b>Desarrollo del Proyecto</b>	25
<b>Coste del Proyecto</b>	47
<b>Resumen, conclusiones y trabajos futuros</b>	51
<b>Bibliografía</b>	55
<b>Anexo</b>	65

# Índice detallado

1	Introducción.....	9
2	Objetivo.....	11
3	Estado del Arte.....	13
3.1	Historia de la Realidad Aumentada.....	13
3.2	Definiciones de la Realidad Aumentada.....	14
3.3	Diferencia entre realidad aumentada y realidad virtual.....	15
3.4	Beneficios y Aplicaciones de la Realidad Aumentada.....	16
3.5	Actualidad y Futuro de la RA.....	20
3.6	¿Por qué Unity?.....	22
3.6.1	¿Por qué Vuforia?.....	23
3.7	Proyectos que aplican RA al soporte.....	23
4	Desarrollo del Proyecto.....	25
4.1	Experimentación con Vuforia.....	25
4.1.1	Interactividad con Vuforia.....	27
4.1.2	Problemas de rendimiento.....	28
4.1.3	Tracking de los targets.....	28
4.1.4	Añadir información a targets.....	29
4.2	Vídeo remoto.....	30
4.3	Generación dinámica o aleatoria de targets.....	31
4.4	Reconocimiento de Objetos.....	33
4.5	Chat VoIP.....	37
4.6	Arquitectura y funcionalidades finales.....	39
4.7	Interfaz de Usuario.....	39
4.7.1	Menús.....	40
4.7.2	Capturas de pantalla.....	43
4.7.3	Rejillas de coordenadas.....	43
4.7.4	Accesibilidad Básica.....	45
5	Coste del Proyecto.....	47
5.1	Presupuesto de Ejecución Material.....	47
5.1.1	Coste de equipos.....	47
5.1.2	Coste por tiempo de trabajo.....	48
5.1.3	Coste total de ejecución material.....	48
5.2	Gastos Generales y Beneficio Industrial.....	48
5.3	Presupuesto de Ejecución por Contrata.....	49
5.4	Honorarios.....	49
5.5	Importe total del presupuesto.....	49
6	Resumen, conclusiones y trabajos futuros.....	51
6.1	Resumen.....	51
6.2	Conclusiones.....	52
6.3	Trabajos futuros.....	53

7	Bibliografia.....	55
8	Anexo .....	65



# 1 Introducción

En la actualidad millones de usuarios utilizan una gran variedad de aplicaciones de smartphone, cada una con funcionalidades y objetivos distintos. Algunas de ellas nos permiten controlar remotamente otros dispositivos, incluso gracias a Google Cardboard, un visor de realidad virtual hecho con cartón, la realidad virtual en Android es muy asequible. Estas aplicaciones nos permiten dar solución a muchos problemas cuyo propósito no depende tanto de la potencia del hardware, sino de la movilidad y disponibilidad que supone tener un smartphone o cualquier otro dispositivo móvil.

Un problema en concreto, el cual se pretende paliar con este proyecto, es el de la comunicación de instrucciones entre dos personas que no se encuentran físicamente en el mismo lugar. La solución que se suele emplear habitualmente es usar un teléfono o algún tipo de comunicación a distancia que cualquier persona pueda tener, ya sea mensajería instantánea o cualquier método de comunicación oral. Aunque esto sea una solución aceptable, una imagen vale más que mil palabras, por lo tanto, en muchas ocasiones, indicar sobre una imagen o tomar una fotografía de lo que se pretende indicar sería mucho más rápido que estar describiendo con palabras el entorno y lo que se pretende comunicar.

Los smartphones son buenas herramientas para resolver este problema el cual es muy común entre muchas profesiones. Aunque actualmente no hay aplicaciones capaces de resolverlo con cierta facilidad, la realidad aumentada ha empezado a abrirse un hueco dentro de este campo, y es que ser capaz de mezclar elementos virtuales con la realidad es un punto muy positivo en este tipo de casos. La realidad aumentada es una tecnología que a la par de vistosa puede ser muy útil, esto implica un nuevo abanico de posibilidades, sobre todo si se habla de dispositivos móviles, ya que son una plataforma que está muy bien capacitada para esta finalidad, debido a que tienen la potencia necesaria para implementar visión artificial, consiguiendo así una gran precisión en la detección de objetos e incorporando componentes como la cámara o la pantalla táctil, que permiten interactuar de forma cómoda y muy visual a cualquier usuario.

Poco a poco las empresas realizan investigaciones respecto a la realidad aumentada, creando hardware adaptado a su uso y reduciendo costes con el tiempo, tanto de hardware como a la hora de programar estas soluciones. También su relación tan estrecha con la realidad virtual ha supuesto la experimentación con la realidad mixta, una mezcla de ambas tecnologías que amplían el alcance de estas.

Es una tecnología muy interesante tanto a nivel personal como profesional, ya que, a diferencia de otras tecnologías más maduras, la realidad aumentada permite ser creativo obteniendo



resultados más rápidamente y que además son bastante satisfactorios, es decir, en un período de tiempo relativamente corto puedes tener una función con resultados gráficos de gran utilidad sin necesidad de conocimientos tan profundos como en otras áreas. La posibilidad de que la realidad interactúe con un mundo virtual es, como mínimo, interesante, ya que abre la puerta a posibilidades que antes eran teóricas o su perfeccionamiento no era posible. A esto hay que añadirle su estrecha relación con la visión artificial, que es otro campo de la informática, cuyos resultados son sorprendentes.

Personalmente, este proyecto supone la aplicación de los conocimientos obtenidos en el grado de ingeniería informática para realizar una aplicación móvil que no solo es útil, sino que además, explora algo tan novedoso como la realidad aumentada y todo lo que ello implica. La cantidad de conocimientos extraídos de dicho proyecto, dan una nueva perspectiva de lo que las nuevas tecnologías, como esta, pueden hacer para cambiar o mejorar el mundo en el que vivimos.

## 2 Objetivo

El objetivo principal de este proyecto de fin de grado consiste en desarrollar una aplicación para Android que intente resolver el problema de dar indicaciones remotas de una forma sencilla de cara al usuario y que no implique un gasto elevado a la hora de implementarse. Al apoyarse en un hardware tan diversificado, como puede ser un smartphone, siempre cabe la posibilidad de problemas de incompatibilidad, de potencia insuficiente o problemas relacionados a estos, incluido el consumo de batería o datos móviles.

Desde el comienzo se descarta la realidad mixta, ya que las pantallas de los smartphones más asequibles no suelen tener la calidad necesaria para que pudiera realizarse una realidad mixta más cómoda, además de que su utilidad en este caso sería como mínimo cuestionable. Un visor VR es bastante voluminoso, aunque el peso también supone una mayor incomodidad en su uso.

Se desarrollarán las siguientes funcionalidades:

- Streaming de la pantalla de un dispositivo a otro, lo que comúnmente se da a conocer como “compartir pantalla”.
- Detección de objetos 3D y marcación de los mismos.
- Pequeñas marcas detectables que den información preestablecida.
- La posibilidad de asignar a una marca cualquiera o patrón un objeto virtual que transmita información editable por el usuario.
- Una rejilla de coordenadas, similar a la que emplean juegos de mesa como Hundir la Flota o el ajedrez para dar indicaciones más precisas.
- Un chat VoIP (Voz por IP) que permita una comunicación fluida entre dos o más personas.



## 3 Estado del Arte

En este capítulo se profundizará en el campo de la realidad aumentada. Comienza con la historia más notoria, mostrando la evolución que ha seguido a lo largo de los años. Después, se presentan definiciones sobre la RA y se marca la diferencia que existe entre RA y realidad virtual. También se expondrán los beneficios que puede aportar, además de la situación en la que se encuentra esta tecnología actualmente y su posible futuro. Finalmente, acaba con una muestra de proyectos que siguen objetivos similares a los establecidos en este.

### 3.1 Historia de la Realidad Aumentada

Aunque el uso de la realidad aumentada se haya popularizado y extendido recientemente, realmente el término se acuñó en los años 90. El crédito de usar el término por primera vez se le otorga a Tom Caudell de Boeing Computer Services, el cual, en búsqueda de facilitar las tareas de manufactura, empezó con un proyecto de realidad virtual que acabó resultando en la superposición de elementos virtuales sobre la realidad.

Alrededor del año 1992 un equipo creó el primer sistema funcional de realidad aumentada que fue usado por las fuerzas aéreas de los Estados Unidos para el entrenamiento de pilotos. Otro, al mismo tiempo que el anterior, publicó un artículo llamado KARMA (Knowledge-based Augmented Reality for Maintenance Assistance o Realidad Aumentada basada en el Conocimiento para la Asistencia en el Mantenimiento) en el cual se empleó una pantalla montada como un visor con sensores para indicar cómo cambiar los cartuchos de tinta de una impresora. Esto último sería la versión primitiva de lo que hoy conocemos como Google Glass o HoloLens.

Hasta 1999 era una tecnología que ni acababa de ser comprendida ni se usaba con frecuencia, donde el software complejo y el alto precio del hardware necesario supusieron un problema. Esta situación cambió al crearse el ARToolKit por Hirokazu Kato y el Instituto Nara de la Ciencia y la Tecnología, el cual permitió seguir movimientos del mundo real que solo necesitaba un dispositivo como una cámara y una conexión a internet.

En el año 2000 Bruce Thomas de Wearable Computer Lab enseñó el primer videojuego para móviles, empleado en exteriores, que usaba realidad aumentada, llamado ARQuake. El videojuego permitía a sus usuarios andar sin necesidad de usar palancas o botones. Para su

funcionamiento aún hacía falta un ordenador a la espalda y giroscopios, pero consiguió que, con un simple visor, el usuario tuviera vistas completamente distintas en función de su posición física en el mundo real.

El gran salto de la realidad aumentada a los smartphones en 2008 permitió a todo el mundo la posibilidad de probar de primera mano esta tecnología. La primera aplicación fue un indicador de puntos de interés, creado para el sistema operativo Android, que empleaba la cámara del teléfono. Más tarde se relanzó para iPhone como una aplicación de navegación con el nombre de Wikitude Drive.

Con el paso de los años más videojuegos y aplicaciones han ido implementando la realidad aumentada. Un gran ejemplo reciente fue el de Pokémon Go, el cual consiguió, con un videojuego de capturar criaturas o Pokémon, enganchar a millones de usuarios en todo el mundo haciendo que estos tuvieran que interactuar con puntos geográficos urbanos, empleando el GPS del teléfono, y con otros usuarios para conseguir ciertos objetivos que se les presentaba.

## **3.2 Definiciones de la Realidad Aumentada**

Entorno al año 1992, Tom Caudell acuñó el término realidad aumentada que más adelante tuvo definiciones relativas.

La primera de las definiciones dada por Ronald Azuma afirma que “la realidad aumentada combina elementos reales y virtuales, está registrada en 3D y además es interactiva en tiempo real”.

Paul Milgram y Fumiso Kishino en 1994 la definen como “una realidad más cercana al mundo real en un continuo que abarca desde el entorno real a un entorno virtual”.

La realidad aumentada también “supone la incorporación de datos e información digital en un entorno real, por medio del reconocimiento de patrones que se realiza mediante un software”.

Con la información vista, podemos extraer una definición común. La realidad aumentada consiste en añadir información digital sobre el mundo real mediante un software, siendo esta interactiva tanto para el usuario como para el mundo en el que se soporta, a diferencia de la realidad virtual, la cual transporta al usuario a un mundo completamente virtual para que interactúe con este.

Aun teniendo una clara definición de realidad aumentada, es fácil llegar a confundirla con la realidad virtual, sobre todo si no se han tenido experiencias con estas tecnologías ya que tienen muchas similitudes. En ocasiones se implementan realidades mixtas, lo que dificulta la diferenciación de ambas.

### **3.3 Diferencia entre realidad aumentada y realidad virtual**

Aunque son dos tecnologías cuyo resultado puede ser bastante opuesto entre ellas, sus amplias similitudes, tanto a nivel de desarrollo como de dispositivos, pueden llegar a complicar su diferenciación, además, si se tiene en cuenta el desconocimiento de estas tecnologías por buena parte de la población, hace que su diferenciación se convierta en una tarea de mayor relevancia.

Dado que previamente se ha tratado la definición de realidad aumentada ahora se tratará la realidad virtual y sus principales diferencias.

La realidad virtual es completamente inmersiva, es decir, es una tecnología que permite al usuario, mediante un headset, percibir e interactuar con un mundo completamente virtual. Esto nos lleva a la principal diferencia entre la RA y la RV, la RA tiene como mundo la realidad mientras que la RV tiene como mundo uno digital.

Otra gran diferencia, aunque esta sea más técnica, es la capacidad de computación necesaria para poder usarlas adecuadamente sin incomodidades, mareos, etc. La RA, aunque sea muy recomendable que fluya a la mayor cantidad de FPS (imágenes por segundo) posible, en el caso de bajos FPS no se ve tan penalizada como en el caso de la RV, que al tener una tasa de FPS inferior a 60 o inestable puede llegar a provocar mareos y malestares con facilidad, lo que es un indicativo de que a nivel de hardware es más exigente la realidad virtual que la aumentada.



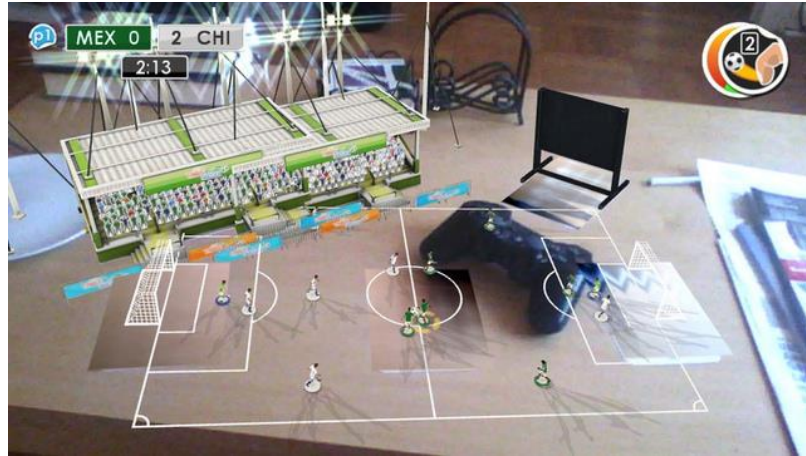
Descripción con imágenes de lo que diferencia a la RA y la RV.

A pesar de las limitaciones que puedan tener, sus beneficios mejoran muchos campos profesionales, los cuales se presentan en detalle en el siguiente punto.

### 3.4 Beneficios y Aplicaciones de la Realidad Aumentada

Dadas las características de la realidad aumentada puede llegar a ser complicado entender la posible utilidad de la misma. Se listan las aplicaciones más prominentes de la realidad aumentada:

- **Videojuegos:** Los videojuegos han hecho su propio uso de la realidad aumentada, empleando tarjetas físicas como referencia o introduciendo elementos en el mundo captados con una cámara, como en el famoso Pokémon Go. Su uso es meramente con el objetivo de entretener, con lo cual el beneficio que ofrece la realidad aumentada suele reducirse a la posibilidad de interactuar de formas distintas con los objetos virtuales.



RA en los videojuegos.

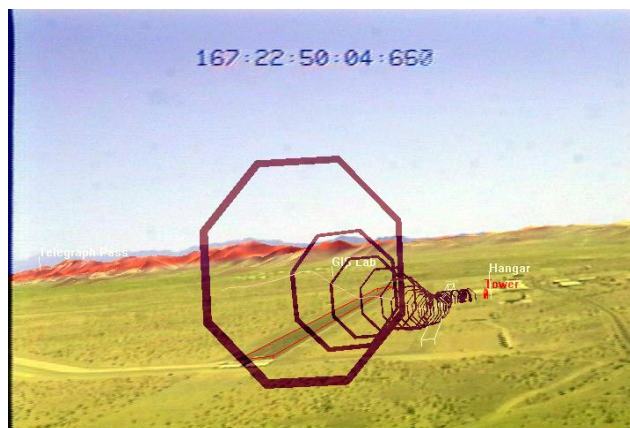
- **Diseño industrial:** Al igual que el modelaje 3D, la realidad aumentada ha ampliado el diseño de productos, permitiendo realizar comparaciones más precisas entre el resultado físico y el diseño virtual. Además, puede emplearse sobre planos para mostrar un producto acabado en base a ellos, entre otras funcionalidades.



Empleando RA para visualizar piezas.

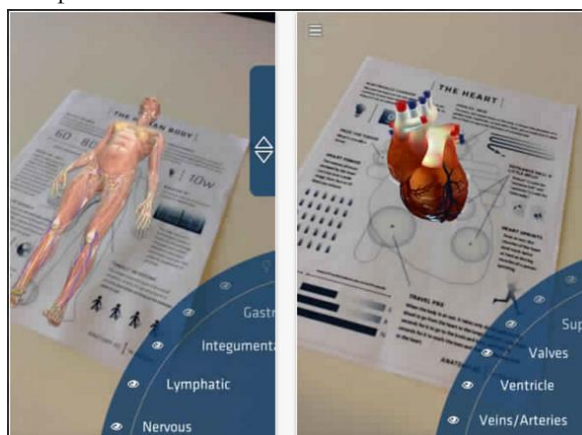
- **Navegación, aviación o militar:** La superposición de mapas, trayectorias y otros elementos informativos permiten al conductor o piloto una mayor precisión en su percepción del estado del vehículo.





Navegación de un avión en el aterrizaje marcado con RA.

- **Aprendizaje y educación:** Mostrar los pasos para alcanzar un objetivo o mostrar un proceso más allá de la pura teoría es una de las principales ventajas de la realidad aumentada en estos campos.



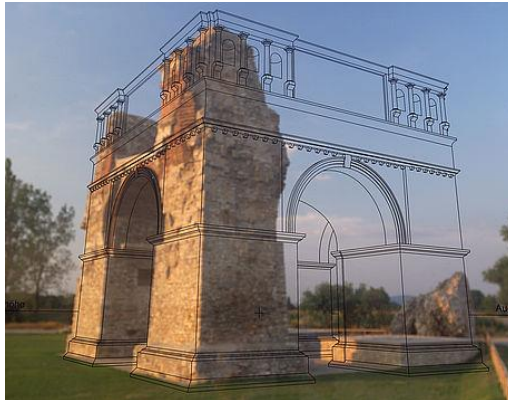
Muestra del cuerpo humano a la izquierda y de un corazón a la derecha, ambos en RA.

- **Arquitectura:** En lugar de maquetas para mostrar el resultado final de un proyecto, se puede emplear modelaje 3D y mostrarlo mediante realidad aumentada, habilitando la posibilidad de ver interiores a voluntad e interactuar con ellos.



Visualización de planos arquitectónicos en 3D con RA.

- **Arqueología:** La reconstrucción de paisajes, construcciones o incluso seres vivos, cuyos fósiles han sido hallados, son los principales usos que recibe la realidad aumentada.



Reconstrucción de edificios históricos mediante RA.

- **Comercio:** Poder mostrar un producto antes de comprarlo, añadir información sobre este más allá de su etiqueta o mostrar el resultado de su montaje.



RA aplicada al comercio.

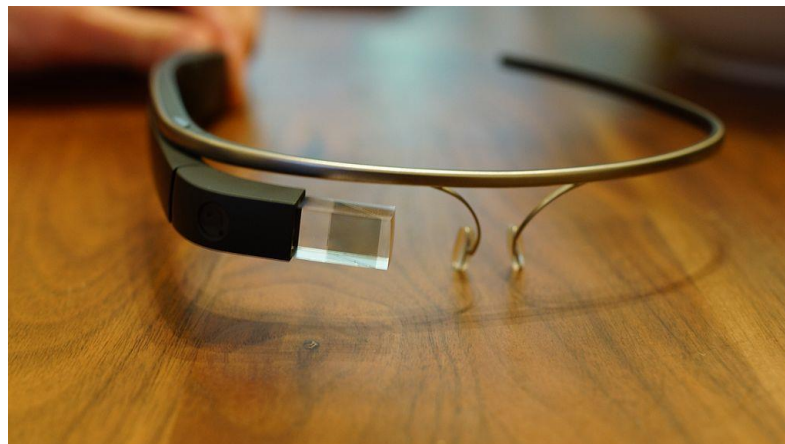
La realidad aumentada está estrechamente ligada al reconocimiento de objetos e imágenes, incluso caras, lo que implica que las condiciones en las que se emplee, como la luz, puedan dar

resultados variables en calidad, al igual que problemas de privacidad al emplearse cámaras, lo que, a diferencia de otras tecnologías, implica una limitación de su eficacia en función de varios factores principalmente relacionados con el entorno y la capacidad de reconocimiento de objetos.

### **3.5 Actualidad y Futuro de la RA**

Actualmente las grandes compañías que impulsan esta tecnología y pretenden de ella hacer un futuro cercano son Google, Apple y Microsoft, cada uno con sus propias apuestas en este campo.

Google se adelantó mucho en esta tecnología al crear las Google Glass, unas gafas de realidad aumentada que suscitaron mucho revuelo en el público, sobre todo en el campo de la privacidad, ya que tener unos “ojos” digitales que vean lo que vemos, capaces de tomar imágenes o vídeos sin siquiera darnos cuenta, fue un factor a tener en cuenta para la aplicación masiva de esta tecnología. Tras esto Google creó Project Tango que buscaba implementar de serie esta tecnología en nuevos smartphones, y que tras 2 dispositivos costosos y con poca utilidad, tanto por problemas de tamaño, potencia o catálogo de aplicaciones, fue cancelado a finales del 2017. Google continuará con la realidad aumentada apostando por ARCore, un framework que permite experimentar la realidad aumentada mediante software compatible con multitud de dispositivos móviles, lo cual lo aproxima bastante al ARKit de Apple.



Google Glass Explorer Edition

Apple, continúa con ARKit para crear software con realidad aumentada para iOS. Según patentes y otras noticias de la actualidad Apple podría estar trabajando en gafas de realidad aumentada, aunque no se debe dar por hecho.

Microsoft, por su parte, tiene las HoloLens, que son sus propias gafas de realidad aumentada. También tiene planes a futuro para esta tecnología, aunque como otros startups, apuntan más a la realidad mixta. Microsoft a finales del 2017 lanzó la plataforma Windows Mixed Reality y headsets inmersivos para su uso con la última actualización de Windows 10, los cuales son compatibles también con aplicaciones de realidad virtual.



Microsoft HoloLens y cómo se verían puestas en la cabeza.

También existen visionarios y startups con una mayor ambición en estos campos, como el famoso Elon Musk, que fundó Neuralink, o la empresa Neurable Inc., ambas apuntando a mezclar el cerebro humano con la realidad aumentada y la realidad virtual. En la actualidad, Neurable ya consiguió crear la primera interfaz cerebro-ordenador para el headset de realidad virtual HTC Vive.

Puede concluirse, por los avances más recientes, que hay una tendencia a mezclar la realidad aumentada y la realidad virtual, o incluso con otros campos como la neurociencia, tratando de acercar aún más ese futuro propio de la ciencia ficción.

## 3.6 ¿Por qué Unity?

Unity es una plataforma diseñada para el desarrollo de videojuegos, permitiendo agilizar el desarrollo facilitando un gran abanico de herramientas. En la actualidad, es una de las plataformas más usadas.

Unity nos ofrece la posibilidad de desarrollar tanto en 2D como en 3D, obteniendo una calidad gráfica impecable junto a la facilidad de programación que tiene tanto en C#, Unity Script o Boo. La implementación que realiza de funcionalidades básicas, pero que llevarían mucho tiempo programarlas desde cero, acelera rápidamente el desarrollo de cualquier proyecto.

A pesar de parecer que su uso se limite al desarrollo de videojuegos, realmente puede aplicarse a una gran cantidad de campos, entre los cuales se encuentra la realidad aumentada y la realidad virtual. Esto es posible gracias a la implementación de assets o paquetes descargables en su tienda, que ya sean de pago o gratuitos, pueden llegar a ser sistemas complejos completamente programados y listos para usar en un par de clicks. También hay frameworks que se integran con Unity dotando de mayor polivalencia a este.

La comunidad que se ha creado en torno a Unity es de las más grandes, debido a que es una plataforma de fácil uso y aprendizaje, una gran cantidad de usuarios comienzan a aprender con esta plataforma. La gran cantidad de documentación que ofrece es uno de los grandes atractivos de esta plataforma, desde descripciones de clases hasta vídeos perfectamente explicados y divididos en dificultades.

Se usa Unity para este proyecto principalmente por su facilidad de uso, ya que, al no tener experiencia, una plataforma tan intuitiva como esta es un gran beneficio. También hay que tener en cuenta la documentación y su comunidad a la hora de obtener información, dejando atrás a opciones como UnrealEngine o CryEngine. A esto se le suma la compatibilidad con el framework Vuforia, el cual será tratado en el siguiente apartado.

### **3.6.1 ¿Por qué Vuforia?**

Para desarrollar un proyecto de realidad aumentada es necesario depender de un framework externo que se integre con Unity, debido a que este último da soporte a la tecnología, pero no la implementa de base.

Vuforia es el framework de realidad aumentada a emplear en el proyecto, siendo de los más asequibles en un comienzo, ya que buena parte de sus características las ofrece de forma gratuita. Su extendido uso, sobre todo en aplicaciones móviles, hace que sea un gran candidato por la gran cantidad de información que se tiene al alcance, comparado con otros.

Ofrece una gran variedad de objetivos de realidad aumentada para su implementación, desde simples imágenes planas hasta la detección de objetos tridimensionales, pasando por cilindros como objetivo u objetos formados por múltiples imágenes. Todo esto puede añadirse a un proyecto con tan solo un objeto, con unas pocas propiedades muy intuitivas que permiten poner en funcionamiento la realidad aumentada en unos minutos.

También ofrece detección basada en la nube o sets de datos alojadas en ella, además de herramientas para digitalizar objetivos de objetos reales u obtener dichos objetivos mediante un plano, siendo compatible con múltiples formatos como cad, fbx u obj.

Su rendimiento y su capacidad multiplataforma con todos los headsets de realidad virtual o aumentada hacen que destaque, ya que no es necesario por parte del desarrollador invertir tiempo en optimizaciones o en compatibilidades, ya que de todo esto se encarga Vuforia sin problemas.

Vuforia es un framework cuyo soporte por parte de los desarrolladores y de la comunidad ofrece un buen punto de partida, que junto a su gran facilidad de uso y la variedad de funcionalidades que ofrece, lo hacen muy adecuado para este proyecto.

## **3.7 Proyectos que aplican RA al soporte**

En la actualidad, este proyecto ya ha sido llevado a cabo, ya que podemos encontrar ejemplos como Skylight, el cual es un producto diseñado para empresas con unas necesidades

parecidas a las que se han mencionado anteriormente, además de enfocarse a dispositivos concretos como Google Glass y con un coste asociado.

Otro proyecto muy similar, desarrollado por Vuforia, es Vuforia Chalk, el cual consiste en emplear una aplicación para dar cierto soporte remoto, dando la posibilidad de dibujar sobre lo que muestra la cámara del usuario remoto.



Imagen de la presentación de Vuforia Chalk.

En caso de encontrar algún otro proyecto que realice tareas similares, a diferencia de los anteriormente mencionados, su uso es muy concreto y personalizado.

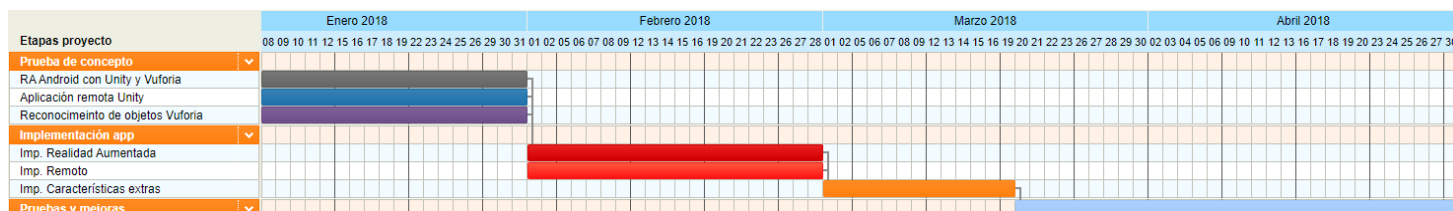
## 4 Desarrollo del Proyecto

En los siguientes puntos, se expondrá el proceso que ha seguido el desarrollo del proyecto, y las consideraciones que se han ido teniendo en cuenta en el transcurso de este. Se encuentra dividido en varias partes, cada una correspondiendo a una funcionalidad o sistema que se ha implementado o probado.

Antes de comenzar, se exponen brevemente los objetivos a alcanzar:

- Streaming de la pantalla de un dispositivo a otro
- Detección de objetos 3D y marcación de los mismos.
- Pequeñas marcas detectables que den información preestablecida.
- Marcas de RA asignables en ejecución.
- Una rejilla de coordenadas.
- Un chat VoIP (Voz por IP).

Para el desarrollo de dichos objetivos se pretendía seguir la siguiente planificación:



En la planificación se tiene en cuenta como características extra tanto la rejilla de coordenadas como el chat VoIP.

En la práctica, todos los tiempos estimados se redujeron en torno a la mitad y la parte del desarrollo del streaming remoto quedó anulada debido a causas que se verán a continuación en posteriores apartados.

### 4.1 Experimentación con Vuforia

Antes de comenzar a desarrollar, había que tener en cuenta las posibilidades que brinda Vuforia y Unity, y dada la inexperiencia, era muy conveniente una fase de experimentación con



diversas características para alcanzar cierto grado de experiencia, el indispensable para poder llevar a cabo el desarrollo de una aplicación con los objetivos que se han establecido.

En esta primera fase se comenzó con un Hello World en Unity, empleando Vuforia con una de sus imágenes de ejemplo. A partir de este punto, se experimentaron las posibilidades que ofrece el framework relativas a la realidad aumentada.

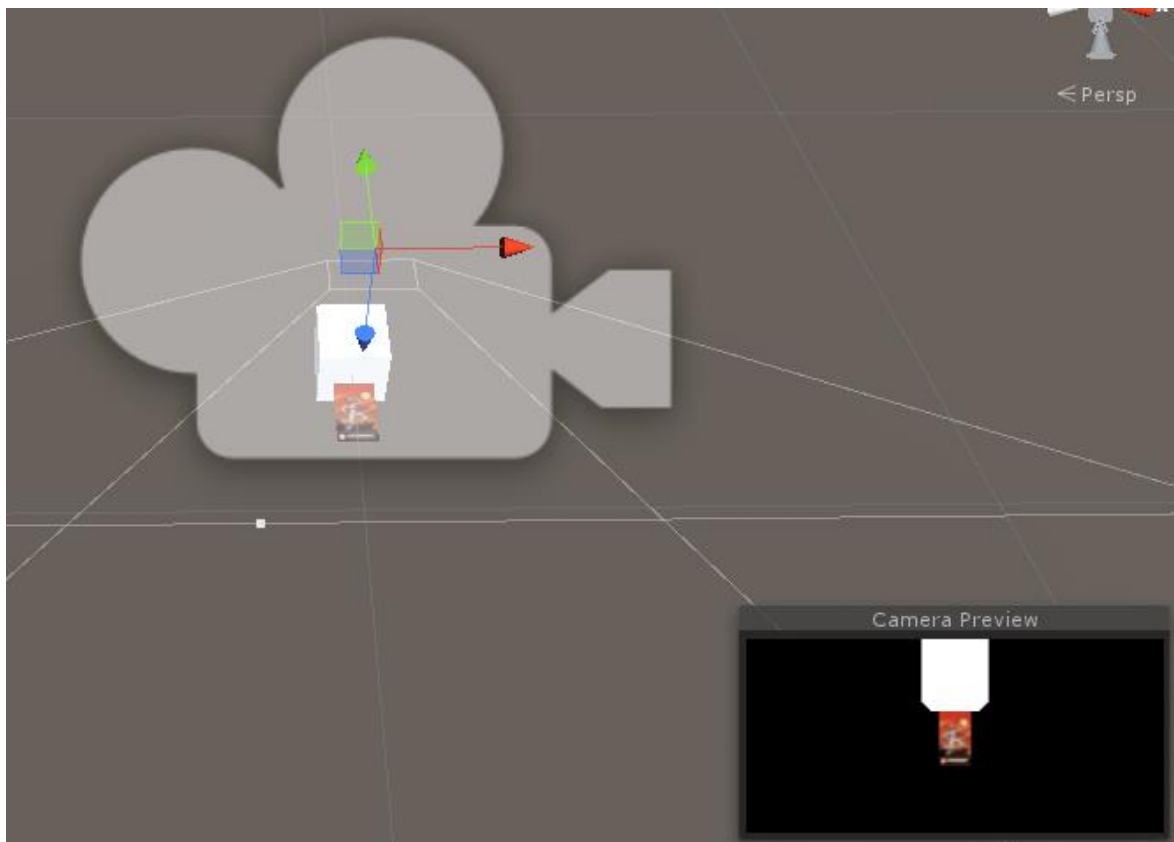


Fig 1. “Hello World” con Vuforia, asignando a una imagen detectable un cubo virtual.

Dado que la plataforma objetivo es Android, se configuró Vuforia acorde con este requerimiento, debido a que las opciones varían dependiendo del sistema operativo que se use, ya que se producen problemas como obtener la pantalla en negro si no se configura adecuadamente. La raíz de estos problemas con Android, suelen ser incompatibilidades con el smartphone usado o simplemente con todo Android, sin importar la versión.

Cabe destacar la definición de target, tanto de un ModelTarget como de un ImageTarget, en el contexto de Vuforia. Un target, sin importar el tipo, es un objeto que Vuforia toma como referencia a la hora de emplear su sistema de detección, es decir, si detecta un target con la cámara, mostrará los objetos virtuales asignados a este y/o realizará las acciones que se le hayan indicado de

antemano. Estos targets pueden tener asignados comportamientos y cualquier objeto que soporte Unity.

#### 4.1.1 Interactividad con Vuforia

En la búsqueda de permitir que el usuario interactúe con la aplicación, para ser capaz de resaltar información y objetos que sean de relevancia para su labor, se tuvo la idea inicial de generar marcadores de realidad aumentada, como flechas o círculos, y que el usuario los moviera y transformara para ajustarlos a sus necesidades.

Para dotar de interacción a los objetos 3D instanciados en la detección de un target, se empleó el asset LeanTouch, mediante el cual se permitía mover hacia arriba y hacia abajo el objeto 3D usando los dedos. Además de la traslación, se añadió también la rotación y la posibilidad de escalar su tamaño con los gestos de los dedos típicos del zoom o la rotación de los mismos.

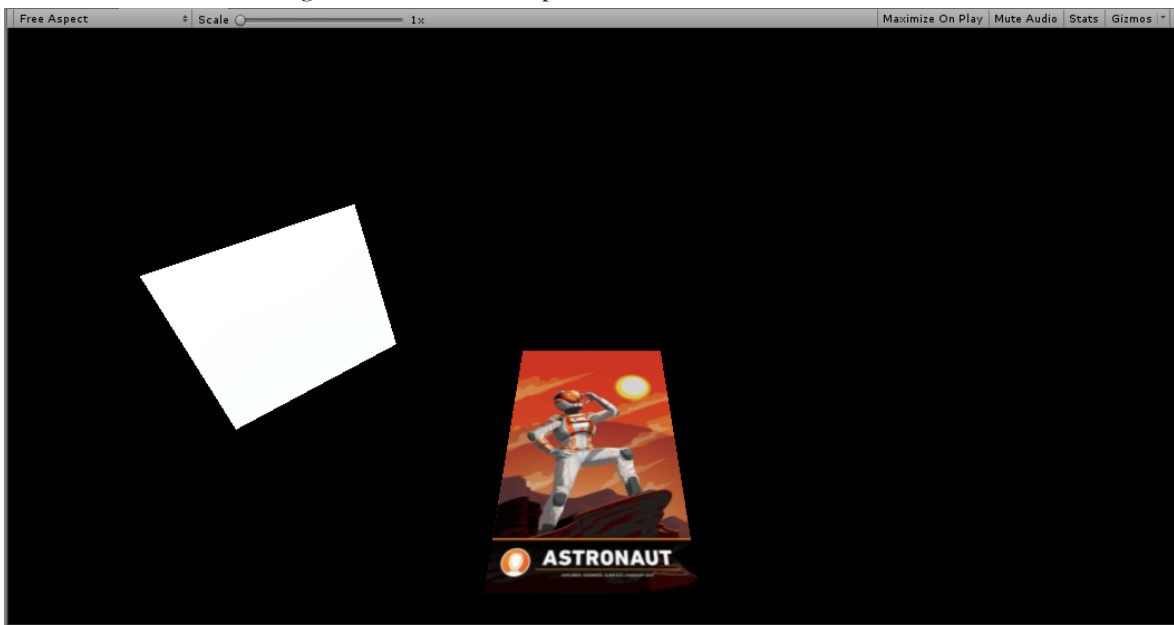


Fig 2. Plano generado sobre la imagen y posteriormente trasladado con el gesto de un solo dedo.

Además, se experimentó también con la adición de eventos que permitieran cambiar de color los objetos con un toque o tap.

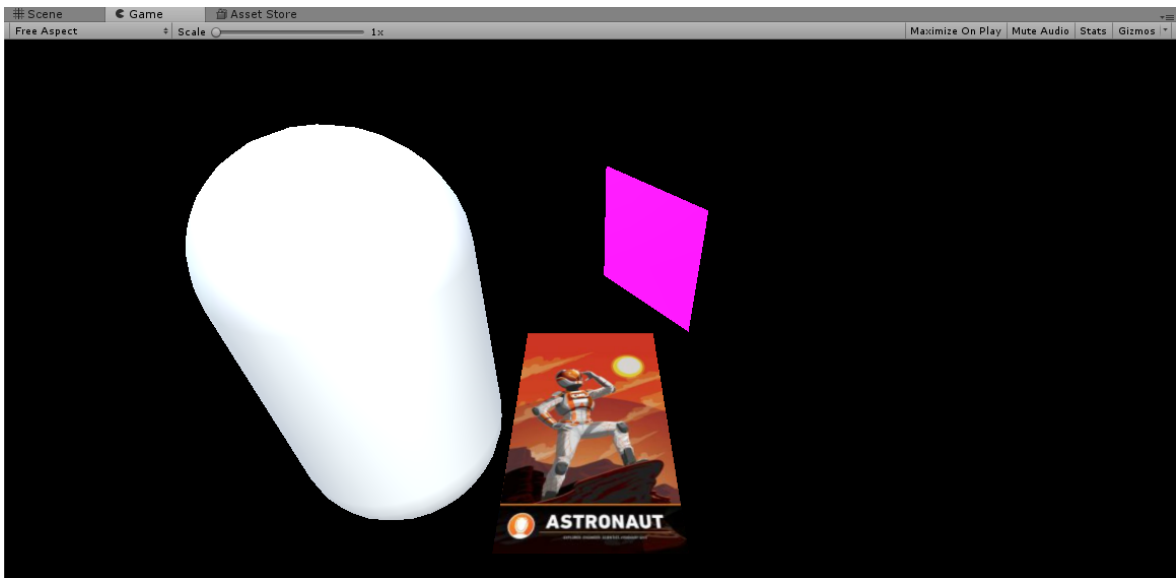


Fig 3. Plano cambiado de color y rotado junto a un objeto con forma de píldora escalado.

El siguiente paso fue más allá de la interacción con los objetos creados en la detección de targets, y se probó la interfaz básica de Unity 3D para superponer texto y botones a la cámara del teléfono. En su propio apartado se hablará del proceso que siguió la interfaz de usuario y las decisiones de experiencia de usuario que fueron tomadas a lo largo de este.

### 4.1.2 Problemas de rendimiento

Todo esto hizo surgir un problema común si no se trata como es debido, que es el rendimiento de la aplicación. Debido al uso del asset LeanTouch para las interacciones con los objetos, cuando se movía la cámara, la aplicación seguía ese movimiento con un par de segundos de retraso, consiguiendo una experiencia de usuario un tanto incómoda. La solución consistió en cambiar los eventos de LeanTouch por otros realizados nativamente en Unity sin emplear assets.

Aunque el rendimiento se vea directamente afectado por el terminal usado, Vuforia nos deja priorizar sus recursos, en nuestro caso, priorizar los FPS (imágenes por segundo) no afecta notoriamente a la calidad de imagen y conseguimos un ligero aumento de fps que hace que nos olvidemos completamente del retraso que se comentó anteriormente.

### 4.1.3 Tracking de los targets

Un detalle que Vuforia consigue resolver con buenos resultados, es ponerle solución al problema del “popping” o la carga repentina de los objetos asociados a un target. ¿Qué ocurre si apartamos la cámara del target?, dichos objetos desaparecerán de nuestra pantalla y por ende

dejarán de estar cargados, lo que ocasiona que al volver a entrar el target en la cámara aparezcan de repente completamente visibles o que nos sea necesario apuntar bien al target para que lo detecte.

Vuforia mejora este apartado con el extended tracking o seguimiento extendido, disponible con un simple click en su configuración. El extended tracking sigue manteniendo los objetos, aunque no aparezcan en pantalla, y calcula respecto al movimiento del teléfono en qué lugar deberían encontrarse, haciendo que parezcan progresivamente por un borde de la pantalla si poco a poco entra el target de nuevo en ella, simulando de cierta forma lo que nos ocurre con cualquier objeto real al apuntarlo con una cámara.

Empleando esta opción, se mantiene de forma mucho más fiable el tracking y permite mover la cámara sin necesidad de estar reconociendo a la perfección el target, simplemente en caso de no encontrarlo calcula la posición en la que debería estar. Esto podría significar una reducción del rendimiento de la propia aplicación, pero en caso de ser así no es notable sin emplear métricas y comparaciones exactas. El extended tracking, como cualquier sistema, no es perfecto lo que implica que pueda calcular erróneamente la posición del target y generar situaciones en las que se incline el objeto virtualizado o problemas similares.

#### 4.1.4 Añadir información a targets

Una vez que se ha experimentado con los objetos, se decide probar interfaces asignadas a targets, haciendo que estos sean fuentes de información de fácil acceso, y dado que podemos usar extended tracking, no debería ser un problema ver interfaces de un tamaño mayor.

El proceso es el mismo, solo que el objeto es de interfaz de usuario, con el texto que se desee, en nuestro caso se decidió emplear un texto blanco sobre un fondo negro semitransparente, para que haga como si fuese un holograma con texto sólido.

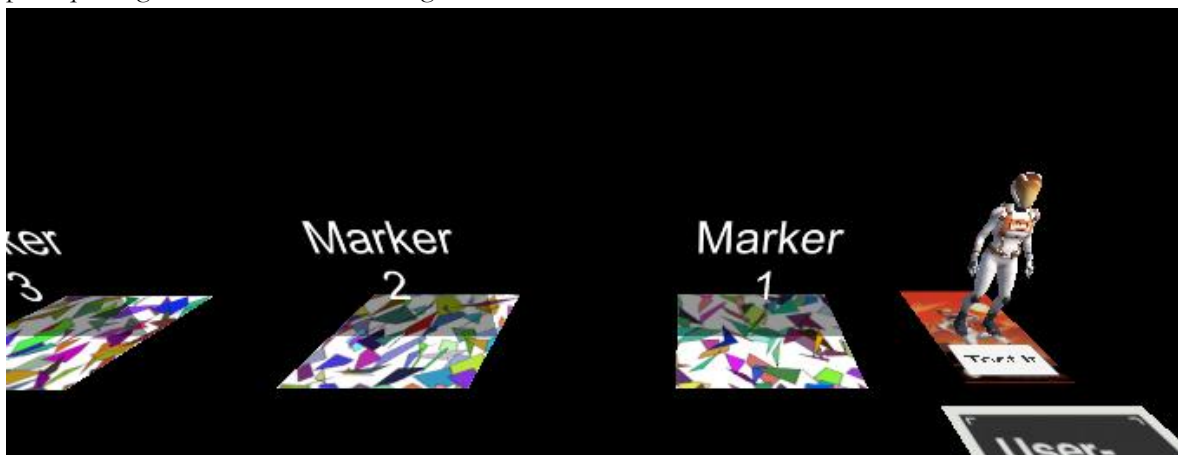


Fig 4. Textos asignados a targets generados aleatoriamente. A un lateral se encuentra un ejemplo de botón por encima del target que el usuario puede pulsar para que el personaje se oriente de cara al usuario.

## 4.2 Vídeo remoto

Una de las primeras ideas fue la de enviar vídeo en tiempo real de un teléfono a otro, dado que pueden darse indicaciones muy fácilmente si en tiempo real puede verse lo que ocurre y el estado del lugar.

Primero analizaremos las limitaciones existentes con respecto a la comunicación de red aportada por Unity. Hay dos formas de enviar imágenes, una de ellas es mediante TCP y otra mediante UDP. UDP permite enviar poca cantidad de información (imágenes de 128x128 píxeles), lo que implica un tamaño de imagen demasiado pequeño como para poder ver con claridad lo que ocurre, y reconstruir una imagen de mucho mayor tamaño con pequeños fragmentos ralentizaría demasiado el vídeo. Esto nos deja con la opción de emplear TCP como transporte de la información, lo que nos limita a un tamaño de imagen de 512x512 píxeles, algo con lo que podría trabajarse. También hay que mencionar que unas pocas personas han intentado esto mismo, tanto uniendo imágenes más pequeñas como hacerlo directamente empleando TCP, y el resultado es desalentador, ya que en caso de que funcionase daba como consecuencia un consumo de RAM descontrolado o una ralentización considerable del vídeo transmitido.

A continuación, trataremos las limitaciones que nos impone Vuforia, ya buena parte de las funciones las lleva a cabo este framework. Vuforia, el cual toma el control de la cámara del smartphone, no permite obtener la cámara o el flujo de vídeo que procesa, con lo cual nos deja con una caja negra en la cual sabemos lo que entra, la cámara, y sabemos lo que sale, las imágenes con realidad aumentada. Esto dificulta la posibilidad de conseguir implementar el streaming de vídeo, ya que por este lado es imposible hacer nada.

Finalmente, Unity no está diseñado para hacer streaming de vídeo, además hay que añadirle que no existen assets que sean capaces de esto mismo, lo que nos dejaría con un complejo desarrollo de un sistema de streaming para android fuera de Unity. Esta situación implica el descarte de la idea, por ser imposible con las herramientas propuestas.

## 4.3 Generación dinámica o aleatoria de targets

Hay varias formas de obtener targets tanto de objetos como de imágenes, en este apartado se tratará la generación dinámica y la generación aleatoria de targets de imágenes.

Las imágenes de la figura 4 empleadas como targets para los textos han sido generadas aleatoriamente por una herramienta online dedicada a generar imágenes para realidad aumentada de forma aleatoria en función de parámetros como tipo de figuras a emplear y su cantidad. Esto permite generar muchos targets sin que sean iguales unos a otros y de manera que sean fácilmente reconocibles.

Respecto a la generación dinámica de targets, Vuforia tiene un constructor de targets que mediante un solo botón mientras apuntamos a una imagen o patrón diferenciable de sus alrededores podemos asignarle unos objetos preestablecidos.

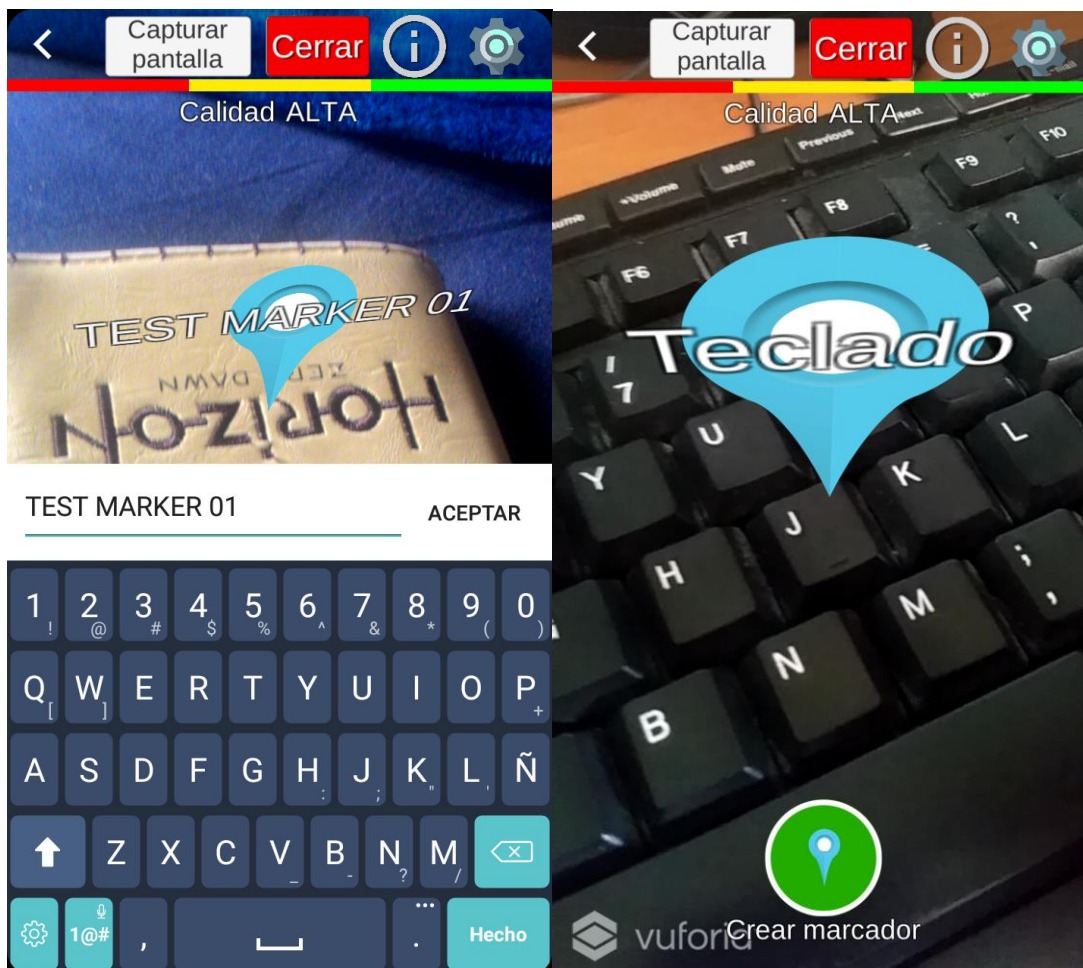


Fig 5 y 6. Visión de la aplicación en modo generación de targets. El botón verde comienza la generación sobre lo que se esté apuntando siempre y cuando el medidor de calidad de características (barra de colores en la imagen) muestre su segundo segmento de color, es decir, calidad media.

La calidad de características se encuentra indicada por la barra superior que se va coloreando según la cantidad de características que detecte. Dicha calidad depende de factores como el enfoque y la luz, siendo capaz de distinguir entre una buena superficie detectable y una mala.

Como ejemplo se utiliza una bombona de oxígeno como objeto a añadir a dichos targets generados, aunque es posible asignar otro tipo de objetos como interfaces.

Un buen uso de esta funcionalidad, el cual puede verse implementado en las figuras 5 y 6, es el de poder generar targets con texto editable y guardarlos posteriormente en el teléfono para ser

cargados en cada ejecución. A la hora de la práctica la carga de targets es viable y se encuentra soportada por la documentación de Vuforia mientras que la posibilidad de guardarlos es nula, es decir, a pesar de poder generarlos y cargarlos Vuforia no permite guardarlos para futuras sesiones.

A pesar de los inconvenientes puede llegar a ser útil en caso de querer realizarse alguna indicación temporal ya que mientras no se cierre la aplicación estos seguirán en memoria.

Finalmente, se probó a generar targets aleatorios y reducirlos a un tamaño de 3x3 cm para utilizarlos como si fuesen pegatinas pequeñas. El resultado fue satisfactorio, no tiene problema en detectar los targets tanto creados como predefinidos. Esto posibilita tener pegatinas sin asignar o asignadas para ponerlas donde el usuario quiera, con la finalidad que se quiera.

## 4.4 Reconocimiento de Objetos

El reconocimiento de objetos es la característica más llamativa de todo lo que nos puede ofrecer Vuforia, dando la posibilidad de no solo reconocer imágenes planas, sino que también podemos detectar objetos 3D previamente incluidos en el proyecto.

Vuforia emplea el nombre de ModelTarget a aquellos targets que son objetos 3D, y la diferencia de un target común o de un conjunto de targets. Para crear dichos targets, Vuforia da la posibilidad de hacerlo de dos formas:

- Escanear manualmente un objeto 3D real con la cámara del smartphone, para ello crearon una aplicación llamada Scanner que nos guía durante el proceso de toma de datos desde todos los ángulos posibles. Mediante puntos, define como es la forma del objeto y luego se basa en ellos para su detección.
- Cargar un modelo 3D o plano CAD para obtener su forma y cargarlo a Unity. Vuforia pone a nuestra disposición un programa con uso limitado (10 proyectos) para poder generar los ModelTargets a partir de modelos 3D, permitiendo elegir la posición que se tomará como referencia en la detección.



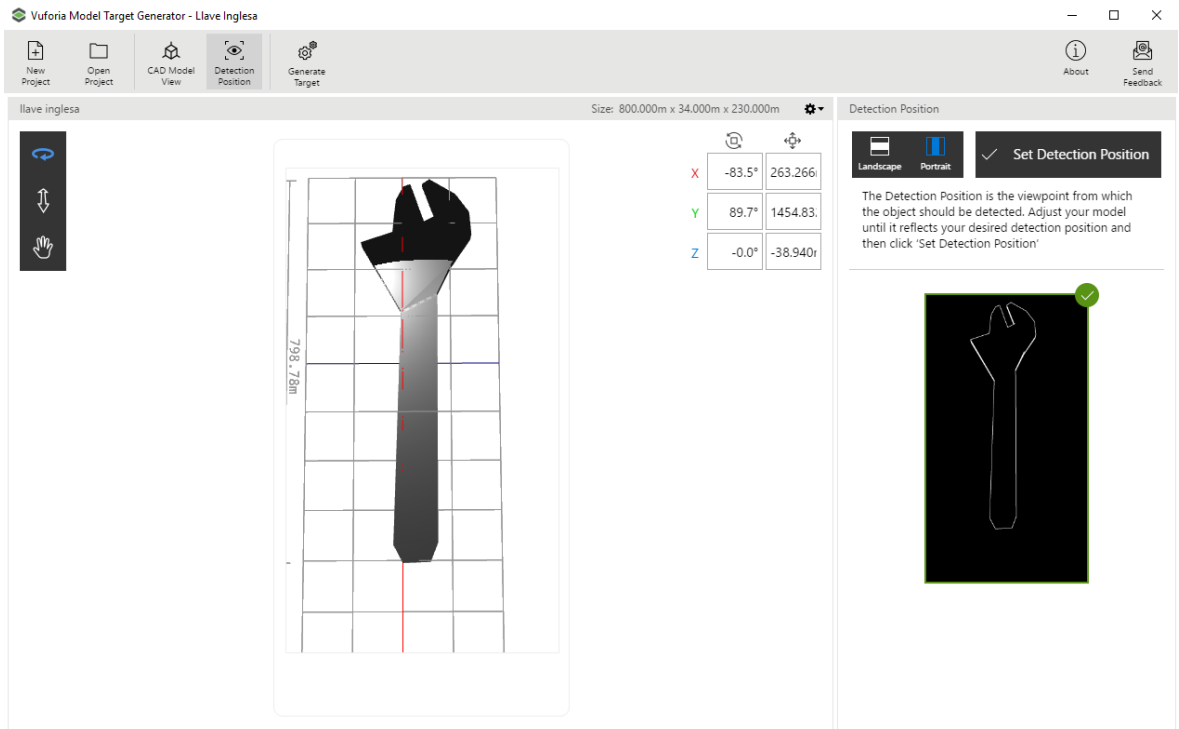


Fig 7. Vista del programa de generación de targets a la hora de definir la posición de detección.

Ante estas opciones, se hicieron pruebas para ver cuán precisos eran los resultados y tomar el que mejor detección ofrezca en el proyecto.

Al realizar pruebas con la primera opción desde objetos pequeños como conectores, hasta objetos grandes, los resultados eran demasiado negativos y el proceso era demasiado tedioso como para poder ser implementados de esta forma. A mayor cantidad de puntos más difícil le era detectar un objeto desde una posición que no fuera la inicial. Con un smartphone con una mejor cámara (Samsung Galaxy S7) este resultado se repetía continuamente. La única forma en la que fue capaz de detectar apropiadamente un objeto fue tomando puntos de la parte frontal de este y no de las demás perspectivas, esto implica que el extended tracking debe estar activado para seguir con relativa precisión al objeto detectado.

Dado los problemas que ha dado el primer método se procedió a probar el segundo con los siguientes objetos:

- **Figura de un Tie Fighter de las películas Star Wars:** Se tomó este ejemplo principalmente debido a que tiene una forma que es un conjunto de formas geométricas bien definidas y posee detalles que lo diferencian ampliamente de su entorno. También se tomó como ejemplo dado que al pertenecer a una película

tan famosa encontrar modelos 3D precisos y gratuitos es mucho más fácil que en otros casos.

- **Llave inglesa de 10 pulgadas:** Esta llave tiene una forma normalizada muy común y con un tamaño mayor al que tienen otros objetos empleados lo que hace que se creen situaciones distintas en la detección ya que no es tan fácil de encuadrar en la cámara o evitar brillos indeseados, y dependiendo de la calidad de esta y de su entorno puede llegar a afectar negativamente a la detección. También tiene peso el hecho de ser una herramienta, algo que es más cercano a un caso real.
- **Tuerca:** A pesar de ser un objeto pequeño, detalles como los surcos para enroscarlo y su forma hexagonal ayudan mucho a la hora de detectarla incluso a cierta distancia, cuando llega a ser un objeto pequeño en pantalla.

Los resultados en esta ocasión fueron muy satisfactorios, ya que sin extended tracking Vuforia era capaz de mantener la detección hasta los 45° de desviación con respecto del objeto detectado. La limitación de este método reside en que toma una perspectiva para tomar como referencia para la detección lo que implica que la mejor forma de detectar inicialmente un objeto es desde dicha perspectiva. La velocidad de detección es muy buena, apenas tarda 1 segundo en mostrar el modelo de referencia una vez detectada la forma, lo que lo hace muy cómodo al usarse. También, ya que emplea la forma de los bordes del objeto como referencia, diferencia fácilmente entre el objeto y el ruido de fondo.



Fig 8. Llave inglesa detectada al ser sujetaada delante de la cámara. Modelo de referencia superpuesto con un tamaño inferior, al ser la superficie detectada menor que la real.

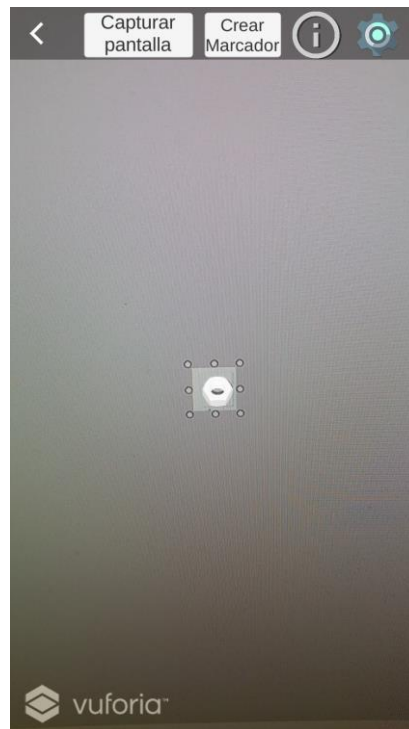


Fig 9. Tuerca muy pequeña detectada sobre una pantalla. El modelo se superpone correctamente.

En el caso del Tie Fighter, por su complejidad y elementos que tapan a otros, la perspectiva desde la cual se detecta con una iluminación adecuada es bastante menor que en la que tenemos con la llave inglesa, lo que supone que objetos con características que a veces son visibles y otras no en su representación 2D puedan dar problemas, aunque también los resultados se encuentran fuertemente ligados a la calidad de la cámara y a la iluminación. En el caso de la iluminación puede llegar a ser bastante permisiva la detección ya que se ha conseguido incluso detectar el Tie Fighter en plena noche con la simple luz del flash de la cámara del mismo smartphone, lo que ayuda enormemente a la hora de ponerlo en práctica.

En el caso de la tuerca, puede deducirse que el tamaño del objeto no importa, aunque este limita la cantidad de características en las cuales puede apoyarse la detección. Todo esto implica que algunos objetos pequeños no lleguen nunca a ser detectados, al ser demasiado simples y carentes de características que los diferencien de otros.

Un problema que surgió durante las pruebas con varios ModelTargets activados simultáneamente es la limitación que impone Vuforia sobre estos, ya que solo permite tener activo un ModelTarget a la vez, debido a que si hay más, todos los modelos serán cargados sobre cualquier objeto que detecte, haciendo muy incómoda esta característica. Para solucionarlo se intentó crear un script para cambiar el ModelTarget activo de forma dinámica con un toggle u otro elemento de interfaz que lo permita. Se concluyó que no es posible realizar este script ya que el objeto ObjectTracker no puede ser obtenido durante ejecución ya que siempre devuelve null por motivos que se desconocen. A día de hoy, nadie ha sido capaz de conseguir este objeto para las versiones actuales de Vuforia, aunque puede que en el futuro esto se pueda solventar fácilmente.

## 4.5 Chat VoIP

Una de las funcionalidades establecidas para la aplicación fue la de un chat de voz mediante IP.

Existen dos posibles implementaciones de este tipo de servicios:

- Emplear P2P para la comunicación directa entre dos usuarios: Esto implicaría conocer información sobre otro usuario como su IP para poder establecer comunicación entre ambos dispositivos. Puede resultar más complejo en uso y en implementación, ya que no hay assets o APIs que permitan una implementación sencilla.
- Comunicación cliente-servidor: A diferencia del P2P, este método implica tener un intermediario, aunque posibilita la creación de salas de conversación de una forma más sencilla y no es necesario conocer la IP de los usuarios para contactar con ellos.

Se decidió seguir la implementación de C/S, ya que se encontró un asset que permite realizar llamadas VoIP de forma fácil y rápida, con soporte para salas de conversación o conversación directa entre los usuarios conectados. El asset en cuestión es Photon PUN Voice, y dado que se basa en la nube, ofrecen una versión gratuita hasta para 20 usuarios simultáneos, más que suficiente para el caso que estamos tratando.

El servidor se encuentra en la nube, y el asset se encarga de automatizar todo lo relacionado con las conexiones a este, facilitando enormemente la implementación de esta funcionalidad.

En la práctica se comporta correctamente y con una calidad de audio al nivel de servicios como Skype o Discord tanto por wifi como por datos móviles, lo que garantiza una fluida conversación sin interrupciones.

Cuando se inicia la aplicación, Photon intentará conectarse a una sala, si esta está llena o no existe ninguna, crea una nueva.

Siguiendo el ejemplo de los desarrolladores, cuando se une a una sala aparecen a un lado botones push-to-talk o “pulsar para hablar”. Por defecto aparece un botón para hablar a todo el que esté en la sala, y si hay otros usuarios en la misma sala aparecerán botones con sus correspondientes nombres para poder hablar solo al que pulses de forma privada. Por defecto se ha dejado abierto el micrófono para toda la sala, lo cual hace más sencillo probarlo.

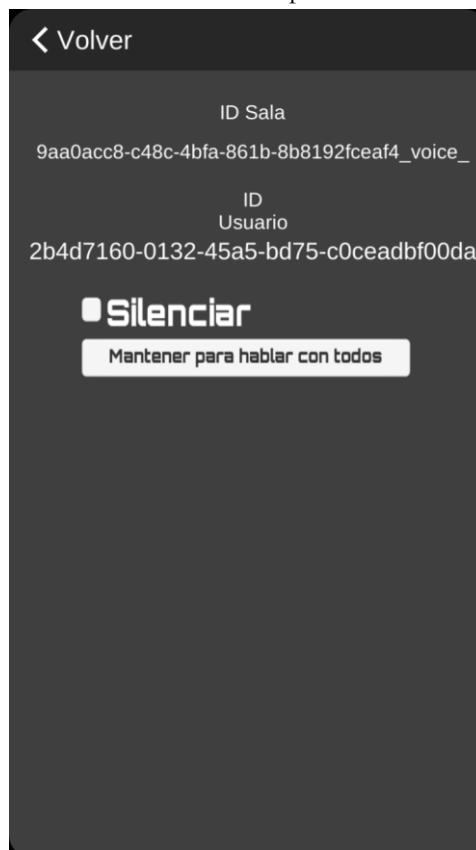


Fig 10. Menú del chat VoIP con opción para silenciar a todos o hablar a todos al mantener pulsado.

## 4.6 Arquitectura y funcionalidades finales

Una vez vistas las principales funcionalidades de la aplicación, se procederá a establecer cuál ha sido la arquitectura final y las funcionalidades que se han implementado.

Para comenzar, en lugar de tener un streaming de vídeo de un dispositivo a otro, se optó por mandar capturas de pantalla con un botón, las cuales, al ser compartidas por otros medios nos evaden la dependencia de servidores propios. Esto se combina con las rejillas de coordenadas, las cuales se tratarán junto a la interfaz final en los siguientes apartados.

Para dar la posibilidad a los usuarios de comunicarse entre ellos se implemento un chat VoIP mediante el framework Photon PUN Voice, el cual se basa en la nube. Esto supone que, en lugar de tener servidores propios, los cuales necesitan mantenimiento y generan gastos, se dependa de la nube que ofrece este servicio VoIP, otorgando una buena calidad de llamada incluso en redes móviles. A esto hay que añadirle la personalización mediante salas o conversaciones privadas que permite.

Finalmente, la detección de targets, objetos reales o la generación de targets dinámica, son funcionalidades que vienen dadas por Vuforia, que en este caso, trata todo de forma local ya que las funcionalidades que requieren de la nube no se han implementado en el proyecto.

Todo esto consigue que no se necesiten servidores propios, ya que en el único caso en el que se necesitarían se depende de la nube, que ofrece un servicio en parte gratuito con un rendimiento y personalización muy superiores al que se conseguiría creando un servidor propio con la experiencia actual.

## 4.7 Interfaz de Usuario

La interfaz de usuario es la que permite que todas las funcionalidades, y otras de las que se hablarán más adelante, puedan usarse de forma cómoda sin superponerse unas a otras o sin saturar la vista del usuario.

## 4.7.1 Menús

La aplicación se divide en varias escenas, y cada escena tiene una función concreta:



- **Escena de splash:** Muestra el logo de Vuforia, lo que indica en qué tecnología se soporta. Para un inicio más rápido puede obviarse esta escena. El splash nos lleva al menú principal cuando acaba.
- **Menú principal:** Aunque minimalista, el menú principal nos da a elegir dos opciones, una de ellas es la del usuario que está en remoto (Cámara + Voz), y la otra es solo chat de voz, ideada más para la persona que da indicaciones y no quiere gastar batería empleando la cámara si no la va a usar.

Fig 11. Menú principal con ambas opciones.



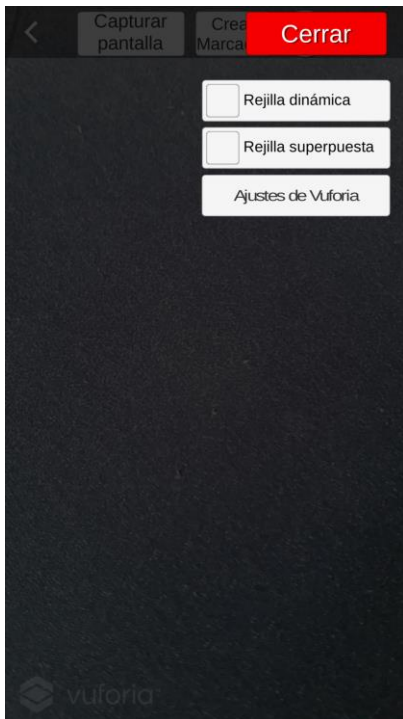
- **Escena Cámara + Voz o Main:** Esta escena es la principal dentro de toda la aplicación, ya que es la que engloba todas las características disponibles para el remoto. Cuenta con un botón para volver al menú principal.
- **Escena Chat de voz o MainVoice:** Escena que sólo permite hablar mediante VoIP. Su apariencia se puede ver en la figura 10. Al igual que la escena Main, tiene un botón para volver al menú principal. También aporta información sobre la ID del usuario y la sala en la que se encuentra, además de los controles “push-to-talk” previamente mencionados.

Fig 12. Escena Cámara + Voz o Main. Aspecto inicial.

Cabe destacar que la escena Main tiene su propia interfaz, la cual es copiada para la escena MainVoice. Dicha interfaz proporciona lo siguiente:

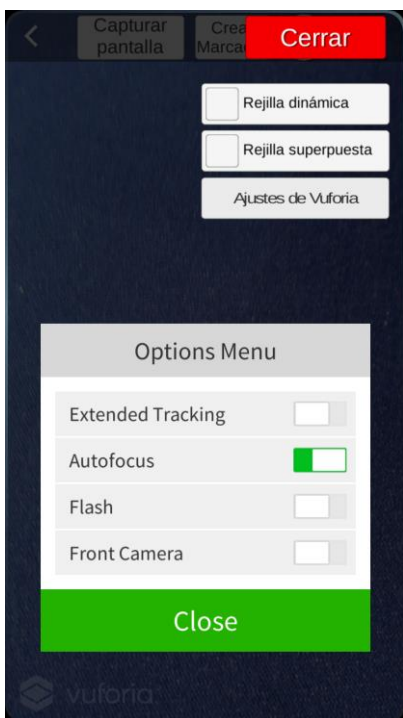
- **Un botón de ajustes:** Representado con un engranaje, un icono usado comúnmente para indicar los ajustes del teléfono en el dock de aplicaciones. Los ajustes permiten activar o desactivar cualquiera de las dos rejillas de coordenadas implementadas de las que se comentará al respecto en apartados siguientes. Además, incorpora un botón para abrir el menú de configuración de Vuforia.





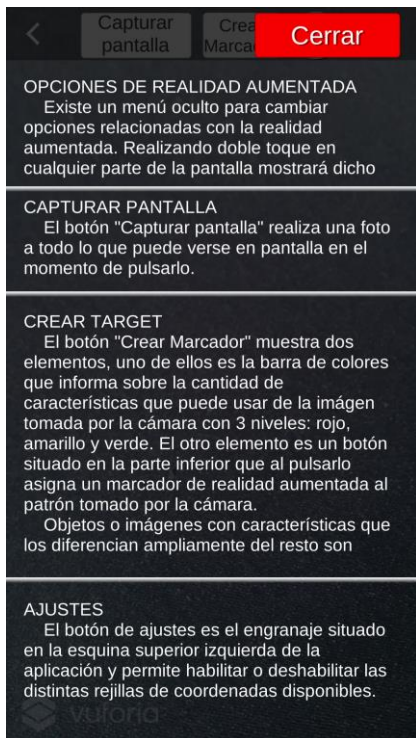
- **Un menú de configuración Vuforia:** Tomado de los ejemplos de Vuforia, este menú se activa haciendo doble toque en cualquier parte de la pantalla, abriéndose un menú para cambiar ajustes de Vuforia como el uso de linterna, auto enfoque, extended tracking o cambiar de cámara.

Fig 13. Menú de ajustes con fondo semitransparente y ambas opciones sin activar. También se muestra el botón para abrir la configuración de Vuforia..



- **Botón de captura de pantalla:** Botón que permite realizar capturas de pantalla y compartirlas por cualquier aplicación que soporte la función Share, permitiendo elegir al usuario el mejor medio del que disponga.
- **Botón para activar la creación de targets:** Al pulsarse, aparece un botón verde en la parte inferior con un marcador azul que si se pulsa creará un target a partir de lo que esté enfocándose con la cámara, si hay suficiente calidad, es decir, si existe una cantidad aceptable de características para realizar detecciones fiables. La barra indicadora de la calidad aparecerá a la vez que el botón, pero debajo de la barra de opciones superior.

Fig 14. Menú de Vuforia, con auto enfoque activado.



- **Botón de información:** Botón situado junto al de ajustes que informa sobre cómo usar la aplicación.

Fig 15. Menú de información.

## 4.7.2 Capturas de pantalla

Como se ha mencionado anteriormente, se ha implementado un botón que permite hacer capturas de pantalla, y además permite compartirlas por la aplicación que el usuario desee.

Esto permite enviar fotos de lo que el remoto está viendo incluyendo también todo lo que realiza Vuforia en su pantalla. Su principal utilidad se encuentra estrechamente relacionada a las rejillas de coordenadas, que entre ambas funcionalidades, podrán enviarse indicaciones desde cualquiera de los teléfonos, sin importar quién sea el usuario remoto.

## 4.7.3 Rejillas de coordenadas

Finalmente se consideró la adición de una funcionalidad adicional que permitiese indicar sobre las imágenes que toma la cámara. Esto dio lugar a las rejillas de coordenadas.

El principal uso, sin importar en qué lado se esté, es poder resaltar o comunicar la localización de algo en concreto. Un ejemplo sería del usuario remoto buscando un tornillo entre varios, dicho usuario tiene dudas de si realmente es ese tornillo el que tiene que quitar, la aplicación le permitiría, en combinación con la captura de pantalla, mandar una captura con la rejilla superpuesta y decirle al que le da indicaciones si el tornillo en el cuadrante A3 es realmente el que tiene que quitar.

En espacios pequeños puede ser realmente útil ya que por ejemplo en la electrónica los componentes son muy pequeños y se encuentran muy concentrados, y usualmente es fácil perderse si hay componentes que se parecen o dificultan la búsqueda. Es un buen apoyo para la detección de objetos, permitiendo encontrar al usuario lo que la detección no pueda resaltar.

También puede usarse por parte de quien dá las indicaciones, ya que, si tiene algún modelo de referencia o un objeto similar, puede emplear o la captura de pantalla o ambas funcionalidades para indicar que busque X objeto o lugar.

A raíz de esta idea se ideó una rejilla que funcionase como una imagen con transparencias que se adaptase al tamaño de la pantalla.



Fig 16. Rejilla basada en una imagen redimensionable.

A continuación, se pensó que si el usuario pudiera resaltar las casillas que él quisiera de un color distinto podría ser más rápido que decir en qué cuadrante o cuadrantes se encuentra lo que quiere indicar. Esto da lugar a otra rejilla creada dinámicamente que se forma con un número de columnas y filas fijo y se rellena mediante checkboxes, cuyo interior es transparente y sus bordes

unidos forman la rejilla. Al pulsar un cuadrante este colorea sus bordes de color rojo, resaltando lo que hay dentro.

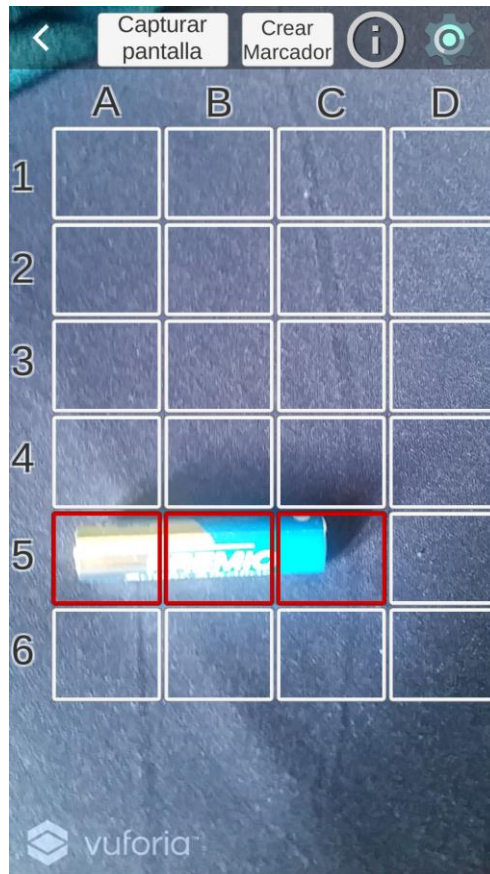


Fig 17. Rejilla dinámica con algunos cuadrantes resaltados.

#### 4.7.4 Accesibilidad Básica

Para facilitar el uso de la aplicación a cualquier persona, se ha tenido en cuenta algunos aspectos relativos a la interfaz:

- **Contraste del texto:** El texto que se encuentra superpuesto a la cámara, al estar expuesto a un fondo cambiante, tiene un contraste variable y esto puede dificultar su lectura. Para solucionar esto se le aplicó un borde blanco que hace que sea legible en cualquier situación en la que se encuentre el fondo.
- **Botones con función explicada con texto:** En este caso el botón en concreto que recibió este ajuste fue el de crear los marcadores, ya que es un botón que no es de uso

extendido como un botón de ajustes o para hacer fotos. Lo que se hizo al respecto fue incluir un texto debajo de este que explica brevemente su función. Aquellos botones que ya cuenten con texto en su interior o cuyo icono sea fácilmente reconocible no han recibido esta característica.

- **Texto para el indicador de calidad de características:** Hay que tener en cuenta que las barras de colores que tan fácilmente nos indican la calidad de las características detectables pueden ser un problema para personas con dificultades para distinguir colores o daltónicas. Para ello se ha añadido texto con contraste debajo de la barra, indicando la calidad actual en tiempo real a medida que se emplea.
- **Botones “Cerrar”:** Los botones para cerrar los menús eran en un principio exactamente iguales a los demás, esto puede dar a entender al usuario que es una funcionalidad más o que al principio no sepa cómo cerrar alguno de estos menús. Para denotar su función de forma clara se colorearon los botones de color rojo, ya que como solemos estar acostumbrados en algunos sistemas operativos o programas, el color rojo suele ser el asignado a los botones de cerrar. Esto acompañado con la palabra “Cerrar” dan la información necesaria al usuario al respecto de su importancia y función.

Una gran adición a la aplicación sería la compatibilidad de esta con TalkBack. TalkBack es una aplicación de Google que lee la pantalla a personas con problemas visuales o ciegas. Esto no fue posible, ya que Unity no ha implementado dicha compatibilidad lo que impide a TalkBack leer la interfaz.

## 5 Coste del Proyecto

El presente apartado contiene la estimación presupuestaria realizada para el desarrollo de este proyecto. La estimación del presupuesto de este proyecto se realiza teniendo en cuenta dos pautas: por un lado, se considera el precio de los equipos necesarios para el desarrollo del proyecto, y por otra parte se tiene en cuenta el coste que supone el tiempo empleado en la ejecución de dicho trabajo.

### 5.1 Presupuesto de Ejecución Material

Se denomina coste total de la ejecución material, a la suma del coste de los equipos y del coste por tiempo de trabajo.

#### 5.1.1 Coste de equipos

<b>EQUIPO</b>	<b>PRECIO</b>	<b>DURACIÓN</b>	<b>USO</b>	<b>TOTAL</b>
Ordenador	600,00 €	3 años	2 meses	33,33 €
Conexión a Internet	29 €	---	2 meses	58 €
Impresora	80,00 €	3 años	0,5 meses	1,11 €
Asus Zenfone 2 Láser	150,00 €	3 años	2 meses	8,33 €

Coste total de equipos	107,77 €
------------------------	----------

### 5.1.2 Coste por tiempo de trabajo

Suponiendo que el salario de un ingeniero en informática es de 12 euros la hora, el coste por tiempo de trabajo es el que se obtiene en esta tabla:

<b>FUNCIÓN</b>	<b>Nº HORAS</b>	<b>EUROS/HORA</b>	<b>TOTAL</b>
Ingeniería	300	12	3.600 €

Coste por tiempo de trabajo	3.600,00 €
-----------------------------	------------

### 5.1.3 Coste total de ejecución material

Es la suma de los importes del coste de materiales y de la mano de obra.

<b>CONCEPTO</b>	<b>COSTE</b>
Coste de equipos	107,77 €
Coste por tiempo de trabajo	3.600,00 €
Coste de ejecución material	3.707,77 €

## 5.2 Gastos Generales y Beneficio Industrial

Normalmente se trata de los gastos necesarios para disponer de instalaciones en las que desempeñar el trabajo, además de otros gastos adicionales. Los gastos generales y el beneficio industrial son el resultado de aplicar un recargo del 20% sobre el Coste Total de Ejecución Material, resultando:

Gastos Generales y Beneficio Industrial..... 741,55 euros

## 5.3 Presupuesto de Ejecución por Contrata

El presupuesto de ejecución por contrata es el resultado de sumar el coste total de ejecución y los gastos generales.

Presupuesto de Ejecución por Contrata..... 4.449,32 euros

## 5.4 Honorarios

Los honorarios facultativos por la ejecución de este proyecto se determinan de acuerdo a las tarifas de los honorarios de los ingenieros en trabajos particulares vigentes a partir del 1 de Septiembre de 1997, dictadas por el Colegio Oficial de Ingenieros Técnicos de Telecomunicación.

Importe	Coficiente reductor	Porcentaje
<i>Hasta 5 millones</i>	$C = 1$	7 %
<i>Desde 5 millones</i>	$C = 0,9$	7 %

Los derechos del visado se calcularán aplicando la siguiente fórmula:  $0,07 \times P \times c$ , donde P es el presupuesto de ejecución material y c es el coeficiente reductor.

DERECHOS DE VISADO	COSTE
$0,07 \times 4.449,32 \times 1$	311,45 €
Total honorarios:	311,45 €

## 5.5 Importe total del presupuesto

El importe total del presupuesto de este proyecto se calcula sumando el presupuesto de ejecución por contrata y los honorarios. A dicho valor se le aplicará el 21% de I.V.A.

*Presupuesto de Ejecución por Contrata* 4.449,32 €

*Honorarios* 311,45 €



SUBTOTAL	4.760,77 €
21% de I.V.A.	999,76 €
IMPORTE TOTAL	5.760,53 €

El importe total del proyecto asciende a la cantidad de NOVECIENTOS CINCUENTA Y OCHO MIL CUATROCIENTOS SETENTA Y UNO PESETAS, o bien CINCO MIL SETECIENTOS SESENTA EUROS CON CINCUENTA Y TRES CÉNTIMOS.958.471

# 6 Resumen, conclusiones y trabajos futuros

## 6.1 Resumen

El proyecto desde su inicio tiene como objetivo facilitar la comunicación desde la distancia entre dos personas, una de ellas con conocimientos técnicos del problema, y otra que se encuentra en el lugar del problema y necesita indicaciones.

Se toma Vuforia y Unity 3D como tecnologías a usar ya que permiten implementar una aplicación para Android y la cual use realidad aumentada con relativa facilidad, además de su extenso uso.

Una de las primeras funcionalidades pensadas para el proyecto fue la de implementar un sistema de streaming de vídeo tomando el feed de la cámara como vídeo a enviar y el otro smartphone como destinatario. Debido a cuestiones técnicas y a la baja experiencia, se tuvo que apartar la idea, ya que se convertía en algo muy difícil de manejar ya que de por sí Unity no está pensado para permitir este tipo de funcionalidad y la información al respecto es prácticamente nula, sin importar la antigüedad.

La alternativa directa a lo comentado anteriormente es el envío de capturas las cuales muestran lo que capta la cámara en un instante determinado y todo lo que se añade por encima con Vuforia y con Unity. Su implementación es mucho más sencilla y con un buen uso puede ser casi tan útil como un vídeo. Además, la consecuencia directa de esta implementación en comparación al streaming es la gran cantidad de datos que no se consumen y que posteriormente se emplean en el chat de audio por VoIP.

El chat de voz mediante VoIP emplea el asset Photon PUN Voice que ofrece una buena calidad de audio sin necesidad de servidores, ya que se basa en la nube y todos los clientes se conectan a su servidor, y emplea salas de chat lo que permite conversaciones de grupo. Trabaja sin problemas tanto en datos móviles como en redes Wi-Fi.

Con respecto a la interfaz, una característica más ideada con el objetivo de guiar con mayor precisión es la adición de rejillas activables para definir un espacio de coordenadas en la cámara. La primera rejilla implementada es una imagen superpuesta que se adapta al tamaño de pantalla, y cabe destacar que es estática. La rejilla final se compone de cuadrantes activables por el usuario para resaltar los que el usuario considerase en un color reconocible.

Finalmente, relativo a Vuforia, nos encontramos con la posibilidad de usar ImageTargets y ModelTargets. Un target es, como su nombre indica, un objetivo reconocible por Vuforia y que, en caso de encontrarse, activará lo que se haya determinado en cada caso, como hacer aparecer modelos 3D en realidad virtual, menús, sonidos u otros elementos. Los ImageTargets son como las tarjetas de realidad aumentada mientras que los ModelTargets son objetos 3D que Vuforia detectará en el mundo real, lo que requieren un proceso para poder digitalizarse para su uso posterior en el proceso.

Los métodos que Vuforia aporta para la detección de objetos 3D son la creación de puntos escaneándolos con el mismo teléfono o la adaptación de planos o modelos 3D de estos mismos objetos. Las pruebas dan un amplio beneficio al uso del segundo método dado que la calidad de detección empleando la aplicación Scanner de Vuforia deja mucho que desear.

## 6.2 Conclusiones

Al conocer en profundidad Vuforia y como trabaja, puede resumirse este framework como “caja negra” a nivel de desarrollo. Esta denominación es bastante precisa con relación a la realidad ya que Vuforia tiene multitud de funcionalidades, pero en las cuales no podemos inmiscuirnos, haciendo que podamos controlar qué input tiene y qué output nos arroja, pero sin ser capaces de cambiar su funcionamiento interno o complementarlo con nuevas funcionalidades que se ajusten a las necesidades del proyecto. Todo esto indica que aún queda mucho recorrido antes de que Vuforia pueda emplearse para proyectos ambiciosos o con necesidades concretas. Un caso que se ajusta a la baja personalización de frameworks como Vuforia se puede encontrar en este mismo proyecto, en la generación de targets dinámicamente. Vuforia permite cargar colecciones de targets, pero no guardarlos en el dispositivo o sincronizarlos con la nube que ofrecen.

El reconocimiento de objetos, es una tecnología fascinante, que es capaz de diferenciar entre distintas formas y reconocer objetos siguiendo unos patrones predeterminados. A pesar de ser tan útil, tiene sus inconvenientes, ya que, dependiendo de la calidad de los datos que recibe la cámara, puede dar resultados muy buenos o un completo desastre. Esto hace que elementos como

la iluminación o la propia apariencia del objeto sean factores muy determinantes de su grado de eficacia.

Al emplear servicios que se apoyan en la nube, también puede verse el potencial que llegan a tener, ya que la nube ha permitido en este proyecto comunicar dispositivos según la arquitectura cliente/servidor, sin necesidad alguna de desarrollar un servidor específicamente para este uso, y con unos costes asequibles para pequeñas o medianas empresas. Esto quiere decir, que los servicios en la nube permiten de forma sencilla implementar sistemas más complejos y además siendo reutilizables para otros proyectos, debido a que la API no cambia. La capacidad de la nube tampoco supone un problema, ya que los balanceos de carga y las redimensiones son llevados a cabo de forma automática por la empresa que otorga el servicio.

Para finalizar, puede reconocerse que la realidad aumentada no solo sirve como un medio de entretenimiento, ya que con trabajo puede llegar a ser una tecnología que puede aplicarse a cualquier campo, aunque esto puede depender de la creatividad del propio desarrollador.

## **6.3 Trabajos futuros**

Este proyecto abre la posibilidad a crear, en base a él, mejores versiones del concepto.

La primera mejora que puede realizarse al proyecto sería la posibilidad de llevar a cabo una librería de objetos detectables personal que puede irse ampliando según el usuario genera nuevos targets. Esto permitiría a cada usuario personalizar un espacio de trabajo adaptado a este y completamente móvil, ya que con apuntar al target deseado se mostraría la información que previamente se ha personalizado. También permitiría la posibilidad de dejar informes, instrucciones y otros grandes textos en simples imágenes detectables en cualquier lugar sobre la marcha, para que otra persona en el futuro no necesite pedir información adicional.

Otro gran paso que puede darse sobre esta base es realizar de forma externa a Unity un streaming de vídeo de la pantalla del teléfono, ya bien en forma de otra aplicación adaptada a las necesidades o desarrollando la misma base en otra plataforma que permita este tipo de funcionalidades de forma nativa.

Finalmente, la mejora del sistema de detección para soportar múltiples objetos sería una labor sencilla, ya que depende más de los desarrolladores de Vuforia que de la dificultad para

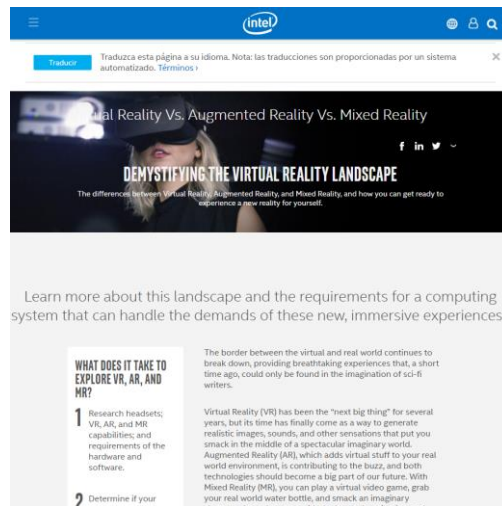
realizarlo. Activar o desactivar los objetos detectables sería una gran suma para obtener una aplicación con mayor utilidad ante todo tipo de casos.

# 7 Bibliografía

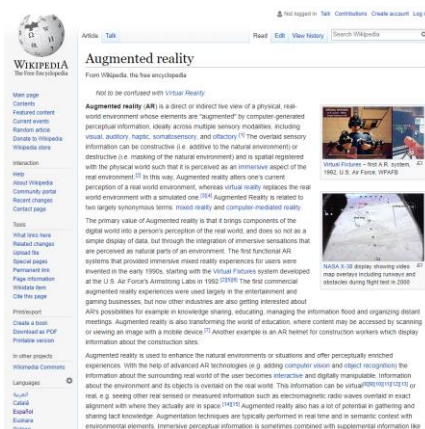
Para la realización de este proyecto se ha utilizado la siguiente bibliografía:

- Virtual Reality vs Augmented Reality vs Mixed Reality:

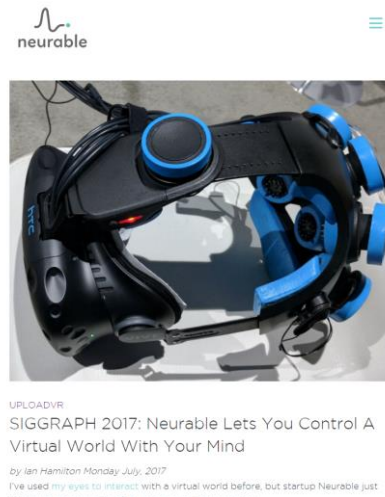
<https://www.intel.es/content/www/es/es/tech-tips-and-tricks/virtual-reality-vs-augmented-reality.html>



- Wikipedia Augmented Reality: [https://en.wikipedia.org/wiki/Augmented\\_reality](https://en.wikipedia.org/wiki/Augmented_reality)



- o Neurable y la interfaz cerebro-ordenador: <http://www.neurable.com/news/siggraph-2017-neurable-lets-you-control-virtual-world-your-mind>



- o Kevin Mise - Big AR: The Future of Augmented Reality: <https://hackernoon.com/big-ar-the-future-of-augmented-reality-c91c267b3d9f>

**HACKERNOON** News | Sign in | Get started

WELCOME! If you are reading this article, it's likely on a mobile device. You see, it is a growing trend that, *in leisure*, citizens of the Internet are using their smartphones and tablets more than their desktops (see below). Even more so, the gap between mobile and desktop *in the workday* continues to grow smaller by the day. It's easy to explain how mobile has slowly crept into the forefront, dominating most of our screen-time: *it's simple, it's portable, and it's accessible within just moments. Your content has never been more available to you than it is today on your mobile device.*

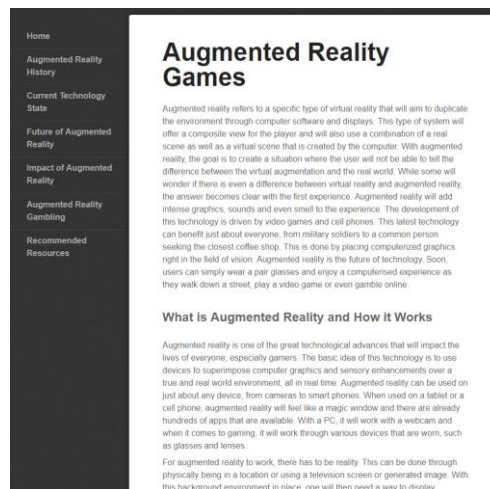
Mobile dominates morning and night. // Source: Smart Insights

In the Big AR series, we have discussed the implications of augmented reality *in its current state, on marketers, on gaming, and on the two current dominating mobile platforms...* But we haven't discussed the future of augmented

- o Wikipedia Google Glass: [https://en.wikipedia.org/wiki/Google\\_Glass](https://en.wikipedia.org/wiki/Google_Glass)



- o Augmented Reality Games. Historia, estado y futuro del de la realidad virtual: <http://www.augmented-reality-games.com/>





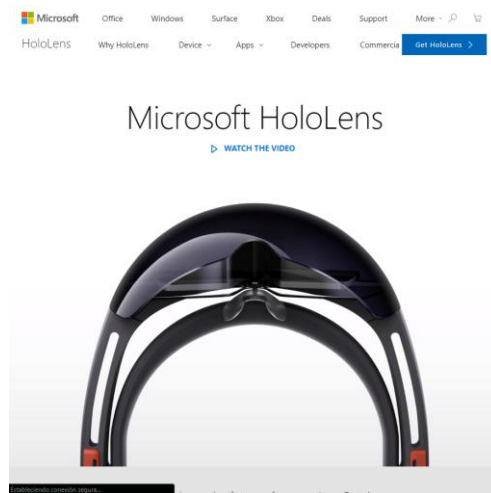
- Wikipedia Windows Mixed Reality: [https://en.wikipedia.org/wiki/Windows\\_Mixed\\_Reality](https://en.wikipedia.org/wiki/Windows_Mixed_Reality)



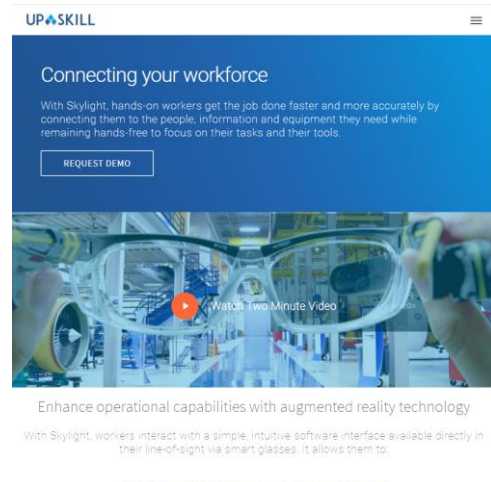
- Augmented Reality State of the Art - Michael R. Blumberg. <http://michaelrblumberg.com/2016/11/29/augmented-reality-state-art/>



- Microsoft HoloLens: <https://www.microsoft.com/en-us/hololens>



- Up-Skill Skylight: <https://upskill.io/skylight/skylight-overview/>

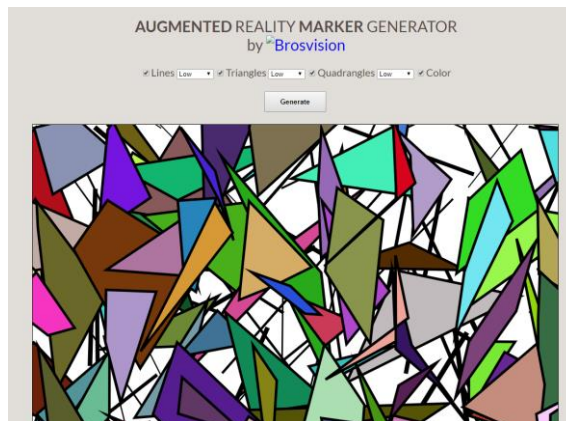


- Vuforia Chalk: <https://chalk.vuforia.com/>



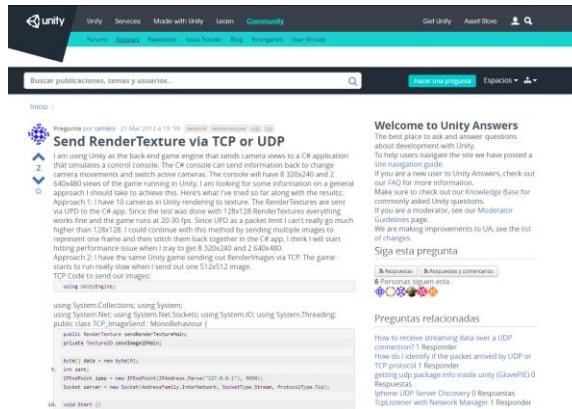
- Desarrollo

- Herramienta online para generar targets: <http://www.brosvision.com/ar-marker-generator>

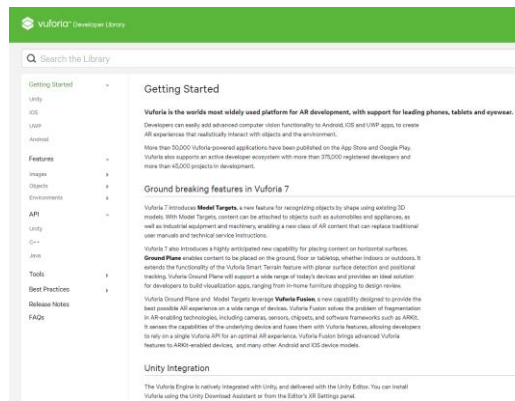


- TCP y UDP para el envío de imágenes:

<https://answers.unity.com/questions/230269/send-rendertexture-via-tcp-or-udp.html>

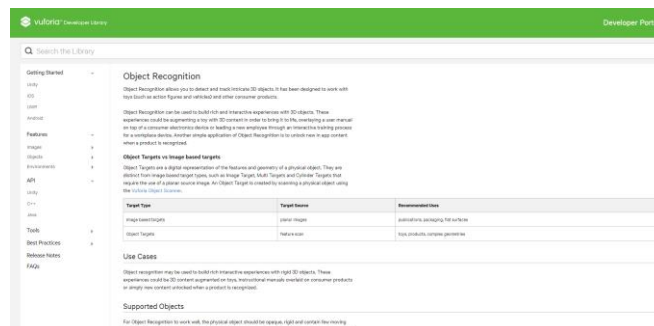


- Vuforia Getting Started: <https://library.vuforia.com/>

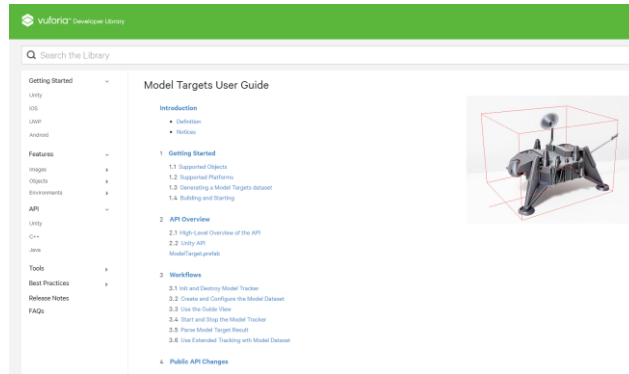


- Vuforia Object Recognition (limitaciones tecnológicas y cómo trabajar con ello):

<https://library.vuforia.com/articles/Training/Object-Recognition>

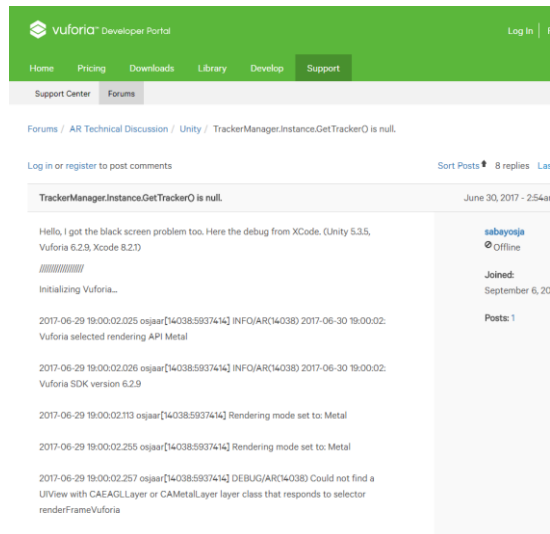


- Vuforia ModelTargets: <https://library.vuforia.com/articles/Solution/model-targets-user-guide.html>



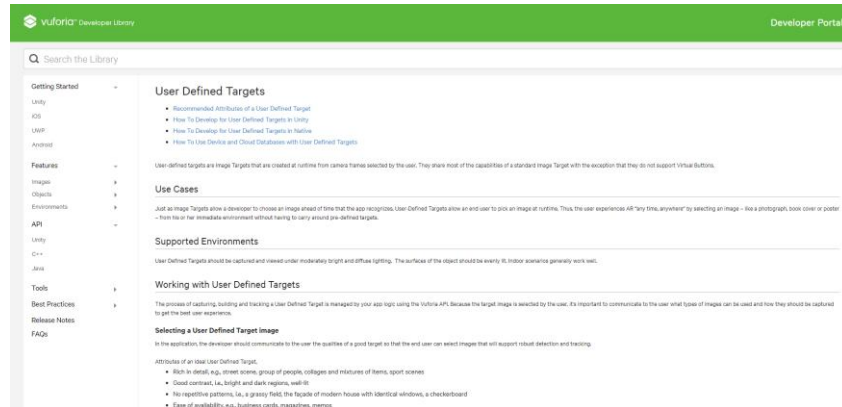
- Foro sobre el problema del ObjectTracker como null:

<https://developer.vuforia.com/forum/unity-extension-technical-discussion/trackermanagerinstancegettracker-null>



- Vuforia User designed Targets:

<https://library.vuforia.com/articles/Training/User-Defined-Targets-Guide>



- Presupuestos

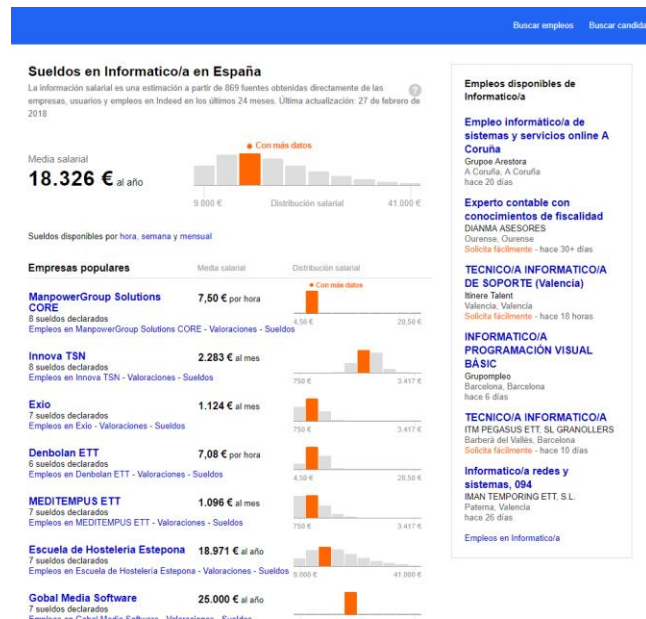
- Conversor de Euros a Pesetas: <https://www.cambioeuro.es/peseta/>

The screenshot shows the CambioEuro.es website. The main heading is "Convertor Euros Pesetas". Below the heading, there are two conversion examples:

Euro	1	=	166,386	Peseta Española
Peseta Española	1	=	0,0060	Euro

Below the conversion examples, there is a section titled "LA PESETA" which provides information about the Spanish Peseta, including its history and its replacement by the Euro. There is also a section titled "MONEDAS Y BILLETES" which lists the denominations of the Spanish Peseta. The website also features a navigation menu and a search bar.

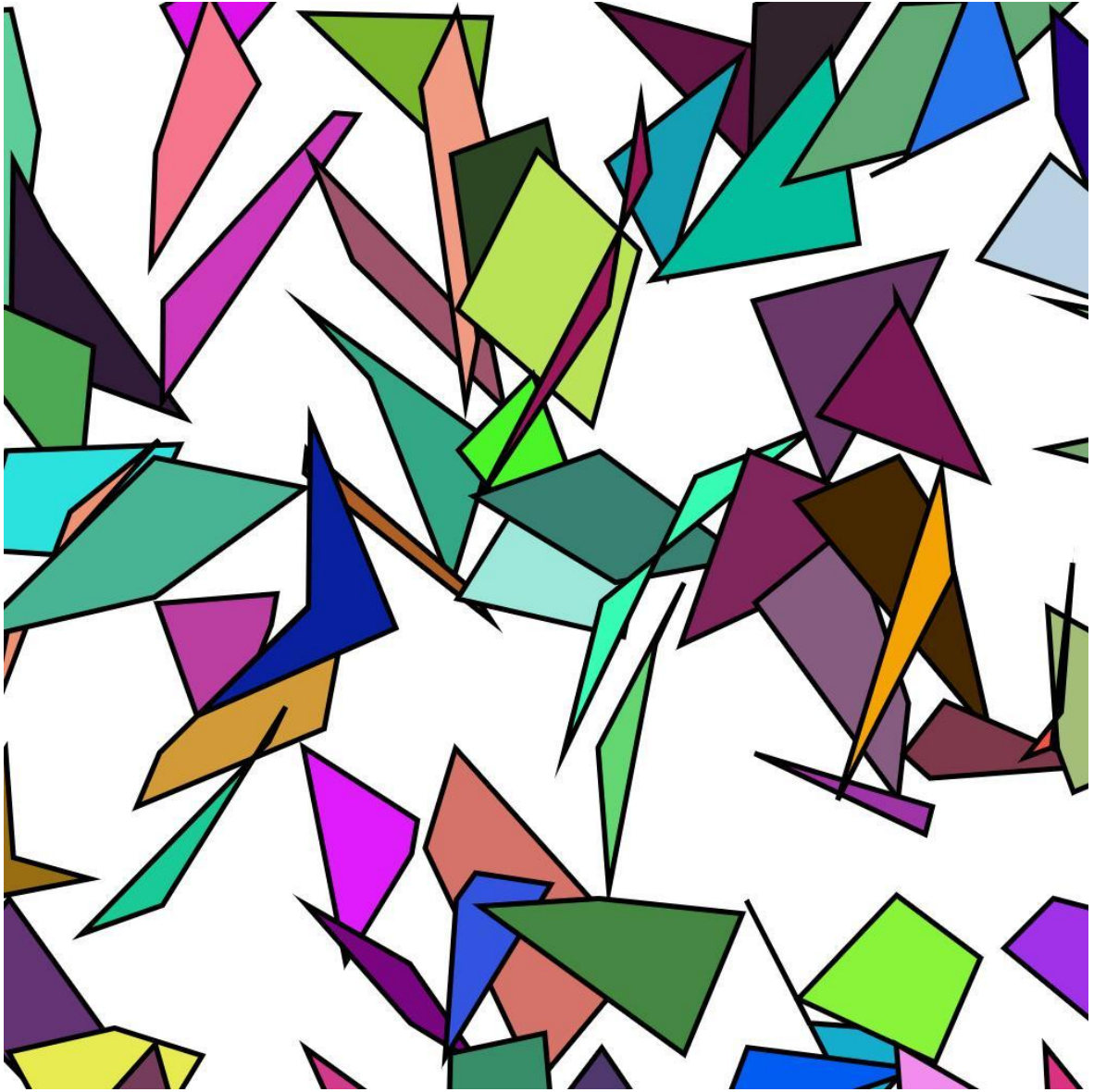
- Sueldos: <https://www.indeed.es/salaries/Informatico/a-Salaries?start=30>

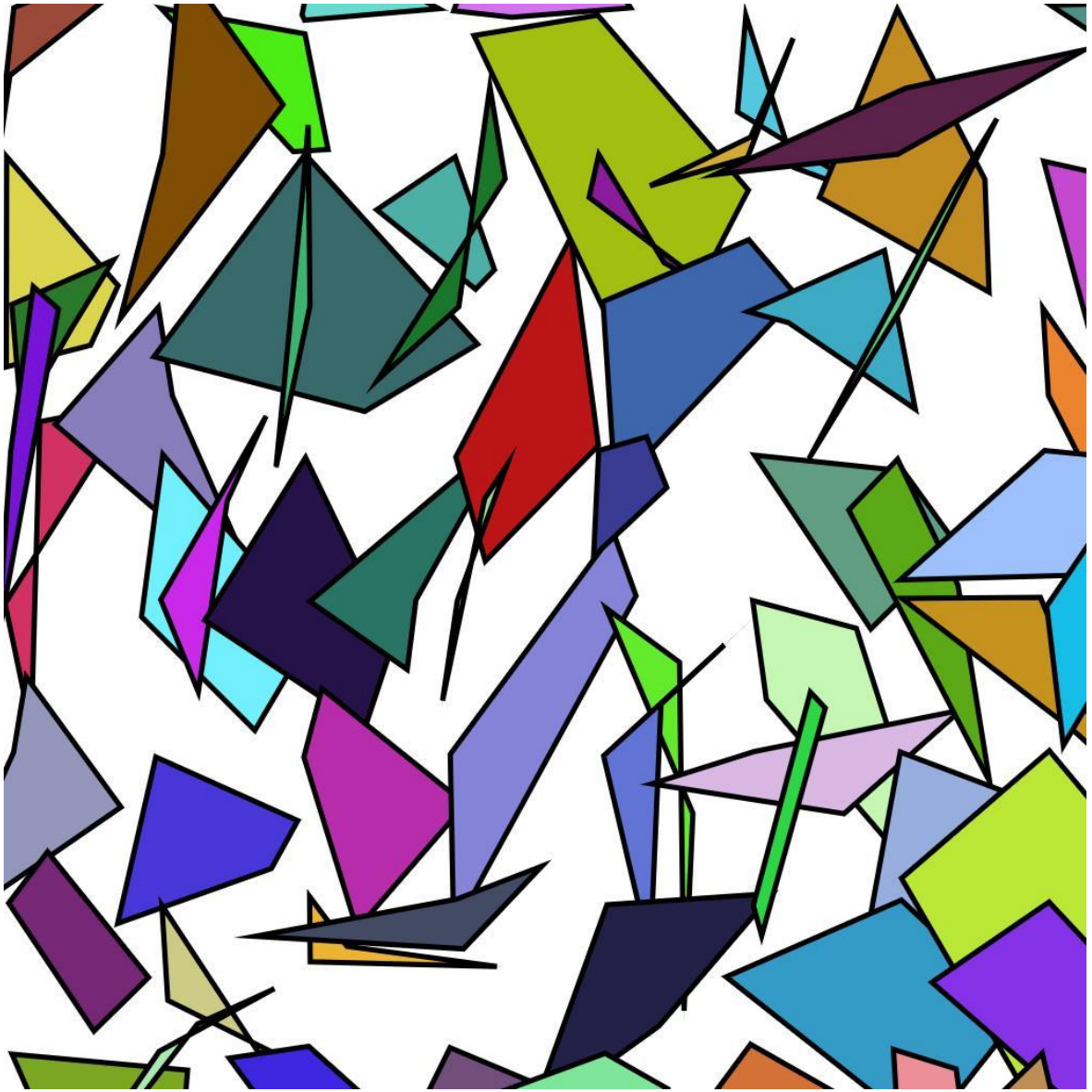


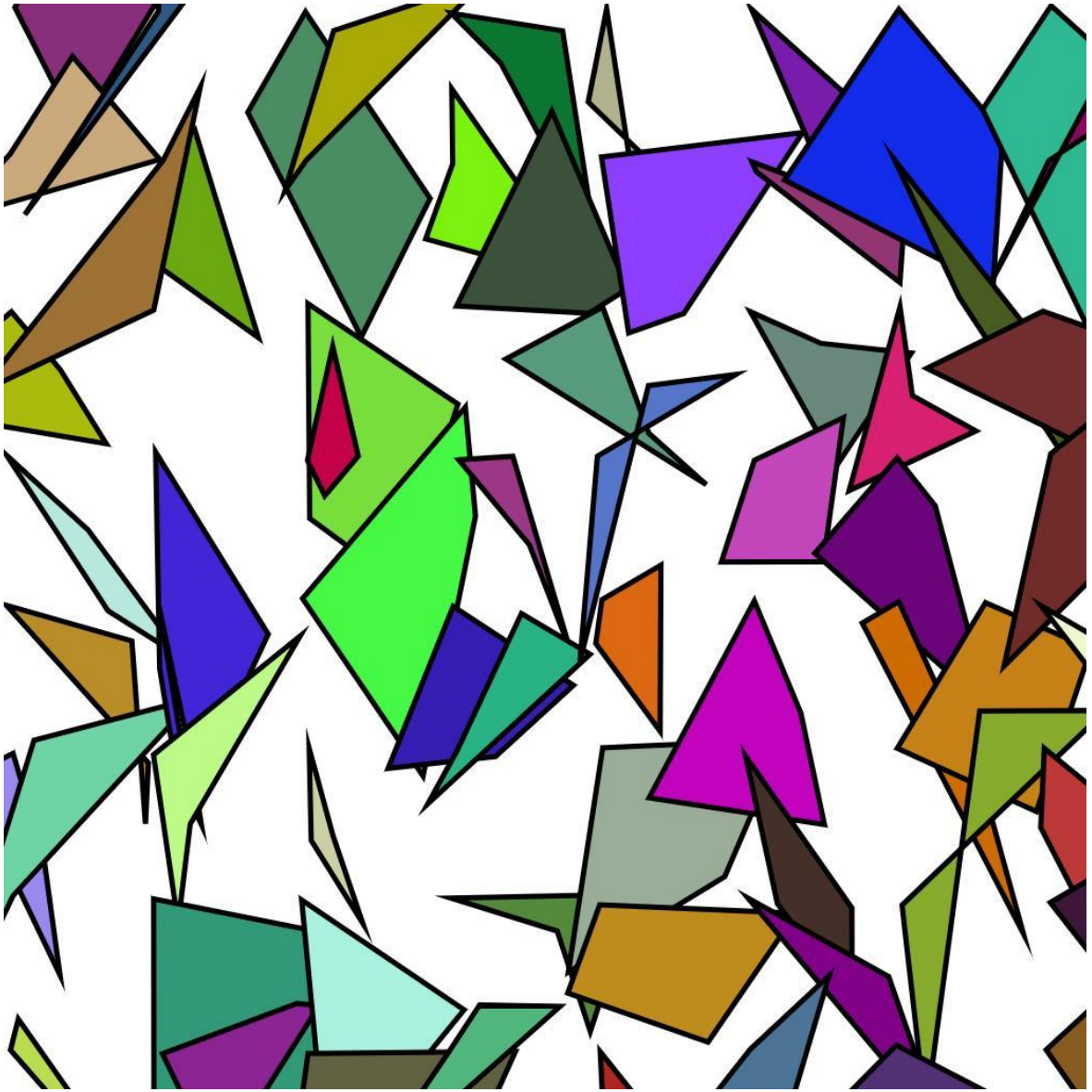
## **8 Anexo**

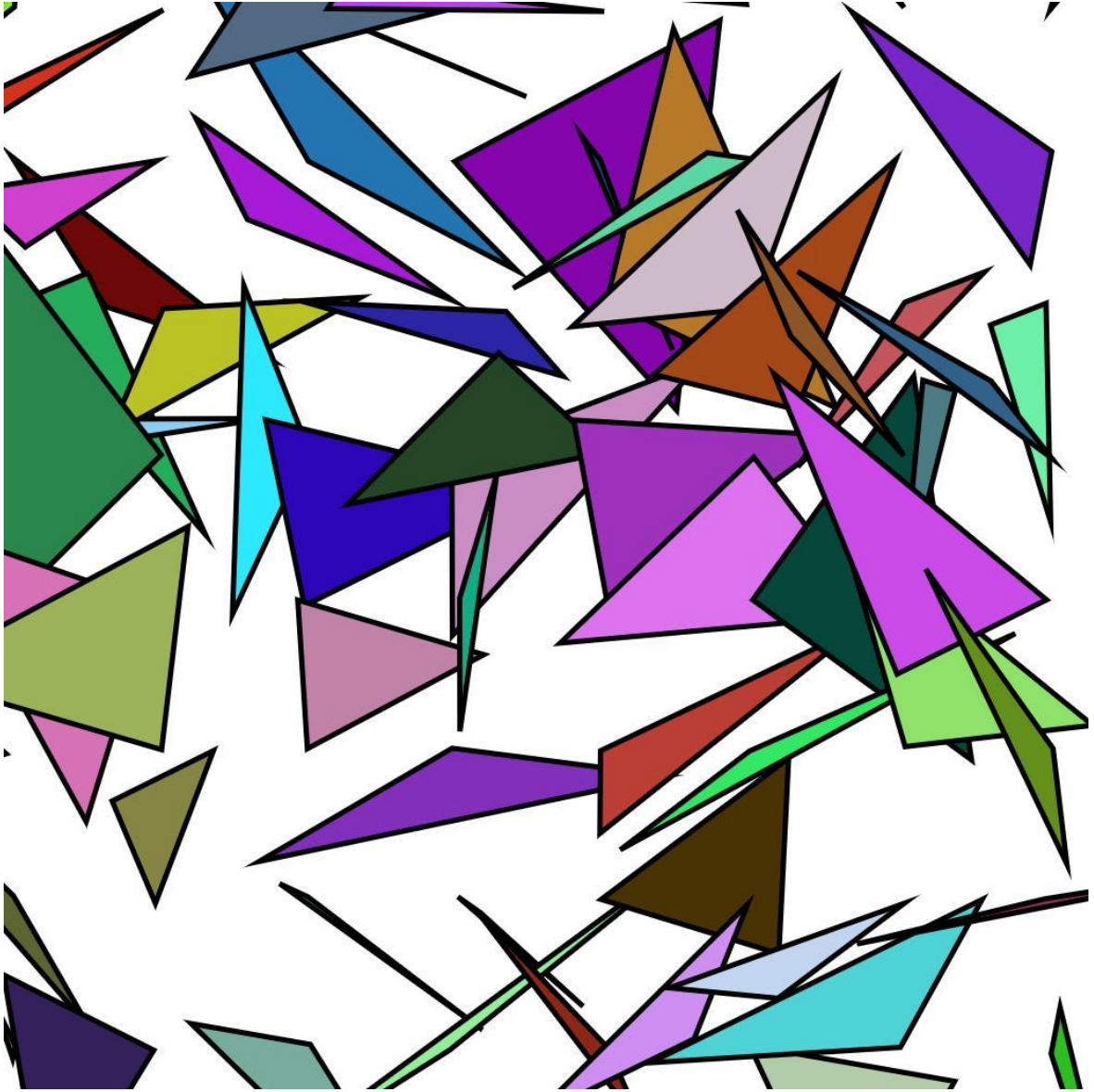
A continuación se mostrarán los targets generados aleatoriamente y empleados en la muestra de interfaces con Vuforia:

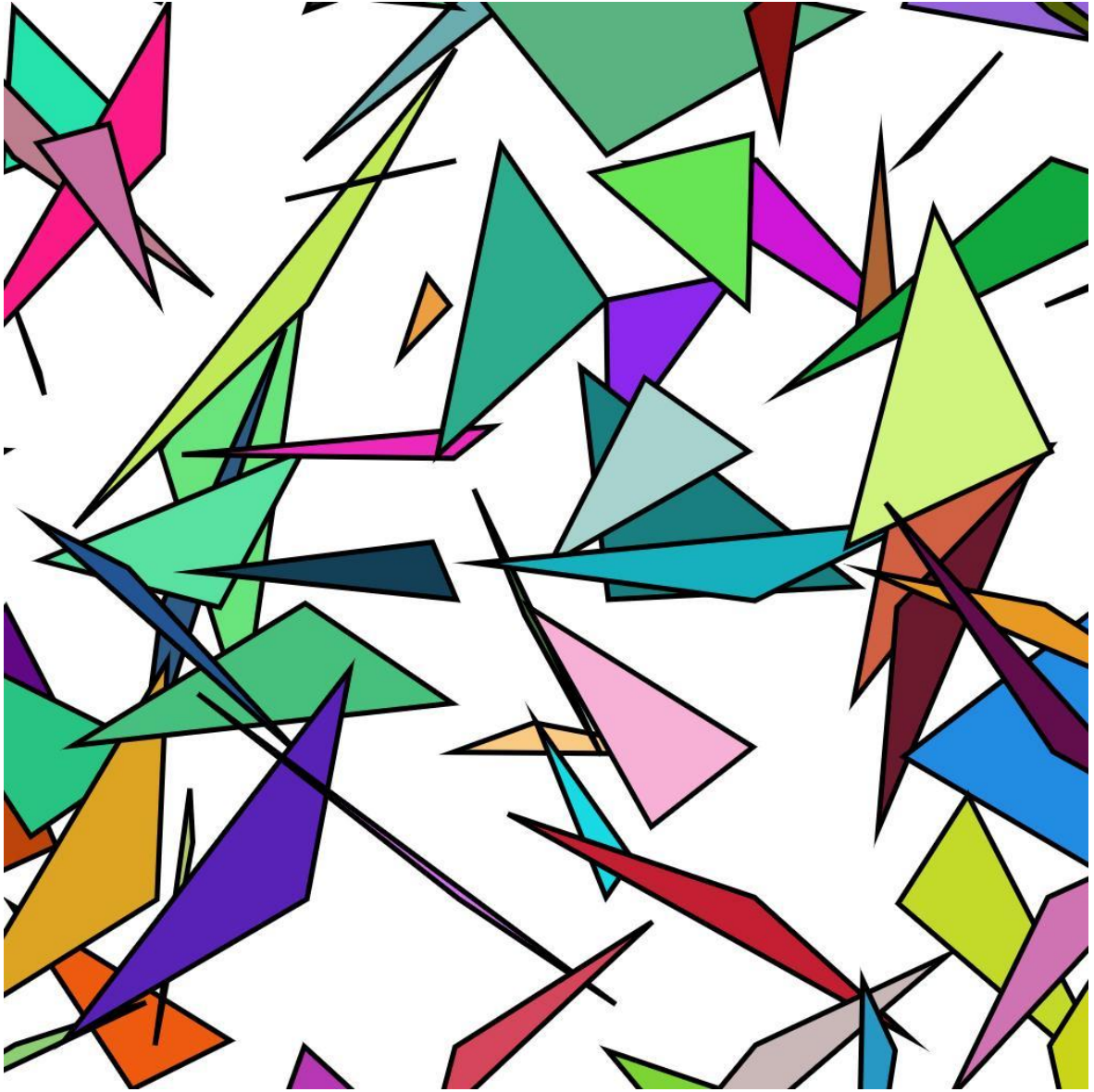


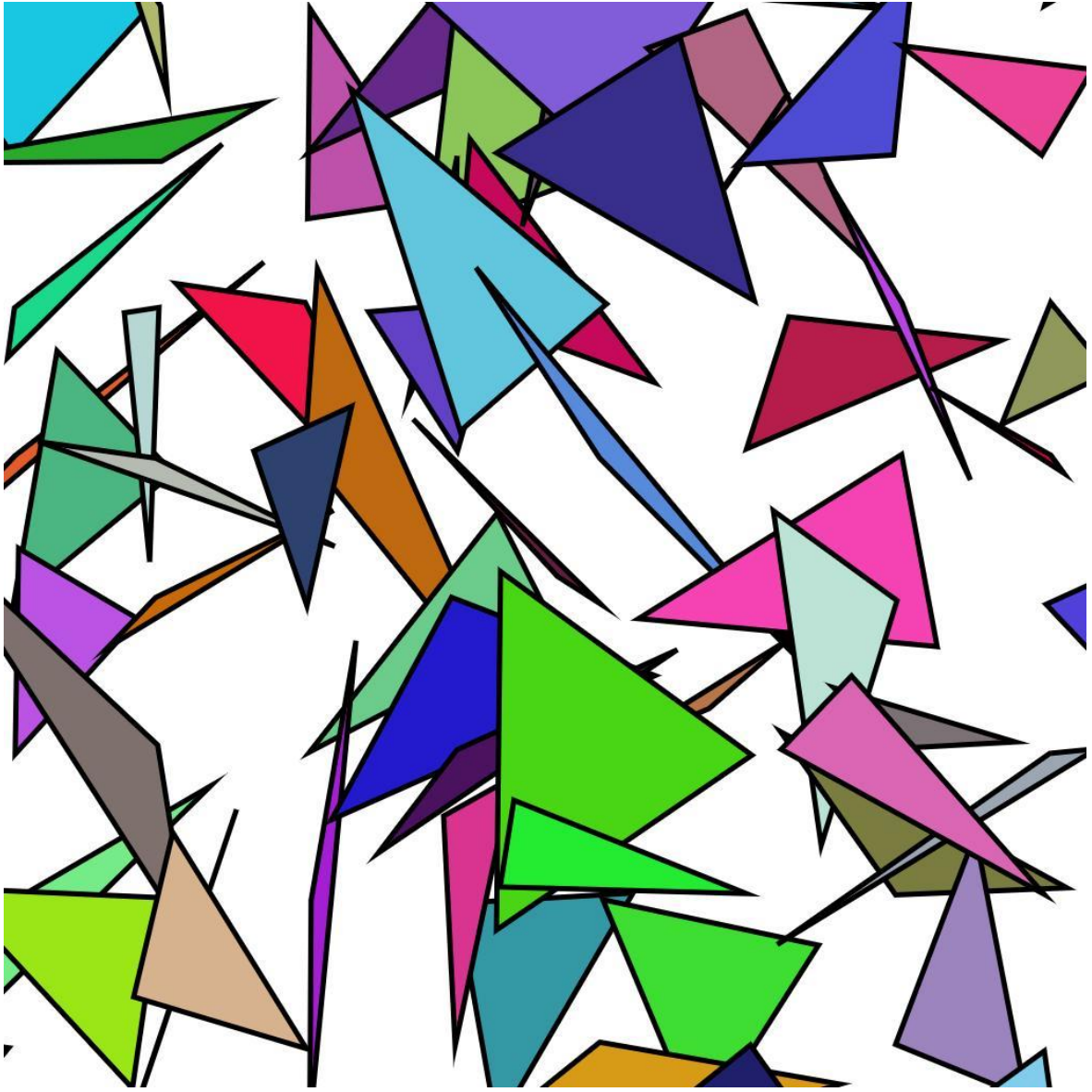










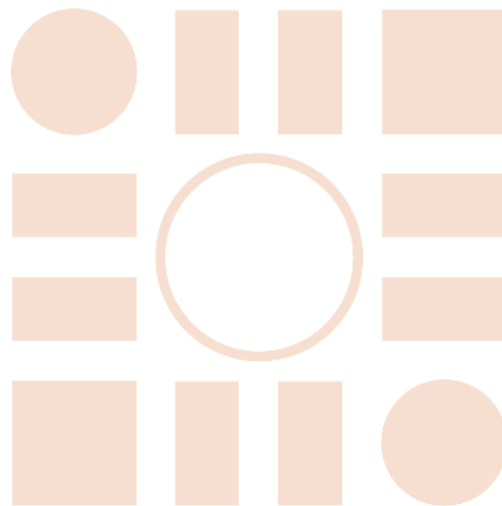








Universidad de Alcalá  
Escuela Politécnica Superior



ESCUELA POLITECNICA  
SUPERIOR



Universidad  
de Alcalá