

Tecnológico de Estudios Superiores de Ecatepec

Ingeniería en sistemas Computacionales

Proyecto: “Clasificador de Imágenes con Búsqueda inversa en tiempo real en el uso en medicamentos en el hogar.”

Equipo de investigación:

Escobar Cerezo Sergio

Orozco Gutiérrez Ricardo

Asignatura: Taller de investigación 2

Grupo: 5851

Profesora: Sonia Cantinca Téllez

Fecha: 17/04/18

índice

1. **REALIDAD AUMENTADA ....................................................................................................................................................4**
   1. TAREAS EN REALIDAD AUMENTADA .........................................................................................................................6
   2. CAPTACIÓN DE LA ESCENA ........................................................................................................................................8
   3. IDENTIFICACIÓN DE ESCENAS .................................................................................................................................10
      1. *Reconocimiento por marcadores ......................................................................................................................... 12*
      2. *Reconocimiento sin marcadores ......................................................................................................................... 14*
   4. TÉCNICAS DE MEZCLADO DE REALIDAD Y AUMENTO ..........................................................................................18
      1. *Fundamentos dimensionales ............................................................................................................................... 18*
      2. *Librerías de aumento ........................................................................................................................................... 18*
      3. *OpenGL ................................................................................................................................................................ 18*
      4. *OpenGL ES .......................................................................................................................................................... 19*
      5. VISUALIZACIÓN DE ESCENA ............................................................................................................................19
      6. *Sistemas de bajo coste ........................................................................................................................................ 19*
      7. *Sistemas de alto coste ....................................................................................................................................... 20*

1.6 RECONOCIMIENTO DE IMÁGENES............................................................................................................................21

*1.6.1. Adquisición de imágenes........................................................................................................................................23*

*1.6.2. Procesamiento digital..............................................................................................................................................23*

*1.6.3. Segmentación........................................................................................................................................................ 23*

*1.6.4. Representación y descripción................................................................................................................................26*

*1.6.5. Reconocimiento e interpretación............................................................................................................................ 28*

*1.6.6. Una alternativa a las técnicas anteriores: similitud de imágenes............................................................................ 30*

* 1. IDENTIFICACIÓN POR POSICIONAMIENTO ...........................................................................................................30

*1.7.1. Sistema de posicionamiento global (GPS)..............................................................................................................30*

*1.7.2. Uso de sensores de movimiento.............................................................................................................................. 31*

1.8 FORMATOS DE SEÑALES DE VÍDEO E IMÁGENES................................................................................................ 31

*1.8.1 Formato RGB........................................................................................................................................................... 31*

*1.8.2 Formato YUV........................................................................................................................................................... 31*

*1.8.3 Formato YCbCr 420.................................................................................................................................................32*

1.9 PLATAFORMAS MÓVILES........................................................................................................................................... 33

* + 1. *Symbian OS ............................................................................................................................................................ 33*
    2. *Windows Mobile ...................................................................................................................................................... 33*
    3. *iPhone OS ............................................................................................................................................................... 34*
    4. *Android ..................................................................................................................................................................... 34*

**2. ANÁLISIS DE SISTEMAS DE REALIDAD AUMENTADA.................................................................................................. 37**

* 1. COMPONENTES HARDWARE Y SOFTWARE .........................................................................................................37
     1. *Componentes Hardware .......................................................................................................................................... 37*
     2. *Componentes Software ........................................................................................................................................... 38*
  2. ANÁLISIS DE SISTEMAS DE REALIDAD AUMENTADA ........................................................................................... 38
     1. *ULTRA.................................................................................................................................................................... 40*
     2. *MARCH ...................................................................................................................................................................40*
     3. *ASTOR.....................................................................................................................................................................40*
     4. *POLAR.................................................................................................................................................................... 42*
     5. *Layar ...................................................................................................................................................................... 43*
     6. *LOCUS .................................................................................................................................................................. 44*
  3. ARQUITECTURAS DE SISTEMAS DE REALIDAD AUMENTADA ........................................................................... 45
     1. *Arquitecturas en sistemas autónomos de realidad aumentada ............................................................................... 46*
     2. *Arquitecturas en sistemas distribuidos de realidad aumentada............................................................................. 47*

1. **IMPLANTACION DE UN SISTEMA DE REALIDAD AUMENTADA ..............................................................................51**
   1. ARQUITECTURAS PARA PROVEEDORES DE REALIDAD AUMENTADA ............................................................. 51
   2. ARQUITECTURAS PARA CLIENTES DE REALIDAD AUMENTADA........................................................................ 52
   3. SISTEMAS AUTÓNOMOS FRENTE A SISTEMAS DISTRIBUIDOS ........................................................................ 52
      1. *Características de un sistema autónomo............................................................................................................... 54*
      2. *Características de un sistema distribuido de realidad aumentada......................................................................... 54*
      3. *Conclusiones ......................................................................................................................................................... 56*

### 3.4 CONFIGURACIÓN POR PARTE DEL PROVEEDOR Y DEL CLIENTE EN LOS SISTEMAS DESARROLADOS EN ESTE TRABAJO……………………………………………………………………………………………………………………...... 57

*3.4.1 Requisitos del proveedor....................................................................................................................................... 58*

* + 1. *Requisitos del cliente ............................................................................................................................................. 60*

1. **APLICACIONES AL E-LEARNING Y A LA DOMÓTICA ..............................................................................................63**
   1. MOTIVACIÓN DE LAS APLICACIONES..................................................................................................................... 63
   2. IDENTIFICACIÓN DE ESCENARIOS ........................................................................................................................ 64
      1. *Técnicas clásicas de visión artificial ..................................................................................................................... 65*
      2. *Uso de redes neuronales....................................................................................................................................... 66*
      3. *Reconocimiento inicial de marcadores .................................................................................................................. 70*
      4. *Búsqueda laberíntica ............................................................................................................................................. 74*
      5. *El problema de las rotaciones y transformaciones: la solución definitiva ................................................................ 77*
   3. TÉCNICAS DE AUMENTO.......................................................................................................................................... 82
   4. ARQUITECTURA DEL SISTEMA ............................................................................................................................... 85
      1. *Interfaz de usuario ................................................................................................................................................... 86*
      2. *Gestión de búsquedas ............................................................................................................................................. 88*
      3. *Gestión de comunicaciones .................................................................................................................................... 90*
      4. *Paquete de utilidades .............................................................................................................................................. 91*
   5. MUSEO DE INFORMÁTICA GARCÍA SANTESMASES ............................................................................................. 91
   6. APLICACIÓN DE DOMÓTICA .................................................................................................................................... 95
2. **CONCLUSIONES Y TRABAJO FUTURO..........................................................................................................................101**

**REFERENCIAS ..........................................................................................................................................................................105**

**ANALYSIS OF AUGMENTED REALITY SYSTEMS ................................................................................................................ 107**

# 1. REALIDAD AUMENTADA

Este capítulo define las tareas que se llevan a cabo en el proceso de aumento de la realidad, también incluye un estado del arte de las tecnologías involucradas en dichas tareas al igual que la identificación de los elementos, tipos de escenarios, reconocimiento por marcadores entre otros.

La sección 1 introduce el concepto de realidad aumentada y describe, a grandes rasgos, el funcionamiento de esta tecnología. Las siguientes cuatro secciones exponen una descripción más amplia de cada una de las tareas expuestas en la sección 1. Así, la sección 2 define las técnicas de captación de escenarios, mientras que en la sección 3 se exponen, a modo de resumen, las técnicas habitualmente empleadas en la identificación de la escena, explicando en qué consiste el reconocimiento mediante marcadores y sin ellos. Las siguientes dos secciones explican las técnicas de mezclado de realidad y aumento, sección 4, y de visualización final, sección 5. En la sección 6 se define en profundidad el proceso de reconocimiento de escenas mediante imágenes, exponiendo a modo de estado del arte las diferentes técnicas existentes en el campo de la visión artificial y su aplicación en los sistemas de realidad aumentada. La sección 7, por su parte, expone las técnicas de reconocimiento de escenas basadas en posicionamiento. En la sección 8 se introducen algunos conceptos básicos sobre las señales de vídeo e imágenes y sus formatos habituales. Por último, puesto que este proyecto pretende construir un sistema para móviles, la sección 9 expone las diferentes plataformas móviles existentes actualmente en el mercado.

## 1.1 Tareas en realidad aumentada

Los sistemas de realidad aumentada son herederos de una tecnología más antigua: la realidad virtual, por tanto, antes de dar una definición y una explicación de la tecnología, es necesario introducir al lector en las bases de lo que es la realidad virtual y sus elementos que lo comprenden, los cuales se retomaran a continuación.

La realidad virtual es un sistema que interactúa con el usuario simulando un entorno real en un entorno ficticio [SOM 02]. Este tipo de sistemas generan una sensación de que lo que está ocurriendo es real, aunque no lo sea. La realidad virtual es, básicamente, una técnica fotográfica de 360 grados, que permite movimientos y desplazamientos tanto verticales como horizontales, ofreciendo libertad absoluta de acción, lo que genera una amplia sensación de realismo en el usuario mediante una estimulación de los cinco sentidos.

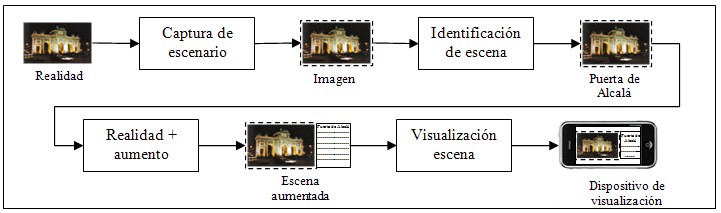
En términos generales, la realidad virtual es un sistema informático que simula un entorno ficticio ofreciendo al usuario la sensación de que es real, así como todo lo que ocurre en él. Una vez explicado de forma breve qué es la realidad virtual, se puede dar una definición más concisa y clara sobre qué es la realidad aumentada.

La realidad aumentada toma parte de la simulación de un entorno ficticio de la realidad virtual para añadirlo sobre un escenario real. De esta manera la escena que nuestros sentidos detectan en un escenario físico real, está enriquecida con información que no existe en el escenario físico original y que es proporcionada por un sistema informático.

Para facilitar la comprensión de la tecnología se expondrá un sencillo ejemplo que ilustre lo visto hasta el momento. Para ello imagine un monumento emblemático como puede ser la puerta de Alcalá en Madrid, España. Sería agradable que al pasar al lado de él se pudiese ver información de cualquier tipo sobre sí mismo como, por ejemplo, un rótulo que indicase que, efectivamente, es la puerta de Alcalá. Siendo más ambiciosos, quizá sería conveniente que apareciese información histórica sobre el monumento. Lamentablemente, a día de hoy es imposible que sobre los objetos físicos pueda aparecer información digital sin ayuda de proyectores 2D o proyectores holográficos. Es más, es posible que a mucha gente no le gustase observar los monumentos cargados de información digital. La solución a este problema la ofrece la realidad aumentada. En la actualidad cualquier persona con un teléfono móvil lo suficientemente avanzado puede capturar la escena con su cámara y visualizar el monumento con la información digital aumentada por la pantalla del teléfono. Incluso es posible que sin una cámara también pudiese ver la realidad mezclada sobre su dispositivo.

Como se ha visto en el ejemplo expuesto, en cualquier sistema de realidad aumentada se requieren dispositivos que identifiquen el escenario real y lo clasifiquen, así como que visualicen tanto entorno real como información digital, más adelante se irán exponiendo y ampliando estos conceptos.

Por otro lado, en todo sistema de realidad aumentada son necesarias, al menos, cuatro tareas fundamentales para poder llevar a cabo el proceso de aumento. Estas tareas se explicarán con mayor profundidad en secciones posteriores, pero básicamente son [LOP 10]: (i) captación de escena; (ii) identificación de escena; (iii) mezclado de realidad y aumento; y (iv) visualización. En la figura 2.1 se puede ver de forma esquemática este proceso.



**Figura 2.1.** Esquema de funcionamiento de un sistema de realidad aumentada típico.

A continuación, se describirán las cuatro tareas necesarias presentes en todo sistema de realidad aumentada. Después se describen técnicas fundamentales para llevar a cabo estas tareas y las principales plataformas móviles que las soportan.

## 1.2 Captación de la escena

Una de las tareas más importantes en cualquier sistema de realidad aumentada es la de identificar el escenario que se desea aumentar. En el caso de los sistemas que utilicen reconocimiento visual, es indispensable contar con algún mecanismo que permite recoger la escena para que pueda ser posteriormente procesada. En esta sección se analizan los diferentes tipos de dispositivos físicos que permiten captar dicho escenario.

Los dispositivos de captura de imágenes son dispositivos físicos que recogen la realidad que deberá ser ampliada [BIM 05]. A grandes rasgos, estos dispositivos se pueden agrupar, principalmente, en dos conjuntos: Dispositivos video-through y Dispositivos see-through

* Dispositivos *video-through*: dentro de este grupo se encuentran aquellos dispositivos que realizan la captura de imágenes o video que se encuentran aislados de los dispositivos de visualización. En este conjunto se encontrarían las cámaras de video o los terminales móviles (siempre y cuando tengan una cámara).
* Dispositivos *see-through*: son los dispositivos que realizan tanto la tarea de capturar la escena real como de mostrarla con información aumentada al usuario. Estos dispositivos acostumbran a trabajar en tiempo real, haciéndolos no sólo más costosos en presupuesto sino también en complejidad. Dentro de este grupo se encontrarían aquellos dispositivos conocidos como *head-mounted*. Cabe remarcar que estos dispositivos see-through llevan años siendo usados, por ejemplo, en los *Head Up Displays* (HUDs) utilizados por los aviones de combate para mostrar información al piloto sobre altura, velocidad, identificación de blancos, y otros sin necesidad de apartar la vista de la zona frontal de la cúpula de su cabina.

La figura 2.2 muestra un ejemplo de este tipo de dispositivos.



**Figura 2.2.** Prototipo de HMD.

## 1.3 Identificación de escenas

El proceso de identificación de escenas consiste en averiguar qué escenario físico real es el que el usuario quiere que se aumente con información digital [BIM 05]. Este proceso puede llevarse a cabo, básicamente, de dos maneras: utilizando marcadores o sin utilizarlos [LOP 10].

### 1.3.1. Reconocimiento por marcadores

En los sistemas de realidad aumentada, un marcador es un objeto cuya imagen es conocida por el sistema [CAW 08, CHO 09, TAT 07]. Las maneras en que el sistema conoce el marcador se pueden agrupar en tres conjuntos, mediante su geometría, su color o mediante ambas características, para llevar a cabo el procesamiento de cualquiera de los dos tipos de marcador se puede escoger cualquiera de los mecanismos que se presentan en la sección 2.6, sin embargo, se debe de tener en cuenta que, dichos mecanismos, suelen implicar una gran capacidad de cálculo y, por tanto, afecta al rendimiento del sistema.

Habitualmente para el reconocimiento de marcadores se utiliza un primer escaneo sobre la imagen más pesado computacionalmente para localizar el marcador que se busca, una vez localizado el mecanismo de actuación suele cumplir con ciertos pasos, que se describen a continuación:

En primer lugar, se establece un rango de variación en el movimiento del marcador para el posterior fotograma. En el procesamiento de dicho fotograma, el rango de búsqueda ya se encuentra acotado a un espacio muy inferior a la inicial, por lo que el tiempo de procesamiento decae considerablemente. Además, por norma general, se utilizan menos técnicas de reconocimiento, empleando el menor número de cálculos para localizar el marcador. Una vez detectado, se procede a las tareas necesarias de mezclado y aumento en los sistemas de realidad aumentada. Este proceso se efectúa de forma iterativa mientras la aplicación esté en ejecución.



**Figura 2.3.** Ejemplo de aplicación de marcador en realidad aumentada.

El proceso recientemente descrito sólo modificará su comportamiento si en algún fotograma en la región de búsqueda no se encontrase el marcador. En esta circunstancia existen diversas posibilidades de actuación:

* Realizar un nuevo escaneo sobre toda la imagen en busca del marcador. Este proceso puede ser efectivo si el marcador ha sido desplazado a una posición alejada de la anterior secuencia o si no se encuentra.
* Buscar de forma recursiva en las regiones vecinas el marcador. Esta solución puede ser óptima si el marcador desplazado se encuentra cerca de la región de búsqueda inicial.
* Utilizar predicción de movimiento. Esta tarea se puede llevar a cabo mediante la variación del movimiento analizando las imágenes o bien mediante el uso de acelerómetros. En este proyecto se utilizará la segunda opción.

Antes de concluir la explicación del reconocimiento por marcadores, es pertinente dar a conocer al lector qué el número de marcadores que puede reconocer este tipo de sistemas, no es ilimitado, ya que en su lugar, el límite de la cantidad de marcadores depende totalmente del algoritmo utilizado.

### 1.3.2. Reconocimiento sin marcadores

De la misma forma, es posible identificar la escena mediante reconocimiento de imágenes o mediante la estimación de la posición. También es posible encontrar sistemas que realicen una combinación de ambas en función de la situación. A este tipo de identificación se le denominará híbrida. Dentro de cada uno de estos dos conjuntos de técnicas se pueden encontrar diversas variaciones que dependerán en gran medida de las prestaciones que deba ofrecer el sistema, así como de sus posibilidades técnicas.

Puesto que las técnicas habituales en este proceso se encuentran centradas en el reconocimiento visual de la escena, se describirá en mayor profundidad ese tipo de técnicas. No obstante, esto no significa que no se puedan utilizar cualquier otro tipo de técnica más apropiada para cada situación, como puede ser el análisis en la intensidad de señales de radiofrecuencia o de señales infrarrojas.

Una vez comprendido lo pertinente sobre el reconocimiento sin el uso de marcadores, es importante denotar que el presente proyecto utiliza, un reconocimiento híbrido, combinando detección de marcadores con un motor de detección de movimiento mediante acelerómetros que permite conocer el desplazamiento ocurrido en cada instante para ayudar al sistema de reconocimiento a realizar una búsqueda más inteligente.

## 1.4 Técnicas de mezclado de realidad y aumento

Una vez descrito el proceso de identificación de escenarios, el siguiente proceso que tiene lugar en los sistemas de realidad aumentada es de sobreponer la información digital que se quiere ampliar sobre la escena real capturada. Cabe resaltar, que esta información digital de aumento puede ser tanto de carácter visual como auditivo o táctil, lo que, por lo general, en la mayoría de sistemas de realidad aumentada sólo se ofrecen los de tipo visual.

### 1.4.1 Fundamentos dimensionales

El primer concepto que hay que diferenciar en el proceso del dimensionado es el tipo de información que se quiere aumentar, ya que de eso dependerá el cómo es que será aumentado, de acuerdo a lo que sea más conveniente. En este apartado se puede distinguir entre dos tipos básicos de información: 2D y 3D.

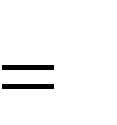
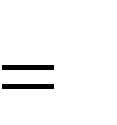
La información bidimensional constituye un plano geométrico digital en el que cada punto del plano se representa por medio de un píxel que, habitualmente, es un vector de colores en escala RGB (rojo-verde-azul). Cada uno de los tres valores que conforman el vector oscila en el rango 0-255, siendo la mezcla de los tres valores 0 la representación del color negro y 255 la del color blanco. Dentro de este tipo de información, no sólo se encuentran las imágenes, sino que también serán clasificados los contenidos textuales, debido a sus características similares de representación.

El otro tipo de información digital anteriormente expuesta es la tridimensional. Este tipo de información de la realidad consiste en un conjunto de vectores multidimensionales para cada punto del plano tridimensional (x, y, z). Al igual que sucedía con la información bidimensional, cada vector uno de estos puntos está formada por un vector RGB o cualquier otro tipo de representación del color.

Una vez definidos los tipos básicos de información digital de aumento, se puede realizar la descripción del funcionamiento real de las técnicas de aumento. En los sistemas de realidad aumentada, excepto en los sistemas que utilizan hologramas tridimensionales o similares, los dispositivos de visualización son en dos dimensiones, como pueden ser las pantallas de ordenadores, teléfonos móviles, etc. Este suceso puede llevar a pensar que sólo es posible representar información bidimensional y, aunque esto es cierto, es posible simular la sensación de tridimensionalidad en un plano 2D.

Para realizar la conversión de una imagen en 3D al plano bidimensional se suele utilizar la técnica de proyección de perspectiva (o proyección de puntos). Esta técnica consiste en simular la forma en que el ojo humano recibe la información visual por medio de la luz y cómo genera la sensación 3D. Este proceso consiste en la superposición de dos imágenes bidimensionales captadas desde diferentes ángulos, dando la sensación de profundidad inexistente en imágenes 2D. La ecuación 2.1 adjunta define la forma de trasladar el plano tridimensional al plano bidimensional.

*f* (*x*) *f* (*y*) *u* ,*v* *z z*



**Ecuación 2.1.** Proyección de plano de 3D (x, y, z) a un plano 2D (u, v).

### 1.4.2 Librerías de aumento

Para el proceso de aumento es necesario disponer de software adecuado para sobreponer a la imagen real la información aumentada deseada. Para este propósito existen diversas librerías disponibles al público. Las más famosas son:

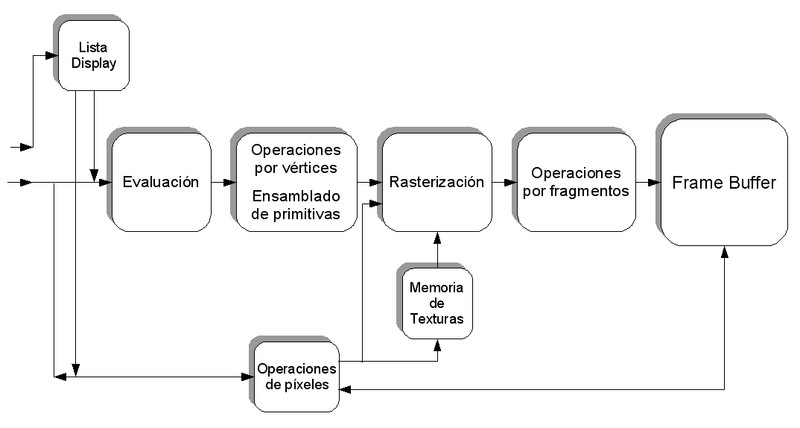
* *ARToolKit*: librería de realidad aumentada que permite la detección de unos marcadores específicos y realiza las tareas de superposición de imágenes [ART 10]. Está desarrollada en lenguaje C++.
* *ARToolKitPlus*: versión más especializada de la librería ARToolKit [ART P 10].
* *JARToolKit*: es la versión de la librería ARToolKit desarrollada en lenguaje Java [JART 10].

Las librerías que se acaban de presentar tienen la ventaja de formar en sí paquetes de realidad aumentada con todas las funcionalidades necesarias para sistemas por marcadores. No obstante, si el reconocimiento de escenarios se realiza por otro mecanismo al de detección de marcadores, estas librerías pierden su utilidad.

### 

### 1.4.3 OpenGL

OpenGL es una especificación estándar que define una API multilenguaje y multiplataforma para escribir aplicaciones que produzcan y consumas gráficos en 2D y 3D. Esta interfaz dispone de más de 250 funciones definidas, de tal forma que el funcionamiento básico consiste en aceptar primitivas tales como puntos, líneas y polígonos y convertirlas en píxeles. Este proceso se lleva a cabo en una pipeline gráfica conocida como la máquina de estados de OpenGL. La figura 2.4 muestra la pipeline que utiliza OpenGL.



**Figura 2.4.** Pipeline de OpenGL.

### 1.4.4 OpenGL ES

OpenGL ES es la variante simplificada de la especificación OpenGL para dispositivos móviles o terminales empotrados [GL ES 10]. Debido al éxito de la especificación en sistemas informáticos tradicionales y a la tendencia actual de trasladar a los dispositivos móviles la funcionalidad de los PCs básica, la especificación OpenGL ES ha tenido una altísima aceptación en los fabricantes de teléfonos móviles, siendo la utilizada por los líderes del mercado. Algunos ejemplos de plataformas que utilizan la tecnología OpenGL ES para manejo de gráficos 2D y 3D son:

* Symbian OS.
* Android.
* Maemo del Nokia N900.
* SDK del iPhone.
* PlayStation 3.

## 1.5 Visualización de escena

En esta sección se describirán los mecanismos de visualización habituales, para ello se realizará una clasificación de estos entre sistemas de bajo coste y sistemas de alto coste. Dentro de los sistemas de realidad aumentada, el último proceso que se lleva a cabo, y quizás uno de los más importantes, es el de visualización de la escena real con la información de aumento. Sin este proceso, la realidad aumentada no tendría razón de ser.

### 1.5.1 Sistemas de bajo coste

Este tipo de sistemas de visualización se caracterizan por estar presentes en la gran mayoría de dispositivos, tanto móviles como fijos y que cualquier usuario puede obtener de forma barata y sencilla. Dentro de este grupo se pueden diferenciar entre dos tipos básicos de sistemas: sistemas móviles y sistemas fijos.

Los sistemas de visualización móviles se refieren a los integrados en dispositivos móviles como teléfonos o PDAs. Estos sistemas se caracterizan por venir integrados en los dispositivos de fábrica, por lo que el usuario no debe preocuparse por adquirir elementos adicionales. En este tipo de sistemas se denominan dispositivos de visualización a las pantallas de dichos terminales móviles. Una de las características más relevantes en este tipo de sistemas es la baja definición y calidad de las imágenes de salida. Esto se debe a que los componentes hardware integrados en ellos encargados de tareas de carácter gráfico no disponen de gran potencia de cálculo ni de altas memorias integradas.

El otro tipo de sistema de visualización de bajo coste anteriormente expuesto son aquellos dispositivos fijos o que, a pesar de su movilidad, se pueden considerar como ordenadores personales. A diferencia de lo que sucede en los sistemas móviles, este tipo de sistemas suelen disponer de hardware adecuado para realizar tareas de visualización más complejas, generando de esta forma imágenes de salida de mayor calidad.

### 1.5.2 Sistemas de alto coste

Los sistemas de alto coste son escasos dentro de la realidad aumentada, aunque se pueden encontrar casos de utilización, como es el caso de los Head Up Displays (HUDs) [OKA 96]. Este tipo de sistemas tienen además la característica de ser interactivos con el usuario desde el punto de vista de que se libera a éste de dispositivos de visualización físicos, pudiendo ver la información aumentada mediante proyecciones sobre elementos físicos reales. Para poder hacer realidad este fenómeno se utilizan dispositivos de proyección en 2D o, sobre todo en los últimos años, en 3D. No obstante, también se pueden encontrar dentro de este grupo aquéllos dispositivos de última generación como los empleados por el ejército o en las simulaciones de vuelo que, debido a la criticidad de su servicio y de las altas prestaciones tanto a nivel hardware como software, conllevan un coste bastante elevado.

Por su parte, en el caso de los sistemas de visualización en 3D parece que se está generando un profundo interés en emplear esta tecnología. Las técnicas de representación holográfica en 3D han avanzado considerablemente en los últimos tiempos, siendo posible representar figuras humanas con una elevada calidad, poder interactuar de forma táctil con el holograma e incluso representar en 3D un objeto sin utilizar electricidad, simplemente mediante el uso de espejos. Evidentemente las empresas han visto una fuente de promoción e ingresos destacada en este sector, por lo que casos como el de Virgin dando una gala con un presentador virtual parecen ser solamente el principio de una nueva época en el sector.



**Figura 2.5.** Holograma interactivo que responde al tacto.

## 

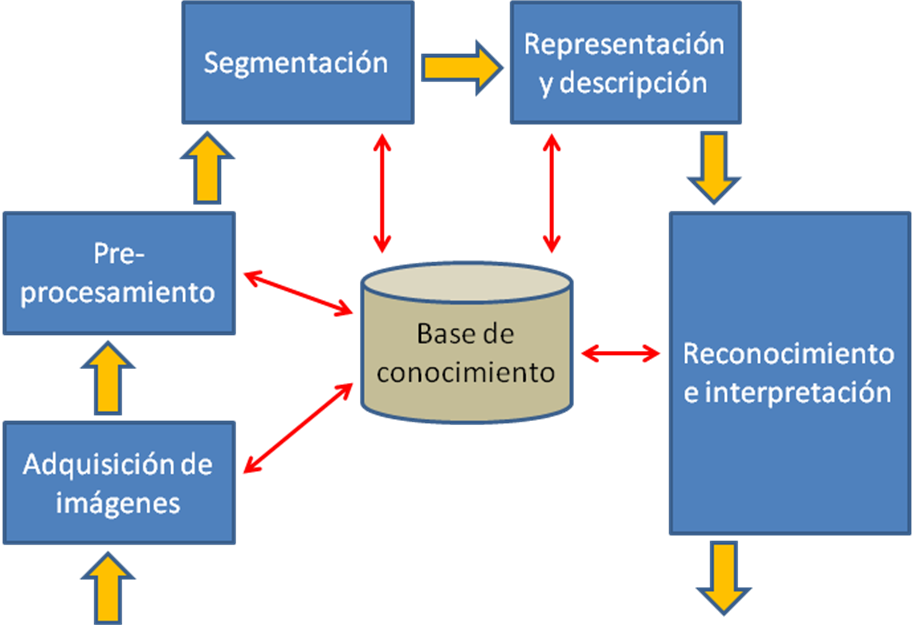
## 1.6 Reconocimiento de imágenes

El reconocimiento de imágenes consiste en extraer de las imágenes que recibe el dispositivo del usuario la información necesaria para identificar el escenario real que se quiere aumentar [JAV 02]. Dentro de este tipo de técnica de identificación, en el mundo de la realidad aumentada se pueden diferenciar claramente dos conjuntos de técnicas de reconocimiento de imágenes, que son el reconocimiento automático de escenarios mediante técnicas de visión artificial, y el reconocimiento por marcadores específicos definidos y conocidos por el sistema.

En esta sección se expondrán en primer lugar las técnicas de visión artificial empleadas para el reconocimiento de imágenes. La razón por la que no se identificarán cuáles de éstas participan en el reconocimiento automático es porque este tipo de sistemas vienen muy influenciados por el entorno sobre el que operan y sería inviable tratar en este trabajo cómo han implementado todos y cada uno de los sistemas de realidad aumentada estos mecanismos. No obstante, la exposición de las diferentes técnicas dará al lector una visión general de estas tecnologías que podrá ayudarle a elegir unas u otras para cada ocasión.

A grandes rasgos, el proceso de reconocimiento de imágenes consiste en los siguientes pasos [PLA 06, ESC 01, PAJ 01]. En primer lugar se adquiere la imagen mediante algún dispositivo preparado para llevar a cabo esta tarea, como puede ser una webcam. Una vez adquirida la imagen se realiza una etapa de procesamiento para eliminar imperfecciones de la imagen tales como ruido. Cuando se ha preprocesado la imagen se procede a su segmentación para buscar información característica en ella que pueda ser de utilidad a posteriores tareas. Tras la fase de segmentación se procede a la búsqueda de características morfológicas tales como perímetros o texturas. A esta etapa se le denomina representación y descripción. Por último, se procede al reconocimiento e interpretación de la escena mediante redes neuronales, lógica borrosa, etc.

Cabe destacar, que no necesariamente todos los sistemas deben implementar las fases anteriormente descritas, la decisión depende de las prestaciones que deba ofrecer y el entorno al que se vaya a enfrentar. La figura 2.6 muestra el proceso que se acaba de explicar.



**Figura 2.6.** Proceso de reconocimiento de imágenes mediante técnicas de Visión Artificial clásica.

A continuación, se explica en qué consiste cada fase y que técnicas se encuentran englobadas en ellas.

### 1.6.1 Adquisición de imágenes

Esta etapa consiste en construir el sistema de formación de imágenes. Con ello se pretende realzar las características visuales de los objetos, como formas o colores, mediante técnicas fotográficas, llamada técnica de preprocesamiento [PLA 06, ESC 01, PAJ 01].

Puesto que esta etapa ya se lleva a cabo de forma automática en los dispositivos de captura de imágenes, no se profundizará más en el tema, ya que sería necesario adentrarse en los fundamentos de la óptica, de la física de sólidos y líquidos, mecanismos de calibración, teoría de señales, etc., y no es el objetivo de este documento explicar estos conceptos que pueden encontrarse en [PLA 06, ESC 01, PAJ 01, JAV 02].

### 1.6.2 Procesamiento digital

En esta fase se trata de cuantificar y codificar la señal de vídeo o fotográfica recibida en forma de imagen, el objetivo es obtener una nueva imagen que o bien mejore su calidad o que destaque algún atributo significativo de ésta [PLA 06, ESC 01, PAJ 01], sin embargo los problemas de calidad pueden ser por falta o exceso de iluminación o por ruido. Al aplicar realce lo que se pretende conseguir es destacar bordes, regularizar colores, etc.

Las técnicas descritas en este capítulo pueden ser agrupadas en dos conjuntos: las que proceden de las señales y aquellas de carácter heurístico. En las técnicas para procesado procedentes de las señales se suelen aplicar los siguientes conceptos: Distancias entre píxeles y Procesos de convolución y operadores de correlación

* *Distancias entre píxeles*. Dentro de esta categoría se pueden encontrar las siguientes relaciones:

o *Relaciones de distancias*: establece la distancia entre píxeles. o *Relaciones de conectividad*: establece que dos píxeles adyacentes pertenecen a un mismo elemento.

* *Procesos de convolución y operadores de correlación*: se utilizan para aplicar filtros sobre las imágenes, por ejemplo, para eliminar ruido de sal y pimienta.

Por su parte, las técnicas de procesado de carácter heurístico se basan en un conjunto de procedimientos sobre el procesamiento digital de las señales y otros tipos de manipulaciones matemáticas. Este tipo de técnicas se pueden agrupar en tres conjuntos: realce o aumento del contraste, suavizado o eliminación del ruido y detección de bordes. Para llevar a cabo estas técnicas conviene hacerlo sobre imágenes en escala de grises, ya que son efectivas sobre la iluminancia de los objetos. A continuación, se describirán brevemente cada conjunto de las técnicas recientemente expuestas.

Las técnicas de realce consisten en aumentar el contraste de las imágenes, este tipo de procesado se basa en los conceptos de histograma, brillo y contraste [PLA 06, ESC 01, PAJ 01]. Por otro lado, las técnicas de suavizado lo que pretenden es eliminar el ruido que pueda tener la imagen. Existen tres tipos básicos de ruido: gaussiano como el que se muestra en la figura 2.7 impulsional y multiplicativo [PLA 06, ESC 01, PAJ 01].

Durante este proceso se utilizan filtros para eliminar ese ruido, los más comunes son [PLA 06]:

* Filtros paso bajo.
* Filtros gaussianos.
* Filtros basados en la mediana.
* Filtros homomórficos.



**Figura 2.7.** Imagen con ruido gaussiano.

La última técnica que se va a describir es la detección de bordes. Esta etapa suele preceder a las tareas de segmentación o a la búsqueda de objetos geométricos más complejos. Un borde se define como una región en la que aparece una fuerte variación de la intensidad en los píxeles adyacentes [PLA 06, ESC 01, PAJ 01]. Las técnicas más comunes sobre detección de bordes son las siguientes:

* Técnicas basadas en el operador gradiente, donde se encuentra, por ejemplo, el método de Sobel.
* Operadores basados en la laplaciana.
* Operador de Canny. En la figura 2.8 se ve el efecto de aplicar este tipo de detección.



**Figura 2.8.** Detección de bordes mediante el operador de Canny.

### 

### 1.6.3 Segmentación

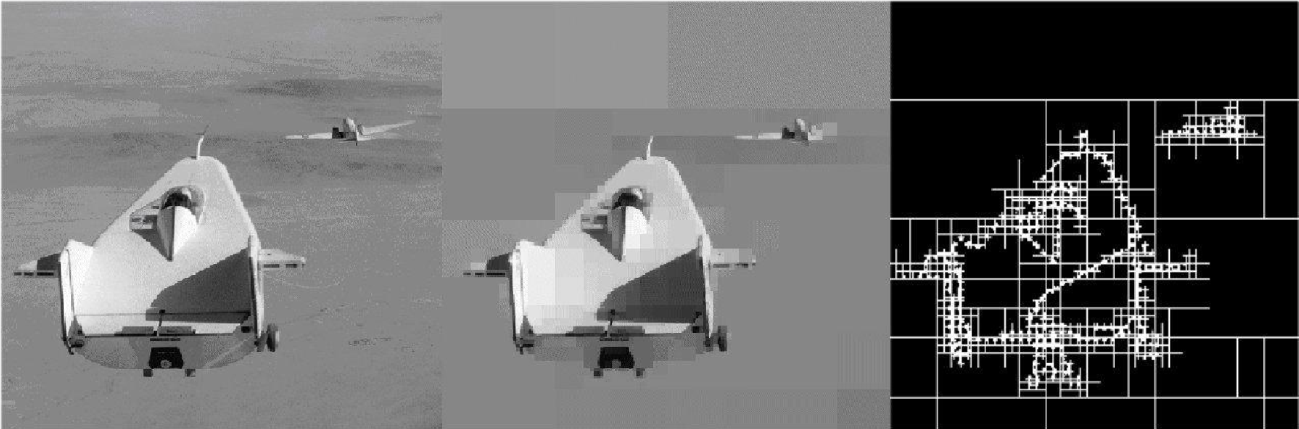
Esta fase tiene como objetivo agrupar los píxeles por algún criterio de homogeneidad para dividir la escena en regiones de interés. La agrupación de los píxeles se hace a razón de la similitud entre vecinos bajo criterios de luminancia, color, bordes, texturas, … Una vez realizada esta tarea la unidad más pequeña de trabajo dejará de ser el píxel para convertirse en agrupaciones de píxeles que constituyen objetos [PLA 06, ESC 01, PAJ 01].

Para este proceso se utilizan tres conceptos básicos:

* Similitud entre píxeles agrupados.
* Conectividad entre píxeles.
* Discontinuidad definida por los bordes de los objetos.

En la práctica imponer estas condiciones resulta casi imposible, por lo que se combinan las técnicas basadas en la localización de regiones uniformes con los contornos de los objetos [PLA 06, ESC 01, PAJ 01]. De tal forma que las técnicas utilizadas en esta fase se pueden agrupar en tres tipos [PLA 06, ESC 01, PAJ 01]: basadas en los contornos de los objetos, basadas en umbrales y en la búsqueda de regiones homogéneas.

El primer tipo de técnicas descritas serán las basadas en los contornos de los objetos [PLA 06]. Para llevar a cabo esta tarea se parte de una detección de bordes anterior que, debido a la presencia de ruido, sombras, etc., no es suficiente para poder segmentar las imágenes. Se trata de agrupar los píxeles etiquetados como bordes empleando la propiedad de conectividad. Habitualmente se utilizan transformadas de Hough para la detección de líneas rectas o círculos [BRA 01]. La siguiente técnica que se describirá es la umbralización [PLA 06, ESC 01, PAJ 01]. Esta técnica es frecuentemente utilizada en entornos industriales, cuando existe una clara diferenciación entre los objetos a extraer y el fondo de la escena. Los principios que rigen son la similitud entre píxeles pertenecientes a un objeto y sus diferencias respecto al resto. Para que se pueda llevar a cabo esta técnica la escena debe tener un fondo uniforme y objetos parecidos. Por último, la técnica basada en la búsqueda de regiones está basada en los criterios de similitud y continuidad de los píxeles que forman una región [PLA 06, PAJ 01]. El proceso suele consistir en una primera etapa de división de la imagen en árboles cuaternarios en los que los nodos son regiones de píxeles con alguna similitud entre ellos y una segunda etapa en la que se fusionan las regiones similares. La figura 2.9 muestra el resultado de aplicar segmentación sobre una imagen.



**Figura 2.9.** División de una imagen en árboles cuaternarios.

### 1.6.4 Representación y descripción

Durante el proceso de segmentación es habitual que el resultado de la delimitación de objetos no sea preciso debido a que se hayan agrupado de forma incorrecta los píxeles, se solapen regiones o los bordes sean imprecisos [PLA 06, ESC 01, PAJ 01]. Es por ello que para poder llevar a cabo un procesamiento de alto nivel es necesario realizar una última etapa de procesado con el objetivo de realzar la geometría y forma de los objetos de la escena. Este tipo de procesamiento se conoce como procesamiento morfológico [PLA 06].

Puesto que las técnicas empleadas en el proceso de visión artificial son muy dependientes del contexto sobre el que se opere, el procesamiento morfológico puede ser empleado en fases anteriores o incluso posteriores en el ciclo de vida del reconocimiento visual, como en restauración de imágenes, detección de bordes, análisis de texturas, aumento de contraste y compresión de imágenes.

Entrando en la descripción de este conjunto de técnicas, el procesamiento morfológico se basa en las nociones básicas de la teoría de conjuntos para extraer estructuras geométricas de ellos. Las técnicas más habituales de este proceso son las siguientes [PLA 06]:

* *Erosión binaria*. Consiste en comprobar si un conjunto (normalmente la representación de una forma geométrica) está completamente incluido en la imagen.
* *Dilatación binaria*. Es la transformación dual a la erosión.
* *Apertura y cierre*. Utilizando los dos operadores anteriores se consiguen diseñar operaciones de realce.
* *Gradiente morfológico*. Se utiliza en la localización de bordes y perímetros de los objetos.
* *Top-hat*. Sirve para descubrir aquellas estructuras de la imagen que han sido eliminadas en el filtrado de apertura y cierre. La figura 2.10 muestra el resultado al aplicar top-hat.



**Figura 2.10.** Top-hat en niveles de grises.

Otra de la forma habitual de realizar el procesado morfológico es mediante las conocidas transformaciones geodésicas [PLA 06]. Este tipo de transformaciones se basa en el concepto de conectividad [PLA 06]. Dentro de este conjunto de técnicas se puede encontrar:

* Dilatación y erosión.
* Reconstrucción.
* Watershead.

### 1.6.5 Reconocimiento e interpretación

Durante esta última fase el objetivo es interpretar automáticamente la escena partiendo de los resultados obtenidos en las anteriores etapas [PLA 06, PAJ 01]. Para llevar a cabo este proceso, se trata de asociar a cada elemento segmentado con un conjunto de valores numéricos o de atributos, denominados vectores de características. Estos vectores sirven de entrada a los sistemas de clasificación de objetos, que les asignarán una etiqueta cualitativa. Por tanto, este proceso, tal y como se describirá en esta sección, consta de las siguientes tareas: en primer lugar se le asocia a cada objeto una etiqueta; una vez etiquetada la imagen se procede a extraer las características del objeto para, por último, asignarle a cada uno una etiqueta cualitativa.

Como se ha explicado, la primera tarea es la de asociar a cada objeto segmentado un valor numérico. Esta etapa se fundamenta en la continuidad de los objetos en el espacio, dada por la relación de conectividad de los píxeles adyacentes. La salida de esta fase es una matriz en la que a los píxeles de cada objeto se les asigna un valor numérico.

La siguiente fase es la de extraer de los objetos características cuantificables para obtener el vector de características. Las propiedades de los objetos pueden ser clasificadas en relación con el contorno de los objetos y con las características propias del conjunto del objeto en su totalidad.

Para poder realizar la cuantificación se utilizan los siguientes descriptores [PLA 06]:

* Descriptores de regiones. - Descriptores topológicos.
* Texturas como se muestran en la figura 2.11.
* Descriptores de fronteras.



**Figura 2.11.** Imágenes con diferentes texturas.

La última tarea de la interpretación automática es el reconocimiento de patrones, para realizar la clasificación es necesario disponer del vector de características y de un amplio conocimiento del universo del problema, ya que para asignar una etiqueta de clasificación a un objeto es necesario disponer de reglas de pertenencia a cada clase. Existen fundamentalmente tres métodos de clasificación [PLA 06]:

* Métodos sintácticos.
* Redes neuronales. Un ejemplo del uso de este tipo de técnicas se puede ver en la figura 2.12.
* Métodos estadísticos.



**Figura 2.12.** Ejemplo de clasificación mediante una red neuronal.

### 1.6.6 Una alternativa a las técnicas anteriores: similitud de imágenes

Antes de concluir con la descripción del reconocimiento automático de imágenes es necesario describir una técnica bastante más ligera en términos computacionales que las que han sido expuestas hasta el momento.

Cabe mencionar que debido a que las aplicaciones que trabajan en tiempo real requieren de dispositivos con recursos de hardware nada limitados, es poco viable que dispositivos con hardware limitado tengan aplicaciones las cuales apliquen alguna de las técnicas que se han visto en este capítulo [CHO 09]. Una solución a este problema consiste en utilizar técnicas de similitud de imágenes. Para ello es necesario disponer de un número de muestras lo suficientemente alto del escenario sobre el que contrastar las imágenes que el dispositivo envía a un servidor dedicado a realizar este proceso si se trata de un dispositivo móvil con un hardware limitado, o bien desde el mismo terminal si sus prestaciones son las adecuadas.

En este proyecto, de hecho, se intentó utilizar este mecanismo, pero fue desechado debido a las limitaciones en cuanto al procesamiento de los dispositivos móviles, ya que la aplicación realiza todos sus cálculos en local. Este tipo de técnicas son útiles cuando se dispone de una infraestructura de red capaz de computar a altísima velocidad mediante clústeres la similitud de imágenes.

## 1.7 Identificación por posicionamiento

La identificación de escenarios mediante posicionamiento se realiza gracias a la estimación en la posición y orientación del usuario, deduciendo de esta forma qué es lo que está viendo [AML 08, DEY 00].

En la mayoría de aplicaciones de realidad aumentada para teléfonos móviles ésta suele ser la solución más extendida [WIKTD 10, LAY 10]. Esto se debe a que las operaciones de cálculo necesarias para realizar reconocimiento de imágenes suelen ser bastante costosas en cuanto a recursos y tiempo, haciendo inviables las aplicaciones en tiempo real. Sin embargo, utilizando técnicas de posicionamiento se consigue descargar a los terminales de grandes operaciones de cálculo, delegando éstos a los servidores [WIKTD 10, LAY 10].

En aplicaciones más sofisticadas, como el sistema POLAR [OLW 05], se utilizan dispositivos de captura del movimiento del usuario, así como cámaras que detectan el movimiento ocular de éste. De esta manera se consigue también estimar la escena que el usuario está visualizando.

De nuevo se aprecia cómo no existe una técnica general para todos los sistemas de realidad aumentada, sino que, cómo se verá en un capítulo posterior, los requisitos dependen estrictamente de las necesidades del sistema. A continuación, se presentan los distintos sistemas de posicionamiento que pueden ser utilizados en el proceso de identificación de escenas.

### 1.7.1 Sistema de posicionamiento global (GPS)

El sistema de posicionamiento global es un servicio de los EE.UU. que proporciona a los usuarios información sobre posicionamiento, navegación y cronometría [XU 07, GPS 10]. El sistema GPS está formado por tres segmentos, el espacial, el de control y el de usuario [XU 07, GPS 10]:

* El *segmento espacial* consiste en un grupo de veinticuatro satélites que envían información sobre la posición y hora de cada satélite.
* El *segmento de control* realiza las tareas de seguimiento y manejo de los satélites desde la tierra.
* El *segmento de usuario* se encuentra localizado en los terminales receptores GPS y determina la posición tridimensional y la hora precisa.



**Figura 2.13.** Satélite del segmento espacial del Sistema de Posicionamiento Global (GPS).

Para un dispositivo receptor GPS el funcionamiento es el siguiente. Cuando quiere determinar su posición, al menos debe recibir tres señales de tres de los veinticuatro satélites disponibles, que contendrán información sobre la posición del satélite y la hora a la que fue enviada la señal. Con estas señales el receptor sincroniza el reloj GPS y calcula el retraso de las señales, o lo que es lo mismo, la distancia al satélite. Mediante triangulación con las tres señales recibidas, se estima fácilmente la posición relativa respecto a los tres satélites. Gracias al envío de la posición absoluta de cada satélite, se calcula, una vez averiguada la posición relativa del receptor, su posición absoluta sobre el planeta.

Extendiendo la tecnología GPS a los sistemas de realidad aumentada, se puede apreciar claramente cómo, en entornos al aire libre y con ayuda de una simple brújula digital, esta técnica es una solución fiable para determinar la escena del usuario. Sin embargo, el uso de GPS en sistemas de realidad aumentada en entornos cerrados no resulta aplicable debido al error de estimación que el propio sistema trae consigo. El error medio de estimación es del rango de dos metros y medio, por lo que se hace evidente como, en entornos cerrados en los que es necesaria generalmente una precisión que ronde los centímetros, esta técnica no puede llevarse a cabo y se suela optar por técnicas de reconocimiento de imágenes ligeras.

### 1.7.2 Uso de sensores de movimiento

Otra de las técnicas utilizadas para la estimación de la posición del dispositivo se realiza mediante el uso de sensores de movimiento [OLW 05]. Por sí sola, este tipo de técnica no es capaz de dar una posición global del dispositivo, aunque en ciertas ocasiones no tenga porqué ser necesario, sino que la aplicación sólo necesite conocer su posición relativa desde el momento que inicia su ciclo de vida. En este tipo de contexto, sensores de movimiento como acelerómetros o giroscopios aportan información manejable para estimar la posición local del dispositivo en cada momento.

## 1.8 Formatos de señales de vídeo e imágenes

En esta sección se explicarán los formatos empleados habitualmente en los datos multimedia, tanto en imágenes como en vídeo. Existen múltiples tipos de espacios de colores, pero en este documento sólo se abordarán los formatos RGB y YUV, que son aquellos que han sido objeto de estudio para poder llevar a cabo el proyecto.

### 1.8.1 Formato RGB

El modelo RGB define un espacio de color formado por tres componentes cromáticas [RGB 98]. Estas componentes son los tres colores primarios, el rojo, el verde y el azul. Mediante la unión de estos tres colores es posible formar cualquier color reconocible por el ojo humano.

Cada una de las tres componentes tiene un valor comprendido entre 0 y 255, siendo el blanco la unión de las tres componentes cuando éstas tienen como valor 255 o, lo que es lo mismo, cuando alcanzan su mayor señal de luminosidad, creando una saturación de colores que se transforma en el blanco; y siendo el negro la ausencia de intensidad en la luminosidad de estas tres componentes, es decir, cuando éstas valen 0. Por otro lado, cualquier color de la escala de grises comprendido entre 0 y 255 se forma cuando las tres componentes tienen el mismo valor. Así pues, cuando tanto la intensidad del rojo, como del verde, como del azul valen 25, se corresponderían con el valor 25 de la escala de grises.

En el ámbito del color en los ordenadores, habitualmente se encuentra una variación del RGB tradicional, el ARGB. Este espacio de color define una cuarta componente alfa (A) que especifica la presencia o no del canal transparente sobre el espacio de color.

### 1.8.2 Formato YUV

El modelo YUV define un espacio de color en tres componentes, una de luminancia (valores en escala de grises) y dos de crominancia [YUV 95, YUV 02]. El modelo YUV imita el modelo de visión humana. Este formato fue originalmente pensado para habilitar compatibilidad entre los televisores en blanco y negro y en color, ya que la señal en escala de grises se transmite por una frecuencia diferente a los valores de crominancia que definen el color de la escena. El espacio de color de la familia de los YUV para las imágenes digitales es el espacio YCbCr, mientras que para las televisiones analógicas se utiliza el espacio YPbPr. Este documento se centrará en el espacio YCbCr que es el utilizado por las cámaras de vídeo de los teléfonos móviles de última generación.

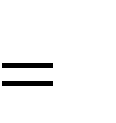
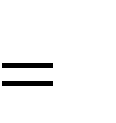
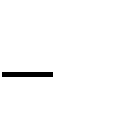
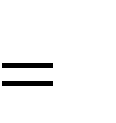
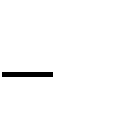
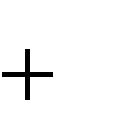
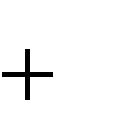
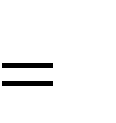
Las siglas YCbCr significan, cada una:

* Y es la componente lumínica. Es un valor entre 0 y 255 que define la intensidad de la luz en cada punto o píxel de la escena. En términos de crominancia, este valor indica el tipo de gris del punto.
* Cb y Cr son las componentes cromáticas. Cb corresponde con la componente de color U del espacio YUV y Cr con la componente V.

De acuerdo al estándar, la conversión entre RGB y YCbCr se encuentra definida en la ecuación 2.2.

*K*

1



*by*

*gy*

*ry*

*r*

*b*

*by*

*gy*

*ry*

*K*

*K*

*Y*

*R*

*C*

*Y*

*B*

*C*

*B*

*K*

*G*

*K*

*R*

*K*

*Y*

**Ecuación 2.2.** Conversión entre espacio RGB y espacio YCbCr

La tabla 2.1 muestra los valores de los coeficientes según los diferentes estándares.

|  |  |  |
| --- | --- | --- |
| **Estándar** | *Kry* | *Kby* |
| ITU601 / ITU-T 709 1250/50/2:1 | 0.299 | 0.114 |
| ITU709 / ITU-T 709 1250/60/2:1 | 0.2126 | 0.0722 |

**Tabla 2.1.** Coeficientes para calcular la conversión entre RGB e YCbCr.

Por último cabe resaltar que cada componente Cb y Cr definen la intensidad cromática de dos valores Y, consiguiendo de esta forma una compresión mayor en los datos.

### 1.8.3 Formato YCbCr 420

Este tipo de espacio de color es el utilizado por el sistema operativo Android para codificar la señal de vídeo recibida por las cámaras de vídeo. Pertenece a la familia de los YUV.

El formato de presentación de los datos de la señal viene codificado en un array de bytes en el cual, si la imagen tiene nxm píxels, los nxm primeros bytes contienen la componente Y de la escena, mientras que el resto de bytes contienen las componentes cromáticas Cb y Cr.

## 1.9 Plataformas móviles

En los últimos años se ha producido un notable crecimiento en el número de plataformas software pensadas en dar soporte a los dispositivos móviles, tales como teléfonos o PDAs, así como en las prestaciones que éstas aportan. Hoy en día, estas plataformas se aproximan más a los sistemas operativos tradicionales que a simples interfaces de acceso a los recursos de los dispositivos móviles. De hecho, la gran limitación que tienen estas plataformas se debe más al hardware al que dan soporte que a sus propias prestaciones.

### 1.9.1 Symbian OS

Symbian OS es un sistema operativo para terminales móviles [SYMB 10]. Este sistema da soporte a los teléfonos móviles de las empresas Nokia, Sony Ericsson y Motorola, entre otros. Adema es un sistema multi-tarea, que realiza protección de memoria. La aproximación a la multi-tarea está basada en un sistema de memoria virtual sobre eventos hacia un servidor asíncronos. De hecho, Symbian es un sistema operativo, en muchos aspectos, muy similar a cualquier otro que se pueda encontrar para ordenadores de escritorio.

El lenguaje de programación nativo de Symbian es C++, por lo que, aunque no sea estrictamente necesario el usarlo, la práctica habitual es programar aplicaciones en ese mismo lenguaje de programación por las prestaciones que ofrece. Este sistema operativo está compuesto de las siguientes capas, ordenadas desde más alto nivel a más bajo:

* UI Framework.
* Servicios de aplicación.
* Servicios OS.
* Servicios base.
* Servicios del Kernel e interfaz de hardware.

### 1.9.2 Windows Mobile

Windows Mobile es el sistema operativo para teléfonos móviles de Microsoft [MCPH 06]. Está basado en Windows CE 5.2, sistema optimizado para dispositivos con almacenamiento mínimo. Es un sistema en tiempo real que utiliza 256 niveles de prioridad. Su unidad fundamental es el hilo, lo que simplifica el interfaz de trabajo en tiempo de ejecución.

Al igual que ocurre con el sistema operativo Symbian, Windows Mobile está desarrollado en el lenguaje C++, por lo que, por norma general, el desarrollo de aplicaciones en esta plataforma se realizará también sobre ese mismo lenguaje. En la figura 2.14 se puede ver un escritorio típico de la plataforma Windows Mobile.



**Figura 2.14.** Captura de un dispositivo móvil que utiliza Windows Mobile.

### 1.9.3 iPhone OS

iPhone OS es el sistema operativo de los dispositivos móviles de Apple iPod Touch y iPhone [IPH DEV 10]. Está basado en una variante del Mach Kernel [SING 06] que se encuentra en Mac OS X, el sistema operativo de Apple para ordenadores de escritorio.

Para poder desarrollar aplicaciones para este tipo de dispositivos se ha de utilizar el iPhone SDK [IPH DEV 10], un kit de desarrollo específico para esta plataforma. Por otro lado es necesario conocer que el sistema consta de cuatro capas de abstracción, que son:

* La capa del núcleo del sistema operativo.
* La capa de Servicios principales.
* La capa de Medios de comunicación.
* La capa de Cocoa Touch.

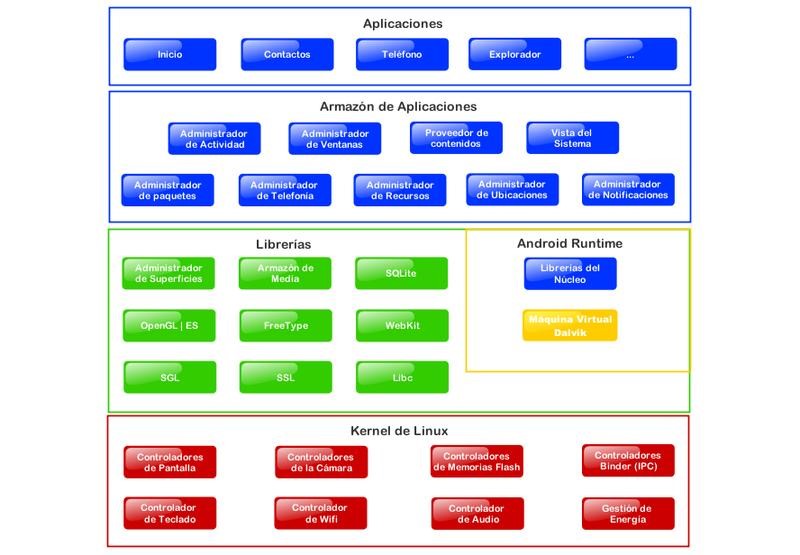
### 1.9.4 Android

Android es un sistema operativo basado en una modificación del Kernel de Linux para dispositivos móviles [AND SDK 10]. Aunque el desarrollo inicial fue llevado a cabo por la empresa Android Inc., actualmente la plataforma pertenece a Google.

Android permite el desarrollo de aplicaciones mediante una SDK en lenguaje Java y un kit nativo (NDK) para lenguaje C. Una de las características que hace de Android un entorno atractivo para los desarrolladores es que el código fuente se encuentra bajo licencias de software libre y código abierto.

La arquitectura de Android se puede ver en la figura 2.15 y está formada por los siguientes componentes:

* Aplicaciones.
* Framework de aplicaciones.
* Bibliotecas.
* Runtime de Android.
* Núcleo Linux.



**Figura 2.15.** Arquitectura de Andro

# 2. ANÁLISIS DE SISTEMAS DE REALIDAD AUMENTADA

Este capítulo analiza algunos de los sistemas de realidad aumentada más destacados, incluyendo los componentes hardware y software de los mismos, así como sus arquitecturas. Además en esta sección se definen los componentes hardware y software necesarios para la implantación de un sistema de realidad aumentada. Después se analizan algunos de los principales sistemas de realidad aumentada identificando los componentes hardware y software de los mismos, así como sus arquitecturas.

La primera sección identifica los componentes hardware y software que habitualmente son necesarios para implantar un sistema de realidad aumentada. En la segunda sección se traslada la identificación de la primera sección a un análisis de diferentes sistemas de realidad aumentada ya desarrollados, obteniendo información sobre el hardware y software empleado. La tercera y última sección define las diferencias entre un sistema autónomo y uno distribuido desde el punto de vista arquitectónico de los sistemas y aplicaciones de realidad aumentada. Las definiciones aportadas se acompañan de un análisis de qué arquitecturas han implantado los sistemas expuestos en la sección 2.

## 2.1 Componentes hardware y software

Esta sección identifica los componentes hardware y software necesarios para la implantación de un sistema de realidad aumentada. En este apartado, estos componentes sólo se identifican de manera genérica (p.e. librería gráfica).

### 2.1.1 Componentes Hardware

Como cualquier otro sistema informático, los sistemas de realidad aumentada requieren de componentes hardware sobre los que se vayan a llevar a cabo las acciones deseadas.

Los elementos básicos necesarios para implantar un sistema de realidad aumentada son similares a los de cualquier sistema informático tradicional. Así pues, son necesarios componentes de procesamiento, almacenamiento, comunicaciones, etc. De hecho, la diferencia en el equipamiento básico se encuentra en la finalidad del sistema.

En consecuencia, los elementos hardware necesarios para cualquier sistema de realidad aumentada son: una cámara, una pantalla de visualización o equipos holográficos en 2D o 3D, un equipo de procesamiento y cálculo, componentes de almacenamiento y, por último, dispositivos de transmisión y recepción de datos en caso de que sean necesarios. No hay que olvidar que las prestaciones de cada uno de ellos dependen fundamentalmente del tipo de sistema que se vaya a implantar. La tabla 1 [LOP 10] recoge un resumen de los componentes hardware habituales en un sistema de realidad aumentada.

**Tabla 3.1.** Resumen de los componentes hardware de un sistema de realidad aumentada**.**

**Componente hardware**

**Técnica**

|  |  |
| --- | --- |
| Captura *video-through* | Todo tipo de cámaras de vídeo |
| Captura *see-through* | Cámaras integradas de vídeo, HMD (*Head Mounted Display*) |
| Reconocimiento geoposicionamiento | Antena GPS, sensores de movimiento |
| Reconocimiento híbrido | Antena GPS, sensores de movimiento, |
|  | dispositivos de captura de imágenes | |
| Tratamiento imágenes | Procesadores con velocidad de procesamiento de, al menos, 2 GHz | |
| Almacenamiento | Disco duro convencional | |
| Comunicaciones locales | Tarjeta de red, conectores RJ45, antena wireless, punto de acceso | |
| Comunicaciones móviles | Equipamiento GSM | |
| Visualización *video-through* | Pantallas de vídeo, monitores, proyectores | |
| Visualización *see-through* | HMD, teléfonos móviles | |

### 2.1.2 Componentes Software

Una vez descritos los componentes hardware se expondrá el software necesario para implantar un sistema de realidad aumentada. Al igual que sucede a la hora de elegir el hardware adecuado, la decisión de utilizar un software u otro viene determinado por las prestaciones y características del sistema.

En términos generales, un sistema de realidad aumentada necesita software de reconocimiento, ya sea por posicionamiento o por imágenes (con o sin marcadores), librerías de procesamiento de imágenes (si es que lo necesitase), al menos una base de datos, librerías de aumento de contenidos digitales y, si fuese necesario, software de comunicaciones. La tabla 2 [LOP 10] muestra un resumen del software genérico necesario para cada etapa.

Cabe destacar que la mayoría de librerías de procesamiento de imágenes, ARToolkit [ART 10], Java3D [J3D 10], implementan tanto funciones de reconocimiento como de procesado y manipulación de contenidos visuales.

**Tabla 3.2. Resumen de los requisitos software de un sistema de realidad aumentada. Componente software**

**Técnica**

|  |  |
| --- | --- |
| Captura escena | Controladores de cámara de vídeo |
| Reconocimiento visual | Librerías de reconocimiento de imágenes |
| Reconocimiento geoposicionamiento | Brújula digital, software GPS |
| Reconocimiento híbrido | Brújula digital, librerías de reconocimiento de imágenes, software GPS |
| Tratamiento imágenes | Librerías de tratamiento de imágenes |
| Almacenamiento | Base de datos |
| Comunicaciones locales | Controlador de red |
| Comunicaciones móviles | Controlador GSM |
| Visualización contenidos | Software de reproducción de contenidos multimedia, librerías de tratamiento de imágenes |

## 2.2 Análisis de sistemas de realidad aumentada

En esta sección se expondrán una serie de proyectos de realidad aumentada que ya han sido o están siendo llevados a cabo con el fin de exponer la diferencia de arquitecturas y componentes empleados según cada situación.

### 2.2.1 ULTRA

El primer proyecto que se va a analizar es el sistema ULTRA [ULT 10]. El objetivo de este proyecto fue crear una plataforma de realidad aumentada ultra ligera y compacta que permitiese a los desarrolladores crear sus propios sistemas. La idea fue originalmente creada para que fuese utilizada en dispositivos móviles, en concreto para PDAs.

El reconocimiento del escenario se realiza mediante marcadores. Esto es posible debido a que la plataforma fue ideada para dar soporte a sistemas en entornos industriales y, por tanto, cerrados. Por ello, los requisitos aliviaron la carga del trabajo del sistema, ya que el reconocimiento por marcadores es un proceso más ligero. La plataforma fue pensada para que todo el proceso de realidad aumentada tuviera lugar en el terminal. Debido a esto, el hardware debe poseer una serie de propiedades predeterminadas. En primer lugar, es necesario que la PDA esté conectada a una cámara de 800x600 píxeles. En segundo lugar, debe proveer de aceleración de video y gráficos 3D para poder realizar las operaciones necesarias. La librería gráfica que se escogió fue OpenGL Es, por lo que fue necesario utilizar una implementación específica de EGL como interfaz entre OpenGL ES y la plataforma nativa del sistema de ventanas de la PDA.

Por otro lado, ULTRA permite la creación de contenidos off-line desde un servidor remoto. Por ello, se requieren dos requisitos hardware más. Primero, el dispositivo móvil debe poder establecer comunicaciones W-LAN para comunicarse con el servidor. Segundo, es necesario disponer de un equipo que actúe como servidor en este caso.

Por último, para poder llevar a cabo la función de creación de contenidos es necesario disponer de un experto que se encargue de realizar las imágenes 2D y 3D, texto o contenidos multimedia que se vayan actualizando en el sistema. El motivo de esta citación es que hasta ahora no se había comentado en ningún momento nada sobre los requisitos de recursos humanos necesarios en un proyecto de realidad aumentada.

**Tabla 3.3.** Resumen de las técnicas empleadas en ULTRA**.**

**Técnica**

**Componente**

**software**

**Componente**

**hardware**

Reconocimiento marcadores OpenGL ES Cámara de teléfono móvil o PDA

Tratamiento imágenes OpenGL ES Procesador y memoria

del teléfono móvil o

PDA

Visualización Delegado en software Pantalla del teléfono del dispositivo móvil o PDA

### 2.2.2 MARCH

MARCH es un proyecto que permite la visualización de contenidos digitales sobre grabados de cuevas prehistóricas [CHO 09]. Este proyecto fue creado para que funcionase en terminales móviles. En la figura 3.2 se puede ver una representación del funcionamiento del sistema.

El primer requisito que aparece es la del reconocimiento de escenas mediante marcadores. Las razones del uso de este mecanismo son dos: poder utilizar el sistema en cuevas reales y aligerar la carga de procesamiento y cálculo del terminal.



**Figura 3.1.** Funcionamiento del sistema MARCH.

El sistema consta de dos partes, una off-line y la aplicación en el móvil. La parte off-line se encarga de registrar las imágenes que serán aumentadas en el terminal. Este proceso consiste en añadir sobre la imagen del grabado la interpretación del experto. Una vez realizado este proceso se procede a extraer en una capa de fondo transparente la interpretación del experto. Un requisito por tanto de este sistema es contar con un experto sobre interpretaciones prehistóricas.

Por último, existe un requisito que restringe enormemente las posibilidades de esta aplicación. El sistema utiliza la librería gráfica NokiaCV [NOK CV 10], por lo que sólo los terminales Nokia que implementen esta API pueden soportar la ejecución de esta aplicación.

**Tabla 3.4.** Resumen de las técnicas empleadas en MARCH**.**

**Técnica**

**Componente**

**software**

**Componente**

**ha**

**rdware**

Reconocimiento marcadores Nokia CV Cámara de teléfono

móvil o PDA

Tratamiento imágenes Nokia CV Procesador y memoria

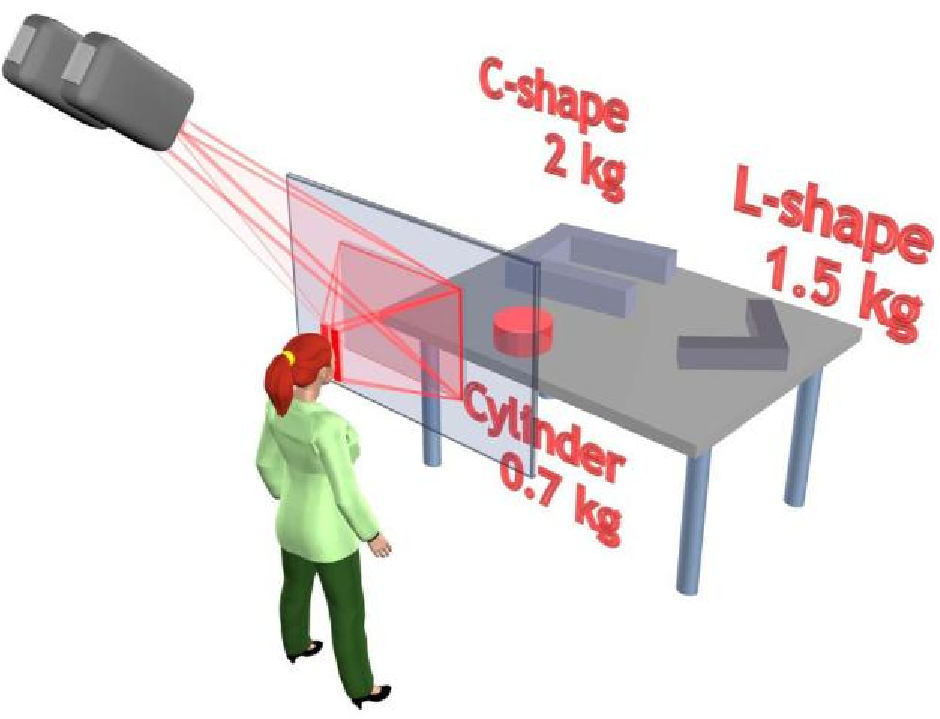
del teléfono móvil o

PDA

Visualización Delegado en software Pantalla del teléfono del dispositivo móvil o PDA

### 2.2.3 ASTOR

ASTOR es un prototipo de sistema de realidad aumentada en el que la información digital se muestra en unas pantallas transparentes sin que el usuario tenga que utilizar ningún dispositivo [OLW 05]. El sistema utiliza representación holográfica para mostrar la información aumentada al usuario. Para ello se sirve de dos o más proyectores que reflejan la imagen sobre elementos holográficos ópticos (HOE) transparentes tal y como se aprecia en la figura 3.



**Figura 3.2.** Representación del sistema de visualización del sistema ASTOR.

En primer lugar, el sistema ASTOR necesita unos HOE transparentes sobre los que representar la imagen holográfica. Los HOE son unos platos de cristal fino o incluso unas láminas de plástico transparente que pueden ser acopladas sobre otras superficies, como puede ser una ventana o un expositor.

En segundo lugar, ASTOR está pensado para trabajar en entornos pequeños, aunque puede operar en entornos más grandes, aumentando considerablemente los costes.

En tercer lugar, puesto que el usuario no requiere de ningún dispositivo, todo el hardware se centra en el lado del proveedor. Al ser todavía un prototipo los gráficos no son excesivamente complejos, por lo que no se requiere de un dispositivo especialmente potente. El prototipo ha sido desarrollado utilizando la tarjeta gráfica NVIDIA GeForce4 MX 440-SE y 256 MB de RAM. También necesitaron de proyectores con control remoto para poder guiarlos hacia la escena correctamente.

En el apartado software emplearon la librería gráfica Java3D [J3D 10] y JRMI [JRMI 10] para las comunicaciones cliente-servidor.

**Tabla 3.5.** Resumen de las técnicas empleadas en ASTOR**.**

**Técnica**

**Componente**

**software**

**Componente**

**hardware**

Reconocimiento imágenes Java3D Cámara de vídeo

|  |  |  |
| --- | --- | --- |
| Tratamiento imágenes | Java3D | NVIDIA GeForce MX  440-SE, 256 MB de  RAM |
| Comunicaciones | JRMI | Arquitectura clienteservidor |
| Visualización | Delegado en software del dispositivo | Proyectores, pantallas  HOE |

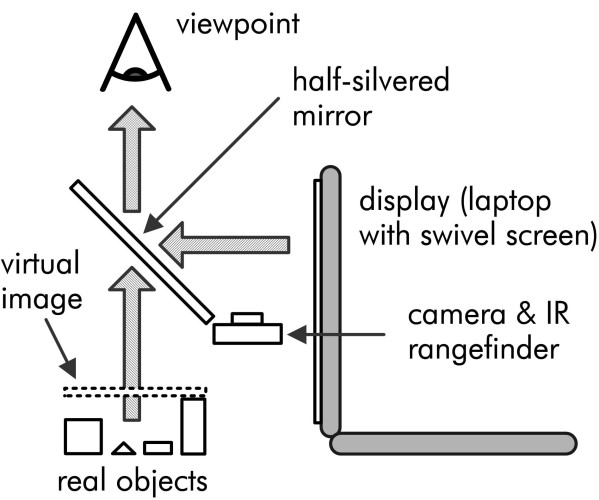
### 2.2.4 POLAR

POLAR es un sistema de realidad aumentada de bajo coste que permite superponer información a una escena real mediante un mecanismo diferente a los expuestos en este documento [OW 05].

El sistema utiliza un espejo medio-plateado que refleja, por un lado los objetos reales que se desean ampliar, y por otro lado la información digital que se va a mostrar. Para ello juega con los diferentes ángulos de visión del espejo, creando el efecto de superposición de las imágenes virtuales a la escena real. Una representación del funcionamiento se puede ver en la figura 4.

Los requisitos hardware son bastante limitados, siendo necesario para llevar a cabo su implantación de los siguientes elementos: un espejo medio-plateado, un ordenador donde se llevan a cabo los cálculos, una pantalla convencional desde la que se genera la imagen virtual que se refleja en el espejo y una cámara y dispositivo de detección de distancia que detecta el movimiento de la vista del usuario para cambiar el enfoque. Además, se puede añadir dispositivos de audio para mejorar la experiencia del usuario.

Desde el punto de vista del software requerido, el sistema necesita de un módulo de comunicación entre los dispositivos y el PC. En el apartado gráfico se emplea la librería OpenGL ES para llevar a cabo las operaciones pertinentes.



**Figura 3.3.** Vista general del funcionamiento del sistema POLAR.

**Tabla 3.6.** Resumen de las técnicas empleadas en POLAR**.**

**Técnica**

**Componente**

**so**

**ftware**

**Componente**

**hardware**

Reconocimiento imágenes OpenGL ES WebCam, sensor de movimiento

Tratamiento imágenes OpenGL ES Procesador y memoria del PC

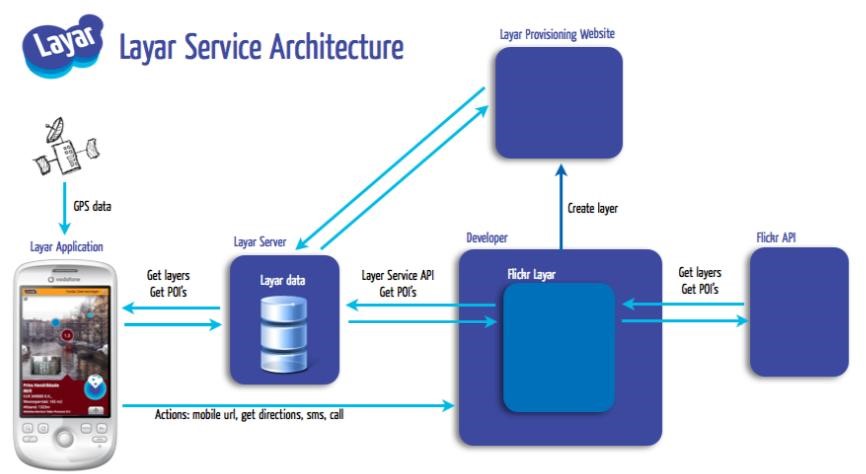
Visualización Delegado en software Pantalla del PC, espejo del dispositivo medio reflectante

#### 2.2.5 Layar

Layar es una aplicación para Android que permite la visualización de información digital sobre un entorno real en teléfonos móviles [LAY 10] Para analizar los requisitos para implantar este sistema se hará una división entre los requisitos del cliente o usuario y los requisitos del proveedor.

El funcionamiento del sistema se explica a continuación. El terminal muestra la escena real mediante su cámara al usuario. Mediante el GPS del móvil, se envía a un servidor la localización del usuario junto a la orientación extraída con la brújula del mismo. Con estos datos, el servidor responde con una capa con la imagen o el texto del lugar que se está enfocando con el teléfono. La figura 5 esquematiza el funcionamiento del sistema.

En primer lugar se explican los requisitos por parte del cliente. Para que el cliente pueda usar la aplicación es necesario que disponga de un teléfono móvil que utilice Android y que tenga una cámara de video integrada en el móvil. Por lo general, todos los teléfonos móviles de última generación disponen de una. Por último, es también indispensable que el terminal tenga localización GPS.



**Figura 3.4.** Arquitectura del sistema Layar.

En el apartado de requisitos por parte del proveedor existe una enorme diferencia. Para poder dar este tipo de servicio a los usuarios es necesario disponer de un servidor que reciba las peticiones de los terminales y les envíe la información que deberá ser ampliada en el mismo. Por otro lado, hace falta una base de datos que almacene las diferentes capas junto a sus respectivas localizaciones. Esto implica, como en cualquier arquitectura de este tipo, un nuevo requisito, que es el de personal de mantenimiento y actualización de los datos.

**Tabla 3.7. Resumen de las técnicas empleadas en Layar.**

**Técnica**

**Componente**

**software**

**Componente**

**hardware**

|  |  |  |
| --- | --- | --- |
| Reconocimiento geoposicionamiento | Brújula digital del  dispositivo, software GPS | Cámara de teléfono móvil o PDA, antena GPS |
| Almacenamiento de imágenes | Gestor de base de datos | Servidor de base de datos |
| Visualización | Delegado en software del dispositivo | Pantalla del teléfono móvil o PDA |

### 

### 2.2.6 LOCUS

Es un sistema de realidad aumentada desarrollado para terminales móviles, especialmente para PDAs y teléfonos 3G [LOC 10]. El usuario puede ver información ampliada sobre determinados lugares con su dispositivo gracias a las técnicas de geo-posicionamiento del sistema.

Para hacerlo posible, han desarrollado modelos 3D para aumentar la escena real. Por otro lado, otra de las tareas que se propusieron fue la de crear herramientas con las que pudieran caracterizar el contexto espacial en el que los usuarios iban a requerir realidad aumentada.

Los requisitos del sistema eran los siguientes: almacenar y gestionar la información espacial y posicional de los lugares; disponer de herramientas de modelado urbano con las que pudieran realizar las imágenes en 3D; poder utilizar herramientas de navegación y estimación de rutas; proveer de un sistema de posicionamiento del dispositivo móvil; poseer un interfaz para trabajar con realidad virtual y otro para realidad aumentada.

**Tabla 3.8.** Resumen de las técnicas empleadas en Locus**.**

**Técnica**

**Componente**

**software**

**Componente**

**hardware**

Reconocimiento geo- Brújula digital del Cámara de teléfono posicionamiento dispositivo, software móvil o PDA, antena

GPS GPS

Visualización Delegado en software Pantalla del teléfono del dispositivo móvil o PDA

Una vez analizados diferentes proyectos y sistemas de realidad aumentada se ofrecerá, a modo de resumen, una serie de componentes hardware y software específicos habitualmente utilizados en este tipo de sistemas.

Dentro de los componentes de hardware se encuentran:

* Dispositivos de captura de imágenes: HMD modelo CyberMind, Visette45 SXGA; WebCam SPC 1300 NC.
* Procesamiento y reconocimiento de imágenes: Intel Core 2 Duo 2.33 GHz, 2 Gb RAM, NVIDIA GeForce MX 440-SE.

El resto de hardware no será citado, debido a que o bien se encuentra integrado en los dispositivos (caso de los componentes de comunicación GPS en los teléfonos móviles) o porque el modelo no es relevante (como sucede con el disco duro, pantalla de ordenador, etc.).

Desde el punto de vista software, se pueden encontrar los siguientes modelos:

* Procesamiento y reconocimiento de imágenes: ARToolkit, Java3D.
* Almacenamiento: PostGreeSQL, MySQL.

El resto de componentes no serán descritos por los mismos motivos que en el análisis del hardware.

## 2.3 Arquitecturas de sistemas de realidad aumentada

Los sistemas de realidad aumentada son sistemas informáticos en los que existe una interacción entre el sistema y el usuario. Como en cualquier otro sistema de estas características, el tiempo de respuesta es un componente crítico que determina en gran medida el éxito o fracaso de una aplicación. En función de las prestaciones del sistema y los componentes hardware disponibles y la complejidad de procesamiento de sus procesos, es normal encontrar dos tipos de arquitecturas básicas [LOP 10]: los sistemas autónomos y los sistemas distribuidos. En esta sección se explicarán sin entrar en demasiado detalle estos dos conceptos.

Los *sistemas autónomos* son aquellos en los que la totalidad de las operaciones y tareas del sistema se llevan a cabo en un terminal único tal y como se muestra en la Figura 3.1. Este tipo de aplicaciones suelen ejecutarse en entornos muy limitados en cuanto a espacio y cuya información cabe en un terminal. También puede darse el caso de que el número de muestras a aumentar y la información que se añade sea suficientemente pequeña para poder ser procesada y mostrada en un solo dispositivo. Dentro de este grupo se pueden encontrar desde ordenadores convencionales conectados con una webcam, hasta teléfonos móviles de última generación, pasando por equipos que están compuestos por dispositivos de visualización y procesamiento conectados entre sí. Por lo general, este tipo de sistemas no requieren de equipamiento demasiado costoso debido a lo limitado de sus prestaciones, pudiendo llevarse a cabo en ordenadores domésticos. Este último caso suele ser más propenso a ser utilizado en entornos de prueba o en primeras fases del proyecto.



**Figura 3.5.** Imagen del videojuego Invizimals [INVZ 10] que utiliza técnicas de realidad aumentada en un solo dispositivo, una PSP.

Los *sistemas distribuidos* son aquellos en los que se delega en un equipo denominado servidor la tarea de realizar la mayor parte de los cálculos, rebajando en los terminales de visualización la carga de procesamiento. Habitualmente en este tipo de arquitectura los terminales se encargan de la recepción de la escena por cualquiera de los métodos ya expuestos en la sección primera de este documento y de la visualización de la realidad aumentada. El proceso intermedio se realiza en el servidor, aunque existen excepciones en las cuales parte de este trabajo se lleva a cabo en los propios terminales. Esta responsabilidad depende también de la aplicación que se está llevando a cabo, así como de las posibilidades del terminal. Por ejemplo, en aplicaciones implantadas en terminales móviles de última generación una posible estrategia es la de realizar parte del procesado de la imagen como detección de bordes mediante transformadas de Hough o métodos más ligeros e imprecisos. De esta forma, se envía al servidor los datos sobre los bordes de la imagen, evitando sobrecargas en la red. No obstante, no suele ser lo habitual por lo que se suele enviar la imagen al completo o comprimida al servidor. Por otra parte, aquellas aplicaciones que envíen datos al servidor de forma inalámbrica requieren de dispositivos de envío y recepción de señales. En entornos cerrados lo habitual suele ser utilizar conexiones W-LAN, por lo que los terminales necesitan antenas de transmisión y recepción y en el lado del servidor hace falta un punto de acceso. En aplicaciones para teléfonos móviles se requieren de equipos de transmisión y recepción GSM en ambos lados.

### 

### 3.3.1 Arquitecturas en sistemas autónomos de realidad aumentada

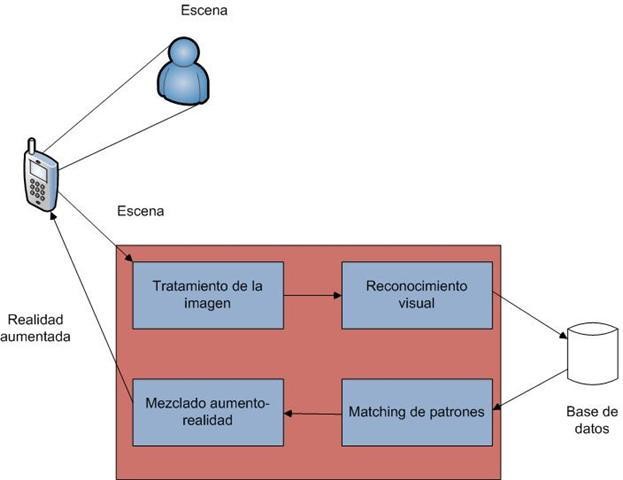
En esta sección se muestran las conclusiones alcanzadas tras analizar la arquitectura interna de las aplicaciones: ULTRA [ULT 10], MARCH [CHO 09], POLAR [OW 05] y ARToolKit [ART 10].

Este tipo de aplicaciones coincidían todas en que el reconocimiento está basado en la detección visual de algún elemento significativo. Puesto que las librerías de reconocimiento y tratamiento de imágenes han sido estudiadas en el capítulo 3, en esta sección se enfocara en la perspectiva funcional de cada uno de los módulos que componen este tipo de aplicaciones.

La arquitectura interna de este tipo de aplicaciones suele estar compuesto por, al menos, los siguientes módulos:

* Captación de escenario.
* Tratamiento de imágenes.
* Reconocimiento visual.
* Matching de patrones.
* Mezclado de realidad y aumento.
* Visualización.

La figura 3.6 muestra una arquitectura interna típica de una aplicación de realidad aumentada que funciona mediante reconocimiento visual.



**Figura 3.6.** Arquitectura típica de una aplicación de realidad aumentada con reconocimiento visual.

Para terminar, cabe destacar que si no se describe la arquitectura de una aplicación con reconocimiento por geo-posicionamiento es porque, a nuestro juicio, no tiene sentido implantar una aplicación local con estas características. Las razones de estas conclusiones son básicamente dos:

* La cantidad de información sobre lugares suele ser bastante amplia, siendo necesario el uso de bases de datos clusterizadas.
* La información sobre lugares suele ser cambiante. Por ejemplo, una aplicación puede tener configurado que en cierto punto del mapa hay un restaurante que por algún motivo ha dejado de existir, ofreciendo información falsa.

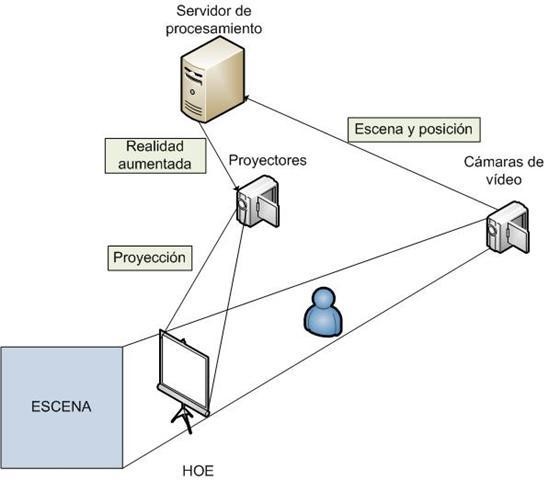
### 2.3.2 Arquitecturas en sistemas distribuidos de realidad aumentada

En esta sección se analizarán las arquitecturas de los sistemas distribuidos considerados en esta sección.

#### ASTOR

El proyecto ASTOR consiste en un sistema distribuido en el que contribuyen al funcionamiento de éste varios elementos. Estos elementos son:

* Cámaras de vídeo. Se encargan de percibir la escena y la posición del usuario.
* Un equipo de procesamiento. Su función es la de realizar las tareas de tratamiento de imágenes. Actúa de servidor.
* Módulo de visualización. Compuesto por una serie de proyectores y pantallas HOE.

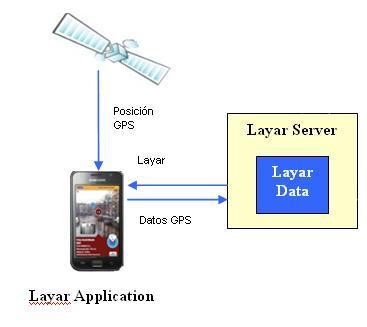


**Figura 3.7.** Arquitectura del sistema ASTOR.

#### LAYAR

El sistema LAYAR tiene una naturaleza distribuida en la que se pueden añadir contenidos o Layares al sistema, haciéndolo todo lo escalable que se desee.

Para su funcionamiento, LAYAR basa las tareas de reconocimiento por geo-posicionamiento en el teléfono móvil, delegando la elección de la capa o Layar de superposición a los servidores dedicados.



**Figura 3.8.** Arquitectura del sistema LAYAR.

#### LOCUS

El último sistema cuya arquitectura se va a analizar es el sistema LOCUS. Este sistema tiene dos componentes:

* La aplicación del teléfono móvil. Se encarga de recibir las señales GPS y de superponer la información.
* Un sistema de almacenamiento de datos sobre los lugares que se pueden aumentar.

Puesto que el funcionamiento de este sistema es muy similar al del sistema LAYAR, su arquitectura es también bastante parecida.

¿Se emplea el uso del clasificador de imágenes en la vida cotidiana?

**Físico Geográfico**: Dar a conocer, y así implementar el desarrollo, uso y relevancia en la vida diaria en la Ciudad de México.

**En el tiempo**: Se dio a conocer en 1990 y a raíces de eso a tenida relevancia en la actualidad, hoy en día en el año 2018

**Significado**: La Realidad Aumentada (RA) consiste en sobreponer objetos o animaciones generadas por computadora sobre la imagen en tiempo real que recoge una cámara web.

De esta manera se puede "aumentar" en la pantalla, la realidad que mira la cámara con los elementos de una realidad virtual "Es el entorno real mezclado con lo virtual".

A diferencia de la realidad virtual, la RA es una tecnología que complementa la percepción e interacción con el mundo real y permite al usuario estar en un entorno aumentado con información generada por una computadora.

<http://www.cedicyt.ipn.mx/RevConversus/Paginas/RealidadAumentada.aspx> (En español) CEDICYT IPN desde el 27 de noviembre del 2016.

Significante: La realidad Aumentada es un uso o software que utiliza como sensor la cámara para poder reconocer una imagen u objeto y así transmitir un dato o visualizar un objeto.

**Oraciones Tópicas**:

* La investigación pretende tener una visión amplia del tema en aspectos a otros campos y para así lograr su implementación en la vida diaria.
* El interés de la siguiente investigación es conocer la aplicación, en que beneficiara en el aspecto cotidiano en la vida diaria del mexicano.

**Recursos**:

* Humanos: En el campo se solicitara apoyo de expertos en el campo, como investigadores de Realidad Aumentada, desarrolladores de la misma.
* Materiales: Se utilizaran materiales como computadora, software de manejo para Realidad Aumentada y un dispositivo móvil para hacer las pruebas.
* Financieros: Se implementarán gastos de licencias de software, compra de dispositivos para el desarrollo de la misma.

**Objeto de Estudio**

“Clasificador de Imágenes con Búsqueda inversa en tiempo real y sus beneficios.”

**Problema de investigación**

Se verá la importancia o la relevancia de ver si se conoce su desarrollo e implementación del clasificador de imágenes en la web en México y que tan avanzada es para su implementación y accesible para el uso cotidiano.

## REFERENCIAS

**[SUT 68]** Sutherland, I. 1968. “A head-mounted three-dimensional display. In Proceeding of the Fall Joint Computer Conference”. AFIPS Conference Proceedings, vol. 33. AFIPS, Arlington, VA., 757- 764.

**[VRES 10]** http://www.virtualresearch.com/

**[DOD 08]** Department of Defense, United States of America, GLOBAL POSITIONING SYSTEM STANDARD POSITIONING SERVICE PERFORMANCE STANDARD, 2008.

**[GPS 10]**<http://www.gps.gov/>

**[SOM 02]** J. A. Somolinos Sánchez, “Avances en robótica y visión por computador”, 2002.

**[CAW 08]** S. Cawood, M. Fiala, “Augmented Reality: A practical guide”, 2008.

**[BIM 05]** O. Bimber, R. Rakar, “Spatial Augmented Reality. Merging Real and Virtual Worlds”, 2005.

**[LOP 10]**

**[CHO 09]** O. Choudary, V. Charvillat, R. Grigoras, P. Gurdjos, “MARCH: Mobile Augmented Reality for Cultural Heritage”, 2009.

**[TAT 07]** K. Tateno, I. Kitahara, Y. Ohta, “A Nested Marker for Augmented Reality”, 2007. **[RGB 98]** Multimedia systems and equipment - Colour measurement and management - Part 2-1: Colour management - Default RGB colour space – sRGB, 1998.

**[SHR 05]** D. Shreiner, M. Woo, J. Neider, T. Davis, “OpenGl Programming Guide: The Official Guide to Learning OpenGL“, 2005.

**[SHR 00]** D. Shreiner, “OpenGl Reference Manual: The Official Reference Document to OpenGL“, 2000.

**[GL ES 10]**<http://www.khronos.org/opengles/>

**[JIA 09]** H. Jiang, S. X. Yu, “Linear Solution to Scale and Rotation Invariant Object Matching “, 2009.

**[ART 10]** http://[www.hitl.washington.edu/artoolkit/](http://www.hitl.washington.edu/artoolkit/)

**[ART P 10]**<http://www.studierstube.icg.tu-graz.ac.at/handheld_ar/artoolkitplus.php/>

**[JART 10]** <http://www.c-lab.de/jartoolkit/>

**[OKA 96]** S. Okabayashi, “Visual Optics of Head-Up Displays (HUDs) in Automotive Applications“, 1996.

**[JAV 02]** B. Javidi, “Image Recognition and Classification. Algorithms, Systems and Applications”, 2002.

**[PLA 06]** C. Platero, “Apuntes de visión artificial”, 2006.

**[ESC 01]** A. Escalera, “Visión por Computador: Fundamentos y Métodos”, 2001.

**[PAJ 01]** G. Pajares, “Visión por computador”, 2001.

**[BRA 01]** T. Bräunl, S. Feyrer, W. Rapf, M. Reinhardt, “Parallel Image Processing”, 2001. **[AML 08**] K. Amlacher, L. Paletta, “Geo-Indexed Object Recognition for Mobile Vision Tasks”, 2008.

**[DEY 00]** A. K . Dey, G. D. Abowd, “Towards a Better Understanding of Context and ContextAwareness”, 2000.

**[WIKTD 10]**<http://www.wikitude.org/>

**[LAY 10]**<http://www.layar.com/>

**[OLW 05]** A. Olwal, C. Lindfors, J. Gustafsson, T. Kjellberg, L. Mattsson, “ASTOR: An Autostereoscopic Optical See-through Augmented Reality System“, 2005.

**[XU 07]** Xu, “GPS. Theory, Algorithms and Applications“, 2007.

**[SYMB 10]**<http://www.symbian.org/>

**[MCPH 06]** F. McPherson, “How to Do Everything with Windows Mobile“, 2006. **[IPH DEV 10]**<http://developer.apple.com/iphone/index.action>**[SING 06]** A. Singh, “Mac OS X Internals. A Systems Approach“, 2006. **[AND SDK 10]**<http://developer.android.com/sdk/index.html>

**[J3D 10]**<http://java.sun.com/javase/technologies/desktop/java3d/>

**[ULT 10]** http://www.ist-ultra.org/

**[NOK CV 10]** http://research.nokia.com/research/projects/nokiacv/

**[JRMI 10]**<http://java.sun.com/docs/books/tutorial/rmi/index.html>

**[OW 05]** A. Owal, T. Höllerer, “POLAR: Portable, Optical see-through, Low-cost Augmented Reality“, 2005.

**[INVZ 10]**<http://www.invizimals.com/>

**[GOGG 10]**<http://www.google.com/mobile/goggles>

**[G SANT 10]** http://www.fdi.ucm.es/migs/presentacion.html

**[AND TTS 10]**<http://developer.android.com/resources/articles/tts.html>

**[HIL 95]** J. R. Hilera, V. J. Martínez, “Redes neuronales artificiales. Fundamentos, modelos y aplicaciones”, 1995.

**[PAT 96]** D. W. Patterson, “Artificial neural Networks. Theory and Applications”, 1996.

**[HAY 94]** S. Haykin, “Neural Networks – A Comprensive Foundation”, 1994.

**[MAT 10]** H. Demuth, M. Beale, M. Hagan, “Neural Network Toolbox 6. User’s guide”, 2010. **[LEV 44]** K. Levenberg, “A Method for the Solution of Certain Non-Linear Problems in Least Squares”, 1944.

**[ZAK 06]** N. C. Zakas, J. McPeak, J. Fawcett, “Professional AJAX”, 2006.

**[JSON 10]** http://www.json.org/

**[XPATH 10]** http://www.w3.org/TR/xpath/

**[IPV 10]**<http://www.upv.es/>

**[SEU 09]** S. J. Kim, A. K. Dey, “Simulated Augmented Reality Windshield Display as a Cognitive Mapping Aid for Elder Driver Navigation”, 2009.

**[AVE 08]** S. J. Kim, A. K. Dey, “Simulated Augmented Reality Windshield Display as a Cognitive Mapping Aid for Elder Driver Navigation”, 2009.

**[YUV 95]** RECOMMENDATION ITU-R BT.601-5, 1982-1995.

**[YUV 02]** RECOMMENDATION ITU-R BT.709-5, 1990-2002.

**An Analysis of Augmented Reality Systems**

|  |  |  |
| --- | --- | --- |
| Héctor López  Telefónica I+D  Madrid, Spain hlp@tid.es | Antonio Navarro  Univ. Complutense de Madrid  Madrid, Spain anavarro@sip.ucm.es | José Relaño  Telefónica I+D  Madrid, Spain joserg@tid.es |