

Tesis de Máster



Universidad del País Vasco Euskal Herriko Unibertsitatea

Sistema de localización y seguimiento de personas en interiores mediante cámara PTZ basado en las tecnologías *Kinect y Ubisense*

Aritz Legarretaetxebarria

Donostia - San Sebastián, Septiembre 2011

Universidad del País Vasco / Euskal Herriko
Unibertsitatea

Departamento de Ciencia de la Computación
e Inteligencia Artificial

Directores: Elena Lazkano Ortega (UPV/EHU)
Maider Laka Iñurrategi (Vicomtech)

www.ccia-kzaa.ehu.es

RESUMEN

Los entornos inteligentes están cada vez más orientados a apoyar la actividad humana. Uno de sus objetivos se centra en la obtención de la y la actividad de las personas en entornos variados. Para ello, el estado del arte ofrece numerosas técnicas y soluciones. Uno de estos entornos, donde más actividad está habiendo en la investigación de estos sistemas, es el de la televisión. Conociendo la posición de las personas que aparecen en ella, es posible realizar un seguimiento mediante una cámara robotizada y añadir una información virtual adicional de manera dinámica. En esta tesis de máster se revisará un estado del arte de los sistemas de localización inalámbricos en interiores existentes hasta la fecha. A partir de estas evaluaciones se realizará un desarrollo experimental con el fin de localizar los objetos mediante algunos de los sistemas y evaluar la fiabilidad de estos. Posteriormente se realizará un prototipo funcional capaz de realizar seguimientos a estos objetos mediante una cámara de video robotizada PTZ (pan-tilt-zoom), con un algoritmo desarrollado para este fin.

Palabras clave: Localización en interiores, localización inalámbrica, *Kinect*, *Ubisense*, segmentación de imagen

LABURPENA

Gaur egun, ingurune adimendunak, gizakien eguneroko jardunean laguntza emateko geroz eta bideratuagoak daude, pertsonen posizio eta jarduerak antzemateko beharra gero eta handiagoa dela ikusten delarik. Horretarako teknika eta soluzio ugari erabili da. Ildo honetako inguruneen artean, ikerkuntza gehien jasaten ari den arloa telebistarena da. Bertan agertzen diren pertsonen posizioa jakinez gero, posible izango zen robotdun kamera batekin jarraipen bat egitea eta dinamikoki alegiazko informazioa gehitzea. Masterreko tesi honetan, gaur egunera arte garatu diren haririk gabeko barrualdeko sistemetan oinarrituriko teknologiaren egoeraren berrikuspena egingo da. Egindako ebaluazioetan oinarrituta, aztertutako sistema batzuen bidez objektuen kokalekua aurkitu eta hauen fidagarritasuna ebaluatzeko asmoz, garapen esperimentalak egingo da. Ondoren, prototipo funtzional bat garatuko da, robotdun PTZ (pan-tilt-zoom) kamera batekin kokalekua aurkitutako objektuei jarraipena egiteko, helburu honetarako garatu den algoritmo bat erabiliz.

ÍNDICE

1. CAPÍTULO 1: INTRODUCCIÓN	11
1.1 Justificación	12
1.2 Vicomtech-IK4	14
1.3 Antecedentes	15
1.3.1 Proyecto Tebas:.....	15
1.3.2 Proyecto Pangea:.....	16
1.4 Objetivos	17
1.4.1 Objetivo general.....	17
1.4.2 Objetivos específicos	17
2. CAPÍTULO 2: TÉCNICAS DE LOCALIZACIÓN	19
2.1 Triangulación	19
2.1.1 Técnica basadas en RSS:	19
2.1.2 Técnica TOA:	20
2.1.3 Técnica TDOA:.....	20
2.1.4 Técnica RTOF:	22
2.1.5 Técnica AOA:	22
2.2 Análisis de escena	23
2.2.2 Método k-nn:.....	23
2.2.3 Redes neuronales artificiales:	24
2.2.4 SVM:.....	24

2.2.5 SMP:.....	25
2.3 Proximidad	25
2.4 Técnicas de visión artificial.....	25
3. CAPÍTULO 3: SISTEMAS DE LOCALIZACIÓN.....	27
3.1 Global Positioning System (GPS)	27
3.2 Radio Frequency Identification (RFID).....	28
3.3 Sistema basado en comunicaciones móviles	31
3.4 UWB (Ultra Wave Band)	32
3.5 WLAN (IEEE 802.11).....	34
3.6 Bluetooth (IEEE 802.15).....	38
3.7 ZigBee.....	39
3.8 Ultrasonidos.....	41
3.9 Infrarrojos (IR)	44
3.10 Visión	45
4. CAPÍTULO 4: PARÁMETROS DE EVALUACIÓN.....	49
4.1 Precisión	49
4.2 Requerimientos.....	49
4.3 Coste	50
4.4 Solución escogida.....	50
4.4.1 <i>Ubisense</i> :	52
4.4.2 <i>Kinect</i> :	56

4.4.3 Fusión:	57
5. CAPÍTULO 5: EXPERIMENTOS	59
5.1 Sistema <i>Ubisense</i> :	63
5.2 <i>Kinect</i> :	67
5.3Fusión:.....	73
6. CAPÍTULO 6: CONCLUSIONES	75
7. CAPÍTULO 7: TRABAJOS FUTUROS	77
REFERENCIAS.....	79

Índice de ilustraciones:

Ilustración 1-1: <i>Captura del sistema TEBAS en un momento de las pruebas en la Triatlón de Bilbao</i>	16
Ilustración 1-2: <i>Salida del sistema para la Simulación en Tiempo Real de Regatas para la Televisión.</i>	17
Ilustración 2-1: <i>Método TOA</i>	20
Ilustración 2-2: <i>Método TDOA</i>	21
Ilustración 2-3: <i>Método AOA</i>	23
Ilustración 4-1: <i>Arquitectura del sistema de Localización comercial UBISENSE</i>	54
Ilustración 4-2: <i>Dispositivo comercial Kinect</i>	56
Ilustración 4-3: <i>Funcionamiento del dispositivo Kinect</i>	56
Ilustración 5-1: <i>Laboratorio de televisión de Vicomtech</i>	59
Ilustración 5-2: <i>Arquitectura del sistema de comunicación entre el PC y el robot de la cámara</i>	59
Ilustración 5-3: <i>Trama de datos para la comunicación con el robot de la cámara</i>	60
Ilustración 5-4: <i>Superficie destinada a la detección con el sistema de referencia definido</i>	61
Ilustración 5-5: <i>Configuración de los movimientos del robot en el sistema de referencia definido</i>	62
Ilustración 5-6: <i>Distribución de sensores en la escena</i>	63
Ilustración 5-7: <i>Montaje del sistema de localización Ubisense</i>	64
Ilustración 5-8: <i>Posiciones (X) donde se ha medido la precisión</i>	65
Ilustración 5-9: <i>Dispositivo kinect colocado sobre la cámara de video</i>	67

Ilustración 5-10: <i>Detección del individuo al que se le realizara el seguimiento</i>	68
Ilustración 5-11: <i>Filtro de seguimiento del individuo</i>	69
Ilustración 5-12: <i>Profundidad de la imagen representada por diferentes colores</i>	69
Ilustración 5-13: <i>Relación del sistema de referencia del escenario con el de la imagen</i>	70
Ilustración 5-14: <i>Vector de movimiento entre dos fotogramas</i>	71
Ilustración 5-15: <i>Tendencia del error en (puntos rojos) en cada punto de calibración (puntos negros)</i>	73

Índice de tablas:

Tabla 4-1: <i>Tabla comparativa de los diferentes sistemas de localización</i>	51
Tabla 5-1: <i>Comparativa de los valores del robot con los valores de los ángulos</i>	62
Tabla 5-2: <i>Errores de precisión en los diferentes puntos del escenario para el sistema Ubisense</i>	65
Tabla 5-3: <i>La media de error del sistema Ubisense</i>	66
Tabla 5-4: <i>Errores de precisión en los diferentes puntos del escenario para el sistema Kinect</i>	72
Tabla 5-5: <i>La media de error del sistema Kinect por eje y su media</i>	72

CAPÍTULO 1: INTRODUCCIÓN

En esta tesis de máster, se pretende llevar a cabo un proyecto de desarrollo experimental para diseñar y desarrollar un sistema de adquisición de contenidos de bajo coste que sirva para las retransmisiones televisivas que no estén respaldadas por grandes índices de audiencia, como pueden ser programas de entretenimiento de bajo seguimiento que se llevan a cabo en platós de televisión, para los cuales un gran despliegue en infraestructura y personal no está justificado.

En concreto se trata la problemática del seguimiento automático de objetos o individuos a través del movimiento de cámaras robotizadas, cámaras PTZ (Pan, Tilt y Zoom). El empleo de cámaras robotizadas que trabajan de una forma autónoma contribuiría directamente en reducir los costes de una producción con pocos recursos y asegurar así su retransmisión. El objetivo es diseñar un software capaz de hacer mover una cámara robotizada, en función del movimiento realizado por el individuo al que ésta debe seguir. Para ello, se debe hallar la posición del individuo en un entorno definido y en tiempo real, y alimentar dicho software con estos datos para el correcto movimiento del robot de la cámara.

Esta memoria está estructurada de la siguiente manera:

En el capítulo 1, se realizará una introducción al entorno en el que se ha desarrollado el trabajo de la tesis de máster y a la problemática, definiendo los objetivos del proyecto. Una vez establecidos los objetivos se realizará una exposición de las diferentes técnicas de localización.

En el capítulo 2, se hará un resumen de las técnicas de localización inalámbricas más comunes y algunos métodos que se utilizan para la localización de dispositivos.

En el capítulo 3, se describirá brevemente un estado del arte de los diferentes sistemas de localización existentes hasta la fecha, que sean capaces de obtener la posición de un objeto en un entorno definido, evaluando las características técnicas de cada uno de ellos. Esta sección es muy importante, porque basándose en este estudio se definirán los sistemas que mejor se amoldarán a las necesidades del proyecto para poder elaborar los diferentes experimentos a realizar.

En el capítulo 4, se realizara un estudio de los parámetros de evaluación que se tendrán en cuenta para seleccionar las mejores opciones para llevar a cabo el desarrollo experimental deseado.

En el capítulo 5, se describirán detalladamente las pruebas que se llevaran a cabo detallando las arquitecturas y configuraciones para la implementación de los diferentes módulos y los equipos utilizados para ello. Además, se especificaran las incidencias técnicas superadas para el correcto funcionamiento de cada módulo y el conjunto de éstos. Por último, se detallarán los ensayos pertinentes para las diferentes arquitecturas descritas.

Una vez realizados estos ensayos se procederá a documentar todos los resultados extraídos de los mismos, plasmándolos en tablas y gráficos, las cuales se cotejaran con el fin de obtener las conclusiones a la problemática planteada. Una vez extraídas las conclusiones se comentaran algunas líneas futuras mediante las cuales se trataría de abrir nuevas vías tratando de mejorar lo realizado en este trabajo.

1.1 Justificación

La llegada de la televisión digital terrestre (TDT) ha conducido a una mayor segmentación de canales, propiciando la creación de nuevas opciones de canales temáticos menos generalistas. Algunos de estos nuevos canales no gozan de grandes índices de audiencia, por ejemplo competiciones deportivas de bajo seguimiento, en los cuales se tiende a producciones cada vez de menor coste. Incluso hay ciertos eventos que, aunque tienen una demanda importante por parte de la audiencia, resulta imposible emitir debido al alto coste que supone.

Es por ello que empiezan a aparecer nuevos productos en el mercado orientados a ajustar costes, reduciendo poco a poco el material, el personal y el espacio necesario, pero siempre sin sacrificar la calidad del resultado final.

En este sentido, en este proyecto se estudiará la automatización de las cámaras que cubren tales eventos. Para ello, se utilizarán cámaras robotizadas situadas en puntos estratégicos, teniendo conocimiento de la posición en la que esta la cámara en relación al entorno y que sean capaces de funcionar automáticamente. Estas cámaras tienen que

ser capaces de realizar un seguimiento del individuo y de esta manera tener la opción de presentar una información adicional sobre ese individuo de forma automática, bien en el caso de un presentador o en un futuro, en el caso de un deportista. La información para los espectadores estará garantizada pero se reducirán los costes sujetos al personal y a la infraestructura en comparación con los grandes acontecimientos.

En la última década ha habido mucha actividad en el desarrollo de sistemas automáticos para el seguimiento de personas. En la mayoría de los casos, el *tracking* de personas se ha realizado mediante la visión por computador, en los que se buscaba seguir los movimientos de uno [1] o varios individuos [2][3][4][5][6], con una sola cámara o con varias [7], y detectar las caras [8][9][10][11] mediante cámaras PTZ. Además, se han realizado infinidad de estudios en seguimientos de deportistas con una sola cámara [12][13][14][15][16] y también con múltiples cámaras [17][18][19], siendo la finalidad de estos la obtención de información de los movimientos de los deportistas para su posterior análisis.

La mayoría de los estudios analizados en la literatura se basan en la información de color extraída de la imagen, las formas y los vectores de movimiento. Con una sola cámara puede resultar un tanto complicado realizar el seguimientos de diferentes personas que pertenecen al mismo equipo, porque su vestimenta será idéntica, y en cuanto se produzca alguna oclusión puede que el nuevo individuo al que el sistema siga no sea el deseado. En este trabajo se busca que la cámara realice este seguimiento individualmente, sin necesidad de asociación de información procedente de diferentes cámaras, ya que el campo de visión de estas puede no guardar ninguna relación. Por este motivo, se estudiará el empleo de sistemas de localización, para que con la suma de esta información adicional en tiempo real, la cámara pueda superar éstas limitaciones.

Existen deportes en los que se emplean dispositivos hardware, con los cuales se obtienen las posiciones de los participantes. Por ejemplo, en emisiones en exteriores, estas posiciones se recogen mediante un sistema de GPS (Global Positioning System), como pueden ser la Formula1, Moto GP y Regatas de Traineras. Pero en los eventos desarrollados en interiores no se podrían obtener estos parámetros mediante este sistema, por su falta de cobertura. Se deberá buscar alguna solución para las necesidades

de eventos en interior y pero también para eventos en exteriores donde no se puede utilizar GPS, por lo complejo que esto pueda resultar.

A su vez, las técnicas de realidad virtual están siendo continuamente extendidas a los más diversos entornos. La televisión es uno de los ámbitos donde más se aprovechan estas técnicas que permiten abaratar costes de producción, generar efectos especiales, enriquecer el grafismo, generar nuevas modalidades de ingreso, etc. Las emisiones deportivas se sirven de las posibilidades que ofrecen las técnicas de *realidad virtual* o *realidad aumentada* para realzar la información o mostrar datos de una forma mucho más intuitiva. Para ello hay estudios en los que se buscan la posición, orientación y la distancia focal de las cámaras, en tiempo real, para generar contenidos virtuales de forma dinámica [20]. Teniendo la posibilidad de realizar un seguimiento automático a los deportistas, las emisoras podrán presentar un mejor contenido y, por otro lado, los televidentes podrán acceder a emisiones más agradables, con mayor cantidad de información, en tiempo real.

1.2 Vicomtech-IK4

Esta tesis de máster se ha llevado a cabo en Vicomtech-IK4, bajo el seguimiento de un tutor de la UPV, con el fin de aplicar y poner en práctica los conocimientos y aptitudes técnicas adquiridos en el Máster Universitario en Ingeniería Computacional y Sistemas Inteligentes impartido en la UPV/EHU.

Vicomtech-Ik4 (Centro de Tecnologías de Interacción Visual y Comunicaciones) es un centro de investigación aplicada que trabaja en el área de gráficos por ordenador interactivos y tecnología multimedia, localizado en el Parque Tecnológico de San Sebastián.

Es miembro de la Alianza Tecnológica Ik4, compuesta por 8 centros tecnológicos punteros del País Vasco y pertenece a la red internacional de investigación GraphicsMedia.net, a la cual pertenecen prestigiosos centros de investigación aplicada a nivel internacional totalmente alineados con las tecnologías de Computer Graphics y Multimedia, lo que le da también un perfil activo y estratégico de internacionalización de su actividad investigadora.

La entidad Vicomtech-IK4 está distribuida en cinco áreas tecnológicas. Esta tesis de máster está enmarcada dentro del área *Televisión Digital y Servicios Multimedia*. Esta área, orientada al sector audiovisual, trabaja principalmente con todas aquellas tecnologías de radiodifusión digital (DVB-T,C,S,H – MHP, MPEG), análisis de contenido multimedia (para procesos de indexación automática, gestión y recuperación) y producción profesional de servicios interactivos, realidad virtual, HDTV, audio y broadcast para móviles.

El esfuerzo en investigación que realizan los componentes de este grupo de veinte investigadores (ingenieros e ingenieras de diferentes disciplinas: telecomunicaciones, industrial, informáticos, electrónicos) viene validado por el hecho de contar con un doctor y dos investigadores trabajando en su tesis doctoral. Este grupo participa en proyectos de investigación y desarrollo en el ámbito local, estatal e internacional, colabora en actividades docentes en la Universidad del País Vasco y ha participado en las labores de consultoría técnica y dirección de la división de TV en una de las *spin-offs* creadas por Vicomtech-IK4.

1.3 Antecedentes

La línea de producción televisiva del área Televisión Digital y Servicios Multimedia lleva años colaborando con productoras tanto a nivel local y nacional y está centrada en generar tecnología que aporte valor añadido a sus productos manteniendo siempre la máxima del “bajo coste”. Esta estrecha colaboración con las productoras y con entidades radiodifusoras hace que desde Vicomtech-IK4 se tenga buen conocimiento de las necesidades del sector audiovisual. A continuación se presentan los proyectos más significativos de la línea. Los resultados obtenidos de los mismos se han integrado dentro de la oferta de las productoras e incluso se han retransmitido a través de EiTb.

1.3.1 Proyecto Tebas:

Como resultado de este proyecto que se prevé finalizar a lo largo de este año 2011, se está obteniendo una plataforma de mezcla de video que a partir de varias entradas de calidad profesional genera una salida a la que se le puede añadir rotulación y grafismo apta para ser contribuida en tiempo real a las entidades

radiodifusoras. Este sistema además genera en tiempo real simulaciones tridimensionales a partir de la información de localización de los objetos de interés. La plataforma completa de bajo coste puede ser empleada como una mesa de realización televisiva, como se muestra en la Ilustración 1-1. A destacar, que esta mesa no contempla más de cuatro entradas. Los resultados preliminares de este proyecto se probaron en la Triatlón de Bilbao, cuya retransmisión se emitió en directo a través de la cadena ETB Sat.



Ilustración 1-1: Captura del sistema TEBAS en un momento de las pruebas en la Triatlón de Bilbao

1.3.2 Proyecto Pangea:

Este proyecto consistió en generar en Tiempo Real simulaciones de competiciones de regatas de traineras para ser integradas en las plataformas televisivas profesionales y emitidas a todos los usuarios. Entre las ventajas de este sistema cabe destacar que facilita la comprensión que los televidentes tienen del estado de la competición debido a que se pueden mostrar vistas diferentes a las tomadas con las cámaras colocadas en los helicópteros o embarcaciones. Además, como se ve en la Ilustración 1-2, se puede añadir información a la escena, tal como los canales por donde las embarcaciones deben navegar, la distancia recorrida, etc. Este sistema adapta la señal de realidad virtual al sistema analógico PAL (Phase Alternating Line).



Ilustración 1-2: Salida del sistema para la Simulación en Tiempo Real de Regatas para la Televisión.

En el desarrollo de estos proyectos, se ha visto la necesidad de utilizar cámaras que sean capaces de realizar seguimientos automáticos a los diferentes objetos o individuos que toman parte en estos eventos y poder añadir información adicional en tiempo real al contenido obtenido.

1.4 Objetivos

1.4.1 Objetivo general

Desarrollar una solución a nivel de hardware y software de un sistema que sea capaz de hacer que una cámara PTZ siga a objetos o individuos en el momento en el que este se está desplazando, con una calidad de panorámica aceptable, dentro de un área limitada en interior.

1.4.2 Objetivos específicos

Calibrar la cámara mediante un software para que funcione en esa área limitada, sabiendo cuál es su propia posición sobre el entorno para controlar los movimientos de ésta.

1. Se obtiene la posición 3D del individuo mediante un sistema de localización por hardware. La superficie en el que se realiza la búsqueda del individuo será limitada en función a la cobertura de la tecnología que se utiliza. Una vez obtenida esta información se le hace llegar al programa que controla la cámara para que ésta siga al individuo.

2. Se probará el sistema de movimiento de la cámara con otro método de obtención de la información de posicionamiento de los individuos, con el *tracking* de individuos en imágenes.
3. Combinando las dos técnicas anteriores, se tratará de mejorar el sistema de seguimiento de la cámara.

CAPÍTULO 2: TÉCNICAS DE LOCALIZACIÓN

En este capítulo se hará un resumen sobre las técnicas inalámbricas más utilizadas, y algunos métodos y algoritmos que se utilizan para la localización de objetos. Estas técnicas inalámbricas se basan en señales electromagnéticas. No es fácil realizar un modelado de las propagaciones de radio en interiores, debido a la baja probabilidad de visión directa entre dispositivos y algunas características físicas del entorno, como pueden ser la distribución del mismo, los objetos en movimientos y las numerosas superficies reflectantes. No existe un modelo estándar para los diferentes entornos de interior [21]. Para disminuir estos errores de medición se utiliza la triangulación tradicional, algoritmos de posicionamiento basado en análisis de la escena, los de proximidad y las técnicas de visión. Dependiendo de las características del entorno, cada uno de ellos presenta ventajas y desventajas, pero utilizando la combinación de los mismos se podrían obtener sistemas más robustos.

2.1 Triangulación

La triangulación utiliza las propiedades de los triángulos para obtener la posición del objeto, mediante el cálculo de distancias de varios puntos de referencia, estos puntos deben ser un mínimo de tres. Esto lo hace midiendo las distancias utilizando el RSS (Received Signal Strength), el TAO (Time Of Arrival), TDOA (Time Difference Of Arrival), el RTOF (Round Trip Time Of Flight) o le AOA (Angle Of Arrival). Este tipo de técnicas estiman la posición por medida de la distancia y a continuación se hará una descripción de cada una de ellas.

2.1.1 Técnica basadas en RSS:

Esta técnica se basa en la atenuación sufrida por la señal y requieren medidas temporales precisas, lo que supone un inconveniente en el entorno de interiores. En este escenario, el canal de comunicaciones sufre lo que se conoce como la multitrayectoria de la señal en el receptor, por las reflexiones que existen se producen. Y si encima la densidad de receptores no es muy alta, resulta difícil

encontrar canales entre los terminales móviles y alguno de los receptores, por lo que la estimación precisa de los tiempos de llegada resulta complicada. Es posible superar este inconveniente estimando la atenuación sufrida por la señal con modelos empíricos o teóricos que relacionen las pérdidas sufridas por la señal en función de la posición.

2.1.2 Técnica TOA:

Esta técnica se basa en la medición del tiempo de que emplea la señal en llegar a los receptores. La distancia entre dos terminales receptoras es directamente proporcional al tiempo de propagación de la señal entre ellos. Obteniendo dicha distancia para los tres nodos receptores, es posible determinar la posición del terminal móvil, tal y como se observa en la Ilustración 2-1. Esta aproximación tiene dos problemas: primero, todos los transmisores y receptores deben estar sincronizados perfectamente, lo que supone una complejidad muy alta en el sistema; segundo, es necesario incluir la información del instante de transmisión en la propia trama.

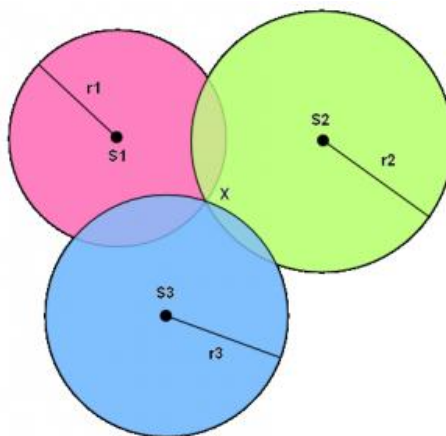


Ilustración 2-1: *Método TOA*

2.1.3 Técnica TDOA:

Estas técnicas basan la triangulación en las distancias a los emisores obtenidas de forma relativa (diferencia entre unas señales y otras). Las ondas de los diferentes emisores llegan al receptor con una diferencia de tiempo en función de las distancias del receptor respecto a los emisores de las mismas. Así, el receptor, tras

identificar los emisores que generaron las señales, sólo tiene que medir el tiempo transcurrido entre la llegada de cada una de las diferentes señales y calcular los puntos que cumplen la diferencia de tiempo obtenida. De este modo se genera, por cada diferencia de tiempos tomada, una hipérbola si la medición se realiza en 2 dimensiones o un hiperboloide si se hace en 3 dimensiones. La posición final se basa en el cálculo de la intersección de todas las hipérbolas o hiperboloides Ilustración 2-2.

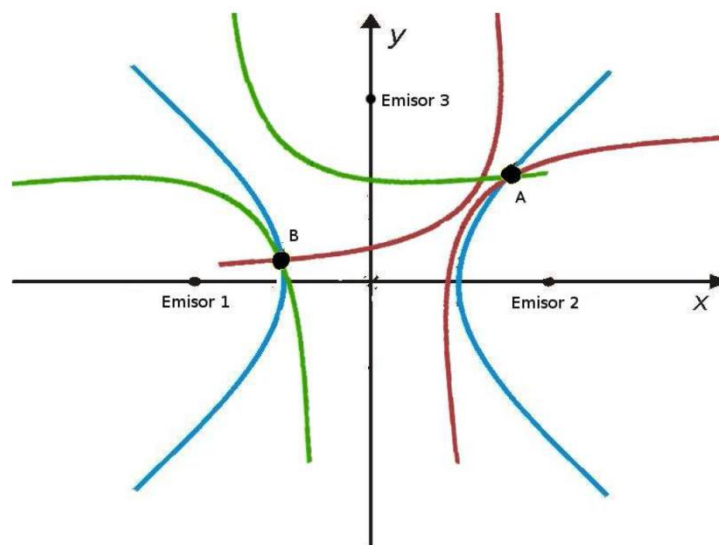


Ilustración 2-2: Método TDOA

Así, si todos los emisores emiten la señal al mismo tiempo, el receptor capturará primero la señal procedente del emisor que tenga más próximo (iniciando un contador para calcular la diferencia al resto de señales). Posteriormente, con una pequeña diferencia de tiempo, la señal del segundo emisor más cercano alcanzará al receptor (capturando el valor del contador, permitiéndole calcular la hipérbola azul), y así sucesivamente (obteniendo las diferencias de tiempo a los diferentes emisores, posibilitándole calcular las hipérbolas roja y verde) hasta que reciba la señal de todos los emisores implicados. Tras todo esto, el robot debería hallar el punto de intersección de todas las hipérbolas generadas para conocer su posición.

2.1.4 Técnica RTOF:

Esta técnica se basa en la medida del tiempo total que transcurre desde que un nodo envía un paquete de datos al terminal móvil, hasta que dicho nodo recibe la respuesta. El terminal móvil debe responder rápidamente a un determinado tipo de mensaje, actuando así a modo de radar. De esta forma, es posible calcular la distancia entre ambos y se puede aplicar un algoritmo similar a los utilizados en las técnicas TOA para estimar la posición del terminal móvil, si bien ahora los requisitos de sincronización son menores. El problema principal de esta técnica es que el nodo no conoce el tiempo que tarda el terminal móvil en procesar del mensaje. Para distancias grandes, este tiempo puede ser despreciable, pero no para distancias cortas, por lo que la precisión de los sistemas RTOF es baja.

2.1.5 Técnica AOA:

Esta otra técnica estima la posición a partir de los ángulos relativos de llegada de la señal proveniente de diferentes terminales fijos o nodos. La estimación de la posición del objeto se calcula en función del ángulo que forma la línea que une el terminal móvil y cada nodo con un ángulo de referencia. Esta situación se muestra en la Ilustración 2-3, en la que se observa que en esta ocasión solamente es necesaria la presencia de dos nodos para estimar la posición del terminal móvil. La estimación de los ángulos θ_1 y θ_2 se puede realizar mediante antenas directivas o con arrays de antenas.

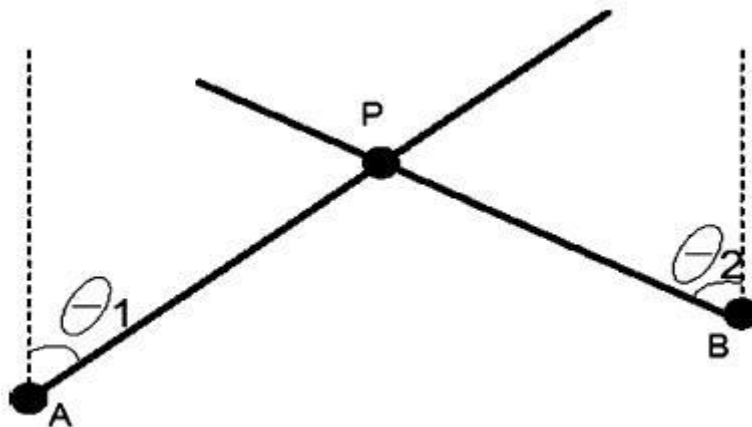


Ilustración 2-3: Método AOA

Esta técnica tiene dos ventajas importantes: son necesarios muy pocos nodos para estimar la posición del usuario, y no es necesaria ninguna sincronización entre los sistemas. La desventaja principal es que se requiere hardware costoso, a veces de gran tamaño, y que la precisión de la medida se reduce al alejarse el terminal móvil del nodo. De nuevo, dicha precisión se ve afectada por los efectos del multitrayecto y por los obstáculos presentes en la zona.

2.2 Análisis de escena

Esta técnica se basa en reconocer la escena recogiendo los parámetros de esta y luego ubica el objeto haciendo coincidir las mediciones del momento con las que tiene recogidas. El RSS basado en la ubicación es comúnmente utilizado en la técnica de análisis de escena. En el análisis de las escenas se utilizan técnicas basadas en el reconocimiento de patrones como pueden ser el método k-nn (K nearest neighbors), paradigmas de aprendizaje y procesamientos automáticos de redes neuronales artificiales, SVMs (Support Vector Machines) y SMP (smallest M-vertex poligon).

2.2.2 Método k-nn:

Es un método de clasificación supervisada que sirve para estimar la función de densidad $F(x / C_j)$ de las predictoras x por cada clase C_j . En el caso del análisis de escena este método se basa el valor RSS momentáneo del nodo, para buscar los k

valores o vecinos más cercanos a éste valor RSS, desde una base de datos previamente construida. De esta manera se conoce la ubicación aproximada. En este caso, k es el parámetro de adaptación para un mejor rendimiento, pero depende fundamentalmente de los datos.

2.2.3 Redes neuronales artificiales:

Las redes neuronales artificiales son un paradigma de aprendizaje y procesamiento automático que se modelan matemáticamente para un problema en concreto y posteriormente formulan una solución (programa) mediante un algoritmo codificado que tenga una serie de propiedades que permitan resolver dicho problema. Una vez programado el sistema para localizar objetos basándose en el RSS que reciben los nodos, se entren el sistema cogiendo como entradas las coordenadas de posición y los RSS. Después del entrenamiento de las redes neuronales, estas obtienen los pesos adecuados y cuando algún dispositivo móvil se encuentre en el área cubierta por estos nodos, se podrá obtener la posición del dispositivo en función al RSS que reciban los nodos y los pesos adquiridos por la red en la etapa de entrenamiento. Siendo la salida del sistema un vector de dos elementos para una posición estimada en 2-D, o un vector de tres elementos, para una posición estimada en 3-D.

2.2.4 SVM:

Los SVMs son un conjunto de algoritmos de aprendizaje supervisado para la clasificación y regresión de datos. Es una herramienta utilizada para el análisis estadístico y para el aprendizaje de las máquinas, y funciona muy bien en aplicaciones de clasificación y regresión. Los SVMs han sido muy utilizados para una amplia gama de aplicaciones en el mundo de la ciencia, medicina e ingeniería con un rendimiento empírico excelente. Se han utilizado con éxito varias clases de SVCs (Support Vector Classifier) y SVRs (Support Vector Regression) en la localización por identificadores digitales [22],[23].

2.2.5 SMP:

Son uno de los conjuntos de algoritmos más utilizados para estimar la posición, como se puede ver en este estudio [24]. En este caso, para cada punto de acceso hay M posiciones posibles, cuya distancia se establece en función al RSS momentáneo comparado con los que se habían almacenado en la base de datos. Se forman polígonos de M vértices para seleccionar al menos un candidato para cada punto de acceso. El polígono mas pequeño es el que tiene menor perímetro, y calculando el promedio de las coordenadas del polígono mas pequeño, se obtiene la estimación de la posición final. El SMP se ha utilizado en para la localización en interiores.

2.3 Proximidad

Esta es técnica que se basa en la proximidad. En este estudio en particular se hace referencia a la proximidad a la que detectan las antenas receptoras un dispositivo móvil. Se extiende una red de antenas cada una con una posición bien definida y cuando el dispositivo móvil es detectado por una sola antena se considera que está en el mismo emplazamiento. Cuando lo detectan más de una antena se considera que esta en el emplazamiento de la antena que reciba la señal más fuerte.

2.4 Técnicas de visión artificial

Las técnicas de visión artificial se pueden aplicar a infinidad de propósitos. Alguno de ellos es el de la detección, segmentación, localización y reconocimiento de ciertos objetos en imágenes. Dependiendo del tipo de objeto que se quiere detectar, el entorno en el que se encuentra, la cantidad de objetos que se quieren detectar y reconocer, y la finalidad del sistema, existen infinidad de algoritmos que se adaptan de diferente manera a cada una necesidades [25].

CAPÍTULO 3: SISTEMAS DE LOCALIZACIÓN

A continuación se revisará al estado del arte de los actuales sistemas de localización de objetos móviles. Hay diferentes enfoques para el estudio de los múltiples sistemas de posicionamiento, como las formas de propagación de las señales, las tecnologías de las redes inalámbricas, los algoritmos de posicionamiento, infraestructuras de sistemas existentes o dedicados.

3.1 Global Positioning System (GPS)

El GPS se utiliza para ubicar dispositivos móviles en cualquier punto de la tierra, con una precisión de pocos metros, 15 metros en el 95% del tiempo. En caso de que el sistema WAAS/EGNOS/MSAS esté activado, la precisión del sistema asciende de 1 a 2 metros. Para obtener la posición del receptor, éste debe de estar al aire libre y, a su vez, debe rastrear las señales de 4 o más satélites para que estos estimen su posición mediante la técnica de la trilateración. En entornos en los que existen obstáculos es más difícil rastrear las señales de los satélites y se crean efectos de múltiples reflexiones, los cuales introducen errores adicionales.

La Unión Europea está desarrollando el sistema de navegación por satélite llamado Galileo se espera poner en marcha en 2014 con el objeto de evitar la dependencia del GPS, y que será de uso civil. Estos satélites transmitirán señales con más potencia que los actuales, por lo que sus señales serán más fáciles de detectar y más resistentes a los efectos múltiples que degradan significativamente la precisión en entornos con obstáculos.

Sin embargo, la baja cobertura de la señal de satélite para entornos de interior disminuye su precisión y lo hace inadecuado para estimar la posición en interiores.

SnapTrack, una empresa de *Qualcomm*, es pionera en el GPS asistido mediante redes inalámbricas (A-GPS) para superar las limitaciones del GPS convencional. Es capaz de proporcionar a la técnica GPS una precisión de entre 5-50 m en la mayoría de los ambientes interiores. La tecnología A-GPS utiliza un servidor de localización con un receptor GPS, que es capaz de detectar simultáneamente los mismos satélites que el

teléfono móvil (o estación móvil), para ayudar al receptor GPS a encontrar señales débiles. El teléfono inalámbrico recoge las mediciones tanto de la constelación GPS como de la red móvil inalámbrica. Estas medidas se combinan con la ubicación del servidor para obtener la estimación de la posición.

Locata Corporation ha inventado una nueva tecnología de posicionamiento llamada *Locata* [26], para los posicionamientos precisos tanto en el interior y el exterior. Parte de la tecnología *Locata* se compone de un tiempo sincronismo con el tranceptor “pseudo-satellite” llamado *LocataLite*. Una red de *LocataLites* forma la *LocataNet*, que transmite las señales GPS, que permiten obtener una posición utilizando las mediciones de fase de portadora de un dispositivo móvil (un *Locata*). El grupo de navegación y posicionamiento por satélite (SNAP) de la Universidad de Nueva Gales del Sur ha colaborado en el desarrollo del sistema *Locata* y ha llevado a cabo pruebas con esta nueva tecnología. Los resultados de los experimentos demuestran que con la portadora de fase, es posible obtener la posición (sin radio módem de enlace de datos) con una precisión de centímetros.

En cuanto al mundo del deporte J. Barnes *et al.* [27] han realizado un estudio basado en el sistema de localización por satélite en el que realizan seguimientos a jugadores de la liga de fútbol australiano. Los jugadores que han participado en este estudio se colocan los dispositivos que reciben las señales de los satélites y van guardando la información de sus posiciones en cada momento para un posterior análisis de las estadísticas del jugador en cuestión. La precisión del sistema es muy baja, de un 5% sobre las bases del fabricante.

3.2 Radio Frequency Identification (RFID)

El sistema RFID se utiliza para almacenar y recuperar datos a través de la transmisión electromagnética de un circuito integrado mediante radiofrecuencia y cada vez más, se utiliza para mejorar los procesos de manipulación de datos. El propósito fundamental es de la tecnología RFID es la de transmitir la identidad de un objeto. Un sistema de RFID consta de lectores y etiquetas RFID, los cuales se comunican entre ellos. El lector RFID es capaz de leer los datos emitidos por las etiquetas. Los lectores y las etiquetas RFID

utilizan frecuencias definidas y un protocolo para la transmisión y recepción de datos, y las etiquetas pueden ser pasivas o activas.

Las etiquetas pasivas de RFID funcionan sin batería. Se utilizan principalmente para reemplazar la tecnología tradicional del código de barras y son mucho más ligeros, más pequeños en volumen, y menos costosas que las activas. La señal que les llega desde los lectores es suficiente para hacer funcionar el circuito integrado del tag, de forma que pueda generar y transmitir una respuesta. Sin embargo, el radio de acción es muy limitado, entre 1-2 m, y el costo de los lectores es relativamente alto. Estos sistemas pasivos suelen usar cuatro bandas de frecuencia: LF (125 kHz), HF (13,56 MHz), UHF (433, 868-915 MHz), y la frecuencia de microondas (2,45 GHz, 5,8 GHz).

Las etiquetas RFID activas son pequeños transceptores que funcionan con batería, necesaria para alimentar sus circuitos y propagar la señal al receptor. Son más fiables que las pasivas y su rango de acción abarca desde decenas hasta cientos de metros. Pero son más grandes, más caros y tienen una vida útil más corta, hasta 10 años. Los rangos de frecuencia utilizados son similares a los de las etiquetas pasivas.

Un sistema de localización que utiliza esta tecnología es *SpotOn* [28]. *SpotOn* utiliza un algoritmo para la detección de la posición en 3-D basado en el análisis de la potencia de las señales de radio. Los investigadores de *SpotOn* han diseñado y construido un hardware que sirve para la localización de objetos mediante las etiquetas. *SpotOn* plantea obtener la posición de los objetos mediante una red de sensores distribuida homogéneamente sin ningún control central, de manera ad hoc. Este sistema utiliza el valor RSS recibido de cada etiqueta para estimar la distancia entre etiquetas. Utiliza la densidad de las etiquetas y la correlación de múltiples mediciones para mejorar la exactitud y la precisión. Otro sistema que se conoce como *Landmarc* (posicionamiento de sensores en interiores utilizando la RFID activa) [29]. Este prototipo utiliza lectores de RFID que operan a 300MHz. Con el fin de aumentar la precisión, sin tener que colocar más lectores, el sistema utiliza referencias de etiquetas fijas adicionales para ayudar a la calibración de la localización. Estas etiquetas sirven como puntos de referencia en el sistema. El sistema *Landmarc* requiere la información de intensidad de la señal de cada etiqueta en los lectores. Después se utiliza el método k-nn para calcular

la posición de las etiquetas RFID. En el 50% de los casos tiene un error de alrededor de 1 m, mientras que los errores máximos en el sistema *Landmarc* son menores a los 2 m.

La tecnología *WITRACK*¹ desarrollada por Fraunhofer IIS [30], con el fin de proporcionar a los espectadores y los organismos de radiodifusión nuevos tipos de información, hace posible determinar con exactitud las posiciones no sólo de los jugadores, sino también de sus brazos y piernas en tiempo real. El sistema opera en la banda de 2,4 GHz y es capaz de calcular la posición de cada transmisor del terreno de juego hasta 200 veces por segundo, con una precisión de unos pocos centímetros. Se utilizan un amplio número de transmisores, por ejemplo en un partido de fútbol, los jugadores llevan 132 transmisores, mas 12 transmisores incrustados en balón) ofrecen una visión general completa del juego al instante.

El sistema se puede configurar con diferentes técnicas de localización para hallar las posiciones del objeto, pudiendo elegir entre RTOF y TDOA.

El sistema LPM (Local Position Measure) de fútbol en 3D desarrollado por *InMotion2*, es una manera de registrar los movimientos de todos los jugadores durante una sesión de entrenamiento. El desarrollo de LPM Soccer 3D comenzó en 2002 cuando se instaló la tecnología en el campo de fútbol del club PSV en Eindhoven. Los sistemas de seguimiento de *InMotion* se basan en un sistema de comunicación inalámbrica, siendo éste un sistema único para obtener las posiciones en tiempo real y las mediciones de la actividad del corazón. La red está conectada a las diferentes estaciones base que hay alrededor del campo. Los atletas usan un chaleco especial con un transpondedor (transmisor) que envía los datos de rendimiento en tiempo real a través de las antenas de las estaciones de base.

¹WITRACK, <http://www.iis.fraunhofer.de/en/bf/ln/anwendungen/sport/index.jsp>

² InMotion, <http://www.inmotio.nl/media/390/voetbaltrainer-june-2010---english-translation.pdf>

La precisión del sistema es de 5 cm y la frecuencia de medición es de hasta 1000 veces por segundo, con un porcentaje de error menor al 1% (según un informe de validación de la Universidad de Groningen). Las medidas de posición se muestran tanto en 2D como en 3D por medio de una aplicación software. El entrenador puede ver los datos de varios jugadores en diferentes posiciones en tiempo real en un Tablet PC y por lo tanto ser capaz de dirigir el entrenamiento del jugador o del equipo.

3.3 Sistema basado en comunicaciones móviles

Muchos sistemas utilizan el acceso múltiple por división de código (GSM / CDMA) en las redes para móviles, para calcular la ubicación de los clientes móviles al aire libre. Sin embargo, la precisión del método E-OTD, utilizando para localizar los terminales móviles mediante la técnica TDOA, es muy baja (50-200 m), dependiendo del tamaño de la celda. Por lo general, es mayor la precisión en las zonas con mayor densidad de antenas (por ejemplo, centros urbanos) y mucho menor en el entornos rurales [31].

La localización en interiores basada en la red de telefonía móvil es posible siempre y cuando el edificio este cubierto por una o varias estaciones base, que sean capaces de emitir señales lo suficientemente potentes para ser recibidas en el interior por los clientes móviles. Hay un sistema de localización en interiores [32], que se basa en el sistema GSM. El dispositivo móvil se conecta a las seis celdas GSM con mayor potencia y es capaz de leer 29 canales adicionales GSM, la mayoría de los cuales son lo suficientemente fuertes como para ser detectados, pero demasiado débiles para utilizarlos para una comunicación eficiente. El aumento de dimensiones introducido por los canales adicionales aumenta considerablemente la precisión de la localización. En los resultados de los experimentos realizados en tres edificios con varias plantas utilizando la técnica de k-nn, muestran que este sistema de localización en interiores es capaz de obtener la posición de los diferentes pisos con una precisión de 2,5 m. El mismo método podría ser aplicado en el IS-95 CDMA y la red móvil 3G.

3.4 UWB (Ultra Wave Band)

UWB es una tecnología de comunicación de la que existen diferentes variantes, adaptadas a diferentes aplicaciones. La más extendida, conocida como OFDM UWB es una implementación optimizada para comunicaciones de alta velocidad de transmisión, con un gran ancho de banda y con un rango de alcance corto. Básicamente, pretende ser un USB wireless para la transmisión rápida de ficheros, música, etc.

Sin embargo, a la hora de ofrecer soluciones de localización, es necesario realizar una adaptación específica que consiste en la emisión de pulsos muy cortos (normalmente, <1 ns), con un ciclo muy bajo (por lo general 1:1000). En el dominio de la frecuencia, el sistema, utiliza un gran ancho de banda (> 500 MHz de ancho). Los sistemas de localización basados en UWB tienen las siguientes ventajas [33][34]. A diferencia de los sistemas de RFID, que funcionan en las bandas individuales del espectro radioeléctrico, UWB transmite una señal a través de múltiples bandas de frecuencias al mismo tiempo, 3,1 a 10,6 GHz. Las señales UWB se transmiten por una duración tiempo mucho más corta que las utilizadas en las RFID convencionales. Las etiquetas UWB consumen menos energía que las etiquetas de radiofrecuencia y puede operar en una amplia zona del espectro radioeléctrico. Las señales UWB se pueden utilizar simultáneamente en los mismos entornos que las señales de radiofrecuencia, sin sufrir interferencias en los diferentes tipos de señales y en el espectro radioeléctrico utilizado. Estos pulsos de corta duración son fáciles de filtrar con el fin de determinar qué señales son las correctas y cuáles son las que se crean mediante las reflexiones del entorno. Al mismo tiempo, la señal pasa fácilmente a través de paredes, equipos y ropa. Sin embargo, los materiales metálicos y líquidos causan interferencias en las señales UWB. Existen soluciones comerciales de sistemas de localización en interiores que utilizan esta tecnología tal y como se explica a continuación.

A la hora de utilizar UWB para ofrecer soluciones de localización es preferible la utilización de pulsos cortos de transmisión en lugar de otras formas de onda más complejas. La razón es sencilla: la medida del TDOA de la señal será más precisa si la señal llega en su totalidad al mismo tiempo. Dicho de otro modo, es necesario que la señal sea corta en el tiempo.

Además se precisa que el ciclo de trabajo sea bajo, y por tanto los pulsos son cortos pero el espacio entre ellos es largo. Una vez más la razón es sencilla: es mejor para el receptor no recibir muchos pulsos simultáneamente.

Existen diferentes formas de implementar transmisores de pulso corto con ciclo de trabajo corto, desde soluciones sencillas a base de diodos de conmutación rápida, a sintetizadores digitales de pulsos con una sincronización precisa, dependiendo de las necesidades concretas de la aplicación [35]. Del mismo modo las implementaciones de los receptores UWB, que son el elemento clave del sistema, pueden variar en función de los requisitos de prestaciones, coste, consumo de energía y tamaño de cada aplicación.

Los receptores deben tener una muy buena sincronización para lograr una precisión muy alta en interiores (hasta 15 cm). La localización de la etiqueta en 2-D y en 3-D se realiza en tiempo real de mediante el uso de dos diferentes soluciones que combinan TDOA con AOA. La ventaja de utilizar ambas soluciones conjuntamente, es que la ubicación se puede determinar a partir de sólo dos sensores, reduciendo la densidad de sensores necesarios sobre los sistemas que sólo utilizan TDOA.

Hasta la fecha, se han desarrollado varios sistemas de localización comerciales basados en UWB. El sistema desarrollado por *Ubisense*³, trabaja con UWB unidireccional con un control de canal bidireccional por división de tiempo de múltiple acceso (TDMA). El sistema de posicionamiento *Ubisense*, utiliza tanto la técnica TDOA como la AOA, para obtener las posiciones de las etiquetas. Las etiquetas transmiten señales a los receptores UWB que se encuentran conectados en red. Cada célula requiere un mínimo de cuatro receptores. En el interior de los edificios o conjuntos de edificios, se pueden conectar en red un número ilimitado de lectores. Los lectores recibirán los datos de las etiquetas,

³ Ubisense, <http://www.ubisense.net/en>

desde una distancia de 45 metros, y los enviarán a través de la plataforma *Ubisense Smart Space*.

Ubisense ha desarrollado un sistema que es capaz de obtener la posición de pequeñas etiquetas que emiten pulsos de radio UWB de hasta 15 cm en el 95% de los casos en los que no hay obstáculos en las líneas de visión entre las etiquetas y receptores. Esta precisión se degrada hasta cerca de 1 metro cuando algunas de las señales han traspasado paredes. Sincronizando el sistema UWB, del cual obtiene la posición de la etiqueta, con el sistema de RFID, se obtiene la identificación única de cada etiqueta localizada.

La compañía *TimeDomain* ha desarrollado un sistema UWB posicionamiento similar al de *Ubisense* llamado *PLUS*⁴. *PLUS* es un sistema de localización en tiempo real de alta precisión que puede proporcionar información de localización en interiores con una precisión inferior a un metro. Para ello también utiliza las técnicas de TDOA y AOA, pero no llega a obtener la posición 3D de la etiqueta, solo obtiene las posiciones (x, y) y el número de planta del edificio donde se encuentra esta.

3.5 WLAN (IEEE 802.11)

Una WLAN es un sistema de comunicaciones de datos que transmite y recibe datos utilizando ondas electromagnéticas, en lugar del par trenzado, coaxial o fibra óptica utilizados en las LAN convencionales, proporcionando conectividad inalámbrica dentro de un área de cobertura.

Con velocidades de transmisión de 11, 54 o 108 Mbps y un rango de entre 50-100 m, IEEE 802.11 es actualmente el dominante estándar de redes locales inalámbricas.

⁴ PLUS, <http://www.plus-ls.com>

Existen distintos métodos para realizar una localización mediante una red inalámbrica, con los vectores de potencia, triangulación de la potencia y métodos heurísticos.

El método de vectores de potencia se basa en realizar una primera fase de entrenamiento o calibración, dividiendo la zona de interés en celdas y realizar medidas en todas ellas, para obtener la información de la señal de los puntos de acceso en cada celda. Dicha información se almacena en una base de datos en forma de vector, donde cada celda conoce la potencia que le llega al usuario desde cada punto de acceso a su posición. Durante la operación normal del sistema, conociendo la señal que recibe el objeto a localizar desde cada punto de acceso, y consultando la base de datos, puede deducirse su posición.

Mediante la triangulación, el dispositivo mide la potencia recibida de tres puntos de acceso y realiza los cálculos necesarios para obtener su posición.

Los métodos heurísticos se pueden complementar a otros métodos de localización para mejorar la precisión de éstos, o bien pueden calcular la posición del dispositivo por sí solos. Estos métodos pueden estar basados en técnicas de proximidad, métodos k-nn y redes neuronales entre otros. La precisión de estas mediciones es de entre 1 y 5 metros.

Place Lab [36] es un software libre destinado a aplicaciones de cálculo de la posición de un usuario. Está capacitado para obtener la localización del cliente tanto en interiores como en exteriores siempre que tenga cobertura Wireless. El sistema se caracteriza porque los clientes pueden determinar su ubicación sin necesidad de interactuar constantemente con un servicio central, es decir, el cálculo de la posición se realiza desde el propio dispositivo.

El sistema se basa en la posición de los puntos de acceso para calcular la posición del usuario. Los dispositivos que utilizan *Place Lab* sólo necesitan interactuar con los puntos de acceso con el fin de adquirir sus identificadores de los cuales las conocen las coordenadas. Los dispositivos son capaces de posicionarse basándose en la monitorización pasiva del entorno de localización. Su precisión teórica es de entre 15 y 20 metros. La cobertura y la exactitud de *Place Lab* dependen del número y del tipo de punto de acceso que hay alrededor del dispositivo cliente.

Permite la localización de los dispositivos mediante Wi-Fi (balizas 802.11 o puntos de acceso), torres GSM de teléfonos móviles, dispositivos fijos de Bluetooth y GPS.

El software desarrollado por la *Ekahau*⁵ se basa en la calibración de la fuerza de señal y código de identificación que emiten las radiobalizas de la red Wireless. Funciona tanto en interior como en exterior siempre que haya cobertura Wireless.

Se trata de un sistema que es compatible tanto con los dispositivos pasivos de *Ekahau* (etiquetas o Tags), como con dispositivos activos (ordenadores portátiles, PDA's o cualquier elemento tecnológico con conectividad Wireless). Goza de una precisión de hasta 1 metro en las mejores condiciones de la red y del estado de los puntos de acceso, lo que permite localizar muchos dispositivos a la vez y sobre el mismo mapa de situación, ya sean estos dispositivos activos o pasivos. La información de posición incluye las coordenadas (x, y), el edificio, el número de planta y la habitación, y funciona sobre cualquier estándar 802.11.

Ekahau utiliza el método del vector de potencia para calcular la posición del usuario a partir de la potencia de los puntos de acceso, además de una heurística de movimiento para mejorar la precisión del sistema.

Herecast [37], es sistema que proporciona servicios de posición basados en Wi-Fi. La idea básica del cálculo de la posición es que *Herecast* crea una base de datos donde se guarda la información para cada punto de acceso. Cada vez que el objeto se acerca a un punto de acceso desconocido se guardan sus datos (país, ciudad, provincia, edificio...) en la base de datos, asociándole una posición. La posición que estima el sistema no son unas coordenadas concretas sino un área y unos datos que el usuario define en el entrenamiento y se guardan en la base de datos.

Para calcular la posición, *Herecast* utiliza un sistema de nombramiento simbólico para reconocer un punto de acceso cercano, y mediante una heurística de proximidad, como método principal de localización, sabemos la posición en la que nos encontramos.

⁵ Ekahau, <http://www.ekahau.com>

El sistema RADAR [38] propone una localización de usuarios en edificios, basado en el método k-nn. Los autores proponen dos enfoques diferentes para determinar la ubicación del usuario. La primera se basa en la medición empírica de la intensidad de la señal del punto de acceso en una fase de entrenamiento. En estos experimentos se especifica que a la hora de intentar obtener la posición del usuario, las orientaciones de usuario, el número de vecinos más cercanos utilizados, el número de puntos de datos, y el número de muestras recogidas en tiempo real, pueden afectar la exactitud del sistema de localización. El segundo, es un modelado de propagación de la señal. Utiliza un modelo de propagación con el factor de atenuación de las paredes (WAF) y de los suelos (FAF), en lugar de los modelos de dispersión *Rayleigh* y el de distribución *Rician*, que se utilizan en escenarios exteriores. WAF tiene en cuenta el número de paredes (obstrucciones). La precisión del sistema de RADAR es de unos 3.2 m, pero en su siguiente trabajo [39], han mejorado la precisión mediante el algoritmo de Viterbi. Mejorando el resultado hasta un rango de entre 2,37 y 2,65 metros.

El sistema *Horus* [40], que ofrece una técnica de agrupación conjunta para la estimación de la posición, utiliza un método probabilístico. Cada coordenada candidata a la posición, es considerada como una clase o categoría. Los resultados del experimento muestran que esta técnica puede adquirir una precisión de más del 90% dentro de 2,1 m. Aumentar el número de muestras en cada punto de muestreo podría mejorar su exactitud, porque el aumento del número de muestras podría mejorar la estimación de medias y desviaciones estándar de la distribución de Gauss.

Haeberlen et al. [41] presentaron un método Bayesiano para la localización topológica sobre la totalidad de una red 802.11, desplegada dentro de un edificio de oficinas de varios pisos. Ellos han demostrado que el uso de un modelo topológico puede reducir drásticamente el tiempo necesario para entrenar al sistema localizador, mientras que la precisión resultante es aún suficiente para muchas aplicaciones de localización.

Aunque la mayoría de los sistemas basados en WLAN utilizan la potencia de la señal, *AeroScout* (antes *Bluesoft*) [42] se basa en el método TDOA. Se necesita recibir la misma señal de radio en tres o más puntos, mucha precisión en el tiempo (nanosegundos) y realizar un proceso utilizando el algoritmo de TDOA para determinar la ubicación.

3.6 Bluetooth (IEEE 802.15)

Bluetooth opera en la banda ISM a 2.4 GHz. En comparación con WLAN, la velocidad de transmisión es menor (1 Mbps), y el rango es más corto (10-15 m). Por otro lado, es un estándar muy "ligero", con una alta ubicuidad (integrado en la mayoría de los teléfonos, asistentes digitales personales (PDAs), etc. y es compatible con varios servicios de red, además de otros de IP. Las etiquetas de Bluetooth son transceptores de pequeño tamaño. Como cualquier otro dispositivo, cada etiqueta tiene un identificador único, y éste puede ser utilizado para la localización de la etiqueta.

La solución de posicionamiento *Topaz*⁶ está basada en la infraestructura y en los accesorios de Bluetooth de *Tadlys*. Esta solución de posicionamiento modular se compone de tres elementos: el servidor de posicionamiento, los puntos de acceso inalámbricos y las etiquetas inalámbricas. El funcionamiento del sistema lo hace adecuado para el seguimiento de seres humanos y objetos. Este sistema proporciona una exactitud de 2 metros en un 95% de los casos. El retardo del posicionamiento es de 15-30 s. Este sistema obtiene la posición del objeto triangulando las señales RSSI obtenidas en los puntos de acceso más cercanos. El rendimiento es aún mejor en el nuevo sistema *Topaz*, el cual integra sensores de infrarrojos y otros transductores, con el posicionamiento y las capacidades de comunicación de Bluetooth.

Antti et al. presentan el diseño e implementación de una aplicación Bluetooth para localización local (BLPA) [43]. En primer lugar, convierten los niveles de la señal recibida para poder calcular las distancias de acuerdo con un modelo de propagación simple, y usan el filtro de Kalman extendido (EKF) para calcular la posición en 3-D,

⁶ TOPAZ: <http://www.tadlys.com>

sobre los cálculos de distancias anteriormente realizados. La precisión de BLPA es de 3,76 m.

3.7 ZigBee

ZigBee es el nombre de la especificación de un conjunto de protocolos de alto nivel de comunicación inalámbrica para su utilización con radios digitales de bajo consumo, basada en el estándar IEEE 802.15.4 de redes inalámbricas de área personal (Wireless Personal Area Network, WPAN). Su objetivo son las aplicaciones que requieren comunicaciones seguras con baja tasa de envío de datos y maximización de la vida útil de sus baterías.

Por un lado están los algoritmos de localización basados en RSS. Este parámetro se puede medir en cada uno de los paquetes recibidos (tanto el nivel RSSI como el instante en el que se recibió el paquete son accesibles para las capas MAC, NWK y APL). La accesibilidad al RSSI implica que los nodos que forman la red pueden realizar estimaciones de localización sin necesidad de hardware adicional.

El modo más simple de obtener la localización de un nodo es solicitar que envíe una señal y considerar que la localización de dicho nodo, será la localización del nodo de referencia que reporte un nivel RSSI superior. La ventaja de este método es que puede ser implementado fácilmente en nodos con pocas prestaciones, pero la precisión de la estimación realizada puede resultar inadecuada para muchas aplicaciones.

Otro método de localización basado en RSSI es realizar la estimación empleando trilateración. Los errores en la medición de la RSSI pueden resultar en una precisión inaceptable de la estimación. Esto sucedería por ejemplo en entornos en los que se produzcan muchas señales multitrayecto, como sucede con la mayoría de los entornos de interior.

En cuanto a los algoritmos basados en tiempo (TOA, TDOA) la sincronización es crítica. Cuanto más rápido viaje la señal en el aire, el error de tiempo será más significativo, es decir, la sincronización de reloj será más crítica si se emplean señales RF en lugar de ultrasonidos. Además, las señales de gran ancho de banda, como UWB, mejoran la estimación de localización significativamente frente a la estimación realizada

con señales IEEE 802.15.4. Para entornos de interior y empleando redes *ZigBee*, las estimaciones de localización basadas en tiempo con señales de radio frecuencia son menos probables que las basadas en RSSI, ya que el pequeño ancho de banda de las señales IEEE 802.15.4, si las comparamos con UWB, podría no ser suficiente para mejorar el error de TOA.

En [44] comparan técnicas convencionales para conocer la localización de objetos móviles usando módulos *ZigBee*. La información de localización se infiere de la RSS de la señal recibida y han clasificado los mecanismos en dos categorías. Técnicas de localización basadas en mapas y localización basada en distancias.

La intensidad del RSS de la señal es inestable, ya que en el caso de *ZigBee*, la señal puede verse afectada por la presencia de redes Bluetooth, WLAN que funcionen en la banda de 2,4GHz. En el paper comparan la precisión de los dos mecanismos.

La primera de las técnicas planteadas, asume que todos los nodos menos uno están fijos y su posición es conocida. Se divide el área que se quiere analizar en celdas y se realiza una medición previa de los valores de RSS recibidos en cada celda, lo que se denomina *fingerprint map*. Para poder identificar el emisor cada dispositivo tiene un ID, y del mismo modo, cada paquete posee un índice para distinguir la secuencia. El nodo móvil, recibe el RSS, el identificador del emisor, y el índice del paquete y envía la información a la estación base. Ésta deduce la localización evaluando la menor diferencia posible con el *fingerprint map* almacenado. La mayor limitación de los algoritmos de *fingerprint* es que requieren que el espacio radioeléctrico sea lo más estable posible, si la naturaleza del entorno es dinámico no es posible aplicarlo.

El otro caso es la localización basada en distancias empleando Cadenas de Markov para deducir la localización. Se aplica trilateración para inferir la distancia respecto a nodos fijos. Se plantea la ecuación de relación entre distancias reales y los niveles RSS recibidos desde los nodos fijos. De forma similar al caso de *fingerprint*, cada nodo móvil, situándose en puntos concretos recibe señales desde los nodos fijos y los indexa. Definen una variable, Bel (belief), que es la posibilidad de que un objeto móvil esté en una posición en el momento t . Lo primero que hacen es inferir las distancia de los nodos fijos al nodo móvil mediante la ecuación de relación y empleando trilateración. A

medida que pasa el tiempo, la variable Bel se actualiza de acuerdo a las Cadenas de Markov, y se emplea para ir actualizando la posición del nodo móvil.

Tras realizar todas las mediciones han comprobado que con el modelo de localización basado en distancias se reduce el error de estimación de posición. Además han podido comprobar como el RSSI cambia dinámicamente dependiendo de los obstáculos, del lugar, y cómo la dirección de la antena afecta al nivel de señal recibido.

3.8 Ultrasonidos

Los sistemas de posicionamiento por ultrasonidos son capaces de estimar la posición de los objetos en tres dimensiones, con una precisión del orden de centímetros.

A día de hoy, ya existen diferentes publicaciones que tratan el tema de localización por ultrasonidos. Mencionamos algunos sistemas implementados de los cuales tenemos constancia.

El método de *The Bat System* [45] se basa en la localización según el modelo TOA dentro de una red de nodos fijos instalados en el techo que actúan como receptores, los cuales se encargan de calcular el tiempo de vuelo de la señal emitida por los nodos móviles y enviar dicha información a una unidad central. El nodo central se implementa en un PC (lo cual le permite realizar cálculos complejos) y se comunica por un protocolo serie con los nodos fijos y por radiofrecuencia con los nodos móviles. Cumple las funciones de coordinador, encargándose de gestionar qué nodo móvil debe emitir la señal, enviándole la activación (señal con el identificador del nodo que debe comenzar a emitir el tren de pulsos de ultrasonidos) mientras que por el protocolo serie indica a las balizas que inicialicen los contadores. Tras dejar pasar cierto tiempo, hace ‘polling’ (pregunta a las balizas si han recibido la señal y el tiempo transcurrido desde la activación hasta que recibieron la señal de ultrasonidos), triangula la posición del nodo móvil y envía dicha posición al nodo móvil.

Por otro lado, el *CRICKET* [46] desarrollado por Nissanka B. Priyantha et al. se basa en la localización según el modelo TOA, pero, a diferencia con el método *The Bat System*, los nodos móviles son los encargados de actuar como nodos receptores y la red de nodos fijos son los encargados de emitir las diferentes señales de ultrasonidos. Cabe

mencionar que sigue existiendo un nodo central, al cual se pueden conectar los nodos móviles y el cual se encarga de triangular la posición. Los nodos sólo envían las medidas de la distancia a cada baliza por un puerto RS232. El nodo central tiene implementado un sistema de mapas con la posición exacta de las balizas, lo cual unido a las distancias calculadas a cada una de ellas por el nodo móvil, le permite implementar algoritmos de localización, permitiendo triangular la posición del nodo móvil.

El sistema está pensado para proveer privacidad a los diferentes nodos del sistema. Cada baliza emite el trigger por radiofrecuencia junto con la información de la habitación en la que se encuentra la baliza y su ID. Las balizas no están coordinadas entre sí, por lo que pueden existir colisiones. En caso de producirse colisiones, las balizas implicadas esperan un tiempo aleatorio antes de volver a emitir su señal.

El nivel de lógica de los nodos receptores es bastante reducido, implementado sólo un mecanismo para conocer la baliza más próxima al nodo (conociendo así en qué habitación se encuentra el nodo).

El sistema de localización 3-D en interiores [47] que utiliza una estación base compacta tiene como novedad frente a los sistemas anteriormente descritos es la posición de los nodos emisores. Estos, en vez de encontrarse distribuidos por la habitación, se encuentran concentrados en una única baliza (con tres emisores). Este hecho repercute negativamente en la precisión del sistema, pero facilita la instalación del mismo, la generación de las coordenadas (permite establecer el origen de coordenadas en la propia baliza) y tiene un coste menor por estación emisora. Como en los casos anteriores, el cálculo de la posición se realizaba fuera de los nodos receptores, en un PC.

Cabe destacar, que la emisión de la señal de ultrasonidos es simultánea para los tres emisores, utilizando un código CDMA para lograr diferenciar la señal de cada uno de los emisores.

El sistema *ISENSE*⁷ es un modelo comercial desarrollado por la compañía *Intersense*. Utiliza una tecnología híbrida para conseguir una localización precisa, combinando la localización mediante ultrasonidos con un sistema de acelerómetros. El sistema despliega una red de emisores de ultrasonidos en el techo de la habitación y utiliza la localización del nodo aplicando el modelo TOA. Para minimizar error se ayuda de un filtro Kalman y la información obtenida de los acelerómetros.

Cabe destacar que tanto los nodos como la red de emisores están conectados a un nodo central, el cual se encarga de coordinar los nodos emisores y calcular la posición de los nodos. Además, cabe resaltar que el coste del conjunto es de 9.950 dólares si el sistema es cableado y de 36.600 dólares si es inalámbrico.

Este sistema [48] ha sido desarrollado por McCarthy, M.R. Muller, H.L. Su principal característica es que prescinde del sistema de radiofrecuencia para conocer el tiempo exacto de emisión de la onda de ultrasonidos. En su lugar, las balizas emiten en un orden secuencial, dejando entre cada emisión un tiempo de guarda. Utilizando la ecuación de la esfera, nos permite formular una ecuación con cinco variables aleatorias (las tres coordenadas, el instante de emisión de la señal de ultrasonidos y el periodo de emisión entre las balizas). Para poder despejar estas incógnitas se utiliza un filtro de Kalman para linealizar la ecuación antes de calcular la posición del receptor.

Cabe destacar que todos los cálculos necesarios para obtener la posición del receptor se realizan en un PC de baja potencia.

⁷ ISENSE, www.isense.com

3.9 Infrarrojos (IR)

La tecnología de la radiación inalámbrica infrarroja (IR) se utiliza en los dispositivos o sistemas que transmiten datos a través de luz infrarroja. Las compañías *Radianse*⁸ y *Versus*⁹, utilizan una combinación de señales de radiofrecuencia con infrarrojos para obtener la posición del objeto. Sus etiquetas emiten señales de infrarrojos y de radiofrecuencia que contienen un identificador único para cada persona o activo, al que se realiza el seguimiento. El uso de RF permite el posicionamiento más general (por ejemplo, la planta), mientras que las señales de IR proporcionan una resolución adicional (por ejemplo, la habitación). Si la señal infrarroja transmitida por un dispositivo se recibe en otro, es muy probable que se encuentren en la misma habitación, porque la luz infrarroja no traspasa las paredes. Además la luz infrarroja se refleja en las paredes por lo tanto la línea de visión entre dos dispositivos no debe ser necesariamente directa.

El sistema basado en tarjetas activas [49], usó una red de receptores infrarrojos para detectar unas insignias, cada uno de estas emite regularmente un único código de infrarrojos. Los empleados del laboratorio de investigación de Olivetti, en el que se implementó el sistema, utilizaban estas insignias. La aplicación inicial del sistema era ayudar a los recepcionistas, para que estos fueran capaces de enrutar las llamadas telefónicas entrantes a las habitaciones a las cuales se encontraban sus destinatarios. Tras un período de prueba de dos semanas, se observó que se había reducido considerablemente el número de llamadas que no llegaban a sus destinatarios.

⁸ Radianse, <http://www.radianse.com>

⁹ Versus, <http://www.versustech.com>

3.10 Visión

Se han desarrollado una amplia gama de sistemas de localización basados en cámaras que se utilizan para *trackear* objetos o personas. El seguimiento de personas con varias cámaras es un problema bien documentado, utilizando técnicas de sustracción de fondo y reconocimiento de regiones [50]. El proyecto *EasyLiving* utiliza este tipo de algoritmos, junto con dos cámaras estéreo (es decir, cuatro puntos de vista en total) a fin de rastrear a los usuarios en un entorno inteligente, con una precisión de alrededor de 10 cm en el plano horizontal [51]. Esto se logra sin necesidad de que el usuario se coloque ningún dispositivo de rastreo o marcador visual. Aunque esto sería muy beneficioso para la identificación de cada usuario, sin ello resulta muy difícil identificarlos. Este sistema mantiene los histogramas de color de cada usuario y para conseguir identificarlos.

A los usuarios previamente identificados por el sistema se les asigna un identificador para cuando estos salgan y vuelvan a entrar en el escenario. Este no funciona muy bien cuando existen oclusiones parciales o totales y cuando hay cambios bruscos de iluminación.

Los sistemas basados en cámaras de posicionamiento abarcan un amplio rango de precisiones y costos de implementación. Existen sistemas de captura de movimiento muy caros, en los que se utilizan varias cámaras para seguir pequeños marcadores pasivos o activos. Los sistemas comerciales de captura de movimiento como Vicon MX¹⁰ pueden seguir marcadores pasivos con una precisión de milímetros a frecuencias superiores a los 500 Hz. Los marcadores están diseñados para ser fáciles de detectar, sin embargo, la identificación de marcadores no es un problema fácil de resolver.

¹⁰ Vicon MX, <http://www.vicon.com>

El sistema de captura de movimiento comercial *IMPULSO*, desarrollado por *Phasespace*¹¹, usa marcadores LED activos, situados en diferentes puntos de cuerpo humano, los cuales emiten señales en frecuencias moduladas únicas, para permitir su identificación y capturar todos sus movimientos.

En el campo del deporte también se han desarrollado muchos sistemas de tracking de jugadores basados en la visión. Por ejemplo, hay sistemas para el fútbol que se basan en imágenes de televisión obtenidas desde una sola cámara [52][53]. Una vez realizado el tracking, utilizan algoritmos de posicionamiento basándose en la posición y parámetros de la cámara para ubicar a cada jugador en el campo. No llegan a identificar cada jugador, diferencian los jugadores de cada “equipo A”, “equipo B” y el árbitro.

El sistema *ASPOGAMO*[54][55], utiliza imágenes de cámaras *broadcast* en las que cada imagen de la secuencia se procesa en tres etapas: la segmentación de las regiones, que posiblemente contienen los jugadores, la localización del jugador dentro de estas regiones y el mapeo de las posiciones de los jugadores en la imagen. Para estimar estas posiciones en el mundo real, se requieren los parámetros de la cámara, además de las posiciones en la imagen. Además, este sistema es capaz de presentar diferentes estadísticas de los jugadores que han intervenido en este proceso.

En otro artículo [56], lo primero que hacen es segmentar el campo y sus líneas, para posteriormente segmentar los jugadores y el balón. Se realiza el tracking del balón y los jugadores mediante el análisis de su contorno, teniendo en cuenta la información del fotograma anterior para analizarlo en el actual. La segmentación precisa de las líneas, es la base que permite encontrar en la imagen un modelo de transformación bueno, que es esencial para obtener la posición real de los jugadores y el balón en el campo.

¹¹ Phasespace, <http://www.phasespace.com>

En el artículo presentado por GuoJun Liu et al. [57] utilizan una técnica para hacer un seguimiento a un patinador basándose en el histograma de color y aplicándole a continuación un filtro de Kalman para obtener un buen tracking. Su posición la obtienen referenciando los parámetros de la imagen a 14 bloques fijos que se encuentran en la escena, en posiciones conocidas.

En cuanto a sistemas de localización que utilizan varias cámaras, se han convertido en actores importantes en aplicaciones de radiodifusión televisiva de eventos deportivos, sobre todo porque permiten superar el problema de la oclusión, que es uno de los retos más importantes a los que se que enfrentan los sistemas de localización y seguimiento. De esta manera no es necesario que los jugadores lleven sensores dedicados, para poder obtener información acerca de sus movimientos.

En estos artículos orientados al mundo del fútbol [17][58] se describen técnicas de homografía basadas en los parámetros de las cámaras para el pre-procesado, aunque en el segundo de ellos también se realiza la sustracción de fondo y extracción de la región donde se encuentra el jugador. A continuación el primero de ellos utiliza descriptores locales que se extraen a partir de los puntos de interés de los objetivos a seguir, para poder utilizar imágenes de cámaras en movimiento. El sistema guarda estas distancias espaciales, a partir de una secuencia de entrenamiento de corta duración, creando un modelo de distribución por puntos, que permitirá realizar el seguimiento del jugador. Y por último, utiliza un filtro de kalman para realizar la fusión de la información de las diferentes cámaras. El segundo artículo en cambio, se hace un seguimiento del deportista en la región que se ha extraído anteriormente y se extrae el color y el movimiento del jugador con cada cámara para su posterior fusión.

Hay otros dos estudios en los que los deportistas a seguir son jugadores de baloncesto. En el primero de ellos [34], se lleva a cabo una sustracción de fondo básica y una homografía con los parámetros de la cámara. A continuación se realiza una proyección de la escena, tras haber extraído el fondo de esta, al mapa de ocupación. Y por último, se realiza el tracking del jugador definiendo las aéreas en función a la velocidad máxima de jugador. En [35], también se lleva a cabo una sustracción del fondo y su posterior proyección de escena al mapa de ocupación. En el momento de hacer el tracking de la

jugadora, también se le realiza una segmentación de la camiseta para conseguir reconocer el número que lleva, llegando a reconocer la jugadora en el 73% de los casos. Recientemente ha salido al mercado un interesante sensor, comercialmente llamado *kinect*. Es un dispositivo desarrollado por Microsoft para la videoconsola Xbox 360, el cual permite a los usuarios controlar e interactuar con la consola sin necesidad de tener contacto físico con un controlador de videojuegos tradicional, mediante una interfaz natural de usuario que reconoce gestos, comandos de voz, y objetos e imágenes. Este dispositivo dispone de una cámara de video, un emisor IR y un receptor IR. Con la combinación de información recibida mediante estos sensores es capaz de posicionar a individuos dentro de su campo de visión. Con el sensor de IR crea un mapa de profundidad de la escena que combina con la imagen captada por la cámara de video y obtiene la información de posición. Además utiliza técnicas de reconocimiento facial para la identificación automática de los usuarios.

CAPÍTULO 4: PARÁMETROS DE EVALUACIÓN

Con el fin de evaluar las diferentes características de las técnicas de posicionamiento en interiores, se ha realizado una evaluación comparativa para determinar cuál de estas técnicas sería la óptima para nuestros objetivos. Se han tomado en cuenta para la evaluación de la precisión, complejidad y costo de las técnicas.

4.1 Precisión

La precisión (o error de localización) suele ser el requisito más importante en los sistemas de posicionamiento. Por lo general, se toma como métrica de rendimiento el error de distancia media, que viene dado por la media de la distancia entre la posición estimada y la real. A menudo se tiene en cuenta esta única característica para evaluar un sistema. Cuanto mayor es la precisión mejor es el sistema, pero suele haber un equilibrio entre la precisión y otras medidas. Por ello, es necesario medir en otras características del sistema para buscar ese equilibrio necesario.

4.2 Requerimientos

La complejidad de un sistema de posicionamiento puede depender de varios factores como pueden ser el hardware, el software, y la latencia del sistema. En el caso del hardware, la complejidad vendría dada por el montaje físico del sistema, en caso del software, por el cálculo del algoritmo de posicionamiento. Y por último, la latencia, que será la suma de retardos temporales desde que un blanco móvil se mueve a una nueva ubicación hasta que la información de esa nueva ubicación es calculada por el sistema.

Básicamente, cuanto menor sea la cantidad de dispositivos que hagan falta, sean estos lo menos intrusivos posible y menor sea su consumo, mejor será el hardware de este sistema. En cuanto al software, la complejidad del sistema viene dada por el tiempo de computación necesario para el cálculo del algoritmo de posicionamiento. Si el cálculo del algoritmo de posicionamiento se realiza en un servidor centralizado, el posicionamiento se puede calcular rápidamente debido a la gran capacidad de

procesamiento. Si se lleva a cabo en una unidad móvil, el tiempo de procesamiento puede aumentar. La mayoría de las unidades móviles carecen de una alta capacidad de procesamiento y de una batería de mucha capacidad. Este es un factor que influye en la latencia del sistema, así que, se prefieren los algoritmos de posicionamiento de baja complejidad. Por último, quedaría la latencia, la cual, en éste sistema tiene que ser mínima, lo más cercano posible al tiempo real. Sería el requisito más importante junto con la precisión, porque dependiendo de este requisito se podrá optar a conseguir un seguimiento inmediato del objeto, que de no ser así, podríamos perder el objeto del encuadre de la cámara.

4.3 Coste

El coste de un sistema de posicionamiento puede depender de muchos factores. Los factores importantes incluyen dinero, tiempo, espacio, peso y energía. El factor tiempo está relacionado con la instalación y el mantenimiento. Las unidades móviles pueden tener espacio reducido y limitaciones de peso. En cuanto a la energía, algunas unidades móviles (etiquetas pasivas de RFID) son completamente pasivas. Estas unidades sólo responden a señales externas y, por tanto, podrían tener una vida útil ilimitada. Otras unidades móviles (por ejemplo, dispositivos con batería recargable) tienen una duración limitada, que puede llegar a varias horas, sin ser recargadas.

4.4 Solución escogida

Se ha realizado un comparativa con las tecnologías más interesantes y los parámetros de evaluación que mejor se ajustan al objetivo de este estudio, mostrándolas en la Tabla 4-1.

Sistema	Técnica	Algoritmo	Precisión	Complejidad	Coste
Microsoft RADAR	WLAN, RSS	k-nn, Viterbi-like	5 m 2,37 – 2,65 m	Media	Bajo
Horus	WLAN, RSS	Métodos Probabilísticos	2 m	Media	Bajo
EKAHAU	WLAN, RSSI	Métodos Probabilísticos	1 m	Media	Bajo
HERECAST	WLAN	Heurística de proximidad	Depende del tamaño de celda	Media	Bajo
SnapTrack	GPS Asistido, TDOA	Combinación de mediciones	5 – 50 m	Alta	Media
UBISENSE	UWB, TDOA, AOA	Mínimos cuadrados	15 cm	< 1 s	Alto
SpotON	RFID, RSS	Ad-hoc Lateración	Depende del tamaño de celda	Medio	Bajo
Landmarc	RFID, RSS	k-nn	< 2 m	Medio	Bajo
TOPAZ	Bluetooth (RSSI)	Triangulación	2 m	15 – 30 s	Medio
MultiLoc [24]	WLAN (RSS)	SMP	2,7 m	1 s	Medio
GSM- Fingerprinting [32]	Redes Móviles GSM (RSS)	k-nn	5 m	Medio	Medio
Witrack	UWB, RTT, TDOA	Triangulación	8 cm	1/200 s	Alto
InMotion	RFID	LPM	5 cm	1/1000 s	Alto
Visión	Procesamiento digital de imágenes	Filtrado Detección Procesado	Milimétrica	Framerate	Bajo
kinect	Procesamiento digital de imágenes, IR	Disparidad entre cada píxel	Milimétrica	Framerate	Bajo

Tabla 4-1: *Tabla comparativa de los diferentes sistemas de localización*

Tras realizar la comparativa con las tecnologías más interesantes se ha observado que dos de los métodos de localización más importantes y ampliamente utilizados en interiores son los de localización con (múltiples) cámaras de video y localización basado en radio. Se basan en dos tecnologías diferentes y cada método tiene sus ventajas y desventajas. Las personas presentes en la escena pueden ocluir parcialmente o totalmente a la persona a la que se le está realizando el *tracking*, lo que causa dificultades en la estimación de la posición exacta y fiable basada en cámaras. Por otro lado, las reflexiones de radio que se pueden producir en un entorno de interior es posible que lleven a múltiples problemas relacionados con la localización basada en radio, disminuyendo la precisión del sistema.

Teniendo en cuenta que el objetivo es hacer un seguimiento a un individuo en el momento que esto se esté moviendo y poder mover la cámara con precisión para seguir el movimiento del individuo seleccionado en cada momento, se plantean como soluciones los siguientes dos sistemas, uno mediante hardware y el otro mediante software.

4.4.1 *Ubisense*:

Para el caso del método de localización por software, se ha escogido uno basado en la técnica UWB con una complejidad muy baja, que es el sistema comercial de *Ubisense*. Además, *Ubisense* y el centro de tecnologías aplicadas de *McLaren* han colaborado en el desarrollo de un sistema integrado¹² con la tecnología de RTLS de *Ubisense*, que es capaz de hacer un seguimiento de la posición de los atletas en tiempo real ofreciendo información sobre la ubicación exacta del deportista en cada momento. Este sistema se puede utilizar tanto en interior como en exterior.

¹² Sistema integrado: <http://www.ubisense.net/en/news-and-events/press-releases/mclaren-applied-technologies-and-ubisense-deliver.html>

Esta técnica deriva del conocimiento y utilización de datos en tiempo real de la plataforma utilizada para la formula1 del centro de tecnologías de *McLaren*. La finalidad de esta nueva plataforma es la de realizar un análisis de rendimiento del atleta basándose en su ubicación en cada momento. La técnica ya ha sido utilizada en deportes como el baloncesto, tenis, fútbol, hockey sobre hielo y la hípica, demostrando ser una herramienta de gestión del rendimiento muy valiosa en varios entornos diferentes.

La tecnología UWB es una solución excelente para la localización mediante sistemas inalámbricos, debido a su capacidad de alta resolución en el dominio del tiempo. Por ello, las técnicas basadas en UWB, suelen proporcionar mayor precisión que las demás técnicas. Su capacidad para resolver problemas derivados de las múltiples trayectorias de las señales, hace que sea posible obtener estimaciones exactas de la ubicación, sin necesidad de utilizar algoritmos más complejos. Esta capacidad de localización de una estimación precisa facilita muchas aplicaciones, tales como la vigilancia médica, la seguridad y el seguimiento de activos.

Como se ha descrito anteriormente el sistema de localización de *Ubisense* basado en tecnología de radio UWB, comprende una red de sensores sincronizados en tiempo entre los receptores y las etiquetas emisoras, colocadas en los objetos o personas que van a ser rastreados.

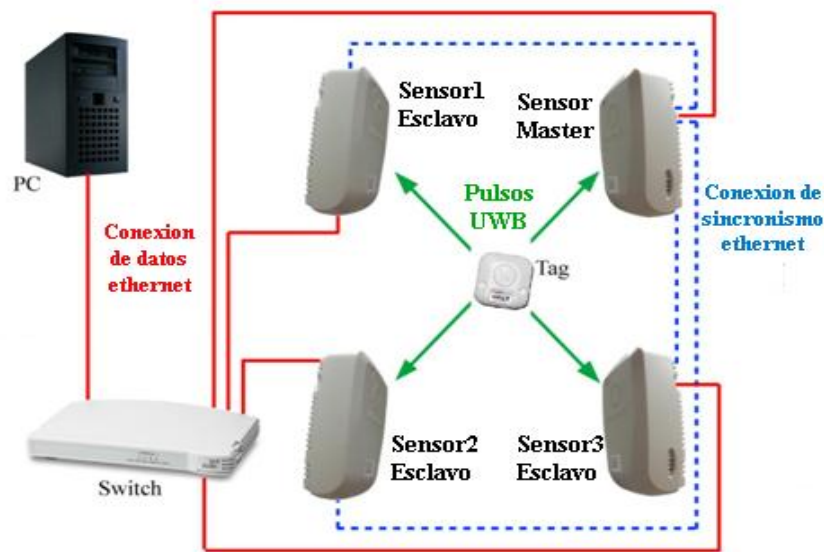


Ilustración 4-1: Arquitectura del sistema de Localización comercial UBISENSE

Los sensores se colocan en la zona en la que se procederá a la identificación de posición de las etiquetas. Estos sensores deben estar alimentados, conectados en red y conectados con un cable de sincronización. Los sensores se agrupan en conjuntos de cooperación que se llaman celdas de localización, y cada celda tiene un sensor maestro y una serie de sensores de esclavos.

Una celda de localización consta de uno o varios sensores conectados entre sí. Los sensores esclavos comparten los datos obtenidos con el sensor maestro, para que este calcule la posición de la etiqueta. Los sensores se configuran de manera que la sincronización de tiempo parte del sensor máster hasta cada uno de los esclavos.

Cada etiqueta se registra con la celda que contiene su ubicación, y se inserta en el programa de la celda. El programa determina el momento en el que la etiqueta debe emitir la señal modulada por UWB, la cual contiene el identificador de cada etiqueta. Cuando una etiqueta emite la señal UWB, la señal es recogida por uno o varios de los sensores de la celda. Los sensores de esclavos decodifican la señal UWB y envían el ángulo de llegada y la información de tiempo al sensor principal vía ethernet. El sensor principal acumula todos los datos recibidos y calcula la ubicación de la etiqueta, proporcionando la posición de la dirección de la etiqueta al programa de configuración de la celda que se encuentra en el PC.

Las etiquetas funcionan con dos canales de radio diferentes: un canal bidireccional de radiofrecuencia convencional y otro canal que solo emite en la banda ultra ancha (UWB).

Dado que hay un único canal UWB, sólo una etiqueta puede ser localizada en cada momento en cada celda de localización. Esta celda divide el tiempo en una serie de franjas temporales, y les asigna a las etiquetas intervalos de tiempo, de modo que su tasa de actualización es lo más cercana posible a la solicitada por la etiqueta. El canal de radiofrecuencia convencional se utiliza para administrar este programa, para enviar las pulsaciones de los botones de las etiquetas y para informar sobre el estado de la batería. El canal UWB sólo se utiliza cuando a cada etiqueta le corresponde el tiempo de emisión establecido.

La posición de la etiqueta se determina mediante la diferencia de tiempo de llegada (TDOA) y el ángulo de llegada (AOA), y las mediciones que se combinan con un algoritmo de mínimos cuadrados. La información de las posiciones de las etiquetas se obtiene a través de la plataforma del software propietario del sistema. La precisión del sistema es de 15 cm, con un 99% de los errores dentro de los 30 cm.

Por otro lado uno de los requisitos de este prototipo era el de utilizar una sola cámara para obtener la posición del individuo y realizar el seguimiento de éste. Con una sola cámara de video es posible obtener la posición de un individuo con una calibración dedicada para una cámara fija, pero en el momento que sea necesario mover la cámara para realizar el seguimiento, no se podría obtener la posición. Para superar esta limitación habría que calibrar la cámara para todas sus orientaciones posibles.

En cambio, el dispositivo *Kinect*, es capaz de obtener una información espacial de la posición del individuo, sin depender de su orientación hacia la escena.

4.4.2 Kinect:

El sensor *Kinect* es una barra horizontal de aproximadamente 23 cm conectada a una pequeña base circular con un eje de articulación de rótula, y está diseñado para ser colocado longitudinalmente por encima o por debajo del televisor.



Ilustración 4-2: *Dispositivo comercial Kinect*

Kinect se compone principalmente de una cámara RGB tradicional (Resolución 640x480 RGB 30fps VGA), un emisor de infrarrojos, una sensor CMOS monocromático de infrarrojos, 4 micrófonos (16bit velocidad de muestreo: 16Hz) y un motor.

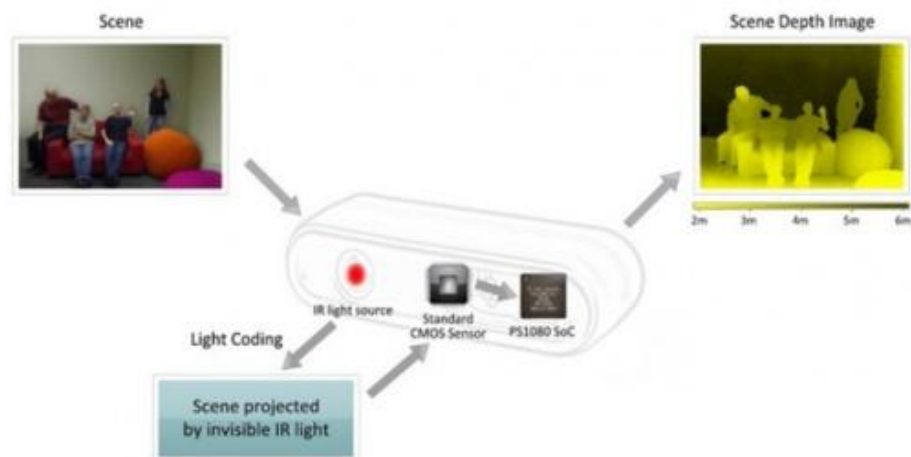


Ilustración 4-3: *Funcionamiento del dispositivo Kinect*

La cámara permite generar una imagen tridimensional de lo que tiene delante y además reconocer partes del cuerpo humano. Para ello utiliza un sónar de luz infrarroja, gracias al tiempo que tarda en rebotar y volver al dispositivo puede saber la distancia.

Para conocer la distancia a la que se encuentra cada píxel de la imagen de profundidad se emite una constelación de puntos con el emisor infrarrojo. Entonces, la cámara infrarroja detecta esta constelación y *Kinect* calcula la disparidad para cada píxel (la diferencia entre donde estaba el punto al proyectarlo a donde está en la proyección). Esencialmente, todos los píxeles que *Kinect* recibe como ruido IR son convertidos en una escala de colores, haciendo que los cuerpos, dependiendo de la distancia, se capturen como rojos, verdes, azules hasta llegar a tonos grises, que representan a objetos muy lejanos. A esto se le llama cámara de luz estructurada. *Kinect* gracias a ello tiene gran precisión en la detección de profundidad y no depende tanto de la luz ambiental.

A continuación, el software toma estas imágenes y las hace pasar por una serie de filtros para que *Kinect* determine qué es una persona y que no lo es. Para esto se siguen una serie de parámetros, en los cuales se define la estructura del cuerpo humano, para impedir que los obstáculos presentes en la escena sean reconocidos como otros jugadores.

Microsoft ha hecho público el kit de desarrollo para programar bajo la plataforma de *Kinect*, con lo que cualquiera puede utilizar el kit de desarrollo para programar su propia aplicación para utilizar el dispositivo *Kinect*. Con el SDK oficial del *Kinect* se pueden realizar dos *trackeos* esqueléticos simultáneos y cuatro trackeos de posición. Y también se pueden llevar a cabo reconocimientos de voz. Pero con la utilización de librerías no oficiales o de software libre, se pueden realizar reconocimientos de caras y *trackeo* de dedos.

Cuando se utiliza el software de la Xbox el campo de acción del sensor está limitado entre 1,2 y 3,5 metros pero el sensor es capaz de realizar un *trackeo* en un rango de entre 0,7 y 6 metros aproximadamente.

Por último, se ha realizado una fusión entre los dos sistemas anteriormente descritos.

4.4.3 Fusión:

Cada uno de los métodos descritos anteriormente tiene algunas limitaciones y, para superarlas, se ha realizado una fusión entre ambos. Esto se ha llevado a cabo

tratando los datos procedentes de los diferentes sistemas conjuntamente, con el fin de mejorar los resultados del sistema.

CAPÍTULO 5: EXPERIMENTOS

Los diferentes experimentos se han llevado a cabo en el laboratorio de televisión de Vicomtech, el cual se muestra en la Ilustración 5-1.



Ilustración 5-1: *Laboratorio de televisión de Vicomtech*

En primer lugar se ha procedido al montaje de la arquitectura, para poder realizar el movimiento del robot a través del PC en el que se van a llevar a cabo los cálculos de orientación del robot de la cámara en cada momento, en función a las posiciones obtenidas a través de los sistemas de localización.



Ilustración 5-2: *Arquitectura del sistema de comunicación entre el PC y el robot de la cámara*

El software desarrollado en el PC, se comunica con el robot a través del puerto serie, mediante el estándar de comunicación RS-232. Pero se ha tenido que convertir esta salida al estándar RS-422, mediante el conversor *ADAM-4520*, para que la fuente de alimentación *AutoCam PS27-5R* de Vinten pueda integrar esta señal con la alimentación del robot *Autocam PH-102* de Vinten, y realizar de esta manera los movimientos necesarios.

El protocolo de comunicación del puerto serie se basa en una velocidad de transmisión de 2400 baudios, 8 bits de datos, 1 bit de stop y ningún bit de paridad. El robot espera una trama de basada en 10 bytes, Tabla 5-3. El checksum es el valor de los 8 bits menos significativos de la suma de los 9 primeros bytes.

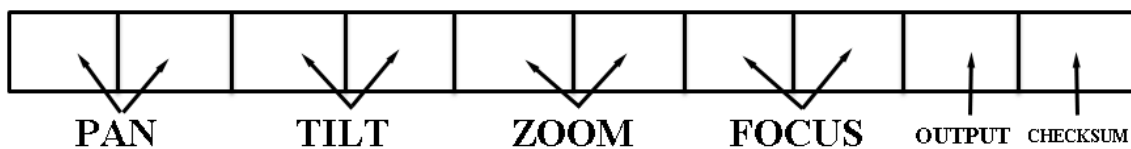


Ilustración 5-3: Trama de datos para la comunicación con el robot de la cámara

Si el valor del checksum no es el esperado toda la trama se interpretará como errónea y no se tendrá en cuenta. El robot de la cámara espera entre 15 y 20 tramas por segundo y la forma los movimientos realizados por la cámara sean lo más suave posibles consiste en enviar el máximo de tramas posible, es decir, 20 tramas por segundo.

Una vez especificado cuál es el funcionamiento del robot, se detallará cómo ha sido la superficie en la que se llevara a cabo la localización y seguimiento de los individuos, y cómo se ha calibrado la cámara en este entorno. Se dispone de un rectángulo de 5x5 metros en el que se coge como centro del sistema de referencia, el punto inferior derecho P (0, 0, 0), Ilustración 5-4.

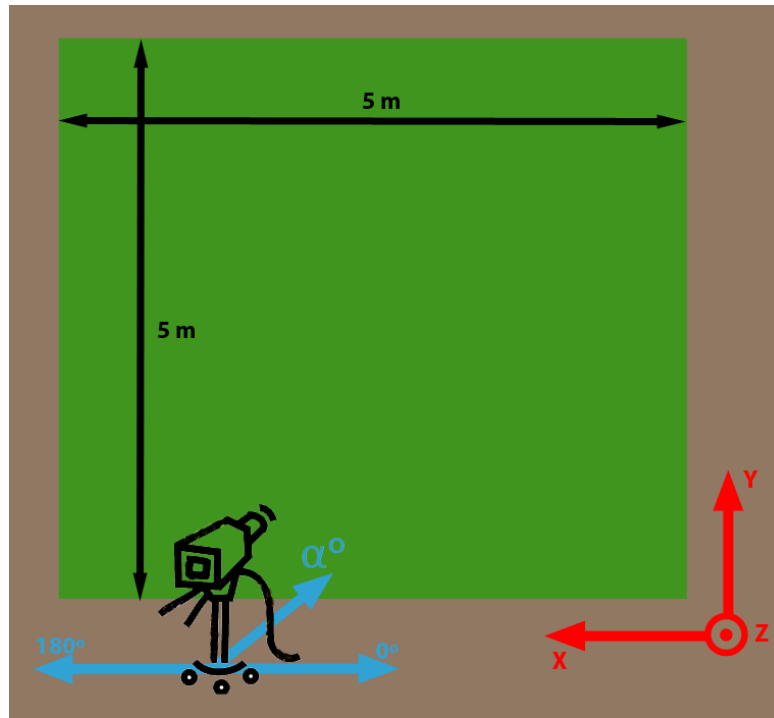


Ilustración 5-4: Superficie destinada a la detección con el sistema de referencia definido

Como se muestra en la Ilustración 5-5, se toma como ángulo de referencia del valor pan de la cámara, el giro que se realiza sobre el eje z, según el sistema de referencia establecido. Del mismo modo, el valor del ángulo de referencia del tilt de la cámara se toma el giro que se realiza sobre el eje x, siendo el valor de 0° cuando la cámara está mirando al suelo y 180° cuando está mira al techo. Se plantea una ecuación matemática que relaciona el ángulo de apertura (α) necesario del robot de la cámara, para orientar el pan a la posición (x, y) en la que se encuentra el objeto y otra ecuación para obtener el ángulo necesario de orientación de la cámara para la posición (y, z).

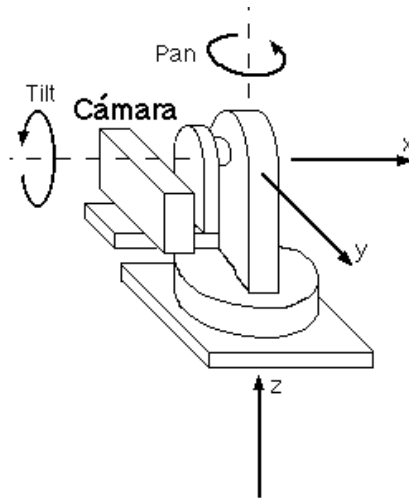


Ilustración 5-5: Configuración de los movimientos del robot en el sistema de referencia definido

Se coge como la orientación del pan para α_{\min} el punto $(0, 0, z)$ y para el α_{\max} el punto $(5, 0, z)$ y sabiendo cuáles son los valores en decimal de pan y los parámetros de posición de la cámara respecto a esos puntos, se realizan los cálculos de trigonometría pertinentes obteniendo los valores de α_{\min} y α_{\max} . La comparativa de los valores se puede observar en la Tabla 5-1. Del mismo modo se realiza la orientación para el tilt cogiendo para β_{\min} el punto $(x, 0, 0)$ y para el β_{\max} $(x, 0, 2)$ y realizando los cálculos de trigonometría obtenemos la relación del valor en decimal del tilt con los grados β_{\min} y β_{\max} .

Pan	α	Tilt	β
6400	23°	19559	28°
28160	139°	42767	135°

Tabla 5-1: Comparativa de los valores del robot con los valores de los ángulos

Basándose en que la relación entre el parámetro pan de la cámara y el ángulo es lineal, se despejan los coeficientes de la siguiente ecuación con los valores obtenidos en la Tabla 5-1.

$$Pan = m + n * \alpha \quad (1)$$

Despejando los valores de la anterior ecuación se obtiene la siguiente fórmula para el cálculo del valor del pan:

$$\text{Pan} = 2085,51 + 187,59 * \alpha \quad (2)$$

El desarrollo del valor del parámetro tilt es análogo al del pan, despejando los coeficientes de la siguiente ecuación con los valores de la Tabla 5-1:

$$\text{Tilt} = k + l * \beta \quad (3)$$

Despejando los valores de la anterior ecuación se obtiene la siguiente fórmula para el cálculo del valor del tilt:

$$\text{Tilt} = 13485,82 + 216,9 * \beta \quad (4)$$

5.1 Sistema *Ubisense*:

Se colocan los sensores del sistema *Ubisense* en las cuatro esquinas que tiene el área de localización, como en la Ilustración 5-6.

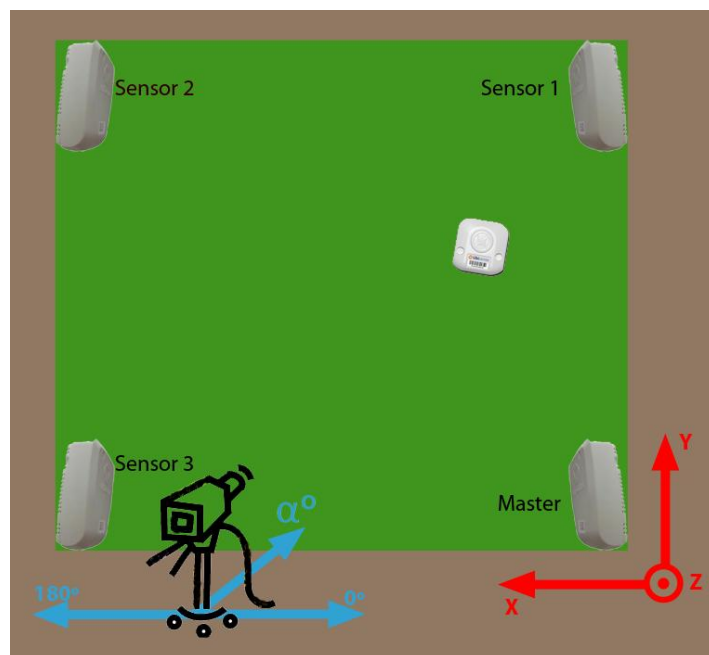


Ilustración 5-6: Distribución de sensores en la escena

Una vez colocados los sensores se procede a la calibración de los mismos Para ello hace falta orientarlos hacia el centro de la escena y medir las distancias entre ellos. Por otro lado, se coloca la etiqueta en una posición definida dentro del escenario. Finalmente, se introducen todos estos parámetros en el software de calibración del sistema *Ubisense*.

Una vez realizado este paso el sistema está listo para ser utilizado y el montaje sería el de la Ilustración 5-7.



Ilustración 5-7: Montaje del sistema de localización Ubisense

Por otra parte se tiene el algoritmo que realiza el movimiento de la cámara en función de la posición de la etiqueta a seguir. Para ello se tiene definida la posición de la cámara respecto al área de localización. De esta manera, cuando el sistema reciba la posición del objeto referida al sistema de referencia del escenario, podrá realizar los cálculos necesarios para indicar a la cámara los movimientos que debe realizar para seguir a la etiqueta.

Los sensores obtienen la posición de la etiqueta basándose en las técnicas de TDOA y AOA, realizando a continuación el cálculo del mínimo cuadrado de los valores recogidos mediante estas dos técnicas. Esta posición se recoge en el algoritmo de movimiento de la cámara mediante unas líneas de código del software del sistema *Ubisense*, y se realiza el cálculo del ángulo α y del β mediante un cálculo trigonométrico para realizar el seguimiento oportuno.

Con el fin de medir la precisión de este sistema, se han realizado pruebas en las que se coloca la etiqueta en diferentes posiciones fijas y conocidas de la escena, Ilustración

5-8, durante 2 minutos en cada una. Posteriormente, se han comparado las localizaciones obtenidas a través del sistema ubisense con las mediciones reales. En la Tabla 5-2 se muestran los resultados de precisión para cada punto del escenario evaluado y en la Tabla 5-3 la media de error del sistema.

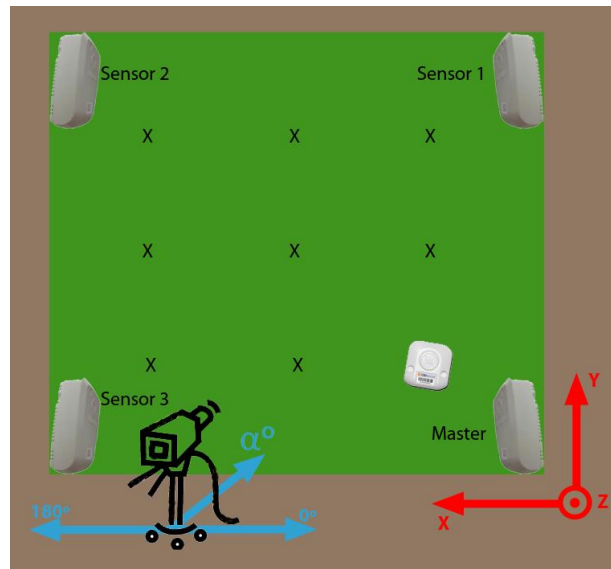


Ilustración 5-8: Posiciones (X) donde se ha medido la precisión

Precisión del sistema por puntos	P1	P2	P3	P4	P5	P6	P7	P8	P9
X	2 cm	6 cm	4 cm	11 cm	10 cm	1 cm	9 cm	18 cm	12 cm
Y	7 cm	3 cm	0 cm	15 cm	4 cm	2 cm	16 cm	6 cm	1 cm
Z	26 cm	2 cm	2 cm	5 cm	13 cm	16 cm	0 cm	5 cm	1 cm
Media	28 cm	8 cm	21 cm	27 cm	20 cm	21 cm	29 cm	20 cm	22 cm

Tabla 5-2: Errores de precisión en los diferentes puntos del escenario para el sistema Ubisense

Precisión del sistema	<i>Ubisense</i>
Media	22 cm

Tabla 5-3: La media de error del sistema *Ubisense*

Como se puede observar en la Tabla 5-2 este sistema tiene una gran precisión y una media de error muy baja, Tabla 5-3. Este error se encuentra dentro de los parámetros especificados por el fabricante, que son de 15 cm para el 95% de los casos y asciende hasta 30 centímetros para el 99% de los casos.

Una vez realizadas las pruebas de precisión del sistema, se ha procedido a realizar un seguimiento a la etiqueta con la cámara, introduciendo las posiciones obtenidas con el sistema de *Ubisense* en el algoritmo de movimiento del robot.

El mayor inconveniente de esta solución es que el sistema suministra la posición cada 850 milisegundos, y esto limita el movimiento de la cámara a la hora de realizar el seguimiento. El objeto a seguir está en movimiento pero la cámara no realiza un seguimiento continuo, sino que cada vez que recibe la información de la posición se mueve a esa dirección, con los golpes movimientos bruscos que ello supone.

Para superar esta limitación se ha desarrollado un filtro de suavizado mediante software, el cual tiene en cuenta las dos últimas posiciones obtenidas para poder hacer un cálculo de la siguiente posición. Esto se hace midiendo la distancia y la dirección en la que se ha movido la etiqueta en las últimas dos muestras y prediciendo cuál puede ser la siguiente posición. Una vez que se obtiene ese valor se divide esa distancia en 17 fracciones, para poder ir enviándole al robot estas nuevas 17 posiciones, 1 cada 50 milisegundos durante los 850 milisegundos, hasta obtener la nueva posición real. Si esta distancia es muy pequeña, del orden de 5 cm, no se tendrá en cuenta, porque puede ser debido a un error en la medición de la posición o a una pequeña oscilación que puede llevar a cabo el cuerpo donde está alojada la etiqueta.

Pero esto tiene la limitación de que cuando el individuo cambia la dirección de su trayectoria este filtro de suavizado no obtendrá una posición estimada aceptable y tardará en reaccionar.

Este sistema también tiene sus puntos fuertes y el más importante es el de poder obtener la identificación de cada etiqueta que se sigue. De ese modo se puede seleccionar cuál de ellas hay que seguir en cada momento y tener identificado el objeto a seguir en todo momento. Además, la precisión que tiene este sistema es más que suficiente para realizar el seguimiento con el robot de la cámara.

5.2 Kinect:

El escenario en el que se han llevado a cabo las pruebas con el dispositivo *Kinect* ha sido el mismo que el que se aprecia en la Ilustración 5-7. La misma que para las pruebas del sistema de localización *Ubisense*. La única diferencia ha sido el acoplamiento del dispositivo *Kinect* sobre una de las cámaras de video. Ver Ilustración 5-9.



Ilustración 5-9: Dispositivo kinect colocado sobre la cámara de video

Una vez colocado el dispositivo *Kinect* sobre la cámara, este dispositivo se conecta al PC con un cable USB. En este PC se llevan a cabo las labores de computación necesarias para el posicionamiento y seguimiento del individuo. Reciben dos flujos de información desde el Kinect, uno es la señal de video en RGB y el otro la señal del receptor infrarrojo, que se interpreta como una imagen de profundidades.

La parte del desarrollo que se encarga de analizar la información recibida desde el dispositivo *Kinect* tiene dos módulos principales, el primero el de calibración y el segundo de seguimiento. En el momento que se lanza el programa se pone en marcha el primer modulo de calibración que trata de detectar la cara del individuo a seguir. Para ello analiza cada fotograma buscando la forma de una cara mediante una función

estándar de reconocimiento facial integrada ya en varias librerías de software libre como función estándar. Una vez que la encuentra lo vuelve a intentar en la misma posición de la imagen en los siguientes cuatro fotogramas y si ha sido detectado satisfactoriamente ya tiene al individuo a seguir localizado. En la imagen inferior izquierda, de la Ilustración 5-10 se observa un círculo rojo que reconoce la cara de la persona en cada fotograma. La distancia en profundidad en la que se encuentra este punto se busca en el mapa de profundidades y así se obtiene la posición del objeto respecto a la cámara.

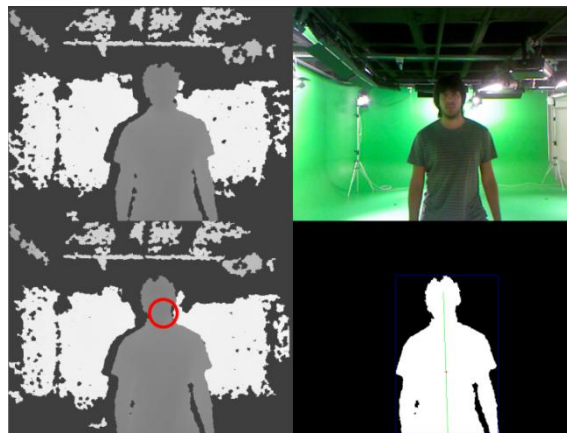


Ilustración 5-10: Detección del individuo al que se le realizara el seguimiento

Una vez localizado el individuo, entra en escena el segundo modulo, de seguirlo. Para ello se filtra la imagen de profundidad utilizando el valor de profundidad en el fotograma anterior (con un offset de seguridad que permita el movimiento mínimo del individuo entre fotogramas). Además se establece una región de interés alrededor de la silueta la cual facilitará la segmentación de la silueta. Ver Ilustración 5-12.

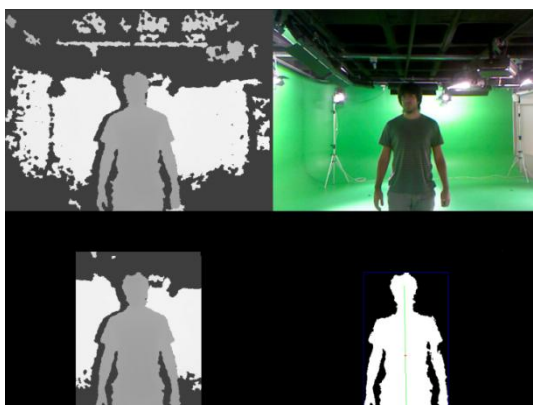


Ilustración 5-11: Filtro de seguimiento del individuo

Una vez realizado este proceso, el individuo puede comenzar a moverse por la escena y se va analizando fotograma a fotograma cual es el movimiento de la persona para actualizar la ventana que determina la región de interés, su posición (el centroide de la silueta) en la imagen y su profundidad. Estos valores son actualizados en cada fotograma posibilitando el tracking del individuo.

Un potencial del sistema es que otro individuo puede irrumpir en la escena y cruzarse con el individuo localizado sin que la oclusión afecte al sistema de seguimiento.

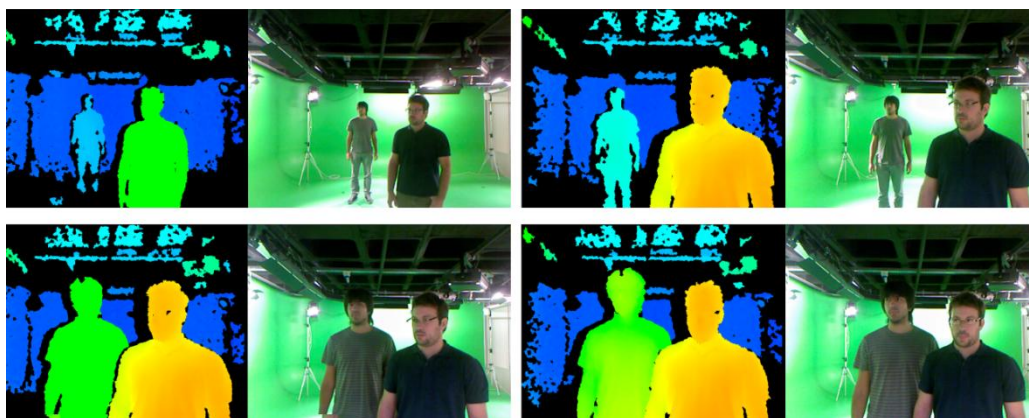


Ilustración 5-12: Profundidad de la imagen representada por diferentes colores

La posición del individuo en la imagen se obtiene en valores de *pixels* del centro de la región-silueta del individuo, y su profundidad en una escala dependiente del rango dinámico que ofrece la *Kinect* para este fin. Por ello hay que realizar cálculos de transformación de coordenadas para obtener la posición en tres dimensiones referenciada a al sistema de referencia de la escena.

Tal y como se aprecia en la imagen Ilustración 5-13, el sistema de coordenadas de la escena real y del plano de la imagen que obtenemos de la cámara motorizada, no se corresponden entre sí. Es necesario un cambio de escala/unidades (pixel a metros) y traslación en los tres ejes.

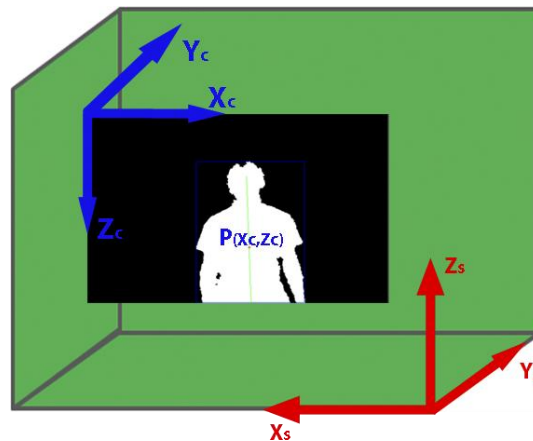


Ilustración 5-13: Relación del sistema de referencia del escenario con el de la imagen

La profundidad ($P_{(y)}$) se relaciona linealmente por un factor de ajuste (f_r) que relaciona directamente el rango dinámico de la cámara *Kinect* dedicada a la profundidad (8 bits, 255 niveles) y la profundidad que dicha cobertura supone en unidades del escenario real. Para este eje únicamente realizaremos el cambio escala/unidades.

$$P_{(y_s)} = P_{(y_c)} * f_r \quad (5)$$

Para que los datos extraídos de la imagen ($P_{(x_c, z_c)}$) se correspondan con las posiciones de la escena real ($P_{(x_s, z_s)}$) (que son las interpretadas por robot de la cámara), realizamos una transformación tanto en escala/unidades como en traslación. Dicha transformación se lleva a cabo pasando la coordenadas de píxeles a metros y sumando el vector de diferencia existente entre los dos sistemas de coordenadas ($\vec{V}_{\text{traslación}}$).

$$O_s = O_{\text{fijo}} + \vec{V}_{\text{traslación}} \quad (5)$$

$$P_{(x_s, z_s)} = P_{(x_c, z_c)} + \vec{V}_{\text{traslación}} \quad (6)$$

$$\vec{V}_{\text{traslación}} \begin{cases} V_{\text{traslación}_x} = O_{s_x} - O_{\text{fijo}_x} \\ V_{\text{traslación}_y} = O_{s_y} - O_{\text{fijo}_y} \\ V_{\text{traslación}_z} = 0 \end{cases} \quad (7)$$

O_{fijo} es el centro de coordenadas del plano de la imagen al inicializar el programa de tracking. Al moverse la cámara en su función de *trackear* personas, el sistema de coordenadas del plano de la imagen también se mueve, por lo que necesita ser actualizada por $\vec{V}_{\text{movimiento}}$, al que se hace referencia en la Ilustración 5-14. Este vector se calcula a partir del movimiento en la imagen de la persona a *trackear* ya que el movimiento de la cámara es la misma que el movimiento del individuo en la imagen en el instante anterior.

$$O_{\text{fijo}} = O_{\text{dinámico}} (n=0) \quad (8)$$

$$O_{\text{fijo}} = O_{\text{dinámico}} + \vec{V}_{\text{movimiento}} \quad (9)$$

$$P_{(x_c, z_c)} = P_{(x_{\text{cdin}}, z_{\text{cdin}})} + \vec{V}_{\text{movimiento}} \quad (10)$$

$$\vec{V}_{\text{movimiento}} = P_{(x_c, z_c)}(n) - P_{(x_c, z_c)}(n-1) \quad (11)$$

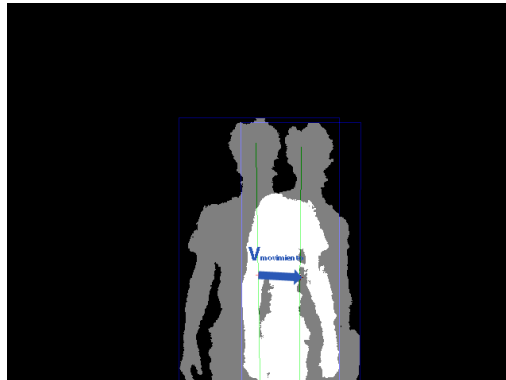


Ilustración 5-14: Vector de movimiento entre dos fotogramas

Se han realizado las mismas pruebas descritas en el primer apartado *Ubisense* del Capítulo 5, pero repitiendo el proceso 5 veces con el fin de obtener la precisión del sistema.

La precisión del dispositivo *Kinect* es milimétrica pero en las condiciones óptimas, que son cuando la persona se encuentra a una distancia de entre 1,5 y 3 metros y el dispositivo no se mueve. En el caso de este experimento, el dispositivo está en movimiento, y hace falta realizar transformaciones para llevar las coordenadas del plano de la imagen a las coordenadas de la escena.

Precisión del sistema por puntos	P1	P2	P3	P4	P5	P6	P7	P8	P9
X	58 cm	13 cm	15 cm	63cm	25 cm	37 cm	72 cm	46 cm	46 cm
Y	16 cm	40 cm	42 cm	87 cm	76 cm	76 cm	54 cm	6 cm	6 cm
Z	20 cm	20 cm	19 cm	19 cm	19 cm	20 cm	19 cm	18 cm	19 cm
Media	85 cm	45 cm	49 cm	102 cm	49 cm	87 cm	91 cm	49 cm	47 cm

Tabla 5-4: Errores de precisión en los diferentes puntos del escenario para el sistema *Kinect*

Precisión del sistema <i>Kinect</i>	Eje X	Eje Y	Eje Z	Total
Media	0,35 cm	0,19cm	0,62 cm	0,73 cm
Desviación Estándar	0,27-	0,05	0,16	0,32

Tabla 5-5: La media de error del sistema *Kinect* por eje y su media

Se ha visto que la tendencia de las mediciones del sistema es siempre la misma, puntos rojos, como se puede apreciar en la Ilustración 5-15.

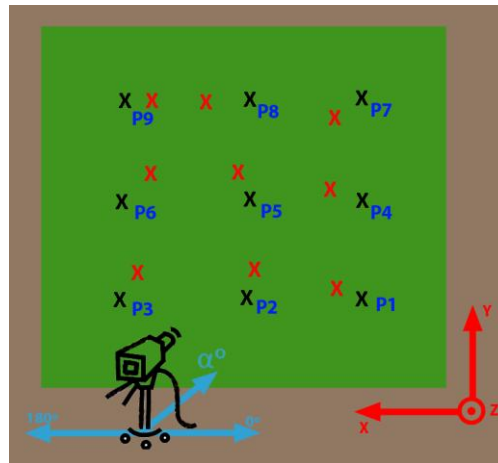


Ilustración 5-15: Tendencia del error en (puntos rojos) en cada punto de calibración (puntos negros)

Las coordenadas que se obtienen mediante este sistema tienden a situarse en los puntos rojos de la Ilustración 5-15, cuando el individuo se encuentra en los puntos negros.

Pero los resultados obtenidos no son los óptimos. Las medias de error de cada punto son bastante altas. Se puede observar que los valores de profundidad son mejores en los puntos más lejanos de la cámara, ver Tabla 4-1, porque en la calibración del sistema se ha buscado cubrir el máximo escenario posible. Esto se ha llevado a cabo a costa de perder precisión de profundidad en los puntos cercanos e intermedios a la cámara.

Parte de este error se estima que es debido a la falta de precisión en la parametrización de la transformación de coordenadas del plano de la cámara al sistema de referencia. Esta parametrización resulta dificultosa debido a que el sistema de referencia de la cámara se mueve con los giros pan y tilt. Por otra parte, correcciones menos importantes pero que no dejan de añadir errores como la falta de corrección de la lente de la cámara y su calibración, afectan directamente en estos resultados. Estas mejoras se contemplarán para trabajos futuros.

5.3 Fusión:

Cuando se habla de la fusión de estos dos sistemas, se dice que el software desarrollado para realizar el seguimiento, es capaz de combinar la información obtenida mediante el sistema de *Ubisense* y por la obtenida mediante el dispositivo *Kinect*. El código ofrece las tres opciones, la de ejecutar los dos sistemas por separado y la de fusionarlos. Para

esto solo hay que cambiar una variable en el código desarrollado, seleccionando cuál será la fuente de datos a tener en cuenta, liberando al programa de realizar cálculos innecesarios en caso de que no se desee realizar el seguimiento mediante la fusión de los sistemas.

Lo que se ha conseguido con esta fusión, ha sido realizar un movimiento de robot más suave a la hora de realizar el movimiento para grabar al individuo. Para ello se le ha dado más peso a la información que se obtiene desde el sistema *Kinect* debido a que este sistema tiene un mayor flujo de datos, aunque sea menos preciso. La precisión del sistema fusionado es la misma que el dispositivo *Kinect* por sí solo, menos cuando el sistema de localización de *Kinect* falla, y se toma en cuenta la posición del sistema *Ubisense*.

No se ha conseguido mejorar la precisión del sistema con la fusión de ambas técnicas, pero si es un sistema robusto ante las oclusiones.

CAPÍTULO 6: CONCLUSIONES

En este capítulo se exponen las conclusiones extraídas de los experimentos realizados y los resultados obtenidos.

Realizando un análisis de la localización de ambos sistemas por separado, se observa que el resultado de la prueba de precisión del dispositivo *Kinect*, en un entorno de interior reducido como el que ha servido para realizar estas pruebas, es bastante menor. Pero aun así, esta falta de precisión es suficiente para realizar un seguimiento del individuo con el robot de la cámara. Es más, su alta frecuencia de obtención de posiciones, compensa con creces su carencia de precisión. Además, como se ha descrito en el capítulo 5.2, el software desarrollado para el seguimiento de individuos basado en la información que proporciona el dispositivo *Kinect*, es capaz de superar las oclusiones parciales, y además si es durante un tiempo limitado también las totales.

En cambio el punto más flojo del sistema es que el alcance del infrarrojo, que llega a cubrir a duras penas el espacio del escenario donde se ha realizado el experimento.

Por el contrario el sistema de localización de *Ubisense* tiene una precisión muy alta, con lo que consigue que el objetivo se encuadre bien en la cámara en el momento que este facilita su posición, pero el retardo que supone la espera de una nueva posición, hace que el objetivo pueda salirse del encuadre de la cámara. Aun ocurriendo esto, se seguiría teniendo al individuo localizado en todo momento mediante el identificador de la etiqueta.

Esta diferencia de precisión entre ambos sistemas es de esperar basándose en la diferencia de precio, siendo mucho más caro el sistema de localización de *Ubisense*.

La solución obtenida mediante este sistema por sí sola, no sería muy agradable para el espectador debido a los golpes bruscos originados por esta baja frecuencia de flujo de información.

Aunque no se haya obtenido una mejora en la precisión de la fusión de los sistemas, se ha conseguido realizar un seguimiento de cámara suave, pudiendo obtener una calidad de *paneo* aceptable. Se tienen identificados a los individuos en todo momento y se puede seleccionar por software, a cuál de ellos se desea realizar el seguimiento,

superando así el problema de las oclusiones. En esta fusión se ha visto que no es necesario el dato de profundidad que proporciona el dispositivo *Kinect* para realizar el seguimiento del individuo, siendo suficiente la información extraída a partir de la imagen de la cámara de video con la fusión del sistema *Ubisense*.

Hoy en día, en algunos eventos deportivos se tiene la posibilidad de conocer la posición del deportista mediante sistemas de localización como pueden ser GPS, *InMotion*, o *Witrack*. El sistema desarrollado sería capaz de obtener los datos de posición espaciales de estos sistemas, y ayudado con el procesado de la imagen de la cámara, realizar un seguimiento a los deportistas, habiendo calibrado el sistema para el entorno en el que se desarrollará la prueba.

CAPÍTULO 7: TRABAJOS FUTUROS

En este capítulo se recogen un conjunto de futuras posibilidades de desarrollo del trabajo iniciado en esta tesis de fin de máster.

Un avance importante para este sistema sería el de realizar una calibración del sistema *Kinect* más exacta y correcciones más leves como la de la lente con el fin de obtener unos mejores resultados en los experimentos realizados, obteniendo una mayor exactitud en los diferentes puntos de la escena.

Por otro lado sería conveniente intentar mejorar la profundidad del sistema basado en el dispositivo *Kinect* para abarcar mas campo de visión y llevar a cabo una implementación de técnicas para hacer *zoom in* o *zoom out*, al individuo al que se le está realizando el *trackeo*. Para ello, se tendría en cuenta el potencial de la función de detección fácil, pudiendo realizar primeros planos del individuo basándose en esta técnica.

Se ha visto la necesidad de realizar una fusión con el sistema de posicionamiento de *Ubisense* y una cámara de video convencional, y evaluar estos resultados con los obtenidos en este trabajo con el fin de completar este estudio.

Otro trabajo que dotaría a este prototipo de mas funcionalidad sería el de introducir información virtual en la imagen basándose en los movimientos de basándose en los movimientos del individuo, teniendo constancia de los parámetros de la cámara en cada momento. Además, basándose en los *trackeos* esqueléticos simultáneos que se pueden realizar con el kit de desarrollo de *Kinect* e incluso realizando el *trackeo* de dedos con librerías de software libre se podría desarrollar una aplicación interactiva.

Por último y como el trabajo más ambicioso de todos, es el de conseguir realizar algunas pruebas de campo para verificar la funcionalidad del prototipo desarrollado en un evento deportivo real, y evaluar sus puntos fuertes y los débiles.

REFERENCIAS

- [1] S. Arseneau, J.R. Cooperstock, "Presenter Tracking in a Classroom Environment", in Proceedings of 1999 IEEE Conference on Industrial Electronics Society pp. 145 - 148 vol.1, November 1999
- [2] T. Yang, S.Z. Li, Q. Pan, J. Li, Real-time multiple objects tracking with occlusion handling in dynamic scenes, in: Computer Vision and Pattern Recognition, San Diego, CA, June 20–25, 2005.
- [3] S. Á. GUÐMUNDSSON, R. LARSEN, H. AANÆS, M. PARDÁS, J. R. CASAS, "TOF imaging in smart room environments towards improved people tracking". In IEEE Conf. on Computer Vision & Pattern Recogn.; Workshop on ToF-Camera based Computer Vision (2008).
- [4] K. Smith, D.G. Perez, J.M. Odobez, Using particles to track varying numbers of interacting people, in: Computer Vision and Pattern Recognition, San Diego, CA, June 20–25, 2005.
- [5] B. Wu and R. Nevatia, "Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet Based Part Detectors," Int'l J. Computer Vision, vol. 75, no. 2, pp. 247-266, 2007.
- [6] T. Zhao, R. Nevatia, and B. Wu, "Segmentation and tracking of multiple humans in crowded environments," in IEEE Trans. Pattern Analysis and Machine Intelligence (to appear), 2008.
- [7] F. Fleuret, J. Berclaz, R. Lengagne, and P. Fua, "Multicamera People Tracking with a Probabilistic Occupancy Map," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 30, no. 2, pp. 267-282, Feb. 2008.
- [8] Z. Zhang, G. Potamianos, A. Senior, S. Chu, and T. S. Huang. "A joint system for person tracking and face detection," Proc. IEEE Int. Work. Human Comp. Interaction, Beijing, 2005.
- [9] Z. Zhang, G. Potamianos, S. M. Chu, J. Tu, and T. S. Huang, "Person tracking in smart rooms using dynamic programming and adaptive subspace learning," in Proc. Int. Conf. on Multimedia and Expo, 2006, pp. 2061–2064.
- [10] Ajmal Mian, "Realtime Face Detection and Tracking Using a Single Pan, Tilt, Zoom Camera" 23rd International Conference Image and Vision Computing, New Zealand, November 2008
- [11] T. Ba Dinh, N. Vo, G. Medioni, "High Resolution Face Sequences from A PTZ Network Camera", IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011), March 2011, pp 531 - 538
- [12] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in Proceedings of 2000 IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head, SC, volume II, June 2000, pp. 142–149.
- [13] Needham, C. J., Boyle, R. D.: "Tracking multiple sports players through occlusion, congestion and scale." British Machine Vision Conference, vol. 1, pp. 93-102 BMVA (2001)
- [14] M. Kristan, J. Perš, M. Peršič, M. Bon, and S. Kovačič. "Multiple interacting targets tracking with application to team sports". In 4th International Symposium on Image and Signal Processing and Analysis ISPA, pages 322–327, September 2005.
- [15] GuoJun Liu, XiangLong Tang, JianHua Huang, JiaFeng Liu, Da Sun, "Hierarchical Model-Based Human Motion Tracking Via Unscented Kalman Filter," iccv, pp.1-8, 2007 IEEE 11th International Conference on Computer Vision, 2007
- [16] S. Kasiri-Bidhendi, R. Safabakhsh, "Tracking of the Players and Ball in Indoor Soccer Games in the Presence Off Occlusion" Computer Conference, 2009. CSICC 2009. 14th International CSI
- [17] Hayet, J.B., Mathes, T., Czyk, J., Piater, J., Verly, J., Macq, B.: A modular multi-camera framework for team sports tracking. In: International Conference on Advanced Video and Signal based Surveillance, Como, Italy (2005) 493–498
- [18] W. Du, J.-B. Hayet, J. Piater, and J. Verly, "Collaborative multi-camera tracking of athletes in team sports," in Proc. Workshop Computer Vision Based Analysis in Sport Environments, 2006
- [19] C. Ó Conaire, P. Kelly, D. Connaghan, and N.E. O'Connor. "Tennissense: A platform for extracting semantic information from multi-camera tennis data". In International Conference on Digital Signal Processing (DSP) 2009.
- [20] J. Chandaria, G. Thomas, B. Bartczak, K. Koeser, R. Koch, M. Becker, G. Bleser, D. Stricker, C. Wohlleber, M. Felsberg, J. Hol, T. Schoen, J. Skoglund, P. Slycke, and S. Smeitz. Real-time camera tracking in the matrix project. In Proceedings of International Broadcasting Convention (IBC), pages 321–328, Amsterdam, The Netherlands, 2006.

- [21] K. Pahlavan, X. Li, and J. Makela, "Indoor geolocation science and technology," *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 112–118, Feb. 2002.
- [22] M. Brunato and R. Battiti, "Statistical learning theory for location fingerprinting in wireless LANs," *Comput. Netw.*, vol. 47, pp. 825–845, 2005.
- [23] C. L. Wu, L. C. Fu, and F. L. Lian, "WLAN location determination in ehome via support vector classification," in *Proc. IEEE Int. Conf. Netw., Sens. Control*, 2004, vol. 2, pp. 1026–1031.
- [24] P. Prasithsangaree, P. Krishnamurthi, and P. K. Chrysanthis, "On indoor position with wireless LANs," in *Proc. IEEE Int. Symp. Pers. Indoor, Mobile Radio Commun.*, Sep. 2002, vol. 2, pp. 720–724.
- [25] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [26] J. Barnes, C. Rizos, J. Wang, D. Small, G. Voigt, and N. Gambale. "Locata: The positioning technology of the future?" presented at 6th Int. Symp. Satellite Navig. Technol. Incl. Mobile Positioning Location Services, Melbourne, Australia, 2003
- [27] Artículo: B. Wisbey, P. G. Montgomery, D.B. Pyne, and B. Rattray, "Quantifying movement deands of AFL football using GPS tracking", in *J Sci Med Sport Journal Science and Medicine in Sport*. doi:10.1016/j.jsams.2009.09.002, 2009
- [28] Jeffrey Hightower, Roy Want, and Gaetano Borriello. *SpotON: An indoor 3d location sensing technology based on RF signal strength*. UW-CSE 00-02-02, University of Washington, Department of Computer Science and Engineering, Seattle, WA, February 2000.
- [29] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil, "LANDMARC: Indoor location sensing using active RFID," *Wireless Netw.*, vol. 10, no. 6, pp. 701–710, Nov. 2004.
- [30] M. Beetz, B. Kirchlechner, and M. Lames, "Computerized Real-Time Analysis of Football Games", *IEEE pervasive computing mobile and ubiquitous systems*, vol. 4(3), pp.33-39, July–September 2005.
- [31] J. J. Caffery and G. L. Stuber, "Overview of radiolocation in CDMA cellular system," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 38–45, Apr. 1998.
- [32] V. Otsason, A. Varshavsky, A. LaMarca, and E. de Lara, "Accurate GSM indoor localization," *UbiComp 2005, Lecture Notes Computer Science*, Springer-Verlag, vol. 3660, pp. 141–158, 2005.
- [33] S. Gezici, Z. Tian, G.B. Giannakis, H. Kobayashi, A.F. Molisch, H.V. Poor, Z. Sahinoglu, "Localization via ultrawideband radios," *IEEE Signal Process. Mag.* 22 (4) (2005) 70–84.
- [34] A. Bensusky, *Wireless Positioning Technologies and Applications*. GNSS Technology and Applications Series, Artech House, 2008.
- [35] R. J. Fontana, "Recent system applications of short-pulse ultra-wideband (UWB) technology," *IEEE Trans. Microw. Theory Tech.*, vol. 52, no. 9, pp. 2087–2104, Sep. 2004.
- [36] A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, J. Tabert, P. Powledge, G. Borriello, and B. Schilit. *Place Lab: Device positioning using radio beacons in the wild*. In *proceedings of Pervasive*, 2005.
- [37] Paciga, M. and Lutfiyya, H. *Herecast: An Open Infrastructure for Location-Based Services Using WiFi*, *IEEE Wireless and Mobile Computing (WiMob'2005)*, Aug. 2005.
- [38] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM 2000*, Mar., vol. 2, pp. 775–784.
- [39] P. Bahl and V. N. Padmanabhan, "Enhancements to the RADAR user location and tracking system," *Microsoft Corp., Tech. Rep. MSR-TR- 2000-12*, Feb. 2000.
- [40] M. Youssef and A. K. Agrawala, "Handling samples correlation in the Horus system," *IEEE INFOCOM 2004*, Hong Kong, vol. 2, pp. 1023– 1031, Mar. 2004.
- [41] A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach, and L. E. Kavraki, "Practical robust localization over large-scale 802.11 wireless networks," in *Proc. 10th ACM Int. Conf. Mobile Comput. Netw.*, Philadelphia, PA, Sep. 26–Oct. 1, 2004, pp. 70–84.
- [42] AeroScout Company. [Online]. Available: <http://www.aeroscout.com/>
- [43] A. Kotanen, M. Hannikainen, H. Leppakoski, and T. D. Hamalainen, "Experiments on local positioning with Bluetooth," in *Proc. IEEE Int. Conf. Inf. Technol.: Comput. Commun.*, Apr. 2003, pp. 297–303.

- [44] A. Song-Ie Noh, et al., "Comparison of the Mechanisms of the Zigbee's Indoor Localization Algorithm," vol. 0, ed. Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, ACIS International Conference, 2008, pp. 13-18.
- [45] Andy Ward, Alan Jones, Andy Hopper "Performance Analysis of UPS (Ultrasonic Positioning System) Using DGPS" 30th Annual Conference of the IEEE Industrial Electronics Society, Busan, South Korea, November 2004
- [46] N. Priyantha, A. Chakraborty, and H. Balakrishnan, "The Cricket Location-Support System" in Proc. ACM 6th Int. Conf. on Mobile Computing and Networking. (MOBICOM), August 2000.
- [47] Dijk, E.O. van Berkel, C.H. Aarts, R.M. van Loenen, E.J. "A 3-D Indoor Positioning Method using a Single Compact Base Station" Pervasive Computing and Communications. Second IEEE Annual Conference on 2004.
- [48] McCarthy, P. Muller, H.L. "RF-free ultrasonic positioning" in Proc. Wearable Computers, Seventh IEEE International Symposium on 2003.
- [49] R. Want, A. Hopper, V. Falcao, and J. Gibbons. "The active badge location system". Technical Report 92.1, Olivetti Research Ltd, 1992.
- [50] A.R. Dick and M.J. Brooks. "A stochastic approach to tracking objects across multiple cameras", *Advances in Artificial Intelligence*, 3339:160–170, 2004
- [51] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale, and S. Shafer. "Multi-camera multi-person tracking for EasyLiving". In *Proceedings of the 3rd IEEE International Workshop on Visual Surveillance (VS 2000)*, page 3, 2000.
- [52] J. Liu et al. "Automatic player detection, labeling and tracking in broadcast soccer video". *Pattern Recognition Letters*. New York, USA: Elsevier Science Inc., vol. 20(2), pp. 103-113, 2009.
- [53] M. Pallavi, J. Mukherjee, A.K. Majumbar, S. Sural, "Graph-Based Multiplayer Detection and Tracking in Broadcast Soccer Videos". *IEEE Transactions on Multimedia*, vol.10(5), pp. 794-805, 2008.
- [54] M. Beetz et al., "Visually tracking football games based on TV broadcasts", in *Proceedings of the 20th international joint conference on Artificial intelligence*, pp. 2066-2071, 2007.
- [55] M. Beetz et al., "ASPOGAMO: Automated sports games analysis models (ASPOGAMO: Automatisierte Sportspielanalyse-Modelle). *Int. J. Computer Sci. Sport*, vol. 8 (1), 4-12, 2009.
- [56] S.K., Bidhendi, and R. Safabakhsh, R., "Effective Tracking of the Players and Ball in Indoor Soccer Games in the Presence of Occlusion", in *Proceedings of the 14th International CSI Computer Conference*, pp. 524-528, 2009.
- [57] G. Liu, X. Tang, J. Huang, J. Liu, and D. Sun, "Hierarchical modelbased human motion tracking via unscented Kalman filter", 2007 IEEE 11th International Conference on Computer Vision, pp. 889-896, 2007.
- [58] S. Iwase and H. Saito, "Tracking soccer players based on homography among multiple views". *Visual Communications and Image Processing*, vol. 5150, pp.283-292, 2003.
- [59] A. Alahi, Y. Boursier, L. Jacques, and P. Vanderghyest, "Sport Players Detection and Tracking with a Mixed Network of Planar and Omnidirectional Cameras", in the *Third ACM/IEEE International Conference on Distributed Smart Cameras*, Como, 2009.
- [60] D. Delannay, N. Danhier, and C. D. Vleeschouwer, "Detection and Recognition of Sports(women) from Multiple views", in the *Third ACM/IEEE International Conference on Distributed Smart Cameras*, Como, 2009.