

SISTEMA DE VISIÓN ARTIFICIAL PARA EL SEGUIMIENTO DE USUARIOS MEDIANTE MARCADORES E ILUMINACIÓN INFRARROJA EN ENTORNOS DE GRAN ACTIVIDAD FÍSICA

Rojas Barrera, Adrian

Curs 2013-2014

Director: Narcís Parés Burguès

GRAU EN ENGINYERIA INFORMÀTICA



Universitat
Pompeu Fabra
Barcelona

Escola
Superior Politècnica

Treball de Fi de Grau

SISTEMA DE VISIÓN ARTIFICIAL PARA EL
SEGUIMIENTO DE USUARIOS MEDIANTE
MARCADORES E ILUMINACIÓN INFRARROJA EN
ENTORNOS DE GRAN ACTIVIDAD FÍSICA

Adrian Rojas Barrera

TRABAJO FINAL DE GRADO

GRADO EN INFORMÁTICA

UNIVERSIDAD POMPEU FABRA

2013/2014

DIRECTOR DEL TRABAJO

Narcís Parés

Agradecimientos

Primeramente darle las gracias a mi tutor del trabajo final de grado Narcís Parés por hacer posible que haya podido realizar este proyecto, sin él no hubiese aprendido lo que he aprendido ni habría madurado tanto, a nivel de ingeniero, como lo he hecho.

Agradezco enormemente la ayuda aportada por todas las personas que han estado a mi alrededor a lo largo del desarrollo de este proyecto y a las que me han acompañado a lo largo de todo mi Grado en Informática en la Universidad Pompeu Fabra de Barcelona.

A toda mi familia que siempre estuvo ahí para apoyarme en los momentos más difíciles.

Especialmente me gustaría agradecer la colaboración de Adrià Ruiz, su apoyo a lo largo de la implementación del proyecto y sus conocimientos me han ayudado muchísimo en la elaboración del mismo. También a su grupo de trabajo que siempre ha estado ahí para sacarme una sonrisa en los buenos y malos momentos.

Por último, pero no menos importante, agradecerle a una persona que ha estado a mi lado de inicio a fin en este largo camino, Pascal Landri, ha sido un estupendo coordinador, que ha estado siempre apoyándome y preocupándose para hacerme avanzar, que ha confiado en mí y que me ha ayudado en lo que ha podido y en lo que le he pedido. *“Merci beaucoup Pascal”/ “Thank you very much Pascal”*.

Resumen

Es por todos conocido el efecto de los videojuegos en los niños tanto en aspectos sociales como saludables, el presente proyecto forma parte del área de investigación del grupo CMTech del DTIC sobre la potenciación de la actividad física en niños mediante *Exergames* (plataformas que fusionan videojuegos y actividad física).

Se trata de la implementación de un sistema para detectar diferentes cuerpos mediante el uso de una cámara y luces infrarrojas, además, hacer una recopilación del movimiento de los mismos (tracking) para saber desde su localización hasta el recorrido que han hecho a tiempo real. Este sistema mejora sensiblemente el seguimiento de la actividad física de los niños en una plataforma de *Exergames* llamada “Tobogán Interactivo” desarrollada y patentada por el CMTech.

No obstante, este sistema de seguimiento es generalizable a cualquier entorno en el que se desarrolle una actividad física intensa, o simplemente para cualquier otro tipo de proyecto que requiera de la detección y recolección de datos de diferentes cuerpos en un mismo entorno.

Resum

És per tots conegut l'efecte dels videojocs en els nens tant en aspectes socials com saludables, el present projecte forma part de l'àrea de recerca del grup CMTech del DTIC sobre la potenciació de l'activitat física en nens mitjançant *Exergames* (plataformes que fusionen videojocs i activitat física).

Es tracta de la implementació d'un sistema per detectar diferents cossos mitjançant l'ús d'una càmera i llums infraroges, a més, fer un recull del moviment dels mateixos (tracking) per saber des de la seva localització fins el recorregut que han fet en temps real. Aquest sistema millora sensiblement el seguiment de l'activitat física dels nens en una plataforma de *Exergames* anomenada "Tobogan Interactiu" desenvolupada i patentada pel CMTech.

Tanmateix, aquest sistema de seguiment és generalitzable a qualsevol entorn en què es desenvolupi una activitat física intensa, o simplement per a qualsevol altre tipus de projecte que requereixi de la detecció i recollida de dades de diferents cossos en un mateix entorn.

Abstract

It is known by all the negative effect of video games on children behavior, both social and healthy aspects, this project is part of the research area of CMTech group of DTIC about enhancing physical activity in children with exergames (platforms that fuse video games and physical activity).

This project is the implementation of a system for detecting specific tracking markers, using a vision system based on an infrared light in order to monitor the movement of participants to know their location and their travel in real time. This system significantly improves the tracking of physical activity of children in exergame platform called “Interactive Slide” deployed and patented by CMTech.

However, this monitoring system is generalizable to any environment in which an intense physical activity is developed, or just for any other project that requires the detection and collection of data from different objects in the same environment.

Prólogo

Muchos estudios remarcan que con el crecimiento de la industria de los videojuegos, ha aumentado el sedentarismo infantil, y con ello, ha aumentado también la obesidad infantil y disminuido su capacidad para socializarse. Algunos padres son partidarios de que esto suceda ante el hecho de que sus hijos estén fuera de casa sin supervisión de un adulto, además, apelan al aumento que proporcionan algunos videojuegos en cuanto a la destreza manual, la coordinación ojo-mano y la agudeza mental.

Para solucionar este fenómeno, se han diseñado y se están estudiando muy a fondo en los últimos años los *Exergames*, o videojuegos que requieren de actividad física individualmente o en grupo.

Una mejora necesaria para que este tipo de sistemas no caigan en la monotonía o sean predecibles y poco eficientes, es la detección en tiempo real de los cuerpos que participan en él, y la interacción del videojuego con dichos cuerpos. Y esto es lo que se realiza en este proyecto.

El sistema implementado en este proyecto consiste en que los usuarios se coloquen unos petos, cada uno con símbolos reflectantes diferentes, para posteriormente y mediante una cámara y luces infrarrojas poder diferenciarlos individualmente y poder registrar así su recorrido, las zonas que frecuentan más y su localización, todo en tiempo real.

Se tiene que recordar en todo momento que se trata de un sistema que reconoce cuerpos, o marcadores, y realiza un tracking¹ de los mismos, en ambientes de alta intensidad física, para poder recopilar datos de los cuerpos individualmente, por lo que se podría aplicar a cualquier ámbito en el que fuese necesario un sistema de estas características.

¹ Reconstrucción de la trayectoria dejada por el movimiento de un cuerpo a través de un entorno controlado.

Índice

	Pág.
Resumen.....	v
Prólogo.....	ix
Lista de figuras.....	xiii
Lista de tablas.....	xv
1. DEFINICIÓN DEL PROYECTO.....	1
1.1. Oportunidad, Antecedentes.....	1
1.2. Objetivos.....	2
1.3. Alcance.....	3
1.4. Calendario.....	4
1.5. Riesgos.....	5
1.6. Contribución.....	6
1.6.1. En conjunto.....	6
1.6.2. Objetivos alcanzados.....	7
2. ESTADO DEL ARTE.....	9
2.1. Actividad física para los niños.....	10
2.2. Campos de juego interactivos.....	13
2.3. Videojuegos para el ejercicio.....	14
2.4. Sistemas de visión.....	17
2.5. Tobogán interactivo.....	26
3. ANÁLISIS DE REQUERIMIENTOS.....	29
3.1. Requerimientos funcionales.....	29
3.2. Requerimientos no funcionales.....	30
4. DISEÑO.....	31
4.1. Estructura del sistema.....	31
4.2. Diseño final.....	32
4.3. Decisiones de diseño.....	34

5. IMPLEMENTACIÓN.....	35
5.1. Arquitectura del sistema.....	35
5.2. IDE y Librerías.....	35
5.3. Módulos del sistema.....	36
5.3.1. Extracción del fondo.....	37
5.3.2. Substracción de fondo y <i>Thresholding</i>	37
5.3.3. Detección de los <i>blobs</i>	38
5.3.4. Procesado de la substracción.....	39
5.3.5. Normalización.....	40
5.3.6. Identificación del objeto.....	41
5.3.6.1. Localización del objeto.....	42
5.3.6.2. Seguimiento del objeto.....	43
5.3.7. Refinamiento y salvado de datos.....	50
5.4. Diagrama de flujo.....	52
5.5. Diagrama de clases.....	54
6. CONCLUSIÓN.....	57
6.1. Trabajo futuro.....	58
6.2. Valoración personal.....	59
Bibliografía.....	61
ANEXO.....	63

Lista de Figuras

	Pág.
Fig. 1 - Sobrepeso y obesidad (adulta e infantil) ahora y en un futuro próximo.....	11
Fig. 2 - Porcentaje de obesidad adulta en las diferentes partes del mundo.....	11
Fig. 3 - Porcentaje de obesidad infantil en las diferentes partes del mundo.....	12
Fig. 4 - Boceto funcionamiento Exergame de Atari Joyboard.....	14
Fig. 5 - Dance Dance Revolution 1998.....	15
Fig. 6 - Running Exergame, expectativa vs realidad del sensor wiimote.....	15
Fig. 7 - El tobogán interactivo (izquierda) con el proyector, el sistema de computación, la cámara y luz infrarroja (esquina inferior derecha).....	17
Fig. 8 - Escena de gente pasando por la puerta de seguridad de un edificio.....	19
Fig. 9 - Funcionamiento <i>Background Substracción y Threshold</i>	20
Fig. 10 - Todas las representaciones de un objeto.....	24
Fig. 11 - Representación gráfica de las diferentes categorías de seguimiento estudiadas.....	26
Fig. 12 - Diagrama del Tobogán Interactivo: Frente, Lateral y Vista aérea.....	31
Fig. 13 - Petos con las diferentes formas reflectantes.....	32
Fig. 14 - Delimitación por zonas para distancia (izquierda) y esfuerzo físico (derecha).....	33
Fig. 15 - Arquitectura general del proyecto.....	35
Fig. 16 - Arquitectura de flujo de datos del sistema.....	36
Fig. 17 - Imagen reservada "fondo".....	37

Fig. 18 - <i>Background Substraction + Threshold</i> = imagen binaria.....	37
Fig. 19 - <i>Blob</i> detección, <i>bounding box</i> y extracción de imágenes del fotograma actual.....	38
Fig. 20 - <i>Bounding Box</i> antes y después de la modificación.....	39
Fig. 21 - Procesador de substracciones.....	39
Fig. 22 - Ejemplo de funcionamiento de algoritmo SVM.....	41
Fig. 23 - Funcionamiento procesado de identificación/posición del objeto.....	43
Fig. 24 - Resultados del algoritmo de seguimiento.....	50
Fig. 25 - Refinado de los datos previos a la interpolación para un mejor resultado.....	51

Lista de Tablas

	Pág.
Tabla 1 - Riesgos del proyecto con: identificador, impacto en el proyecto y estrategia para prevenirlos.....	5

1. DEFINICIÓN DEL PROYECTO

En este capítulo se hará una definición detallada de todo el proyecto, explicando las oportunidades, antecedentes, objetivos, alcance, calendario y riesgos del mismo.

1.1 Oportunidad, Antecedentes

El proyecto se origina en la necesidad de medir la actividad física individual de los niños que participan en el proyecto del tobogán interactivo (*Interactive Slide*²).

Debido a la creciente evolución de los *Exergames*³ y su mayor accesibilidad por parte de los usuarios que los necesitan, la necesidad de sistemas y métodos para calcular y monitorizar la cantidad de actividad física también aumenta [7, 8].

En el desarrollo de *Exergames*, los expertos han conseguido definir variables que relacionan la cantidad de actividad física con el ritmo cardíaco, y han podido implementar dichas variables con éxito. El único problema que se ha tenido, es que este tipo de sistemas implementados medían la actividad del grupo, en lugar de la actividad del usuario individualmente, por lo tanto, no se detectaba si se había dejado de participar y no se podía motivar al usuario para su continuidad en el ejercicio, haciendo al sistema menos interactivo y por lo tanto menos eficiente.

El reto tecnológico que se aborda en este proyecto final de grado, nos deja la oportunidad de poder implementar *Exergames* más completos y que por lo tanto:

- Sean mucho más accesibles a cualquier tipo de usuario
- Tengan una mayor interacción y que, por lo tanto, sean más eficientes
- Puedan tratar otros problemas de la sociedad actual como por ejemplo, la cada vez más creciente falta de socialización, y/o la exclusión que sufren algunos niños a temprana edad, y que afecta a su posterior entrada en sociedad

² <http://www.dtic.upf.edu/~npares/projectes/InteractiveSlide/InteractiveSlide.htm>

³ <http://en.wikipedia.org/wiki/Exergaming>

1.2. Objetivos

El objetivo principal de este proyecto es el de mejorar la experiencia interactiva de los usuarios en este tipo de plataformas, y la recopilación de datos por parte de los expertos para poder perfeccionar y ampliar sus sistemas.

Lo que se pretende conseguir es la medición de todos estos datos mencionados en tiempo real, poder utilizarlos junto con datos cardiovasculares y fórmulas predefinidas por los expertos en educación física para, también en tiempo real, hacer *Exergames* que interactúen con el usuario.

Otro de los objetivos es el de conseguir integrar este sistema en una plataforma existente, para poder realizar pruebas reales y conseguir resultados válidos para poder implantar este tipo de tecnologías y hacerlas útiles para su uso en escuelas, centros especializados o cualquier lugar donde sea necesario.

Un objetivo muy remarcado a lo largo de la implementación de este proyecto, ha sido el de diseñarlo de tal forma que sea escalable y reutilizable. Esto es debido a que no se pretende que su utilización se limite únicamente al desarrollo de *Exergames*, sino que toda la comunidad pueda utilizarlo en proyectos que requieran del seguimiento individual de cuerpos moviéndose (bruscamente o no) en un espacio controlado.

Por otra parte, lo que se busca en esta primera toma de contacto con este tipo de tecnología, es que su funcionamiento sea óptimo para los requerimientos especificados, y como en todo buen proyecto, queremos realizar avances paso a paso, por lo que principalmente buscamos la detección de los diferentes cuerpos, realizar el historial de su movimiento por el espacio delimitado por zonas de más o menos actividad, y la utilización de estos datos recopilados para mostrar datos de su actividad física a tiempo real.

Más adelante un objetivo clave será crear *Exergames* que utilicen este proyecto para mejorar la interacción con el usuario, poder implementar métodos de aprendizaje, y dinamizar la experiencia con el usuario.

1.3. Alcance





Primeramente, se realizó un estudio/comprensión de lo que se tenía que llevar a cabo en este proyecto, posteriormente se fueron estudiando los diferentes tipos de tecnologías que se utilizarían en el desarrollo de cada una de las partes, a continuación se fue desarrollando modulo a modulo todo el sistema:

- Extracción de fotogramas o *frames*⁴
- Diferenciación de las zonas de interés con el resto del *frame*
- Substracción, a partir de estas, de la zona de interés del *frame* original
- Tratamiento de la substracción
- Procesado mediante un modelo de entrenamiento vasado en *support vector machine*
- Implementación de la inteligencia artificial necesaria para controlar los diferentes factores de ambiente
- Integración en una interfaz para poder monitorizar y observar todos los datos obtenidos

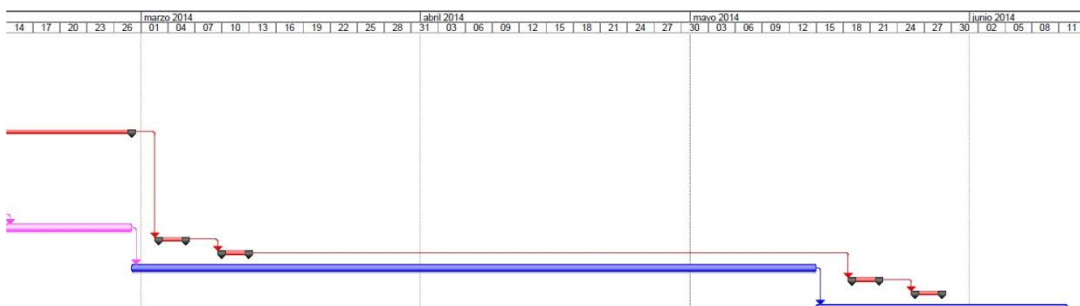
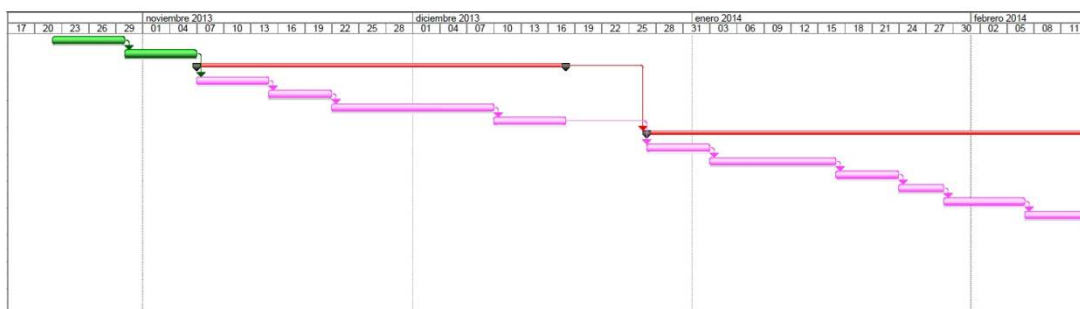
Finalmente, se llevó a cabo el refinamiento del software a base de pruebas estáticas con videos. Una vez el sistema era sostenible se dio un paso más allá probándolo en un escenario real, con cuerpos reales y variables que no controlábamos. Una vez testado y viendo los posibles errores más graves que podíamos tener, se pasó a la segunda fase de refinamiento para depurar dichos errores y hacer el sistema completo para su utilización.

⁴ <http://es.wikipedia.org/wiki/Frame>

1.4. Calendario

Proyecto: SISTEMA DE VISIÓN ARTIFICIAL PARA EL SEGUIMIENTO DE USUARIOS MEDIANTE MARCADORES E ILUMINACIÓN INFRARROJA EN ENTORNOS DE GRAN ACTIVIDAD FÍSICA	 Material sobre el que se trabajará.
	 Fase inicial.
	 Implementación.
	 Fase final.

Nombre de tarea	Duración	Comienzo	Fin
Comprensión del proyecto	6 días	mar 22/10/13	mar 29/10/13
Definición de los métodos	6 días	mié 30/10/13	mié 06/11/13
Trabajo con fotos	29 días	jue 07/11/13	mar 17/12/13
Background Substraction + Thresholding + Detección de Blobs	6 días	jue 07/11/13	jue 14/11/13
Substracción imágenes interes	5 días	vie 15/11/13	jue 21/11/13
Procesado de imágenes	12 días	vie 22/11/13	lun 09/12/13
Identificación con SVM	6 días	mar 10/12/13	mar 17/12/13
Trabajo con videos (tiempo real)	45 días	vie 27/12/13	jue 27/02/14
Detección cuerpos en movimiento	5 días	vie 27/12/13	jue 02/01/14
Algoritmo IA tracking de objetos	10 días	vie 03/01/14	jue 16/01/14
Refinado de los datos	5 días	vie 17/01/14	jue 23/01/14
Salvado y devolución de datos	3 días	vie 24/01/14	mar 28/01/14
Refinado código para convertirlo en Addon	7 días	mié 29/01/14	jue 06/02/14
Introducir Addon en Interfaz "Pascuza"	5 días	vie 07/02/14	jue 13/02/14
Últimos ajustes para pruebas en espacio real	10 días	vie 14/02/14	jue 27/02/14
Trabajo en espacio real	3 días	lun 03/03/14	mié 05/03/14
Trabajo en tiempo real (II)	3 días	lun 10/03/14	mié 12/03/14
Recopilación de información y mejora del código + Memoria	54 días	vie 28/02/14	mié 14/05/14
Trabajo en tiempo real(III)	3 días	lun 19/05/14	mié 21/05/14
Trabajo en tiempo real (IV)	3 días	lun 26/05/14	mié 28/05/14
Memoria	20 días	jue 15/05/14	mié 11/06/14



1.5. Riesgos

Los riesgos en este proyecto venían dados principalmente por el tiempo, la decisión del software adecuado para garantizar el funcionamiento en cada modulo, y sobre todo, el funcionamiento final también ha sido un aspecto crítico en este proyecto.

ID	Riesgo	Impacto	Estrategia
R1	Algún modulo anterior requiere de cambio en su lógica y/o implementación	Alto	Definir módulos pequeños y escalables que tengan un impacto mínimo entre sí.
R2	Escoger una tecnología inadecuada en una de las partes del proyecto	Alto	Estudio previo a fondo de la tecnología a utilizar teniendo en cuenta cualquier variable independiente que interactúe con ella.
R3	Falta de personal para realizar correctamente las pruebas en espacio real.	Alto	Planificación con mucha antelación y posible ampliación de personal antes de la prueba.
R4	Falta de conocimientos/tiempo para el desarrollo de algún modulo.	Alto	Consulta a expertos en la materia y disponibilidad de tecnología alternativa.
R5	Fallo en la devolución final de los datos.	Crítico	Testeo exhaustivo de la devolución de datos modulo a modulo.
R6	Confiar en las pruebas hechas con videos (fuera del espacio real).	Bajo	Extraer nuevo material para trabajar en la implementación aprovechando cada prueba en el espacio real.
R7	Posponer su prueba en el espacio real.	Crítico	No dejar nada al azar y planificar a la perfección la prueba.
R8	Rendimiento.	Crítico	Refinar cada módulo para que no afecte a las imágenes por segundo.

Tabla 2 - Riesgos del proyecto con: identificador, impacto en el proyecto y estrategia para prevenirlos

Es vital tener en cuenta el riesgo R5, debido a que un fallo de esta índole puede hacerte pensar que el sistema funciona bien, y cometer fallos a posteriori en la creación de módulos que interactúen con este.

El riesgo R7 es crítico debido a que las pruebas que requiere este proyecto en un espacio real son muy costosas a nivel de organización de preparativos y personal, es necesaria la participación de niños de primaria, un espacio muy grande (sala polivalente), que no está libre siempre, y un despliegue de medios bastante costoso.

El *frame rate*⁵ es un aspecto muy importante a tener en cuenta cuando se trata de trabajar con sistemas que funcionan a tiempo real, por eso el riesgo R8 es un punto clave a tener en cuenta si queremos que el proyecto funcione y se pueda implantar “fuera del laboratorio” en un futuro.

1.6. Contribución

En este apartado se expondrá la contribución aportada por este proyecto al mundo de los *Exergames* o de los sistemas que requieran de este tipo de funcionalidad.

1.6.1. En conjunto

En conjunto este proyecto ha aportado un sistema de visión escalable que puede compartirse en la comunidad de OpenFrameworks y que expone toda una lógica que se puede exportar a muchos otros campos.

Centrándonos en los *Exergames*, este sistema conseguirá que la funcionalidad de los mismos se mejore, que puedan tener aprendizaje automático⁶ y que puedan centrarse en otros aspectos y no solo en el del ejercicio físico.

Es un sistema de bajo costo y un rendimiento aceptable que puede identificar y hacer un seguimiento a un máximo de cuatro cuerpos en un espacio real, no obstante el sistema es escalable y con sencillas modificaciones se podría ampliar a más o menos cuerpos de ser necesario.

⁵ http://en.wikipedia.org/wiki/Frame_rate

⁶ http://es.wikipedia.org/wiki/Aprendizaje_autom%C3%A1tico

Además se ha conseguido un sistema no invasivo para monitorizar actividad física intensa a tiempo real, cosa que hace que el sistema tenga un valor añadido muy importante.

1.6.2. Objetivos alcanzados

Los objetivos alcanzados han sido los siguientes:

- Sustracción de imágenes de interés *frame a frame* utilizando el sistema de detección de *blobs* (región de la imagen diferente al fondo después de hacer *Background Subtraction*⁷ y *Thresholding*⁸)
- Generación de un modelo entrenado correctamente basado en SVM (Support Vector Machines) que pueda identificar las formas independientemente de su posición, deformidad y/o cantidad de movimiento
- Implementación de un pequeño algoritmo de inteligencia artificial que tenga en cuenta las variables que no controlamos:
 - Cruce de objetos
 - Desaparición del *blob* en la escena
 - Fallo del modelo de entrenamiento (solo para fallos leves)
- Tracking de los objetos y la posibilidad de obtener estos datos a tiempo real para poder tratarlos y hacer cambios en el flujo de la actividad

⁷ http://en.wikipedia.org/wiki/Background_subtraction

⁸ [http://en.wikipedia.org/wiki/Thresholding_\(image_processing\)](http://en.wikipedia.org/wiki/Thresholding_(image_processing))

2. ESTADO DEL ARTE

Uno de los pasos fundamentales previos al diseño de cualquier aplicación o sistema, es el estudio de todas las tecnologías y conocimientos que intervienen en o alrededor del sistema que quieres diseñar. Con este estudio previo se asegura la correcta comprensión de todo lo que envuelve al sistema y por lo tanto de una mejor implementación del mismo. A todo este estudio previo se le llama estudiar el estado del arte.

Para ello, se analizarán a fondo todo tipo de tecnologías y estudios que se relacionen con este proyecto. Los sistemas para controlar la actividad física de los niños, campos de juego interactivos, los *Exergames*, los sistemas de seguimiento por visión o Vision Tracking Systems y, cómo no, el tobogán interactivo creado por el grupo CMTech del DTIC de la universidad Pompeu Fabra.

A lo largo de la realización del estudio del estado del arte, se ha podido observar que en los últimos años y con la aparición de sistemas como el control remoto de Wii, PlayStation Eye, el Kinect de la Xbox, etc... Han aumentado los proyectos enfocados a la detección de cuerpos mediante sistemas de visión y la utilización de los mismos para crear *Exergames*.

Este fenómeno también se ha visto muy abordado debido a la facilidad por parte de los usuarios para adquirirlos, y a los aspectos negativos que pueden solucionar como la obesidad infantil o la falta de socialización provocada por la cada vez más creciente obsesión por realizar tareas estáticas frente a computadoras o videoconsolas.

2.1. Actividad física para los niños

A lo largo de este apartado se documentarán los estudios que hacen pensar que el enfoque del proyecto que se trata aquí es correcto y beneficioso para el futuro de mucha gente, y que además, puede ser beneficioso para que muchas personas puedan llevar una vida más cómoda una vez se pasa esa etapa en la que no se tiene en cuenta lo que pasara después, la infancia. Después se explicarán las tecnologías que hacen posible la medición de la actividad física de los niños o dicho de otra forma, la relación entre actividad física y ritmo cardíaco.

Primeramente, se ha estudiado el trabajo de un científico llamado Jules Hirsch [10], un Doctor en Medicina de The Rockefeller University, especializado en biología y en el campo de la conducta humana y su metabolismo. Hirsch ha realizado estudios en los que demuestra que nuestra conducta cuando somos niños es determinante no solo en nuestro metabolismo cuando somos adultos, sino en la dificultad que podemos tener al intentar cambiarlo en el futuro.

A nivel científico, cuando nos preguntamos porque una persona u objeto es más grande, nos viene a la mente la pregunta de si es porque las células son más grandes o simplemente porque tiene más células de lo normal, en el caso de las personas, es una combinación de ambas. El científico demostró los siguientes hechos:

- Las células de las personas con sobrepeso son mayores a las de las personas delgadas
- Cuando una persona sobrepasa los 27Kg de sobrepeso, el cuerpo genera más células deteniendo el crecimiento de las existentes
- En la niñez es cuando se determina generalmente el número de células grasas, es decir, un niño obeso tiende a multiplicar exponencialmente su número de células. Haciendo así que cuando esta persona sea adulta, este mayor número de células al hacerse de mayor tamaño generen una tendencia a la obesidad
- Las personas con un mayor número de células (obesidad infantil) tienen una dificultad mucho mayor al intentar perder peso que las personas que tienen células más grandes pero en menor cantidad (infancia sana)

No se está diciendo que si una persona ha tenido una infancia sana en la que no ha sufrido de obesidad infantil pueda llevar una vida insalubre en su mayoría de edad, sino que esta persona tendrá una mayor

calidad de vida pudiendo tener la libertad de no “vigilarse” tanto como una persona que sufrió de obesidad infantil. Y no solo estamos hablando de la posibilidad de tener una mejor calidad de vida, sino de evitar enfermedades como la diabetes tipo 2.

Como es bien sabido, la obesidad es una barrera que muy poca gente puede romper, algo que más vale prevenir, y que uno de sus principales enemigos es la actividad física. Como dicen en la página oficial de la organización mundial de la salud o World Health Organization (WHO)⁹, la malnutrición y falta de actividad física en un niño, es un error que se puede catalogar de abandono, puesto que para nosotros es muy simple de ver, pero hay que hacerles ver a ellos que el “mañana” es “hoy”.

Por todos estos estudios realizados que nos demuestran que el sobrepeso es un problema que nos afecta a todos y que si no se frena podría aumentar exponencialmente, sobretodo en Europa y EEUU (países consumidores de últimas tecnologías) he visto una oportunidad y un paso más hacia el

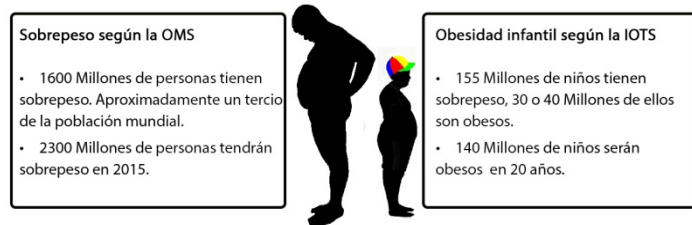


Figura 1 - Sobrepeso y obesidad (adulta e infantil) ahora y en un futuro próximo.

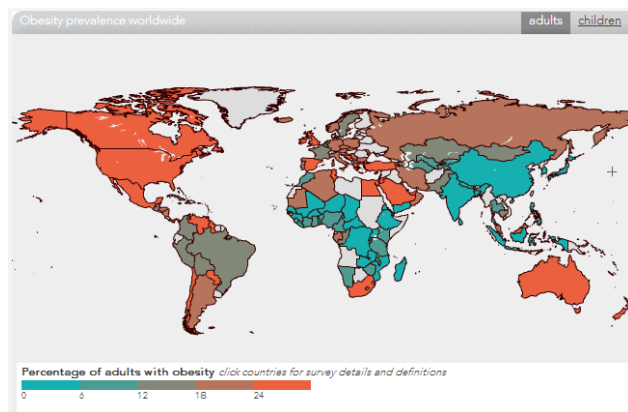


Figura 2 - Porcentaje de obesidad adulta en las diferentes partes del mundo.

⁹ <http://www.who.int/es/>

camino correcto, hacer un aporte a todo este tipo de tecnología para ayudar a reducir todas estas cifras. Una mejora en los sistemas que reenfozan lo que comprendíamos como sedentarismo (videojuegos) hacia actividad física para ayudar a prevenir el sobrepeso (*Exergames*) puede ser beneficiosa a nivel mundial.

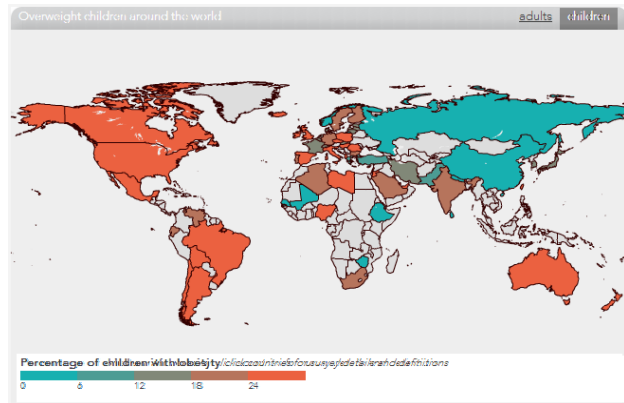


Figura 3 - Porcentaje de obesidad infantil en las diferentes partes del mundo.

Un factor fundamental para poder implementar estos sistemas es la correcta medición de la actividad física en los niños. Tener la posibilidad de variar y monitorizar esta actividad da la opción de crear *Exergames* con multitud de ejercicios diferentes, entretenidos y/o con diferentes ritmos de entrenamiento. Hasta hace unos años no se empezó a pensar en una variable para tener el control de la cantidad de actividad física, pero hace unos tres años el grupo CMTech del DTIC definió y demostró una variable que llamaron Tempo de Interacción o “Interaction Tempo”, con ella se consigue tener un control en la actividad física que llevan a cabo los niños en los experimentos.

Esta variable hoy en día ha demostrado que puede controlar la cantidad de actividad física o “amount of physical activity” (APA), pero no se ha podido utilizar al 100% puesto que los sistemas de visión actuales podían hacer un tracking aproximado del grupo de personas, con el sistema que se implementa en este proyecto se podrá hacer un tracking individual a tiempo real que explotará esta variable hasta llevarla a cualquier tipo de sistema en el que se necesite tener un control de la actividad física del usuario individual en un grupo de acción.

Así pues, este proyecto servirá para mejorar los sistemas que hacen que los niños o adultos aumenten su actividad física.

2.2. Campos de juego interactivos

Los campos de juego o parques han sido un motor para todos los niños del pasado, una herramienta con la que, utilizando la imaginación, podías ser o hacer casi cualquier cosa. En la actualidad, con toda la tecnología que hay al alcance de los niños, esta necesidad de potenciar la imaginación ha quedado en un segundo plano, por lo que los campos de juego como se conocían han quedado obsoletos, y los niños y adolescentes prefieren estar en casa jugando con videojuegos.

Para poder luchar contra esto, se idearon unos campos de juego que interactuaban con el usuario llamados campos de juego interactivos o *Interactive Playgrounds*¹⁰.

Principalmente estos campos de juego interactivos, eran fundamentalmente estructuras inflables con alguna finalidad deportiva: correr, saltar, deslizarse, llegar del punto A al punto B sorteando obstáculos, etc. Pero en la actualidad, y con la enorme necesidad de hacer que los niños realicen más actividad física, se ha ido un paso más allá, mezclando este tipo de estructuras con videojuegos. Esto aumenta exponencialmente su usabilidad, lo dinamiza y lo más importante de todo, lo hace más atractivo para los niños.

Con la utilización del sistema ideado en este proyecto, se irá más allá, haciendo que los campos de juego interactivos puedan serlo aún más si cabe, pudiendo aprender de el comportamiento de los usuarios que juegan en él y, por lo tanto, generando una interacción persona-máquina mucho más sofisticada.

¹⁰ <http://www.interactive-playground.com/> (ejemplo)

2.3. Videojuegos para el ejercicio

Como ya sabemos, y ya hemos comentado en puntos anteriores, desde la aparición de los videojuegos, la necesidad de salir a hacer ejercicio jugando para entretenerse se ha visto seriamente reducida gracias, entre otros, a la industria de la animación digital.

Para intentar solventar esto, en el mundo de la animación digital encontramos una variante llamada videojuegos para el ejercicio o *Exergames*, los cuales tratan de hacer que el usuario mueva todo el cuerpo mediante la utilización de ambientes interactivos con experiencias inmersas que simulan sensaciones de presencia. Los *Exergames* se han convertido en el enfoque tecnológico de la actividad física (AF) que consigue promover el ejercicio entre los usuarios mediante el juego y la competencia.

Este tipo de tecnología es, sobre todo, versátil. Se han hecho diversos estudios a todos los niveles, en niños, adolescentes y adultos de temprana y larga edad. Se han realizado *Exergames* dedicados a multitud de fines como el *fitness*¹¹, entretenimiento con ejercicio físico, mejora de movilidad, *running*, mejora de flexibilidad, etc. En general, todo lo basado en ejercicios cardiovasculares como aeróbic, flexibilidad y resistencia.

Una de las primeras interfaces de Exergame diseñadas fue la de Atari Joyboard [8] lanzado en 1982. Se trataba de unos sensores que se colocaban en el suelo en los que el usuario se subía y con los que haciendo presión en uno u en otro se podía simular el movimiento de un esquiador que se reflejaba en el monitor.

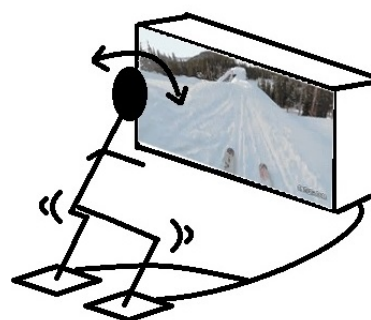


Figura 4 - Boceto funcionamiento Exergame de Atari Joyboard.

Sin embargo esta plataforma no tuvo mucho éxito, y salieron posteriores plataformas de las cuales la primera que empezó a gustar fue la de Dance Dance Revolution de Konami¹² (DDR), que se publicó 15 años después y que consistía en un juego rítmico

¹¹ <http://es.wikipedia.org/wiki/Fitness>

¹² <https://www.konami.com/ddr/>

que dispone de pulsadores en el suelo con forma de flechas (↑, ↓, ←, →) y un monitor en el que se muestra el juego, el cual consiste en unos símbolos (↑, ↓, ←, →) que pasan por una zona que podríamos llamar “de control”, justo cuando el símbolo pasa por esa zona has de pulsar el símbolo correspondiente en la zona de pulsadores.



Figura 5 - Dance Dance Revolution 1998

Una cosa curiosa es que simplemente fue diseñado como un juego interactivo en el que divertirse, no con la idea de ser usado como *Exergame* pero muy pronto se convirtió en una referencia para el resto de los videojuegos para el ejercicio.

Más recientemente y por todos conocido, Nintendo™, introduce la videoconsola Wii™, afirmando que dispone de una funcionalidad muy rica en actividad física, y que dispone de una amplia variedad de videojuegos deportivos. Sin embargo, debido al tipo de sensores utilizados hay muchas formas de realizar la actividad que te demanda el videojuego de una forma que no requiere de esfuerzo físico, y así se pierde el sentido de la experiencia interactiva, y con él la esencia de los *Exergames*.

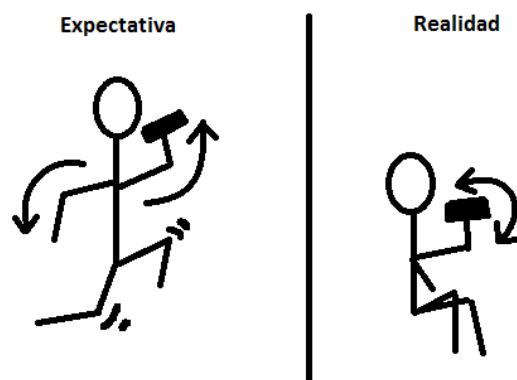


Figura 6 - Running Exergame, expectativa vs realidad del sensor wiimote

En los últimos 10 años se han hecho numerosas investigaciones en la rama de los *Exergames*, centrándose en la perspectiva de la salud, y se han encontrado tres aspectos principales que se han de medir para su correcta funcionalidad:

1. Si hay un gasto de energía significativamente más alto jugando *Exergames* que en reposo
2. Si el ejercicio efectuado con el *Exergame* equivale a 60 minutos de actividad física en niveles que van de moderada a fuerte
3. Si el *Exergame* puede ser utilizado como sustituto de las actividades físicas habituales

Teniendo en cuenta estas tres medidas, los estudios han demostrado que los *Exergames* aumentan significativamente el gasto energético, pero también, que no se llega a los niveles óptimos de ejercicio efectuado.

En cuanto al efecto de los *Exergames* a largo plazo solo ha habido un estudio que ha podido comprobar su efecto en 24 semanas, es el estudio de Maddison et al.¹³ [8], que afirma un cambio pequeño pero firme sobre el índice de masa corporal, y una disminución del sedentarismo de los niños con sobrepeso que realizaron el experimento.

Un aspecto muy importante de estos estudios fue la motivación de los participantes, Warburton et al., han demostrado que después de un programa de 6 semanas utilizando *Exergames*, una bicicleta estática con interfaz aumentaba un 30% la asistencia y registraba una mayor mejoría en la salud que una bicicleta estática convencional.

Chin et al., observaron en sus estudios, como la participación en este tipo de sistemas por parte de grupos mejoraba drásticamente todos los resultados, el doble de tiempo aplicado al juego, mayor participación, más rendimiento y mejoras no solo físicas sino también sociales. El problema que no pudieron resolver fue si se garantizaría la continuidad en la utilización de los *Exergames*, por lo que deducimos que tenemos que realizar juegos mejores, más dinámicos, con aprendizaje automático, y que garanticen una mayor y más larga utilización del *Exergame*. Como mínimo hay que garantizar su

¹³ http://www.wikilengua.org/index.php/et_al.

utilización ininterrumpida durante aproximadamente un año hasta que salga la siguiente versión del juego (aproximadamente una versión por año es lo habitual).

Así pues, la conclusión a la que llegamos es a la de que es necesario crear Exergames entretenidos, dinámicos, con aprendizaje automático y que se jueguen en grupo. Estamos de acuerdo en que el ser entretenido y dinámico depende del diseño pero en cuanto al aprendizaje automático y juego en grupo entran en juego otras tecnologías, entre ellas la tratada en este proyecto final de grado, el sistema de visión artificial para el seguimiento de usuarios o marcadores, es una pieza clave para que estos puntos se cumplan satisfactoriamente y poder explotar estas dos características al máximo.

2.4. Sistemas de visión

El sistema de visión o *Vision System* es un tipo de tecnología que se ha puesto muy de moda en los últimos años, se trata de sistemas de visión artificial, que permiten a un computador procesar imágenes que le llegan a través de una cámara y poder realizar diferentes funciones dependiendo de lo que se quiera conseguir.

En la actualidad hay multitud de implementaciones de sistemas de visión, a continuación se explicaran los más novedosos y que ayudaron a implementar el que se trata en este proyecto final de grado.

Seguridad

El seguimiento del flujo de gente a tiempo real es un recurso muy utilizado para aplicaciones de seguridad, turismo o marketing.

Una de las funcionalidades más comerciales de este tipo de sistemas es la de utilizarlos para temas de seguridad. Habitualmente se trata de la detección de cualquier cuerpo sin necesidad de identificar su identidad, sistemas que permiten saber si cualquier cosa esta entrando en una zona de control en un momento determinado.

A continuación se estudia un artículo [6] que explica un método para la detección de cuerpos y su recuento cuando pasan por una puerta. Es interesante no solo por el hecho de su amplia gama de posibilidades, sino porque su rendimiento es muy óptimo, se nos dice que es posible la detección de hasta 6 cuerpos a tiempo real con un promedio de 10 fotogramas por segundo con la utilización de un Pentium IV.

Este sistema consta de una cámara situada en el techo para controlar a la gente que pasa por una puerta de seguridad, la imagen captada por la cámara es llamada “región de seguimiento”, esta imagen es procesada y se calcula el número de personas que pasan por ella. También es delimitada una “línea de recuento” en la que se determina la gente que entra y sale por la puerta una vez estas personas la cruzan en una dirección o en otra.

El problema de recuento y seguimiento de las personas ha sido descompuesto en 3 fases:

1. Fase de alerta: determina si cualquier objeto importante ha entrado en escena
2. Fase de seguimiento: seguir su movimiento hasta que se alcance la línea de recuento
3. Fase de interpretación: Establecer cuantos de los objetos rastreados son personas.

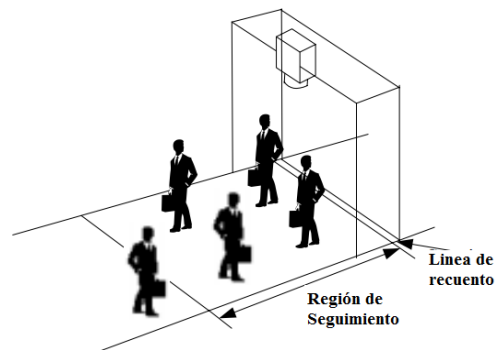


Figura 8 - Escena de gente pasando por la puerta de seguridad de un edificio.

Este método utiliza dos tecnologías empleadas en este proyecto y en la mayoría de sistemas de visión: la substracción de fondo (*Background Substraction*) y el método del valor umbral (*Thresholding*) [1].

El *Background Substraction*, como su propio nombre indica, es la substracción de un fondo. Se diferencian los pixeles de dos imágenes haciendo una resta de las mismas para identificar los píxeles que pertenecen al fondo y los que no.

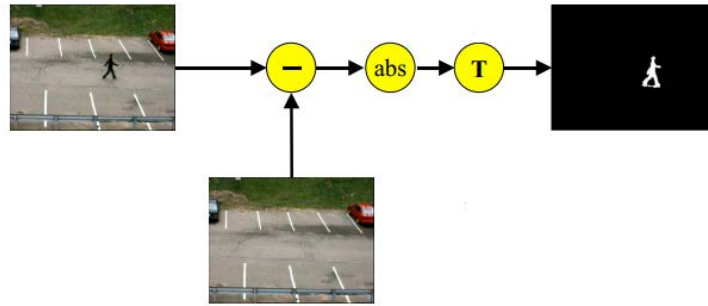


Figura 9 - Funcionamiento *Background Substraction* y *Threshold*

Y posteriormente el *Thresholding*, permite diferenciar los píxeles en la imagen resultante de la resta y coger los que más nos interesen a partir de un umbral que le indicamos, etiqueta con un 1 (objeto) y con un 0 (no objeto) para así generar una imagen binaria en la que podemos diferenciar claramente al objeto que entro en nuestro campo de visión, a estas imágenes se las llamamos *blobs*.

El siguiente paso del sistema es determinar lo que llamamos caja delimitadora o *bounding box* con el que se delimita un cuadrado o caja alrededor del objeto detectado en el paso anterior y con la que se puede realizar el seguimiento del objeto. Esta tecnología también es utilizada en este proyecto de final de grado.

Realizando estos pasos en muchos fotogramas se puede realizar un seguimiento muy preciso de un objeto, puesto que podemos basarnos en la distancia entre centroides de las cajas delimitadoras que se van determinando y así no tener una necesidad de conocer la identidad del objeto detectado para definir su recorrido.

Control de calidad en industrias

Para este apartado se ha estudiado una empresa innovadora y de las más importantes del mercado en sistemas de visión artificial dedicado al procesamiento de imágenes para el control de calidad, lectura de códigos e identificación de objetos, llamada *Industrial Vision System Ltd (IVS)*¹⁴. Fundada en 2000, se ha convertido en una de las compañías líderes en Europa en la proporción de sistemas de visión. Proporcionan los sistemas más innovadores y avanzados en el mercado y los distribuyen al mundo entero.

¹⁴ <http://www.industrialvision.co.uk/>

IVS es el ejemplo de esta tecnología llevada al límite, con maquinaria de alta precisión y software desarrollado para infinidad de posibilidades, la diferencia y el punto fuerte del tipo de sistema de visión que utiliza esta empresa, es su necesidad de conocer la identidad en todo momento de los objetos que están detectando.

La base es realmente parecida a la de este proyecto de final de grado, pero sin la necesidad de realizar un seguimiento del objeto.

No nos dan mucha información, puesto que su software es marca registrada, pero básicamente por lo que podemos leer en su documentación [16] nos hacemos a la idea de que se realiza una detección de la imagen que aparece en el fotograma, se tiene un modelo entrenado para la diferenciación de las formas detectadas, y se ejecutan unas funciones predeterminadas dependiendo del objeto que ha sido detectado.

En este tipo de sistemas se utiliza hardware de la más alta calidad, por lo que los factores de rendimiento de software no son tan decisivos como por ejemplo en el sistema que tratamos en este proyecto final de grado, en el que el hardware es de bajo costo y se ha de tener muy en cuenta el rendimiento.

Robots autónomos

Este tipo de tecnología surge de la necesidad del ser humano de acceder a sitios que a primera vista son inaccesibles, o para cumplir funciones de carácter social.

Estos sistemas se llaman neuro-difusos¹⁵, e incorporan las capacidades de las redes neuronales en cuanto a su capacidad de aprender y generalizar, utilizando sistemas de visión para reconocer el entorno.

Este estudio [5] puede ser tomado como punto de partida para futuras aplicaciones del sistema ideado en este proyecto final de grado. El sistema consta de una cámara, motor a pasos y el software MATLAB que procesa las imágenes y da las órdenes al motor, y pretende aprender de la escena que capta la cámara, y poder ajustar la posición angular de esta con respecto a la información visual proveniente de la escena detectada.

¹⁵ <http://taee.euitt.upm.es/actas/2010/papers/2010S13B01.pdf>

Este tipo de sistema de visión, es un primer paso hacia el desarrollo de un sistema de navegación para los robots autónomos. Es un claro ejemplo de lo que se pretende en los *Exergames* en un futuro, que el juego aprenda de su entorno y pueda interactuar más y mejor con los jugadores. Y esto se conseguirá con la incorporación de la tecnología implementada en este trabajo final de grado.

Conducción automática

Es un sueño desde la aparición del coche fantástico¹⁶ conseguir coches que conduzcan por si solos, ese sueño cada vez se está viendo más cerca gracias a numerosos estudios que intentan dar este paso. Para este punto me gustaría tratar uno de los más novedosos y actuales estudios sobre el tema, el coche auto conducido de Google [13].

Estos primeros prototipos que ha lanzado Google¹⁷ son un claro ejemplo de cómo la interacción de los sistemas de visión junto con sistemas de radar y otros componentes, pueden llegar a romper barreras tecnológicas. Barreras tecnológicas que ayudarían a que, por ejemplo, un invidente tenga una calidad de vida y una independencia que pocas veces habría soñado.

El prototipo de Google es una modificación de un Toyota Prius que utiliza una serie de sensores para navegar por la vía pública sin conductor.

Actualmente ya hay prototipos de coches auto conducidos de Google que circulan por algunas ciudades de Nevada. El coche auto conducido de Google dispone de los siguientes componentes:

- LIDAR: Un sensor de rotación en el techo que escanea más de 200 metros en todas las direcciones para generar un mapa tridimensional preciso de objetos alrededor del coche
- Cámara de Video: Una cámara montada cerca del espejo retrovisor detecta las luces de tráfico y ayuda al ordenador a bordo del coche a reconocer obstáculos en movimiento, como peatones y/o ciclistas
- Radar: Cuatro sensores de radar para automóviles estándar, tres de frente y uno en la parte trasera, ayudan a determinar las posiciones de los objetos distantes

¹⁶ http://es.wikipedia.org/wiki/Knight_Rider

¹⁷ http://en.wikipedia.org/wiki/Google_driverless_car

- Estimador de posición: Un sensor montado en la rueda trasera izquierda que mide los pequeños movimientos hechos por el coche para ayudar a localizar con precisión su posición en el mapa
- Receptor GPS
- Sensor de movimiento inercial

Actualmente, Google ya ha lanzado al mercado su primer coche automático¹⁸, es un poco diferente a sus prototipos más novedosos integrados en un coche convencional, se trata de un automóvil diseñado completamente por Google, electrónico y de pequeño tamaño, al cual mediante una aplicación le indicas donde quieres ir y el coche automáticamente se dirige al objetivo.

Seguimiento o Tracking de objetos

Esta es una parte importante dentro de la tecnología de visión por computador o sistemas de visión, y podría decirse que es la clave para este proyecto final de grado, y todos los sistemas de visión.

Con la potencia que han alcanzado las computadoras de hoy en día, las cámaras de vídeo a bajo coste y la creciente necesidad de análisis de video automatizado, se ha generado un gran interés por los algoritmos de seguimiento de objetos [17]. Hay tres pasos clave en el análisis de video:

- Detección de objetos interesantes en movimiento
- Seguimiento de dichos objetos fotograma a fotogramas
- Análisis del movimiento del objeto para reconocer su comportamiento

Por esto, el uso del seguimiento de objetos puede ser necesario en sistemas de visión para la detección automática de objetos, vigilancia automática en sistemas de seguridad, indexación de videos, interacción hombre-máquina (reconocimiento de gestos, mirada, etc.), monitorización del tráfico y navegación de vehículos.

La forma más simple de definir el seguimiento de objetos, sería la de: estimador de trayectoria de un objeto en el plano de la imagen, que se mueve alrededor de una escena. Aunque además de realizar el rastreo, estos algoritmos pueden proporcionar

¹⁸ <http://appskyworld.com/2014/05/google-presenta-el-primer-automovil-sin-conductor/>

información del objeto (forma, tamaño, orientación). El seguimiento de objetos puede ser complejo por los siguientes factores:

- Pérdida de información por el paso del mundo 3D a una imagen 2D
- Ruido de la imagen
- Movimiento complejo de un objeto
- Naturaleza o ritmo del articulado de los objetos
- Desaparición parcial o total del objeto en la imagen (se esconde detrás de otro objeto o parecido)
- Cambios en la iluminación de la escena
- Requisitos de procesamiento en tiempo real

Dos de las tareas imprescindibles a realizar con el objeto cuando lo estamos detectando, y que resuelven muchas de los factores de riesgo especificados anteriormente son:

- Construir un modelo de lo que queremos detectar (necesario si queremos identificar objetos específicos)
- Usar la información extraída acerca del objeto del fotograma anterior para hacer predicciones sobre el fotograma actual y restringir la búsqueda

Durante la implementación de todo este proyecto final de grado, efectivamente se han encontrado todos estos factores que se han tenido que resolver, ya sea con condicionantes del experimento (luz, ángulo de la cámara y normativas de comportamiento) o con implementaciones internas a base de testeo (variable de aproximación de cuerpos, división de la escena por zonas, implementación de una pequeña parte de inteligencia artificial y creación de modelo entrenado para identificar los diferentes objetos).

La representación de objetos la podríamos definir como la suma entre la forma y la apariencia del mismo. Hay diferentes representaciones para la forma y también para la apariencia del objeto, a continuación presentaremos todas las variantes:

Representaciones de la forma del objeto:

- **Puntos:** El objeto es representado por un punto (el centroide) o por un conjunto de puntos. Esta forma de representación es adecuada para objetos que ocupan una pequeña región de la imagen (a)(b)

- **Formas geométricas primitivas:**

La forma del objeto se representa con un rectángulo, elipse, etc. Aunque esta representación es más adecuada para objetos rígidos simples, también es utilizada para el seguimiento de objetos no rígidos (c)(d)

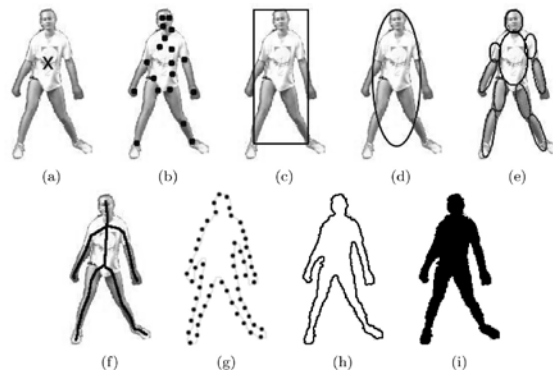


Figura 10 - Todas las representaciones de un objeto.

- **Siluetas de objetos y contorno:** El contorno es el perímetro de un objeto, y la silueta es la región dentro de ese contorno. Este tipo de representación es adecuado para hacer seguimiento de formas no rígidas y complejas (g)(h)(i)
- **Modelos de formas articuladas:** Los objetos articulados se componen de partes del cuerpo, modelados por cilindros o elipses, que se mantienen unidos por la articulación (como un cuerpo humano) (e)
- **Modelo de esqueleto:** Esqueleto del objeto se puede extraer aplicando la transformación del eje medial en la silueta del objeto. Esta representación se puede utilizar para modelar objetos tanto articulados como rígidos (f)

Representación de la apariencia de un objeto:

- **Plantillas:** Construcción con formas o siluetas geométricas simples. Adecuado para objetos de seguimiento cuyas posturas no varían considerablemente durante el curso del seguimiento
- **Densidades de probabilidad de la apariencia del objeto:** Las estimaciones de densidad de probabilidad de la apariencia del objeto pueden ser paramétricas o no paramétricas
- **Modelos de apariencia activa:** Se genera mediante el modelado de forma simultánea de la forma del objeto y su apariencia
- **Modelos de apariencia multi-vista:** Codifican diferentes vistas de un objeto

Depende mucho del escenario en el que nos encontramos, pero para una correcta elección de la forma de representación de un objeto, y por lo tanto la correcta elección de las técnicas para representar su forma y su apariencia son:

- Tamaño de los objetos a los que hay que realizar el seguimiento
- Dificultad de la forma del objeto

Hay una serie de características que se han de tener en cuenta y estudiar a la hora de realizar el seguimiento de un objeto que son: el color, los bordes, la textura y el flujo óptico. Un buen estudio de estas características ayudará a obtener una mayor precisión en la detección del objeto bajo cambios de iluminación o intentos de confundirse con el entorno.

Todos los métodos de seguimiento requieren un mecanismo de detección del objeto, hay tres mecanismos que son considerados los más populares:

- **Detectores de puntos:** Encuentra puntos interesantes en imágenes que tienen una textura expresiva en sus respectivas localidades.
- **Segmentación:** Partición de la imagen en regiones porcentualmente similares.
- **Substracción de fondo (*Background Substraction*):** Ya explicado anteriormente y utilizado en este proyecto final de grado, es la diferencia entre una imagen que se toma de referencia (el fondo vacío por ejemplo) y la de los fotogramas que vienen a continuación, con lo que se consigue abstraer los píxeles que no concuerdan, y así detectar el objeto que ha invadido la escena.

Utilizando todo este estudio y tecnología previamente explicada, se puede finalizar con el seguimiento o tracking del objeto, el cual consiste en establecer correspondencia entre instancias de objetos que se encuentran en los diferentes fotogramas de la escena. Dicha tarea se puede realizar por separado, ejecutando los algoritmos de detección de objetos, almacenando los datos y haciendo un rastreo a posteriori de los datos, o conjuntamente, se calcula iterativamente la localización del objeto y se calcula la correspondencia a tiempo real.

Las tres principales categorías de seguimiento son las siguientes:

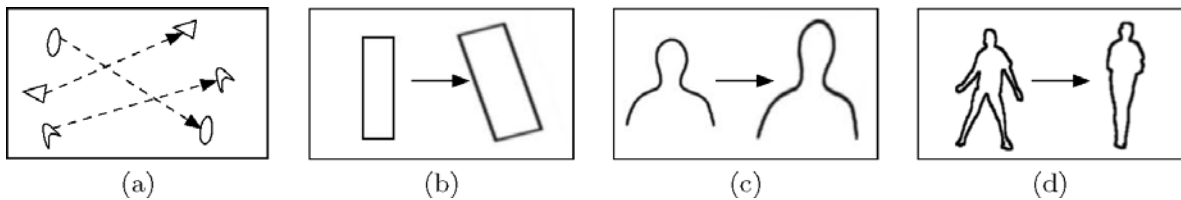


Figura 11 - Representación gráfica de las diferentes categorías de seguimiento estudiadas.

- **Seguimiento del punto (a):** Los objetos detectados en la consecución de los fotogramas son representados por puntos, y se realiza una coincidencia de puntos. Este enfoque requiere de un mecanismo externo para detectar los objetos en cada fotograma
- **Seguimiento del Kernel (b):** Kernel hace referencia a la forma y apariencia del objeto. Ej. Kernel puede ser una plantilla rectangular o una forma elíptica como un histograma asociado. Los objetos son seguidos mediante el cálculo del movimiento del Kernel en fotogramas consecutivos
- **Seguimiento de la silueta (c) + (d):** Estos métodos utilizan la información codificada dentro del objeto (apariencia, densidad y modelos de la forma). Dado el modelo de objeto, la silueta son seguidas por cualquier forma coincidente (c) o la evolución del contorno (d)

2.5. Tobogán interactivo

El tobogán interactivo o *Interactive Slide* es un proyecto del departamento de estudios audiovisuales de la Universidad Pompeu Fabra. Mezcla todos los conceptos explicados anteriormente, es un campo de juego interactivo para el desarrollo de actividad física para niños en el que se proyectan *Exergames* y se utilizan sistemas de visión.



Figura 7 - El tobogán interactivo (izquierda) con el proyector, el sistema de computación, la cámara y luz infrarroja (esquina inferior derecha)

Como se puede apreciar en la Figura 7 es un tobogán inflable, en el que se le proyectan juegos para el ejercicio, consta de una superficie deslizante de 6.5 metros de largo x 8 metros de ancho y 3.85 metros de altura, que actúa también como pantalla de proyección de imágenes. Un sistema de visión por ordenador detecta los movimientos de los niños que se deslizan por la pantalla para realizar acciones en el juego.

Este proyecto pretende resolver dos de los problemas que también hemos comentado anteriormente, el sobrepeso infantil y la falta de socialización de los niños en la actualidad.

Cumple con todos los requisitos mencionados hasta ahora:

- Es dinámico, puesto que se pueden proyectar en el infinito de juegos
- Se juega en grupo, preparado para 4 niños, por lo que gracias a estudios anteriores sabemos que el rendimiento de los niños será mayor y su participación más continua y activa
- Entretenido y novedoso
- Utiliza tecnología adecuada para poder correr *Exergames* con aprendizaje automático y seguidor de usuarios y marcadores infrarrojos

En el tobogán interactivo, se ha validado su medida de la actividad física, sin embargo, esta medida se hace a partir de la actividad del grupo en lugar de la actividad de cada niño en particular [7, 8]. Por lo tanto, el sistema no detecta si uno de los niños ha dejado de participar activamente, y de suceder esto, no puede indicar al niño que vuelva a jugar. Este proyecto final de grado pretende justamente corregir esto, se podrá detectar la actividad y hacer un seguimiento o tracking de cada niño individualmente, haciendo así que los datos sean más precisos, los juegos más completos y que las posibilidades crezcan. Además esto funcionara a tiempo real, con unos fps (fotogramas por segundo) altos y con una exactitud aceptable.

3. ANÁLISIS DE REQUERIMIENTOS

En este capítulo se explicarán los servicios que tiene que tener el sistema y sus restricciones. Se definirá lo que se espera que tiene que hacer el sistema: interacción entre el sistema y el entorno, y detalle de los servicios y funciones que debe tener. También se definirá cómo tiene que ser el sistema: se describirán restricciones que limitaran las elecciones para su implementación (atributos relacionados con la calidad)

3.1. Requerimientos funcionales

- El sistema deberá ser capaz de detectar cuatro formas diferentes y poder hacer un seguimiento de las mismas
- El sistema deberá almacenar los datos de seguimiento de los usuarios
- El sistema deberá hacer una devolución a tiempo real de las posiciones de todos los usuarios
- El sistema deberá ser capaz de hacer un tratado de un video fotograma a fotograma extrayendo las imágenes de interés
- Deberá también poder modificar estas imágenes para pasarlas al modelo de entrenamiento y que este pueda detectar de qué imagen se trata
- El sistema deberá tener en cuenta cuando hemos perdido a un objeto, y definir por donde ha pasado y a qué velocidad ha ido desde esa posición hasta la “actual”
- Devolución cada 10 segundos del seguimiento del usuario, para poder hacer una estimación de su actividad física y conseguir que el juego interactúe con él
- El sistema no deberá verse afectado por actividades físicas intensas

3.2. Requerimientos no funcionales

- El programa estará programado en C++
- Se utilizarán librerías de `openCv`¹⁹ para el tratamiento de imágenes
- El sistema deberá tener el rendimiento suficiente para hacer el tratado de fotogramas, procesado de imágenes y recopilación de datos en tiempo real
- El software tiene que estar implementado de una forma compacta, es decir, tiene que funcionar como un `addon`²⁰, llamado por una función a la que le pasamos todas las variables de interés
- Se realizará interpolación en los puntos de la recopilación de seguimiento de los objetos cuando este se haya perdido
- El sistema deberá introducir un pequeño algoritmo de “inteligencia artificial” para posibles problemas de ruido en la imagen, pérdida de objetos, movimientos inesperados de los objetos y en general, cualquier factor que pueda alterar la recopilación de datos
- El programa estará desarrollado en el IDE de *CodeBlocks*²¹ basado en *OpenFrameworks*²², por lo que tendrá que seguir las buenas prácticas que exige la comunidad *OpenFrameworks*
- El modelo entrenado funcionará con librerías y un algoritmo SVM²³ (*Support Vector Machine*), pasándole una imagen devolverá un vector²⁴ de 5 posiciones, identificando en cada posición, el porcentaje que tiene la imagen de ser de la forma 1, 2, 3, 4 o ninguna
- El sistema deberá ser lo más preciso posible, y no deberá ser sensible a datos inesperados, lo procesará e intentará interpretarlos

¹⁹ <http://opencv.org/>

²⁰ <http://ofxaddons.com/>

²¹ <http://www.codeblocks.org/>

²² <http://www.openframeworks.cc/>

²³ [http://www.csie.ntu.edu.tw/~cjlin/libsvm/\[9\]](http://www.csie.ntu.edu.tw/~cjlin/libsvm/[9])

²⁴ [http://es.wikipedia.org/wiki/Vector_\(inform%C3%A1tica\)](http://es.wikipedia.org/wiki/Vector_(inform%C3%A1tica))

4. DISEÑO

En este capítulo se explicara la estructura que debe tener el diseño para que sea funcional y pueda integrarse con el tobogán interactivo, como ha quedado el diseño definitivamente y finalmente, las decisiones que se han tenido que tomar para llegar a las conclusiones finales.

4.1. Estructura del sistema

Como planteamiento inicial se explicará la estructura para la que se tenía que adaptar este proyecto [14, 15]:

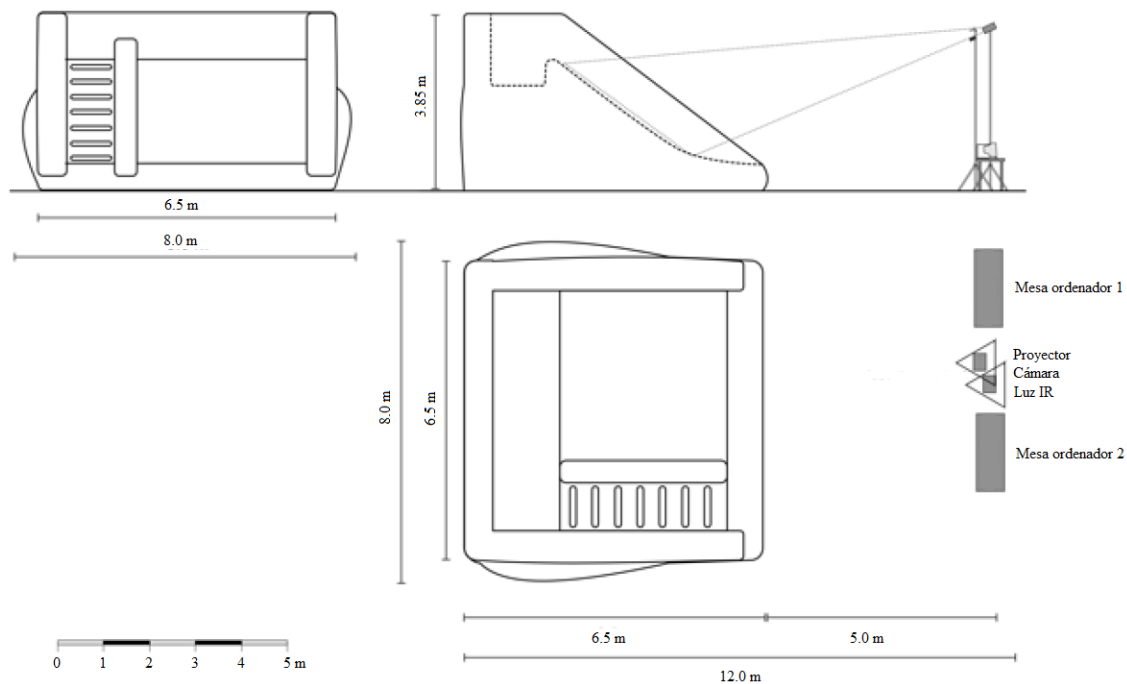


Figura 12 - Diagrama del Tobogán Interactivo: Frente, Lateral y Vista aérea.

Se trata de un tobogán inflable de 3,85 metros de alto y 6,5m de ancho y largo, el proyector, cámaras y la luz infrarroja están situadas a unos 5 metros del tobogán. El experimento está enfocado a niños, y participan 4 a la misma vez en el juego. Por estos condicionantes se plantean los siguientes factores a tener en cuenta:

- Se tienen que detectar formas que están, la mayor parte del tiempo, a una media de 9 metros. Puesto que es la zona donde se preparan los usuarios para deslizarse por el tobogán

- Al tratarse de un experimento enfocado y testeado por niños, se necesitan formas no invasivas de detectarlos y hacer un seguimiento de sus movimientos, para que sean más proactivos en el ensayo
- Los objetos a reconocer han de ser de unas formas específicas que no hagan caer en errores a la hora de definir de cual se trata, además el modelo que entrenemos ha de ser lo más robusto posible para tener en cuenta todas las posibles formas del objeto durante el ensayo (niños de costado, mezcla de objetos en imagen 2D o torsión del cuerpo)
- Hay que tener en cuenta las diferentes zonas de detección, puesto que hay dos factores clave que afectan de diferente forma a la recopilación de datos:
 1. La distancia entre la cámara y las diferentes zonas
 2. La cantidad de actividad física que se ejerce en las diferentes zonas

4.2. Diseño final

Para resolver todos los factores planteados en el apartado anterior, el diseño resultante fue el siguiente:

Para detectar objetos a grandes distancias se decidió aprovechar la luz infrarroja, con la que, colocando formas reflectantes en los participantes del experimento, se conseguía una forma nítida y compacta de píxeles al hacer el *Background Substraction* y el *Thresholding*.

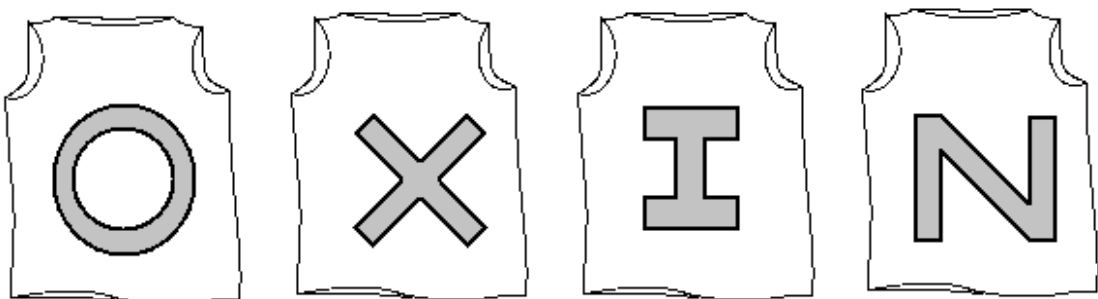


Figura 13 - Petos con las diferentes formas reflectantes.

Para evitar que la forma de detección de los usuarios sea invasiva, se decidió poner las formas reflectantes en la parte delantera y trasera de unos petos deportivos (Figura 13).

Las formas escogidas, han sido O, X, I y N puesto que después de numerosos intentos han sido las formas que mejor han reaccionado con el modelo de entrenamiento.

Finalmente, se decidió hacer dos delimitaciones distintas de la imagen captada del tobogán, una para solventar el problema de distancia entre cámara y objetivos y otra para delimitar las aéreas en las que se realiza un mayor o un menor esfuerzo físico.

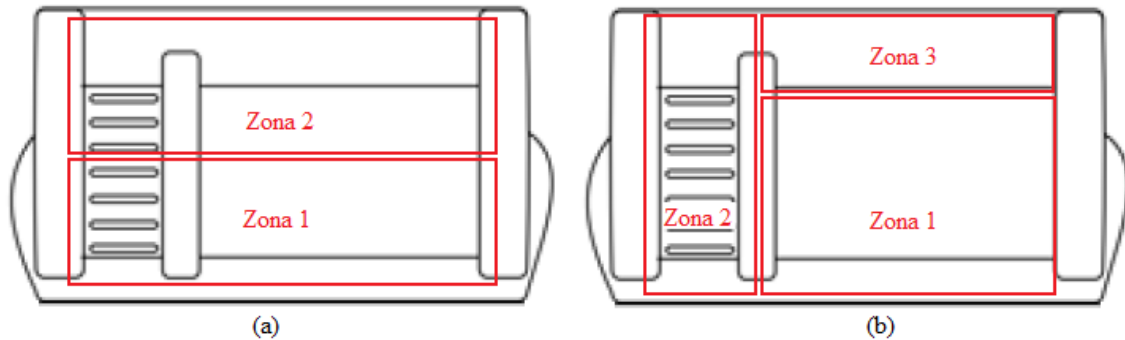


Figura 14 - Delimitación por zonas para distancia (izquierda) y esfuerzo físico (derecha)

Para la delimitación dedicada a la distancia de los objetos **(a)**, se decidió hacer una división en dos zonas, la primera se trata de los objetos que están a una media de entre 5 y 6 metros y son los que consideramos “objetos cercanos”. En la zona 2 es donde se determinó que deben estar los “objetos lejanos”, y suelen estar a una distancia de entre 8 a 9 metros. Con estas delimitaciones se puede ser más preciso con la distancia entre objetos (1 metro en la zona 2 corresponde a menos píxeles que en la zona 1).

Para la delimitación dedicada al ejercicio físico de los niños **(b)**, se ha hecho un estudio más a fondo con expertos en educación física, que ha determinado las 3 zonas de acción que encontramos en el tobogán, se determinó que la zona 2 es la zona en la que los niños realiza el mayor esfuerzo físico, la zona 3 es considerada la zona de descanso puesto que es en la que los niños están más estáticos, y la zona 1 se considera una zona intermedia puesto que los niños se deslizan y corren hacia las escaleras para volver a subir.

4.3. Decisiones de diseño

Las decisiones tomadas en el diseño de este proyecto se tomaron en base a la practicidad, el limitado presupuesto y la comodidad de los participantes.

Creemos haber tomado las decisiones correctas que hacen que estas tres características se cumplan.

Sabemos que con una mayor calidad de los componentes utilizados, y quizás con la integración de algún tipo de tecnología de radar, este proyecto sería algo más preciso en los resultados y su rendimiento sería mayor, pero la diferencia es menos apreciable si hacemos un balance en cuanto a gastos-mejoras. Por todo esto, se cree que este diseño es el adecuado, y que con mejoras gracias a la continuidad de trabajo en el mismo se conseguiría que fuese igual o más preciso que sistemas mucho más costosos.

5. IMPLEMENTACIÓN

En este capítulo se explicará en profundidad todas las partes de la implementación del software de este proyecto. Se explicarán todas las tecnologías utilizadas y la interacción entre ellas.

5.1. Arquitectura del sistema

Se comenzará explicando la arquitectura general del sistema y a continuación detallaremos cada una de sus partes en profundidad.

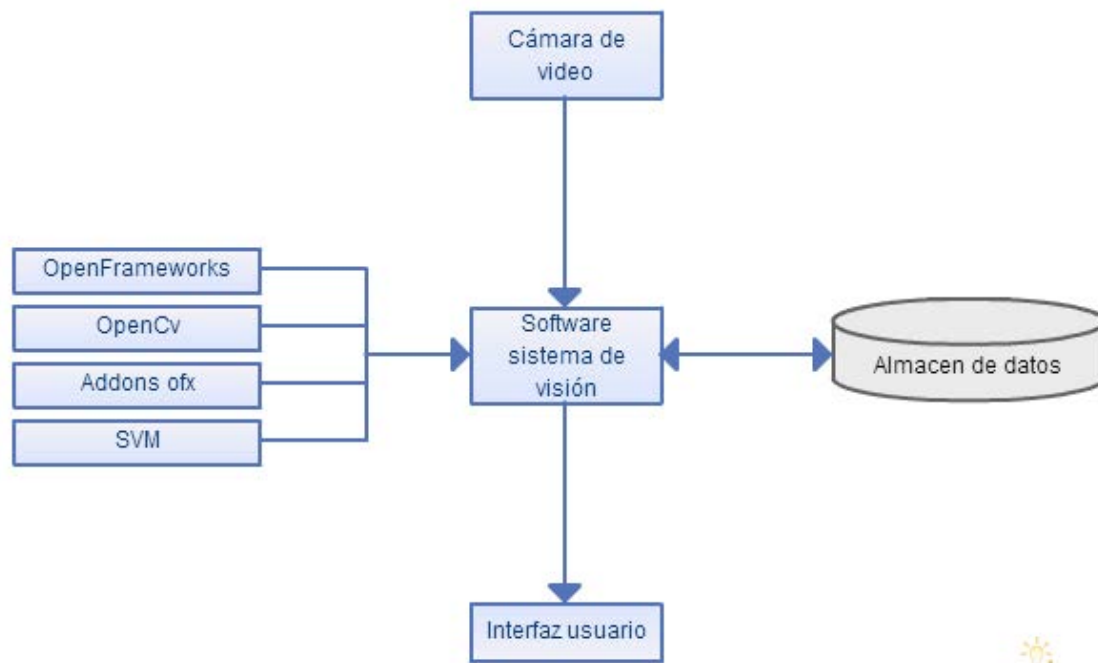


Figura 15 - Arquitectura general del proyecto.

5.2. IDE y Librerías

En cuanto a la arquitectura del sistema, se ha programado en C++, y se ha utilizado el IDE *Code::Blocks*. Para su implementación se han utilizado las librerías de *OpenFrameworks*, algunas de *OpenCv* [11], y *SVM* [9] para la parte de *machine learning* (identificación de objetos). Al software de sistema de visión le llegan las

imágenes de la cámara de video, las procesa, y almacena datos acerca de las imágenes procesadas, que utiliza para devolverlos a la interfaz o generar estadísticas.

5.3. Módulos del sistema

Profundizando un poco más, podemos saber el proceso que sigue la arquitectura especificada para el procesamiento de los datos. La arquitectura de flujo de datos sería la siguiente:

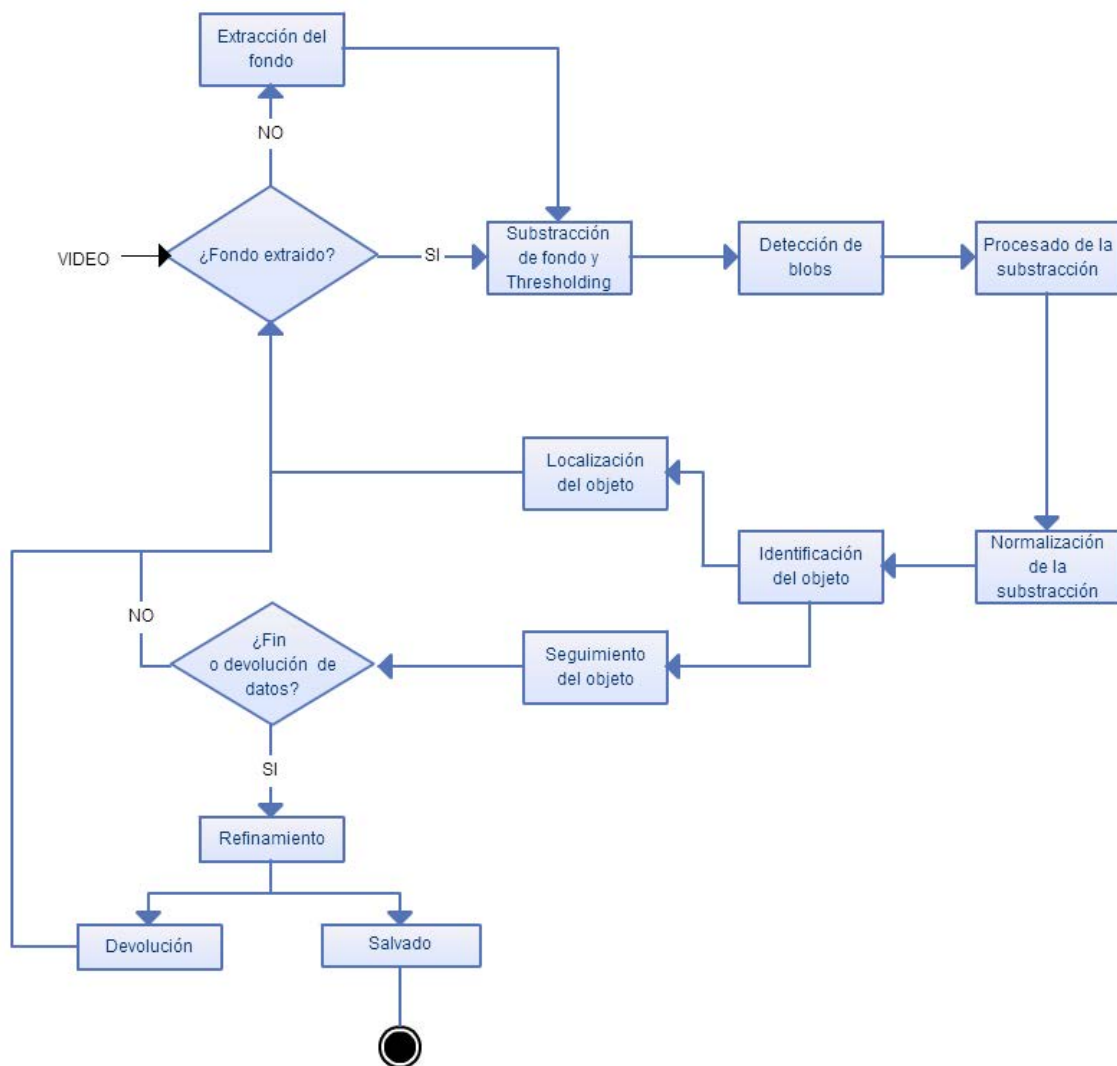


Figura 16 - Arquitectura de flujo de datos del sistema.

5.3.1. Extracción del fondo

Lo primero que se encuentra el sistema es la imagen proveniente del video, en una primera instancia, extraemos la imagen que consideraremos “fondo” y la reservamos para poder hacer el *Background Subtraction* a lo largo de todo el funcionamiento del sistema.

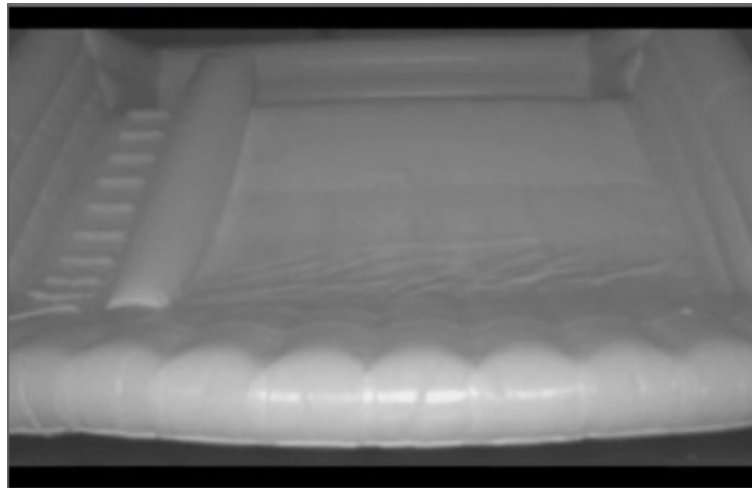


Figura 17 - Imagen reservada "fondo".

5.3.2. Substracción de fondo y *Thresholding*

El siguiente paso es el de realizar el *Background Subtraction* entre la imagen de fondo y el siguiente fotograma y de el resultado hacer el *Thresholding* para obtener una imagen binaria y así poder identificar más fácilmente los píxeles que han cambiado de un fotograma a otro.

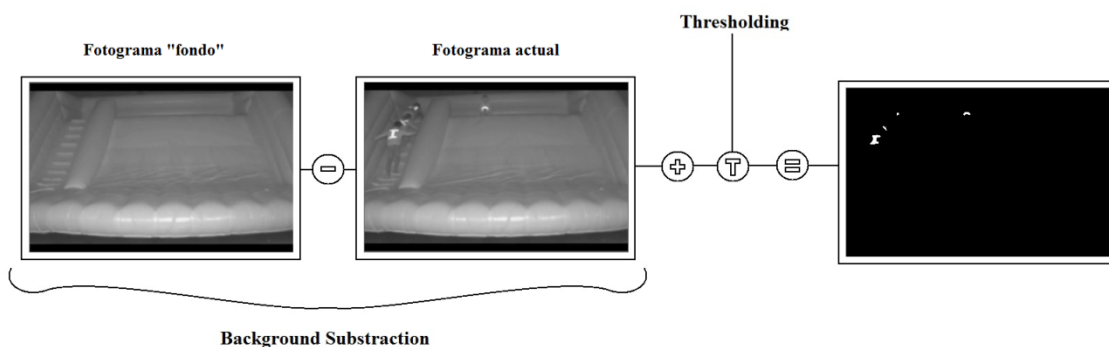


Figura 18 - *Background Subtraction* + *Thresholding* = imagen binaria

5.3.3. Detección de los *blobs*

El siguiente paso es realizar lo que llamamos *Blob Detection*²⁵, este método se encarga de localizar y diferenciar las zonas más claras o más oscuras de la imagen que se le proporciona. Con la imagen binaria que se le pasa, esta tarea se le hace mucho más fácil, y podemos detectar las formas o “*blobs*” que encontramos en la imagen.

Posteriormente se realiza el *Bounding Box* de los *blobs* detectados, con este proceso podemos enmarcar en un cuadrado el *blob* detectado, pudiendo saber así su altura, posición en la imagen centroide, área, etc. Así también tenemos un marco que podemos utilizar para extraer la imagen del fotograma actual.

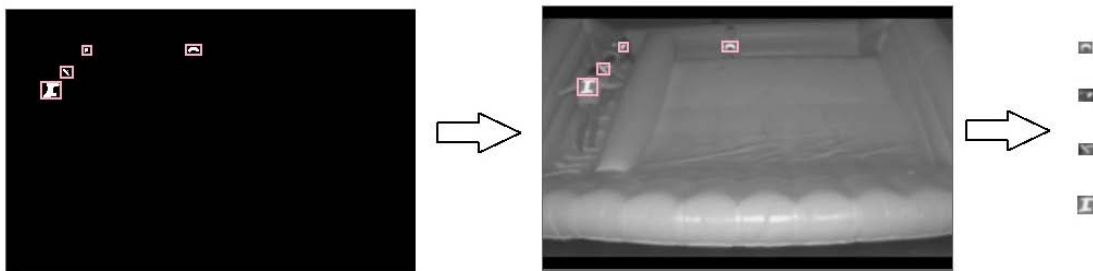


Figura 19 - *Blob Detection*, *Bounding Box* y extracción de imágenes del fotograma actual.

Cabe decir que para realizar este proceso se utiliza una función llamada “*findContours*”²⁶ de la librería *ofxOpenCv*²⁷, y que se tuvo que modificar internamente para que fuera más precisa a la hora de recopilar las imágenes que posteriormente se tenían que procesar.

Modificación: había muchas veces en las que al detectar una forma se detectaban varios “*blobs*” debido a que, por ejemplo, el brazo del niño cruzaba la imagen captada por la cámara, detectando así la forma en dos partes distintas, o dicho de otra forma, dos *blobs*.

²⁵ http://en.wikipedia.org/wiki/Blob_detection

²⁶ http://openframeworks.cc/documentation/ofxCvContourFinder.html#show_findContours

²⁷ <http://openframeworks.cc/documentation/>

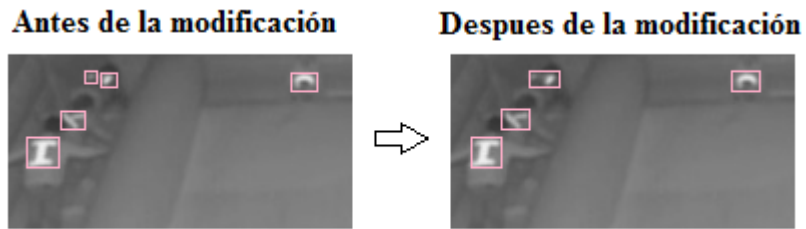


Figura 20 - *Bounding Box* antes y después de la modificación.

Así que se añadió otra variable a los parámetros de la función *findContours* llamada “maxDistance”, con la que se determinó que cuando los centroides de dos *blobs* se encontraran a una distancia inferior a dicha variable, estos se unieran en uno solo, así, al hacer la substracción de la imagen, cogemos la forma completa, y no la misma forma en dos partes.

5.3.4. Procesado de la substracción

Una vez se tienen las imágenes del fotograma actual o substracciones, dicho de otro modo, las imágenes pertenecientes a los objetos de los cuales queremos saber la identidad, las procesamos.

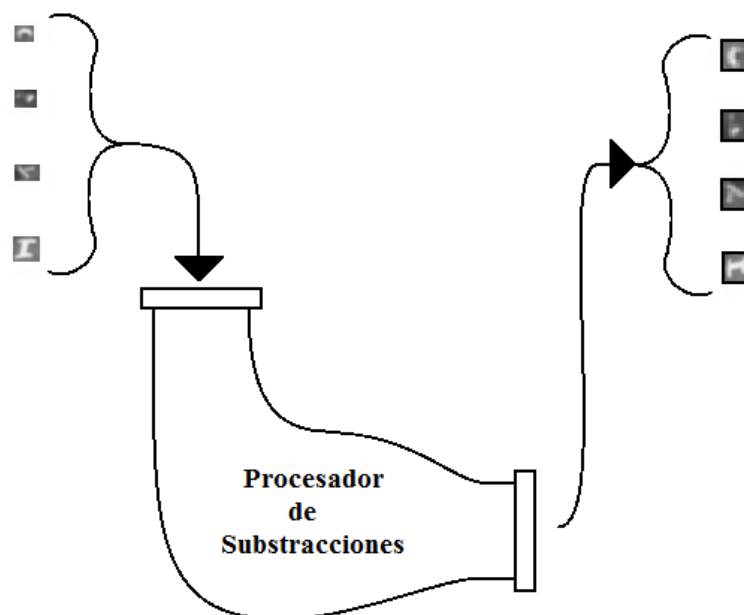


Figura 21 - Procesador de substracciones

Lo que hace este procesador de subtracciones es:

- Hace un *resize* o redimensionado de la subtracción a 100 píxeles x100 píxeles
- Añade un borde negro de 5 x 5 píxeles
- Redimensionado de el conjunto (105 x105 píxeles) a 30 x 30 píxeles
- Realiza una rotación de 90° a la derecha
- Realiza un espejo (volteo horizontal) de la imagen²⁸

Para que el modelo entrenado pueda ser efectivo se tienen que cumplir todas las condiciones con las que fue entrenado. Este procesado se ejecuta debido a que el entrenamiento del modelo para el algoritmo SVM se realizó en Matlab, con la característica de ser imágenes de 100x100 píxeles, con un borde negro extra de 5 píxeles, todo redimensionado a 30x30 píxeles (tres primeras operaciones del procesador), también, la manera de vectorizar²⁹ las matrices respecto a openCv es diferente, por lo que se hacen las dos últimas operaciones para “ordenar” los píxeles.

5.3.5. Normalización

Después de realizar el procesado de la imagen, se normaliza, esto lo realiza una clase llamada “Normalizer”, y lo que realiza es lo siguiente:

$$\text{Para todos los píxeles de la imagen: } pixel(i,j) = \frac{pixel(i,j) - Media}{Desviación\ estandar}$$

Siendo:

Media = suma de todos los píxeles dividido entre el número de píxeles de la imagen.

$$\text{Desviación estándar} = \sqrt{\frac{\sum_{\text{para todos los píxeles}} (pixel(i,j) - Media)^2}{N^{\circ} \text{ píxeles de la imagen} - 1}}$$

Y se devuelve la imagen con los píxeles modificados en un vector (*array*) de tantas posiciones como píxeles tiene la imagen (30x30 en este caso).

²⁸ <http://docs.gimp.org/2.8/es/gimp-tutorial-quickie-flip.html>

²⁹ [http://en.wikipedia.org/wiki/Vectorization_\(mathematics\)](http://en.wikipedia.org/wiki/Vectorization_(mathematics))

5.3.6. Identificación del objeto

El *array* resultante (vector) es entregado al algoritmo *Support Vector Machine* (SVM). SVM es un algoritmo que, a partir de un modelo entrenado con multitud de imágenes, te devuelve probabilidades acerca de qué imagen es la que le muestras.

Más específicamente, SVM funciona con vectorización, todas las imágenes que le pasamos al modelo para entrenarlo se las pasamos vectorizadas, a partir de estos vectores se crean una serie de puntos en el espacio, con estos puntos se calculan lo que llamamos “hiperplanos³⁰”, que delimitan los espacios pertenecientes a cada clase o símbolo, a continuación cuando le pasamos una imagen vectorizada, a partir de toda esta información previa, el algoritmo puede indicarnos porcentajes acerca de que símbolo o clase se trata, por eso es tan importante tratar a las imágenes de igual manera que se hizo a la hora de entrenar el modelo que utiliza el algoritmo SVM.

En nuestro caso tenemos 5 posibilidades, de la posición 0 a la 4, que el objeto sea una O, una X, una I, una N o NADA respectivamente, por lo que cuando le pasamos una imagen nos devuelve un *array* de cinco posiciones con las probabilidades de que sea una u otra opción.

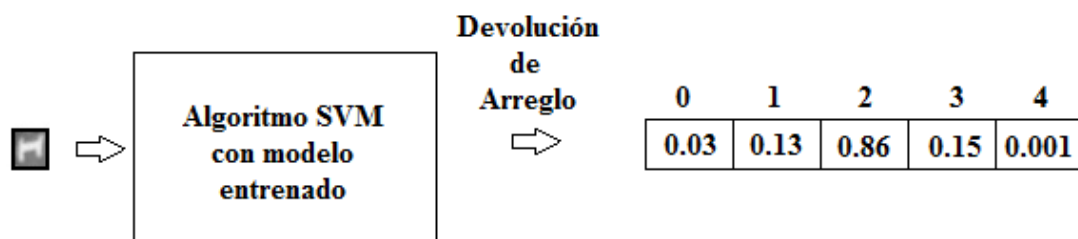


Figura 22 - Ejemplo de funcionamiento de algoritmo SVM.

Una vez tenemos la posición del objeto en la escena (localización), y hemos identificado a dicho objeto (forma aproximada de la imagen por algoritmo SVM), tenemos dos formas de actuar:

³⁰ <http://es.wikipedia.org/wiki/Hiperplano>

1. Procesar la identificación/posición exacta del *blob* y devolverla para saber a tiempo real donde se encuentra el objeto y que objeto es
2. Almacenar la posición y la identidad del objeto para calcular, almacenar y devolver el recorrido de dicho objeto

5.3.6.1. Localización del objeto

Para el procesamiento y devolución de la posición exacta del objeto identificado, se analiza su posición basándose en las estadísticas devueltas por el algoritmo SVM, una distancia mínima entre objetos de X píxeles y una variable de precisión Y (*accuracy*). Si el algoritmo SVM devuelve una probabilidad mayor a la precisión que se ha especificado anteriormente (Y), se considera óptimo y se procede a la devolución sin realizar ninguna otra operación. De otro modo si el algoritmo SVM devuelve una probabilidad menor se procesa el objeto identificándolo como el objeto que tiene a una distancia inferior a la distancia mínima X especificada anteriormente. Este algoritmo nos ayuda a evitar dos cosas: Fallos en la devolución de probabilidades por parte del algoritmo SVM y evitar la relación de los objetos únicamente por sus distancias.

A continuación se verá un ejemplo que ilustra cómo funciona este pequeño algoritmo de identificación.

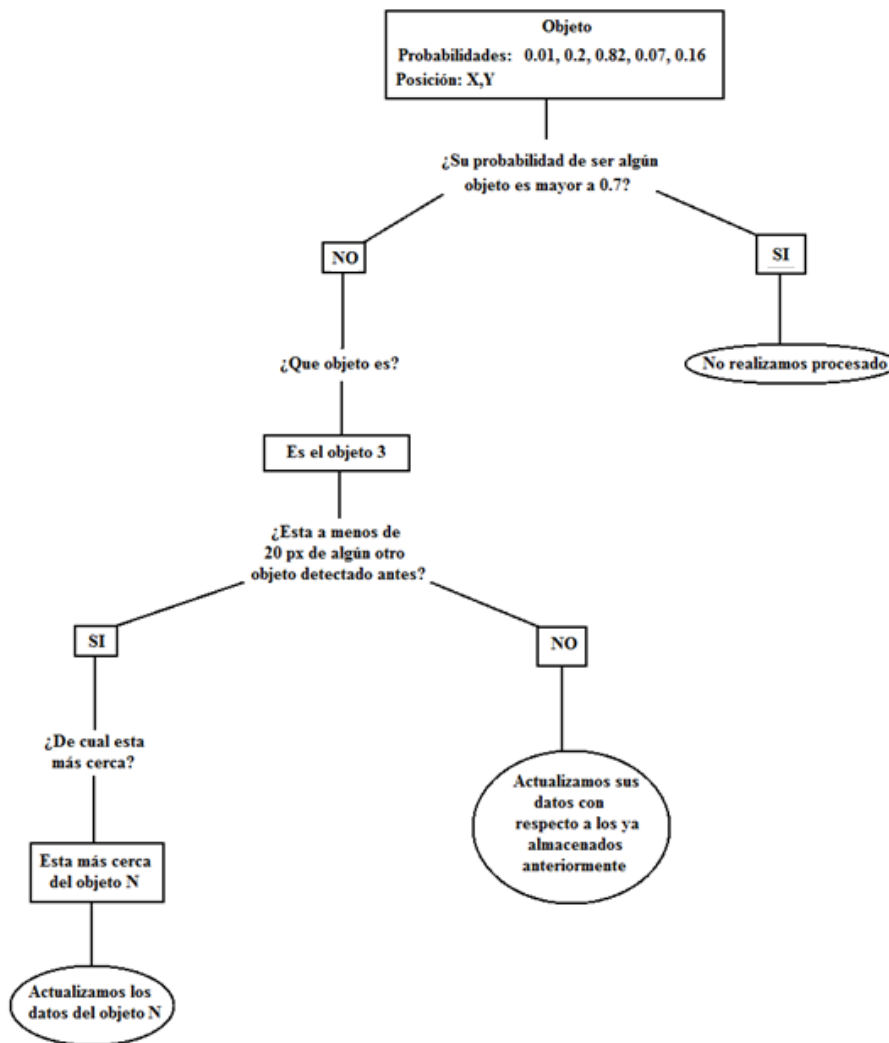
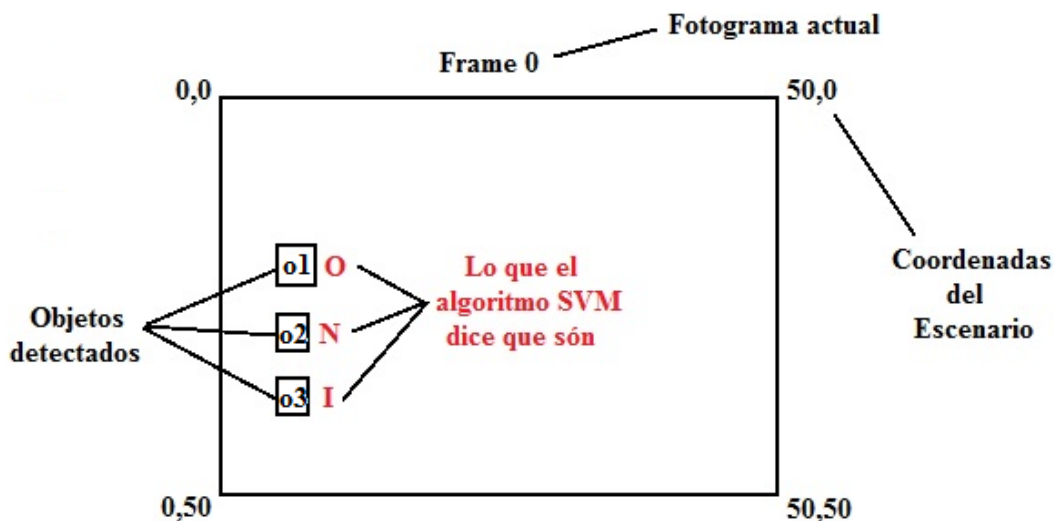


Figura 23 - Funcionamiento procesado de identificación/posición del objeto.

5.3.6.2. Seguimiento del objeto

Para realizar el seguimiento o tracking de los objetos (segunda forma de actuación), se utiliza un algoritmo que incorpora inteligencia artificial. Con este algoritmo, se consigue evitar fallos del sistema cuando los objetos se cruzan, desaparecen (ocultación detrás de un obstáculo del escenario) o el algoritmo SVM se confunde de objeto (Véase anexo para ver resultados reales). El ejemplo que se presenta a continuación pretende ilustrar cómo funciona el algoritmo:



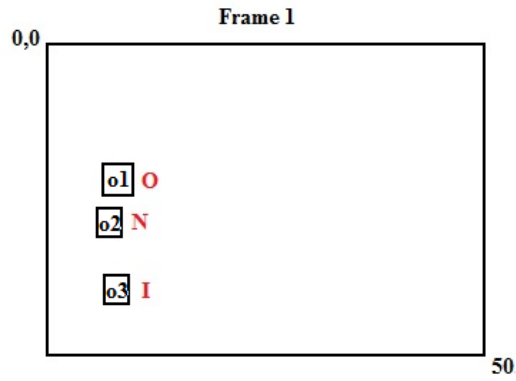
Vector BlobsDetectados:

Pos	o1	o2	o3
SVM	O	N	I
Frame	0	0	0

Vector SeguimientoObjeto:

	0
O	
N	
I	

El vector de objetos detectados “*BlobsDetectados*” tiene tantas posiciones como objetos/*blobs* detectados haya en ese momento en el *frame*, y almacena posiciones, salida del algoritmo SVM y el primer fotograma en el que fue detectada la forma. El vector de seguimiento “*SeguimientoObjeto*” tiene tantas posiciones como fotogramas hayamos analizado del video, por lo que crece 1 posición por cada fotograma que pasa, y almacena las posiciones de los objetos que analizamos, en este ejemplo las posiciones del objeto O, del objeto N y del objeto I.



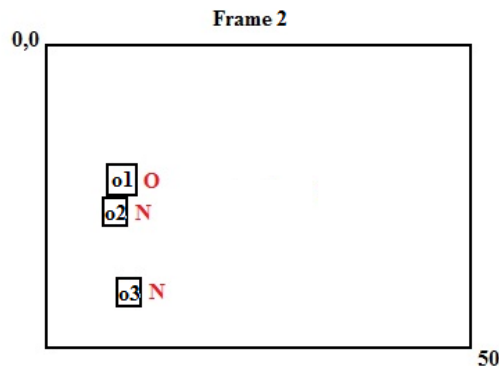
Vector BlobsDetectados:

	o1	o2	o3
Pos	5,25/7,26	5,35/6,33	5,45/7,47
SVM	O/O	N/N	I/I
Frame	0	0	0

Vector SeguimientoObjeto:

	0	1
O		
N		
I		

Para poder rellenar el vector “*BlobsDetectados*”, el algoritmo asocia los objetos detectados en el siguiente fotograma por la distancia que hay de los anteriores a los actuales.

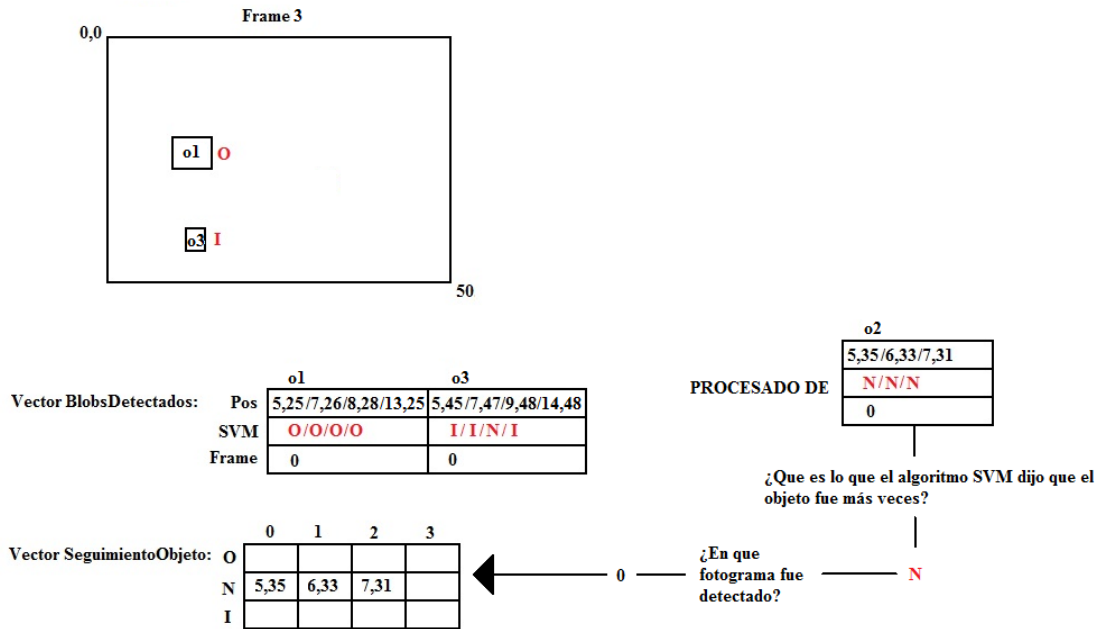


Vector BlobsDetectados:

	o1	o2	o3
Pos	5,25/7,26/8,28	5,35/6,33/7,31	5,45/7,47/9,48
SVM	O/O/O	N/N/N	I/I/N
Frame	0	0	0

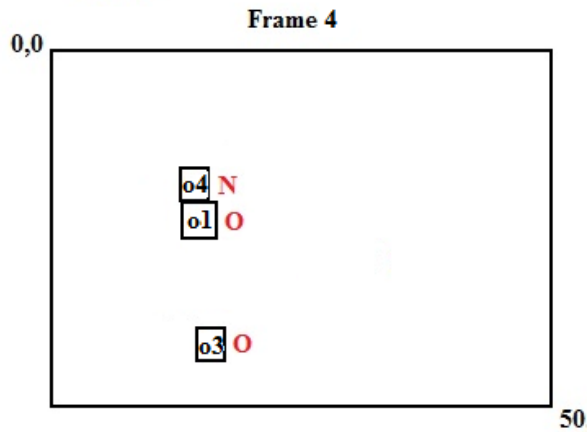
Vector SeguimientoObjeto:

	0	1	2
O			
N			
I			



El sistema ha unido los objetos o1 y o2 en un solo objeto o “blob” porque estaban muy cerca el uno del otro (ha habido un cruce de objetos), y nuestro algoritmo ha asociado este “nuevo” objeto con el objeto o1 por cercanía. Por lo que interpreta que el objeto o2 ha desaparecido y lo procesa.

Como podemos ver, con este almacenamiento de datos y posterior procesado evitamos fallos con los cruces de los objetos (cosa muy frecuente en este tipo de experimentos).



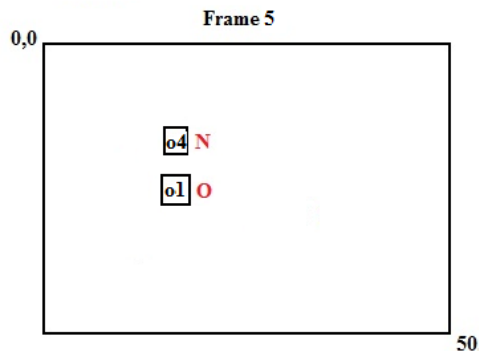
Vector BlobsDetectados:

	o1	o3	o4
Pos	5,25/7,26/8,28/13,25/15,30	5,45/7,47/9,48/14,48/16,48	14,25
SVM	O/O/O/O/O	I/I/N/I/O	N
Frame	0	0	4

Vector SeguimientoObjeto:

	0	1	2	3	4
O					
N	5,35	6,33	7,31		
I					

Se ha vuelto a separar en 2 el *blob* que antes se había unido en uno solo, así que reconocemos lo que antes veíamos como o2, como un objeto nuevo (o4), y lo añadimos al vector de *blobs* detectados con: su posición, lo que el algoritmo SVM ha dicho que es y el fotograma en el que lo hemos detectado.



Vector BlobsDetectados:

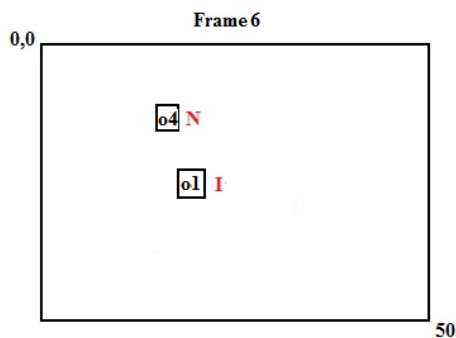
	o1	o4
Pos	5,25/7,26/8,28/13,25/15,30/18,32	14,25/18,22
SVM	O/O/O/O/O/O	N/N
Frame	0	4

Vector SeguimientoObjeto:

	0	1	2	3	4	5
O						
N	5,35	6,33	7,31			
I	5,45	7,47	9,48	14,48	16,48	

← Fué más veces **I** y se detecto en el fotograma 0

Podemos comprobar que en este nuevo fotograma se ha perdido el objeto o3, por lo que podemos intuir que se haya detrás de algún obstáculo del escenario, por lo que lo procesamos igual que hicimos con o2.

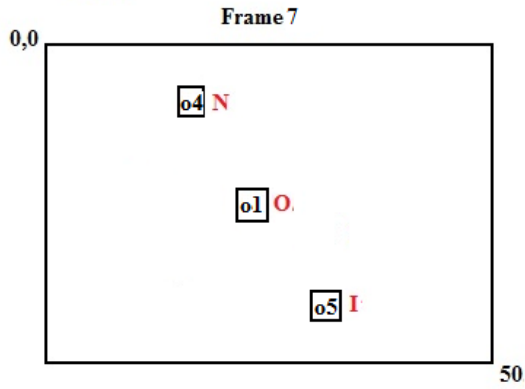


Vector BlobsDetectados:

	o1	o4
Pos	5,25/7,26/8,28/13,25/15,30/18,32/23,32	14,25/18,22/18,16
SVM	O/O/O/O/O/O/I	N/N/N
Frame	0	4

Vector SeguimientoObjeto:

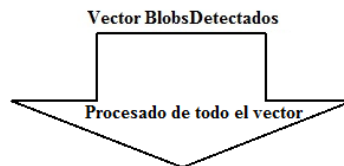
	0	1	2	3	4	5	6
O							
N	5,35	6,33	7,31				
I	5,45	7,47	9,48	14,48	16,48		



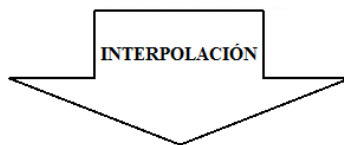
	o1	o4	o5
Vector BlobsDetectados: Pos	5,25/7,26/8,28/13,25/15,30/18,32/23,32/29,32	14,25/18,22/18,16/18,10	35,48
SVM	O/O/O/O/O/I/O	N/N/N/N	I
Frame	0	4	7

Vector SeguimientoObjeto:	0	1	2	3	4	5	6	7
O								
N	5,35	6,33	7,31					
I	5,45	7,47	9,48	14,48	16,48			

FIN DEL PROGRAMA:



Vector SeguimientoObjeto:	0	1	2	3	4	5	6	7
O	5,25	7,26	8,28	13,25	15,30	18,32	23,32	29,32
N	5,35	6,33	7,31		14,25	18,22	18,16	18,10
I	5,45	7,47	9,48	14,48	16,48			35,48



Para rellenar las posiciones en las que no hemos detectado a los objetos, simplemente realizamos una interpolación sencilla y los rellenamos.

Vector SeguimientoObjeto:	0	1	2	3	4	5	6	7
O	5,25	7,26	8,28	13,25	15,30	18,32	23,32	29,32
N	5,35	6,33	7,31	10,28	14,25	18,22	18,16	18,10
I	5,45	7,47	9,48	14,48	16,48	22,48	28,48	35,48

Una vez finaliza la ejecución del sistema global, añadimos todas las posiciones que se habían almacenado en el vector de *blobs* detectados para obtener un seguimiento completo de todas las formas analizadas en el proceso.

Resultados:

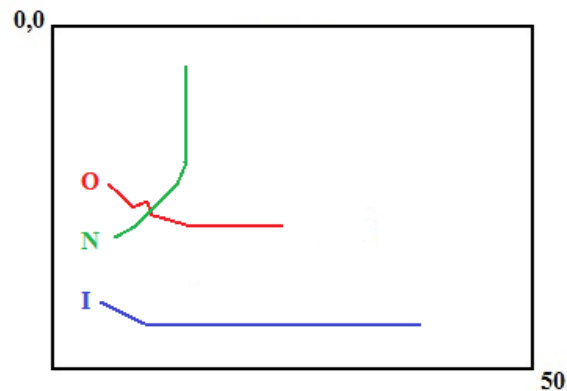


Figura 24 - Resultados del algoritmo de seguimiento

5.3.7 Refinamiento y salvado de datos

Finalmente, cuando nos disponemos a utilizar los datos que hemos ido almacenando tenemos dos opciones: salvarlos o utilizarlos a tiempo real para realizar operaciones con ellos.

En cualquiera de los dos casos, se tienen que “refinar”, esto se debe a que hay momentos en los que se pierden los datos, por ello realizamos (como se muestra en el ejemplo anterior) una interpolación, pero esto no es suficiente, antes de realizar la interpolación se debe tener en cuenta donde perdimos el objeto y donde lo encontramos para poder realizar un refinamiento que consiste en la colocación de puntos “imaginarios” que hacen que cuando se realiza la interpolación, los resultados finales tengan más sentido (Véase anexo para ver resultados reales).

A continuación se muestra un ejemplo para la mejor comprensión de esta parte:

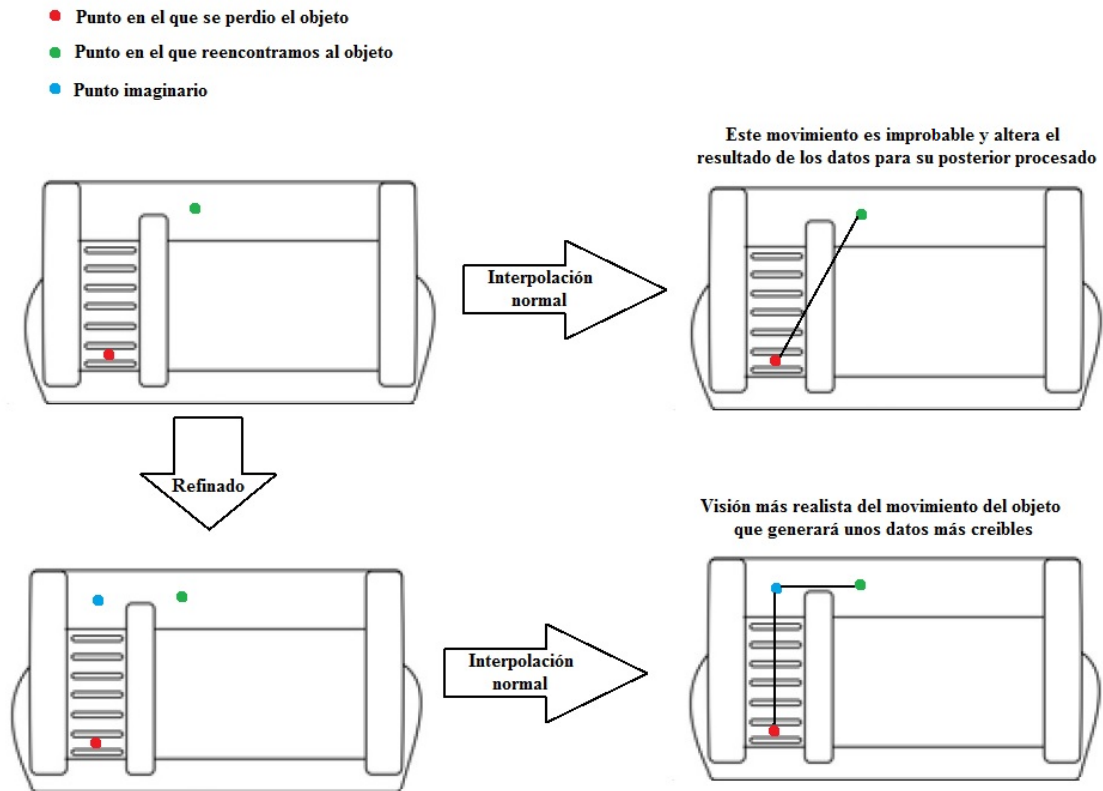
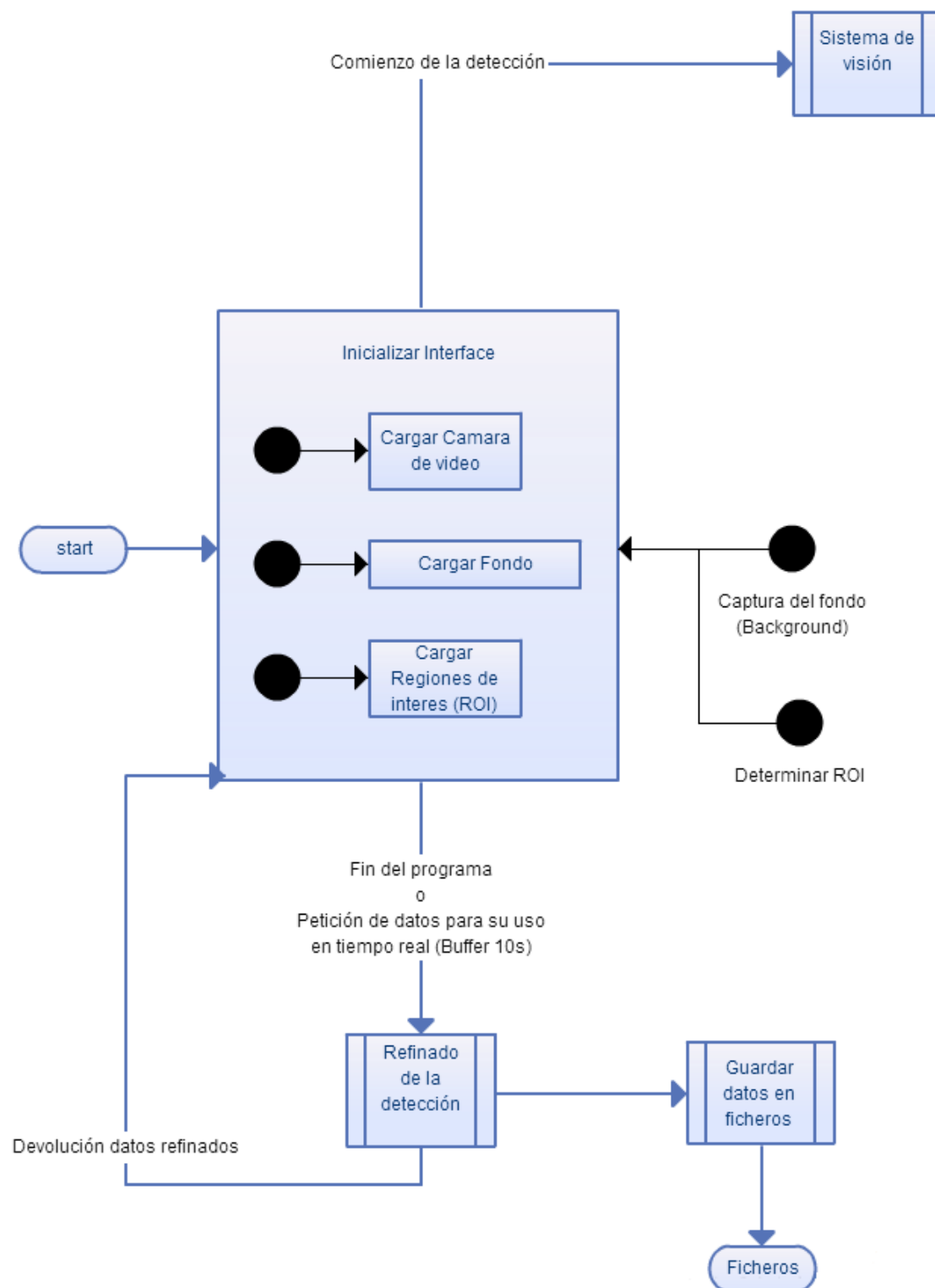


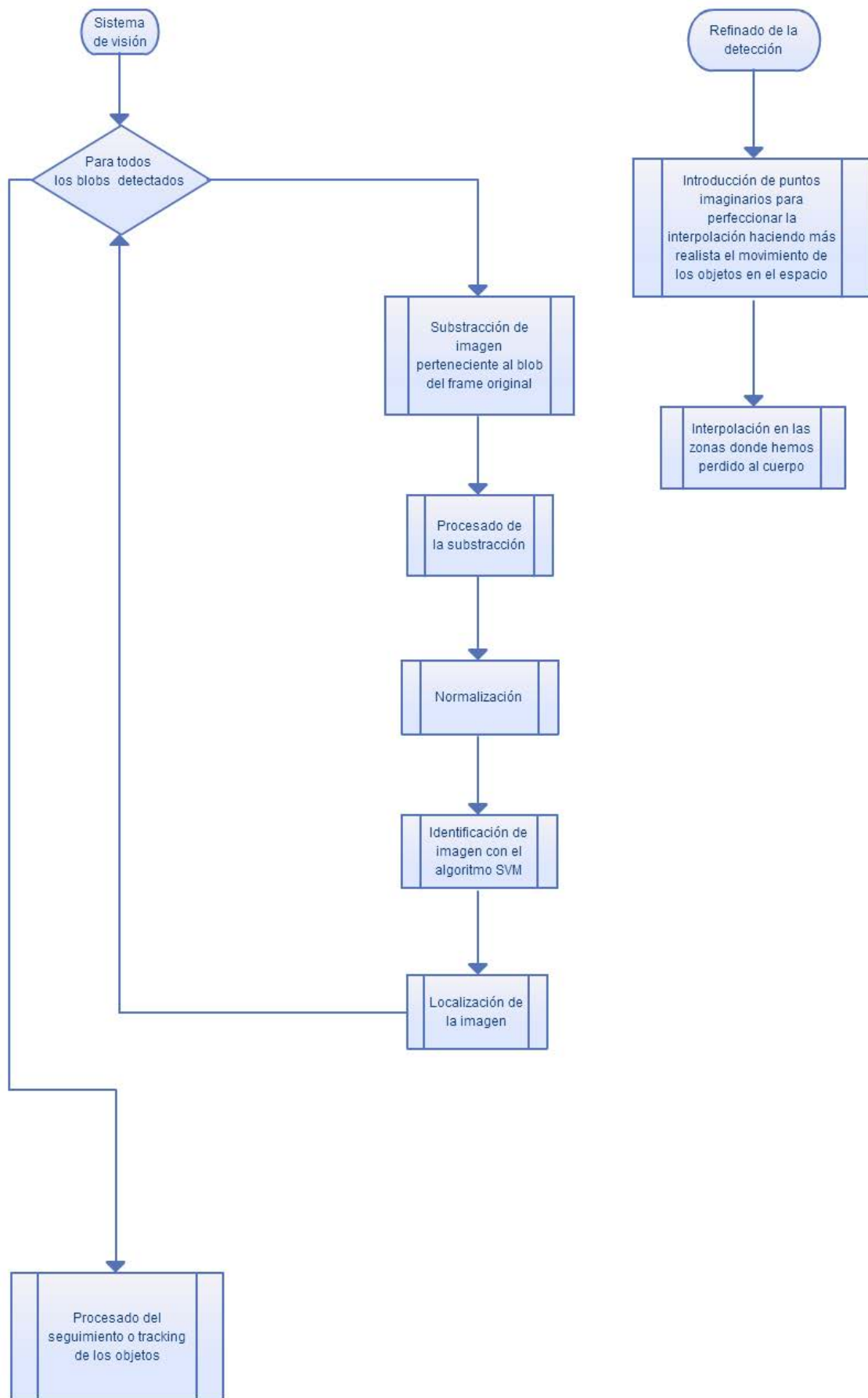
Figura 25 - Refinado de los datos previos a la interpolación para un mejor resultado.

Para el salvado de datos, se colocan todos los datos almacenados en ficheros independientes (uno por cada cuerpo al que le hemos hecho el seguimiento), identificando fotograma a fotograma:

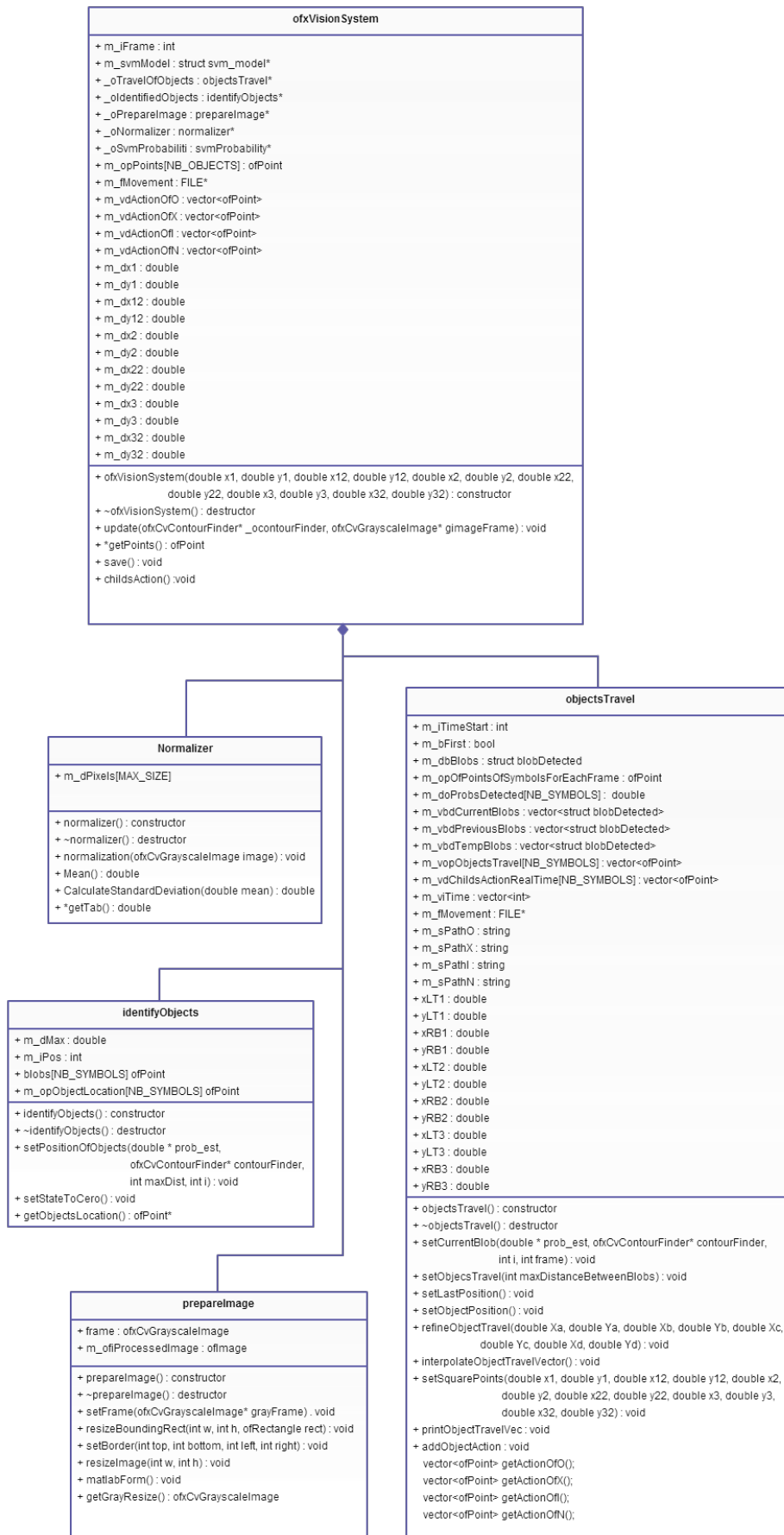
- La localización (X, Y)
- El tiempo de la localización
- La zona en la que ha estado en el momento de la localización

5.4. Diagrama de flujo





5.5. Diagrama de clases



La clase `ofxVisionSystem` tiene relación de composición con todas las otras clases para poder generarse. Esta clase es el addon que consigue realizar todo el proceso de identificación y seguimiento de los objetos pasándole el video fotograma a fotograma y el conjunto de *blobs* detectados en cada uno de ellos.

Las clases que utiliza el sistema de visión “`ofxVisionSystem`” son:

- **prepareImage**: Que subtrae la zona de interés del fotograma y la prepara para su procesamiento por el algoritmo SVM para su identificación.
- **Normalizer**: Que normaliza dicha imagen para que el algoritmo SVM tenga una mayor precisión
- **identifyObject**: Que utiliza un algoritmo sencillo para identificar con más precisión la localización del objeto
- **objectTravel**: Que utiliza el algoritmo de inteligencia artificial para determinar y almacenar con precisión el recorrido de los objetos individualmente

6. CONCLUSIÓN

Me llevo un muy buen sabor de boca con este proyecto, he podido poner en práctica la mayoría de conocimientos adquiridos a lo largo del grado, y he podido aprender de un tema que en mi titulación simplemente se roza, los sistemas de visión y todo lo que los envuelven.

Por otra parte he podido comprobar que toda esta tecnología tratada en este proyecto es fascinante, tienen multitud de utilidades y una gran proyección en trabajos futuros.

Además está el factor de haber podido aportar mi granito de arena al proyecto del equipo de Narcís Parés (*The Interactive Slide*), cosa que me hace pensar que algún día podré decir que colabore en un proyecto que está ayudando a cambiar vidas.

Dicho esto, me gustaría decir que creo que se tendría que trabajar más en proyectos como este, puesto que tienen mucha proyección de futuro al enfocarse en niños, ¿y que son los niños sino nuestro futuro?

En las dos oportunidades que me han brindado para probar en espacio mi sistema con más de 50 niños, he visto una reacción muy positiva por parte de ellos, además el sistema ha podido reconocer sin apenas problemas a todos los niños, y ha podido hacer un seguimiento de todos ellos con una precisión aceptable. Por lo que con esto, y la imaginación de la gente con la que he trabajado, seguro que se podrán crear *Exergames* más funcionales, con mayor interacción persona-máquina y con aprendizaje automático para una mejor y más buena participación y monitoreo de los niños al jugar.

Como conclusión final, decir que me ha parecido que los resultados han sido óptimos, y que esta tecnología con un poco más de trabajo se podría aplicar a cualquier sistema de visión para hacer futuros experimentos, además, la relación rendimiento-precio del proyecto es excepcional.

6.1. Trabajo futuro

El trabajo futuro que se realizará a partir del punto en el que nos encontramos, y en el que espero poder participar, es el de la realización de un *Exergame* enfocado a explotar al 100% el sistema que trata este proyecto final de grado.

Un *Exergame* que interactúe directamente con los niños, que les diga el orden de lanzamiento por el tobogán interactivo, que haga actuar a los niños con una menor actividad, y que fomente (más de lo que ya se hace) además del ejercicio físico, la socialización de los participantes, haciendo que estos tengan la necesidad de tener que trabajar en grupo para pasar de nivel o conseguir más puntuación.

Además como ya se ha dicho a lo largo de esta memoria, utilizar el sistema ideado en este trabajo final de carrera única y exclusivamente para *Exergames* en el tobogán interactivo, considero que sería un error, puesto que este proyecto podría ser utilizado para muchos otros ámbitos como podrían ser:

- Vigilancia automática y control de actividad de niños en guarderías
- Abaratar costes en el reconocimiento automático de objetos
- Otros *Exergames*
- Control nocturno de ganado
- Aviso de emergencia en autopista
- *Coaching* interactivo con un entrenador personal para clases de spinning, aerobic, etc.

6.2. Valoración personal

El desarrollo de este proyecto ha sido para mí un gran reto que me ha hecho aprender día a día tanto nuevos conceptos como la forma de ser un mejor ingeniero. Este proyecto final de carrera me ha enseñado que la moral de un ingeniero es algo muy importante, y lo que verdaderamente te puede hacer disfrutar de tu profesión.

Trabajando con niños me he dado cuenta de que mis estudios no solo se limitan a ser empleados para la programación, sino que pueden llegar a tener recompensas que jamás pudiese haber esperado.

En conclusión, esta experiencia ha sido de las mejores de mi vida estudiantil, una experiencia en la que he aprendido muchas cosas, no solo desde el punto de vista del programador ni de ingeniero, sino también desde el punto de vista social.

En cuanto a la parte más técnica, el hecho de haber trabajado con tantas variables de entorno que no controlaba, y que provocaban problemas en la recopilación de datos, ha hecho que dé un paso más allá en mi concepto de un desarrollador, debido a que muchas veces he tenido que pensar en soluciones, como el algoritmo de inteligencia artificial implementado, para solucionar dichos problemas.

Bibliografía

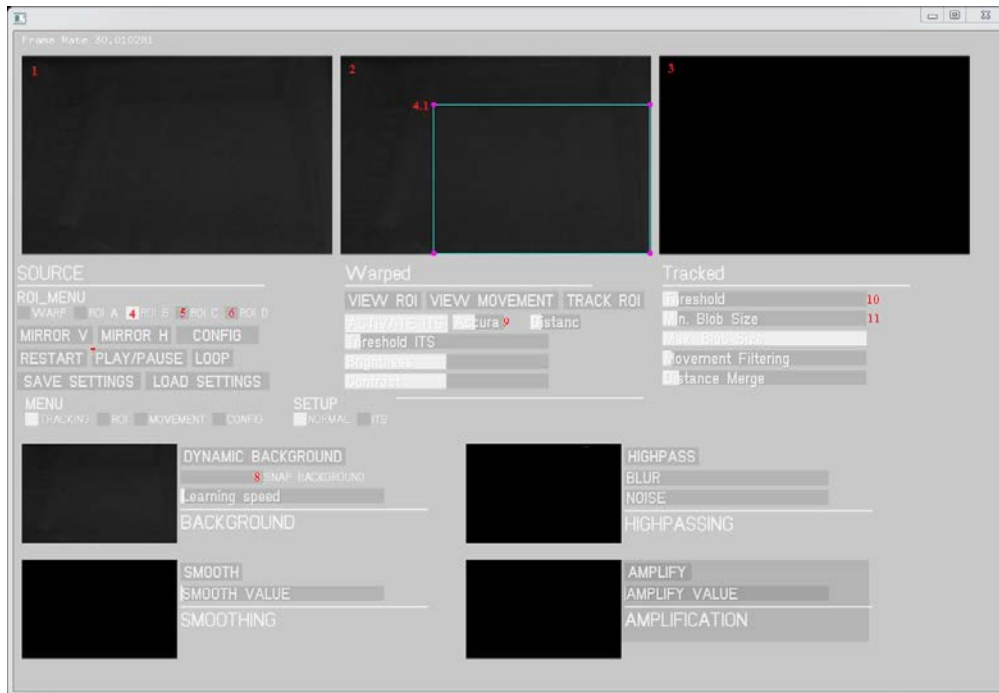
- [1] Change, V. (n.d.). Lecture 24 Background subtraction Video Change Detection.
- [2] Chaparro Valderrama, J. 1956-, Lema, M. il, & Unisur, F. D. C. S. Y. H. (n.d.). 8. Formulación del proyecto | banrepcultural.org. Retrieved from http://www.banrepcultural.org/blaavirtual/sociologia/met1/8.htm#_ftn1
- [3] Consejos para el desarrollo de tu proyecto . (n.d.). Retrieved May 17, 2014, from <http://users.dsic.upv.es/~fjabad/pfc/estructura.pdf>
- [4] Ingeniería del software - Monografias.com. (n.d.). Retrieved May 20, 2014, from <http://www.monografias.com/trabajos15/ingenieria-software/ingenieria-software.shtml>
- [5] Josefina, I., Enríquez, G., Manuel, J., Cortés, R., Natalia, M., Bonilla, I., & Gómez, P. (2008). Seguimiento Autónomo de la Posición de un Objeto por Visión y Control Neuro- difuso en MATLAB, (November), 335–340.
- [6] Kim, J., Choi, K., Choi, B., & Ko, S. (2002). Real-time vision-based people counting system for the security door. ... *Technical Conference on Circuits Systems, ...*, 3672. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Real-time+Vision-based+People+Counting+System+for+the+Security+Door#0>
- [7] Landry, P., & Pares, N. (2012). Controlling the amount of physical activity in a specific exertion interface. *Proceedings of the 2012 ACM Annual Conference Extended Abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12*, 2393. doi:10.1145/2212776.2223808
- [8] Landry, P., Parés, N., Fabra, U. P., & Boronat, R. (n.d.). Controlling and Modulating Physical Activi- ty through Interaction Tempo in Exergames : A Quantitative Empirical Analysis.
- [9] LIBSVM -- A Library for Support Vector Machines. (n.d.-a). Retrieved May 20, 2014, from <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- [10] Los enigmas de la Obesidad: Sus causas (página 2) - Monografias.com. (n.d.). Retrieved May 23, 2014, from

<http://www.monografias.com/trabajos87/enigmas-obesidad-sus-causas/enigmas-obesidad-sus-causas2.shtml>

- [11] Medrano, C. (n.d.). Tutorial de OpenCV.
- [12] Muñoz, J. E., Villada, J. F., Carlos, J., & Trujillo, G. (2013). Exergames : una herramienta tecnológica para la actividad física, *19*(3), 126–130.
- [13] Self-driving car pushes sensor technology | EDN. (n.d.). Retrieved May 26, 2014, from <http://edn.com/electronics-blogs/automotive-innovation/4402800/Self-driving-car-pushes-sensor-technology>
- [14] Soler-adillon, J. (n.d.). Interactive Slide : an Interactive Playground to Promote Physical Activity and Socialization of Children, 1–10.
- [15] Soler-adillon, J. (2009). A Novel Approach to Interactive Playgrounds : the Interactive Slide Project.
- [16] Vision, I. (n.d.). IVS Vision Systems - An introduction to our machine vision technology.
- [17] Yilmaz, A., Javed, O., & Shah, M. (2006). Object tracking. *ACM Computing Surveys*. doi:10.1145/1177352.1177355

ANEXO

Explicación Interfaz



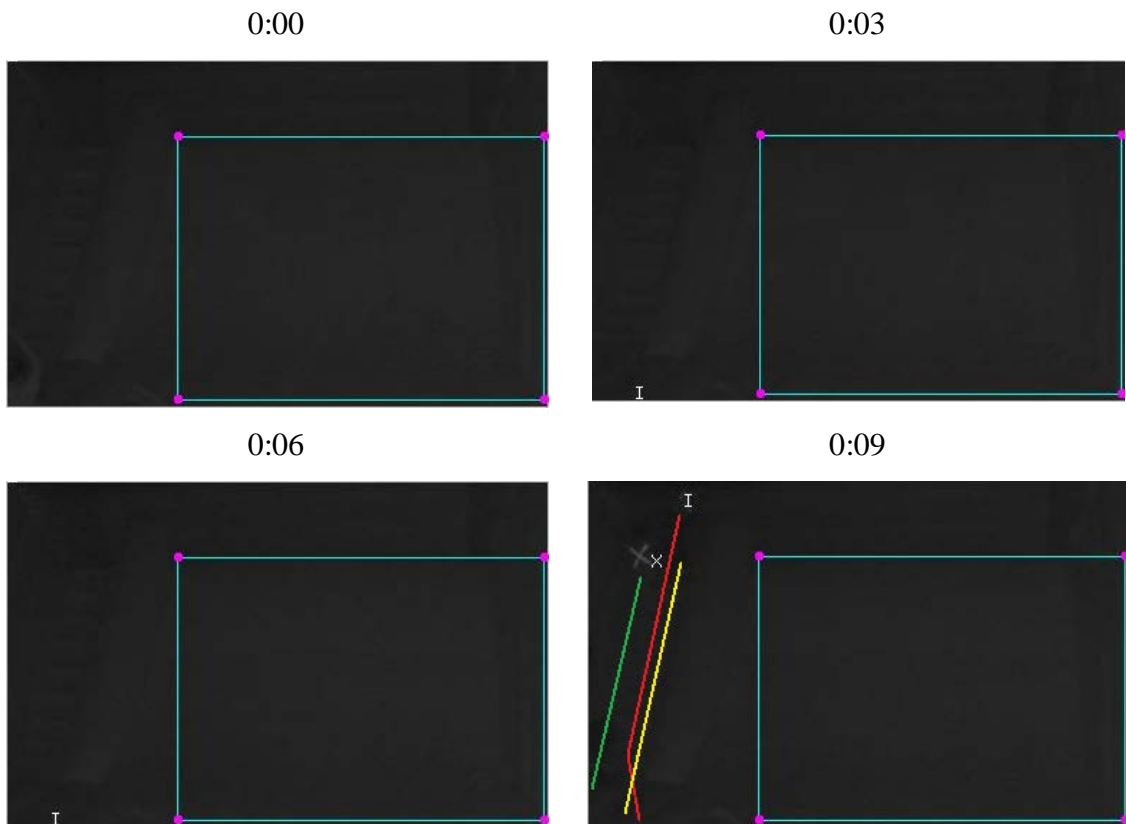
1. Salida de video normal
2. Salida de video + localización a tiempo real de los diferentes objetos
3. Imagen resultante de la Substracción de fondo, el *thresholding*, la detección de *blobs* y el *bounding box*
4. Zona de detección A. Cuadrado que delimita una zona de detección para clasificar las diferentes zonas con más o menos actividad física
 - 4.1 Cuadrado delimitador variable a tiempo real arrastrándolo por las esquinas
5. Zona de detección B
6. Zona de detección C
7. Botón Play/Pause del video para reproducir o pausar el video
8. Botón para realizar una captura del fondo “*Snap Background*”
9. Regulador de precisión para el algoritmo de identificación de objetos (Apartado 5.3.6.1)
10. Regulador del umbral del *thresholding* (Ver subapartado de “Seguridad” del apartado 2.4)
11. Regulador distancia entre *blobs* de la modificación realizada explicada en el apartado 5.3.3

Resultados

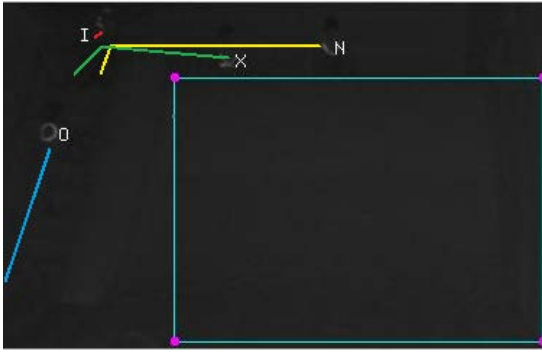
Para hacer el análisis de los resultados, se cogerá como referencia, para tener una mejor visualización de los mismos, la salida de video (2) que muestra los resultados a tiempo real. Se han realizado capturas del experimento cada tres segundos en un tiempo de 2 minutos y 45 segundos, con esto se puede apreciar los resultados obtenidos por el sistema.

En cada imagen se aprecia como el programa nos devuelve la forma y la situación de la imagen, esto se realiza imprimiendo en pantalla una letra identificadora a la derecha del objetivo (Ver primeros cuatro párrafos del apartado 4.2).

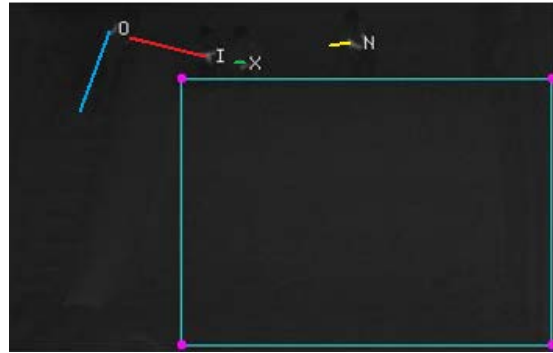
Además, se han marcado con líneas de diferentes colores los recorridos almacenados por el sistema de tracking de objetos con el refinado de los mismos ya implementado, al final se expondrán algunos casos particulares para ver el efecto del refinado en los resultados del sistema.



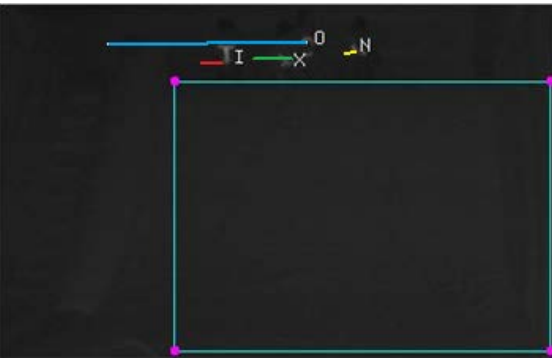
0:12



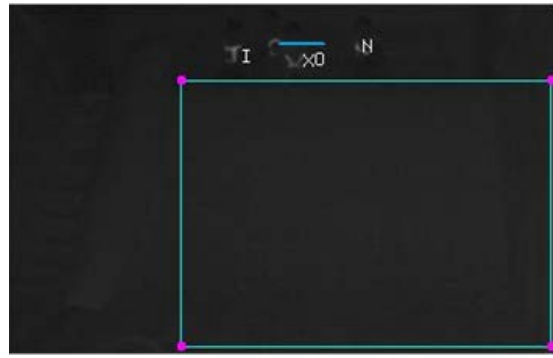
0:15



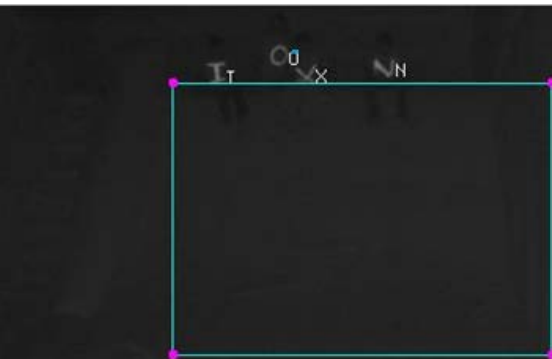
0:18



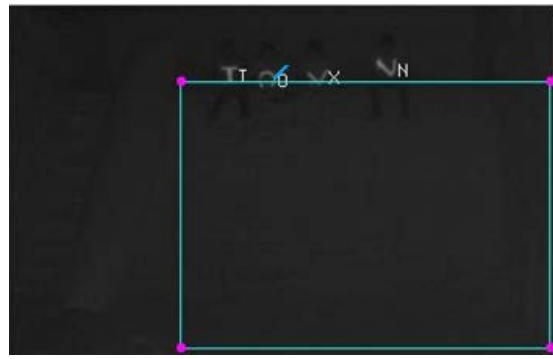
0:21



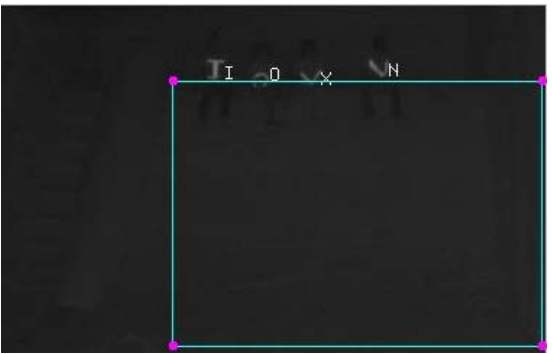
0:24



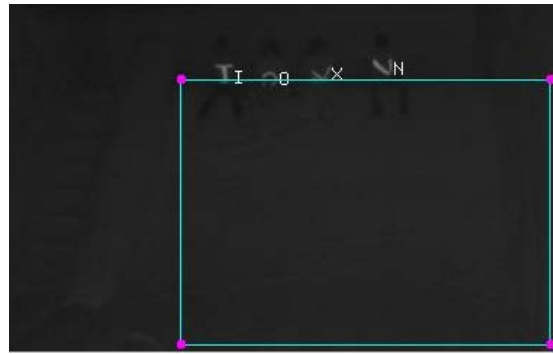
0:27



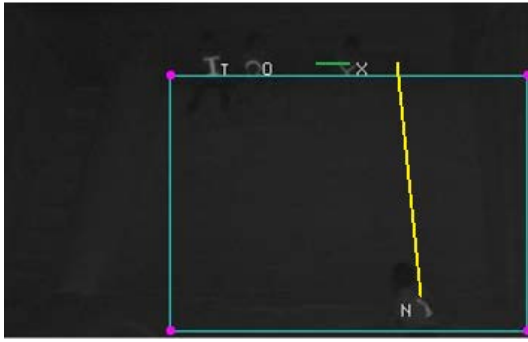
0:30



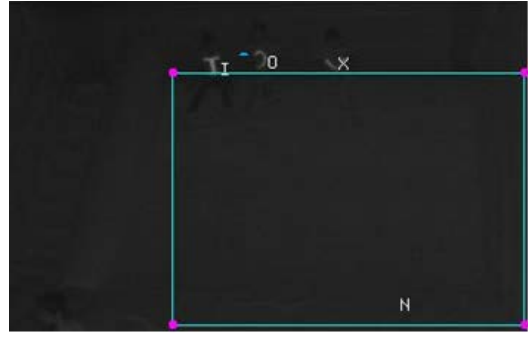
0:33



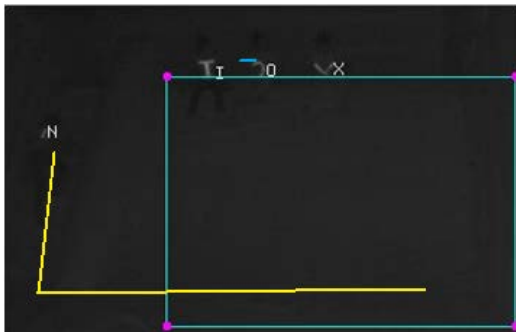
0:36



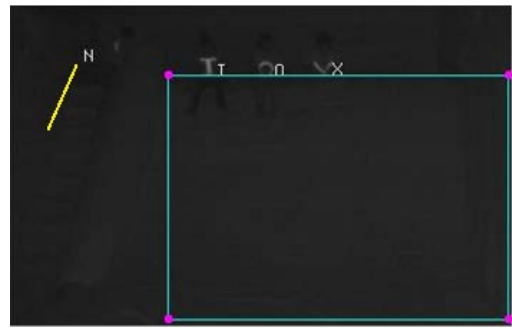
0:39



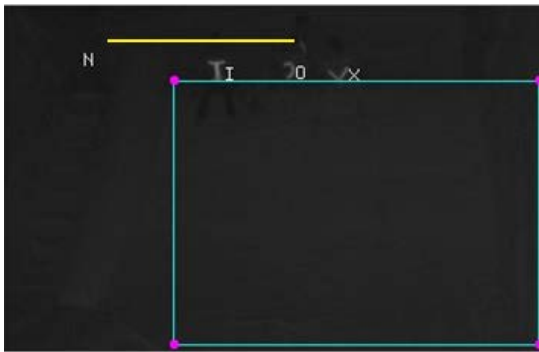
0:42



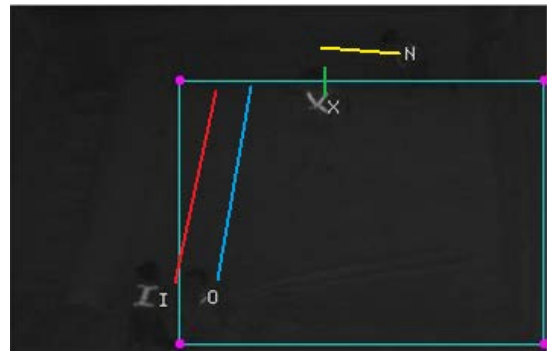
0:45



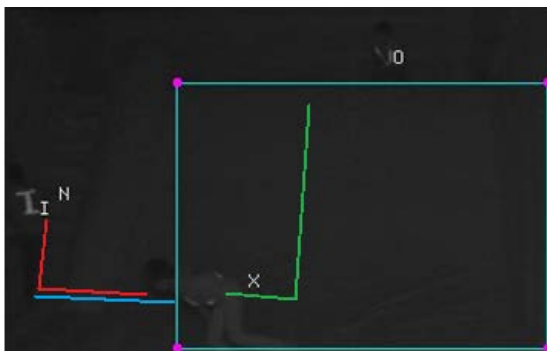
0:48



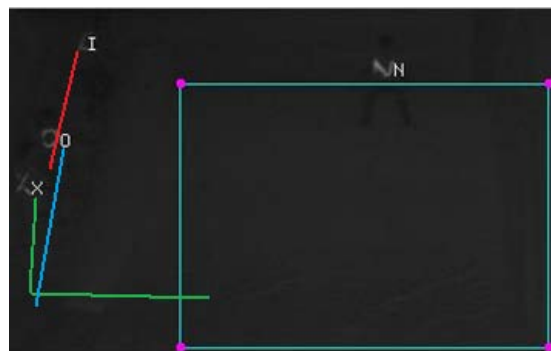
0:51



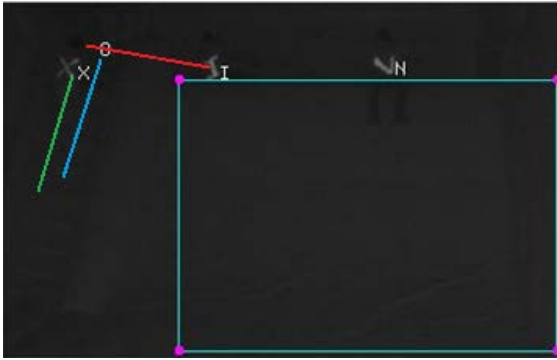
0:54



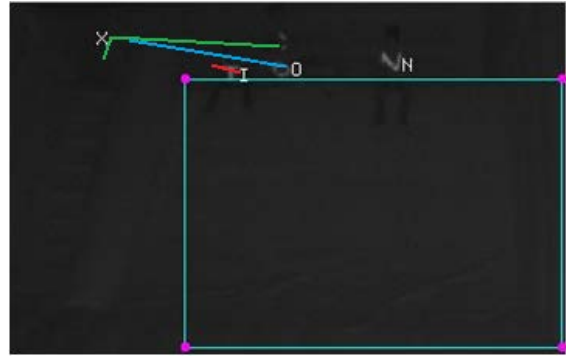
0:57



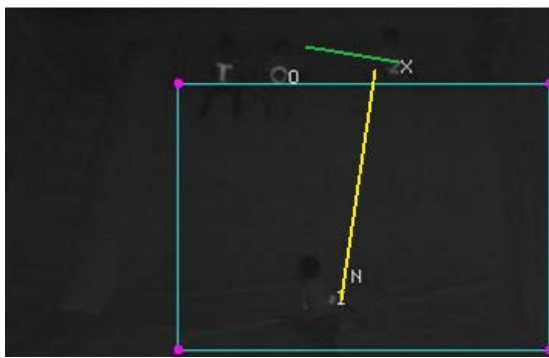
1:00



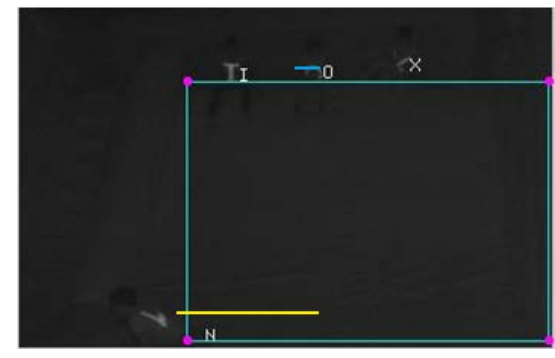
1:03



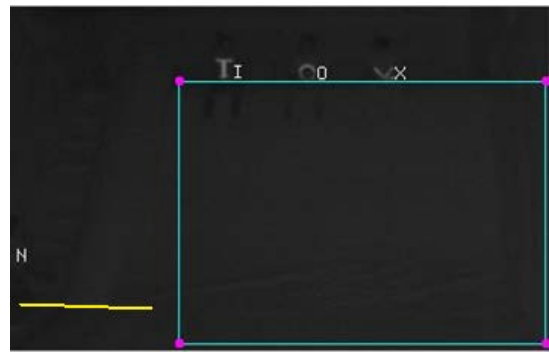
1:06



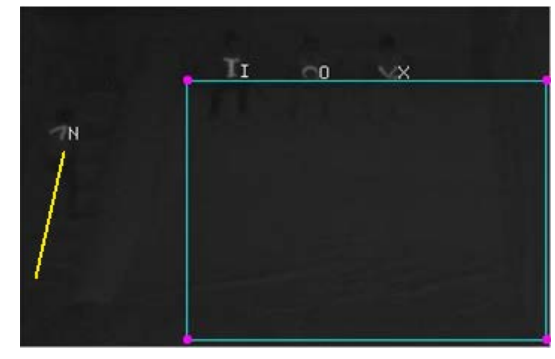
1:09



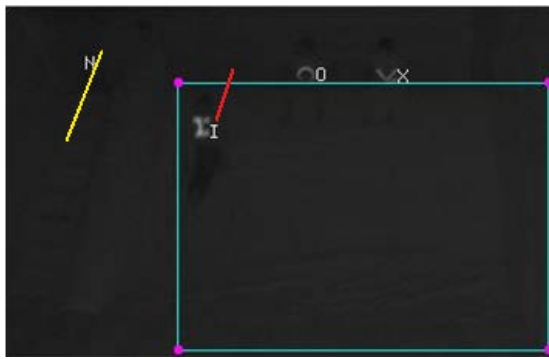
1:12



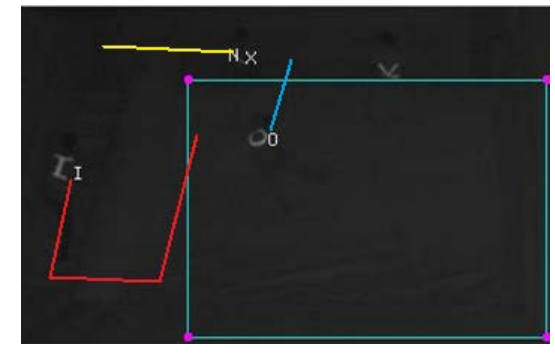
1:15



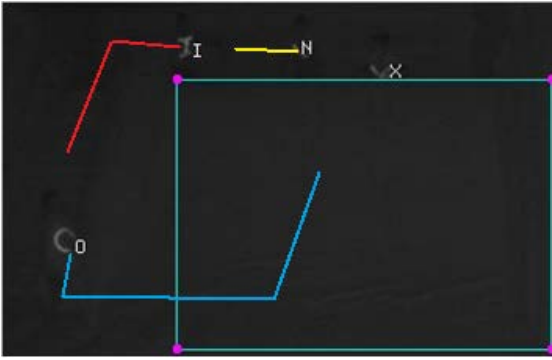
1:18



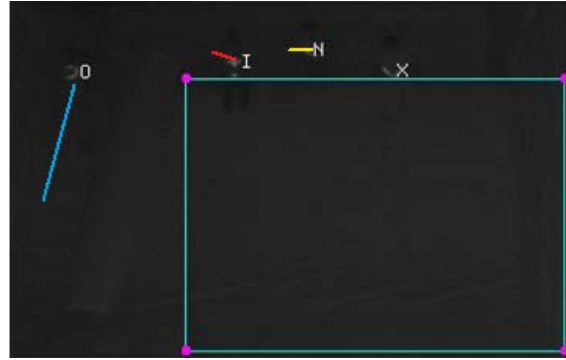
1:21



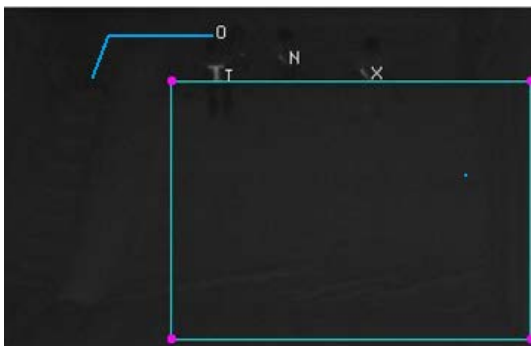
1:24



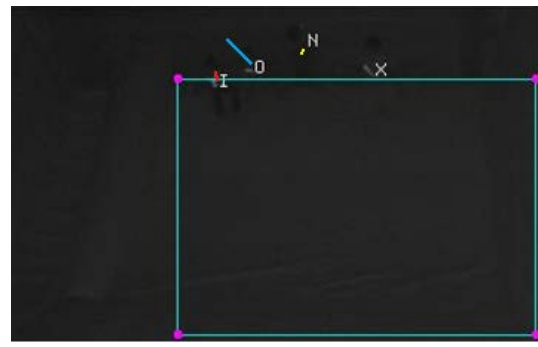
1:27



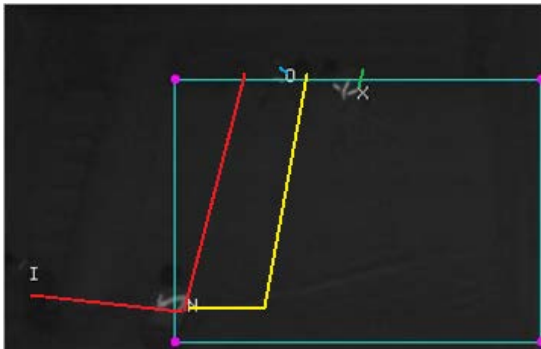
1:30



1:33



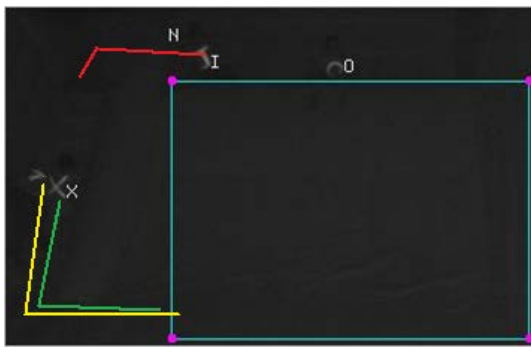
1:36



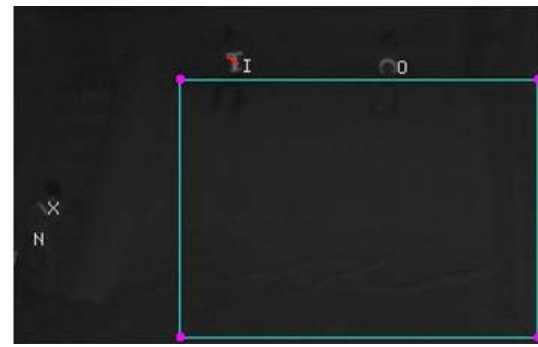
1:39



1:42



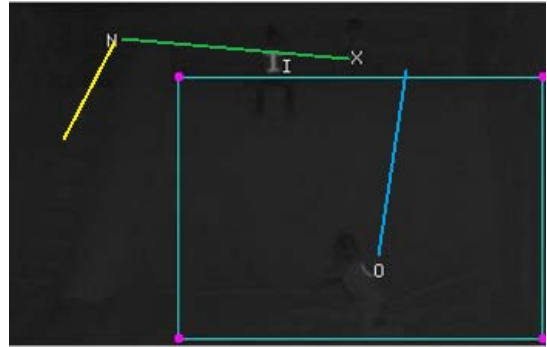
1:45



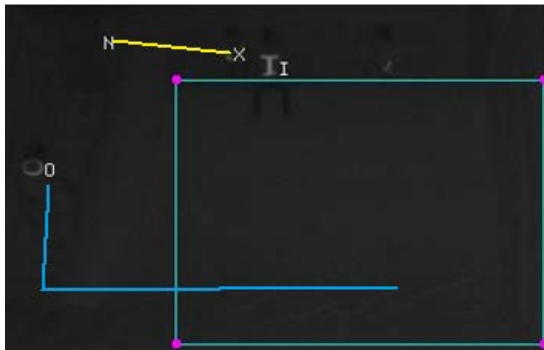
1:48



1:51



1:54



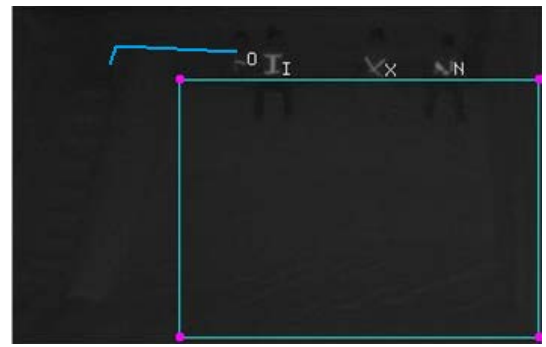
1:57



2:00



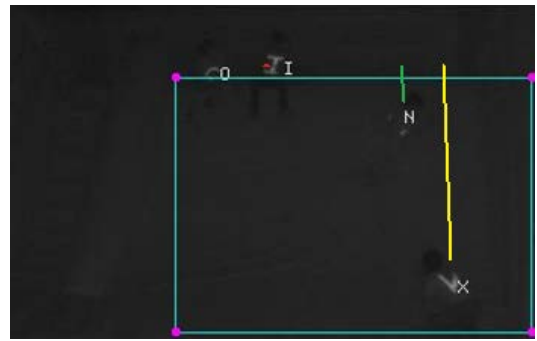
2:03



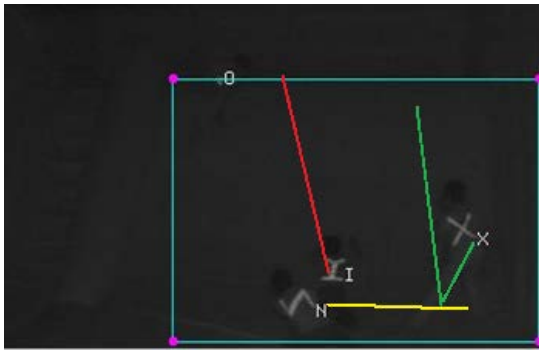
2:06



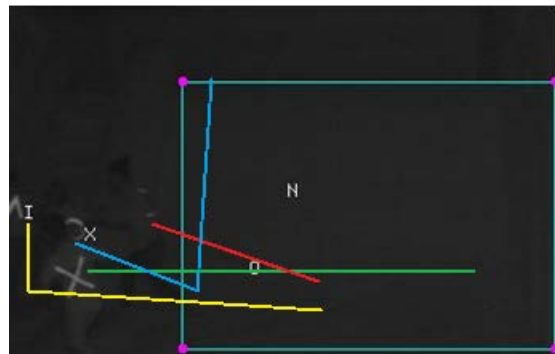
2:09



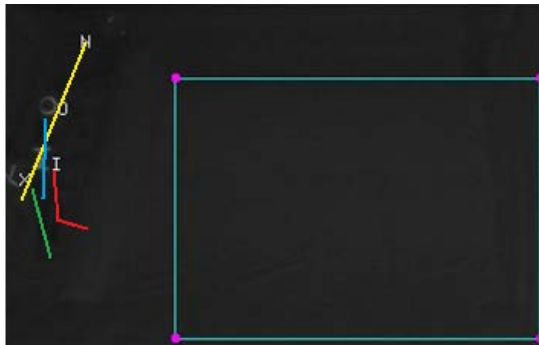
2:12



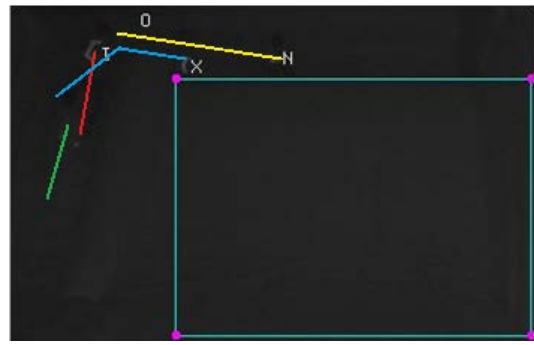
2:15



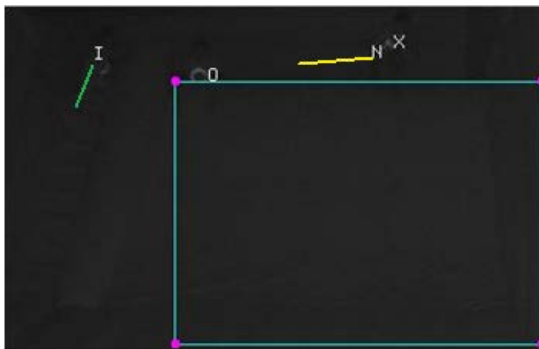
2:18



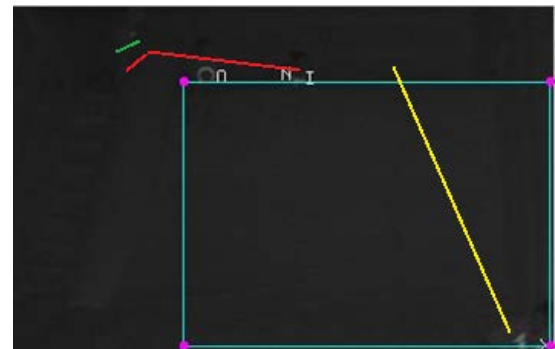
2:21



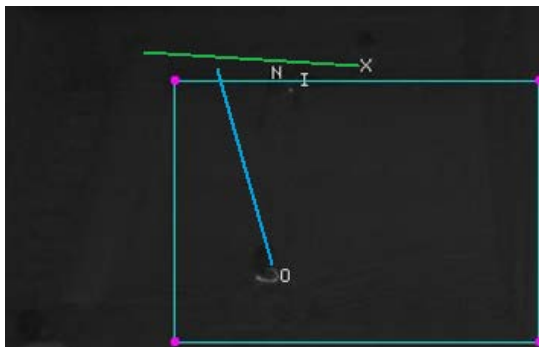
2:24



2:27



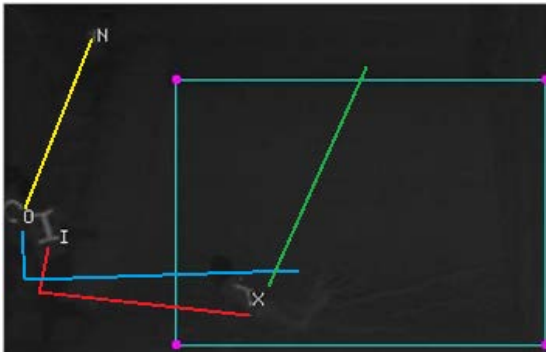
2:30



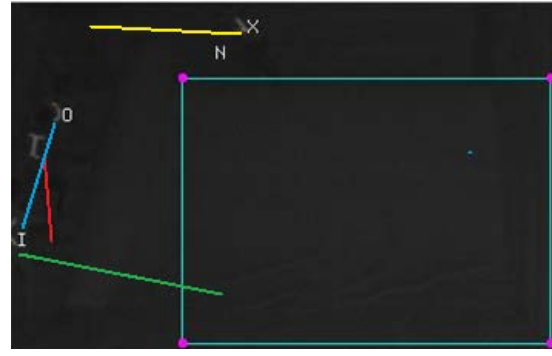
2:33



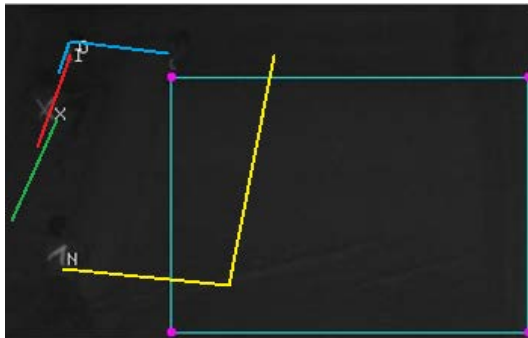
2:36



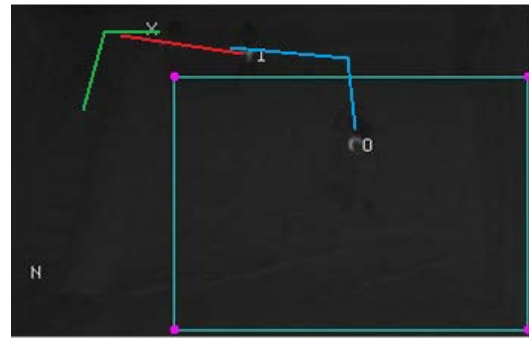
2:39



2:42



2:45



Como se ve a lo largo de todo este experimento en particular, el usuario que corresponde a la letra N tiene tendencia a desaparecer de la imagen una vez se ha deslizado por el tobogán, en estos casos, obviamente, perdemos el rastro del sujeto. Como un ejemplo claro de ello podemos tomar las imágenes en los minutos 0:36 y 0:42, en las que se pierde al sujeto y se reencuentra subiendo por las escaleras, si no se hubiese realizado un correcto refinado de la interpolación se habría producido una línea recta desde el punto A (minuto 0:36) al punto B (minuto 0:42) que no sería realista con respecto a lo que realmente está sucediendo. Otro ejemplo donde se puede ver este mismo caso es el de los minutos 2:27 y 2:33.