

Autor: Karim Muñoz Fernández
Tutor: Dr. Fernando García Fernández
Fecha: 27 de Febrero de 2015



UNIVERSIDAD CARLOS III DE MADRID

ESCUELA POLITÉCNICA SUPERIOR

Departamento de Ingeniería de Sistemas y Automática

PROYECTO FIN DE CARRERA

**Fusión sensorial para la detección y seguimiento de
peatones.**

Índice general

| | |
|--|----|
| 1. Introducción..... | 9 |
| 2. Propósito | 10 |
| 3. Estado del arte | 11 |
| 3.1 Fusión Ladar y LWIR..... | 11 |
| 3.1.1 Algoritmo Ladar..... | 11 |
| 3.1.2 Algoritmo LWIR | 15 |
| 3.1.3 Fusión | 16 |
| 3.2 Fusión láser y cámara | 17 |
| 3.2.1 Sistema de arquitectura | 18 |
| 3.2.2 Detección probabilística basada en el filtro de partículas | 21 |
| 3.3 Seguimiento de trayectorias | 24 |
| 3.3.1 Arquitectura y procesamiento..... | 25 |
| 3.3.2 Fusión de datos Multi-Sensor..... | 26 |
| 3.3.3 Conclusiones | 27 |
| 3.4 La fusión de múltiples sensores de video..... | 27 |
| 3.4.1 Seguimiento..... | 28 |
| 3.4.2 Proyección de la aproximación de Gauss | 29 |
| 3.4.3 Proyección de probabilidad mapa completo..... | 31 |
| 3.5 Conclusiones | 32 |
| 4. Herramientas | 34 |
| 4.1 Hardware del sistema..... | 34 |
| 4.1.1 Telémetro Láser..... | 34 |
| 4.1.1.1 Especificaciones técnicas..... | 35 |
| 4.1.2 Sistema de Visión..... | 35 |
| 4.1.2.1 Especificaciones técnicas..... | 35 |
| 4.1.3 GPS | 35 |
| 4.2 Software del sistema..... | 36 |
| 4.2.1 Matlab R2008b | 36 |
| 5. Descripción detallada del sistema | 37 |
| 5.1 Datos | 39 |
| 5.1.1 Datos GPS | 39 |
| 5.1.2 Datos Láser | 40 |

| | | |
|-------|----------------------------------|----|
| 5.1.3 | Datos Cámara | 41 |
| 5.2 | Detecciones..... | 42 |
| 5.2.1 | Detección Telémetro láser. | 42 |
| 5.2.2 | Detección Cámara..... | 43 |
| 5.2.3 | Fusión de detección..... | 44 |
| 5.3 | Cálculos realizados..... | 47 |
| 5.4 | Impresión..... | 50 |
| 6. | Resumen general del sistema..... | 52 |
| 7. | Análisis de pruebas..... | 54 |
| 7.1 | Caso 1. | 54 |
| 7.1.1 | Resultados caso 1 a 5m..... | 55 |
| 7.1.2 | Resultados caso1 a 10m..... | 55 |
| 7.2 | Caso 2. | 56 |
| 7.2.1 | Resultados caso 2 a 5m..... | 57 |
| 7.2.2 | Resultados caso 2 a 10m..... | 58 |
| 7.3 | Caso 3. | 58 |
| 7.3.1 | Resultados caso 3 a 5m..... | 59 |
| 7.3.2 | Resultados caso 3 a 10m..... | 60 |
| 7.4 | Caso 4. | 60 |
| 7.4.1 | Resultados caso 4 a 5m..... | 61 |
| 7.4.2 | Resultados caso 4 a 10m..... | 62 |
| 7.5 | Caso 5. | 62 |
| 7.5.1 | Resultados caso 5 a 5m..... | 63 |
| 7.5.2 | Resultados caso 5 a 10m..... | 64 |
| 7.6 | Caso 6. | 64 |
| 7.6.1 | Resultados caso 6 a 5m..... | 65 |
| 7.6.2 | Resultados caso 6 a 10m..... | 66 |
| 7.7 | Caso 7. | 66 |
| 7.7.1 | Resultados caso 7 a 5m..... | 67 |
| 7.7.2 | Resultados caso 7 a 10m..... | 68 |
| 8. | Discusión de los resultados..... | 69 |
| 9. | Acciones futuras..... | 71 |
| 9.1 | Cámaras..... | 71 |

| | | |
|------------|---------------------------|----|
| 9.2 | Láser | 71 |
| 9.3 | Cálculos | 71 |
| 9.4 | GPS | 71 |
| 10. | Presupuesto | 73 |
| 11. | Conclusiones | 74 |
| 12. | Referencias | 75 |

Índice de imágenes

| | |
|--|----|
| <i>Imagen 1. Vista de un conductor.</i> | 9 |
| <i>Imagen 2. Sistema LWIR utilizado</i> | 12 |
| <i>Imagen 3. Resultado del documento.</i> | 17 |
| <i>Imagen 4. Resultado de detección.</i> | 20 |
| <i>Imagen 5. Resultado de seguimiento.</i> | 23 |
| <i>Imagen 6. Detección visual</i> | 24 |
| <i>Imagen 7. Detección infrarroja</i> | 24 |
| <i>Imagen 8. Trayectoria tras la fusión</i> | 25 |
| <i>Imagen 9. Ejemplo de sistema multisensor</i> | 28 |
| <i>Imagen 10. Sistema de fusión propuesto en el documento.</i> | 32 |
| <i>Imagen 11</i> | 34 |
| <i>Imagen 12</i> | 34 |
| <i>Imagen 13. Bumblebee 2</i> | 35 |
| <i>Imagen 14. Bumblebee XB3</i> | 35 |
| <i>Imagen 15. Ejemplo de sistema GPS.</i> | 36 |
| <i>Imagen 16. Matlab</i> | 36 |
| <i>Imagen 17. Perfil del sistema presentado.</i> | 37 |
| <i>Imagen 18. Detección de la cámara.</i> | 38 |
| <i>Imagen 19. Detección del Láser</i> | 38 |
| <i>Imagen 20. Doble detección.</i> | 39 |
| <i>Imagen 21. Detección de láser desde el láser y la cámara.</i> | 43 |
| <i>Imagen 22. El recuadro rojo representa a la detección de la cámara.</i> | 44 |
| <i>Imagen 23</i> | 48 |
| <i>Imagen 24. Curva de la campana de Gauss.</i> | 49 |
| <i>Imagen 25. Secuencia del seguimiento de la detección.</i> | 50 |
| <i>Imagen 26. Detección completa.</i> | 51 |
| <i>Imagen 27. Detección caso 1 a 5m.</i> | 54 |
| <i>Imagen 28. Detección caso1 a 10m.</i> | 54 |
| <i>Imagen 29. Resultados caso 1 a 5m</i> | 55 |
| <i>Imagen 30. Resultados caso 1 a 10m.</i> | 56 |
| <i>Imagen 31. Detección caso 2 a 5m.</i> | 56 |
| <i>Imagen 32. Detección caso 2 a 10m.</i> | 57 |
| <i>Imagen 33. Resultados caso 2 a 5m.</i> | 57 |
| <i>Imagen 34. Resultados caso 2 a 10m.</i> | 58 |
| <i>Imagen 35. Detección caso 3 a 5m.</i> | 58 |
| <i>Imagen 36. Detección caso 3 a 10m.</i> | 59 |
| <i>Imagen 37. Resultados caso 3 a 5m.</i> | 59 |
| <i>Imagen 38. Resultados caso 3 a 10m.</i> | 60 |
| <i>Imagen 39. Detección caso 4 a 5m.</i> | 60 |
| <i>Imagen 40. Detección caso 4 a 10m.</i> | 61 |
| <i>Imagen 41. Resultados caso 4 a 5m.</i> | 61 |
| <i>Imagen 42. Resultados caso 4 a 10m.</i> | 62 |
| <i>Imagen 43. Detección caso 5 a 5m.</i> | 62 |
| <i>Imagen 44. Detección caso 5 a 10m.</i> | 63 |
| <i>Imagen 45. Resultados caso 5 a 5m.</i> | 63 |
| <i>Imagen 46. Resultados caso 5 a 10m.</i> | 64 |

| | |
|--|----|
| <i>Imagen 47. Detección caso 6 a 5m.</i> | 64 |
| <i>Imagen 48. Detección caso 6 a 10m.</i> | 65 |
| <i>Imagen 49. Resultados caso 6 a 5m.</i> | 65 |
| <i>Imagen 50. Resultados caso 6 a 10m.</i> | 66 |
| <i>Imagen 51, Detección caso 7 a 5m.</i> | 66 |
| <i>Imagen 52. Detección caso 7 a 10m.</i> | 67 |
| <i>Imagen 53. Resultados caso 7 a 5m.</i> | 67 |
| <i>Imagen 54. Resultados caso 7 a 10m.</i> | 68 |

Índice de ecuaciones

| | |
|---|----|
| <i>Ecuación 1</i> | 12 |
| <i>Ecuación 2</i> | 13 |
| <i>Ecuación 3</i> | 14 |
| <i>Ecuación 4</i> | 14 |
| <i>Ecuación 5</i> | 14 |
| <i>Ecuación 6</i> | 14 |
| <i>Ecuación 7</i> | 14 |
| <i>Ecuación 8</i> | 15 |
| <i>Ecuación 9</i> | 15 |
| <i>Ecuación 10</i> | 15 |
| <i>Ecuación 11</i> | 18 |
| <i>Ecuación 12</i> | 18 |
| <i>Ecuación 13</i> | 19 |
| <i>Ecuación 14</i> | 19 |
| <i>Ecuación 15</i> | 19 |
| <i>Ecuación 16</i> | 20 |
| <i>Ecuación 17</i> | 20 |
| <i>Ecuación 18</i> | 20 |
| <i>Ecuación 19</i> | 21 |
| <i>Ecuación 20</i> | 21 |
| <i>Ecuación 21</i> | 21 |
| <i>Ecuación 22</i> | 21 |
| <i>Ecuación 23</i> | 22 |
| <i>Ecuación 24</i> | 22 |
| <i>Ecuación 25</i> | 23 |
| <i>Ecuación 26</i> | 23 |
| <i>Ecuación 27</i> | 27 |
| <i>Ecuación 28</i> | 27 |
| <i>Ecuación 29</i> | 29 |
| <i>Ecuación 30</i> | 29 |
| <i>Ecuación 31</i> | 30 |
| <i>Ecuación 32</i> | 30 |
| <i>Ecuación 33</i> | 31 |
| <i>Ecuación 34</i> | 31 |
| <i>Ecuación 35</i> | 32 |
| <i>Ecuación 36. Fórmula de la investigación 1</i> | 47 |
| <i>Ecuación 37. Fórmula de la investigación 2</i> | 47 |
| <i>Ecuación 38. Fórmula de la investigación 3</i> | 47 |
| <i>Ecuación 39. Fórmula de la investigación 4</i> | 48 |
| <i>Ecuación 40. Fórmula de la investigación 5</i> | 48 |
| <i>Ecuación 41. Fórmula de la investigación 6</i> | 48 |
| <i>Ecuación 42. Fórmula de la investigación 7</i> | 48 |
| <i>Ecuación 43. Fórmula de la investigación 8</i> | 48 |

Índice de Tablas

| | |
|--|----|
| <i>Tabla 1</i> | 10 |
| <i>Tabla 2</i> | 10 |
| <i>Tabla 3</i> | 55 |
| <i>Tabla 4</i> | 55 |
| <i>Tabla 5</i> | 57 |
| <i>Tabla 6</i> | 58 |
| <i>Tabla 7</i> | 59 |
| <i>Tabla 8</i> | 60 |
| <i>Tabla 9</i> | 61 |
| <i>Tabla 10</i> | 62 |
| <i>Tabla 11</i> | 63 |
| <i>Tabla 12</i> | 64 |
| <i>Tabla 13</i> | 65 |
| <i>Tabla 14</i> | 66 |
| <i>Tabla 15</i> | 67 |
| <i>Tabla 16</i> | 68 |
| <i>Tabla 17. Presupuesto de Material</i> | 73 |
| <i>Tabla 18. Presupuesto de Personal</i> | 73 |
| <i>Tabla 19. Presupuesto total</i> | 73 |

1. Introducción.

En los últimos años, se ha mostrado un gran interés en la necesidad de mejorar la capacidad de seguridad y la vigilancia para una infinidad de situaciones y áreas. En este proyecto vamos a investigar cómo crear una mayor seguridad mediante la vigilancia en la parte delantera de un coche. Lo principal en la seguridad, es detectar la variedad de múltiples objetos que se puedan dar en el campo de visión, y seguir los posibles movimientos y trayectorias para la toma de una decisión final.



Imagen 1. Vista de un conductor.

La detección y el seguimiento, los vamos a realizar mediante una cámara situada en la parte más alta del vehículo que nos ofrecerá una detección de posibles peatones con una cierta probabilidad para los distintos puntos de nuestro plano de visión, seguido de un láser colocado en la parte frontal que ofrece otra detección de posibles peatones con probabilidad seguramente distinta a la anterior. Con la fusión de estas dos señales obtendríamos una nueva probabilidad que nos ofrecerá mayor exactitud a la hora de detectar y seguir a los peatones con lo que seremos capaces de realizar un control fiable y en tiempo real del objetivo de esta investigación.

Inicialmente, se van a utilizar un sistema de fusión de múltiples sensores, donde hace que sea una característica muy interesante dando robustez, buen rendimiento y fiabilidad al sistema.

2. Propósito

Todas las personas hoy en día, buscan la máxima seguridad en todos los ámbitos y aspectos de la vida. Constantemente lo buscamos en casa, en el trabajo, en un viaje, etc... para todo buscamos seguridad.

Múltiples son los datos que vemos cada día en las noticias de situaciones en las que se ha ignorado el hecho de estar o no lo suficientemente seguros en cada momento. Siempre se ven robos, accidentes, muertes infortunadas, etc... que no son para nada agradables y aun mucho menos si se sufren con un familiar, con alguien conocido, o incluso sólo con llegar a presenciarlo.

Este motivo es el que nos mueve a llevar a buscar un sistema en un coche, que nos ayude a obtener dicha seguridad a la hora, en concreto, de la conducción. Muchos son los despistes que se pueden producir en trayectos cortos, conocidos, o cuando estamos cansados. Y es prácticamente imposible, por no decir imposible, que se eliminen con un sistema que evite todo tipo de situaciones peligrosas.

Tabla 1

| | Año 1993 | Año 2003 | Año 2012 |
|--------------------------------------|------------|------------|------------|
| Parque de vehículos | 17.809.987 | 27.006.203 | 32.962.502 |
| Accidentes por cada 10.000 vehículos | 45 | 40 | 27 |
| Muertos por cada 10.000 vehículos | 4 | 2 | 1 |
| Muertos por cada 1.000 accidentes | 80 | 54 | 23 |
| Heridos por cada 1.000 accidentes | 1.466 | 1.507 | 1.394 |
| Muertos por cada 10.000 habitantes | 1,63 | 1,28 | 0,41 |

El propósito de este proyecto está en intentar ayudar a mejorar dicha seguridad.

Muchos son los accidentes producidos por atropellos a lo largo de los años en distracciones, despistes o faltas incluso de atención, y muchas han sido las mejoras que se han hecho, se hacen, y se harán para evitarlos como ABS (Antibloqueo de frenos), BAS (Asistencia en frenada de emergencia), TCS (Control de tracción), ESP (Control de estabilidad), ACC (Control de crucero adaptativo), LDWS (Sistema de detección de carril), etc... y muchos más tanto activos como pasivos que podríamos nombrar.

El propósito de este proyecto va a ser, ya que no se pueden eliminar por completo todos los accidentes ni los heridos, ni fallecidos; intentar ayudar aún más a evitar estos fatídicos acontecimientos y seguir contribuyendo a la tendencia descendente de estos acontecimientos.

Tabla 2

| | Año 2003 | Año 2008 | Año 2012 |
|------------------------------------|----------|----------|----------|
| Accidentes con víctimas | 99.987 | 93.161 | 83.115 |
| Fallecidos en accidentes | 5.399 | 3.100 | 1.903 |
| Letalidad en Accidentes de tráfico | 3,46 | 2,31 | 1,62 |

3. Estado del arte

En este apartado vamos a comentar una serie de documentos, en los cuales explican cómo puede haber distintas formas para realizar una detección múltiple y además poder realizar el seguimiento de dichas detecciones a través de una serie de algoritmos. Todo esto realizado en los últimos años, nos va a servir como base para nuestro proyecto.

3.1 Fusión Ladar y LWIR

Para comenzar en este apartado hemos encontrado un primer documento [1] en el que se centraron en la detección de personas en movimiento y su seguimiento desde un vehículo no tripulado creado por Susan M. Thornton, Mike Hoffelder, and Daniel D. Morris, “Multi-sensor Detection and Tracking of Humans for Safe Operations with Unmanned Ground Vehicles”. Para ello han usado la mezcla de datos de un láser (3D LADAR); y una cámara de infrarrojos de onda larga (LWIR). Con estos elementos han intentado tener en cuenta varios problemas reales a la hora de la detección; problemas como personas paradas, distintos ambientes de detección, distintas apariencias que pueden adoptar las personas detectadas, también tienen el desafío de diferenciar a una persona propensa de una estructuras de personas...

En resumen, estos problemas pueden dar lugar a falsas alarmas debido a las limitaciones de los algoritmos. Su objetivo principal es aplicarlo a un ambiente militar en el que puedan conseguir un vehículo totalmente autónomo capaz de vigilar los 360° de su alrededor. Para ello deben de utilizar varias modalidades de sensores que les permitan adaptarse a varios ambientes tanto de día como de noche. Otro problema es la gente parada que lo intentará adaptar como una mejora a las características de los datos del algoritmo del Ladar.

A la información del Ladar le complementan la información de la cámara para conseguir estos logros en una gran variedad de contextos. Presentan un modelo estadístico y morfológico para la detección de humanos en las imágenes de la cámara con algoritmos que hacen más simple la detección siempre que no sean comparados con el medioambiente, que no es lo que se proponen; lo que proponen es hacerlo robusto frente a los desafíos mencionados.

Van a utilizar un par de Ladar que cada uno abarque unos 180° del campo de visión y una serie de LWIR cámaras. Cada uno de los Ladar recoge una serie de datos, a unos 10Hz, en una red basada en un mapa de 2D sincronizado con el otro sensor, mientras tienen decidido qué hacer con la concatenación de los mapas debido a que se les superpone entre ellos un pequeño ángulo.

3.1.1 Algoritmo Ladar

Usan un algoritmo para detectar el movimiento de vehículos y personas con uno de los Ladar. Para los datos de configuración del dúo de Ladar aparece un algoritmo como un singular mapa de profundidad, cuyo enfoque es aplicado al nuevo dúo de datos sin modificación.

Los dos objetivos clave de este documento son buscar objetos en las escenas y analizar sus movimientos. La detección de los objetos la hacen en dos pasos, primero eliminan las devoluciones de tierra y segundo aplican una técnica de la región de edificio contiguo que aprovecha la información de adyacencia del ángulo creado por el Ladar en el mapa de profundidad; cuyos requerimientos computacionales son bajos para cualquier objeto. La efectividad de hallar objetos humanos depende de cómo el algoritmo elimine la tierra, para ello encontraron un modelo que usaba planos horizontales que trabajaba bien eligiendo el rango, pero usaba puntos irreales, algunos demasiado lejanos que eran indeseables por añadir una complejidad computacional.

La tierra es etiquetada por un ángulo elevado de computación entre los puntos vecinos y el ángulo del mapa de profundidad. Usando el mapa de tierra en adición al ángulo del mapa de profundidad, una altura desde el suelo que vuelve es estimada. Usan esta altura para eliminar los puntos menores de 0.25 m. Este enfoque tiene la ventaja de remover toda la hierba y la vegetación, así como los bancos en un ambiente urbano, todo lo que puede dificultar su algoritmo si no es eliminado.



Imagen 2. Sistema LWIR utilizado

Utilizan un modelo de densidad de probabilidad de superficie para el registro de objetos en 3D. El enfoque del registro se basa en modelar explícitamente la superficie del objeto como una mezcla de gaussianas en 3D, centradas en cada punto muestreado:

$$\rho_s(fX) = \sum_i N(x_i, \sigma_i^2) / n.$$

Ecuación 1

La función que usan es un sumatorio de cada uno de los puntos en una distribución normal de media X y covarianza σ . Las covarianzas son proporcionales a la densidad de muestreo, y a su vez a la distancia del Ladar. Los modelos están registrados y calificados por la optimización de la medida de similitud Bhattacharya [2], que compara dos funciones de densidad y da una estimación absoluta, que permite la mejor calidad de un punto para ser evaluado. La medida de similitud les sirve también para resolver ambigüedades en puntos coincidentes y evitar oclusiones. Una implementación discreta mediante el filtrado de convolución permite el registro en tiempo real sin ser atrapados por mismos locales.

La mayor parte del trabajo de la detección de objetos en movimiento se consigue con la agrupación y registro, pero hay un número de fuentes de desorden, así como un número de objetos que aparecen y desaparecen debido a la oclusión. Estos efectos se dan en estimaciones falsas de movimientos y por lo tanto falsos positivos. El uso de un seguidor del filtro de Kalman [3] puede minimizar estos efectos, mediante la aplicación de coherencia del movimiento.

Para la clasificación de características y con el fin de hacer detecciones en objetos en movimiento y estacionarios, los datos del Ladar proporcionan información importante a cortas distancias. Implementaron una nueva característica que elimina las falsas alarmas, ya que hallaron desordenes aleatorios, cosas como vegetación, hierba o cosas que se presentan en el medio.

La forma detectarlas era agrupando puntos en los objetos a través de su mapa de profundidad de crecimiento en la región; es decir, que proyectaban los puntos agrupados de cada región en 3D en planos 2D. Así creaban una plantilla de 32*16 (binario) que se alinea con el eje mayor de agrupación. Una medida de cómo se distribuyen uniformemente a través de una rejilla en 2D es lo que se refieren como el factor de relleno.

$$ff = 1 - \frac{(\# \text{ empty bins})}{(\text{total} \# \text{ bins})}.$$

Ecuación 2

El resultado que las plantillas de binarios en 2D de detección de verdaderos humanos, será más o menos uniformemente distribuida, por el contrario una plantilla de desorden aleatorio se forma irregular con puntos concentrados en pequeñas agrupaciones del mapa binario. Cuando varias de estas pequeñas agrupaciones se juntan en el conjunto de limitaciones de las detecciones humanas dan lugar a los errores y a que el valor del factor de relleno sea mucho menor. El valor de este factor de relleno disminuye también a medida que se alarga la distancia de la detección, incluso posiblemente por debajo del valor del umbral especificado; pero sin embargo para poder compensar esta disminución, lo calculan mediante la acumulación de rendimientos a lo largo de varios fotogramas, que incluso mejora enormemente la forma en detalle del conjunto proyectado. Aunque el factor de relleno era capaz de eliminar muchas alarmas y desordenes, no era capaz de eliminar objetos con el tamaño similar al de un ser humano, objetos como barriles o postes; pero para ello desarrollarán otro modelo en 2D Ladar con un modelo humano determinado.

Han desarrollado una efectiva fuerza de detección (strength-of detection (SoD)), cuyo valor lo asocian a cada grupo en cada fotograma de los datos del Ladar. El factor de relleno no refleja el aumento de la confianza en la clasificación que resulta de varias detecciones como un ser humano en varios cuadros de datos. Debido a esto implementan una medida que tuviera en cuenta si el objeto detectado es del tamaño adecuado para un ser humano, si se mueve a una velocidad realista o está parado y si la detección persiste en el tiempo.

La confianza la definen en cuatro factores: Cs: tamaño ; Cf: forma ; Cv: velocidad y Cl: vida. El tamaño (Cs), está vinculado estrechamente con el ancho (w), la altura (h) y la profundidad (d), de cada grupo de puntos que definen a los humanos.

$$C = c_s * c_f * c_v * c_l,$$

Ecuación 3

$$c_s = \frac{1}{3}(t_w + t_d + t_h)$$

Ecuación 4

$$t_i = \begin{cases} 1.0, & \text{if } i \leq i_{\max} \\ i_{\max} / i, & \text{otherwise} \end{cases}, \quad i = w \text{ or } d$$

Ecuación 5

$$t_h = \begin{cases} 1.0 & \text{if } h_{\min} \leq h \leq h_{\max} \\ h_{\max} / h & \text{if } h > h_{\max} \\ h / h_{\min} & \text{if } h < h_{\min} \end{cases}.$$

Ecuación 6

Para definir la forma utilizan una fórmula ponderada del factor de relleno, donde aparece el propio factor de relleno (ff) calculado a partir de la densidad combinada 2D, el número de cuadros (Lh) y una constante determinada experimentalmente (Wf). El valor de la forma aumenta para grupos que están constantemente en el tamaño humano.

$$c_f = \frac{1.0}{1.0 + \exp(-\omega_f - ff * l_h)},$$

Ecuación 7

Puesto que los seres humanos están limitados en la velocidad, la usan como otra variable a su medida de confianza. En el caso de que la velocidad sea alta, es probable que se trate de otro objeto que no sea humano, como un coche; por eso la confianza bajaría.

$$c_v = \begin{cases} 1.0, & \text{if } v \leq v_{\max} \\ v_{\max} / v, & \text{otherwise} \end{cases}$$

Ecuación 8

Además cuanto más tiempo pasa de la detección de un mismo grupo de puntos más seguros están de que se trata de la identificación de un ser humano, donde se compone de la vida de agrupación del número total de fotogramas que han sido rastreados (l), el número de fotogramas que ha sido de tamaño humano (l_h), y una variable tomada experimentalmente (W).

$$c_l = \frac{1.0}{1.0 + \exp(-\omega_l - l_h * \frac{(l_h - 1)}{l})}$$

Ecuación 9

La fuerza de detección que caracteriza al algoritmo proporciona un medio para hacer una operación completa del receptor analizando el rendimiento de los algoritmos. Utiliza un conjunto de base de datos para el análisis del Ladar cuyo algoritmo consta de varios escenarios y cada uno de ellos de más de un minuto de duración. En estos casos los seres humanos en movimiento y otros estacionarios siempre realizan las mismas trayectorias, para hacer difícil su detección variaron la velocidad de los humanos en movimiento y la del sensor, además de poner obstáculos como otros coches. Al medir el rendimiento del algoritmo definieron dos medidas, una para detectar caso ciertos y otra para falsos; el conjunto de ellas no podía superar el de humanos. Todo objeto que fuera detectado mayor que el dato umbral era humano, además que el dato umbral era modificado para cada escena. Su rendimiento al final fue del 99%, con menos de una alarma falsa por segundo.

3.1.2 Algoritmo LWIR

El algoritmo produce regiones de interés humano mediante un proceso de dos etapas en que primero agrupa los extractos de las primeras detecciones y después las clasifica en base a un número de características geométricas. Calculan la desviación de intensidad de las imágenes globales y locales, donde X_{ij} en la intensidad de los píxeles y m y σ son la media y la desviación estándar que se han calculado de forma global o en una pequeña ventana alrededor del píxel:

$$n_{ij} = (x_{ij} - m) / \sigma$$

Ecuación 10

La información del borde se obtiene mediante el gradiente, además con el uso de umbrales determinados empíricamente se crean las imágenes binarias a partir de las imágenes de la desviación local y global, así como en el mapa de borde. Para un trabajo futuro lo orientarían a la selección probabilísticamente de los umbrales establecidos en cada caso. La dilatación morfológica y la limpieza se usan antes de la agrupación del vecino más cercano.

La segunda parte del algoritmo calcula simples características geométricas que se utilizan para retener a solo aquellos grupos que tienen parecidos geométricos a los humanos en la naturaleza. Las dos características geométricas son la relación del eje mayor para la agrupación del eje menor y la relación entre el número de píxeles del perímetro en el grupo para el número de píxeles del borde. Definen dos medidas umbrales, E_t y A_t ; tales que si las medidas si son menores que la relación de los ejes A_t y mayor que la relación del borde E_t , son considerados como humanos. Las características y los umbrales fueron elegidos para no eliminar medidas grandes y otras medidas no ideales de seres humanos.

Para establecer un rendimiento básico para el algoritmo, se centraron en analizar conjuntos de datos que son ideales, tales como que los seres humanos tienden a ser más calientes que sus alrededores, y esto puede llevar a confusión.

La variación en la resolución con un rango es un reto para el algoritmo LWIR, puesto que detectar a una persona en un rango corto puede ser más difícil que en el rango normal de medida. Sin embargo, una resolución baja a largas distancias puede dar una imagen homogénea térmicamente que no distinga nada; o una alta resolución a distancias cortas puede dar una imagen con mayor variación térmica en la que se distingan muy claramente los detalles. Además un ser humano tiene cara, extremidades, etc; que sin información adicional no podrían unirse a la detección del ser humano, para ello la incorporación del conocimiento sobre el rango del objeto desde el sensor, la unión de las partes del ser humano puede realizarse.

Su algoritmo inicial para el LWIR lo basaron en obtener una medida mayor del ser humano detectado que de su entorno, además que esta medida era aceptable para entornos nocturnos o climas más fríos; pero lo era lo que querían, robusto y poder usarlo en diferentes condiciones medioambientales. Aprovechando la calibración térmica del sensor quieren evaluar los umbrales térmicamente, en un análisis preliminar demuestran que los humanos generan un rango de emisiones térmicas a través de su cuerpo, que varía debido al material de las ropas, así como los reflejos del medio ambiente.

3.1.3 Fusión

En este primer apartado en cuanto a lo que nos centramos de este trabajo, podemos decir que la fusión la realiza en base a una mezcla de las dos partes anteriormente descritas (Ladar y LWIR), uno por la detección rápida y fiable acortas distancias y el otro por la detección térmica a largas distancias. La forma más común de fusión son técnicas de visión estéreo [4], pero estas pueden llevar a métodos computacionalmente densos y requerir procesamientos especiales con el fin de llegar a obtener los porcentajes que se requieren en este documento. Entonces, para ello aprovechan la información de profundidad del Ladar para reducir la carga computacional eliminando todas las regiones cercanas de las imágenes LWIR del procesamiento estéreo. Cualquier acercamiento a la fusión depende de las medidas registradas con precisión de los sensores que intervienen, tanto intrínsecamente como extrínsecamente.



Imagen 3. Resultado del documento.

En conclusión de este documento, desarrollaron un sistema que permitía detectar personas a cortas (Ladar) y largas distancias (LWIR), con dos sistemas que aunque fueran distintos les permitía complementarlos; es decir, lo que uno no conseguía lo hacía el otro.

3.2 Fusión láser y cámara

En el documento de Xuan Song, Jinshi Cui, Huijing Zhao and Hongbin Zha, “A Bayesian Approach: Fusion of Laser and Vision for Multiple Pedestrians Tracking” que tiene relación con esta investigación [5].

Intenta plantear el problema de la interacción múltiple a la hora de la detección y seguimiento. En este documento presentan un sistema con el uso de un láser que ofrece un mapa robusto para la detección; y la visión de una cámara que ofrece información a color para dar un control fiable con un alto rendimiento y en tiempo real de la detección y seguimiento de personas entre multitudes. La fusión de este sistema contiene dos aspectos: las características de la fusión de nivel y la estrategia.

La primera parte se encarga de extraer de forma automática la información útil del láser y de la visión combinándola en un marco bayesiano, para luego formar un máximo a posteriori y obtener un resultado de rastreo. En la segunda parte como el escáner láser proporciona detecciones robustas, pueden realizar la detección a base de seguimiento visual que lo hace más fiable. Además con el fin de calcular la estimación del mapa y teniendo en cuenta las interacciones y oclusiones de los peatones, proponen la detección en una probabilidad basada en un filtro de partículas que tiene incorporado la detección de las interacciones e información dinámica del modelo mediante el ajuste de los pesos del mismo. Utilizan el filtro de partículas para cada objetivo, de tal manera que las partículas se muestran en un espacio de estados pequeño; con esto se obtiene una mejor aproximación y una complejidad computacional aceptable.

En este trabajo se han ido basando en una serie de documentos que compartían problemas. Problemas como la condición cambiante de la luz, la interacción y oclusión de objetos además de un vasto uso de tiempo. Todas estas complejidades hacían difícil utilizar métodos de los que parten en sistemas medioambientales reales, debido a que no los controlaban. Por otro lado se han basado en métodos de detección y seguimiento mediante el láser que puede realizar detecciones de humanos automáticas en escenas variables. Esos métodos son insensibles a los cambios de luz y los procesados de datos del láser no son muy largos [6], [7], [8]. Muchos sistemas fueron inventados en el pasado pero sus limitaciones los

hacen inherentes y obvios, ya que no pueden dar cierta información como el color de los objetos. Es difícil obtener grupos de características que hagan distinguir unos objetos de otros.

Según comentan, recientemente han salido a la luz varios documentos [9], [10], que detectan y siguen personas que combinan la información, según las distancias, del láser y la información de visión obtenida desde las cámaras. Mientras que todos estos sistemas detectan a una o un par de personas, no puede ser aplicado en sistemas de vigilancia y seguimiento real en el medio, debido a que se centran en la detección y seguimiento de múltiples peatones en un ambiente real.

3.2.1 Sistema de arquitectura

El objetivo de su sistema es incorporar la visión del láser y la cámara en un mismo marco para obtener resultados en el seguimiento más precisos. Pueden aprovechar la capacidad de detección del escáner láser para facilitar la extracción de la información de visión de cada objeto. Primero se realiza la detección con el láser y después se fusiona con la de los sensores de visión en un mismo marco bayesiano, donde usan un método de máximo a posteriori eficiente.

Hay dos subsistemas. El subsistema del escáner láser contiene un modelo LMS291, que tiene un alcance de 80 metros con un error de 3centímetros, situado a una altura de 1,20 metros para una exploración horizontal; donde los datos de los objetos como su movimiento se extraen en un sistema de coordenadas rectangulares. El subsistema de visión utiliza los datos del láser, para localizar la región del cuerpo en la imagen de vídeo en función de la altura, de la posición del cuerpo y del calibrado del modelo de la cámara; en el cuál usan menos de 11 puntos de control [11]. El rectángulo utilizado para modelar la región en el plano de imagen es descrito por (p^B) el punto central del cuerpo en el plano de la imagen y (s) el factor ancho y alto del rectángulo. Estos dos parámetros pueden ser estimados por la cámara fácilmente, puesto que dependen de la diferencia de anchura y de altura entre humanos con una medida constante.

Formulan el problema de seguimiento secuencial, como calcular el máximo a posteriori (x^*) tal que $x^* = \arg \max_{x \in X} P(x|y)$, donde $x = (x_0, \dots, x_t)$ es la secuencia de estado y $y = (y_0, \dots, y_t)$ es la secuencia de observación. Una base de estimación secuencial bayesiana puede ser descrita como una recursividad de dos pasos: predicción ($p(x_t|y_{t-1})$) y filtrado ($p(x_t|y_t)$).

$$p(\mathbf{x}_t | \mathbf{y}_{t-1}) = \int p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{y}_{t-1}) d\mathbf{x}_{t-1}$$

Ecuación 11

$$p(\mathbf{x}_t | \mathbf{y}_t) = \frac{p(\mathbf{y}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{y}_{t-1})}{p(\mathbf{y}_t | \mathbf{y}_{t-1})}$$

Ecuación 12

La recursividad requiere la especificación de un modelo dinámico describiendo el estado de evolución $p(x_t|x_{t-1})$ y un modelo de probabilidad a la luz de las medidas de corriente $p(y_t|x_t)$. La recursividad es iniciada con alguna distribución para el estado inicial $p(x_0)$. Para modelar el estado de una persona individual en un tiempo (t), se modela usando un vector de estado de 5 dimensiones; en donde k es el identificador de la persona, $p_{k,t}^L$ es el baricentro que denota el grupo de puntos del láser, y $p_{k,t}^B$ denota la posición del cuerpo. Cada uno de ellos son vectores de 2 dimensiones excepto s que es el factor de escala para cada persona.

$$x_{k,t} = [p_{k,t}^L, p_{k,t}^B, s]$$

Ecuación 13

El modelo dinámico en su sistema de seguimiento, utiliza un modelo de velocidad constante que se describe mejor por una ecuación de segundo orden autorregresiva; en donde las matrices A, B, C y el Σ serían aprendidas de un grupo de secuencias representativas donde las pistas correctas han sido obtenidas del experimento. La distribución normal $N(0, \Sigma)$ es una gaussiana de ruido con media 0 y desviación típica estándar de 1.

$$x_{k,t} = Ax_{k,t-1} + Bx_{k,t-2} + CN(0, \Sigma)$$

Ecuación 14

El modelo de observación del láser es la “elipse humana” que se genera por una agrupación de puntos. Para finalmente extraerla, primero tienen que obtener la imagen del plano del escaneo del láser a través de la sustracción del fondo, después para agrupar los puntos utilizan la ventana Parzen [12] cuya estimación de densidad es un conocido método no paramétrico que estima la distribución de los datos de la muestra, la cual utilizan para convertir los puntos simples discretos en la función de densidad continua.

$$p(\mathbf{I}) = \frac{1}{n} \sum_{i=1}^n G(\mathbf{I} - \mathbf{I}_i, h)$$

Ecuación 15

En donde $\{I_1, \dots, I_n\}$ es un conjunto de muestras d-dimensionales, G es la ventana gaussiana de la función y h es el parámetro de anchura que depende del tamaño de la región del cuerpo. La interacción entre personas que están cerca, causa pérdidas en las detecciones de algunas de ellas. Para evitar esto introducen una función que es capaz de crear dos elipses en vez de una para dos personas que están cercanas. Para una persona aislada cuyo estado es $x_{k,t}$ evalúan la probabilidad de observación y_t , donde $I(p)$ es la intensidad para p de la imagen del láser después del proceso de la ventana de Parzen y $\varphi(p_{k,t}^{(i)}, p_{k,t}^{(j)})$ es la función potencial:

$$P_{laser}(y_t | \mathbf{x}_{k,t}) = I(\mathbf{p}_{k,t}^L) \varphi(\mathbf{p}_{k,t}^{(i)}, \mathbf{p}_{k,t}^{(j)})$$

Ecuación 16

$$\varphi(\mathbf{p}_{k,t}^{(i)}, \mathbf{p}_{k,t}^{(j)}) = \begin{cases} \exp\left(-\frac{\max|\mathbf{p}_{k,t}^{(i)}, \mathbf{p}_{k,t}^{(j)}|}{w}\right), & \max|\mathbf{p}_{k,t}^{(i)}, \mathbf{p}_{k,t}^{(j)}| > w \\ 1, & \text{otherwise} \end{cases}$$

Ecuación 17

Donde $p_{k,t}^{(i)}$ es el punto de muestreo en el grupo, $|p_{k,t}^{(i)}, p_{k,t}^{(j)}|$ es la distancia entre dos puntos del grupo de muestreo, y w es el límite de la distancia dependiendo de la diferencia del ancho del ser humano.



Imagen 4. Resultado de detección.

Para la observación de la visión, adoptan un modelo multicolor; basado en un histograma de color Tono-Saturación-Valor, porque el espacio de color HSV puede separar la información cromática de los efectos de sombreado. La observación de la persona es representada por un histograma de color N-bin extraído de la región $R(p_{k,t}^B)$, centrada en la ubicación $p_{k,t}^B$ de imagen de la cámara. Es denotada como:

$$Q(\mathbf{p}_t) = \{q(n; \mathbf{p}_{k,t}^B)\}_{n=1, \dots, N}$$

Ecuación 18

$$q(n; \mathbf{p}_{k,t}^B) = K \sum_{k \in R(\mathbf{p}_{k,t}^B)} \delta(b(k) - n)$$

Ecuación 19

Donde δ es la función de Kronecker [13] para decidir si k es la ubicación en la región $R(\mathbf{p}_{k,t}^B)$, K es una constante de normalización que asegura $\sum_{n=1}^N q(n; \mathbf{p}_{k,t}^B) = 1$, $b(k) \in \{1, \dots, N\}$ es el índice binario asociado con el vector de color en la posición k del píxel. Esta ecuación define $q(n; \mathbf{p}_{k,t}^B)$ como la probabilidad de un color binario n en un tiempo t . La probabilidad de los datos debe favorecer al histograma de color candidato $Q(\mathbf{p}_{k,t}^B)$ cerca para la referencia del histograma $Q(\mathbf{p}_0)$; debemos elegir una distancia (D) en las distribuciones de color HSV. D es derivada del coeficiente de similitud Bhattacharyya, y definida como:

$$D(q^*, q(\mathbf{p}_{k,t}^B)) = \left[1 - \sum_{n=1}^N \sqrt{q^*(n; \mathbf{p}_0) q(n; \mathbf{p}_{k,t}^B)} \right]^{\frac{1}{2}}$$

Ecuación 20

La probabilidad es evaluada como:

$$P_{vision}(\mathbf{y}_t | \mathbf{x}_{k,t}) \propto e^{-\lambda D^2(\mathbf{p}_{k,t}^B, \mathbf{p}_0)}$$

Ecuación 21

Por tanto la observación conjunta para el láser y la visión es:

$$P_{joint}(\mathbf{y}_t | \mathbf{x}_{k,t}) \propto P_{laser}(\mathbf{y}_t | \mathbf{x}_{k,t}) P_{vision}(\mathbf{y}_t | \mathbf{x}_{k,t})$$

Ecuación 22

3.2.2 Detección probabilística basada en el filtro de partículas

El filtro de partículas debido a su eficiencia, flexibilidad y fiabilidad, ha sido una técnica de aproximación numérica exitosa para la aproximación bayesiana secuencial, sin modelos lineales o gaussianos. Como ya se dijo propusieron un filtro de partículas que incorporaba la detección del láser y la información de la visión para una correcta detección y seguimiento ante múltiples objetos, el cual no sólo se ocupa de hacer una buena aproximación a los objetos sino también de las interacciones entre estos y las oclusiones.

La idea tras el filtro de partículas es simple. Comienzan con un conjunto ponderado de muestras $\{w_t^n, x_t^n\}_{n=1}^N$, distribuidos aproximadamente acordes a $p(x_{t-1}|y_{t-1})$, donde se generan nuevas muestras a partir de una distribución propuesta diseñada adecuadamente, que puede depender del antiguo estado y de nuevas medidas. Para mantener un ejemplo consistente, la nueva importancia de los pesos se establecen como:

$$w_t^{(n)} \propto w_{t-1}^{(n)} \frac{p(y_t | \mathbf{x}_t^{(n)})p(\mathbf{x}_t^{(n)} | \mathbf{x}_{t-1}^{(n)})}{q(\mathbf{x}_t^{(n)} | \mathbf{x}_{t-1}^{(n)}, y_t)}, \quad \sum_{n=1}^N w_t^{(n)} = 1$$

Ecuación 23

De vez en cuando es necesario volver a muestrear las partículas para evitar la generación de pesos de poca importancia. Uno de los temas cruciales es la elección de de la distribución $p(x_t|x_{t-1}, y_t)$. En su sistema pueden ejecutar un filtro para cada persona detectada y su distribución es:

$$q_k(\mathbf{x}_t | \mathbf{x}_{t-1}, y_t) = \sum_{j=1}^{n_k} \beta_{kj} q_D^j(\mathbf{x}_t | \mathbf{x}_{t-1}, y_t) + (1 - \sum_{j=1}^{n_k} \beta_{kj}) \times p_k(\mathbf{x}_t | \mathbf{x}_{t-1})$$

Ecuación 24

Donde $q_D^j(x_t|x_{t-1}, y_t)$ es la información de la detección que tiene influencia en el objeto k ; β_{kj} sus pesos y $p_k(x_t|x_{t-1})$ el modelo dinámico. Cuando un objetivo no interactúa con los demás, tenemos una propuesta de alto peso, que es como una detección simple; pero cuando se producen interacciones, se producen varias detecciones que compiten en la propuesta, por tanto tiene bastante eficacia en las interacciones y oclusiones entre objetos. Esto se convierte en una distribución de pico múltiple, aunque pueden compartir las mismas detecciones en la interacción son diferentes entre sí por dos razones: los pesos β_{kj} son diferentes y los objetivos tienen una información dinámica diferente. Para cada objetivo las partículas tienen un alto peso en su estado real, y la posterior es aproximada a su optimización. Cuando fusionan dos objetivos, no pierden ningún objetivo porque una detección específica tiene un gran peso en la propuesta respectiva de cada objetivo. Por otro lado debido a cada paso del filtro de partículas por un pequeño espacio, pueden obtener una mejor aproximación y una reducción computacional significativa.

El peso β_{kj} de una detección debe ser una probabilidad que refleje similitud con el objeto, “detecciones probabilísticas”. Con el fin de considerar las interacciones entre los objetos, utilizan el modelo de probabilidad conjunta (JPDA) [14], para calcular β_{kj} definen un evento de asociación (θ) expresado como un vector de dimensión n_d , que son el número de detecciones. Cada θ determina de forma única cómo cada una de las detecciones se asigna a una persona específica. Este vector se puede extraer de un grupo de número como $\{0,1,2, \dots\}$ y $\theta(j) = k$ significa que la detección j es para el objeto k . Entonces θ para cada probabilidad conjunta y β_{kj} son:

$$L(\theta) = \prod_{j=1}^{n_s} P_{\text{joint}}(y_t^j | \mathbf{x}_{k,t}^{\theta(j)})$$

Ecuación 25

$$\beta_{kj} = \alpha \sum_{\theta \in \Theta_{kj}} L(\theta) \times P_{\text{prior}}$$

Ecuación 26

Donde Θ_{kj} es el grupo de asociaciones conjuntas de eventos que incluye todos los casos de detección j siendo desde el objeto k , α es el factor de normalización que asegura las sumas de todo β_{kj} hasta θ y el conocimiento previo de P que depende de la exactitud de las detecciones.



Imagen 5. Resultado de seguimiento.

Experimentaron con una cámara situada a unos 5 metros de altura y el láser a 1,20 metros, capturando datos a una frecuencia de 15fps. Hicieron varias pruebas en las que había fuerte y poca luz y varios puntos con oclusiones y cruces, en los lugares donde el láser detectaba y la cámara ayudaba a su detección se producía un alto rendimiento, en el resto había muchas dificultades. Por tanto en las zonas donde su sistema completo actuaba no había fallos. Utilizaron además otros dos métodos distintos para compararlo (la detección basada en filtro de partículas que es el que ellos proponen, un conjunto de filtros de partículas y filtros independientes de partículas) y su resultado fue que con las interacciones sólo el método que ellos proponen era capaz de hacer el seguimiento, además el tiempo promedio utilizado era menor.

En conclusión han realizado la detección a través de un escáner láser y la visión por medio de una cámara; para extraer la información útil y combinarla en un marco bayesiano, además de proponer una estimación eficaz del método probabilístico de detección basado en filtro de partículas.

3.3 Seguimiento de trayectorias

La relevancia del documento de Lauro Snidaro y Gian Luca Foresti, “Sensor Fusion for Video Surveillance”, [15], es debido a que se centraron en la importancia de vigilar espacios públicos, para evitar ciertos acontecimientos que recientemente han marcado la historia del mundo. Desean mejorar la capacidad de vigilancia y control de operadores humanos en espacios de gran tamaño. Además se pueden realizar por el interés puesto en una fusión en tiempo real con una nueva generación de sistemas de vigilancia.

Para esta investigación nos interesa el sistema infrarrojo y el uso de un sistema óptico, que extrae las imágenes con movimiento para calcular el seguimiento, la clasificación y los procedimientos de fusión de datos. Esta técnica se basa en la localización y el reconocimiento de objetos en varios niveles; en el primer nivel detecta los objetos sobre un plano 2D y transforma estas posiciones a un plano 3D que hace más fácil su seguimiento en niveles posteriores, además de que cada nivel se compromete a la vigilancia de un sub-área del medio supervisado.



Imagen 6. Detección visual



Imagen 7. Detección infrarroja

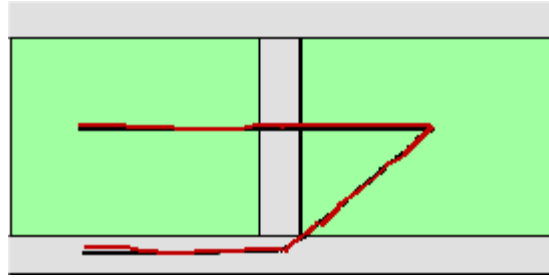


Imagen 8. Trayectoria tras la fusión

En este sistema han dado mucha importancia al ruido producido y a su medición a la hora de seguir la trayectoria del objeto, debido a que la medida puede ser muy variante en estados posteriores en función de esta medida. Este problema se puede reducir con la utilización de más sensores, de hecho en los últimos avances en cámaras han hecho a éstas más deseables que los sensores.

En la fusión de datos que es lo que nos interesa para este proyecto, hay que decir que la robustez no nos la dan un conjunto de cámaras poco fiables y sí un sistema de sensores altamente capacitados para estas funciones.

3.3.1 Arquitectura y procesamiento

Proponen un sistema de procesamiento diferente para cada área, con sus propios sensores estáticos y nodos de procesamiento. Un nodo es responsable del funcionamiento de los algoritmos correspondientes y de la detección de objetos ya sea por diferenciación de imágenes, el filtrado de éstas o por la burbuja de extracción. Este último paso es más para el seguimiento, se produce a nivel del sensor de movimiento e identifica las regiones de la imagen a través de los algoritmos de detección de cambios.

La detección de movimiento y la extracción de burbuja se explotan siguiendo un planteamiento de la sustracción de la capa de fondo. La detección de cambios se realiza mediante un algoritmo para el cálculo del umbral automático basado en números de Euler. La capa de fondo es actualizada mediante un filtro de Kalman. Marco a marco se aplica la sustracción también en relación a mejorar los resultados de detección, además filtros morfológicos también se aplican para mejorar la calidad de las notas extraídas mediante la eliminación de píxeles falsos debido al ruido, y por mejorar la conectividad de las regiones. Cada región tiene dos sensores (cámara óptica e infrarroja) y monitorizan la misma región, cada uno actúa independientemente para la misma zona, y en una zona en la que quizá puede haber más de un objeto o poca iluminación, entre ambos sensores son capaces de extraer el objeto perfectamente.

En un sistema de vigilancia de vídeo pueden existir varios objetivos en movimiento y para seguir a todos los objetos simultáneamente se necesita un sistema multisensor que sea capaz de fusionar las medidas independientes de cada uno de los objetos. Durante el seguimiento producido a nivel local utilizan un algoritmo de asociación de objetos entre las imágenes previas y las posteriores. La técnica seguida es la mostrada en el documento [16].

Para llevar a cabo la fusión de datos que se ha mencionado antes, es necesario un marco común además de los sensores registrados en él. Un mapa 2D con vista desde arriba se puede tomar como un marco común de coordenadas, o incluso un GPS desde una vista más amplia. Como resultado de una geometría

proyectiva entre un píxel y una superficie plana viene dada por una homografía plana (para saber a qué distancia están los obstáculos con una sola cámara [17]). Dicho píxel elegido para representar un objeto y ser transformado a las coordenadas del mapa es la proyección del centroide del objeto. Cada procedimiento local sólo tiene en cuenta los objetos en el campo de visión del sensor correspondiente. Para cada objeto conocido en el instante de tiempo anterior, sólo las mediciones que entran en una distancia que bloquea se consideran. La distancia de Mahalanobis se puede utilizar para determinar la región de validación, paso que reduce la probabilidad de asociaciones erróneas debido al ruido (la distancia de Mahalanobis está relacionada con la distancia Euclídea). La trayectoria en el mapa de arriba que vista de todos los objetos se modela a través de un filtro Kalman lineal, donde el vector de estado $x = (x, V_x, y, V_y)$ está constituido por la posición y velocidad del objeto en el mapa. En cada cuadro (el sistema procesa 25 fotogramas por segundo) donde se recibe un nuevo indicador de la posición.

Las estimaciones de la posición de los diferentes sensores se fusionan de una manera centralizada. La fusión de datos se realiza teniendo en cuenta la fiabilidad del sensor en cada instante de tiempo.

3.3.2 Fusión de datos Multi-Sensor.

La fusión que se lleva a cabo en este documento usa un enfoque de filtro de Kalman con el fin de obtener mejores estimaciones de la posición de los objetos observados. Sin embargo, dos sistemas de fusión fueron considerados durante los experimentos: la fusión de las mediciones y el seguimiento de la fusión.

El algoritmo de fusión de medición es teóricamente óptimo, mientras que el seguimiento de la fusión tiene unos requisitos menos automatizados, aunque son sub-óptimos para regiones con baja iluminación y en sensores de baja calidad (ruidosos). La ejecución de un filtro de Kalman para cada pista, obtiene una estimación de filtrado de la posición del objetivo que permite suavizar las variaciones extremas debido a los errores de segmentación. El régimen actual empleado fue track-to-track dado por las limitaciones computacionales. De hecho, durante los experimentos, la alta frecuencia de las medidas y requisitos de tiempo real no permiten tener en cuenta la información de la regeneración.

El proceso para cada objetivo comprende 3 pasos: la recopilación de mediciones disponibles de los sensores locales, agrupación y asignación de las medidas para cada destino conocido en el instante de tiempo anterior y la actualización del estado de cada objetivo por la alimentación de las estimaciones asociadas al filtrado para el algoritmo de fusión. Tienen en cuenta que el segundo paso se realiza con la restricción de que sólo una única medida de un sensor determinado se va a asociar con un único objetivo en la lista mantenida por el procedimiento de fusión. Para regular el proceso de fusión de forma automática en función del rendimiento de los sensores, presentan una medida de confianza al peso de las estimaciones locales.

La explotación de fusión de datos puede mejorar la precisión de seguimiento por la arquitectura de múltiples cámaras que se ha propuesto. La fusión de datos recogidos de diferentes sensores requiere la determinación de mediciones de precisión para que puedan ser fusionadas en una manera ponderada. Sin hacer distinción entre las medidas, esto podría llevar a la inestabilidad del filtro con estimaciones erróneas y sobre todo al mal funcionamiento de los sensores. Los fallos de hardware o de las condiciones desfavorables de iluminación podrían ofrecer un rendimiento muy pobre y generar errores de segmentación. Por lo tanto, teniendo en cuenta todas las mediciones el objetivo es obtener una medida mejor que la de un único sensor. La idea fundamental es obtener del filtro de Kalman una estimación fusionada más sesgada por las medidas exactas y no afectada por las inexactas. La capacidad de respuesta del filtro para las mediciones se puede ajustar a través de una matriz de covarianza de error de medición R . Si los valores propios de la matriz particular R son más pequeños que los de la otra, la medición correspondiente tendrán un peso mayor.

La siguiente medida llamada Relación de Aspecto (AR) da un valor al grado de confianza asociado con la j -ésima del objeto extraído en un tiempo t del sensor s :

$$AR(\mathbf{B}_{j,t}^s) = \frac{\sum_{x,y \in \mathbf{B}_{j,t}^s} D(x,y)}{|\mathbf{B}_{j,t}^s|c}$$

Ecuación 27

Donde $D(x,y)$ es el mapa de diferencia obtenida como la diferencia absoluta entre la imagen actual y la referencia, y c es una constante de normalización en función del número de tonos de color utilizados en la imagen. La AR es, pues, un número real que va de 0 a 1 que da una estimación del nivel de rendimiento de cada sensor para cada nota extraída. Los valores de AR de las notas extraídas del sensor de infrarrojos son considerablemente más altos que los extraen de uno óptico. Los valores de AR se utilizan para regular el error de medición de la matriz de covarianza de los datos de posición de peso en el proceso de fusión. La siguiente función fue desarrollada para el error de medida de la posición:

$$r(\mathbf{B}_{j,t}^s) = GD^2(1 - AR(\mathbf{B}_{j,t}^s))$$

Ecuación 28

Donde GD es la distancia “gating”. La función por tanto se utiliza para ajustar el margen de error de medición a fin de que las posiciones del mapa calculado para objetos con valores de AR de alta son de confianza más (cercano a cero), mientras que los objetos mal detectados (valor bajo de AR) son de confianza inferior (el error de medición es igual a la distancia de activación periódica).

3.3.3 Conclusiones

En este trabajo, la fiabilidad del sensor se consideraba explícitamente en un sistema múltiple de cámaras de video-vigilancia de ambientes al aire libre. La confianza se ha definido de forma automática a la redundancia de la medida de la fuerza de la ubicación de los objetivos, procedente de los diferentes sensores del proceso de fusión de datos. De esta manera los errores de localización debido a la incorrecta segmentación de los objetivos se ha reducido, así como la calibración debido a las transformaciones de la perspectiva de los errores. Los resultados experimentales les han mostrado la eficacia de la confianza eligiendo la medida de ponderación del sensor automático y la mayor precisión alcanzable por los datos propuestos en la fusión. En particular, en el procedimiento de fusión se han producido trayectorias que son más continuas y por lo tanto útiles para un sistema de vigilancia.

3.4 La fusión de múltiples sensores de video

En este último documento de Lauro Snidaro, Ingrid Visentini y Gian Luca Foresti, “Fusing multiple video sensors for surveillance” [18], vamos a tratar la detección en tiempo real, el seguimiento, reconocimiento y comprensión de la actividad de los objetos en movimiento en múltiples sensores que representan las cuestiones fundamentales que deben resolverse con el fin de desarrollar sistemas de vigilancia que sean capaces de monitorizar autónomamente ambientes amplios y complejos. En este contexto, la fusión de sensores desempeña un papel fundamental en la gestión de la información, además de mejorar el rendimiento del sistema. Aquí se presenta un marco de fusión novedoso para combinar los datos provenientes de sensores múltiples y heterogéneos observando un área de vigilancia.

La fusión de los sensores se puede aplicar de diferentes maneras y en diferentes niveles:

Señal fusión nivel. Este es el más bajo nivel en el que la fusión puede llevarse a cabo. En el ámbito de la videovigilancia la fusión de la imagen es muy conocida en la comunidad de la teledetección, en las imágenes pancromáticas y multispectrales se fusionan para obtener resultados de alta resolución espacial y espectral. La señal fusión nivel es la más restrictiva, ya que generalmente sólo permite que los sensores del mismo tipo que se fusionen (por ejemplo, sensores de imagen).

Característica fusión nivel. Se trata de un tipo de nivel más alto de fusión con respecto a la anterior. La fusión de datos a nivel de característica facilita la combinación de tipos heterogéneos de sensores (por ejemplo, audio y video). Puede ser el caso por ejemplo de las señales de color y el borde que puede ser fusionado con el seguimiento de un objeto.

La fusión también se puede realizar en los niveles superiores, pero en este trabajo se centran en la característica fusión nivel. Comentan que existe bastante sobre la fusión de imágenes, pero no se ha hecho mucho por la fusión de múltiples sensores para la vigilancia de vídeo con la no superposición de campos totalmente de visión. En la mayoría de los casos, de matrices de sensores para la vigilancia se despliegan para maximizar la cobertura, pero el costo de las cámaras inferiores permite tener múltiples sensores de vigilancia de la misma escena. Además, es necesario extraer las características de cada sensor de visión que tienen que fundirse en un formato de representación común. En particular, presentan una forma aumentada de la fusión de sensores con diferentes condiciones de visibilidad por probabilidad proyectada. La principal novedad que transmite esta obra es la explotación de los mapas de probabilidades obtenidos para cada objetivo de cada sensor y su proyección en un mapa común como medio para llevar a cabo la fusión de sensores.

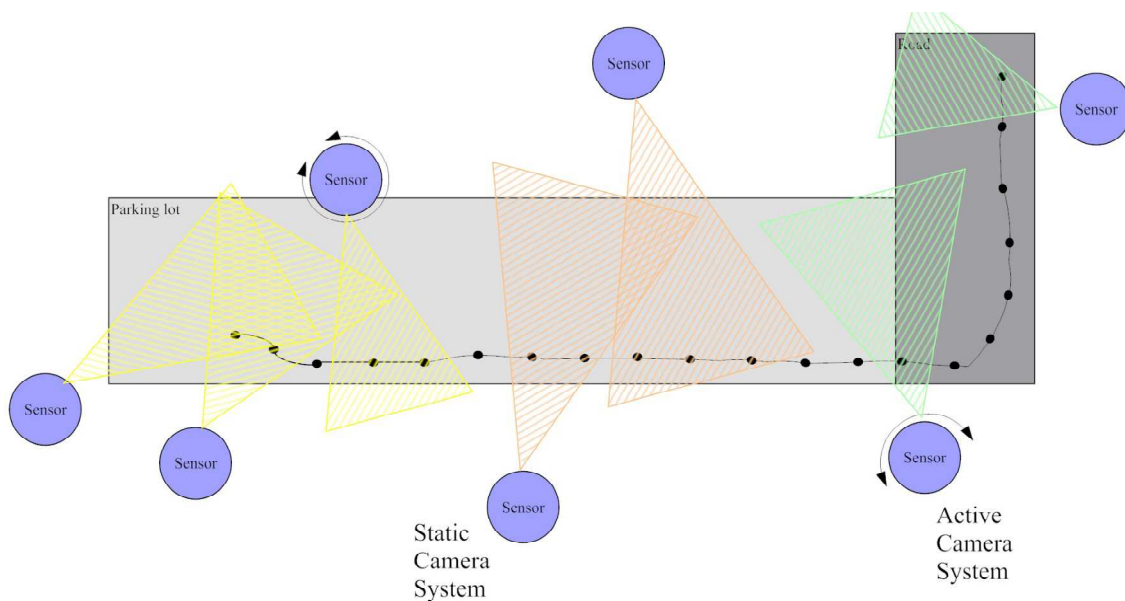


Imagen 9. Ejemplo de sistema multisensor

3.4.1 Seguimiento

Para estos investigadores parte importante fusión fue el paso previo, la detección, el seguimiento y la extracción de las características de los objetos. Tal fue el asunto, que hicieron un barrido de los últimos años, de las investigaciones y de todos los posibles avances que pudieran ayudar a su investigación, pero siempre teniendo en cuenta las situaciones difíciles como las condiciones de variación de la iluminación, eventos o cambios de apariencia ilimitada de fuentes de vídeo.

Seguimiento de un objeto se puede realizar a diferentes niveles. A nivel del sensor, en el plano de la imagen de cada sensor y otro considera que mezcla diferentes características para mejorar la robustez del seguidor; características que traen una gran mejora en la localización de destino y de seguimiento.

Impulsar ya era una técnica atractiva para la clasificación fuera de línea, pero el trabajo seminal de Oza [Oza 2005] dio más impulso a este tema, proponiendo una aplicación en línea que ha abierto un nuevo mundo de posibilidades. Seguimiento a través de la clasificación, especialmente cuando se realiza con un conjunto de clasificadores, se considera más robusto a oclusiones y cambios de iluminación y además se puede lograr una mayor robustez del sistema y del rendimiento con un conjunto de clasificadores [19], [20]. Este enfoque puede ser de especial interés cuando el clasificador es actualizado por el aprendizaje del aspecto actual de la meta. Impulsar la línea, por ejemplo, combina estos dos requisitos.

Consideramos dos clases posibles $\{w_1, w_2\}$ de manera que $w_1 = +1$ que representa el objetivo, y $w_2 = -1$ el fondo y sea f una función de extracción que toma como entrada una imagen parche x . Dado un conjunto de hipótesis binarias débiles $\{h_1, h_2, \dots, h_M\}$ de modo que cada clasificador $h_m: \mathfrak{R}^d \rightarrow \Omega$ se aplica al descriptor d proporcionado por el extractor de características $f: X \rightarrow \mathfrak{R}^d$ donde X es un conjunto de patrones y $\Omega = \{w_1, w_2\}$ donde un patrón x se asigna a la clase más probable de w si:

$$P(\omega | H(f(x))) = \arg \max_k P(\omega_k | H(f(x)))$$

Ecuación 29

Donde H es el conjunto resultante de la combinación de la h_c . Por el teorema de Bayes, tenemos que:

$$P(\omega | H(f(x))) = \frac{P(H(f(x)) | \omega) P(\omega)}{P(f(x))}$$

Ecuación 30

Donde $p(w | H(f(x)))$ representa la distribución de probabilidad conjunta de las medidas extraídas por los clasificadores combinados en H . Consideran un conjunto clasificador de H al que está basado en la combinación de técnicas en línea Impulso. En particular, cuentan con la cascada de marco [21] para acelerar la aplicación del conjunto y la búsqueda de la meta a múltiples escalas. La combinación obtenida del clasificador se refiere a un solo detector en el sistema de múltiples cámaras, cada destino tiene su propio clasificador. Aplicada a toda la imagen de una búsqueda exhaustiva, los fuertes clasificadores de H generan un mapa de riesgo [22].

3.4.2 Proyección de la aproximación de Gauss

El máximo que aparece en el mapa confianza es la posición (x_c, y_c) del centro del rectángulo de selección que incluye el parche x buscado por el conjunto de clasificadores. El máximo representa la posición más probable de la búsqueda de destino en el marco actual. Este punto prevén que esté en la parte inferior del cuadro como límite, así se obtiene la posición del punto (x_g, y_g) que se supone que es el suelo. Con el fin de fusionar la posición del blanco detectado por los múltiples ángulos de visión de los diferentes sensores, es necesario un marco común de coordenadas. Esto se resuelve mediante la proyección de este último punto (x_g, y_g) de cada sensor a través de una transformación homográfica obtenido durante la puesta en marcha del sistema.

En este trabajo, en lugar de proyectar una estimación única del punto de la posición del objetivo en función de cada sensor, se propone proyectar de asignar una función de verosimilitud calculada como una

aproximación única de Gauss al plano de la región de probabilidad de la búsqueda del mapa de plano de la imagen de cada sensor.

Esta elección tiene una aproximación de la probabilidad multimodal pobre, por tanto, decidieron adoptar en esta primera validación experimental del marco propuesto por los siguientes puntos:

- La probabilidad está representada analíticamente, lo que facilita el procesamiento posterior.
- Facilita la fusión, por lo que la fusión simple bayesiana se puede aplicar.
- El empleo de clasificadores on-line aprendidos para cada objetivo y la búsqueda de áreas grandes restringidas, en gran medida reduce la aparición de las distribuciones de probabilidad multimodales fuertes.

Por lo tanto, eligieron el modelo de riesgo en el plano de la imagen dentro de la región de búsqueda de cada sensor en dos variables de distribución normal $N(X|\mu_i, \Sigma_i)$. En aras de la claridad y para mantener la notación sencilla, nos estamos refiriendo a un solo objetivo detectado por el sensor de i -ésimo. La covarianza Σ_i se calcula simplemente como la varianza de probabilidad máxima de muestra con respecto al punto de valor máximo en (x_c, y_c) como:

$$\Sigma_i = \frac{1}{N} \sum_{n=1}^N \left[\begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} x_c \\ y_c \end{pmatrix} \right] \left[\begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} x_c \\ y_c \end{pmatrix} \right]^T$$

Ecuación 31

La suma se realiza sobre las coordenadas (x_n, y_n) , con $n = 1..N$, donde los índices n de los píxeles en el área de búsqueda tienen un valor más alto que el umbral del clasificador de decisión entre la clase de fondo y la clase plano. La media μ_i está fijada para ser (x_g, y_g) que es el punto que se proyecta a través de la eventual transformación homográfica. La matriz de covarianza Σ_i entonces se puede escalar, teniendo en cuenta los parámetros de escala de la matriz de transformación homográfica, para proyectar la distribución al plano del mapa.

La fusión de las funciones de probabilidad proyectada desde el plano de la imagen se realiza a través de "Fusion via independent likelihood pool". Este enfoque ha sido elegido necesariamente ya que el marco elaborado hasta el momento produce sólo funciones de verosimilitud con respecto a la ubicación de la meta de cada sensor y no una distribución posterior. Las funciones de probabilidad utilizadas se generan a través de aproximación con una sola gaussiana de los datos dentro de la región de búsqueda de los mapas de imagen de probabilidad. A su vez, los mapas de riesgo son el resultado directo de la aplicación en la imagen de origen del conjunto clasificador en línea preparado para detectar el objetivo. El grupo de riesgo independiente regula la distribución posterior del estado x dado el conjunto de medidas $Z^s = \{Z_1 \in \mathcal{Z}_1, \dots, Z_s \in \mathcal{Z}_s\}$ hecho por s sensores donde Z_i es el conjunto de medidas hechas por los i -ésimos sensores. La distribución posterior dado el conjunto de observaciones anteriores es:

$$\begin{aligned} P(\mathbf{x}|Z^s) &= \frac{P(Z^s|\mathbf{x})P(\mathbf{x})}{P(Z^s)} \\ &= \frac{P(z_1, \dots, z_s|\mathbf{x})P(\mathbf{x})}{P(z_1, \dots, z_s)} \end{aligned}$$

Ecuación 32

Y suponiendo la independencia condicional de las medidas señaladas del estado, la probabilidad es:

$$P(\mathbf{z}_1, \dots, \mathbf{z}_s | \mathbf{x}) = P(\mathbf{z}_1 | \mathbf{x}) \dots P(\mathbf{z}_s | \mathbf{x}) = \prod_{i=1}^s P(\mathbf{z}_i | \mathbf{x})$$

Ecuación 33

La distribución posterior viene dada por:

$$\begin{aligned} P(\mathbf{x} | \mathbf{Z}^s) &= P(\mathbf{Z}^s)^{-1} P(\mathbf{x}) \prod_{i=1}^s P(\mathbf{z}_i | \mathbf{x}) \\ &= CP(\mathbf{x}) \prod_{i=1}^s \Lambda_i(\mathbf{x}) \end{aligned}$$

Ecuación 34

Donde x ahora indica la posición del objeto en el plano del mapa y $\Lambda_s(x)$ es la función de probabilidad producida para cada sensor. $P(x)$ es la distribución a priori de x y C es una constante de normalización que garantiza que la fórmula integra a 1 con respecto de x . Si la distribución a priori se toma como uniforme, entonces la fórmula combina todas las medidas actuales sin tener en cuenta la historia previa. Si se instaura el anterior se establece en la fusión posterior obtenida en el instante de tiempo anterior, entonces tiene lugar una estimación completa Bayesiana recursiva. Además, se debe señalar que toda la masa de probabilidad de la gaussianas se concentra en la región donde se supone que el objetivo esté mientras que los lugares alejados en el mapa tienen una probabilidad muy baja pero estrictamente mayor que cero como es requerido por las distribuciones de probabilidad.

3.4.3 Proyección de probabilidad mapa completo

El enfoque de proyección descrito proporciona una manera rápida y eficaz para representar la probabilidad de un objetivo por medio de una aproximación gaussiana individual. Esta representación analítica de las parejas de probabilidad del objetivo se ve muy bien con la técnica de fusión de probabilidad independiente de la piscina y por lo tanto ofrece la solución más fácil para sensores efectivos de fusión heterogéneos para el seguimiento del objetivo. El único enfoque gaussiano puede estar limitado en el sentido de que las distribuciones multimodales no pueden ser representadas.

Los mapas de probabilidad producidos por los conjuntos de clasificadores no son multimodales. Puede darse el caso de que durante las oclusiones u otras condiciones ambiguas los clasificadores no sean capaces de detectar correctamente el objetivo, la probabilidad resultante podría ser multimodal. Las imágenes de origen se han tomado de una secuencia de vídeo de un estacionamiento adquiridas en condiciones de poca luz. Además, las superficies de las probabilidades pueden ser también de color a través de MATLAB para facilitar su visualización.

Para abordar el problema de la representación de probabilidad multimodal hay al menos dos posibilidades. La primera sería utilizar una mezcla de gaussianas para representar la probabilidad multimodal, para cada sensor, la mezcla tendría que calcularse a partir de la probabilidad producida por los clasificadores para cada objetivo (poco práctico). La segunda posibilidad, que se describe es más sencilla. La idea es proyectar las probabilidades, todas producidas por los clasificadores de planos de imagen de los sensores, a la visión de un mapa común de la parte superior del área observada, es decir, para cada sensor y el objetivo de cada uno, realiza una proyección de píxeles racional de los mapas de

riesgo producido por los conjuntos, reproduciendo fielmente todas las características de la probabilidad de imagen plano original.

Para superar estos problemas, han decidido emplear una formulación matemática diferente. En esta formulación la parte $P(x|Z^s)$ se obtiene a través de la marginación de los mapas de riesgo de acuerdo con:

$$P(x|Z^s) = \int P(x|L) (L|Z^s) dL \approx \sum_{i=1}^s P(x|L_i) w_i$$

Ecuación 35

Expresa la probabilidad posterior de la posición x dado el conjunto de observaciones Z_s producidas por los sensores s en términos de la probabilidad de mapas de $P(x|L_i)$ y su "importancia" $P(L_i|Z^s)$ de acuerdo con las mediciones recogidas. Este factor actúa como un peso w_i de la suma de los mapas.

Para resumir podemos decir que se utiliza en un sensor multi-configuración para combinar las probabilidades de diferentes sensores en la parte superior de visión común de mapa de la zona observada. Señalan que en su caso los factores de ponderación w_i se toman uniformes. Consideran la adopción de una estrategia de ponderación dinámica como la investigación futura, la marginación prácticamente se dio cuenta de que como suma de los mapas de probabilidad obtenidos de todos los sensores de observación de la zona y por la normalización posterior.

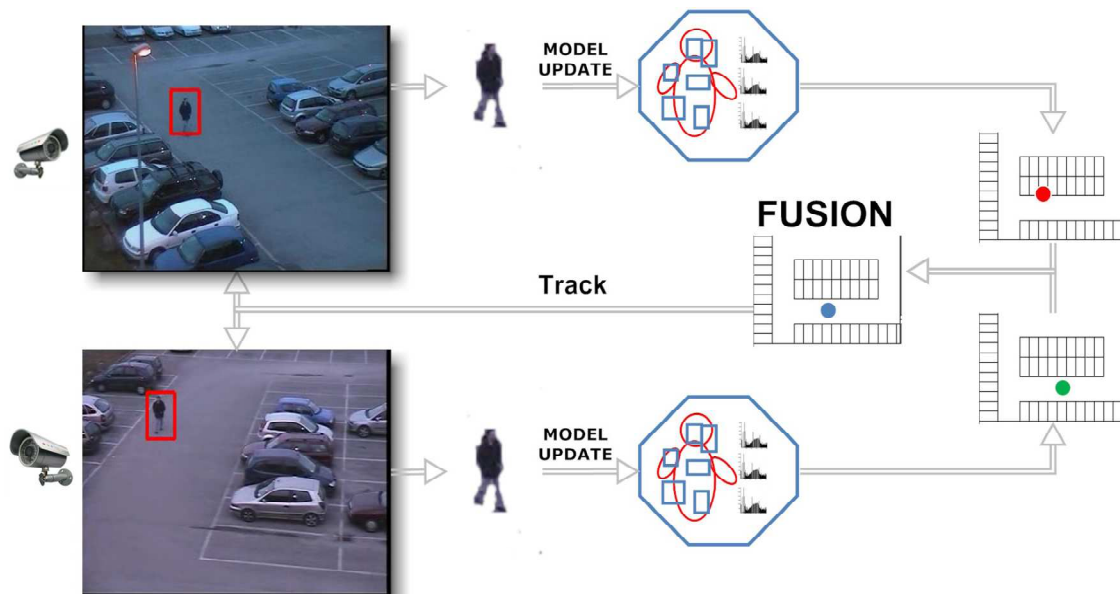


Imagen 10. Sistema de fusión propuesto en el documento.

3.5 Conclusiones

En este trabajo se discuten los beneficios de la fusión de sensores aplicados al dominio de vigilancia de vídeo. A pesar de que la fusión se puede aplicar en muchos niveles diferentes, se han concentrado en la fusión característica de nivel y nos han mostrado cómo es el paso del procesamiento clave para combinar múltiples sensores. Muestran cómo un sistema de vigilancia con varias cámaras pueden explotar la fusión de datos para obtener estimaciones más precisas de la posición de un objetivo. El resultado fue obtenido a través de un novedoso marco que explota la proyección del riesgo producido por los clasificadores en

línea de formación utilizados para el seguimiento del objetivo. Esta proyección permite efectivamente incorporar las incertidumbres de la detección de los diferentes sensores en la estimación final fusionada.

El método tiene la ventaja de la fusión de sensores de forma dinámica, la realización o la selección en función de los datos disponibles. La probabilidad de la fusión propuesta ha demostrado que funcionan mejor que los sensores individuales y de "calidad" media ponderada.

4. Herramientas

4.1 Hardware del sistema

4.1.1 Telémetro Láser

El telémetro láser usado en el proyecto, es el modelo SICK LMS 291-S05, capaz de leer un plano de 2 dimensiones con una resolución desde 1° hasta 0.25° , y con un campo de visión configurable de 100° hasta 180° . La distancia máxima es regulable desde los 8 hasta los 80m y el periodo mínimo de muestreo del láser es de 13ms.



Imagen 11

El funcionamiento de dicho telémetro láser se basa en la tecnología LIDAR (*Light Detection and Ranging* o *Laser Imaging Detection and Ranging*), permitiendo determinar la distancia desde un emisor láser a un objeto o superficie utilizando un haz láser pulsado. El tipo de escaneado es lineal, ya que el telémetro dispone de un espejo rotatorio en un sólo sentido que va desviando el haz láser, produciendo líneas paralelas en el terreno como patrón de escaneado. El inconveniente principal de este sistema es que el giro del espejo es en una sola dirección (no siempre se obtienen detecciones).

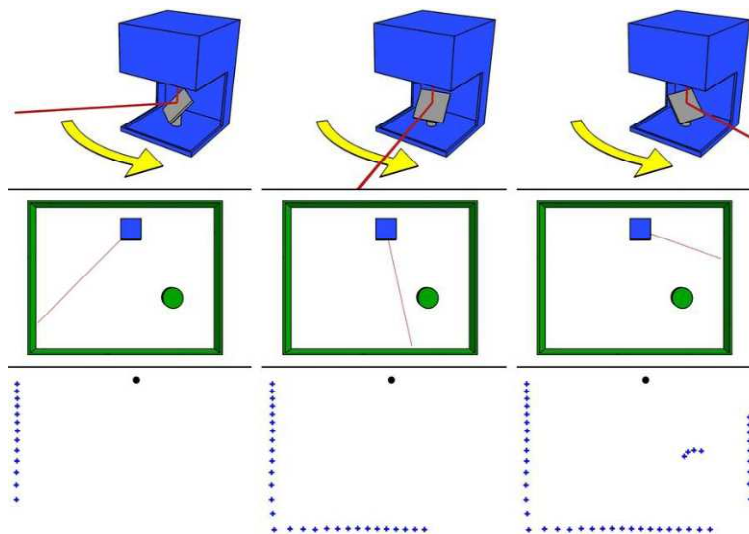


Imagen 12

4.1.1.1 Especificaciones técnicas.

Las especificaciones técnicas del telémetro láser utilizado en dicho proyecto pueden observarse en el Anexo II.

4.1.2 Sistema de Visión

El sistema de visión estéreo que vamos a utilizar es el Bumblebee2 y Bumblebee XB3. Esta cámara implementa tecnología de nueva generación que permite obtener tiempos más rápidos en la adquisición de datos y mejora de la calidad 3D. Incluye la librería Triclops SDK para poder trabajar con la cámara que permite a los usuarios controlar los ajustes de la cámara, ajustar la calidad de la imagen y el acceso en tiempo real a imágenes de rango de profundidad utilizando la tecnología de visión estereoscópica.

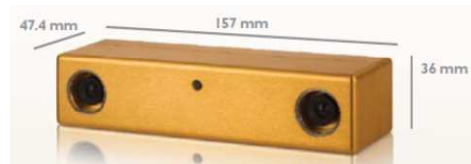


Imagen 13. Bumblebee 2



Imagen 14. Bumblebee XB3

El kit de desarrollo de software FlyCapture (SDK) para la adquisición de imágenes y control de la cámara que se incluye con el Bumblebee2 y Bumblebee XB3, es compatible con Microsoft® Windows® e incluye una cámara controlador de dispositivo, biblioteca de software completa con la programación de aplicaciones Interface (API), programas de demostración y código fuente por ejemplo C / C ++. El SDK FlyCapture también incluye el controlador PGRPRO™, que proporciona el depurado y diagnóstico mejorado y permite a los dispositivos IEEE1394b, ir a 800 Mb / s.

4.1.2.1 Especificaciones técnicas.

Las especificaciones técnicas del telémetro láser utilizado en dicho proyecto pueden observarse en el Anexo II.

4.1.3 GPS

El sistema global de navegación por satélite (GNSS) permite determinar la posición de un objeto, una persona o un vehículo con una precisión hasta de centímetros (si se utiliza GPS diferencial), aunque lo habitual son unos pocos metros de precisión.



Imagen 15. Ejemplo de sistema GPS.

En el presente proyecto se utiliza el GPS como complemento de localización, detección y seguimiento de los peatones. Aunque no se vaya a tener en cuenta realmente para el objetivo fundamental del proyecto, en la fusión, es parte del sistema que integramos.

4.2 Software del sistema

4.2.1 Matlab R2008b

Matlab es un software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje M), y servicio de especie. Además es multiplataforma, ya que es compatible con Unix, Windows, Mac OS y GNU/Linux.

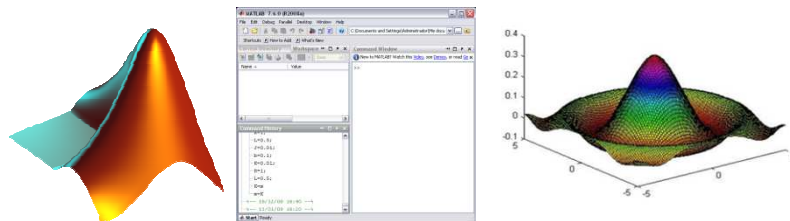


Imagen 16. Matlab

Entre sus prestaciones básicas se hallan la manipulación de matrices, la representación de datos y funciones, la implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con otros dispositivos hardware. Es un software muy usado en universidades y centros de investigación y desarrollo. En los últimos años ha aumentado el número de prestaciones, como la de programar directamente procesadores digitales de señal o crear código VHDL.

El programa Matlab no ha sido utilizado directamente para el desarrollo del presente proyecto, pero sí de una forma indirecta al ser este programa el encargado de crear las bases de datos que almacenan la información proporcionada por los sistemas de detección de obstáculos, así como la información proporcionada por el dispositivo GPS. Dichas bases de datos ofrecen toda la información recogida de los dispositivos de una forma lógica y colocada, algo esencial para poder leer los datos y sobre todo interpretarlos, y con ello poder realizar el cálculo de las zonas de riesgo de los peatones.

5. Descripción detallada del sistema

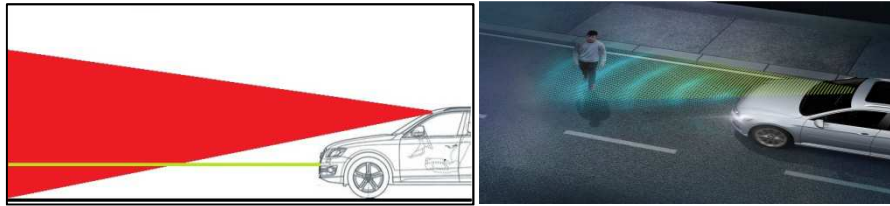


Imagen 17. Perfil del sistema presentado.

En este apartado va a ser explicado más detalladamente un esquema a través del cual será fácil la comprensión de la investigación que vamos a realizar.

El sistema que se pretende mostrar se monta en todo su conjunto en un vehículo de pruebas. Dicho montaje incluye todos los elementos descritos en el apartado 4 más una serie de cálculos previos que serán explicados.

La primera parte consta de la toma de datos. Cada uno de los sistemas tiene su propia serie de datos, tales que nos van a servir en las siguientes partes por ejemplo para hacer los cálculos necesarios, para a hacer también distinción entre ellos, entre los distintos tipos de sensores, diferenciar si ha habido una detección o incluso de si esta es correcta y suficiente como para seguir teniendo en cuenta una detección.

“1 2012.17.1 20:12:36.856 0 01 1209 9.22 23.59 0003 0000 ; 0 01 1209 9.51 28.62 0003 0000 ; 0 01 1209 -6.27 15.62 0002 0000 ; 0 01 1209 -7.42 16.25 0003 0000 ; 0 01 1209 -8.21 18.67 0002 0000 ; 0 01 1209 -6.03 14.30 0002 0000 ; 0 01 1209 -6.29 16.93 0002 0000 ; 0 01 1209 7.62 34.74 0002 0000 ; 0 01 1209 3.20 34.34 0003 0000 ; 0 01 1209 9.94 31.75 0002 0000 ; 1 01 1209 -2.10 10.25 0005 0000 ; 0 01 1209 -2.82 12.47 0002 0000 ;”

La segunda parte es el análisis de estos a priori. Con “A Priori” queremos identificar a la probabilidad de acierto de una secuencia de datos que en teoría tiene cada uno de los elementos. Cada uno tiene una velocidad de toma de datos, por eso habrá más datos de un tipo que de otro. La velocidad de toma de datos es importante porque hará tener un mayor o menor número de detecciones y a su vez un mayor y menor número de detecciones correctas para su análisis.

La tercera parte consta de los cálculos. Se han tenido que calcular tanto las velocidades, como la diferencia de tiempos, así como la probabilidad a priori y a posteriori de las detecciones. Ya que se está haciendo una estimación de una detección futura con cierta probabilidad (A Priori), para luego compararla con la probabilidad de acierto tras la nueva detección (A Posteriori). Esto es lo que nos indica luego el posterior análisis de si se acierta o no y cómo de precisos somos.

La cuarta parte consta del análisis a posteriori. Con “A Posteriori” queremos identificar a la probabilidad que tenemos de haber acertado o no con nuestro diagnóstico previo y de si esa predicción es lo suficientemente buena o no.

La quinta parte consta de la decisión de seguimiento o descarte. Toda detección está sujeta a una serie de reglas que nos dirán si descartamos o no a cada una de las detecciones, se pueden llegar a descartar incluso antes de que se lleguen a hacer unos primeros cálculos. Es como si fuera una constante criba en la sólo las detecciones ideales; es decir, las que queremos que sean buenas según lo que buscamos, serán las que contabilicen para nuestra investigación.

El objetivo de esta investigación es el de detectar, seguir y evaluar si la obtención de datos de la cámara y el láser son los adecuados para la fusión. Para ello vamos a mostrar cómo es una detección de la cámara y el láser.



Imagen 18. Detección de la cámara.

La detección de la cámara la vamos a identificar por el recuadro rojo que se ve en la imagen. Si dicho recuadro rojo está envolviendo al peatón es una detección correcta de la cámara, si no lo hace no lo es.

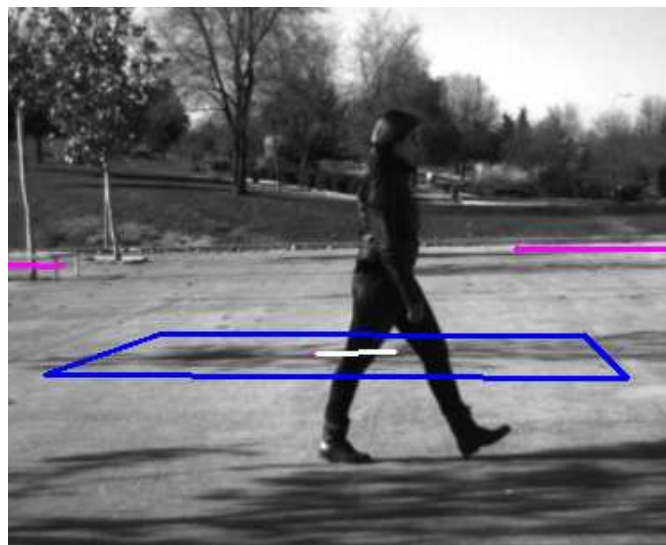


Imagen 19. Detección del Láser

La detección del láser la vamos a identificar por el recuadro azul que se ve en la imagen. Si dicho recuadro azul está envolviendo al peatón es una detección correcta del láser, si no lo hace no lo es.

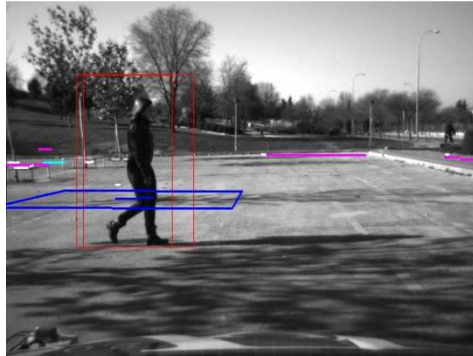


Imagen 20. Doble detección.

Cómo es una detección y cómo las vamos a identificar es importante. No las vamos a analizar visualmente si no que vamos a tomar una serie de datos que vamos a poder analizar y descartar aquellos que no sean necesarios.

5.1 Datos

Se han tomado una serie de lecturas en las que cada uno de los elementos utilizados tiene una serie de datos que conforman los elementos que queremos hacer notar. Por ejemplo el tipo de elemento detectado, la posición, etc... No todos los datos van a ser útiles para esta investigación y pueden ser ignorados.

En todos y cada uno de los tipos de datos mostrados a continuación, se muestra una lectura reducida, la cual podemos identificar claramente por un “;”, mientras que la lectura completa va finalizada en “.”. Esto es importante diferenciarlo porque a la hora de hacer la lectura no es lo mismo una detección en un punto que en otro y todas y cada una de las detecciones van separadas entre sí y finalizadas claramente.

5.1.1 Datos GPS

El GPS es un sistema que a la hora de la fusión, que es de lo que trata en sí este proyecto de investigación, no vamos a tomar en cuenta y va a quedar descartado; pero además de poder incluirlo en posibles trabajos futuros, tiene algunos datos en común con los otros dos sistemas y que lo hace interesante para éste apartado.

“2 1 -9.45 18.16 -9.77 17.27 -9.61 17.71 0.00 0.32 0.89 0.00 0.00”

Primer dato: “2” es el número de identificación del GPS. Cada uno de los tipos de datos tiene uno distinto. Este primer dato es importante para distinguir entre láser, cámara o GPS.

Resto de datos: no son de relevancia para este proyecto, pero son datos como la posición que en caso de ampliar en investigaciones futuras podrían incluirse.

5.1.2 Datos Láser

El sistema de datos que nos proporciona el telémetro láser, es uno de los que sí se va a utilizar para el análisis de si es correcto o no y de si hace al sistema fiable y robusto que es lo que se busca.

“0 2012.17.1 19:6:17.287 259 2 2 2545292 1 29.97 27.22 29.76 26.74 29.86 26.98 0.19 0.22 0.48 0.00 0.00”

Primer dato: “0”, es el número identificativo para el sistema del telémetro láser. Lo vamos a utilizar para distinguir qué tipo de detección tenemos. Toda lectura que comience con este dato en primer lugar nos indicará que son lecturas del telémetro láser.

Segundo dato: “2012.17.1 19:6:17.287”, esto es la fecha y hora a las que se hicieron las pruebas y se tomaron las diferentes lecturas. Va a ser importante porque necesitamos diferenciar entre lectura y lectura y entre secuencia y secuencia.

Tercer dato: “259”, este dato es el número de secuencia. Cada una de las lecturas son identificadas por su secuencia, la cual nos servirá para distinguir entre ellas y para los cálculos.

Cuarto dato: “2 2 2545292 ”, esta serie de dato es ignorada, debido a que no se van a ser utilizados.

Quinto dato: “1”, este dato es el número que no identifica si lo que estamos detectando es un peatón o no. Sólo si aparece un 1 es peatón detectado, resto de valores en esta posición no lo serán.

Sexto dato: “29.97”, es el primer dato con decimales que vamos a tener en cuenta y corresponde a la posición de X del objeto o supuesto peatón detectado.

Septimo dato: “27.22”, es el segundo dato con decimales tomado y nos va a servir para representar la posición de Y del objeto o supuesto peatón detectado.

Resto de datos: “29.76 26.74 29.86 26.98 0.19 0.22 0.48 0.00 0.00”, va a ser ignorados debido a que no nos interesan en esta investigación.

5.1.3 Datos Cámara

El sistema de visión es el tercer y último conjunto de datos a analizar y es la otra parte importante que sí se utiliza para su pre y post análisis en la toma de decisiones de detección y seguimiento como en el apartado de datos del láser.

“1 2012.17.1 19:6:17.329 5 01 259 9.26 23.38 0003 0000”

Primer dato: “1”, es el número identificativo para el sistema de cámaras. Lo vamos a utilizar para distinguir qué tipo de detección tenemos al igual que al telémetro láser y al GPS. Toda lectura que comience con este dato en primer lugar nos indicará que son lecturas de la cámara de visión.

Segundo dato: “2012.17.1 19:6:17.329”, esto es la fecha y hora a las que se hicieron las pruebas y se tomaron las diferentes lecturas. Va a ser importante porque necesitamos diferenciar entre lectura y lectura y entre secuencia y secuencia.

Tercer dato: “5”, este dato es el número que nos identifica si lo que estamos detectando es un peatón o no. Sólo si aparece un 5 es peatón detectado, resto de valores en esta posición no lo serán.

Cuarto dato: “01”, este dato es ignorado, debido a que no va a ser utilizado.

Quinto dato: “259”, este dato es el número de secuencia. Cada una de las lecturas son identificadas por su secuencia, la cual nos servirá para distinguir entre ellas y para los cálculos.

Sexto dato: “9.26”, es el primer dato con decimales que vamos a tener en cuenta y corresponde a la posición de X del objeto o supuesto peatón detectado.

Septimo dato: “23,38”, es el segundo dato con decimales tomado y nos va a servir para representar la posición de Y del objeto o supuesto peatón detectado.

Resto de datos: “0003 0000”, va a ser ignorados debido a que no nos interesan en esta investigación.

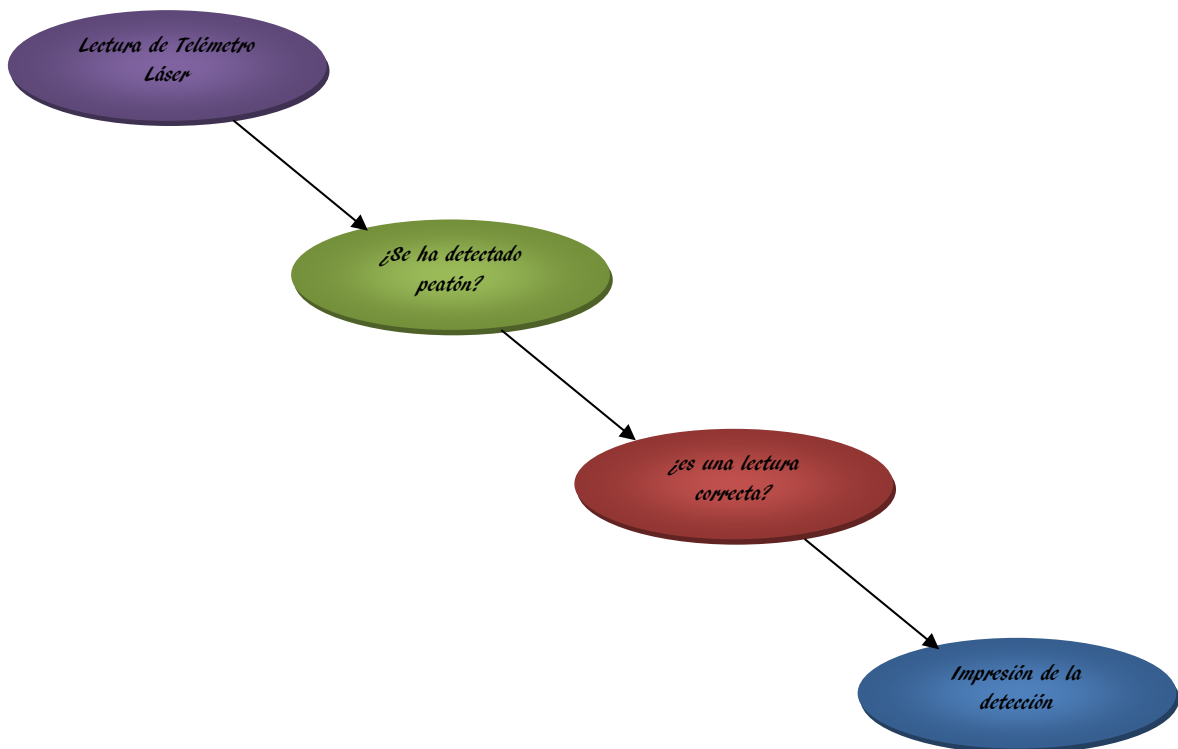
5.2 Detecciones.

El siguiente apartado que vamos a desarrollar son las detecciones, es decir, qué es lo que vamos a considerar en cada uno de los casos y qué tipo de detección es, debido a que vamos a instaurar varios niveles de detección.

5.2.1 Detección Telémetro láser.

Las detecciones del telémetro láser, tal y como se ha identificado en el anterior apartado de Datos, son aquellas que en la lectura se van identificando con el número "0". A partir de aquí nos indica todo lo que detecta como peatones y no peatones y su posición.

Un ejemplo claro de lo que se hace con las lecturas es lo siguiente:



Es un bucle cerrado que constantemente realiza nuevas lecturas. Cuando estas lecturas se toman repetidas veces en el telémetro láser estaríamos haciendo un primer caso de detección. Es el caso de detección simple.

Una detección simple se produce cuando sólo estamos detectando lecturas con uno de los instrumentos o también en el caso de que sean distintas lecturas y no la misma. Esto puede quedar mejor aclarado en el apartado de cálculos.

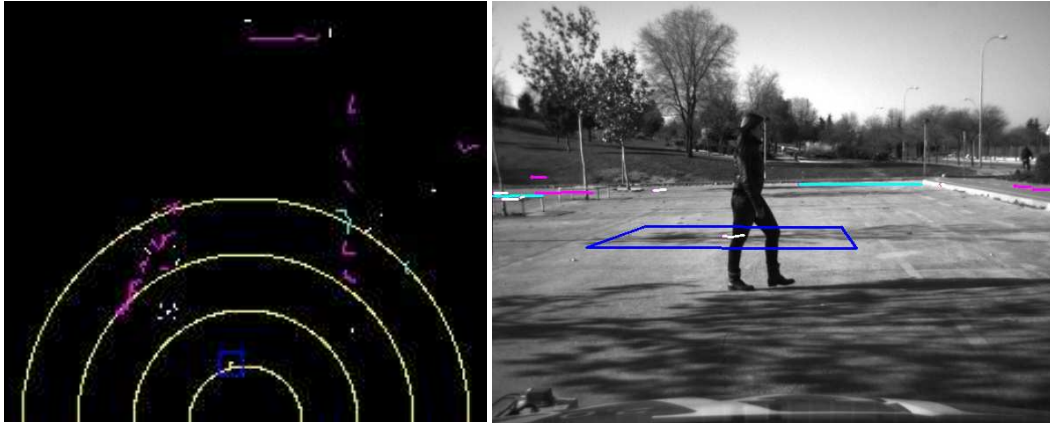


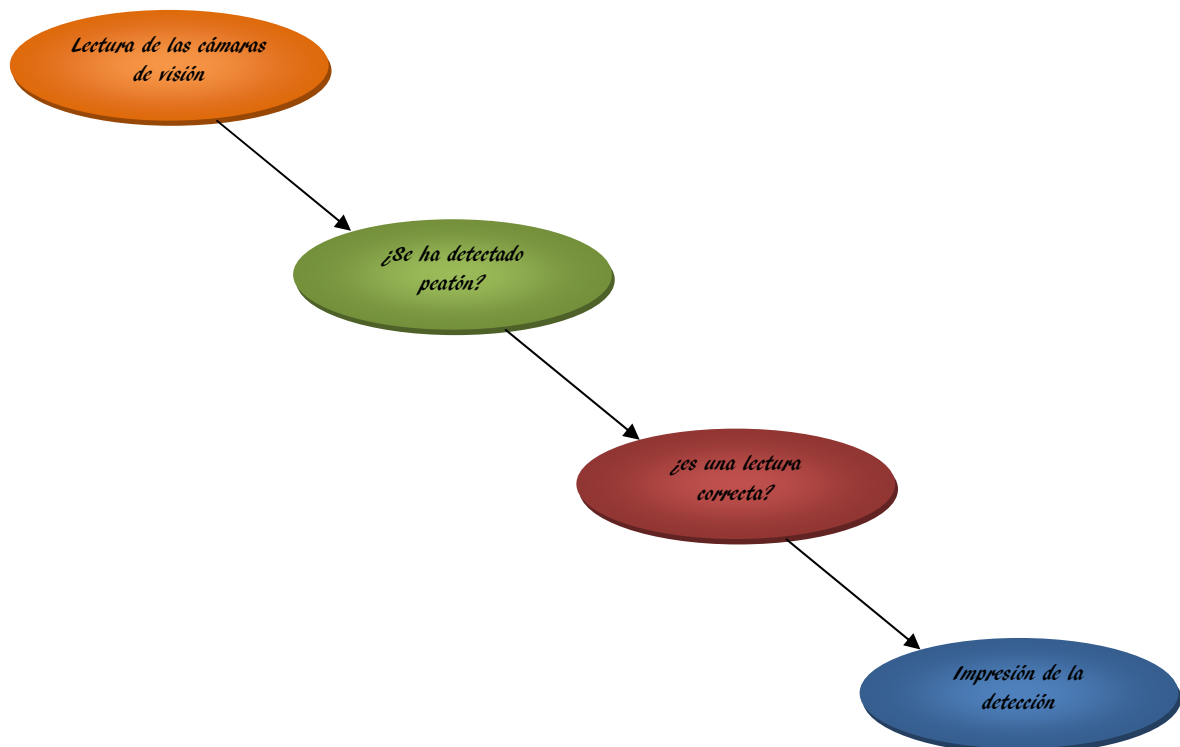
Imagen 21. Detección de láser desde el láser y la cámara.

En estas representaciones podemos observar cómo es una detección del láser en detalle ofrecido por el telémetro láser (recuadro de color azul que se observa en ambas imágenes) de lo que es un peatón y siendo correcta la detección. Es tan importante que haga la detección como que sea realmente cierta.

5.2.2 Detección Cámara.

Las detecciones de la cámara al igual que las del telémetro láser se identifican primero en las lecturas por un número, en este caso son las que comienzan con un “1” tal y como se ha visto en el apartado anterior.

El ejemplo de las detecciones del láser nos sirve perfectamente para describir también la detección de la cámara:



Con este bucle constante lo que se muestra es la detección continua del sistema de cámaras láser o lecturas distintas de las tomadas entre el sistema del láser y el de visión. En concreto estamos hablando de una detección simple nuevamente.

El anterior y este son los dos tipos de detecciones que nos vamos a encontrar en el sistema, el resto de detecciones válidas serán aquellas que llamaremos redundantes, y serán las que hacen a nuestro sistema robusto.

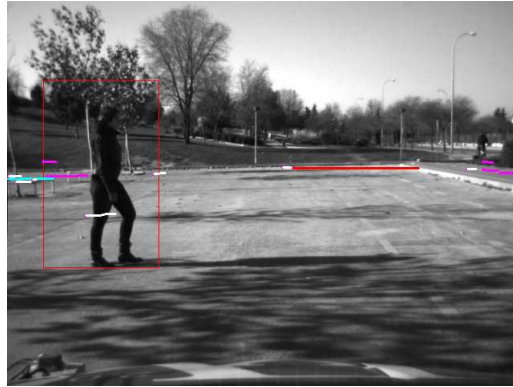
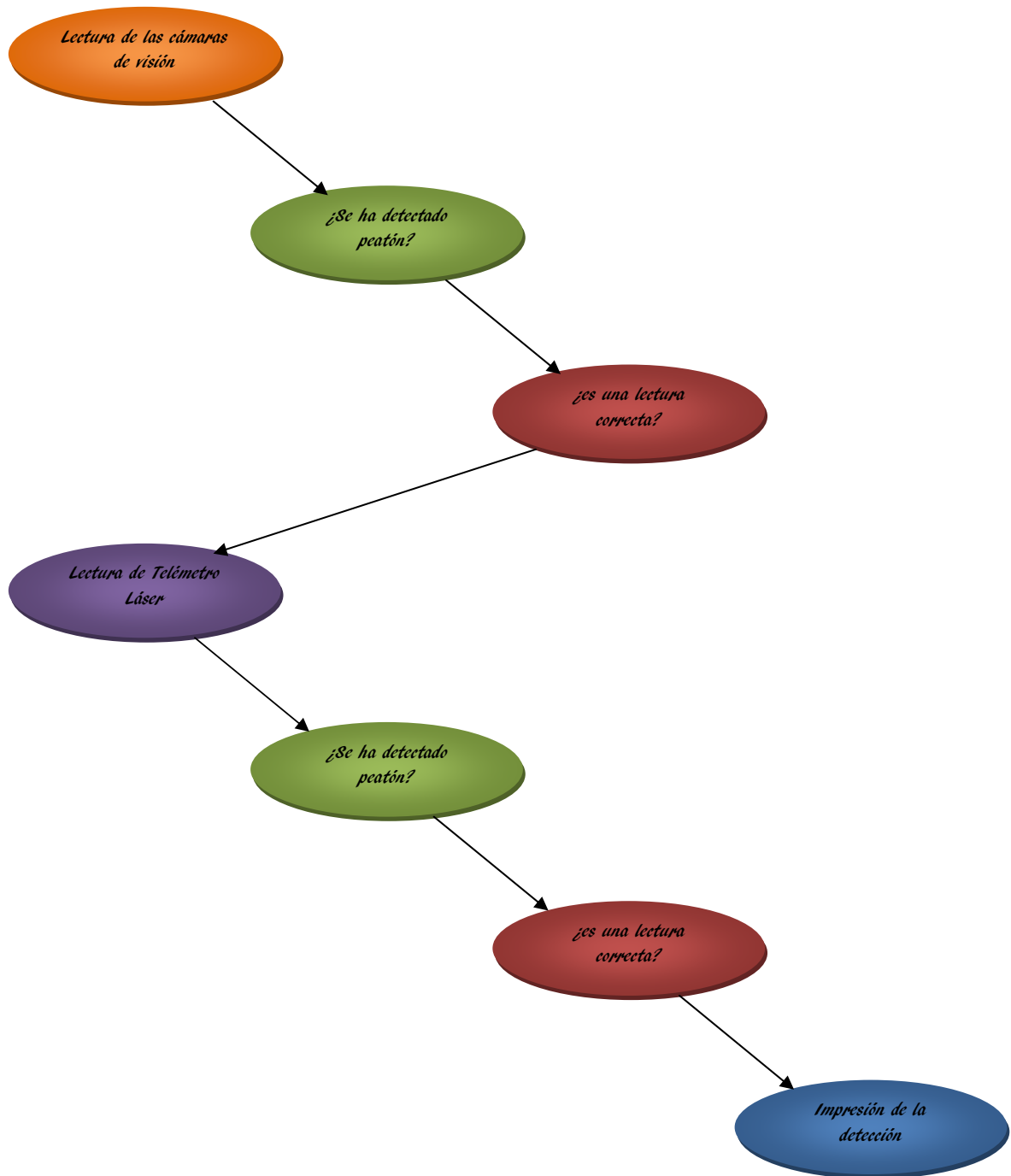


Imagen 22. El recuadro rojo representa a la detección de la cámara.

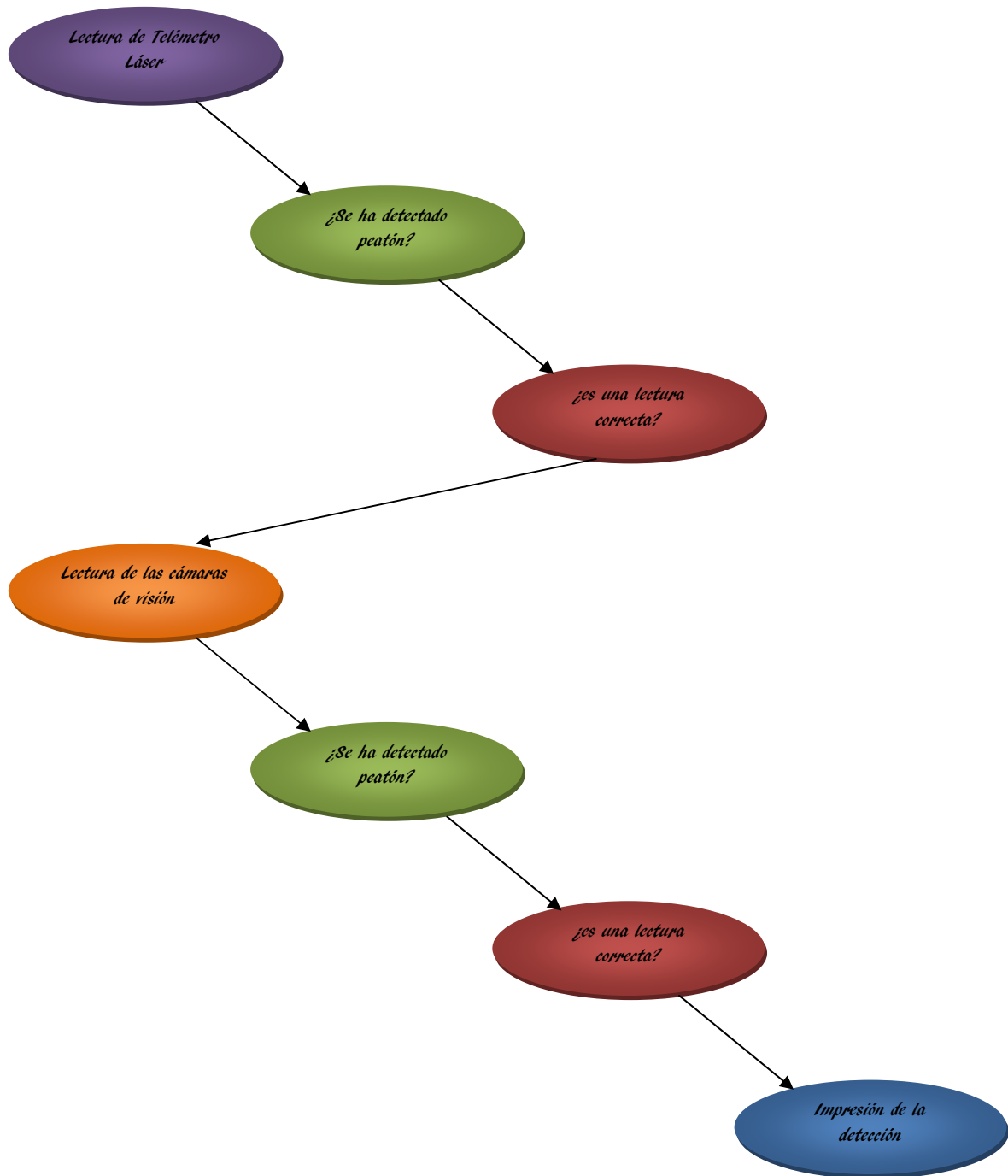
Visualmente una detección del láser es completamente distinta a la del sistema de cámaras, como se puede observar en la anterior imagen, la detección de la cámara la tenemos representada por ese recuadro rojo que se aprecia perfectamente.

5.2.3 Fusión de detección.

Vamos a llamar redundancia de detección a aquella que va a ser detectada por ambos sistemas, es decir; primero será detectada por uno de ellos y después será corroborada como buena, una vez confirmado será detectada casi simultáneamente por el segundo sistema y su confirmación es lo que hará que el sistema sea robusto, porque son dos los sistemas que han confirmado que hay una detección y que se ha hecho en el mismo punto.



Para realizar la redundancia hemos supuesto dos posibles casos, uno en el que primero se haga la lectura con el sistema de cámaras de visión y seguido sea con el telémetro láser, y un segundo que será al revés, primero el telémetro láser y seguido el sistema de cámaras.



Tal y como se puede observar es una detección segura, lo que se acaba de mostrar. Para que sea así se deben dar varios condicionantes, talque sean consecutivas, los puntos sean exactos, haya peatón, etc... Pero son los únicos casos en los que se detecta la doble detección.

5.3 Cálculos realizados.

En este siguiente apartado van a ser explicado los cálculos realizados y luego implementados en el código Matlab de la investigación.

El primer caso que debemos hacer mención es la lectura de tiempos. Todos los registros de datos salen tomados de la misma forma, esto nos ayuda a hacer una comparación fácil.

19:6:17.329

Se trata de identificar la hora, el minuto y el dato más importante que es el segundo con las milésimas identificadas, ya que la toma de datos se realiza muy rápida. Estos tiempos nos sirven primero para comparar entre si son lecturas o no distintas, ya que entre una lectura y otra pueden pasar milésimas. También nos ayudan a saber la diferencia de tiempo que hay entre una lectura (S(1)) y otra (S(2)).

$$S(3)=S(1)-S(2)$$

$$0,022=17,351-17,329$$

Ecuación 36. Fórmula de la investigación 1

El tiempo también es importante para realizar el cálculo de la velocidad, aquí además de la diferencia de tiempos también es necesario la diferencia de distancias entre un punto (X(1)) y otro (X(2)). Habiendo una única excepción; en el caso de que los dos puntos tengan la misma posición, podría darse el caso o no de la doble detección y además la velocidad sería 0.

$$X(3)=X(1)-X(2) \quad 0,09=0,56-0,47$$

Ecuación 37. Fórmula de la investigación 2

$$Vel=X(3)/S(3) \quad 4,1=0,09/0,022$$

Ecuación 38. Fórmula de la investigación 3

Toda lectura como hemos dicho antes debe pasar una serie de condiciones, otra condición son los límites máximos de una zona de interés. Nuestra zona de interés se comprende entre 2 unidades máximo por cada uno de sus ejes tomando como referencia el punto de detección anterior.

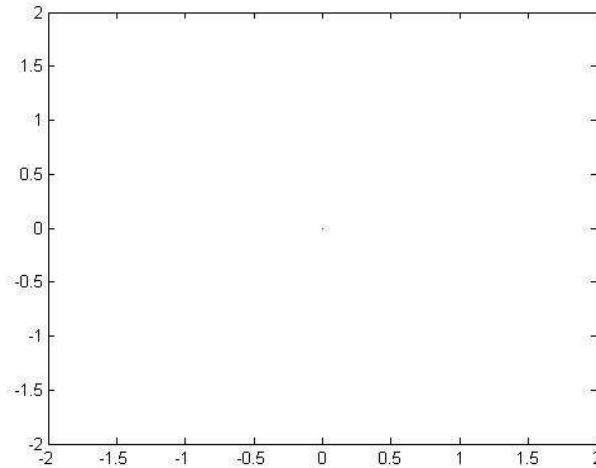


Imagen 23

Una vez realizados una serie de cálculos mínimos nos interesa en especial hacer una estimación de los puntos o lecturas en las que podremos obtener la siguiente lectura del peatón, pero además saber cuál es la probabilidad de que ese punto sea correcto. Esto es lo que venimos comentando de una estimación a posteriori, puesto que la probabilidad a priori se realiza cuando estamos detectando peatón y sabemos que lo es.

Primero hay que obtener la posición del punto en el que estimamos que se podrá obtener la siguiente detección de peatón. Ésta se realiza en base al un nuevo tiempo de detección a la velocidad calculada en el punto anterior.

$$X_{\text{estimado}} = S(3) * \text{Vel} \quad 0,102 = 0,025 * 4;1$$

Ecuación 39. Fórmula de la investigación 4

$$X_{\text{nueva}} = X(1) + X_{\text{estimado}} \quad 0,662 = 0,56 + 0,102$$

Ecuación 40. Fórmula de la investigación 5

Este punto estimado tiene una probabilidad de ser acertado en el punto exacto o no. Para ello lo calcularemos a través de una Gaussiana que nos estime con qué probabilidad es este punto correcto.

$$R = \text{gaussmf}(X(1), [0.2 \ X_{\text{nueva}}]);$$

Ecuación 41. Fórmula de la investigación 6

$$G = P_{\text{laser}} * R; \quad G = P_{\text{cam}} * R; \quad G = P_{\text{las_cam}} * R;$$

Ecuación 42. Fórmula de la investigación 7

$$P(1) = (P(2) * G) / P_B;$$

Ecuación 43. Fórmula de la investigación 8

Gaussmf es una función ya implementada que incorpora Matlab, pero para calcular la probabilidad estimada ($P(1)$) primero debemos de tener en cuenta si lo calculamos en base al telémetro láser (P_{laser}), las cámaras de visión (P_{cam}) o la doble detección de láser y cámara (P_{las_cam}). Después debemos saber que es una probabilidad condicionada a lo calculado en anteriores detecciones tales como sus posiciones o incluso las probabilidades ($P(2)$)

La Gaussiana es una una curva a través de la cual estimamos la probabilidad, pero tenemos que tener en cuenta que cada uno de los aparatos con los que tomamos lecturas tiene una probabilidad de acierto diferente, y además que estamos haciendo una estimación a posteriori; es decir, una estimación de un punto en una lectura futura, por lo que hay que tener muy en cuenta todos los datos tomados anteriormente.

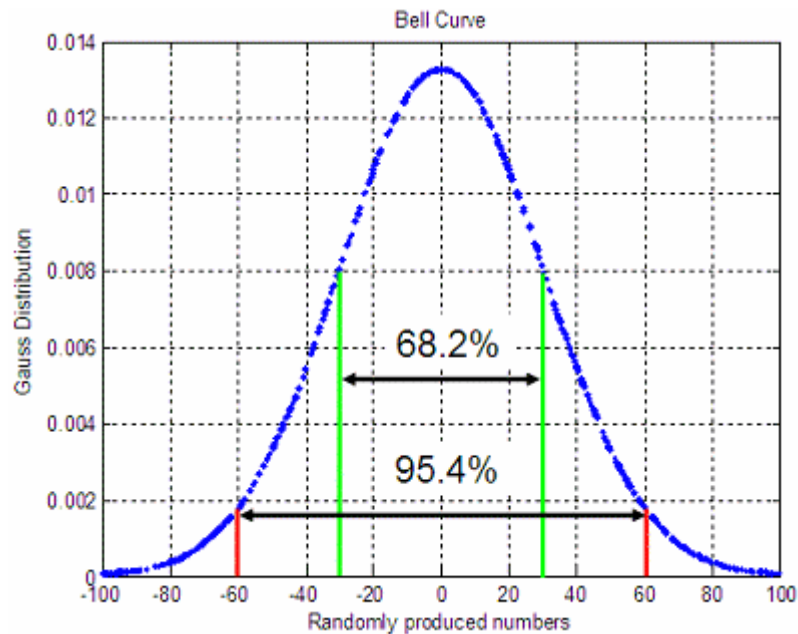


Imagen 24. Curva de la campana de Gauss.

Toda esta serie de cálculos se realiza repetidas veces y en función siempre de unos valores preestablecidos inicialmente, como por ejemplo las posiciones y tiempos iniciales o velocidades o incluso las probabilidades a priori y a posteriori. Siempre se obtienen al principio unos valores desajustados, y hasta que no se toman unas 3 lecturas correctas no se producen valores significativos para tenerlos en cuenta.

Todos los valores iniciales están en un principio en 0 excepto los previamente calculados, como lo son la probabilidad de acierto de la cámara ($P_{cam}=0,751$), la probabilidad de acierto del láser ($P_{laser}=0,886$), y la probabilidad conjunta de acierto del láser y la cámara ($P_{las_cam}=0,676$) y otros valores umbrales como Sigma ($\sigma=0,2$), la probabilidad condicionada ($PB=0,9$).

5.4 Impresión.

Durante toda la investigación se está recalcando la distinción entre las detecciones. Pues bien, durante la impresión, también vamos a distinguirlas además de ver la trayectoria de los peatones.

Para imprimir todos y cada uno de los punto de interés, se ha utilizado la función de Matlab 'Hold on' y 'Hold off'. Con 'Hold on' lo que hacemos es guardar todo lo que queramos imprimir, y si lo utilizamos lo que hacemos es imprimir la secuencia de detecciones anterior y la que estemos usando en ese momento. Y para cerrar la impresión se utiliza la función 'Hold off' que deja de guardar todo.

Impresión antes de 'Hold on'. Como se puede observar en la secuencia de imágenes, según vamos obteniendo detecciones vamos guardando cada una de ellas.

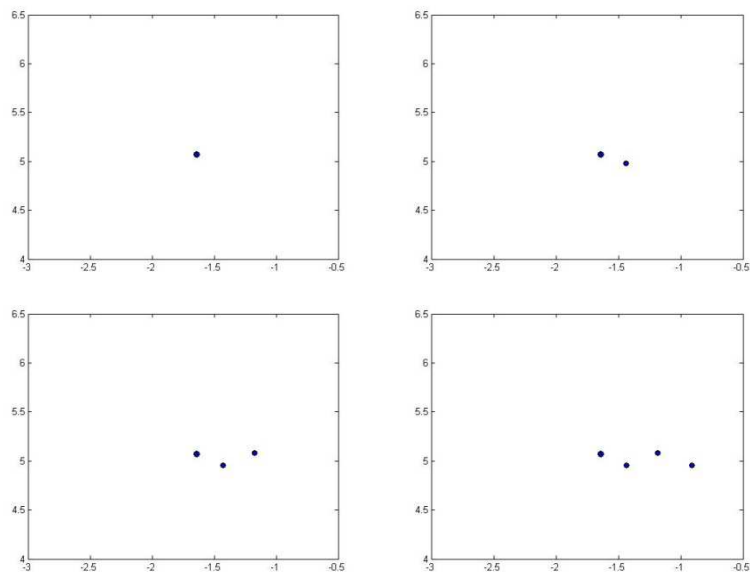


Imagen 25. Secuencia del seguimiento de la detección.

Impresión después de 'Hold on' con las detecciones guardadas anteriormente y antes de 'Hold off'

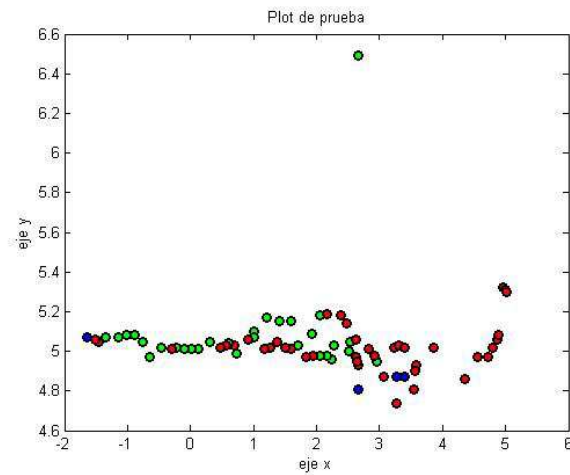


Imagen 26. Detección completa.

Cierre de guardar detecciones con es 'Hold off'.

Tal y como se puede observar en esta última imagen vamos a obtener tres tipos de color distinto en la detecciones, ya sea para la cámara (color verde), caso del Telémetro láser (color rojo); o el caso de las dobles detecciones Láser-cámara (color amarillo) y Cámara-láser (color azul).

6. Resumen general del sistema.

Este es un apartado en el que se va a hacer una explicación periférica del sistema planteado para la fusión. Se pretende que sea más fácil la comprensión del mismo sin entrar en tanto detalle y con la finalidad de clarificar todo lo anteriormente explicado.

La parte más interesante del proyecto es el sistema ideado para la fusión. Para considerar dicha fusión siempre se deben realizar dos lecturas consecutivas y además que cada una sea de un aparato distinto, ya sea el sistema del láser o el de la cámara. El sistema es un bucle cerrado constante en el que se está constantemente leyendo hasta el punto en el que se acaban las lecturas.

Todas las lecturas están mezcladas entre sí y lo primero es detectar qué sistema es el que nos ofrece la primera. Se pasa a comprobar que se detecte peatón y sólo en este caso es cuando realizamos la serie de cálculos anteriormente explicados. Ahora nos encontramos en el punto en el que tenemos un peatón detectado y confirmado.

Lo siguiente es realizar una segunda lectura, que al igual que la primera se pasa por todos los pasos hasta que se detecta peatón y se realizan todo lo cálculos. Es en esta parte en la que tenemos que tomar la decisión de si es una 'simple' o 'doble' detección, debido a que en cada caso habrá una forma distinta de actuar.

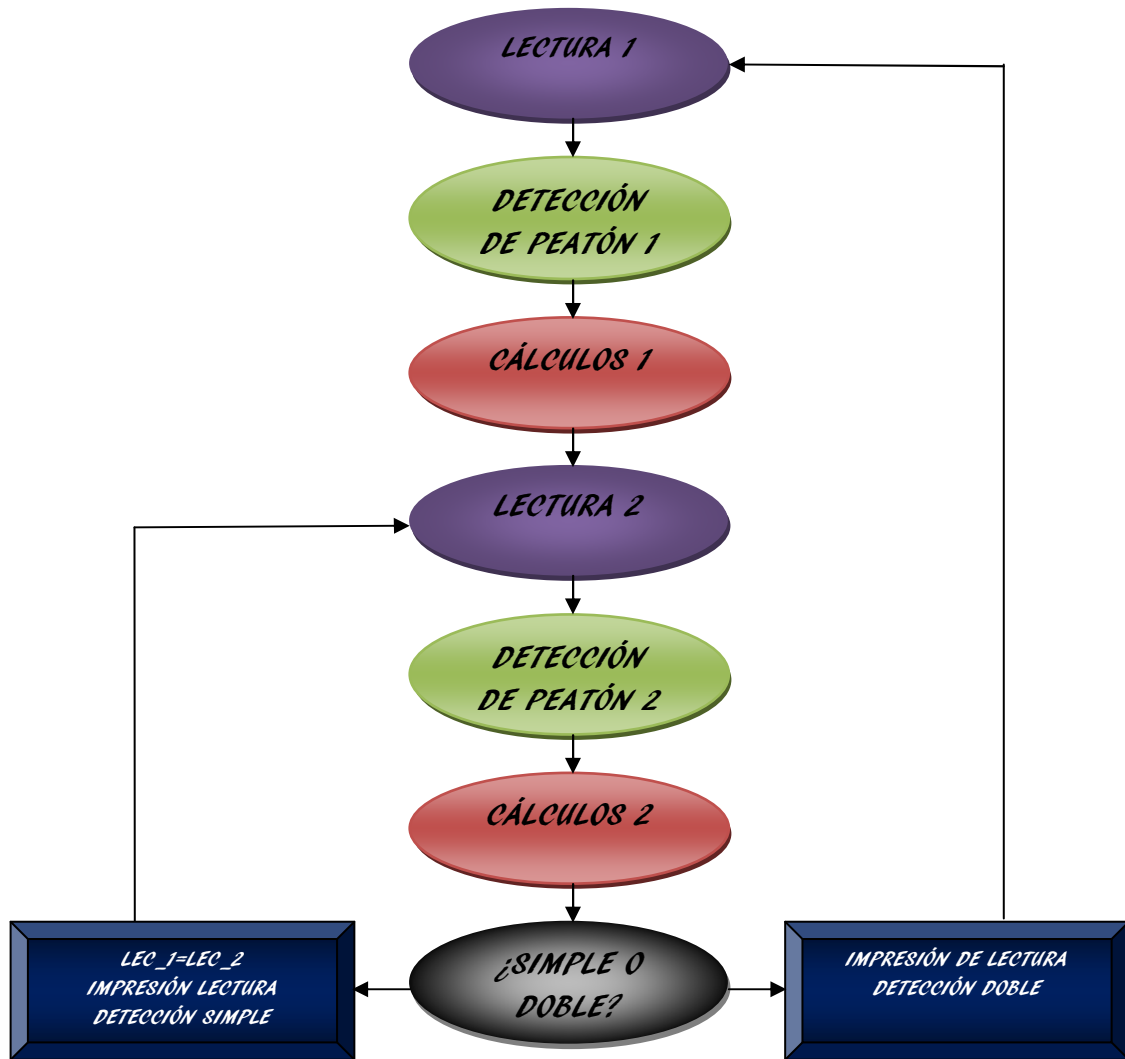
Para que sea una detección simple o doble se ha tomado la decisión durante los cálculos en que se distingue entre el tiempo y la posición del peatón. En el caso de que no coincidan se considera que es una "Detección simple" y en el caso de que sí coincidan sería "Detección doble".

La "Detección simple" se puede dar en cuatro simples casos:

- Caso 1: Sólo se están tomando lecturas de la cámara, es decir, todas las detecciones son sólo de la cámara.
- Caso 2: Sólo se están tomando lecturas del láser, es decir, todas las detecciones son sólo del telémetro láser.
- Caso 3: Se han tomado una lectura del láser primero y otra de la cámara después además de seguidas, pero son distintas en tiempo o posición.
- Caso 4: Se han tomado una lectura de la cámara primero y del láser después además de seguidas, pero son distintas en tiempo o posición.

La "Detección doble" (fusión) se puede dar sólo en dos casos:

- Caso 1: Se han tomado una lectura del láser primero y otra de la cámara después además de seguidas, pero son iguales en tiempo y posición.
- Caso 2: Se han tomado una lectura de la cámara primero y del láser después además de seguidas, pero son iguales en tiempo y posición.



Por último a la hora de imprimir los resultados y dibujar en las gráficas es muy importante la distinción del tipo de detección. Con la detección “Simple” se imprimirán y dibujarán los datos de la primera lectura, además después los valores significativos de los cálculos de la segunda lectura pasan a ser los de la primera y lo que se realiza en una nueva segunda lectura. Sin embargo con la detección “Doble” imprimimos y dibujamos los datos de la segunda directamente, ya que al ser iguales en posición y tiempo no quiere decir que sean iguales los cálculos. Y en esta ocasión empezariamos de nuevo la primera lectura.

7. Análisis de pruebas.

Toda investigación debe tener una serie de pruebas o procesos experimentales los cuales han de planificar, diseñar, controlar, normalizar y estandarizar todo test que se quiera realizar en sus diferentes ámbitos de estudio. Hemos planificado crear esta fusión de sensores para la seguridad, además se a diseñado un sistema para controlar todas las detecciones y sus seguimientos que normalizaremos, para finalmente estandarizar, o mejor dicho, intentar estandarizar los resultados.

7.1 Caso 1.

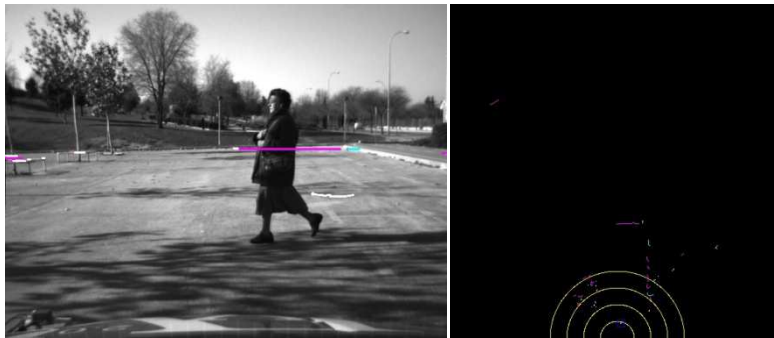


Imagen 27. Detección caso 1 a 5m.

El caso primero se trata de un sujeto (tal y como se ve en la imagen es una señora) el cual se tomarán datos a una distancia de 5 metros. También en este mismo caso se van a tomar más datos con un sujeto (misma señora) a 10 metros de distancia.

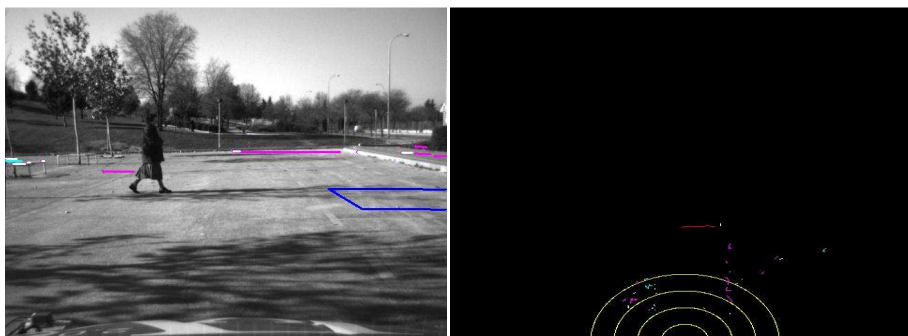


Imagen 28. Detección caso1 a 10m.

7.1.1 Resultados caso 1 a 5m.

Tabla 3

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|-------|-------|-------|--------|--------|-------|---------|---------|--------|-----|
| 1 | -1.1 | -1.1 | 0 | 48.886 | 48.839 | 0.047 | 0 | -1.1045 | 0.8342 | D |
| 0 | -1.29 | -1.1 | -0.19 | 52.396 | 48.886 | 3.51 | -0.0541 | -1.1 | 0.523 | S |
| 1 | -1.29 | -1.29 | 0 | 52.489 | 52.396 | 0.093 | 0 | -1.295 | 0.9342 | S |
| 0 | -1.36 | -1.29 | -0.07 | 54.174 | 52.489 | 1.685 | -0.0415 | -1.29 | 0.7724 | S |
| 0 | -1.56 | -1.36 | -0.2 | 55.672 | 54.174 | 1.498 | -0.1335 | -1.4222 | 0.5998 | S |
| 0 | -1.56 | -1.56 | 0 | 56.608 | 55.719 | 0.889 | 0 | -1.56 | 0.6265 | D |

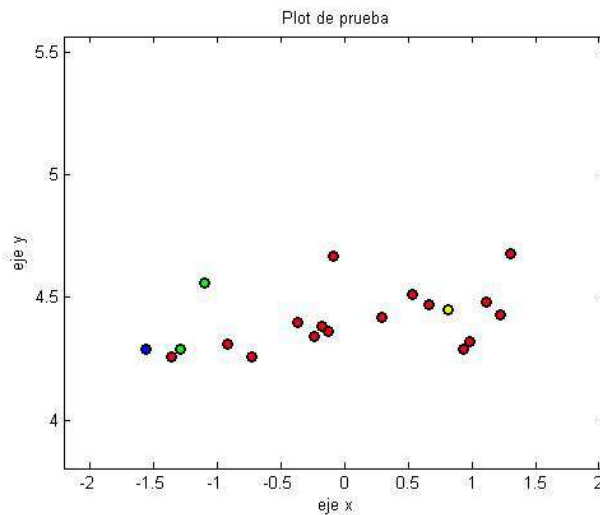


Imagen 29. Resultados caso 1 a 5m

7.1.2 Resultados caso1 a 10m.

Tabla 4

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|------|------|-------|--------|-------|-------|--------|--------|--------|-----|
| 1 | 1.08 | 1.08 | 0 | 4.78 | 4.47 | 0.31 | 0 | 1.1476 | 0.7881 | D |
| 0 | 1.17 | 1.08 | 0.09 | 5.4040 | 4.78 | 0.624 | 0.1442 | 1.08 | 0.7012 | S |
| 0 | 1.38 | 1.17 | 0.21 | 6.59 | 5.404 | 1.186 | 0.1771 | 1.3411 | 0.6773 | S |
| 1 | 1.38 | 1.38 | 0 | 6.652 | 6.59 | 0.062 | 0 | 1.391 | 0.8332 | D |
| 0 | 1.54 | 1.38 | 0.016 | 7.635 | 6.652 | 0.983 | 0.1628 | 1.38 | 0.5956 | S |
| 0 | 1.6 | 1.54 | 0.06 | 8.665 | 7.635 | 1.03 | 0.0583 | 1.7077 | 0.5073 | S |

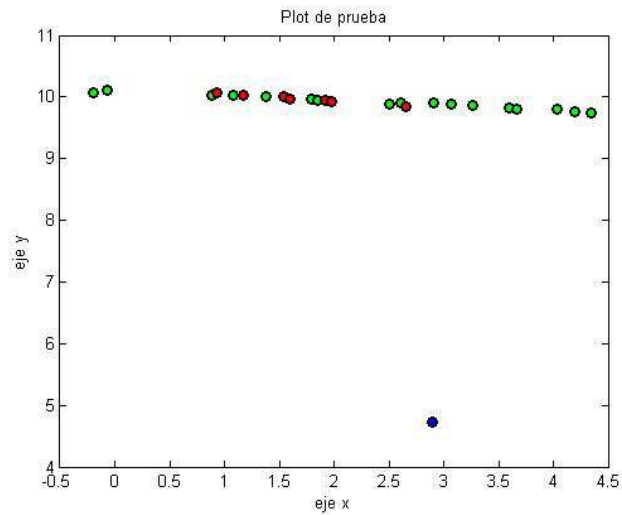


Imagen 30. Resultados caso 1 a 10m.

7.2 Caso 2.

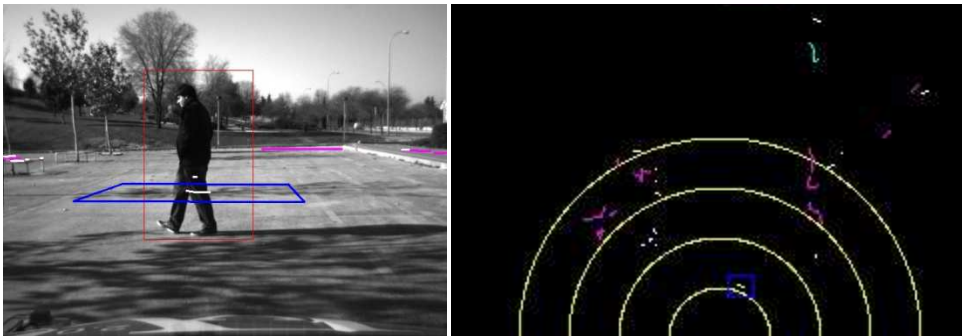


Imagen 31. Detección caso 2 a 5m.

En el segundo caso se tomarán datos con otro sujeto distinto (en este caso es un hombre), pero las distancias serán las mismas a 5 y a 10 metros de distancia.

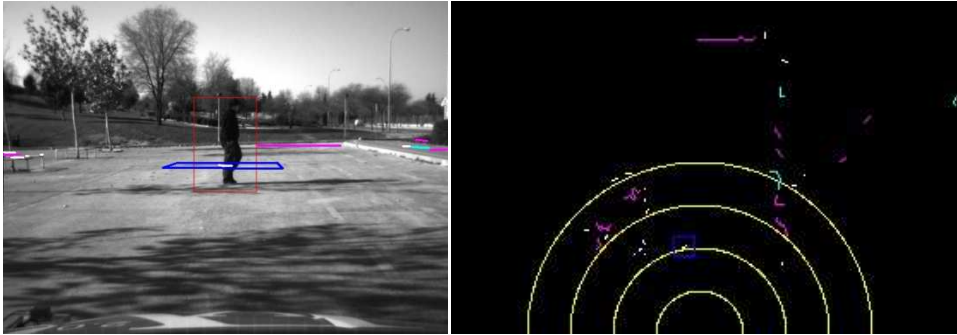


Imagen 32. Detección caso 2 a 10m.

7.2.1 Resultados caso 2 a 5m.

Tabla 5

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|------|------|-------|-------|-------|-------|--------|--------|--------|-----|
| 1 | 1.92 | 1.59 | 0.33 | 5.423 | 3.426 | 1.997 | 0.1652 | 1.934 | 0.8324 | S |
| 0 | 2.06 | 1.92 | 0.14 | 6.484 | 5.423 | 1.061 | 0.1320 | 2.0953 | 0.8068 | S |
| 1 | 2.06 | 2.06 | 0 | 6.53 | 6.484 | 0.046 | 0 | 2.0661 | 0.8341 | D |
| 0 | 2.17 | 2.06 | 0.11 | 7.357 | 6.53 | 0.827 | 0.133 | 2.06 | 0.7058 | S |
| 0 | 2.38 | 2.17 | 0.21 | 7.466 | 7.357 | 0.109 | 1.9266 | 2.1845 | 0.4309 | S |
| 1 | 2.66 | 2.38 | 0.28 | 7.513 | 7.466 | 0.047 | 5.9574 | 2.4706 | 0.5328 | S |

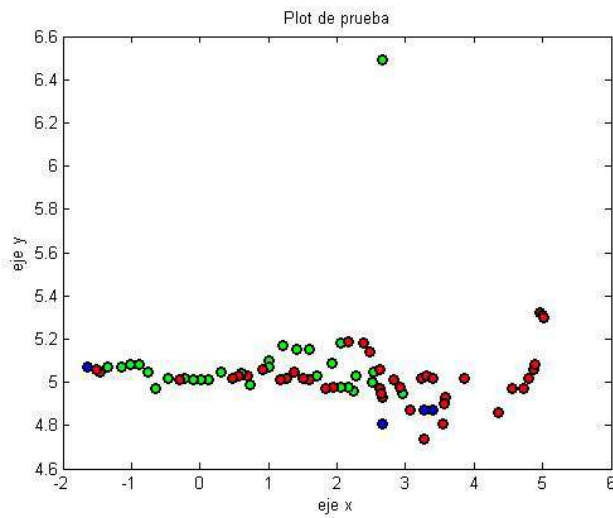


Imagen 33. Resultados caso 2 a 5m.

7.2.2 Resultados caso 2 a 10m.

Tabla 6

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|------|------|-------|--------|--------|-------|--------|--------|--------|-----|
| 1 | 0.27 | 0.27 | 0 | 49.695 | 49.664 | 0.031 | 0 | 0.2757 | 0.8341 | D |
| 0 | 0.4 | 0.27 | 0.14 | 50.678 | 49.695 | 0.983 | 0.1322 | 0.27 | 0.6648 | S |
| 0 | 0.54 | 0.4 | 0.14 | 51.848 | 50.678 | 1.17 | 0.1197 | 0.5547 | 0.6527 | S |
| 1 | 0.54 | 0.54 | 0 | 51.879 | 51.848 | 0.031 | 0 | 0.5437 | 0.8343 | D |
| 0 | 0.67 | 0.54 | 0.13 | 52.831 | 51.879 | 0.952 | 0.1366 | 0.54 | 0.6649 | S |
| 0 | 0.87 | 0.67 | 0.2 | 53.735 | 52.831 | 0.904 | 0.2212 | 0.7934 | 0.6083 | S |

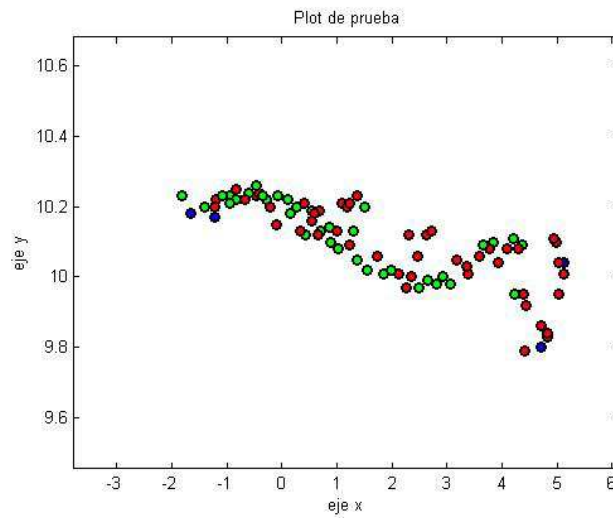


Imagen 34. Resultados caso 2 a 10m.

7.3 Caso 3.

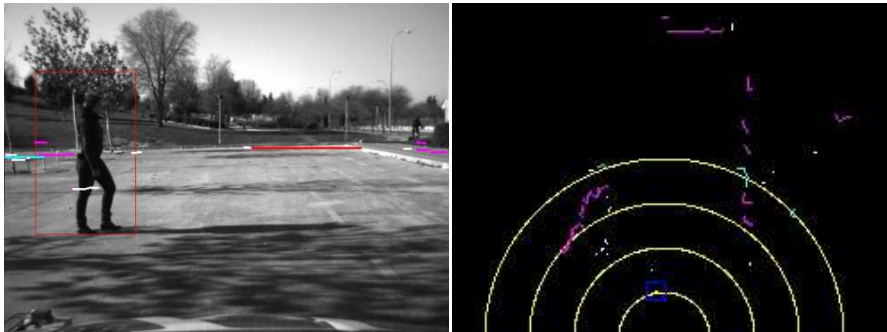


Imagen 35. Detección caso 3 a 5m.

En el caso tercero, al igual que en los otros 2 casos anteriores, se van a tomar lecturas a 5 y a 10 metros de distancia (el nuevo sujeto es ahora una mujer).

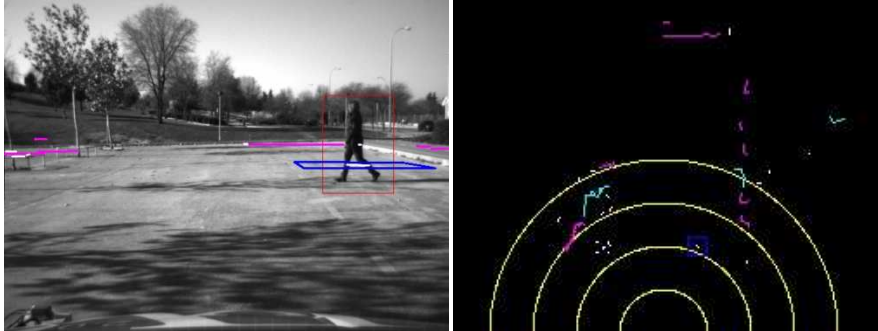


Imagen 36. Detección caso 3 a 10m.

7.3.1 Resultados caso 3 a 5m.

Tabla 7

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|------|------|-------|--------|--------|-------|--------|--------|--------|-----|
| 0 | 0.4 | 0.27 | 0.14 | 50.678 | 49.695 | 0.983 | 0.1322 | 0.27 | 0.6648 | S |
| 0 | 0.54 | 0.4 | 0.14 | 51.848 | 50.678 | 1.17 | 0.1197 | 0.5547 | 0.6527 | S |
| 1 | 0.54 | 0.54 | 0 | 51.879 | 51.848 | 0.031 | 0 | 0.5437 | 0.8343 | D |
| 0 | 0.67 | 0.54 | 0.13 | 52.831 | 51.879 | 0.952 | 0.1366 | 0.54 | 0.6649 | S |
| 0 | 0.87 | 0.67 | 0.2 | 53.735 | 52.831 | 0.904 | 0.2212 | 0.7934 | 0.6083 | S |
| 1 | 0.87 | 0.87 | 0 | 53.767 | 53.735 | 0.032 | 0 | 0.8771 | 0.8339 | D |

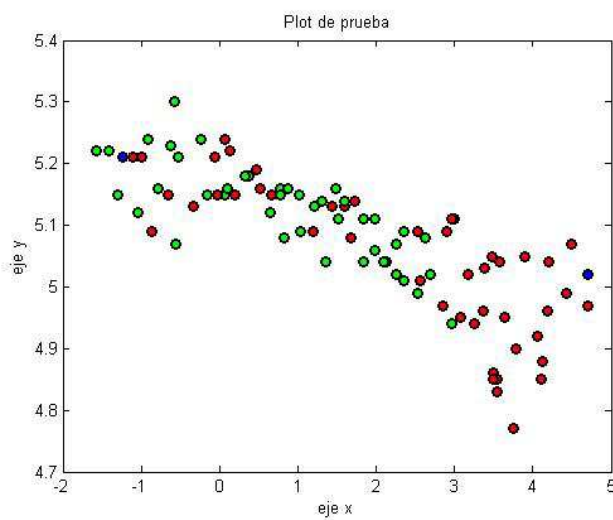


Imagen 37. Resultados caso 3 a 5m.

7.3.2 Resultados caso 3 a 10m.

Tabla 8

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|------|------|-------|--------|--------|-------|--------|--------|--------|-----|
| 1 | 0.46 | 0.46 | 0 | 34.908 | 34.868 | 0.04 | 0 | 0.4671 | 0.8339 | D |
| 0 | 0.64 | 0.46 | 0.18 | 35.988 | 34.908 | 1.08 | 0.1667 | 0.46 | 0.5476 | S |
| 1 | 0.64 | 0.64 | 0 | 36.18 | 35.988 | 0.192 | 0 | 0.672 | 0.8238 | D |
| 0 | 0.77 | 0.64 | 0.13 | 37.306 | 36.18 | 1.126 | 0.1155 | 0.64 | 0.6566 | S |
| 1 | 0.77 | 0.77 | 0 | 37.342 | 37.306 | 0.036 | 0 | 0.7742 | 0.8343 | D |
| 0 | 0.93 | 0.77 | 0.16 | 38.44 | 37.342 | 1.098 | 0.1457 | 0.77 | 0.5964 | S |

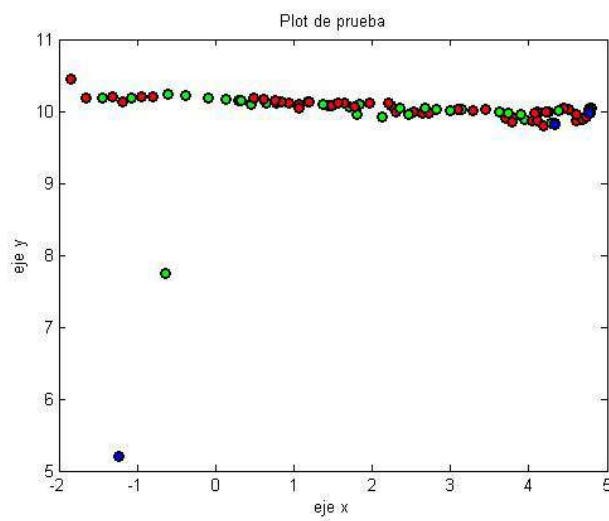


Imagen 38. Resultados caso 3 a 10m.

7.4 Caso 4.

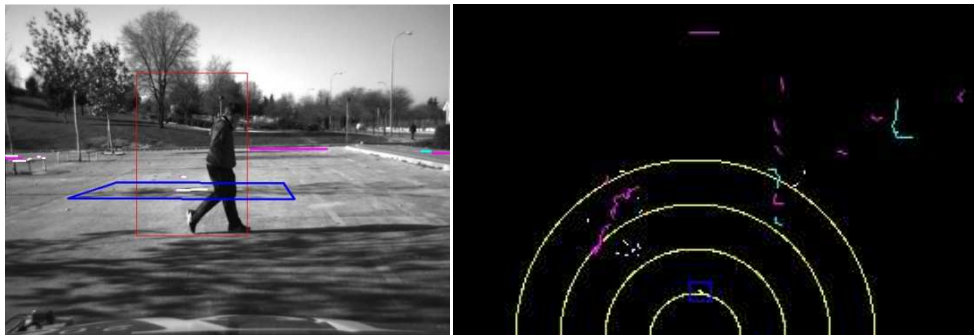


Imagen 39. Detección caso 4 a 5m.

En el caso cuarto vamos a tener al sujeto masculino, en este caso es diferente sujeto; pero se analizará como en todos los caso a una distancia de 5 y de 10 metros.

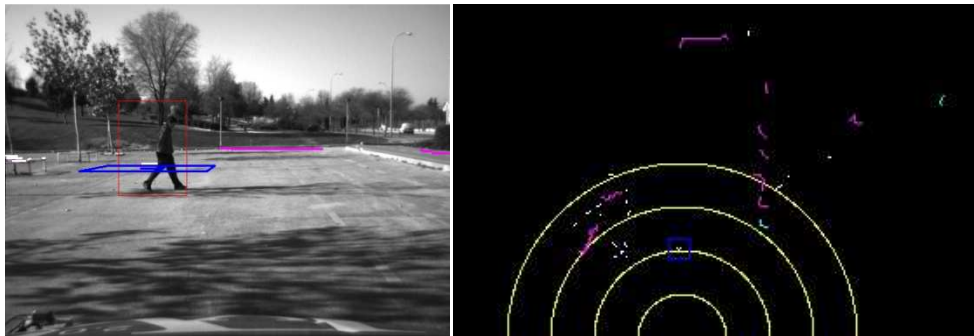


Imagen 40. Detección caso 4 a 10m.

7.4.1 Resultados caso 4 a 5m.

Tabla 9

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|-------|-------|-------|--------|--------|-------|--------|---------|--------|-----|
| 1 | -0.78 | -0.78 | 0 | 38.126 | 38.17 | 0.046 | 0 | -0.7506 | 0.8255 | D |
| 0 | -0.68 | -0.78 | 0.1 | 39.184 | 38.216 | 0.968 | 0.1033 | -0.78 | 0.7172 | S |
| 1 | -0.68 | -0.68 | 0 | 39.215 | 39.184 | 0.031 | 0 | -0.6769 | 0.8343 | D |
| 0 | -0.48 | -0.68 | 0.2 | 40.447 | 39.215 | 1.232 | 0.1623 | -0.68 | 0.495 | S |
| 1 | -0.48 | -0.48 | 0 | 40.478 | 40.447 | 0.031 | 0 | -0.475 | 0.8342 | D |
| 0 | -0.29 | -0.48 | 0.19 | 41.773 | 40.478 | 1.295 | 0.1467 | -0.48 | 0.523 | S |

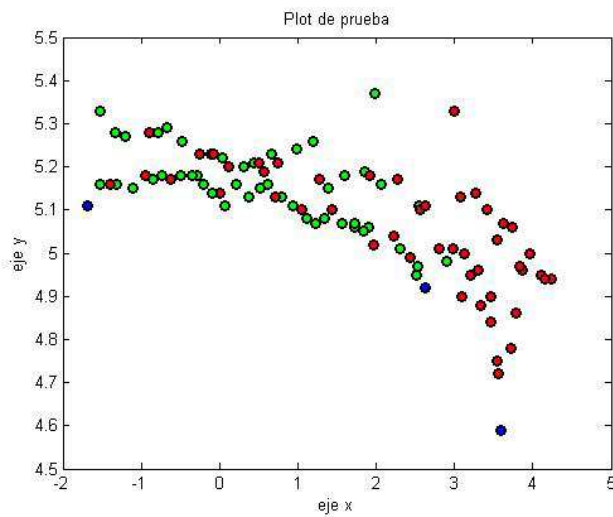


Imagen 41. Resultados caso 4 a 5m.

7.4.2 Resultados caso 4 a 10m.

Tabla 10

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|-------|-------|-------|--------|--------|-------|--------|---------|--------|-----|
| 1 | -0.4 | -0.4 | 0 | 22.94 | 22.909 | 0.031 | 0 | -0.3538 | 0.8125 | D |
| 0 | -0.25 | -0.4 | 0.15 | 23.595 | 22.94 | 0.655 | 0.229 | -0.4 | 0.6038 | S |
| 1 | -0.25 | -0.25 | 0 | 23.642 | 23.595 | 0.047 | 0 | -0.2392 | 0.8332 | D |
| 0 | -0.11 | -0.25 | 0.14 | 24.843 | 23.642 | 1.201 | 0.1166 | -0.25 | 0.642 | S |
| 1 | -0.11 | -0.11 | 0 | 24.874 | 24.843 | 0.031 | 0 | -0.1064 | 0.8343 | D |
| 0 | 0 | -0.11 | 0.11 | 25.545 | 24.874 | 0.671 | 0.1639 | -0.11 | 0.706 | S |

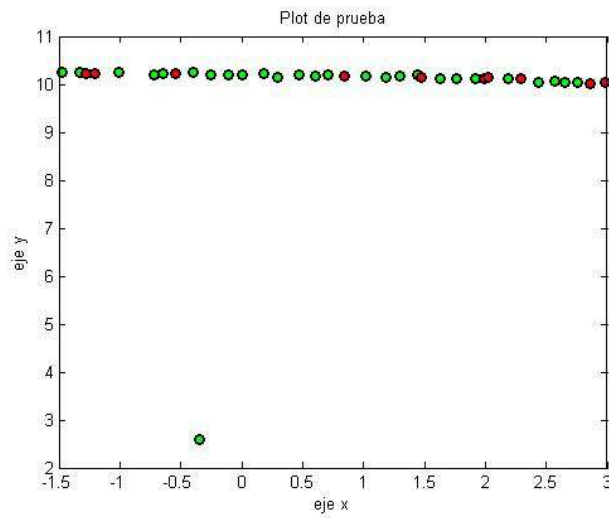


Imagen 42. Resultados caso 4 a 10m.

7.5 Caso 5.

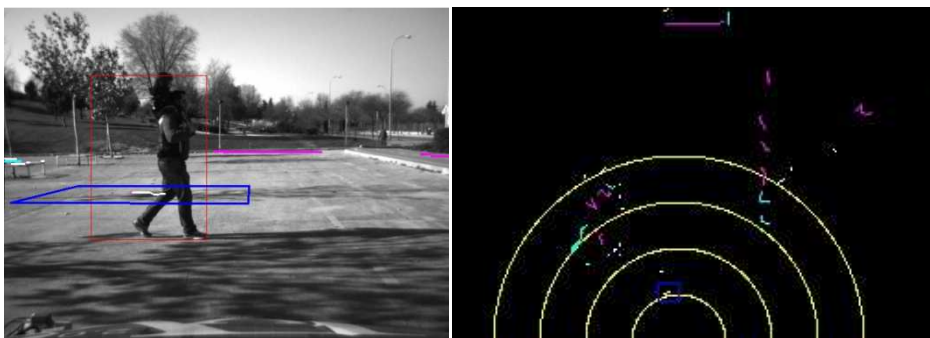


Imagen 43. Detección caso 5 a 5m.

Caso 5. Se ha incluido para este caso, como se puede observar en la imagen, a dos sujetos, pero juntos. En vista de lo que se observa el sistema los considera como si fueran uno solo, pero de mayor tamaño. Serán analizados de nuevo a 5 y a 10 metros de distancia.

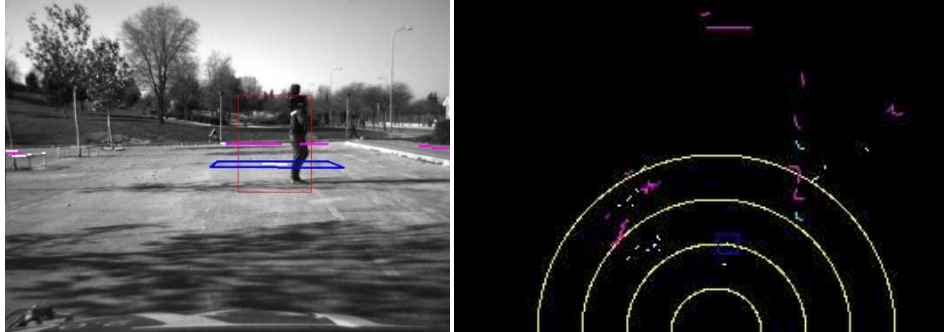


Imagen 44. Detección caso 5 a 10m.

7.5.1 Resultados caso 5 a 5m.

Tabla 11

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|------|------|-------|--------|--------|-------|--------|--------|--------|-----|
| 1 | 0.11 | 0.11 | 0 | 55.195 | 55.164 | 0.031 | 0 | 0.1101 | 0.8344 | D |
| 0 | 0.13 | 0.11 | 0.02 | 56.178 | 55.195 | 0.983 | 0.0203 | 0.11 | 0.8174 | S |
| 0 | 0.25 | 0.13 | 0.12 | 56.443 | 56.178 | 0.265 | 0.4528 | 0.1354 | 0.6828 | S |
| 1 | 0.25 | 0.25 | 0 | 56.49 | 56.443 | 0.047 | 0 | 0.2713 | 0.8297 | D |
| 0 | 0.43 | 0.25 | 0.18 | 57.956 | 56.49 | 1.466 | 0.1228 | 0.25 | 0.5448 | S |
| 1 | 0.43 | 0.43 | 0 | 58.3 | 57.956 | 0.344 | 0 | 0.4722 | 0.816 | D |

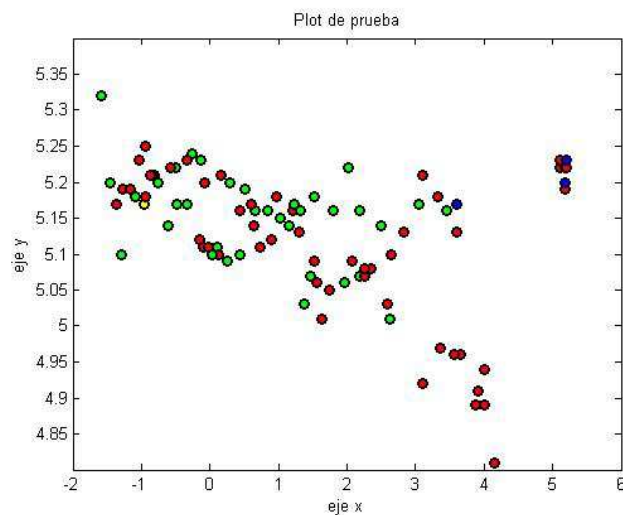


Imagen 45. Resultados caso 5 a 5m.

7.5.2 Resultados caso 5 a 10m.

Tabla 12

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|------|------|-------|--------|--------|-------|--------|--------|--------|-----|
| 1 | 0.24 | 0.13 | 0.11 | 10.95 | 9.81 | 1.14 | 0.0965 | 0.2388 | 0.8344 | S |
| 1 | 0.4 | 0.24 | 0.16 | 11.78 | 10.95 | 0.83 | 0.1928 | 0.3201 | 0.7704 | S |
| 0 | 0.53 | 0.4 | 0.13 | 12.186 | 11.78 | 0.406 | 0.3202 | 0.4783 | 0.7335 | S |
| 0 | 0.62 | 0.53 | 0.09 | 13.168 | 12.186 | 0.982 | 0.0916 | 0.8444 | 0.3847 | S |
| 0 | 0.65 | 0.62 | 0.03 | 13.231 | 13.168 | 0.063 | 0.4762 | 0.6258 | 0.376 | S |
| 0 | 0.81 | 0.65 | 0.16 | 13.309 | 13.231 | 0.078 | 2.0513 | 0.6871 | 0.3065 | S |

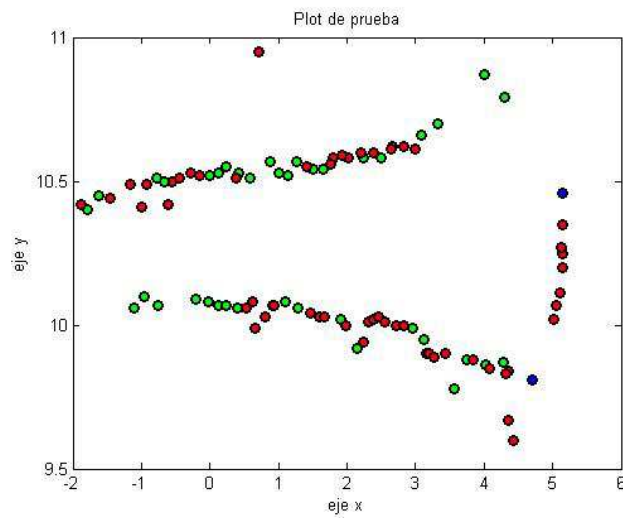


Imagen 46. Resultados caso 5 a 10m.

7.6 Caso 6.

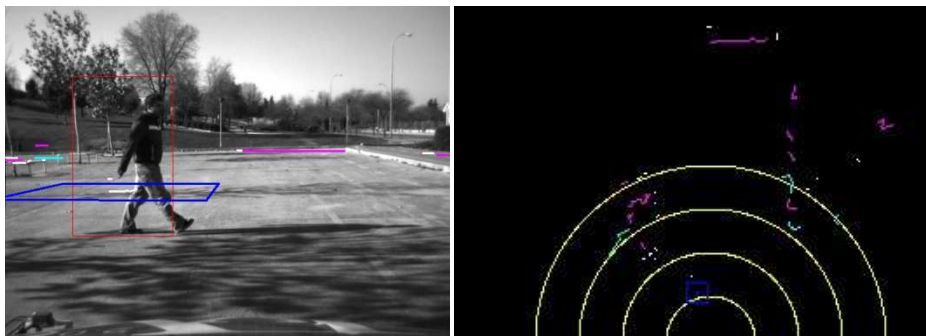


Imagen 47. Detección caso 6 a 5m.

En el caso número 6 vamos a volver a presentar el caso de un hombre, tal y como se ve en la imagen, para la detección y seguimiento a 5 y 10 m.

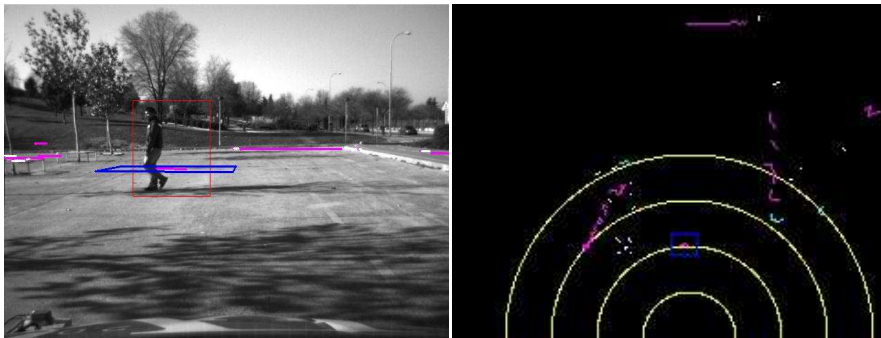


Imagen 48. Detección caso 6 a 10m.

7.6.1 Resultados caso 6 a 5m.

Tabla 13

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|------|------|-------|--------|--------|-------|---------|--------|--------|-----|
| 1 | 2.38 | 2.38 | 0 | 11.704 | 11.666 | 0.038 | 0 | 2.3487 | 0.8243 | D |
| 0 | 2.37 | 2.38 | -0.01 | 13.16 | 11.704 | 1.456 | -0.0069 | 2.38 | 0.8104 | S |
| 0 | 2.14 | 2.37 | -0.23 | 13.358 | 13.16 | 0.198 | -1.1616 | 2.3686 | 0.4151 | S |
| 1 | 2.14 | 2.14 | 0 | 13.395 | 13.358 | 0.037 | 0 | 2.097 | 0.8154 | D |
| 0 | 2 | 2.14 | -0.14 | 14.659 | 13.395 | 1.264 | -0.1108 | 2.14 | 0.6283 | S |
| 1 | 2 | 2 | 0 | 14.722 | 14.659 | 0.063 | 0 | 1.993 | 0.8339 | D |

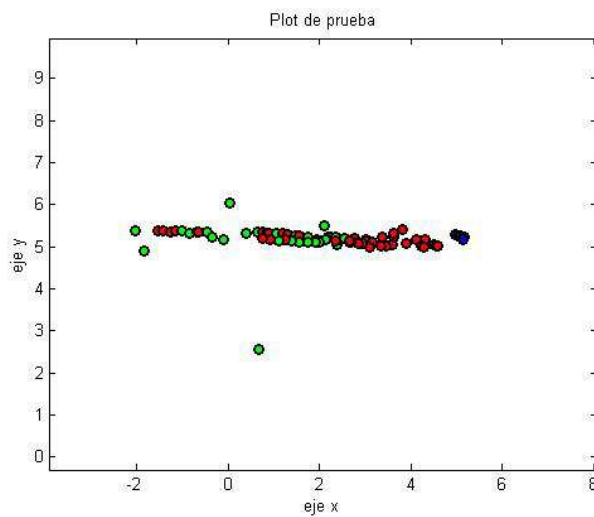


Imagen 49. Resultados caso 6 a 5m.

7.6.2 Resultados caso 6 a 10m.

Tabla 14

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|------|------|-------|--------|--------|-------|---------|--------|--------|-----|
| 1 | 4.11 | 4.11 | 0 | 53.359 | 53.279 | 0.08 | 0 | 4.1142 | 0.8343 | D |
| 0 | 4.3 | 4.11 | 0.19 | 54.239 | 53.359 | 0.88 | 0.2159 | 4.11 | 0.523 | S |
| 0 | 4.4 | 4.3 | 0.1 | 55.162 | 54.239 | 0.923 | 0.1083 | 4.4993 | 0.4552 | S |
| 0 | 4.39 | 4.4 | -0.01 | 55.429 | 55.162 | 0.267 | -0.0375 | 4.4289 | 0.4397 | S |
| 0 | 4.53 | 4.39 | 0.14 | 56.243 | 55.426 | 0.814 | 0.172 | 4.3595 | 0.301 | S |
| 0 | 4.57 | 4.53 | 0.04 | 56.333 | 56.243 | 0.09 | 0.4444 | 4.5455 | 0.2941 | S |

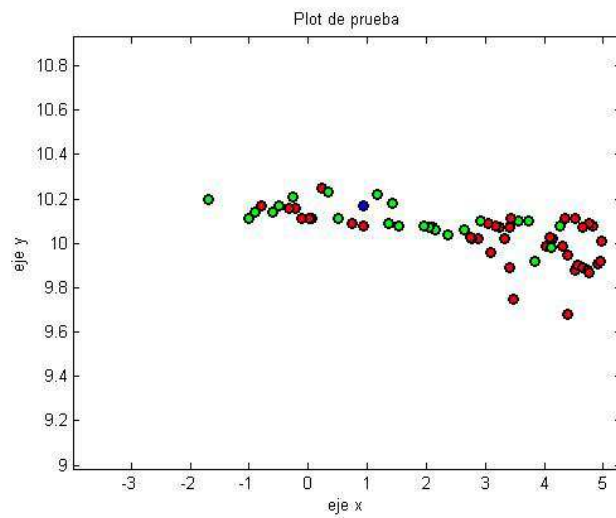


Imagen 50. Resultados caso 6 a 10m.

7.7 Caso 7.

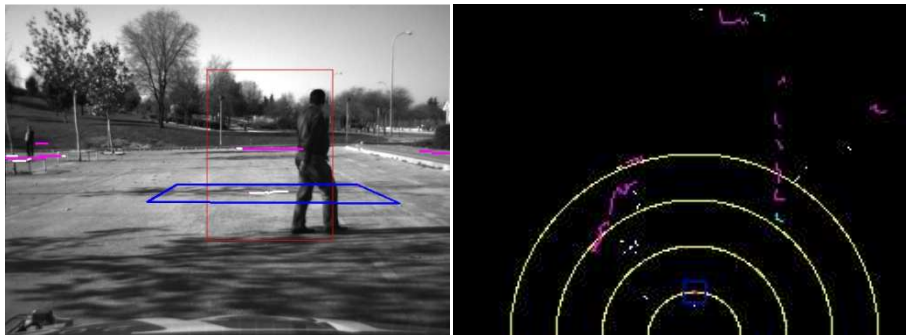


Imagen 51, Detección caso 7 a 5m.

En este séptimo caso vamos a introducir un peatón distinto a los anteriores, para observar como siempre las detecciones y el seguimiento según su movimiento a 5 y 10m. Este peatón es distinto al ser un hombre diferente a los anteriores.

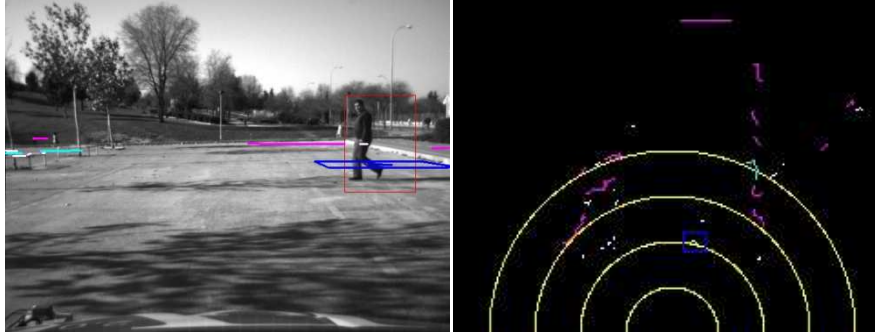


Imagen 52. Detección caso 7 a 10m.

7.7.1 Resultados caso 7 a 5m.

Tabla 15

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|------|------|-------|--------|--------|-------|---------|--------|--------|-----|
| 1 | 2.65 | 2.65 | 0 | 26.296 | 26.256 | 0.04 | 0 | 2.641 | 0.8336 | D |
| 1 | 2.37 | 2.65 | -0.28 | 27.528 | 26.296 | 1.232 | -0.2273 | 2.65 | 0.3132 | S |
| 1 | 2.16 | 2.37 | -0.21 | 28.773 | 27.528 | 1.245 | -0.1687 | 2.087 | 0.7807 | S |
| 1 | 2.05 | 2.16 | -0.11 | 29.798 | 28.773 | 1.025 | -0.1073 | 1.9871 | 0.7942 | S |
| 1 | 1.96 | 2.05 | -0.09 | 30.476 | 29.798 | 0.678 | -0.1327 | 1.9772 | 0.8314 | S |
| 1 | 1.53 | 1.96 | -0.43 | 31.94 | 30.476 | 1.464 | -0.2937 | 1.7657 | 0.4168 | S |

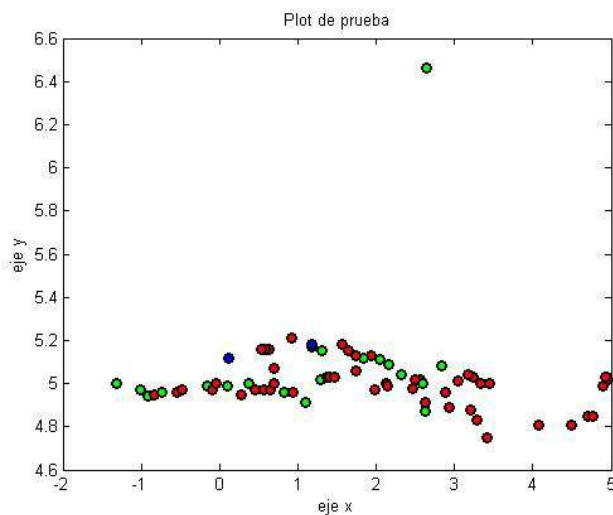


Imagen 53. Resultados caso 7 a 5m.

7.7.2 Resultados caso 7 a 10m.

Tabla 16

| La/Ca | X | X1 | Dis_X | T | T1 | Dis_T | Vel | New_X | Prob | S/D |
|-------|------|------|-------|--------|--------|-------|--------|--------|--------|-----|
| 1 | 1.97 | 1.97 | 0 | 10.463 | 10.432 | 0.031 | 0 | 2.0099 | 0.818 | D |
| 0 | 2.09 | 1.97 | 0.12 | 11.384 | 10.463 | 0.921 | 0.1303 | 1.97 | 0.6727 | S |
| 0 | 2.06 | 2.09 | -0.03 | 11.509 | 11.384 | 0.125 | -0.24 | 2.1063 | 0.6447 | S |
| 0 | 2.28 | 2.06 | 0.22 | 11.649 | 11.509 | 0.14 | 1.5714 | 2.0264 | 0.2841 | S |
| 1 | 2.28 | 2.28 | 0 | 11.805 | 11.649 | 0.156 | 0 | 2.5251 | 0.3937 | D |
| 0 | 2.37 | 2.28 | 0.09 | 12.663 | 11.805 | 0.858 | 0.1049 | 2.28 | 0.3503 | S |

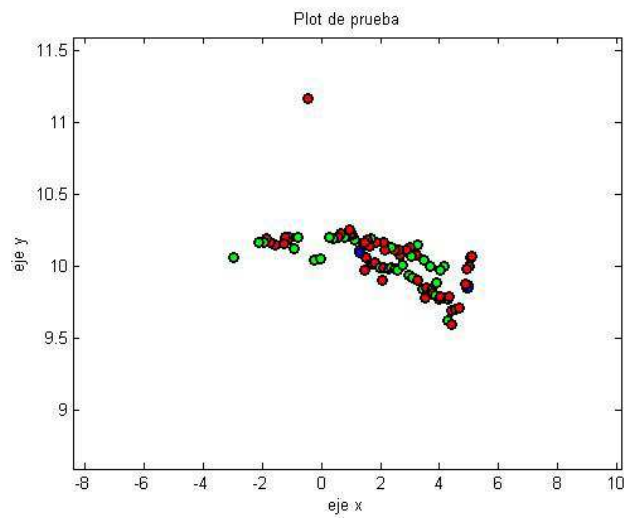


Imagen 54. Resultados caso 7 a 10m.

8. Discusión de los resultados.

Todos nuestros casos han sido tomados a una distancia de 5 ó 10 metros, en unos se pudieron tomar más y en otros se tomaron menos, pero en cualquiera de ellos fueron suficientes para ver que todos respondían de la misma manera y con la misma fiabilidad. Ya que hablamos de datos de fiabilidad es el momento de decir que como se puede observar las detecciones que han sido tomadas, se ha decidido dejar todas en el Anexo I, a pesar de no ser algunas de ellas válidas y las que hay en la memoria son sólo pequeños resúmenes para amenizar. La intención es observar el comportamiento de las detecciones y de sus probabilidades. Pues bien, como vemos en los resultados las detecciones simples son más numerosas que las detecciones dobles (68,2% frente a 31,8% respectivamente). Además se aprecia también que en primer lugar las detecciones dobles llevan una probabilidad de estimación mayor que las simples (0,85 frente a 0,68), también se observa un claro descenso de la probabilidad entre detecciones con las mismas condiciones, es decir, dos detecciones iguales en condiciones, no en valores, pues tiene una probabilidad menor en la detección posterior.

Se puede apreciar claramente que la frecuencia de toma de datos es mayor en la cámara que en el láser (66,4% frente a 33,6% respectivamente). Además se saca en conclusión que según van sucediéndose las detecciones la probabilidad de las mismas va bajando debido a que cada vez es más imprecisa la detección. Dicha detección es imprecisa porque no son siempre las mismas condiciones en las detecciones. Es muy difícil hacer dos detecciones exactas, incluso cuando consideramos que una detección es doble, como podemos apreciar hay una mínima variación de tiempo, Pero es tan insignificante que al sujeto no le da tiempo a cambiar de posición.

Lo que se espera y se puede ver en la mayoría de casos, son detecciones cada poco tiempo, eso quiere decir que las distancias recorridas por los peatones sean cortas, a su vez que estos dos datos sean pequeños implica que las velocidades sean aproximadamente proporcionales. Que uno de estos datos sea desproporcionado o se salga fuera de las proporciones, implica que si no varía todo en la misma proporción, se obtendrán muchas probabilidades fuera de rango y son las que al fin y al cabo estamos observando. Para cada una de estas desviaciones hay un motivo que sin ninguna duda puede ser estudiado y detectado, pero no se ha eliminado debido a que nos interesa saber cuándo y cómo pueden ser observados datos que se escapan fuera de nuestras especificaciones.

Un caso que mencionamos que puede ser estudiado es un largo proceso de detecciones de probabilidad 0. Eso quiere decir que la probabilidad calculada es demasiado pequeña y que puede expresarse directamente con 0. Se ha producido debido a que el espacio de tiempo transcurrido entre la detección anterior y esta es muy grande y esto lleva a que la velocidad sea mínima. Este ejemplo afecta a las detecciones posteriores y es un claro caso de estudio que podría omitirse o adaptarlo al resto para que las siguientes detecciones no se vieran afectadas. También se da el caso, en cuanto a los resultados de probabilidad, que una probabilidad sea tan pequeña que tenga valores exponenciales negativos (p.e. $2,0e-5$). Estos valores pueden darse por el motivo, al igual que los de probabilidad 0, de que un valor en los cálculos se desvíe mucho, es decir, cantidades desorbitadas de tiempo, de distancia recorrida por el peatón, un cambio de dirección momentánea en la detección del peatón, etc... En cualquier caso todos estos tipos de errores podrían ser estudiados para una posible mejora de las detecciones, debido a que estas detecciones afectan en las posteriores y una detección que podría ser buena acaba siendo mala por arrastrar dicho error.

Otros datos que pueden llamar la atención son un buen número de detecciones en las que se tienen valores negativos de posición, o de velocidad, o incluso de nueva estimación de la posición. La conclusión es que son tan válidos como los positivos. Lo único que nos hace indicar es la dirección del peatón en el

momento de la detección, la dirección de la velocidad que lleva y la nueva posición que se estima. En resumen, el peatón va a derecha o izquierda (visto en las imágenes).

Como se puede observar en los casos, hay muchas detecciones dobles que vienen precedidas de casos simples con valores de probabilidad estimada de detección con valores 0 o altamente exponenciales, y que sin embargo al ser una detección doble da una probabilidad buena y muy aceptable. Pues bien esto se debe a que en las detecciones dobles se utilizan valores distintos de probabilidad a priori, y esto hace que dichos valores sean en general tan buenos.

Podemos distinguir entre los dos tipos de detección doble, casos de '1,0' (con un 58,1% del total de las detecciones detectadas primero con láser) o '0,1' (un 2,4% detectadas primero con la cámara), que los casos '1,0' no son tan abundantes como los casos '0,1', se han dado así porque las lecturas de cámara son más abundantes que las del telémetro, pero en la mayor parte de las detecciones dobles, se detectaban primero con el láser. Y también hay que decir que no todas las detecciones dobles han sido consideradas buenas, debido a que hay algunas que han dado probabilidades estimadas muy bajas o 0.

9. Acciones futuras.

9.1 Cámaras.

En el sistema de cámaras podemos hacer varias propuestas para mejoras en el futuro nuestro sistema de detecciones.

Una opción es mejorar las calidades de la cámara, en cuanto a calidad de la visión y calidad de la detección, porque una cámara que nos dé más imágenes y además un mayor porcentaje de detecciones correctas en relación con las detecciones totales, podría suponer que el sistema mejorase.

Otra opción de estudio podría ser la colocación de una mayor cantidad de cámaras que nos ofrezca un mayor campo de visión de los peatones, así podrían ser detectados en un mayor rango de distancia.

Más opciones serían cambiar de posición las cámaras para estudiar que probabilidad de acierto tiene la cámara y el sistema en general.

9.2 Láser.

El sistema del telémetro láser, nos puede ofrecer las mismas opciones de futuro que el sistema de cámaras; es decir, mejorar el telémetro, ponerlo en distinta posición o añadir más telémetros al conjunto del sistema.

9.3 Cálculos.

En cuanto a lo que se refiere a los cálculos, se pueden mejorar o proponer varias opciones.

La primera es, por ejemplo, hacer un estudio de todos aquellos caso en los que la detección no es correcta. Todos esos caso podrían ser eliminados y así mejorar las detecciones futuras.

La segunda, se podrían proponer nuevos sistemas de cálculos; es decir, aquí se ha propuesto realizar las probabilidades a través de Gaussianas, ¿Por qué no usar otro sistema diferente a ver qué resultados da?

Por último, también dentro de los cálculos podríamos incluir la programación. Todo el código se puede depurar y optimizar. Además de poderse programar en cualquier otro lenguaje.

9.4 GPS.

Otra acción futura es la de añadir al sistema de fusión las detecciones del sistema GPS, en nuestro proyecto la fusión es sólo en base al láser y la cámara pero el GPS podría ser un tercer sistema en la

fusión que haría al conjunto más robusto e incluso más fiable ya que un peatón en una posición podría ser detectado por tres sensores diferentes.

10. Presupuesto.

Tabla 17. Presupuesto de Material

| CONCEPTO | COSTE UNITARIO | UNIDADES | IMPORTE |
|----------------------------------|----------------|----------|---------------|
| Telómetro láser SICK LMS 291-S05 | 2.000€/U | 1 | 2.000€ |
| Stereo Vision Bumblebee 2 | 300€/U | 1 | 300€ |
| Licencia de Matlab | 100€/U | 1 | 100€ |
| Total | | | 2.400€ |

Tabla 18. Presupuesto de Personal

| CONCEPTO | COSTE UNITARIO | UNIDADES | IMPORTE |
|------------------------------|----------------|----------|----------------|
| Ingeniero Técnico Industrial | 1.750€/mes | 12 meses | 21000€ |
| Total | | | 21.000€ |

Tabla 19. Presupuesto total

| CONCEPTO | COSTE UNITARIO | UNIDADES | IMPORTE |
|-------------------|----------------|----------|----------------|
| Coste de Material | 2.400€/U | 1 | 2.400€ |
| Coste de Personal | 21.000€/U | 1 | 21.000€ |
| Total | | | 23.400€ |

11. Conclusiones.

Para empezar lo que se ha querido hacer con este proyecto ha sido detectar y realizar un seguimiento de dichas detecciones para probar su fiabilidad; de hecho, estudiar los resultados en cuanto a las detecciones y su seguimiento, nos lleva a estudiar la fiabilidad que nos da el sistema en cuanto a la seguridad, y ese es el objeto real de la investigación.

Los resultados han sido satisfactorios aunque no todas fueron buenas (78,9%) y no todas fueron malas detecciones (21,1%). A través de ellas hemos podido observar que el algoritmo utilizado puede ser una buena forma para hacer fiable el sistema, pero para ello habría que ser aún más estrictos en cuanto a los criterios de búsqueda de datos. En realidad dicha búsqueda y su programación en Matlab han sido las reales dificultades de la investigación, además de las múltiples pruebas que se tuvieron que realizar para llegar a los resultados mostrados.

El estudio de todo dato ha sido muy útil primero para ver en qué puede fallar nuestro sistema, y cómo se puede mejorar, y segundo y más importante que es un sistema que puede ser muy válido para el control de la seguridad en muchos ámbitos de la vida.

12. Referencias.

- [1] Susan M. Thornton, Member, IEEE, Mike Hoffelder, and Daniel D. Morris, Member, IEEE, “Multi-sensor Detection and Tracking of Humans for Safe Operations with Unmanned Ground Vehicles”.
- [2] Frank J. Aherne; Neil A. Thacker; Peter I Rockett. “The Bhattacharyya metric as an absolute similarity measure for frequency coded data”, *Kybernetika*, Vol. 34 (1998), No. 4, [363]—368
- [3] Greg Welch and Gary Bishop, “An Introduction to the Kalman Filter”, Department of Computer Science University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3175
- [4] M. Bertozzi, E. Binelli, A. Broggi, and M. Del Rose, “Stereo Vision based approaches for Pedestrian Detection,” in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.3, June 20-26, 2005.
- [5] Xuan Song, Jinshi Cui, Huijing Zhao and Hongbin Zha, “A Bayesian Approach: Fusion of Laser and Vision for Multiple Pedestrians Tracking”, 2000.
- [6] Xuan Song, Jinshi Cui, Xulei Wang, Huijing Zhao and Hongbin Zha State Key Laboratory of Machine Perception, Peking University, China. “Tracking Interacting Targets with Laser Scanner via On-line Supervised Learning”, 2008.
- [7] H. J. Zhao, R. Shibusaki, “A novel system for tracking pedestrians using multiple single-row laser range scanners”, 2005.
- [8] J. Cui, H. Zhao, R. Shibusaki, “Fusion of Detection and Matching Based Approaches for Laser Based Multiple People Tracking”, 2006.
- [9] Nicola Bellotto and Huosheng Hu, Department of Computer Science, University of Essex Colchester CO4 3SQ, U.K., “Vision and Laser Data Fusion for Tracking People with a Mobile Robot”, 2006.
- [10] M. Scheutz, J. McRaven, Gy. Cserey, Department of Computer Science and Engineering, Department of Electrical Engineering University of Notre Dame, Notre Dame, IN 46556, USA, “Fast, Reliable, Adaptive, Bimodal People Tracking for Indoor Environments”
- [11] Y. Tsai, “A versatile camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses,” IEEE Journal of Robotics and Automation, pp. 323-344, 1987.
- [12] Pascal Vincent and Yoshua Bengio, “Manifold Parzen Window”, Dept. IRO, Université de Montréal, Canada.
- [13] Marco Fontana and K. Alan Loper, “Kronecker function rings: a general approach”.
- [14] S. Cong, L. Hong and D. Wicker, “Markov-chain Monte-Carlo approach for association probability evaluation”, 2004
- [15] Lauro Snidaro and Gian Luca Foresti, Dept. Mathematics and Computer Science, University of Udine (Italy). Ruixin Niu and Pramod K. Varshney, Dept. of Electrical Engineering and Computer Science, Syracuse University (USA), “Sensor Fusion for Video Surveillance”.
- [16] Dorin Comaniciu, Visvanathan Ramesh and Peter Meer. “Kernel-Based Object Tracking”
- [17] Método Pin-hole
- [18] Lauro Snidaro, Ingrid Visentini and Gian Luca Foresti, Dept. of mathematics and computer science, University of Udine (Italy), “Fusing multiple video sensors for surveillance”.
- [19] Nemanja Petrović, Ljubomir Jovanov, Aleksandra Pižurica, and Wilfried Philips. “Object Tracking Using Naive Bayesian Classifiers”, Dept. of Telecommunications and Information Processing (TELIN-IPI-IBBT), Ghent University, Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium.
- [20] Thierry Chateau, Vincent Gay-Belille, Frederic Chausse, and Jean-Thierry Lapresté. “Real-Time Tracking with Classifiers”, Lasmea, UMR6602, CNRS, Blaise Pascal University, Clermont-Ferrand, France.
- [21] Lauro Snidaro and Ingrid Visentini, “Fusion of heterogeneous features via cascaded on-line boosting”, Department of Mathematics and Computer Science, University of Udine, 2008.
- [22] Shai Avidan, “Support Vector Tracking”, IEEE Transactions on pattern analysis and machine intelligence, vol. 26, no. 8, August 2004.