



**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**  
**Facultad de Ingeniería en Electricidad y Computación**

“UTILIZACIÓN DE INFORMACIÓN MUTUA COMO FUNCIÓN DE  
CORRESPONDENCIA EN EL DOMINIO MULTIESPECTRAL”

**TESINA DE SEMINARIO**

Previo a la obtención de los títulos de:

**INGENIERO EN CIENCIAS COMPUTACIONALES**  
**ORIENTACIÓN SISTEMAS MULTIMEDIA**

PRESENTADA POR:

Ayllyn Nathaly Franco Rodríguez

GUAYAQUIL–ECUADOR

2015

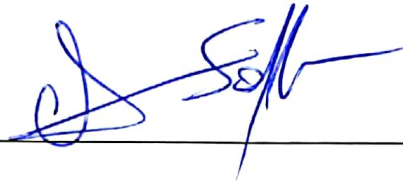
## **AGRADECIMIENTO**

Agradezco a los profesores de la ESPOL por el conocimiento impartido en las aulas. A los profesores Ángel Sappa y Gonzalo Luzardo por la guía en esta tesina. A los amigos por el apoyo recibido. Al disco Superunknown de Soundgarden por acompañarme en el desarrollo y escritura de este proyecto.

## **DEDICATORIA**

A mi madre, que hizo todo lo posible para  
la obtención de este título.

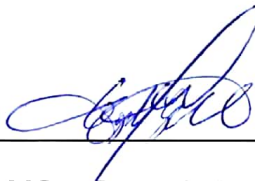
# TRIBUNAL DE SUSTENTACIÓN



---

Ph. D. Angel D. Sappa

**PROFESOR DEL SEMINARIO**



---

MSc. Gonzalo Luzardo Morocho

**PROFESOR DELEGADO**

## DECLARACIÓN EXPRESA

"La responsabilidad del contenido de esta tesina, nos corresponde exclusivamente; y el patrimonio intelectual de la misma a la Escuela Superior Politécnica del Litoral".

(Reglamento de Graduación de la ESPOL)

A handwritten signature in blue ink, appearing to read 'Ayllyn Franco', is written over a horizontal line.

**Ayllyn Franco Rodríguez**

## RESUMEN

El presente trabajo ha sido desarrollado como un proyecto final de graduación para la carrera de grado en Ciencias Computacionales de la Facultad de Ingeniería en Electricidad y Computación – FIEC de la ESPOL.

El propósito de este proyecto es ampliar el campo de trabajo y realizar un análisis de la información mutua como función de correspondencia entre imágenes del espectro visible y del espectro infrarrojo, determinando qué factores afectan este cálculo y de qué modo.

Se han puesto en práctica diferentes conceptos estudiados a lo largo de carrera de ingeniería y el seminario de graduación, del mismo modo de la revisión de literatura hecha.

Para esto, se desarrolló una interfaz gráfica donde se establecen los factores de cómputo, pudiendo obtener a partir de esto, diferentes resultados que se analizaron posteriormente. En este estudio se usaron imágenes infrarrojas (LWIR) y del espectro visible (VS), las imágenes fueron proporcionadas por un par estereoscópico multiespectral (VS-LWIR) del Centro de Visión por Computador de la Universidad Autónoma de Barcelona, España.

La tesis se encuentra organizada de la forma siguiente:

Capítulo 1 – Introducción: Este primer capítulo es introductorio y contiene un resumen acerca del proyecto realizado, sus objetivos, desarrollo, análisis del problema y la solución que se ha decidido plantear.

Capítulo 2 – Estado del arte: En este capítulo se describen conceptos, métodos y algoritmos importantes relacionados al desarrollo del proyecto, este capítulo se encuentra dividido en seis diferentes secciones siendo las de más interés los métodos de correspondencia de imágenes y la información mutua.

Capítulo 3 - Desarrollo: En este capítulo se detallan todos y cada uno de los procedimientos realizados para poder obtener finalmente un mapa de disparidad. Del mismo modo se describe la interfaz gráfica elaborada para el proyecto, la cual permite ingresar y seleccionar diferentes valores pertinentes al análisis de los cuales han dependido los resultados obtenidos.

Capítulo 4 – Análisis y evaluación: En este capítulo se presentaran los resultados obtenidos de las pruebas para poder identificar bajo qué condiciones se obtiene un mapa de disparidad más representativo de la escena.

Conclusiones y Recomendaciones: Se presentaran las conclusiones obtenidas del proyecto, del mismo modo al final se realizan comentarios y sugerencias para proyectos similares.



## ÍNDICE GENERAL

RESUMEN.....	VI
ÍNDICE GENERAL .....	IX
ÍNDICE DE FIGURAS.....	XII
ÍNDICE DE TABLAS .....	XV
ABREVIATURAS Y SIMBOLOGÍA .....	XVI
INTRODUCCIÓN.....	XVII
CAPÍTULO 1	
DESCRIPCIÓN Y JUSTIFICACIÓN DEL PROYECTO.....	1
1.1. Análisis del problema.....	1
1.2. Diseño de la solución.....	2
1.3. Objetivos.....	2
1.3.1. Generales .....	2
1.3.2. Específicos.....	3
1.4. Organización de la Tesis .....	4
CAPÍTULO 2	
ESTADO DEL ARTE.....	6
2.1 .Antecedentes históricos.....	6
2.2. Imágenes Multiespectrales .....	8

2.2.1. Imágenes infrarrojas .....	10
2.2.2. Aplicaciones.....	12
2.3. Sistemas automáticos de análisis.....	13
2.4. Reconocimiento de Patrones.....	14
2.4.1. Patrones .....	14
2.4.2. Reconocimiento de Patrones.....	15
2.5. Correspondencia de imágenes.....	16
2.5.1. Métodos de correspondencia.....	17
2.5.2. Dificultades .....	19
2.5.2.1. Amplitud de la búsqueda.....	19
2.5.2.2. Ambigüedad.....	19
2.5.2.3. Mal condicionamiento. ....	20
2.6. Información Mutua .....	20
2.6.1. Definición de la información mutua .....	21
2.6.2. Medidas de Información.....	22
2.6.2.1. Hartley.....	22
2.6.2.2. C. E. Shannon.....	23
2.6.3. Análisis de las definiciones de la información mutua .....	26
CAPÍTULO 3	
DESARROLLO .....	28
3.1. Discretización de imágenes .....	30
3.2. Selección de valores pertinentes .....	31

3.4. Histograma conjunto .....	32
3.5. Cálculo de la información mutua.....	33
3.6. Disparidad.....	35
3.7. Implementación de la interfaz .....	39
CAPÍTULO 4	
EVALUACIÓN Y ANÁLISIS DE RESULTADOS .....	42
4.1. Resultados Experimentales .....	42
4.1.1. Grupo 1 .....	42
4.1.2. Grupo 2.....	47
4.1.3. Grupo 3.....	50
4.2. Análisis de resultados .....	51
CAPÍTULO 5	
CONCLUSIONES Y RECOMENDACIONES.....	53
Conclusiones .....	53
Recomendaciones .....	55
Anexo A: Calibración de las cámaras para obtener la base de imágenes ....	56
A.1. Parámetros intrínsecos.....	56
A.2. Parámetros extrínsecos.....	60
Anexo B: Tablas de conversión de valores de gris a otra cantidad escogida	64
Bibliografía.....	69

## ÍNDICE DE FIGURAS

Figura 2. 1. Fluido biológico al microscopio .....	9
Figura 2.2. Gráfico de intensidades .....	9
Figura 2.3. Ejemplo de imagen infrarroja .....	10
Figura 2.4. Ejemplo de imagen infrarroja y en escala de grises. ....	13
Figura 3.1. Diagrama de flujo del proceso para análisis. ....	28
Figura 3.2. . Imagen VS con una ventana a buscar seleccionada .....	34
Figura 3.3. Imagen IR con el intervalo de búsqueda de la información mutua mínima.....	35
Figura 3.4. (a) Ejemplo de configuración de las cámaras del par estéreo. (b) Relación geométrica entre los parámetros del par estéreo para obtener la profundidad Z a partir de la disparidad d. ....	36
Figura 3.5. (a) Imagen VS con un punto a buscar, la estrella. Imagen LWIR mostrando el desplazamiento -el punto hallado por cuadros, cada cuadro representa una unidad.....	37

Figura 3.6. Ejemplo. En (a) se encuentra un ejemplo de dos imágenes VS del mismo plano. (b) es el resultado del mapa de disparidad obtenido .....	38
Figura 3.7. Interfaz del sistema elaborado.....	40
Figura 4.1. Parte de los resultados, a la izquierda imágenes VS, columna de en medio imagen LWIR, a la derecha resultado obtenido. ....	44
Figura 4.2. Resultados obtenidos cambiando los niveles de gris a 16. ....	45
Figura 4.3. Resultados obtenidos cambiando el tamaño de la ventana a 21x21 pixeles.....	46
Figura 4.4. Algunos de los resultados obtenidos al trabajar con los datos de la Tabla 5 para el segundo grupo de análisis. ....	48
Figura 4.5. Resultados de trabajar con 16 niveles de gris en el 2do grupo ...	49
Figura 4.6. Resultados de trabajar con 16 niveles y aumentando el tamaño de ventanas a 21x21 pixeles. ....	49
Figura 4.7. Algunos de los resultados obtenidos al trabajar con los datos de la Tabla 5 para el segundo grupo de análisis. ....	51
Figura A.1. Modelo de Calibración pin-hole. ....	57

Figura A.2 Imágenes del patrón tomadas con la cámara VS.....	58
Figura A.3 Rectángulos detectados en el primer set de imágenes VS .....	58
Figura A.4.(a) Imagen original VS. (b) Imagen rectificada VS. ....	59
Figura A.5.(a) Imagen original VS. (b) Imagen rectificada VS. ....	60
Figura A.6. Posición en el espacio de cada imagen capturada con respecto al par estéreo.....	62
Figura A.7. Posición en el espacio de cada imagen capturada con respecto al par estéreo.....	62

## ÍNDICE DE TABLAS

Tabla 1. Descripción de los métodos de correspondencia.....	18
Tabla 2. Distribución de grupos de valores de gris para discretizar a 8 valores de gris.....	31
Tabla 3. Ejemplo de un histograma conjunto para el análisis de dos ventanas con 8 valores de gris diferentes posibles, de tamaño 10x10. ....	33
Tabla 4. Valores escogidos para el primer grupo de análisis.....	43
Tabla 5. Valores escogidos para el segundo grupo de análisis.....	47
Tabla 6. Valores escogidos para el tercer grupo de análisis.....	50
Tabla B.1. Distribución de grupos de valores de gris para discretizar a 16 valores de gris. ....	65
Tabla B.2. Distribución de grupos de valores de gris para discretizar a 24 valores de gris. ....	66
Tabla B.3. Distribución de grupos de valores de gris para discretizar a 32 valores de gris. ....	68

## ABREVIATURAS Y SIMBOLOGÍA

- **LWIR:** Long wave length infrared
- **MWIR:** Mid-Wave InfraRed
- **NIR:** Near InfraRed
- **SWIR:** Short-Wave InfraRed
- **VS:** Visible Spectrum



# INTRODUCCIÓN

Este proyecto consiste realizar un análisis de la información mutua como función de correspondencia entre imágenes del espectro visible y del espectro infrarrojo, determinando qué factores afectan este cálculo y de qué modo, a través de la elaboración de una interfaz y sistema de cómputo capaz de recibir y trabajar con imágenes del espectro infrarrojo y del espectro visible.

Dentro de las áreas de Procesamiento de Imágenes y Visión por computador existen ya muchas aplicaciones que se dedican a diferentes procedimientos como son el realzado y la disminución de ruido en imágenes, detección de puntos de interés, compresión de imágenes, restauración y correspondencia de puntos, aplicaciones que en un principio se desarrollan para trabajar solo en el espectro visible.

La correspondencia de imágenes dentro del dominio del espectro visible sufre ya ciertos problemas y se han establecido distintos métodos para su cómputo. En este proyecto se va a analizar cómo afectan estos factores al momento de trabajar en el dominio multiespectral. La interfaz desarrollada servirá para seleccionar las imágenes, establecer distintos factores de análisis para crear diferentes escenarios y luego mostrará los resultados que deberán ser guardados para poder hallar diferencias entre cada escenario.

# CAPÍTULO 1

## DESCRIPCIÓN Y JUSTIFICACIÓN DEL PROYECTO

### 1.1. Análisis del problema

El problema principal es analizar si la información mutua sirve o no como función de correspondencia entre imágenes del espectro visible y del espectro infrarrojo de la misma escena. Esta labor es fácil para el ojo humano, puesto que a pesar de que las imágenes del espectro visible contienen información diferente a las imágenes infrarrojas, información fotométrica y térmica respectivamente, el cerebro es capaz de relacionar esta información, llenar o ignorar espacios y decidir cuál punto de una imagen del espectro visible corresponde a otro punto de la imagen del espectro infrarrojo.

Sin embargo esta labor puede ser algo complicada de realizarse en un sistema de cómputo, por lo cual se deben definir y analizar los factores que competan al cálculo automático de la correspondencia.

## **1.2. Diseño de la solución**

- Se debe identificar los conceptos y factores pertinentes involucrados en la correspondencia de imágenes.
- Desarrollar una interfaz gráfica (GUI) en el framework QT, en donde se pueda manipular con facilidad las variables pertinentes al análisis, así como también poder observar los resultados obtenidos.
- Elaborar el cálculo automático de la información mutua mínima, para luego obtener un mapa de disparidad que se deberá analizar, utilizando como lenguaje de programación C++ y OpenCV.
- Analizar los resultados obtenidos.

## **1.3. Objetivos**

### **1.3.1. Generales**

El objetivo del proyecto es analizar el uso de la información mutua como función de correspondencia entre imágenes del espectro

visible y el espectro infrarrojo, tomando en cuenta diferentes factores al momento del análisis como son la cantidad de valores de gris con las que deseo tratar las imágenes, el rango de búsqueda para el cálculo del mínimo de información mutua y el tamaño de ventana con el cual recorro ambas imágenes.

### **1.3.2. Específicos**

- Estudiar el estado del arte en información mutua y correspondencia de imágenes.
- Estudiar los distintos factores y conceptos a considerar para el cálculo de información mutua entre una variedad establecida de imágenes o ventanas.
- Adaptar estos factores y conceptos relacionados al cálculo de la información mutua al dominio multiespectral.
- Evaluar el cálculo de la información mutua con distintas imágenes del espectro visible y el espectro infrarrojo, considerando diferentes factores dentro del análisis.
- Fortalecer los conocimientos y habilidades en el ámbito de procesamiento de análisis de imágenes a través del estudio de la teoría y la implementación de proyectos con enfoque de aplicación real.

## 1.4. Organización de la Tesis

La tesis se encuentra organizada de la forma siguiente:

Capítulo 1 – Introducción: Este primer capítulo es introductorio y contiene un resumen acerca del proyecto realizado, sus objetivos, desarrollo, análisis del problema y la solución que se ha decidido plantear.

Capítulo 2 – Estado del arte: En este capítulo se describen conceptos, métodos y algoritmos importantes relacionados al desarrollo del proyecto, este capítulo se encuentra dividido en seis diferentes secciones siendo las de más interés los métodos de correspondencia de imágenes y la información mutua.

Capítulo 3 - Desarrollo: En este capítulo se detallan todos y cada uno de los procedimientos realizados para poder obtener finalmente un mapa de disparidad. Del mismo modo se describe la interfaz gráfica elaborada para el proyecto, la cual permite ingresar y seleccionar diferentes valores pertinentes al análisis de los cuales han dependido los resultados obtenidos.

Capítulo 4 – Análisis y evaluación: En este capítulo se presentaran los resultados obtenidos de las pruebas para poder identificar bajo qué condiciones se obtiene un mapa de disparidad más representativo de la escena.

Conclusiones y Recomendaciones: Se presentaran las conclusiones obtenidas del proyecto, del mismo modo al final se realizan comentarios y sugerencias para proyectos similares.

# CAPÍTULO 2

## ESTADO DEL ARTE

En este capítulo se ha realizado una revisión de literatura sobre los temas pertinentes al desarrollo del presente trabajo.

### 2.1. Antecedentes históricos

La correspondencia de regiones o elementos es una tarea que se lleva a cabo en diferentes disciplinas tales como el Procesamiento de imágenes o la Visión por computador, e incluso en labores tan simples como armar un rompecabezas. Cada una de estas disciplinas posee sus propios problemas y dificultades en la correspondencia de elementos, problemas que pueden afectar ciertos factores como la rapidez de cómputo de los sistemas de análisis o incluso la precisión de los mismos. Por lo tanto la labor de 'matching de elementos' es un campo de investigación que se encuentra en constante evolución y crecimiento.

Hablando de aspectos históricos y de las primeras personas que realizaron trabajos en el 'matching de elementos', los primeros experimentos datan de los años cincuenta cuando Hobrough planteó una solución de naturaleza analógica, en la que los correladores realizados mediante hardware comparaban (correlaban) los niveles de gris de dos imágenes. Luego de eso, llegando a los años ochenta y a lo largo de los mismos, la correspondencia de imágenes y elementos se enfocó en el campo digital. Kreiling (1976), fue una de las primeras personas en aplicar la correlación digital para general algunos MDE, luego de eso se continuaron añadiendo técnicas de correlación a los instrumentos, no obstante no se encontró una solución universal pese al esfuerzo realizado [1].

Para poder realizar estos experimentos, primero se ha debido comprender y descubrir todo aquello que compete al análisis de una escena: el análisis de una imagen busca encontrar y cuantificar propiedades de los objetos presentes en estas, para lo cual es necesario entender cuáles son los elementos que componen estas imágenes, el origen y la naturaleza de las mismas. Los métodos de análisis utilizados sobre una imagen están bastante relacionados con la naturaleza de dichas imágenes.

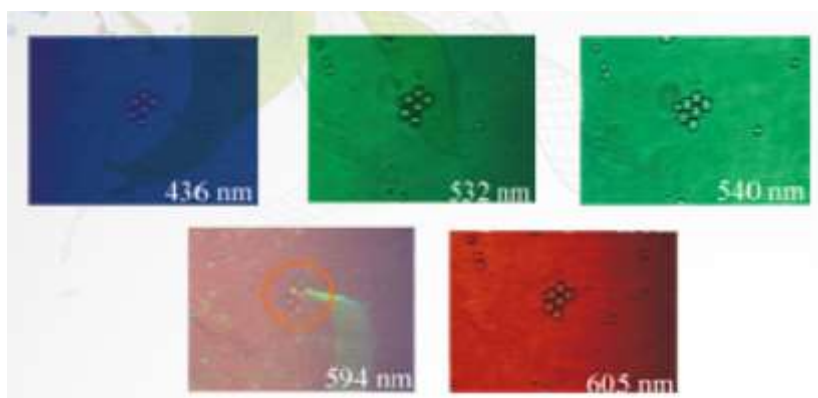


La percepción de la vista humana por su parte es bastante limitada, existen 64 tonos de color perceptibles y 16 de grises, los cuales le permiten al ser humano tomar decisiones en el mundo cotidiano en el que vive al momento de analizar circunstancias, objetos y demás, sin embargo la creación de nuevas tecnologías le ha permitido al ser humano observar imágenes que bajo el sentido normal de la visión, son invisibles ante su apreciación, pudiendo ser estas imágenes acústicas, magnéticas, gravimétricas, neurografías, etc.

## **2.2. Imágenes Multiespectrales**

Una imagen multiespectral es una que captura los datos de la imagen a frecuencias específicas a través del espectro electromagnético. Reproduce la figura del objeto en función de la longitud de onda que este objeto se encuentre reflejando o (emitiendo) [2].

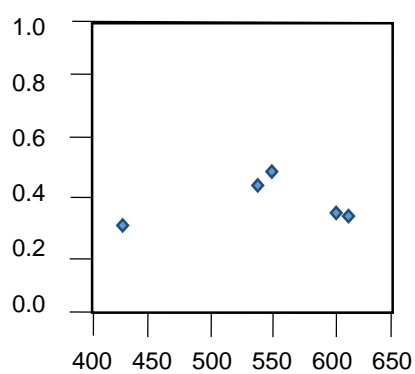
Las imágenes a continuación fueron captadas con una cámara multiespectral con filtros espectrales en 435, 532, 540, 594 y 605nm  $\pm$  20nm (representación en falso color) como vemos en la Figura 2.1.



**Figura 2. 1. Fluido biológico al microscopio**

**Fuente: Headwall Photonics**

Si normalizamos la intensidad y representamos los valores del mismo píxel en las cinco imágenes en un gráfico el resultado sería el de la Figura 2.2.



**Figura 2.2. Gráfico de intensidades**

**Fuente: Headwall Photonics**

Las longitudes de onda pueden estar separadas por filtros o por el uso de instrumentos que son susceptibles a longitudes de onda particulares, incluyendo luz de las frecuencias más allá del rango de luz visible, como la infrarroja [2] (ver Figura2.3).



**Figura 2.3. Ejemplo de imagen infrarroja**

**Fuente: Headwall Photonics**

### **2.2.1. Imágenes infrarrojas**

Las imágenes infrarrojas (IR), nos permiten capturar espectros lumínicos que el ojo humano no puede divisar. La visión humana normal puede ver solo la luz visible que es una pequeña franja del espectro electromagnético [3].

La energía IR se genera por la vibración de los átomos y moléculas y se comporta de manera similar a la luz visible, donde puede ser reflejada, refractada, absorbida y emitida. Cuanto más estos átomos y

moléculas se mueven, mayor es la temperatura del objeto parte del espectro electromagnético [4].

Para ver la luz infrarroja se requieren cámaras infrarrojas con sensores especiales que no necesitan de la luz visible para operar. Las infrarrojas pueden estar divididas dependiendo de su longitud de onda:

- Infrarrojo Cercano (NIR: Near InfraRed): 780–1000nm, usada mayormente para imágenes de vegetación.
- Infrarrojo Corto (SWIR: Short-Wave InfraRed): 1000–3000nm, es usada en imágenes de vegetación, contenido húmedo del suelo y algunos fuegos forestales.
- Infrarrojo Medio (MWIR: Mid-Wave InfraRed): 3000– 5000nm, es usada para imágenes del suelo, humedad, características geológicas, fuegos, arcillas, silicatos.
- Infrarrojo Lejano (LWIR: Long-Wave InfraRed): 8000– 12000nm, usa la radiación emitida en lugar de la reflejada, para imágenes de estructuras geológicas, diferencias térmicas en las corrientes de agua, fuegos y para estudios nocturnos.

### **2.2.2. Aplicaciones**

Debido a que la luz infrarroja penetra en zonas que la luz visible no puede ver y revela objetos oscuros, por esta razón tiene un gran número de usos y aplicaciones nuevas están siendo creadas para extender aún más su alcance. Los primeros usos se dieron originalmente por los militares para producir cámaras de visión nocturna, binoculares y mira de armas, dispositivos de imágenes infrarrojas se utilizan ahora con regularidad por diversos organismos civiles.

La policía, los bomberos y equipos de búsqueda de rescate utilizan cámaras infrarrojas, respectivamente, para atrapar criminales en la oscuridad y para salvar personas, animales y bienes atrapados en catástrofes. Con las imágenes LWIR los técnicos pueden eliminar riesgos potenciales mediante la localización de partes sobrecalentadas, químicos y fugas de sustancias de tratamiento delicado y peligroso.

Dentro de la correspondencia de imágenes se encuentran aplicaciones que nos permiten reconstruir escenas y modelos 3D a través de la correspondencia de texturas obtenidas mediante imágenes infrarrojas. Sin embargo, según la geometría diferente de la adquisición de las

imágenes se requieren métodos variados para la coincidencia de imágenes [5].

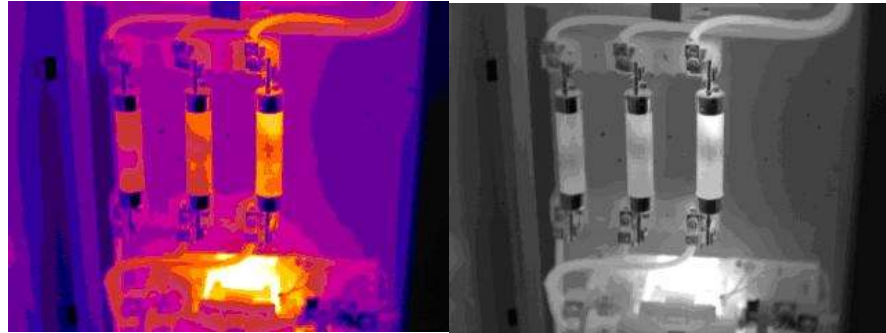


Figura 2.4. Ejemplo de imagen infrarroja y en escala de grises.

Fuente:[<http://www.grainger.com/content/qt-thermal-imaging-applications-uses-features-345>]

### 2.3. Sistemas automáticos de análisis.

Los sistemas automáticos de análisis de imágenes pueden realizar en ciertas ocasiones el trabajo del análisis cerebral humano sobre una imagen, de igual o mejor forma, sin embargo, siguen existiendo casos en los que el análisis humano es mejor que cualquier software de análisis automático debido a que la función cerebral es capaz de trabajar con muy poca información o con información distorsionada debido al ruido presente en una imagen, haciendo posible por ejemplo que al realizar una detección de bordes, el ser humano encaje líneas o descarte otras que no pertenecen a ningún objeto en cuestión sino que

están presentes en una escena debido a algún efecto como el Blur. El ser humano por su parte es capaz de relacionar objetos y establecer correspondencia con información faltante debido a que conforme vive, adquiere una base de conocimientos al respecto de las cosas de forma natural.

Lo que se busca no es lograr que los sistemas automáticos copien las funciones humanas de análisis de imágenes, sino que las entiendan, para de este modo poder interpretar escenas y entender realmente qué es lo que hay y sucede en las mismas.

*"La abstracción va unida a una relación lógica entre las propiedades y la función o naturaleza del objeto; entre más características se conozcan más se entiende la esencia del mismo."* [6].

## **2.4. Reconocimiento de Patrones**

### **2.4.1. Patrones**

Un patrón dentro del campo del análisis de imágenes es un arreglo de descriptores. El nombre característica es usado en ocasiones en la literatura de reconocimiento de patrones para llamar a un descriptor [7].

Una clase de patrones es una familia de patrones que comparten algunas características en común. El reconocimiento de patrones por parte de las máquinas involucra técnicas para asignar patrones a sus respectivas clases, tratando de conseguir esto de forma automática y con el mínimo de intervención humana posible.

Algunos de los arreglos de patrones más usados en la práctica son los vectores (considerados usualmente para descripciones cuantitativas), cadenas de caracteres y árboles (para descripciones estructurales).

#### **2.4.2. Reconocimiento de Patrones**

El reconocimiento de algo por parte del ser humano ocurre aparentemente de forma natural, sin embargo para conocer un objeto es necesario contar con los siguientes elementos:

- a) las propiedades fundamentales que lo distinguen
- b) una etiqueta o nombre asociado
- c) la identificación de sus funciones o propósito [8].

Antes de proceder a una operación de reconocimiento de patrones es necesario haber modelado matemáticamente las propiedades morfológicas y espectrales del objeto que se desea identificar. Un modelo matemático es una forma de describir algo en lenguaje



matemático, proporcionando una vía para la organización, reconocimiento y representación de la información pertinente a un patrón o clase de éstos [5]. A partir del modelo se construye el algoritmo computacional que sirve para ejecutar la tarea de reconocimiento en el análisis de una imagen digital.

Para reconocer un objeto es necesario, entonces, llevar a cabo los siguientes pasos:

- abstracción de sus propiedades fundamentales
- modelación matemática de estas propiedades y
- medida de las mismas de acuerdo con una métrica [9].

## **2.5. Correspondencia de imágenes**

La correspondencia de imágenes es una labor que busca hallar la imagen, región, punto o sección correspondiente de una imagen previamente seleccionada, en otra u otras imágenes, es decir, hallar su entidad homóloga.

En lo que se refiere a la terminología presente en la correspondencia de imágenes, no existe una terminología estandarizada y en ocasiones esto actuará como una fuerte barrera entre las diferentes disciplinas que participan en la correlación de imágenes como por ejemplo la Visión Computacional, el Análisis de Imagen y la Fotogrametría.

En Fotogrametría la correspondencia de imágenes se denomina a veces correspondencia estereoscópica automática o simplemente correlación, mientras que en la Visión Computacional se denomina problema de correspondencia.

Conviene saber algunos conceptos básicos sobre la correspondencia de imágenes:

**Entidad homóloga o conjugada:** es un término más genérico que punto homólogo. Se denominan entidades conjugadas a las imágenes de los elementos del espacio objeto, que incluyen puntos, líneas y superficies.

**Entidad de correspondencia:** es la primitiva, que se compara con primitivas de otras imágenes para buscar entidades homólogas. Las primitivas incluyen niveles de gris, elementos extraídos y descripciones simbólicas.

### **2.5.1. Métodos de correspondencia**

Un método de correspondencia debe calcular la semejanza entre dos entidades posiblemente homólogas [10]. El nombre del método se añade tras el nombre de la entidad de correspondencia. Existen diferentes métodos de correspondencia de imágenes:

<b>Método de Correspondencia</b>	<b>Índice de correspondencia</b>	<b>Entidades de correspondencia</b>
Por áreas	Correlación, mínimos cuadrados	Niveles de gris
Basadas en características	Función de costo	Bordes, regiones
Simbólica	Función de costo	Descripción simbólica

**Tabla 1. Descripción de los métodos de correspondencia.**

- Índice de Semejanza: es una medida cuantitativa de lo buena que es la correspondencia de las entidades entre sí [10]. Generalmente el grado de semejanza se determina mediante una función de costo.
- Estrategia de correspondencia: se utiliza para referirse al esquema de conjunto que tiene la solución del problema de la correspondencia de imágenes [10].

Se debe tener presente que al hablar sobre correspondencia de imágenes, también se tiene la posibilidad de trabajar con múltiples imágenes, no solo un par.

### **2.5.2. Dificultades**

Existen muchos problemas que ocurren al momento de la automatización de búsqueda de puntos homólogo, causados usualmente por los siguientes factores:

#### **2.5.2.1. Amplitud de la búsqueda**

Cuando se tiene que buscar el punto por toda una imagen en lugar de una región de la misma, entre mayor la resolución de dicha imagen, crece el tiempo de procesamiento.

#### **2.5.2.2. Ambigüedad**

Se describe como la posibilidad de que algo pueda interpretarse o entenderse de varios modos. Sucede cuando el sistema encuentra más de un punto de correspondencia en dicha imagen y se produce si la entidad de correspondencia no es suficientemente característica (única). La solución pasará por la selección de más entidades de correspondencia únicas. [9]

### **2.5.2.3. Mal condicionamiento.**

Una imagen se caracteriza por su debilidad geométrica y radiométrica. Esta debilidad se concreta en la posibilidad de que al buscar el punto homólogo exista:

- Solución inexistente: puede no haber punto homólogo debido a ocultaciones, desplazamientos por relieve, diferencias radiométricas.
- Solución múltiple: puede haber más de una solución (aparente) en el caso de patrones repetitivos asociadas a cultivos, edificaciones, etc.
- Solución inestable: la solución puede quedar afectada por condiciones geométricas o radiométricas de manera que en diversas imágenes pueden detectarse puntos homólogos distintos (para un mismo punto de referencia) [11].

## **2.6. Información Mutua**

Conociendo ahora que un método basado en similitud de imágenes consiste en un modelo de transformación que se aplica para hacer referencia de coordenadas de una imagen para localizar sus

coordenadas correspondientes en el espacio de otra imagen, podemos analizar la función de costo a utilizar en este trabajo, siendo en este caso la información mutua.

### 2.6.1. Definición de la información mutua

La información mutua es una medida de información de la dependencia estadística de dos variables, o también puede considerarse como la cantidad de información que una variable contiene acerca de otra. Cualitativamente puede ser considerada como la medida de qué tan bien una imagen explica a otra.

Esto se explica, para dos variables discretas  $X$  e  $Y$  cuya probabilidad de distribución conjunta es  $P_{XY}(x, y)$ , la información mutua entre ellos, que se denota  $I(x, y)$ , está dada por:

$$I(X; Y) = \sum_{x,y} P_{XY}(x, y) \log \frac{P_{XY}(x, y)}{P_X(x)P_Y(y)} = E_{P_{XY}} \log \frac{P_{XY}}{P_X P_Y} . \quad (1)$$

Donde  $P_X(x)$  y  $P_Y(y)$ , son los marginales.

$$P_X(x) = \sum_y P_{XY}(x, y) \quad (2)$$

$Y P_Y(y) = \sum x P_{XY}(x, y)$  y  $EP$  es el valor esperado de la distribución  $P$ . La atención se centra en variables discretas, pero la mayoría de los resultados obtenidos para las variables discretas se extienden muy naturalmente a las continuas, uno puede reemplazar las sumas por integrales. Sin embargo se debe tener en cuenta que la sustitución formal de sumas por integrales esconde una gran cantidad de sutileza, y, para las distribuciones que no son suficientemente suaves, puede incluso no funcionar [12].

A continuación una breve explicación sobre las medidas de información para poder interpretar la ecuación (1)

## **2.6.2. Medidas de Información**

### **2.6.2.1. Hartley**

En el año 1928, Hartley, ante la idea de Nyquist de medir velocidades de señales transmitidas en el telégrafo [13], quiso ir más allá y usó por primera vez la palabra “información” en sus trabajos, dejando claro que es una cantidad medible.

La medida de información Hartley considera lo siguiente:

- Los mensajes con cadenas de caracteres de un alfabeto fijo

- La cantidad de información contenida de ese mensaje debe ser una función del total del número posible de mensajes
- Si existe un alfabeto con  $S$  símbolos, entonces existen  $S^n$  mensajes de longitud  $n$ .
- La cantidad de información contenida en dos mensajes debe ser la suma de información contenida en esos mensajes individuales.

La única función que satisface estos requerimientos es la función logaritmo, por lo cual la relación que usó fue

$$H = n \log S \quad (3)$$

En donde  $n$  es la longitud del mensaje y  $S$  es el número de posibles valores cada símbolo en el mensaje transmitido. Usando esto por lo tanto, la medida información se convertía en un número decimal, luego de un tiempo esta medida de información fue llamada el Hartley en su honor.

#### **2.6.2.2. C. E. Shannon**

En el año de 1948, Shannon propuso una variante, considerando que  $X$  es una variable real discreta con  $n$  posibles resultados  $\{x_1, \dots, x_n\}$ , la



probabilidad de esos resultados fuesen  $x_i$  es  $p_x(x_i)$ . La información contenida en un mensaje acerca de la salida de  $X$  sería

$$-\log p_x(x_i) \quad (4)$$

Y debe cumplir con las siguientes condiciones:

- Debe ser máxima cuando  $p_x(x)$  es uniforme, y en este caso se debe aumentar con el número de valores posibles que  $x$  puede tomar.
- Debe seguir siendo el mismo si reordenamos las probabilidades asignadas a diferentes valores de  $x$ .
- La incertidumbre acerca de dos variables aleatorias independientes debe ser la suma de las incertidumbres acerca de cada uno de ellos.

Luego mostró que la única medida de la incertidumbre que satisface todas estas condiciones es la entropía, definida como:

$$H(X) = - \sum_x P_X(x) \log P_X(x) = -E_{P_X} \log P_X \quad (5)$$

Aunque no es particularmente obvio a partir de esta ecuación,  $H(X)$  tiene una interpretación muy concreta:

Supongamos que  $x$  es elegido al azar de la distribución  $PX(x)$ , y alguien que conoce la distribución  $PX(x)$ , se le pide que adivine cuál  $x$  fue elegido por pedir sólo sí / no. Si esta persona usa la estrategia de las preguntas que formula óptima, que consiste en dividir la probabilidad por la mitad en cada conjetura haciendo preguntas como "¿es  $x$  mayor que  $x_0$ ?", entonces el número medio de preguntas sí / no que se necesita para adivinar  $x$  se encuentra entre  $H(X)$ , y  $H(X) + 1$  [14]. Esto le da un sentido cuantitativo a la "incertidumbre": es el número de preguntas sí / no que se necesitan para adivinar variables aleatorias, dado el conocimiento de la distribución subyacente y tomando la estrategia óptima de las preguntas.

La entropía condicional es la incertidumbre media acerca de  $X$  después de observar una segunda variable aleatoria  $Y$ , y está dada por:

$$H(X|Y) = \sum_y P_Y(y) \left[ - \sum_x P_{X|Y}(x|y) \log(P_{X|Y}(x|y)) \right] = E_{P_Y} \left[ - E_{P_{X|Y}} \log P_{X|Y} \right]$$

(6)

Donde  $P_{X|Y}(x|y) (\equiv \frac{P_{XY}(x,y)}{P_Y(y)})$ , es la probabilidad condicional de  $X$ , dado  $Y$ . Con las definiciones de  $H(X)$ , y  $H(X|Y)$ , la ecuación (1) se puede escribir:

$$I(X, Y) = H(X) - H(X|Y). \quad (7)$$

### 2.6.3. Análisis de las definiciones de la información mutua

La primera definición dada en la ecuación (1), se encuentra íntimamente relacionada a la Divergencia Kullback-Leiber entre dos distribuciones [15], en este caso la medida de información mutua  $I(X, Y)$  se encuentra en dependencia de estas dos distribuciones, y su valor al hacer la correspondencia de imágenes será maximizado cuando las imágenes con las que se hace la correspondencia se encuentren alineadas.

Interpretando esta definición de la información mutua partiendo de la entropía, ya conocimos otra de las definiciones posibles de este valor, siendo esta la ecuación (7) la cual también se puede interpretar como

$$I(X, Y) = H(X) + H(Y) - H(X, Y) \quad (8)$$

Todo esto nos indica lo siguiente:

- Maximizar la información mutua es equivalente a minimizar la entropía conjunta.

- La ventaja en el uso de información mutua sobre la entropía conjunta es que incluye la entropía de la entrada individual.

# CAPÍTULO 3

## DESARROLLO

Basándonos en la literatura propuesta y tomando en cuenta las diferentes definiciones, métodos y posibles factores pertinentes, el trabajo a realizar con una base de imágenes del espectro visible (VS) y del espectro infrarrojo lejano (LWIR) se especifica de la forma siguiente:

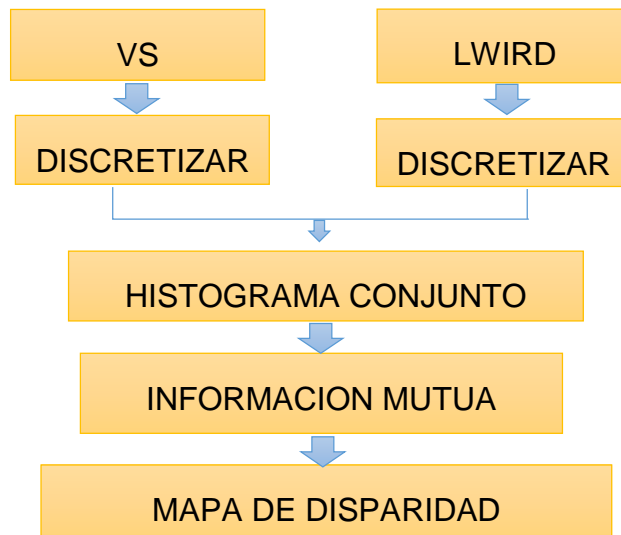


Figura 3.1. Diagrama de flujo del proceso para análisis.

El método consiste en seleccionar imágenes VS y LWIR de una base definida, imágenes que se encuentran en escala de grises y que se deben discretizar a una cantidad de valores de gris nueva escogida entre diferentes valores posibles para luego proceder a calcular el histograma en conjunto de estas dos imágenes.

Estas imágenes serán recorridas usando ventanas con un tamaño de valor también a escoger, buscando una ventana de la imagen VS en la imagen LWIR usando como función de costo la información mutua.

Al escoger una ventana en la imagen LWIR dentro de un rango establecido de búsqueda, como ventana homóloga de la imagen VS, podemos calcular la distancia a la cual se encuentra esta ventana homóloga y con esto obtener la disparidad entre esos puntos.

Al obtener la disparidad de cada punto, podemos obtener una matriz de disparidades mejor conocido como un mapa de disparidad. A continuación se explica en detalle cada etapa.

Las imágenes usadas en este análisis fueron proveídas y su obtención se explica en el Anexo A, para futuras reproducciones del proceso de obtención en trabajos posteriores.

### 3.1. Discretización de imágenes

Se seleccionan dos imágenes del mismo escenario; una imagen VS y una imagen LWIR, estas imágenes serán discretizadas a otra cantidad de valores de gris que el usuario escoja, estas cantidades pueden ser: 8, 16, 24 o 32 valores de gris.

Para esto se dividen los 256 valores posibles de gris actuales para la cantidad  $x$  escogida de valores de gris nuevos, estableciendo una cantidad  $x$  de grupos que representan estos nuevos valores. Cada grupo comprende un rango de valores de tamaño  $256/x$ .

Cada pixel de las imágenes iniciales se discretiza buscando en cuál rango está, hallando así en el número de grupo que se encuentra y con este número su nuevo valor de gris.

Por ejemplo para una imagen de 8 posibles valores de gris, cada grupo tendrá el tamaño de 32, considerando  $256/8$  y cada pixel de las imágenes iniciales tomará el valor del grupo en el que se encuentre, recorriendo toda la imagen y convirtiendo sus valores.

<b>Grupo 1</b>	<b>Grupo 2</b>	<b>Grupo 3</b>	<b>Grupo 4</b>
0-31	32-63	64-95	96-127
<b>Grupo 5</b>	<b>Grupo 6</b>	<b>Grupo 7</b>	<b>Grupo 8</b>
128-159	160-191	192-223	224-255

**Tabla 2. Distribución de grupos de valores de gris para discretizar a 8 valores de gris.**

Por lo tanto un pixel de la imagen VS o LWIR con un valor de gris de 220 pasará a tener el valor de 7 por hallarse en el grupo número 7, solo para el caso expuesto de 8 posibles valores de gris.

Al escoger otra cantidad de valores de gris, los valores cambiarán. Se puede observar este cambio en el Anexo B.

### **3.2. Selección de valores pertinentes**

En esta etapa escogemos algunos valores que afecta la rapidez y los resultados del análisis de las imágenes, estos valores son:

- El tamaño de la ventana con la cual vamos a recorrer toda la



imagen tomando como referencia cada uno de los píxeles de la misma. Los tamaños de ventana establecidos a escoger son: 11, 21, 33 y 51 píxeles.

- El intervalo de búsqueda  $x$ , estableciendo este de donde a dónde vamos a buscar la ventana homóloga de la ventana VS en la imagen LWIR. Este intervalo de búsqueda se puede escoger entre los valores de: 10, 15 y 32 píxeles, considerando que se realiza una búsqueda en  $x+\text{intervalo}$  y  $x-\text{intervalo}$ , antes y después del punto.

### 3.3. Histograma conjunto

Se creó un histograma conjunto seleccionando un conjunto de características locales de píxeles y la construcción de un histograma multidimensional. Cada entrada en un histograma conjunto contiene el número de píxeles de la imagen que se describen por una combinación particular de valores de características [16].

La elaboración del histograma conjunto consiste en tomar dos imágenes discretizadas, en este caso serán dos ventanas, una VS y otra LWIR, y hacer un conteo de pares de píxeles dependiendo de sus valores de gris nuevos.

Esto quiere decir que si la posición  $(a, b)$  en la ventana VS tiene el valor de gris  $X$ , y la misma posición  $(a, b)$  en la ventana LWIR tiene el valor de  $Y$ , perteneciendo  $X$  e  $Y$  al conjunto posibles de valores, en el histograma conjunto, el valor  $(X, Y)$  aumentará una unidad. Este procedimiento se realiza recorriendo todas las coordenadas de las ventanas VS y LWIR.

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>
<b>1</b>	0	0	0	0	0	0	0	0
<b>2</b>	1	2	0	23	10	0	0	0
<b>3</b>	0	0	2	3	40	2	0	0
<b>4</b>	0	0	0	0	0	0	0	0
<b>5</b>	0	0	0	0	0	0	0	0
<b>6</b>	0	0	0	0	0	0	0	0
<b>7</b>	0	0	0	0	0	0	0	0
<b>8</b>	0	0	0	0	0	0	0	0

Tabla 3. Ejemplo de un histograma conjunto para el análisis de dos ventanas con 8 valores de grises diferentes posibles, de tamaño 10x10.

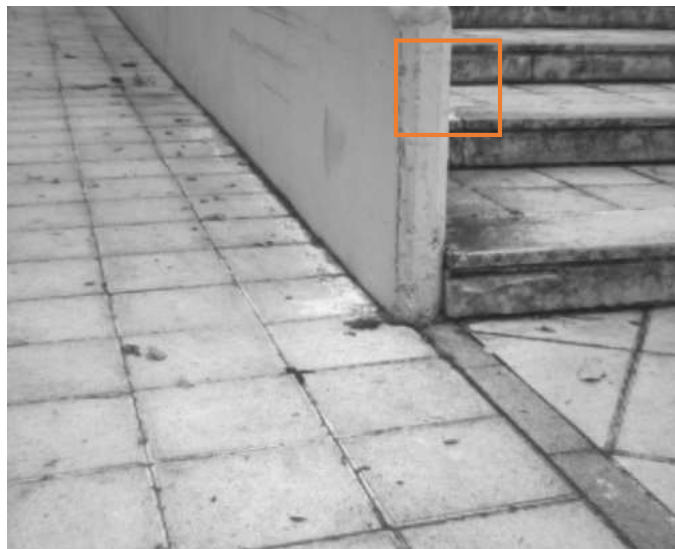
### 3.4. Cálculo de la información mutua

Para el cálculo de la información mutua usamos la ecuación (1) descrita dentro de la revisión de literatura, en conjunto con la ecuación (3) para

el cálculo de las probabilidades a usar, todo esto sobre el histograma conjunto obtenido por cada una de las dos ventanas analizadas, la ventana VS y la venta LWIR respectivamente.

Cabe recordar que el histograma conjunto se da por la suma de cada ocurrencia de un par ordenado, entre más alineado esté más se acercará al máximo de información mutua.

Al realizar el cálculo de la información mutua, buscando una ventana VS entre el rango establecido de ventanas LWIR, escogemos el valor máximo de información mutua entre todos los obtenidos, representando este valor a una de las ventanas LWIR analizadas, definiendo esta ventana como la homóloga de la ventana VS.



**Figura 3.2. Imagen VS con una ventana a buscar seleccionada.**



**Figura 3.3.** Imagen IR con el intervalo de búsqueda de la información mutua mínima.

En la Figura 3.2 podemos observar una imagen VS donde se encuentra seleccionada una ventana que procederemos a buscar en la imagen LWIR. En la Figura 3.3, esta búsqueda debe tener un rango definido, una de las ventanas entre la búsqueda debe ser escogida como su homóloga mediante la maximización de la información mutua. Al escoger esta ventana y su posición, podemos calcular el desplazamiento de la ventana LWIR escogida respecto a la posición de la venta VS. Este desplazamiento será la disparidad.

### **3.5. Disparidad**

La disparidad entre imágenes es una forma de estimar la profundidad de cada uno de los puntos en la escena de estas imágenes. En este

caso asumiremos que la escena es estática, es decir, los objetos visibles en la escena no cambian su posición en la misma ni sufren deformaciones.

La disparidad se refiere a la diferencia en la localización de un objeto en las correspondientes dos (izquierda y derecha) imágenes como se ve por la izquierda y derecha del ojo que se crea debido a la paralaje (separación horizontal de los ojos). En este caso el cerebro utiliza esta disparidad para calcular la mencionada información de profundidad de las imágenes bidimensionales [17].

En la Figura 3.4 tenemos la geometría de la configuración de las cámaras de par estéreo y su relación con el valor de la disparidad.

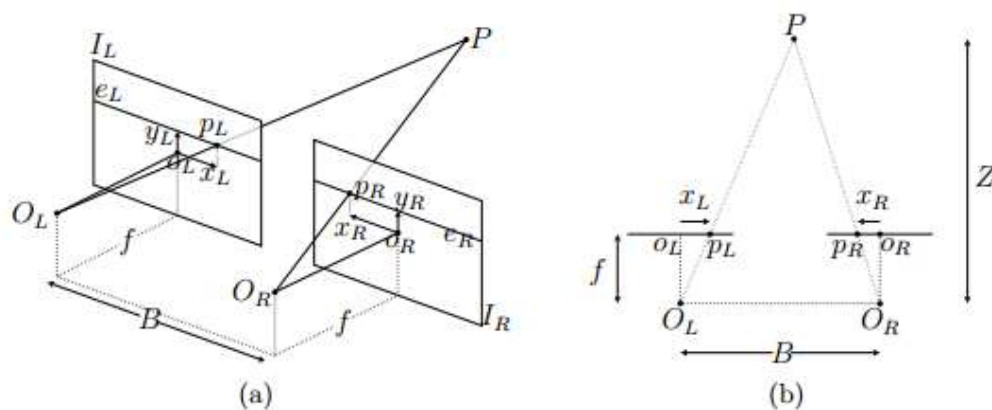
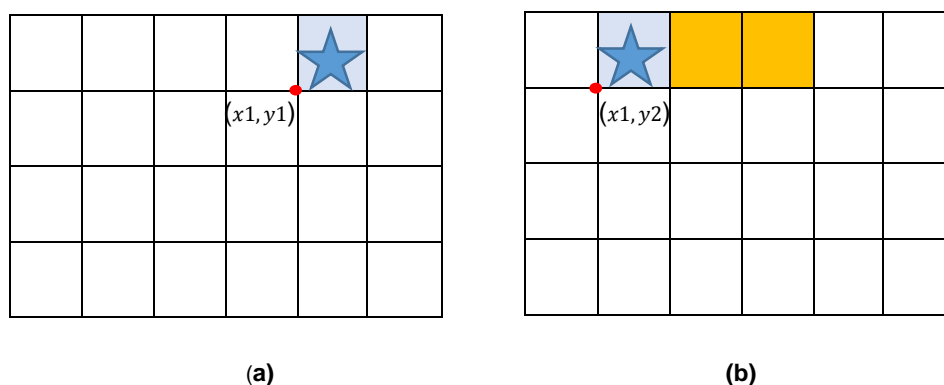


Figura 3.4. (a) Ejemplo de configuración de las cámaras del par estéreo. (b) Relación geométrica entre los parámetros del par estéreo para obtener la profundidad  $Z$  a partir de la disparidad  $d$ .

El cálculo de la disparidad se realiza a partir del desplazamiento obtenido al calcular la ventana homóloga de la imagen VS en la imagen LWIR.



**Figura 3.5. (a) Imagen VS con un punto a buscar, la estrella. Imagen LWIR mostrando el desplazamiento -el punto hallado por cuadros, cada cuadro representa una unidad.**

Mirando la Figura 3.5 tenemos que en (a) representa la imagen VS, la ventana que buscamos ubicada en  $(x1, y1)$ , la hallamos en (b), la imagen LWIR, en los puntos  $(x1, y2)$ , desplazado por  $y1 - y2$ , ese desplazamiento es la disparidad. Se obtiene sacando la diferencia entre su punto en la imagen VS y el hallado en la imagen LWIR.

Ahora esta disparidad se calcula para todos los pixeles que se toman de referencia por ventana, pudiendo ser negativa o positiva, obteniendo así una matriz que contiene todos estos valores por pixel.

Esta matriz representa un mapa de disparidad entre las dos imágenes, VS y LWIR respectivamente.

El mapa de disparidad será mostrado como una imagen en escala de grises que contiene los valores de disparidad para cada pixel.



**Figura 3.6. Ejemplo. En (a) se encuentra un ejemplo de dos imágenes VS del mismo plano. (b) es el resultado del mapa de disparidad obtenido**

En la Figura 3.6 podemos observar un ejemplo de mapa de disparidad obtenido con dos imágenes VS, uno de los objetivos del proyecto es el análisis entre una imagen VS y una LWIR.

La conversión de los desplazamientos obtenidos anteriormente, las disparidades, a un valor de gris, dependerá del rango establecido para la búsqueda del mínimo de información mutua escogido mediante la interfaz.

Pudiendo obtener el mapa de disparidad se procede a la evaluación de las imágenes, analizando los resultados según diferentes valores escogidos como son el rango de búsqueda, el tamaño de la ventana y a cuántos valores de gris las imágenes se discretizaron previamente.

### **3.6. Implementación de la interfaz**

En la Figura 3.7 podemos observar la interfaz gráfica elaborada para el presente proyecto con la cual el usuario puede cargar las imágenes que desea analizar, en este caso el par VS y LWIR de la misma escena, luego de eso se pueden establecer los factores de análisis ya establecidos anteriormente para finalmente hacer el cálculo de información mutua a lo largo de las imágenes para obtener el mapa de disparidad.

Esta interfaz fue desarrollada utilizando el framework gráfico Qt 5.2.1. Las librerías usadas dentro del análisis de las imágenes forman parte de OpenCV 2.4, otros algoritmos y funcionalidades utilizados dentro del análisis fueron implementados en C++ a lo largo del desarrollo.

La descripción de cada uno de los diferentes elementos presentes en la interfaz se detalla a continuación.





Figura 3.7. Interfaz del sistema elaborado.

- (1) y (2): Sirven para seleccionar y cargar las imágenes VS y LWIR que se analizarán.
- (3) y (4): Se muestran las imágenes seleccionadas.
- (5): Selección de la cantidad de niveles de gris a las cuales se van a discretizar las imágenes VS (3) y LWIR (4).
- (6) y (7): Selección del intervalo de búsqueda y el tamaño de ventana que se va a buscar respectivamente. Este intervalo será

para adelante y atrás del eje de referencia a buscar en la medida que la imagen permita.

- (8): Adecuar imagen prepara ciertos valores con los cuales se van a trabajar.
- (9): Realiza el cálculo del máximo de información mutua para cada ventana de la imagen VS en la imagen LWIR, obteniendo el desplazamiento por cada una y guardando todos estos valores
- (10): Elabora el mapa de disparidad y lo presenta en pantalla según el valor escogido en (6)
- (11): Guarda el mapa de disparidad obtenido
- (12): Menú con información sobre el proyecto.

# CAPÍTULO 4

## EVALUACIÓN Y ANÁLISIS DE RESULTADOS

En este capítulo se presenta una evaluación de los resultados obtenidos con la técnica propuesta y un análisis de los mismos.

### 4.1. Resultados Experimentales

Se presentan los resultados experimentales obtenidos usando la base de datos de imágenes del Centro de Visión por Computador<sup>1</sup>. Cada caso de los resultados experimentales está agrupado según los factores de análisis descritos anteriormente.

#### 4.1.1. Grupo 1

El primer grupo de análisis parte con los siguientes valores:

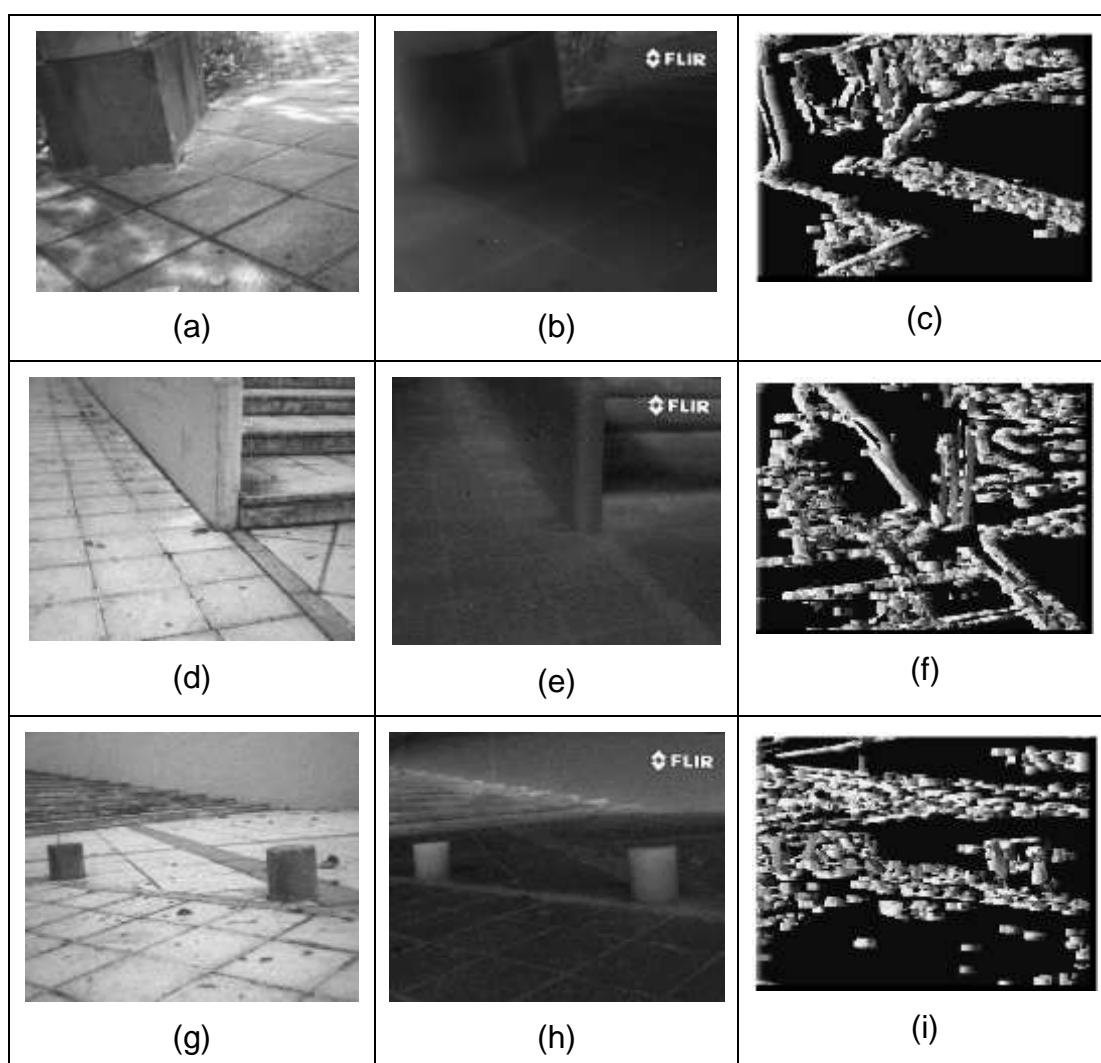
Niveles de Gris	8
-----------------	---

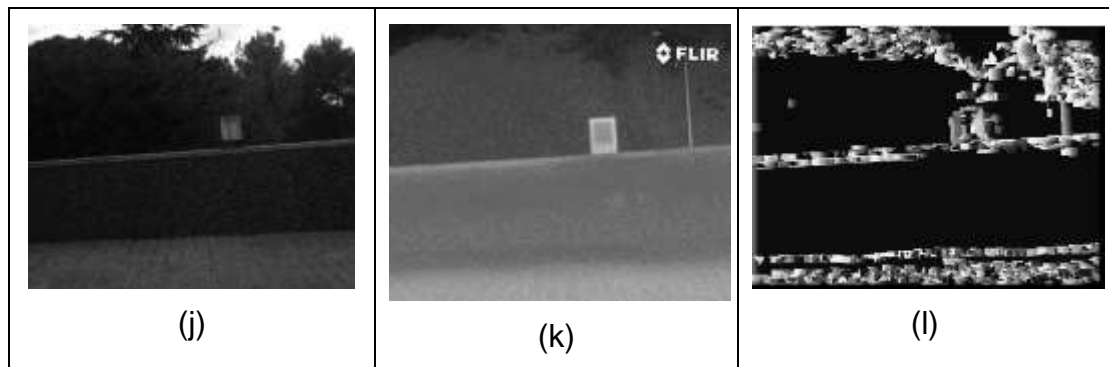
---

<sup>1</sup> <http://www.cvc.uab.es/adas/projects/simeve/>

Intervalo de Búsqueda	10
Tamaño de Ventana	11

Tabla 4. Valores escogidos para el primer grupo de análisis

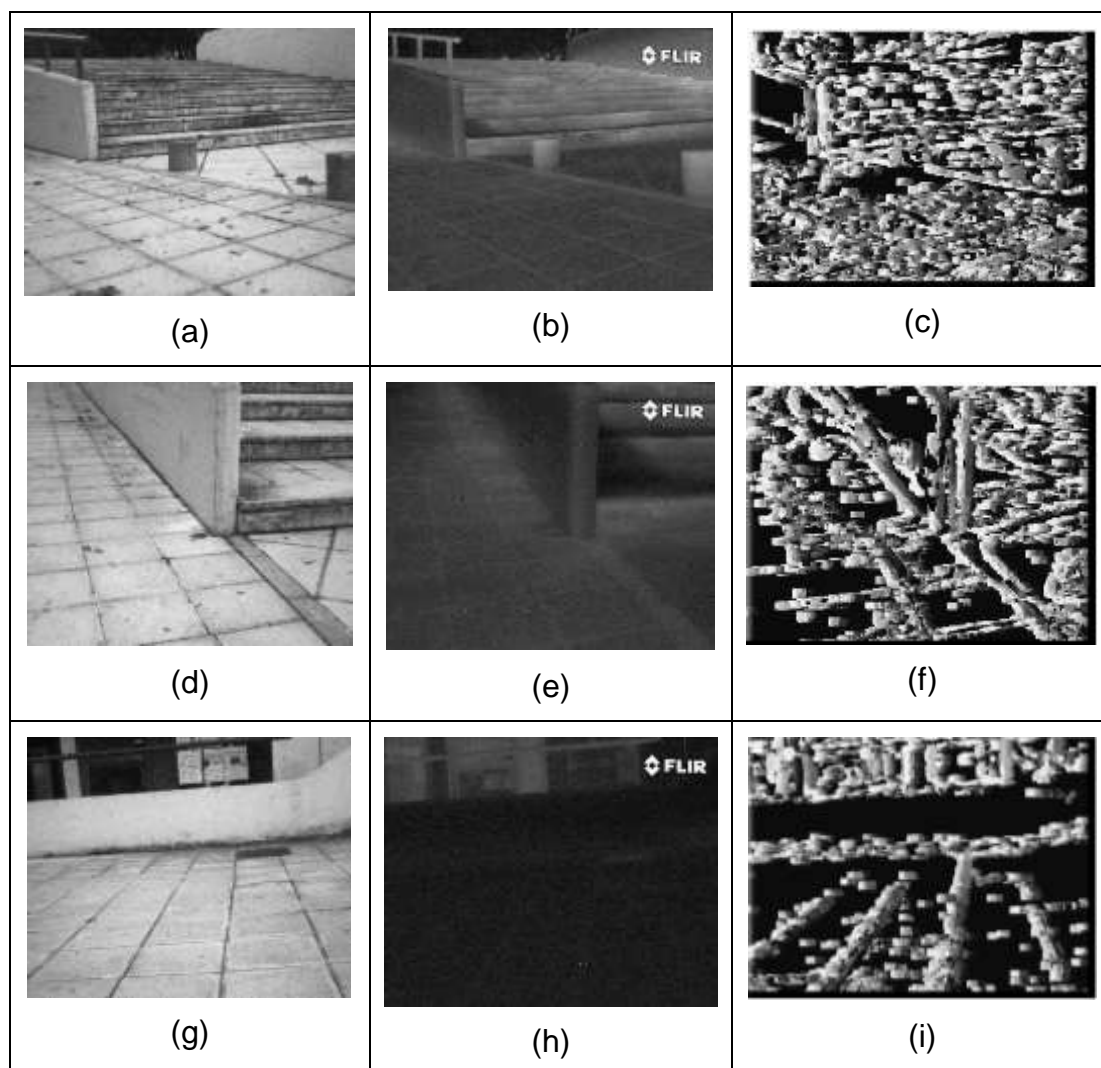




**Figura 4.1. Parte de los resultados, a la izquierda imágenes VS, columna de en medio imagen LWIR, a la derecha resultado obtenido.**

Podemos observar en el mapa de disparidad semejanza con los contornos de los objetos dentro de la escena. Al existir mucha variación en los tonos de gris en la imagen VS o la imagen LWIR, menos suave será el mapa de disparidad obtenido. Si los tonos se mantienen cercanos en partes de las imágenes, es más fácil identificar puntos sin embargo no se puede realizar la distinción de profundidades de la escena.

Las siguientes imágenes fueron analizadas aumentando la cantidad de niveles de gris de 8 a 16.



**Figura 4.2. Resultados obtenidos cambiando los niveles de gris a 16.**

En la Figura 4.2 se puede observar los resultados de aumentar este valor, el mapa de disparidad obtenido sufre una intensificación del ruido, se vuelve menos suave y más difícil identificar objetos y planos.

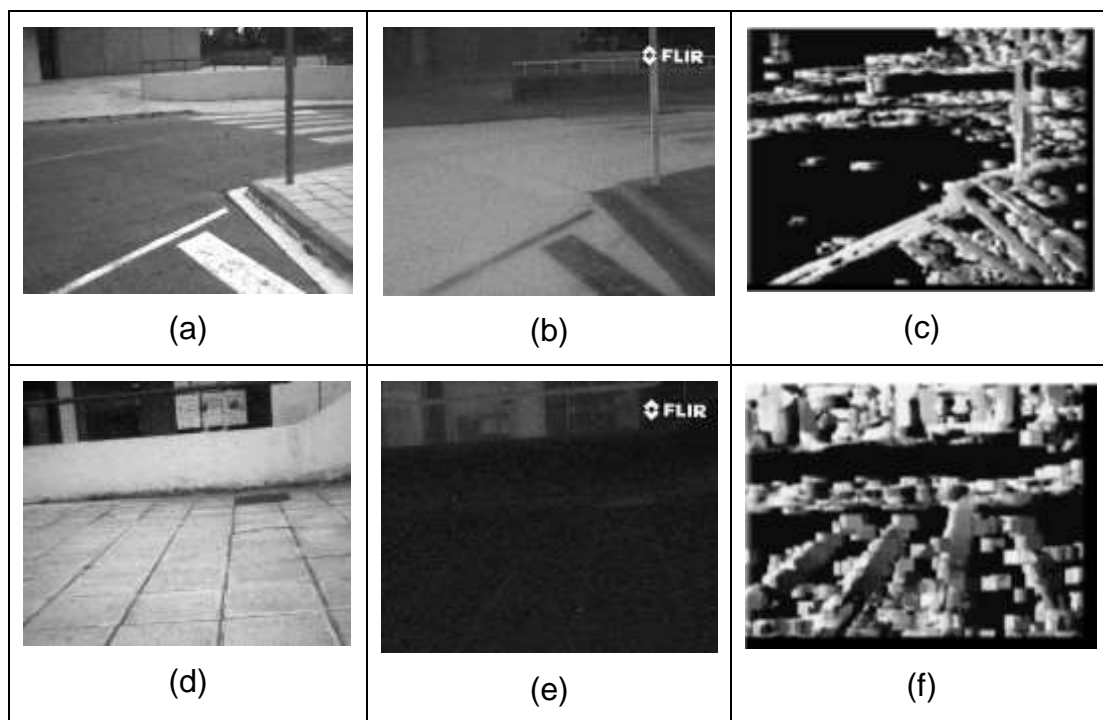


Figura 4.3. Resultados obtenidos cambiando el tamaño de la ventana a 21x21 píxeles.

En la Figura 4.3 se observan los cambios al aumentar el tamaño de la ventana de 11x11 a 21x21. Se mantienen los efectos de aumentar la cantidad de niveles de gris, es decir, las imágenes siguieron aumentando de ruido conforme más niveles de gris se usan pero los contornos de las formas se volvieron más pronunciados. Siguen sin poderse distinguir qué tan distantes están los objetos entre sí.

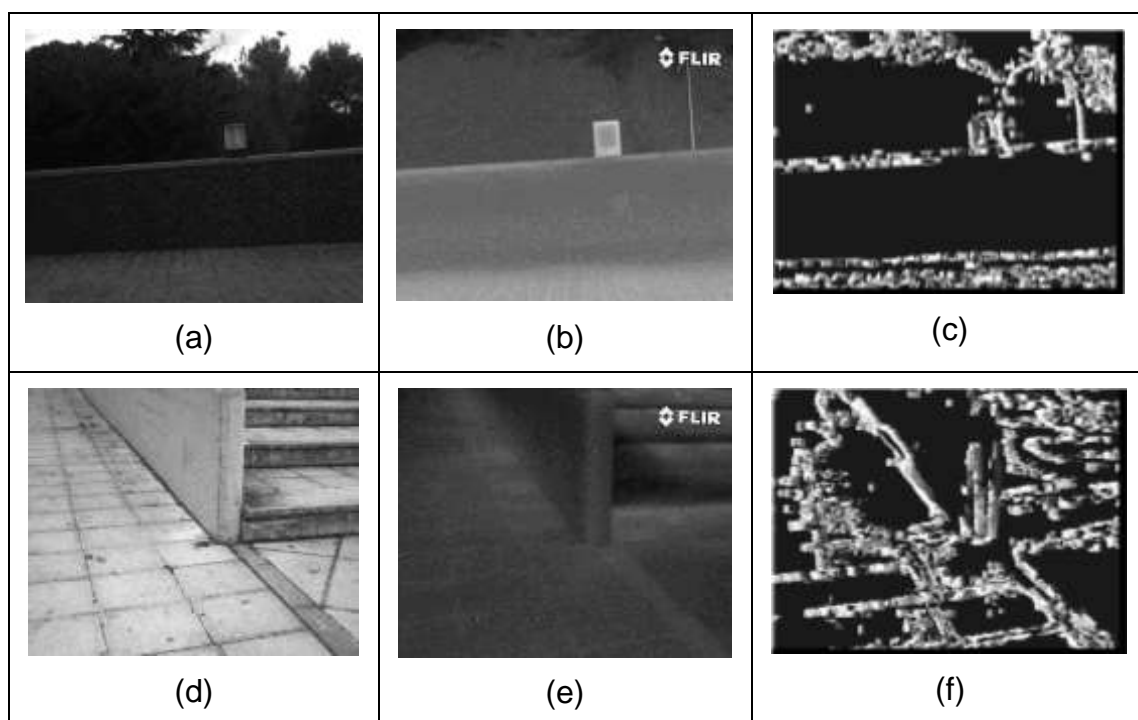
#### 4.1.2. Grupo 2

El segundo grupo de análisis parte con los siguientes valores:

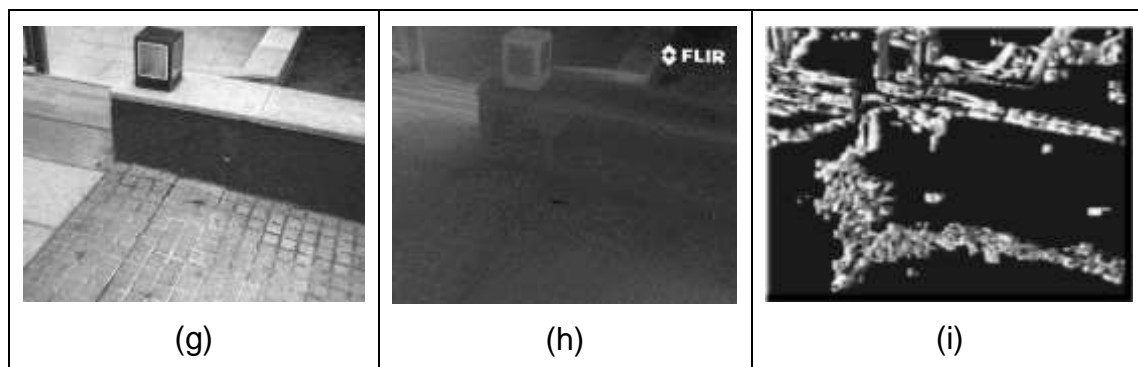
Niveles de Gris	8
Intervalo de Búsqueda	5
Tamaño de Ventana	11

Tabla5. Valores escogidos para el segundo grupo de análisis.

En esta ocasión se ha cambiado el intervalo de búsqueda para la información mutua.



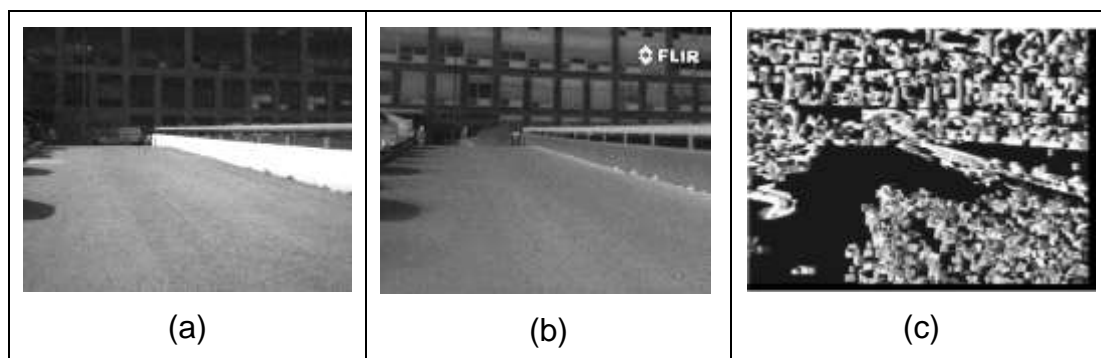




**Figura 4.4.** Algunos de los resultados obtenidos al trabajar con los datos de la Tabla 5 para el segundo grupo de análisis.

Como se observa en la Figura 4.4 se obtuvieron mapas con contornos definidos, al existir un intervalo menor de búsqueda la cantidad de niveles de gris posibles es menor.

En la Figura 4.5, en los resultados, sigue perdiéndose definición al aumentar los valores de gris a los cuales se han discretizado las imágenes previamente al análisis. En este caso se trabajó con 16 niveles de gris, intervalo de 5 y ventanas 11x11 píxeles.



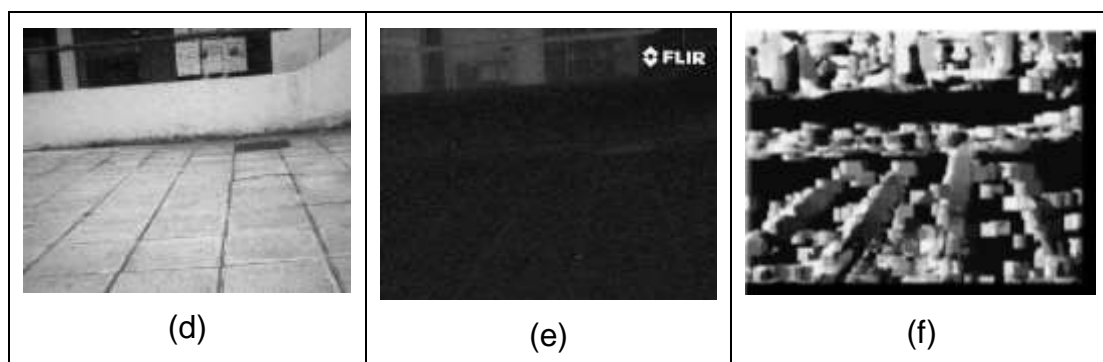


Figura 4.5. Resultados de trabajar con 16 niveles de gris en el 2do grupo

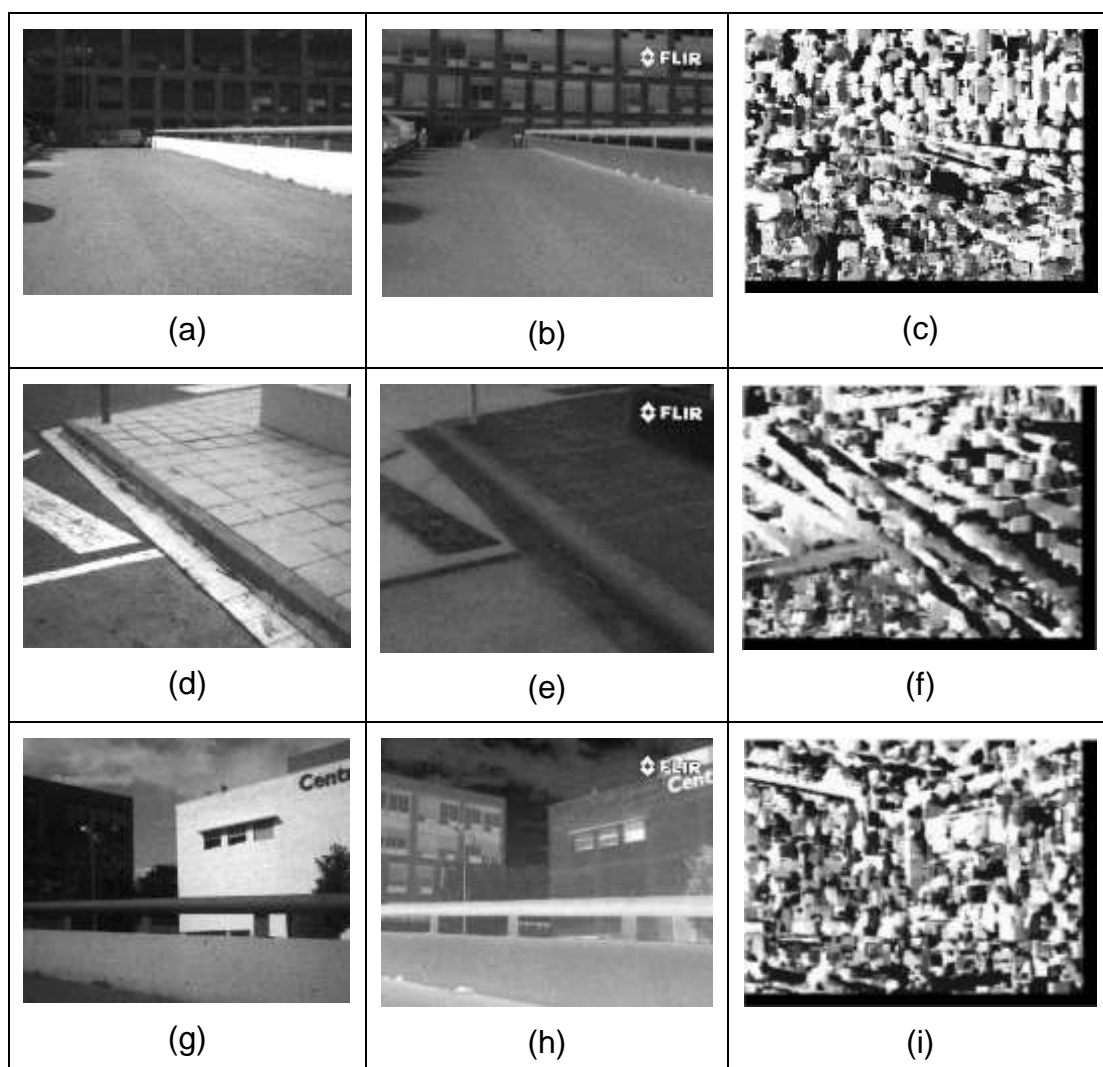


Figura 4.6. Resultados de trabajar con 16 niveles y aumentando el tamaño de ventanas a 21x21 píxeles.

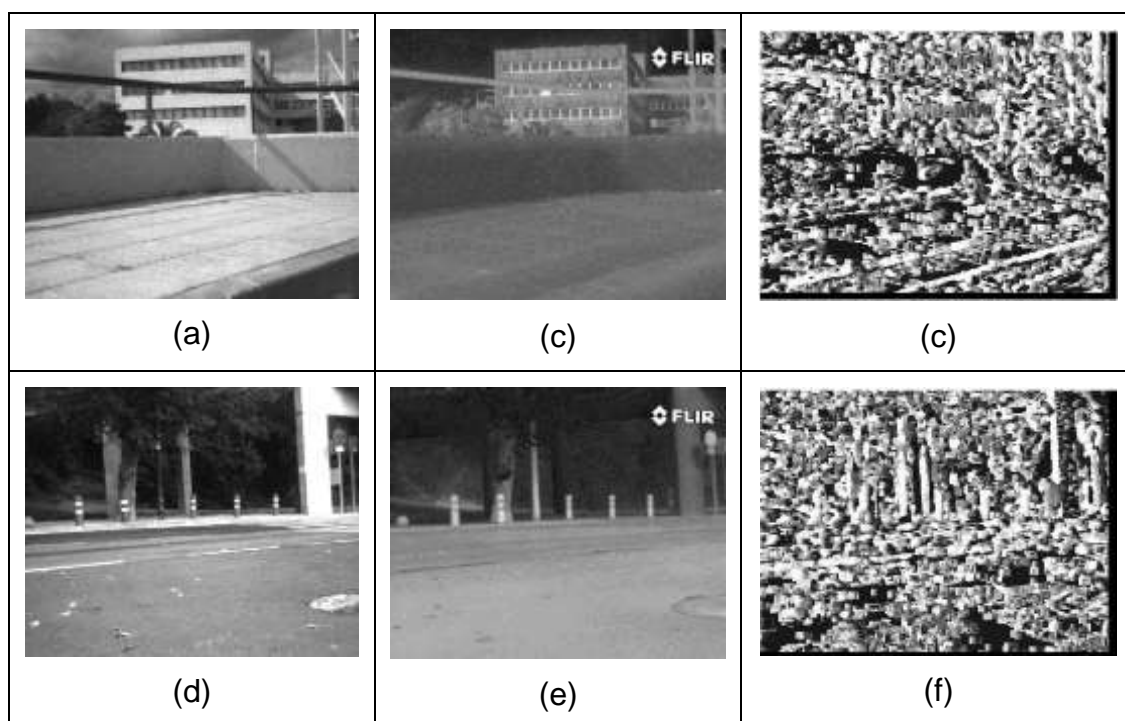
Al aumentar el tamaño de la ventana de búsqueda, los mapas de disparidad se volvieron ilegibles en su mayoría (ver Figura 4.6).

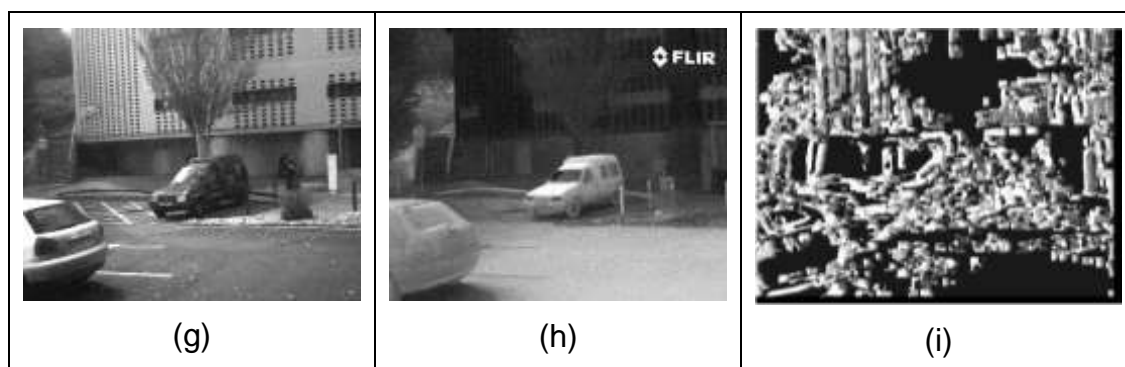
### 4.1.3. Grupo 3

El tercer grupo de análisis parte con los siguientes valores:

Niveles de Gris	24
Intervalo de Búsqueda	5
Tamaño de Ventana	11

Tabla 6. Valores escogidos para el tercer grupo de análisis





**Figura 4.7. Algunos de los resultados obtenidos al trabajar con los datos de la Tabla 5 para el segundo grupo de análisis.**

En esta parte de los resultados se hallan muchos más bordes en comparación a los otros mapas.

## **4.2. Análisis de resultados**

Según los resultados obtenidos sobre el análisis empírico de las imágenes VS y LWIR se considera lo siguiente:

- Los mapas de disparidad obtenidos sí me permiten identificar objetos bajo ciertas condiciones en las que se haga el análisis de información mutua, siendo estas el nivel de gris escogido para la discretización y el tamaño de la ventana de búsqueda en la maximización de la información mutua.
- Estos mapas de disparidad sin embargo no me permiten identificar planos y distancias en la escena con la base de imágenes usada.

- Entre menor es la cantidad de niveles de gris a las cuales han sido discretizadas las imágenes VS y LWIR, menos ruido se obtiene en los mapas de disparidad obtenidos, esto es, se identifican menos bordes.
- Aumentar el intervalo de búsqueda en la maximización de la información mutua no altera en gran proporción los mapas de disparidad obtenidos.
- Aumentar el tamaño de la ventana de búsqueda nos da un mapa de disparidad que presenta los bordes de los objetos en escena más pronunciados y con blur.
- El negro en los mapas de disparidad significa que en la maximización de la información mutua, se obtuvo el mismo resultado para cada ventana analizada del rango de búsqueda, quedándose siempre con la primera coincidencia; la primera ventana de izquierda a derecha analizada.

## CONCLUSIONES Y RECOMENDACIONES

En este capítulo se resumen las conclusiones obtenidas del trabajo realizado sobre las imágenes de espectro visible e infrarrojo provistas por el Centro de Visión por Computación, Barcelona, España. En la segunda parte de este capítulo se presentan las recomendaciones que a nuestro entender serían las líneas de trabajo futuro que permitirían mejorar los resultados aquí presentados.

### Conclusiones

Habiendo realizado el análisis de los resultados obtenidos se puede concluir lo siguiente:

1. La información mutua y su uso como función de correspondencia entre imágenes del espectro visible y del espectro infrarrojo lejano sí sirve para generar un mapa de disparidad que bajo las condiciones planteadas en el apartado 4.2, es útil para identificar objetos en la escena, sin embargo, este mapa no es útil para

detectar empíricamente la distancia entre los objetos de la mencionada escena.

2. De los factores usados en este proyecto para el cálculo de la información mutua, descritos en las secciones 3.1 y 3.2, se obtiene que usando los mínimos establecidos; Valores de gris para la discretización igual a 8, Tamaño de Intervalo igual a 5 y Tamaño de ventana igual a 11, se obtienen mejores resultados.
3. El cálculo de la información mutua entre imágenes puede ser una actividad de cómputo costosa en cuanto a tiempo, por lo cual su uso en aplicaciones de tiempo real puede resultar poco efectivo.
4. Al analizar las imágenes se debe considerar que estamos tratando tanto las imágenes del espectro visible como el espectro infrarrojo, en escala de grises, a pesar de que las imágenes del espectro infrarrojo están en otra frecuencia y su registro de información se convierte en otro valor de gris diferente al espectro visible. Esto produce errores de cálculo de la ventana homóloga real, y por consecuente valores que afectan el mapa de disparidad.
5. La tecnología usada permite la elaboración de proyectos de aplicación real, los lenguajes C, C++ y C# son de rápido cómputo y OpenCV sí posee librerías capaces de ayudar en esta labor.

## Recomendaciones

Este proyecto sirve como base para abrir futuras líneas de trabajo e investigación tanto en la ESPOL como en otros centros de investigación. A continuación se listan posibles mejoras que se proponen realizar en el sistema o recomendaciones para alguien que desea continuar la investigación.

1. Trabajar con otro grupo posible de valores de gris para discretizar las imágenes, diferente al especificado en la sección 3.1.
2. Trabajar con otro grupo posible de tamaño de ventana y de intervalo de búsqueda diferente a los valores establecidos en la sección 3.2.
3. Usar una base distinta de imágenes VS y LWIR que contengan elementos en escena en diferentes planos claramente distinguibles, sin perspectivas.
4. Seleccionar otra tecnología para elaborar el sistema de cómputo, como puede ser Matlab, pudiendo así comparar tiempos en los cálculos hechos.
5. Usar otros métodos de correspondencia de imágenes y realizar una comparación entre estos, estableciendo factores que determinen la eficiencia y eficacia de cada uno.



# Anexo A

## Anexo A: Calibración de las cámaras para obtener la base de imágenes

Este apéndice tiene como fin mostrar cómo se trabajó con las cámaras para obtener la base de imágenes proveída para su uso dentro del proceso de análisis. La calibración de una cámara representa un paso importante ya que se consiguen estimar los parámetros intrínsecos y extrínsecos de la misma los cuales son necesarios para realizar la reconstrucción 3D del entorno y situar la cámara en el mismo.

### A.1. Parámetros intrínsecos

Los parámetros intrínsecos describen el funcionamiento de una cámara, son:

- Distancia focal: Es la distancia entre el sensor y la lente de la cámara. En el modelo pin-hole ilustrado en la Figura A.1 la distancia focal está dada por la distancia entre el centro óptico y el plano imagen.

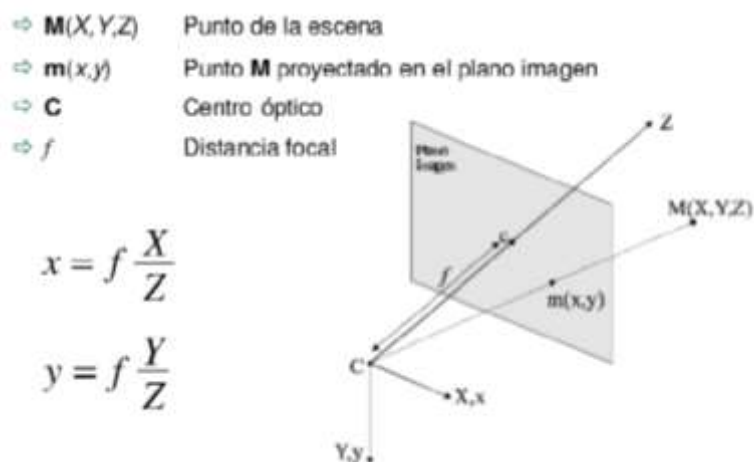
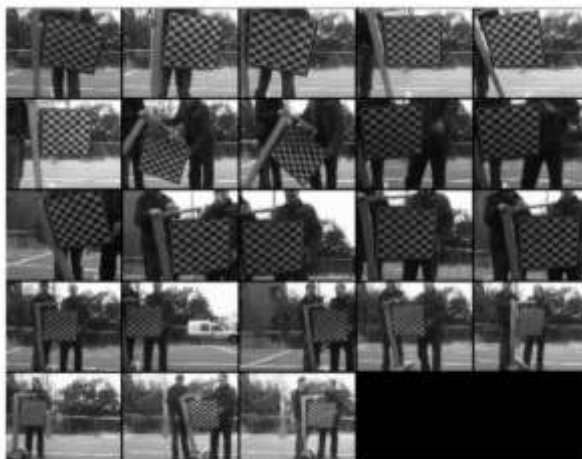


Figura A.1. Modelo de Calibración pin-hole.

- Punto principal: Son las coordenadas en píxeles del punto central de la imagen, es decir, el punto que define el eje óptico en el plano de imagen.
- Sesgo: Representa la diferencia en las dimensiones de los píxeles del sensor de imagen en caso de no ser cuadrados, para lo cual se utilizan factores de escala horizontal y vertical  $k_u$  y  $k_v$  respectivamente.
- Distorsión: Representa la distorsión radial de la lente.

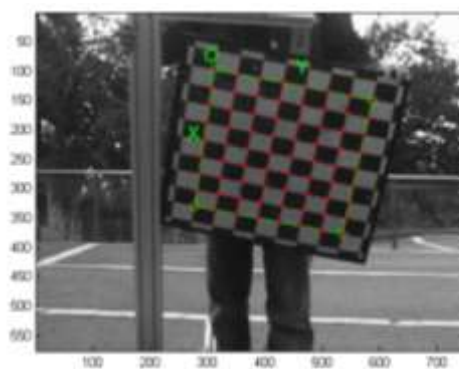
Para la calibración de los parámetros intrínsecos se utilizó el Toolbox de Matlab “Camera Calibration Toolbox”, se cargaron las imágenes tomadas de un patrón de calibración compuesto por rectángulos de 0.047m de alto por 0.061m de ancho, dicho patrón fue capturado por la

cámara a diferentes distancias y desde diferentes ángulos. El set de imágenes tomadas se puede ver en la Figura A.2.



**Figura A.2. Imágenes del patrón tomadas con la cámara VS.**

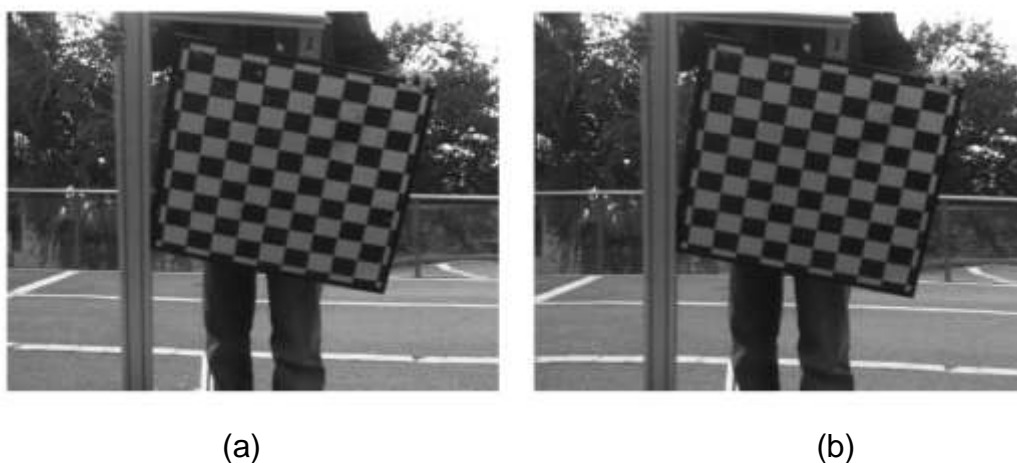
Posteriormente para cada imagen se indicaron las cuatro esquinas correspondientes al patrón de calibración utilizado, a partir de las cuales el Toolbox obtiene las posiciones de todos los rectángulos del patrón como puede verse en la Figura A.3.



**Figura A.3. Rectángulos detectados en el primer set de imágenes VS**

Realizado el paso anterior para la totalidad de las imágenes VS, se obtienen los parámetros intrínsecos antes mencionados, permitiendo la corrección (rectificación) de las imágenes tomadas.

En la Figura A.4, (a) y (b) se puede apreciar la corrección realizada sobre la primera imagen del set VS.



**Figura A.4. (a) Imagen original VS. (b) Imagen rectificada VS.**

A partir de la aplicación de los parámetros intrínsecos obtenidos en el set completo de imágenes tomadas del patrón de calibración, el Toolbox permite obtener la posición relativa de cada imagen capturada con respecto a la posición de la cámara.

En la Figura A.5 podemos ver la posición en el espacio de cada imagen VS tomada del patrón de calibración.

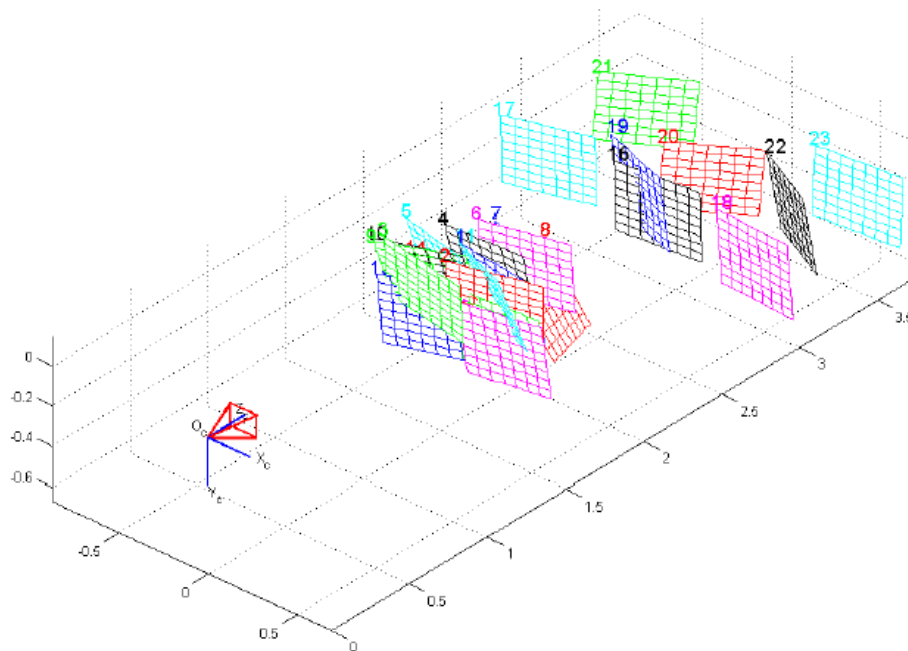


Figura A.5. (a) Imagen original VS. (b) Imagen rectificada VS.

Fuente: [18]

El procedimiento anteriormente descrito fue repetido para la cámara LWIR, obteniéndose los valores de sus parámetros de calibración intrínseca.

## A.2. Parámetros extrínsecos

Los parámetros extrínsecos definen la relación espacial entre la cámara y el mundo real. En el caso de visión estéreo la calibración de parámetros extrínsecos del par de cámaras permite la obtención de los valores de rotación y traslación de una cámara con respecto a la otra.

Siendo estos valores indispensables para la reconstrucción de escenas estéreo a partir de un par de imágenes capturado por las cámaras.

Dichos parámetros son:

- Rotación: relacionan la orientación una cámara respecto de la otra.
- Traslación: relacionan la posición una cámara respecto de la otra.

Para la calibración de los parámetros extrínsecos se utilizó el Toolbox de Matlab “Stereo Camera Calibration Toolbox” , se cargaron los parámetros intrínsecos de la cámara VS ubicada a la izquierda en el par estéreo, y los parámetros intrínsecos de la cámara LWIR ubicada a la derecha, dichos parámetros fueron obtenidos mediante el procedimiento indicado en la sección anterior.

Esto permitió obtener la rotación y traslación de una cámara con respecto a la otra.

En la Figura A.6 podemos ver la posición en el espacio de cada imagen tomada del patrón de calibración con respecto al par estéreo.

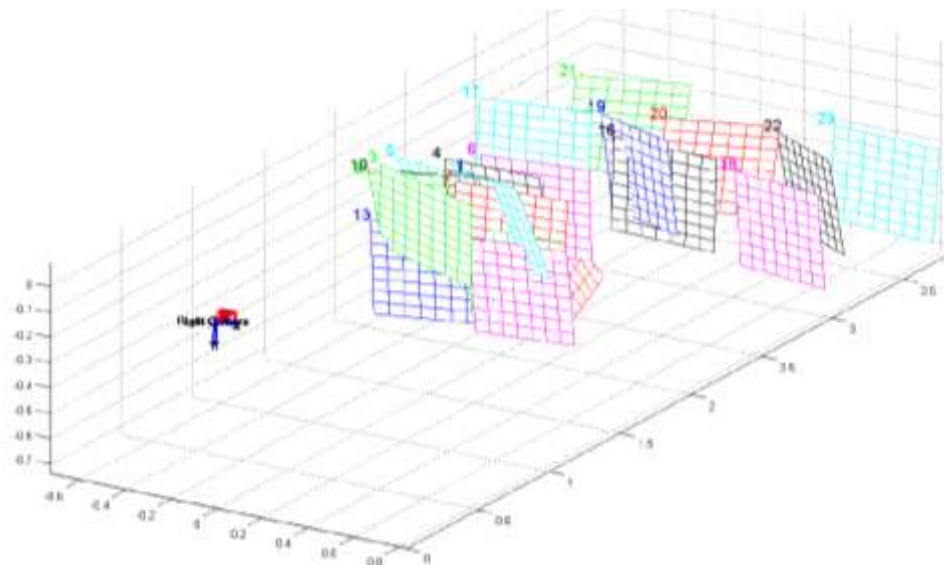


Figura A.6. Posición en el espacio de cada imagen capturada con respecto al par estéreo.

Para poder apreciar de manera más clara la rotación y traslación de una cámara respecto la otra, se efectuó un acercamiento al par estéreo anterior como se puede ver en la Figura A.7.

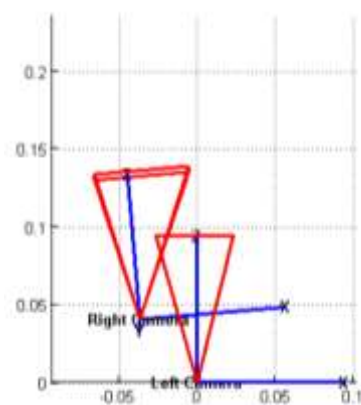


Figura A.7. Posición en el espacio de cada imagen capturada con respecto al par estéreo.

Se puede apreciar de manera clara la posición de la cámara derecha respecto de la izquierda que ha sido tomada como referencia.



## Anexo B

### Anexo B: Tablas de conversión de valores de gris a otra cantidad escogida

En esta sección se explica la distribución de valores de gris hechas para cada cantidad que el usuario escogió mediante la interfaz gráfica. El grupo indica el nuevo valor que tomará el pixel para su posterior análisis de encontrarse en el rango de valores mostrado.

Discretizar a 16 niveles:

<b>Rango</b>	<b>Grupo</b>	<b>Rango</b>	<b>Grupo</b>
<b>0-15</b>	Grupo 1	<b>128-143</b>	Grupo 9
<b>16-31</b>	Grupo 2	<b>144-159</b>	Grupo 10
<b>32-47</b>	Grupo 3	<b>160-175</b>	Grupo 11
<b>48-63</b>	Grupo 4	<b>176-191</b>	Grupo 12

<b>64-79</b>	Grupo 5	<b>192-207</b>	Grupo 13
<b>80-95</b>	Grupo 6	<b>208-223</b>	Grupo 14
<b>96-111</b>	Grupo 7	<b>224-239</b>	Grupo 15
<b>112-127</b>	Grupo 8	<b>240-255</b>	Grupo 16

Tabla B.1. Distribución de grupos de valores de gris para discretizar a 16 valores de gris.

Discretizar a 24 niveles:

<b>Rango</b>	<b>Grupo</b>	<b>Rango</b>	<b>Grupo</b>
<b>0-9</b>	Grupo 1	<b>120-129</b>	Grupo 13
<b>10-19</b>	Grupo 2	<b>130-139</b>	Grupo 14
<b>20-29</b>	Grupo 3	<b>140-149</b>	Grupo 15
<b>30-39</b>	Grupo 4	<b>150-159</b>	Grupo 16
<b>40-49</b>	Grupo 5	<b>160-169</b>	Grupo 17

<b>50-59</b>	Grupo 6	<b>170-179</b>	Grupo 18
<b>60-69</b>	Grupo 7	<b>180-189</b>	Grupo 19
<b>70-79</b>	Grupo 8	<b>190-199</b>	Grupo 20
<b>80-89</b>	Grupo 9	<b>200-209</b>	Grupo 21
<b>90-99</b>	Grupo 10	<b>210-219</b>	Grupo 22
<b>100-109</b>	Grupo 11	<b>220-229</b>	Grupo 23
<b>110-119</b>	Grupo 12	<b>230-255</b>	Grupo 24

**Tabla B.2. Distribución de grupos de valores de gris para discretizar a 24 valores de gris.**

Discretizar a 32 niveles:

<b>Rango</b>	<b>Grupo</b>	<b>Rango</b>	<b>Grupo</b>
<b>0-7</b>	Grupo 1	<b>128-135</b>	Grupo 17
<b>8-15</b>	Grupo 2	<b>136-143</b>	Grupo 18

<b>16-23</b>	Grupo 3	<b>144-151</b>	Grupo 19
<b>24-31</b>	Grupo 4	<b>152-159</b>	Grupo 20
<b>32-39</b>	Grupo 5	<b>160-167</b>	Grupo 21
<b>40-47</b>	Grupo 6	<b>168-175</b>	Grupo 22
<b>48-55</b>	Grupo 7	<b>176-183</b>	Grupo 23
<b>56-63</b>	Grupo 8	<b>184-191</b>	Grupo 24
<b>64-71</b>	Grupo 9	<b>192-199</b>	Grupo 25
<b>72-79</b>	Grupo 10	<b>200-207</b>	Grupo 26
<b>80-87</b>	Grupo 11	<b>208-215</b>	Grupo 27
<b>88-95</b>	Grupo 12	<b>216-223</b>	Grupo 28
<b>96-103</b>	Grupo 13	<b>224-231</b>	Grupo 29
<b>104-111</b>	Grupo 14	<b>232-239</b>	Grupo 30

<b>112-119</b>	Grupo 15	<b>240-247</b>	Grupo 31
<b>120-127</b>	Grupo 16	<b>248-255</b>	Grupo 32

**Tabla B.3. Distribución de grupos de valores de gris para discretizar a 32 valores de gris.**

## BIBLIOGRAFÍA

[1] Hurtado F., Atlas de matemáticas (Primera edición), 1997, Idea Books, S.A., página 8.

[2] Grupo Alava Ingenieros, Artículo técnico de imágenes multiespectrales de Avala Ingenieros, 2010, página 1.

[3] Starr C., Biology: Concepts and Applications, Thomson Brooks/Cole, 2005, páginas 203-205.

[4] Lira J., Tratamiento digital de imágenes multiespectrales, Capítulo I, Aspectos generales (Segunda edición), 2010, página 28.

[5] Iwaszczuk D, Urban Remote Sensing Event, 2011, páginas 25-28

[6] Lira J., Tratamiento digital de imágenes multiespectrales, Capítulo I, Aspectos generales,(Segunda edición), 2010, página 2.

[7] Correspondencia de imágenes,  
<http://ocw.usal.es/eduCommons/enseanzas-tecnicas/procesamiento-avanzado-de-imagenes-digitales/contenidos/Tema3.pdf>, fecha de consulta Julio 2014.

[8] Lira J., Tratamiento digital de imágenes multiespectrales, Capítulo VIII, Reconocimiento de Patrones (Segunda edición), 2010, página 3736.

[9] Bender A., An Introduction to Mathematical Modelling, Capítulo I, página 5.

[10] Correspondencia de imágenes,  
<http://ocw.usal.es/eduCommons/enseanzas-tecnicas/procesamiento-avanzado-de-imagenes-digitales/contenidos/Tema3.pdf>, fecha de consulta Agosto 2014.

[11] González R., Digital Image Processing, Capítulo XII, Object Recognition, 2007, página 884.

[12] Gray, R., Entropy and Information Theory, 1990, Springer-Verlag.

[13] Nyquist H., Certain Factors Affecting Telegraph Speed, 1924.

[14] Cover, T.M. y Thomas, J.A, Elements of information theory, Capítulo 2, 1991, página 45.

[15] Kullback, S.; Leibler, Annals of Mathematical Statistics: On information and sufficiency, 1951, páginas 79–86.

[16] Pass, G. y Zabig, R., Comparing Images Using Joint Histograms, Capítulo 2 Image Summaries, página 3.

[17] Kushalm, A. y Ponce, J., Modeling 3D Objects from Stereo Views and Recognizing Them in Photographs, páginas 563–574. 2006.

[18] Pistarelli M., Detección y Correspondencia de características en imágenes multiespectrales, 2013, páginas 65–66.