

# Utilización de información mutua como función de correspondencia en el dominio multispectral

Ayllyn Franco Rodríguez<sup>(1)</sup> Angel D. Sappa<sup>(1,2)</sup>

<sup>(1)</sup>Facultad de Ingeniería Eléctrica y Computación (FIEC)  
Escuela Superior Politécnica del Litoral (ESPOL)  
Campus Gustavo Galindo, Km 30.5 vía Perimetral  
Apartado 09-01-5863. Guayaquil-Ecuador  
ayllynfranco@gmail.com

<sup>(2)</sup>Centro de Visión por Computador (CVC)  
Universidad Autónoma de Barcelona (UAB)  
Edificio O, Campus UAB  
08193 Bellaterra, Barcelona, España  
asappa@cvc.uab.es

## Resumen

*El propósito de este proyecto es realizar un análisis de la información mutua como función de correspondencia entre imágenes del espectro visible y del espectro infrarrojo, ampliando el campo de trabajo y 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, y de recibir datos que afecten el cálculo de la correspondencia multispectral.*

**Palabras Claves:** *Información mutua; Correspondencia multispectral.*

## Abstract

*The purpose of this project is to make an analysis of the mutual information as a function of correspondence between images of the visible spectrum and the infrared spectrum, expanding the field of work and determining which factors affect this calculation and in which ways, through the development of an interface and a computer system capable of receiving and working with images of the infrared spectrum and visible spectrum, and that also receives the data which affects the calculation of the multispectral correspondence.*

**Keywords:** *Mutual Information; Multispectral Matching.*

## 1. Introducción

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 también 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 multispectral.

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.

## 2. Descripción 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.

## 3. Solución propuesta

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 para finalmente analizar los resultados obtenidos.

## 4. 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 [1].

### 4.1. Reconocimiento de Patrones

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 [2].

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 [3].

## 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 [16] [17].

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.

### 5.1. Métodos de Correspondencia

Un método de correspondencia debe calcular la semejanza entre dos entidades posiblemente homólogas (ejemplo [4], [5], [6]). El nombre del método se añade tras el nombre de la entidad de correspondencia.

Existen distintos algoritmos para describir puntos característicos, desde el más sencillo que podría ser utilizar un valor booleano que describa si el punto característico es una esquina de un objeto de la imagen, hasta usar un vector de 128 dimensiones que caracterice el punto de interés usando información de sus vecinos.

Método de correspondencia	Índice de correspondencia	Entidades de correspondencia
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.

Estos algoritmos suelen constar de partes específicas que se muestran en la Tabla 1 y que se definen como:

- Índice de semejanza o de correspondencia: es una medida cuantitativa de lo buena que es la correspondencia de las entidades entre sí [4]. Generalmente el grado de semejanza se determina mediante una función de costo.
- Estrategia o método de correspondencia: se utiliza para referirse al esquema de conjunto que tiene la solución del problema de la correspondencia de imágenes [4].

## 6. 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. [7]. 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.

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

$$P_Y(y) = \sum_x P_{XY}(x, y) \quad (3)$$

Y  $EP$  es el valor esperado de la distribución  $P$ .

La primera definición dada en la ecuación (1), se encuentra íntimamente relacionada a la Divergencia Kullback-Leiber entre dos distribuciones [8], en este caso la medida de información mutua  $(X, Y)$  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, se puede obtener 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 [13].

Como una medida de similitud, información mutua ha disfrutado de una gran cantidad de éxito, particularmente en el dominio de las imágenes médicas [14].

## 7. Desarrollo de la solución

Basándonos en la literatura propuesta [15] 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 1.** Diagrama de flujo del método propuesto

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, pudiendo así calcular una matriz de disparidades mejor conocido como un mapa de disparidad.

### 7.1. Discretizar 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.

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  $256/x$ .

### 7.2. Selección de Valores pertinentes

En esta etapa escogemos algunos valores que afectan 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.

- 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.

### 7.3. 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 [10].

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.

	1	2	3	4	5	6	7	8
1	0	0	0	0	0	0	0	0
2	1	2	0	23	10	0	0	0
3	0	0	2	3	40	2	0	0
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0

**Figura 2.** Histograma conjunto de dos ventanas con 8 valores de grises diferentes posibles, de tamaño 10x10.

### 7.4. Cálculo de información mutua

Para el cálculo de la información mutua usamos la ecuación (1) descrita en la sección 6, en conjunto con la ecuación (2) y (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 ventana LWIR respectivamente.

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.



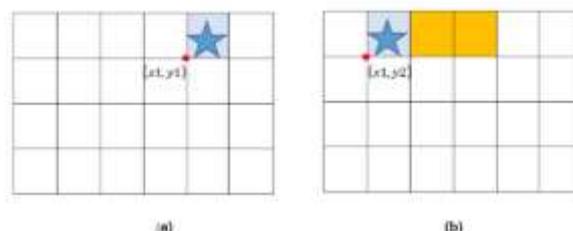
**Figura 3.** Imagen VS con una ventana a buscar. Imagen IR con el intervalo de búsqueda de la información mutua.

Al escoger esta ventana y su posición, podemos calcular el desplazamiento de la ventana LWIR escogida respecto a la posición de la ventana VS. Este desplazamiento será la disparidad.

### 7.5. Disparidad

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 al paralaje (separación horizontal de los ojos). [11]

En este caso el cerebro utiliza esta disparidad para calcular la mencionada información de profundidad de las imágenes bidimensionales [12].

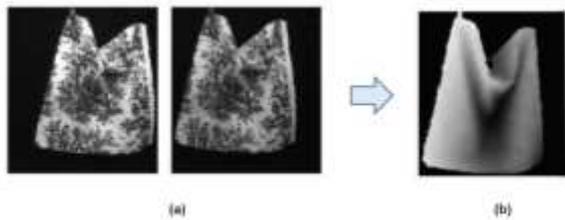


**Figura 4.** (a) Imagen VS con un punto a buscar, la estrella. (b) Imagen LWIR mostrando el desplazamiento -el punto hallado por cuadros, cada cuadro representa una unidad.

Mirando la Figura 4 tenemos que en (a) la imagen VS, la ventana que buscamos ubicada en  $(x_1, y_1)$ , la hallamos en (b), la imagen LWIR, en los puntos  $(x_1, y_2)$ , desplazado por  $y_1 - y_2$ , 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 píxeles que se toman de referencia por ventana, pudiendo ser negativa o positiva, obteniendo así una matriz que contiene todos estos valores por píxel. Esta

matriz representa un mapa de disparidad entre las dos imágenes, VS y LWIR respectivamente [8] [9].



**Figura 5.** (a) Ejemplo de dos imágenes VS del mismo plano. (b) es el resultado del mapa de disparidad obtenido

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.

## 8. Resultados experimentales

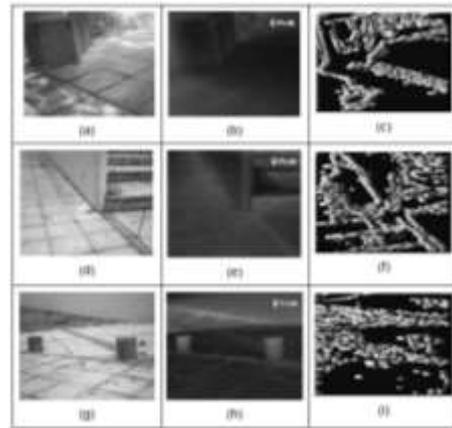
Seguidamente se presentan los resultados experimentales obtenidos de usar la técnica propuesta en la base de datos del Centro de Visión por Computador. En todos los casos se ha utilizado el mismo conjunto de parámetros seleccionados en el GUI.

### 8.1. Grupo 1

El primer grupo de análisis parte con los siguientes valores:

Niveles de Gris	8
Intervalo de búsqueda	10
Tamaño de ventana	11

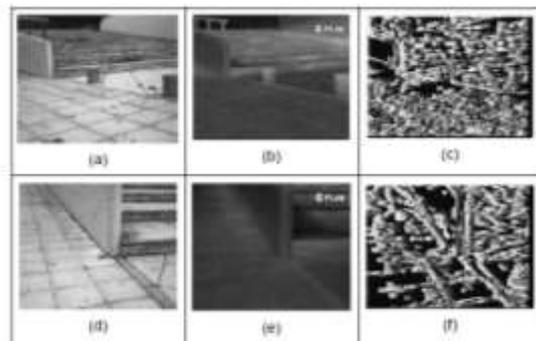
**Tabla 2.** Valores escogidos para el primer grupo de análisis



**Figura 6.** Parte de los resultados, a la izquierda imágenes VS, columna de en medio imagen LWIR, a la derecha resultado obtenido.

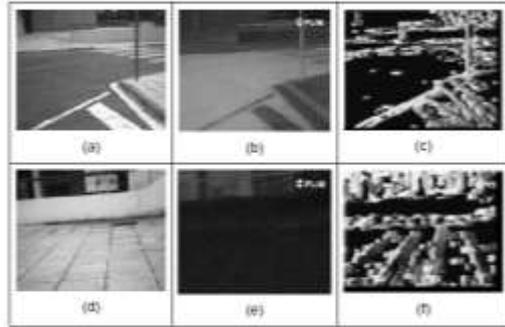
Se observa en la Figura 6, en el mapa de disparidad semejanza con los contornos de los objetos dentro de la escena. Al existir mucha variación en los tonos de gris en la imagen VS o la imagen LWIR, menos suave será el mapa de disparidad obtenido. Si los tonos se mantienen cercanos en partes de las imágenes, 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 7.** Resultados obtenidos cambiado los niveles de gris a 16.

En la Figura 7 se puede observar los resultados de aumentar este valor del tamaño de la ventana, el mapa de disparidad obtenido sufre una intensificación del ruido, se vuelve menos suave y más difícil identificar objetos y planos.



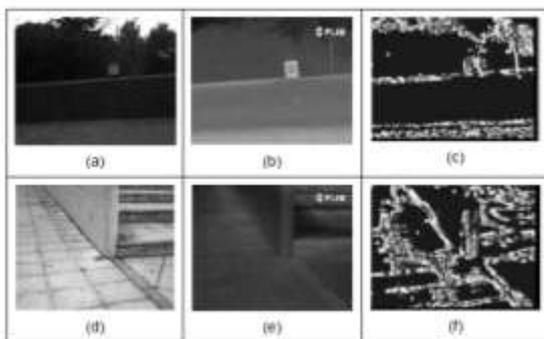
**Figura 8.** Resultados obtenidos cambiando el tamaño de la ventana a 21x21 píxeles.

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 pero los contornos de las formas se volvieron más pronunciados. Sigue sin poderse distinguir qué tan distantes están los objetos entre sí.

## 8.2. Grupo 2

Niveles de Gris	8
Intervalo de búsqueda	5
Tamaño de ventana	11

**Tabla 3.** Valores escogidos para el segundo grupo de análisis



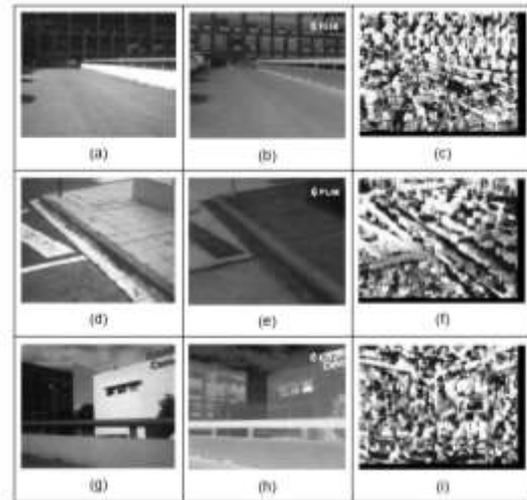
**Figura 9.** Algunos de los resultados obtenidos al trabajar con los datos de la Tabla 3 para el segundo grupo de análisis.

Como se observa en la Figura 9 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 10, 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 10.** Resultados de trabajar con 16 niveles de gris en el 2do grupo



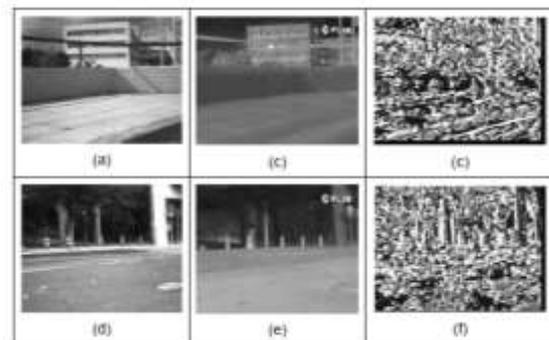
**Figura 11.** Resultados de trabajar con 16 niveles y aumentando el tamaño de ventanas a 21x21 píxeles.

Al aumentar el tamaño de la ventana de búsqueda, los mapas de disparidad se volvieron ilegibles en su mayoría.

## 8.3. Grupo 3

Niveles de Gris	24
Intervalo de búsqueda	5
Tamaño de ventana	11

**Tabla 4.** Valores escogidos para el tercer grupo de análisis



**Figura 12.** Resultados obtenidos al trabajar con los datos de la Tabla 4 para el segundo grupo de análisis.

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

#### 8.4. 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.

#### 9. Conclusiones

A partir del análisis empírico sobre el par de imágenes expuesto en el presente trabajo se llegó a las siguientes conclusiones:

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 ciertas condiciones, 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 7.1 y 7.2, 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 consiguiente 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.

#### 10. Recomendaciones

A continuación se presentan las recomendaciones que a nuestro entender serían las líneas de trabajo futuro que permitirían mejorar los resultados aquí presentados.

1. Trabajar con otro grupo posible de valores de gris para discretizar las imágenes, diferente a los presentados en los resultados experimentales.
2. Trabajar con otro grupo posible de tamaño de ventana y de intervalo de búsqueda diferente a los valores presentados.
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.

## 11. Referencias

- [1] Correspondencia de imágenes, <http://ocw.usal.es/eduCommons/enseñanzas-tecnicas/procesamiento-avanzado-de-imagenes-digitales/contenidos/Tema3.pdf>, fecha de consulta Julio 2014.
- [2] Iwaszczuk D, *Urban Remote Sensing Event*, 2011, pp. 25-28
- [3] Bender A., *An Introduction to Mathematical Modelling*, Capítulo I, pp. 5.
- [4] Correspondencia de imágenes, <http://ocw.usal.es/eduCommons/enseñanzas-tecnicas/procesamiento-avanzado-de-imagenes-digitales/contenidos/Tema3.pdf>, fecha de consulta Agosto 2014.
- [5] Restrepo-Spetch A., Sappa A. and Devy M., *Edge Registration Versus Triangular Mesh Registration, a Comparative Study*, Signal Processing: Image Communication, Vol. 20, No. 9-10, October-November 2005, pp. 853-868.
- [6] Rouhani M. and Sappa A., *Correspondence Free Registration through a Point-to-Model Distance Minimization*, IEEE Int. Conf. on Computer Vision, Barcelona, Spain, November 6-13, 2011, pp. 2150-2157
- [7] Roshini, VS., Dr. Revathy K., *Using mutual information and cross correlation as metrics for registration of images. Journal of Theoretical and Applied Information Technology*, 2008, pp. 476
- [8] Kullback, S.; Leibler, *Annals of Mathematical Statistics: On information and sufficiency*, 1951, pp. 79-86.
- [9] Pass, G. y Zabig, R., *Comparing Images Using Joint Histograms*, Capítulo 2 Image Summaries, pp. 3.
- [10] Yedida, J.S., Freeman, W.T., Weiss Y., *Understanding belief propagation and its generalizations*. Exploring Artificial Intelligence in the New Millennium, 2003, Capítulo 8, pp. 239-236
- [11] Barrera F., Lumbreras F. and Sappa A., *Multimodal Stereo Vision System: 3D Data Extraction and Algorithm Evaluation*, IEEE Journal of Selected Topics in Signal Processing, Vol. 6, No. 5, September 2012, pp. 437-446.
- [12] Kushalm, A. y Ponce, J., *Modeling 3D Objects from Stereo Views and Recognizing Them in Photographs*. 2006, pp. 563-574.
- [13] Viola, P., Wells III, W.M.: *Alignment by maximization of mutual information. International Journal of Computer Vision* 24 (1997) pp 137-154
- [14] Pluim, J., Maintz, J., Viergever, M.: *Mutual information based registration of medical images: a survey*. IEEE Transactions on Medical Imaging, 2003, pp. 986- 1004
- [15] Barrera F., Lumbreras F. and Sappa A., *Multimodal template matching based on gradient and mutual information using scale-space*, IEEE Int. Conference on Image Processing, September 2010, Hong Kong, pp. 2749-2752.
- [16] Aguilera C., Barrera F., Lumbreras F., Sappa A., and Toledo R., *Multispectral Image Feature Points*, Sensors, Vol 12, No. 9, September 2012, pp. 12661-12672.
- [17] Ricaurte P., Chilán C., Aguilera-Carrasco C., Vintimilla B. and Sappa A., *Feature Point Descriptors: Infrared and Visible Spectra*, Sensors, Vol. 14, No. 2, February 2014, pp. 3690-3701.