

Correspondencia multiespectral en el espacio de Hough

Gonzalo Andrade Moreira⁽¹⁾ Angel Sappa^(1,2)

⁽¹⁾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
gonzalo.andrade.moreira@gmail.com

⁽²⁾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 siguiente trabajo tiene como objetivo desarrollar una aplicación que pueda establecer o describir una correspondencia de características (rectas o círculos) entre imágenes del espectro visible y del espectro infrarrojo con el fin de poder ser utilizadas en aplicaciones futuras de estereovisión, por ejemplo. En el proyecto se pondrán en práctica algoritmos ya desarrollados en los cuales se evaluarán los resultados obtenidos con diferentes variables o procedimientos. Para esto, se desarrollará una interfaz gráfica donde se podrán verificar los procedimientos que se realizaron durante la evaluación.

Palabras Claves: Transformada de Hough; Correspondencia Multiespectral; Correspondencia de Características.

Abstract

The purpose of this study is to develop an application that could establish or describe feature matching (lines or circles) between images from the visible and infrared spectrum so they can be used in future stereo vision applications, for instance. The project will implement already known algorithms to be evaluated with different variables or procedures. In order to achieve this, a graphical user interface will be developed in which it will be possible to evaluate and verify the procedures during the evaluation.

Keywords: Hough Transform; Multispectral Matching; Feature Matching.

1. Introducción

En el área de visión por computador existen muchas aplicaciones que sirven para resolver problemas de interés, ya sea para mejorar procesos u obtener resultados a tareas específicas, entre estas aplicaciones podemos nombrar procedimientos de realzado, compresión, restauración, detección, correspondencia y reconocimiento de objetos.

La correspondencia entre dos objetos se define como una función f , que determina alguna relación binaria entre algún elemento del primer conjunto X con algún elemento del segundo conjunto Y . Podemos decir que esa función f definirá una correspondencia [1].

La correspondencia entre imágenes se basa en la búsqueda y detección de entidades [2] como puntos, líneas y áreas entre dos o más fotogramas [3], las cuales son luego utilizadas como referencias durante el proceso de correspondencia.

El propósito de este proyecto es realizar un análisis entre dos imágenes de diferentes espectros utilizando la transformada de Hough, el cual nos permite detectar correspondencias a través de figuras

tales como rectas o circunferencias. Para el estudio mencionado se utilizarán imágenes infrarrojas (LWIR) y del espectro visible (VS). Dichas imágenes fueron proporcionadas por un par estereoscópico multiespectral (VS-LWIR) del Centro de Visión por Computador de la Universidad Autónoma de Barcelona, España [4].

2. Descripción del problema

El principal problema es encontrar correspondencia entre los elementos de las imágenes multiespectrales, los cuales no son obtenibles mediante técnicas clásicas de correspondencia como SURF y SIFT o similares, debido a que ambas imágenes tienen información diferente.

Podemos observar en la imagen a continuación, que el ojo humano detecta correspondencia, sin embargo, ambas imágenes poseen información diferente, en la imagen de la izquierda (espectro visible) tenemos información fotométrica, mientras que en la imagen de la derecha (infrarroja lejano) tenemos información térmica.



Figura 1. A la izquierda: visible; a la derecha: infrarrojo lejano.

3. Solución propuesta

Debido a que el tipo de información de las imágenes es diferente, se deberán transformar ambas a un espacio común, en nuestro caso utilizando la transformada de Hough.

Se desarrollará una interfaz gráfica (GUI), en donde se pueda manipular con facilidad las variables de los algoritmos de detección de características y descriptores, así como también poder observar los resultados obtenidos.

También se desarrollarán los algoritmos de detección de características y correspondencia para finalmente analizar los resultados obtenidos.

4. Puntos de interés

La detección de puntos de interés es un área de la visión por computadora que tiene como objetivo detectar puntos de importancia o interés para un futuro procesamiento. Un punto de interés se encuentra bien definido cuando garantiza una cierta repetitividad en la imagen independientemente de pequeños cambios en la misma (fotométricos o geométricos).

Una esquina puede ser definida como una intersección entre dos bordes, donde un borde es un cambio brusco en el valor de la intensidad de la imagen. Normalmente una esquina es considerada un punto de interés vital ya que es utilizado en el área de visión por computador para obtener diferentes tipos de características de una imagen dada.

Los puntos de interés son utilizados para la detección, tracking y reconocimiento de objetos, generalmente en dos o tres dimensiones. Se puede clasificar los tipos de detección de puntos de interés en dos grupos:

- Basados en la derivada de la intensidad.
- Basados en el valor de la intensidad.

5. Descripción de puntos de interés

Una vez detectado un punto de interés debe ser descrito usando características específicas de la imagen, de manera tal que independientemente del punto de vista e iluminación se lo pueda identificar de manera unívoca.

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.

Una característica importante de los puntos de interés es que las posiciones relativas entre ellos y la escena original no deberían cambiar de una imagen a otra.

Por ejemplo, cuatro esquinas de una puerta son utilizadas como puntos de interés, estos cuatro puntos pueden ser encontrados sin importar la posición de la puerta, sin embargo, el reconocimiento podrá fallar si la puerta se encuentra abierta o cerrada. Los puntos de interés localizados en objetos flexibles podrían no funcionar si ocurre algún cambio en su geometría interna durante el procesamiento.

En la práctica, para reducir errores en el reconocimiento de objetos se describen muchos puntos característicos de cara a minimizar los errores causados por las variaciones locales. Seguidamente se presentan las características principales de los descriptores más utilizados en la literatura.

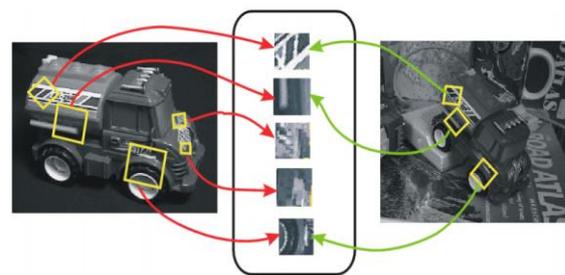


Figura 2. Descriptores invariantes a las transformaciones: traslación, rotación, escalamiento, brillo.

6. Descripción de puntos de interés en imágenes multispectrales VS-NIR, VS-LWIR

El análisis multispectral ha sido objeto de atención durante los últimos años en el área de procesamiento de imágenes y videos [3].

La información que proveen las imágenes de diferentes bandas espectrales puede ser directa o suplementaria a la información proporcionada por el espectro visible, lo que puede ayudar a resolver problemas más eficientemente. El procesamiento de imágenes de bandas espectrales fuera del espectro visible requiere el desarrollo de nuevas herramientas o adaptación de los métodos ya existentes (p. ej. [5], [6], [7]).

El análisis multiespectral ha sido ampliamente estudiado en el área de sensores remotos como lo son los satélites, sin embargo, gracias los avances en la tecnología, ahora las imágenes multiespectrales se pueden utilizar en aplicaciones de cámaras de seguridad o asistencia a la conducción de vehículos. Los puntos característicos son la base de diferentes problemas de visión por computadora y generalmente son estudiados en tres niveles: detección, descripción y correspondencia.

Uno de los enfoques más sobresalientes es el algoritmo SIFT [8]. Durante la última década, se han realizado varias modificaciones al algoritmo original con el objetivo de mejorar las debilidades de SIFT. En el caso en particular en las que las imágenes correspondan a distintas bandas del espectro, descriptores clásicos como SIFT y SURF requieren ser adaptados [9].

La dificultad de hallar correspondencias entre puntos característicos de imágenes VS-LWIR se encuentra en la relación no lineal que existe entre las intensidades de sus píxeles. En las imágenes LWIR la intensidad de brillo se encuentra relacionada a las variaciones de temperatura de los objetos, mientras que en VS el color del objeto proviene del reflejo de luz. Esto produce una falta de correlación en los resultados entre sus respectivos gradientes.

Otra característica de las imágenes LWIR es que poseen pérdida de detalle y textura [10]. También hay que tomar en cuenta que las imágenes LWIR poseen otro tipo de información que no se encuentra presente en la visible, por ejemplo, cuando no hay suficiente iluminación en una escena, como en la oscuridad de la noche. A continuación se exponen diferentes descriptores para la búsqueda de correspondencias entre imágenes de distintas bandas del espectro.

7. Desarrollo de la solución

Basados en la literatura mencionada, se propone un nuevo método para la búsqueda de correspondencias para líneas y círculos entre pares multiespectrales VS-

LWIR, a continuación se presenta el diagrama de flujo del método propuesto.

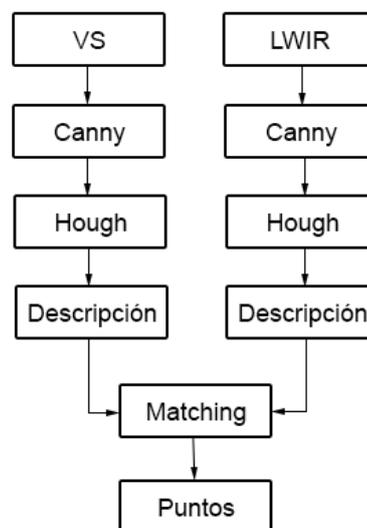


Figura 3. Diagrama de flujo del método propuesto

El método consiste en analizar las imágenes VS-LWIR por separado, utilizando el método de Canny para detección de bordes [11], con el fin de suavizar las imágenes. Debido a que una de las imágenes es de espectro diferente, no se podrá utilizar métodos clásicos a niveles de puntos; para esto utilizamos la transformada de Hough con el objetivo de llevar ambas imágenes a un dominio común donde ambas están representadas. Finalmente se realizará el matching para encontrar correspondencia de líneas y círculos en el par de imágenes.

7.1. Detección de contornos

El detector de contornos Canny es uno de los métodos más utilizados en la actualidad, detectando los bordes de una manera robusta [12]. La detección de contornos de Canny es un proceso de cuatro pasos:

- Suavizado de la imagen
- Cálculo de la magnitud y ángulo del gradiente
- Supresión de no-máximos
- Aplicación de umbral con histéresis

7.2. Transformada de Hough

La transformada de Hough es una técnica para la extracción de características, utilizada en el análisis de imágenes, visión por computadora y procesamiento digital de imágenes. El algoritmo fue presentado por Paul Hough en 1962 [13] para la detección de líneas

en una imagen, pero después fue extendida para la detección de formas arbitrarias, como círculos o elipses. La transformada de Hough está definida para una función $A(x,y)$ como se indica a continuación:

$$H(\theta, \rho) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(x,y) \delta(\rho - x \cos \theta - y \sin \theta) dx dy$$

Donde A es una imagen binaria, cada pixel en (x,y) es transformado en una onda senoidal $\rho = x \cos \theta - y \sin \theta$.

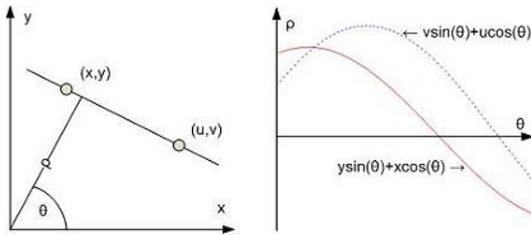


Figura 4. Transformada de Hough

Cada punto del plano $x - y$ es representado por una curva sinusoidal en el plano $\rho - \theta$. Los puntos colineales se mapean en curvas que tienen un punto de intersección común. Por lo tanto las líneas en el plano $x - y$ pueden ser detectadas contando el número de curvas que se intersectan en cada punto del plano $\rho - \theta$. Esto se logra con un arreglo acumulador.

La transformada de Hough no solo se encuentra restringida para la detección de líneas rectas, también puede ser utilizada para detectar otro tipo de curvas en la imagen. Por ejemplo para encontrar círculos, utilizamos la ecuación:

$$(x - a)^2 + (y - b)^2 = r^2$$

Donde el círculo se define con estos tres parámetros: la x e y son las coordenadas de su centro y r su radio. Para encontrar círculos utilizando la transformada de Hough, se necesita un arreglo acumulador de tres dimensiones.

Las líneas y círculos son detectados por la interpretación del acumulador cuando todos los puntos hayan sido transformados. Un ejemplo de este proceso se muestra en la Figura 5. La manera más básica de detectar líneas o círculos es estableciendo un umbral para el acumulador, e interpretar todos los valores por encima de este umbral.

Este umbral deberá tener un valor adecuado para una correcta detección, el cual será variable dependiendo del tamaño y contenido de la imagen. Por lo tanto dos imágenes de igual tamaño pero con diferente composición morfológica y niveles de ruido necesitaran distintos valores de umbral. También hay que tomar en cuenta que los valores máximos globales

no son detectados, ya que tener un umbral fijo genera una serie de máximos locales que implicaría la detección de líneas y círculos erróneos.

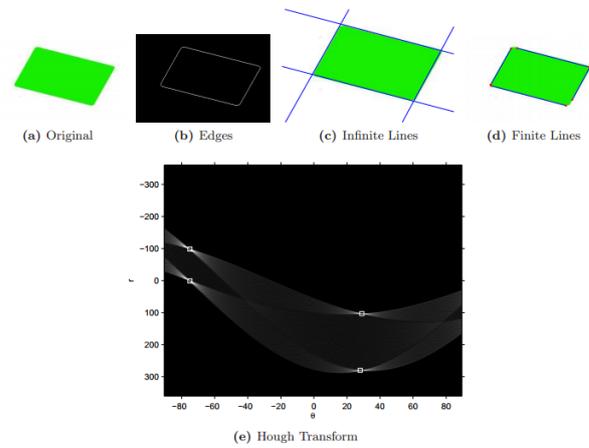


Figura 5. Detección de líneas utilizando la Transformada de Hough.

7.3. Descripción de características

En el siguiente apartado se detallarán los resultados obtenidos por los detectores y descriptores mencionados anteriormente. El análisis se realizó en varios pares de imágenes VS-LWIR de dos tipos de sets diferentes, imágenes que contienen líneas y círculos respectivamente (ver ejemplo en Figura 6).



Figura 6. El set de prueba. Imagen superior: visible. Imagen inferior: infrarrojo.

Para la detección de bordes Canny se utilizó la función $CV:Canny$ presente en las librerías de OpenCV, el cual utiliza dos umbrales; un umbral inferior σ_{low} y uno superior σ_{high} , donde σ_{high} se sugiere que sea un valor 3 veces mayor que el de σ_{low} (siguiendo la recomendación de Canny). Por

esto la interfaz tendrá por defecto $\sigma_{low}=50$, con la opción de asignarle un valor entre 0 y 100, mientras que σ_{high} será un valor constante de 150 (3 veces el valor de σ_{low}).



Figura 7. Canny de imagen VS.

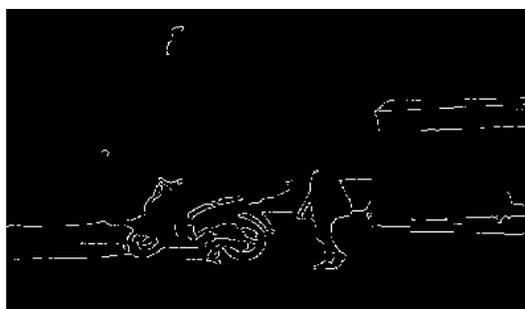


Figura 8. Canny de imagen LWIR.

En la Figura 8 se puede apreciar que en la imagen LWIR se pierde mucha más información en comparación de la imagen VS, debido al aspecto suavizado que esta posee. Por consecuencia de esto, muchos bordes no son considerados como tales. Un acumulador es utilizado en el espacio de Hough para determinar las áreas donde haya una mayor cantidad de intersecciones. Cada vez que un punto es transformado, el contenedor del acumulador se incrementa por todas las líneas que puedan pasar por ese punto. La resolución del acumulador determina la precisión con la que las líneas o círculos puedan ser detectadas.

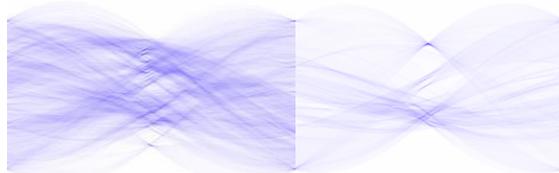


Figura 9. Resultado del cálculo de la transformada de Hough, correspondiente a las Figuras 7 y 8 para

detección de líneas en la imagen VS (izquierda) y LWIR (derecha).

Posteriormente las líneas y círculos son detectados en el par de imágenes como se muestra en la Figura 10. Hay que tomar en cuenta que para obtener una detección robusta se deberá manipular los valores de los umbrales, tanto en el algoritmo Canny como de la transformada de Hough.



Figura 10. Detección de líneas y círculos en el par de imágenes.

7.4. Matching

La búsqueda de correspondencias se realiza comparando cada punto del acumulador en ambos espacios de Hough (VS y LWIR) hasta encontrar puntos de intersección iguales. Si existe coincidencia, solo se mostrará las líneas o círculos que compartan las mismas características. Para realizar un matching robusto se necesitará modificar los valores de los diferentes umbrales.



Figura 11. Correspondencia de líneas y círculos en el par de imágenes.

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. Algunas ilustraciones de estos resultados se presentan en la Figura 12.



Figura 12. Resultados experimentales obtenidos.

8.1. Propuestas de modificaciones

A partir del análisis realizado en las imágenes de prueba, alguna de las cuales fueron mostradas en la sección anterior, se identificaron los siguientes problemas:

- Reducido número de correspondencias.
- Detección de líneas y círculos no presentes en el par de imágenes (falsos negativos).

Para atenuarlas, se proponen las siguientes correcciones:

8.1.1. Rectificación manual de imágenes

Las imágenes proporcionadas por el par estereoscópico multiespectral (VS-LWIR) son imágenes de una misma toma, sin embargo no son exactamente iguales ya que fueron tomadas por cámaras que tenían una pequeña separación entre sus ejes ópticos. Esto significa que el par de imágenes de prueba, a pesar de que fue rectificado utilizando los parámetros obtenidos al inicio del proceso de adquisición, podrá tener un pequeño desfase de píxeles como se puede observar en la Figura 13. Para resolver este problema, se sobreponen las imágenes

con menor opacidad y se cuadran hasta ser lo más precisos posibles. Este proceso manual se realiza solo a los efectos de obtener pares de imágenes perfectamente rectificadas que será usada como referencia para evaluar las implementaciones realizadas.

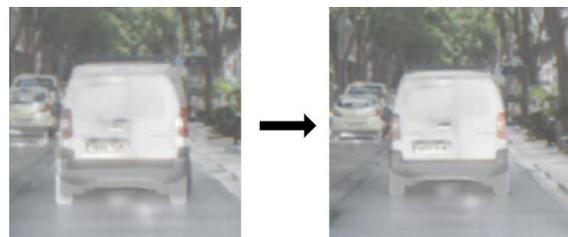


Figura 13. Alineación manual de las imágenes.

8.1.2. Modificación del umbral de Canny para imágenes LWIR

El detector Canny es comúnmente utilizado en imágenes de espectro visibles dando muy buenos resultados, sin embargo, debido a la naturaleza de la imagen LWIR (el suavizado, por ejemplo) mucha información se pierde utilizando los valores de umbrales recomendados por defecto en el algoritmo Canny. Para resolver este problema, se recomienda utilizar valores de umbrales diferentes en las imágenes LWIR. En la Figura 14. se puede observar la diferencia entre detectores de Canny con diferentes valores de umbral máximos σ_{high} , se puede observar que con un menor σ_{high} aparecerán líneas que antes no eran visibles.



Figura 14. Modificación de umbral en la imagen LWIR.

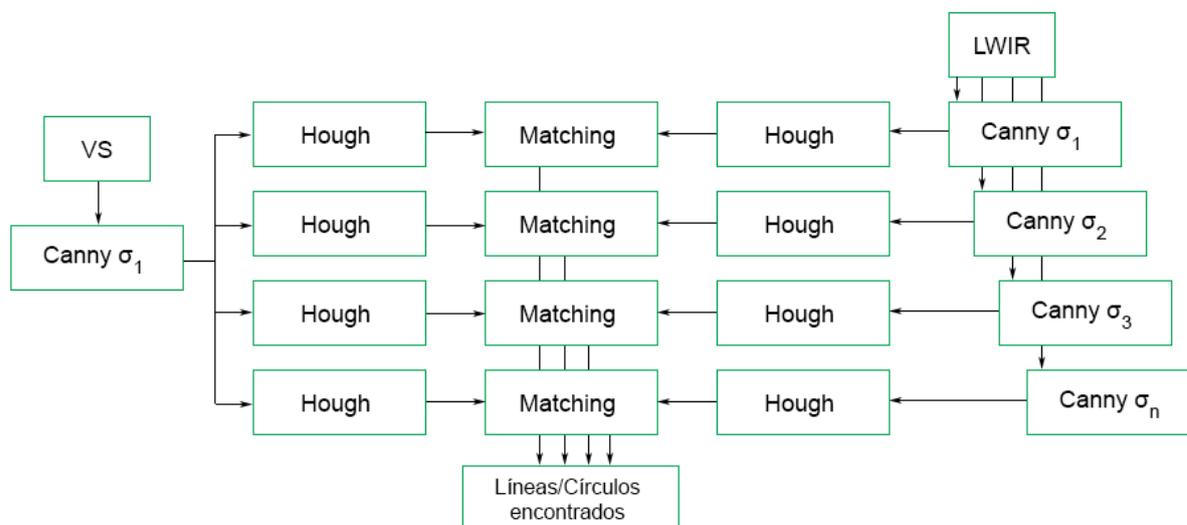


Figura 15. Diagrama de flujo del método propuesto

Es muy importante que la detección de bordes sea óptima en ambas imágenes ya que impacta directamente sobre la transformada de Hough y las correspondencias encontradas. Mientras más robusta sea la detección de bordes, existirá menor cantidad de falsos positivos.

8.1.3. Análisis Espacio Escala

Con el fin de mejorar los resultados se ha utilizado una representación espacio-escala, tanto a la hora de obtener los contornos utilizando el detector Canny como a la hora de extraer las líneas y círculos más representativos del espacio de Hough. A continuación se detalla el enfoque que se evaluó para obtener una mayor cantidad de correspondencias. De cara a mejorar los resultados obtenidos, se propone modificar el diagrama de flujo del método presentado de la Figura 3. El objetivo es utilizar diferentes valores de umbral en el detector de Canny para obtener diferentes contornos de la imagen LWIR. Luego se obtiene la transformada de Hough para la imagen LWIR y se busca correspondencia con la imagen VS. Las correspondencias obtenidas con los distintos valores de umbral son acumuladas en una única representación. En la Figura 15. se muestra el diagrama de flujo del método propuesto.

8. Conclusiones

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

1. Utilizar valores no convencionales en los umbrales en la imagen LWIR permite obtener una detección de bordes más robusta.
2. Utilizando el valor recomendado para el umbral de la transformada de Hough (umbral de 50, en una escala de 0 a 100) se obtendrá una mejor detección tanto en líneas como en círculos.
3. La modificación manual en el proceso de alineación de imágenes mejoró notoriamente la cantidad de correspondencias encontradas.
4. Mientras más grande sea el valor de los umbrales de Canny menor será la cantidad de correspondencias.
5. La transformada de Hough para círculos no es tan robusta en comparación a la de líneas debido a que es más probable que las formas circulares en una imagen sean discontinuadas por objetos presentes en la escena (oclusiones), mientras que las líneas son mucho más fáciles de detectar.
6. Los resultados obtenidos utilizando procedimientos generales no resultaron óptimos al momento de trabajar con imágenes LWIR debido a su naturaleza.
7. Luego de analizar los problemas que ocasionaban el bajo rendimiento de la búsqueda de correspondencias, se recomendaron sus posibles soluciones, las cuales fueron puestas a prueba en el enfoque

representado en la Figura 15, obteniéndose muy buenos resultados.

8. El método propuesto es una solución robusta para la detección de correspondencia en pares de imágenes multispectrales. Sin embargo, estos resultados todavía pueden ser mejorados utilizando alguna técnica de optimización que permita encontrar los mejores parámetros para los algoritmos utilizados en el proceso (umbrales de Canny y Hough).

9. 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. El algoritmo de Canny, a pesar de ser una de las mejores opciones al momento de detectar contornos, sigue dependiendo del ajuste de umbrales y máscaras para su optimización, los resultados obtenidos de este algoritmo pueden ser mejorados si aplicáramos un pre procesamiento al par de imágenes LWIR-VS con el fin de obtener líneas y círculos más pronunciados. De ese modo, la transformada de Hough dará mejores resultados.
2. Se recomienda dividir la imagen en diferentes regiones, y en cada una de ellas utilizar diferentes valores de umbrales dependiendo de su composición morfológica y niveles de ruido presentes. Finalmente, se deberán unir los resultados obtenidos para proceder a aplicar la transformada de Hough.
3. Para mejorar la cantidad de correspondencias encontradas se propuso mejorar la rectificación de imágenes utilizando un proceso manual de imágenes, esta etapa podría ser mejorada con la ayuda de algoritmos de rectificación que simplifiquen el problema al momento de encontrar correspondencias.

10. Referencias

[1] Hurtado, F., Atlas de matemáticas (1 edición). 1997. Idea Books, S.A, página 8.
[2] 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.

[3] Zitová B., Flusser J., Image registration methods: a survey, Image Vision Comput, páginas 977-1000, 2003.
[4] 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.
[5] 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.
[6] 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.
[7] Barrera F., Lumbreras F. and Sappa A., "Multispectral Piecewise Planar Stereo using Manhattan-World Assumption", Pattern Recognition Letters, Vol. 34, No.1, January 2013, pp. 52-61.
[8] Lowe D.G., Local Feature, View Clustering for 3D Object Recognition, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition; Kauai, HI, USA, páginas 8-14, 2001
[9] Lee J., Kim Y., Lee D., Kang D., Ra J., Robust CCD and IR image registration using gradient-based statistical information, IEEE Signal Process; páginas 347-350
[10] Zhang L., Wu B., Nevatia R., Pedestrian Detection in Infrared Images Based on Local Shape Features, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition; Minneapolis, MN, USA, páginas 1-8, 2007
[11] Canny, J., A Computational Approach To Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6) páginas 679-698, 1986
[12] Mai, F., Hung, Y., Zhong, H., Sze, W., A hierarchical approach for fast and robust ellipse extraction, Pattern Recognition, 41(8) páginas 2512-2524, 2008
[13] Hough, P., Method and Means for recognizing complex Patterns, U.S. Patent 3,069,654