



ANÁLISIS COMPARATIVO DE LAS TÉCNICAS SURF Y ORB PARA LA DETECCIÓN DE PUNTOS DE INTERÉS EN FOTOGRAFÍAS AÉREAS

Yesmar Andrés Godoy Olivera

Álvaro Isledier Ducuara Oyuela

**Facultad de Ingeniería
Programa Ingeniería Electrónica**

Ibagué, 2019

ANÁLISIS COMPARATIVO DE LAS TÉCNICAS SURF Y ORB PARA LA DETECCIÓN DE PUNTOS DE INTERÉS EN FOTOGRAFÍAS AÉREAS

Yesmar Andrés Godoy Olivera

Álvaro Isledier Ducuara Oyuela

Trabajo de grado que se presenta como requisito parcial para optar al título de:

Ingeniero Electrónico

Director:

Doctor Manuel Guillermo Forero Vargas

Profesor Universidad de Ibagué

**Facultad de Ingeniería
Programa Ingeniería Electrónica**

Ibagué, 2019

Dedicatoria

“Este proyecto va dedicado primero a Dios por darme el entendimiento y perseverancia, a mis padres por su apoyo incondicional, por el esfuerzo inagotable que realizan cada día por darme una mejor vida, por la sabiduría que me han transmitido y por la actitud que me han inculcado”.

Álvaro Ducuara

“Le dedico este proyecto principalmente a mis padres, sin el apoyo de ellos no hubiera podido sacar mi carrera adelante. El compromiso y responsabilidad de mis padres en sus trabajos ha sido el mejor ejemplo de que el trabajo duro logra cualquier cosa.”

Yesmar Godoy

Agradecimientos

Agradecemos a nuestro tutor de tesis PhD. Manuel Guillermo Forero Vargas por su inmensurable apoyo durante la realización de este estudio. Le agradecemos la paciencia, el tiempo dedicado y los conocimientos transmitidos durante todo el proceso. Podemos concluir que gracias a su pasión por el tema, logró despertar en nosotros un gusto muy grande por el área de procesamiento digital de imágenes.

Resumen

Este proyecto hace parte de una investigación desarrollada dentro del semillero en procesamiento de imágenes y patrones LÚN de la Universidad de Ibagué con el fin de evaluar técnicas para la detección y descripción de puntos de interés para ser utilizadas en el registro de fotografías aéreas tomadas en el espectro visible e infrarrojo cercano adquiridas por parte del CIAT “Centro Internacional de Agricultura Tropical”. Las técnicas evaluadas son dos métodos publicados recientemente: “Speeded Up Robust Features”, conocido como SURF, y “Oriented FAST and Rotated BRIEF”, conocido por su sigla ORB, con el propósito de encontrar puntos de interés en cada imagen que sirvan como puntos de referencia para registrarlas. Al implementar y analizar las dos técnicas independientemente, se modifican las características y se realiza el análisis comparativo que determina cuál de las técnicas es la más apropiada de utilizar para imágenes de este tipo. Adicionalmente, el filtro de Gabor fue empleado para encontrar descriptores de textura en las imágenes, encontrando que la entropía proporciona los mejores resultados.

Palabras clave: Puntos de interés, imágenes en espectro visible, imágenes en espectro infrarrojo cercano, método SURF, detector FAST, descriptor BRIEF.

Abstract

This project is part of a research developed, within the image processing and pattern recognition seedbed Lún of the University of Ibagué, with the goal of evaluating detection and description of interest points techniques to register visible and near infrared aerial photographs. These pictures were acquired by the International Center for Tropical Agriculture "CIAT". The evaluated techniques are two recently published methods: "Speeded Up Robust Features", known as SURF, and "Oriented FAST and Rotated BRIEF", known by its acronym ORB, with the purpose of finding interest points employed as reference to merge them. Both methods were studied and a comparative analysis was carried out to determine which one is the most appropriate to merge this kind of images. In addition, Gabor filters were employed to find texture descriptors in the images, finding that entropy provides the best result.

Keywords: interest points, images in visible spectrum, images in near infrared spectrum, SURF method, FAST detector, BRIEF descriptor.

Contenido

INTRODUCCIÓN	1
1. OBJETO DE ESTUDIO	3
1.1 PLANTEAMIENTO DEL PROBLEMA	3
1.2 JUSTIFICACIÓN	3
1.3 ESTADO DEL ARTE	4
1.4 OBJETIVO GENERAL.....	6
1.4.1 OBJETIVOS ESPECÍFICOS	6
2. MARCO TEÓRICO	7
2.1 PROCESAMIENTO DIGITAL DE IMÁGENES.....	7
2.2 PUNTO DE INTERÉS	8
2.3 ALGORITMO SURF	8
2.3.1 <i>Detector</i>	9
2.3.2 <i>Descriptor</i>	14
2.3.3 <i>Emparejamiento</i>	18
2.4 ALGORITMO ORB.....	19
2.4.1. <i>Detector FAST</i>	19
2.4.2. <i>Orientación por intensidad del centro</i>	20
2.4.3. <i>Descriptor BRIEF</i>	21
2.5 FILTRO DE GABOR.....	22
3. DESARROLLO.....	25
3.1 MÉTODO SURF	25
3.2 MÉTODO ORB.....	26
3.3 PROTOCOLO DE EVALUACIÓN	27
4. RESULTADOS.....	28
4.1 ANÁLISIS DE RESULTADOS MÉTODO SURF	28
4.2 ANÁLISIS DE RESULTADOS MÉTODO ORB	36

4.3 COMPARACIÓN DE LOS MÉTODOS EN PRUEBAS DE RENDIMIENTO	42
4.4 EXPERIMENTACIÓN CON IMÁGENES AÉREAS	43
4.4.1. Análisis De Resultados Del Método SURF con fotografías aéreas.....	44
4.4.2. Análisis De Resultados Del Método ORB.....	49
5. DESCRIPTOR DE GABOR	52
6. CONCLUSIONES Y RECOMENDACIONES.....	55
6.1. CONCLUSIONES	55
6.2. RECOMENDACIONES.....	56
REFERENCIAS BIBLIOGRÁFICAS	57

Lista de figuras

	Pag.
Figura 1. Procesamiento digital de imágenes.....	7
Figura 2. Puntos de interés en dos imágenes	8
Figura 3. Algoritmo SURF	9
Figura 4. Imagen Integral	10
Figura 5. Operaciones Imagen integral	10
Figura 6. Izquierda a derecha: imagen original, derivada en y, derivada en x	11
Figura 7. Comportamiento de la señal a los filtros.....	12
Figura 8. Filtro Binarizado	13
Figura 9. Filtros aproximados de la matriz Hessiana	14
Figura 10. Escala de detección	14
Figura 11. Filtros de Haar	15
Figura 12. Escala de detección S.....	15
Figura 13. Vecindaje circular de tamaño 6S	16
Figura 14. Suma de respuestas de Haar.....	16
Figura 15. Dirección del punto de interés	16
Figura 16. Cuadricula 4x4 en punto de interés	17
Figura 17. Subcuadricula de 5x5.....	17
Figura 18. Emparejamiento de puntos de interés	18
Figura 19. Detección de características mediante círculo de 16 pixeles.....	20
Figura 20. Puntos de interés detectados mediante FAST.	21
Figura 21. Identificación de puntos de interés con el método ORB	22
Figura 22. Forma de la Función de Gabor.....	23
Figura 23. Resultados de la aplicación del método Gabor en distintos ángulos.	24
Figura 24. Detección y descripción del algoritmo SURF.....	26
Figura 25. De Izquierda a derecha: imagen estándar (Clown), imagen NIR, imagen VIS.....	27
Figura 26. Líneas de congruencia entre la imagen original a la imagen blanco y negro.	28
Figura 27. Líneas de congruencia entre la imagen original y la imagen escalada a un tamaño con porcentaje menor.	29
Figura 28. Líneas de congruencia entre la imagen original y la imagen en espejo horizontal....	30
Figura 29. Líneas de congruencia entre la imagen original y la imagen con cambios de rotación	

.....	31
Figura 30. Líneas de congruencia entre la imagen original y la imagen con cambios de traslación en pixeles.....	32
Figura 31. Líneas de congruencia entre la imagen original y la imagen con ruido gaussiano....	34
Figura 32. Grafica resultante de la cantidad de puntos de interés del algoritmo SURF	35
Figura 33. Líneas de congruencia detectadas por el algoritmo ORB.....	36
Figura 34. Similitud de puntos de interés entre imágenes escaldadas por el algoritmo SURF.	36
Figura 35. Similitud de puntos de interés en una imagen con espejo horizontal por el algoritmo ORB.....	37
Figura 36. Similitud de puntos de interés en imágenes con diferentes ángulos de rotación por el algoritmo ORB	38
Figura 37. Similitud de puntos de interés en imágenes con traslación de pixeles por el algoritmo ORB.....	39
Figura 38. Similitud de puntos de interés en imágenes con perturbaciones de ruido gaussiano por parte del algoritmo ORB.....	40
Figura 39. Grafica resultante de la cantidad de puntos de interés del algoritmo ORB	41
Figura 40. Cantidad de puntos de interés detectados en cada parámetro por los algoritmos SURF y ORB	42
Figura 41. Tiempo de respuesta de los algoritmos SURF y ORB por cada parámetro	43
Figura 42. Resultados de las fotografías de cultivo de pasto con los parámetros: Octaves: 3, Layers Per Octaves: 3, Hessian Thresholded: 0.00010, Initial Sampling Step: 1.....	44
Figura 43. Resultados de las fotografías de cultivo de pasto con los parámetros: Octaves: 4, Layers Per Octaves: 4, Hessian Thresholded: 0.00010, Initial Sampling Step: 2.....	45
Figura 44.Resultados de las fotografías de cultivo de pasto con los parámetros: Octaves: 3, Layers Per Octaves: 3, Hessian Thresholded: 0.00010, Initial Sampling Step: 1.....	46
Figura 45. Resultados de las fotografías de cultivo de pasto con los parámetros: Octaves: 4, Layers Per Octaves: 4, Hessian Thresholded: 0.00010, Initial Sampling Step: 2.....	47
Figura 46. Resultados de las fotografías de cultivo de pasto con los parámetros: Octaves: 3, Layers Per Octaves: 3, Hessian Thresholded: 0.00010, Initial Sampling Step: 1.....	48
Figura 47. Resultados de las fotografías de cultivo de pasto con los parámetros: nfeatures: 500, scaleFactor: 1.2, nlevels: 8, edgeThreshold: 31, WTA_K: 2, patchSize: 31	50
Figura 48. Resultados de las fotografías de cultivo de pasto con los parámetros: nfeatures: 500, scaleFactor: 1.2, nlevels: 2, edgeThreshold: 31, WTA_K: 2, patchSize: 31	51
Figura 49. Objetos de referencia para la detección de textura Gabor.....	52

Figura 50. Resultados del algoritmo Gabor en distintos ángulos.....	52
Figura 51. Parámetros del descriptor Gabor	53
Figura 52. Resultado obtenido de la entropía en las imágenes NIR y VIS.....	53
Figura 53. Resultado obtenido de la desviación estándar en cada imagen aérea NIR y VIS.....	54
Figura 54. Resultado obtenido de la media en las imágenes aéreas NIR y VIS	54

Lista de Tablas

Tabla 1. Valores de los parámetros de entrada del método SURF.....	25
Tabla 2. Parámetros de entrada del método ORB.....	26
Tabla 3. Resultados obtenidos del algoritmo SURF	28
Tabla 4. Resultados obtenidos del algoritmo SURF en cambio de escala.....	30
Tabla 5. Resultados obtenidos del algoritmo SURF al cambio de espejo horizontal	30
Tabla 6. Resultados obtenidos del algoritmo SURF al cambio de espejo horizontal	32
Tabla 7. Resultados obtenidos del algoritmo SURF al cambio de traslación de la imagen.....	33
Tabla 8. Resultados obtenidos del algoritmo SURF en respuesta a una imagen con perturbación de ruido gaussiano.....	34
Tabla 9. Resultados Obtenidos del Algoritmo ORB al cambio de niveles de grises ByN	37
Tabla 10.Resultados Obtenidos del algoritmo ORB a cambios de rotación de la imagen.....	38
Tabla 11. Resultados Obtenidos del algoritmo ORB a cambios de rotación de la imagen.....	39
Tabla 12. Resultados Obtenidos del algoritmo ORB a perturbaciones en la imagen como el ruido gaussiano	40
Tabla 13. Resultados de las pruebas del algoritmo ORB	41
Tabla 14. Resultados de la imagen NIR_A_T1 y VIS_A_T1 con parámetros: Octaves: 3, Layers Per Octaves: 3, Hessian Thresholded: 0.00010, Initial Sampling Step: 1	45
Tabla 15. Resultados de la imagen NIR_A_T1 y VIS_A_T1 con parámetros: Octaves: 4, Layers Per Octaves: 4, Hessian Thresholded: 0.00010, Initial Sampling Step: 2.	46
Tabla 16. Resultados de la imagen NIR_C_T2 y VIS_C_T2 con parámetros: Octaves: 3, Layers Per Octaves: 3, Hessian Thresholded: 0.00010, Initial Sampling Step: 1.	47
Tabla 17. Resultados de la imagen NIR_C_T2 y VIS_C_T2 con parámetros: Octaves: 4, Layers Per Octaves: 4, Hessian Thresholded: 0.00010, Initial Sampling Step: 2.	48
Tabla 18. Resultados de la imagen NIR_H_T6 y VIS_H_T6 con parámetros: Octaves: 3, Layers Per Octaves: 3, Hessian Thresholded: 0.00010, Initial Sampling Step: 1.	49

Lista de Símbolos y abreviaturas

Símbolo	Término	Unidad SI	Definición
$I(x)$	Imagen	pixels	unidad básica que forma una imagen digital
N	Numero de pixeles	pixels	unidad básica que forma una imagen digital
$U_{(x,y)}$	Imagen integral		Ecu 1
$q_{(i,j)}$	Valor pixel		Es valor del pixel en la posición (i, j) .
T	Intensidad de una zona		Ecu 2
$g(x, y)$	Funcion Gaussiana		Ecu 4
$H(f)$	Matriz Hessiana		Ecu 3
$L_{xx}(x, \sigma)$	segunda derivada de una gaussiana de tamaño (σ) en x.		Ecu 6
$\Delta H(x, \sigma)$	Determinante de la matriz Hessiana		Ecu 7
$\varphi(t)$	Función de Harr		Ecu 9
$\sum h_x$	Sumatoria de respuestas de filtro de Haar en x		Ecu 10
$\sum h_y$	Sumatoria de respuestas de filtro de Haar en y		Ecu 11
$\sum h_x $	Sumatoria absoluta de respuestas de filtro de Haar en x		Ecu 12
$\sum h_y $	Sumatoria absoluta de respuestas de filtro de Haar en y		Ecu 13
V_d	Vector descriptor SURF		Ecu 15
$h(x, y, \theta_k, f, \sigma_x, \sigma_y)$	Funcion de Gabor		Ecu 24
m_{pq}	momento de un parche		Ecu 16
C	Centroide de detector FAST		Ecu 17

Θ	Orientación del parche		Ecu 18
$T(p, x, y)$	prueba binaria		Ecu 21
$f_n(p)$	vector de n pruebas binarias		Ecu 20
$g_n(p, \theta)$	operador BRIEF dirigido		Ecu 23

Abreviaturas

Abreviatura	Término
ORB	Oriented Fast and Rotated Brief
SURF	Speed Up Robust Features
FAST	Features from accelerated segment test
BRIEF	Binary Robust Independent Elementary Features
CIAT	Centro Internacional De Agricultura Tropical
NIR	Imágenes en infrarrojo cercano
VIS	Imágenes en espectro Visible

Introducción

En procesamiento digital de imágenes, existen técnicas que permiten determinar las características de las mismas; mediante un proceso que parte de la detección de las zonas más relevantes tales como: esquinas, bordes y manchas, continúa la búsqueda de la similitud y finaliza con el reconocimiento entre objetos. La visión artificial, con el fin de automatizar procesos, adopta estas técnicas para reconocimiento y seguimiento de objetos, contribuyendo así en tareas de inspección. Por esta razón, en este trabajo se pretende realizar un estudio comparativo para determinar el método más eficiente para el registro de fotografías aéreas en el espectro visible e infrarrojo cercano de cultivos, tomadas en el Centro Internacional de Agricultura Tropical “CIAT” ubicado en Palmira, Valle del Cauca.

Uno de los pioneros en esta área fue Hans Moravec, quien en 1980 introdujo un algoritmo capaz de extraer información de las esquinas en una imagen como principal y único método de reconocimiento de objetos [11]. Esta técnica no es robusta, pues el método no es isotrópico introduciendo errores en la detección de los puntos de interés cuando los bordes de los objetos no están alineados con las filas y columnas de la imagen. Por tal motivo, se han desarrollado métodos más eficientes que permiten detectar con mayor precisión y robustez las esquinas de los objetos. Algunas de estas técnicas permiten, además, extraer características de la imagen, haciendo posible comparar puntos de interés de un grupo de imágenes con el fin de determinar la coincidencia de objetos. Actualmente, dentro de las técnicas destacadas, se encuentran los algoritmos SURF (Speed Up Robust Features) [1] y ORB (Oriented FAST and Rotated BRIEF) [2]. Las aplicaciones más comunes de estas técnicas se dan en el campo del reconocimiento facial, seguimiento de patrones y calibración fotográfica.

La adquisición de imágenes de diferentes modalidades es otra aplicación de interés para llevar acabo el desarrollo de estas técnicas. En nuestro caso, las fotografías aéreas en el espectro infrarrojo y visible son captadas a través de un dron. Estas imágenes fueron tomadas en el Centro Internacional de Agricultura Tropical (CIAT) con el fin de realizar un análisis de índices de vegetación. Para ello, estas imágenes son normalmente enviadas a un laboratorio de la Universidad de Tokio para su fusión y análisis. Este trabajo hace parte de un proyecto liderado por el Doctor Michael Selvaraj, quien busca analizar el crecimiento de diferentes tipos de pasto y determinar la mejor variedad.

Con el fin de reducir costos y desarrollar una tecnología propia, el semillero en procesamiento de imágenes, visión por computador y reconocimiento de patrones Lún de la Universidad de Ibagué comenzó una colaboración con Selvaraj para desarrollar herramientas que permitan inicialmente realizar el registro de estas imágenes.

El principal objetivo de este trabajo es detectar la técnica más adecuada para la detección y correlación de imágenes de origen de espectro visible e infrarrojo evaluando características como velocidad, recursos computacionales, cantidad y calidad de puntos de interés y robustez a rotaciones y perturbaciones. Aunque no se planteó en los objetivos presentados en el anteproyecto, se obtuvieron además tres descriptores de textura a partir de la transformada de Gabor, con el fin de suplir las desventajas encontradas en los descriptores obtenidos con el algoritmo ORB.

Para el desarrollo de este proyecto, se realizó una búsqueda bibliográfica con el fin de recopilar información teórica sobre las técnicas SURF y ORB, entender su funcionamiento y buscar implementaciones de estos métodos. De esta forma, se encontraron varios archivos fuente, los cuales fueron estudiados y mejorados con el fin de hacerlos más eficientes y apropiados para el análisis. De esta manera, el método SURF fue desarrollado en lenguaje JAVA como plugin de programa de libre acceso Imagej. El método ORB se desarrolló en lenguaje Python, y dado que los resultados no fueron los esperados, no se justificó la implementación del algoritmo en lenguaje JAVA. Una vez se tuvieron las técnicas implementadas se evaluaron empleando imágenes manipuladas en las cuales se introdujeron rotaciones y perturbaciones. Finalmente, se realizaron las pruebas con las imágenes de diferentes espectros entregadas por el CIAT, para encontrar la técnica más adecuada para este tipo de imágenes.

El algoritmo ORB, presentó falencias en el proceso de detección respecto a las congruencias entre imágenes, encontrando un 90% de falsos positivos puntos de interés, por tal motivo, la técnica fue considerada no apta para este tipo de imágenes. Esto conllevo, a la implementación y análisis del descriptor de textura Gabor.

1. Objeto de Estudio

1.1 Planteamiento del Problema

El procesamiento digital de imágenes permite el registro de imágenes en diferentes modalidades. En el Centro Internacional de Agricultura Tropical (CIAT) se trabaja con fotografías aéreas en espectro visible e infrarrojo cercano de un cultivo en pasto tomadas desde un dron, con el fin de estudiar el crecimiento y desarrollo del pasto, y determinar las mejores variedades provenientes de este cultivo. Una vez adquiridas, las imágenes son enviadas a la Universidad de Tokio para su fusión, y realizar un estudio de índices de vegetación, a un elevado costo, dado que no se cuenta con la tecnología adecuada para su análisis en Colombia. Con el fin de reducir los costos e introducir esta tecnología en Colombia, el semillero de procesamiento de imágenes y patrones Lún ha comenzado una investigación en esa área con el fin de determinar la técnica más adecuada para encontrar los puntos de interés en fotografías de ambos espectros con el fin de fusionarlas.

Los puntos de interés son los puntos más representativos de una imagen, que se extraen a partir de esquinas, bordes y manchas de la misma, con el fin de reconocer objetos. Por lo tanto, se requiere el análisis e implementación de las técnicas más recientes y consideradas en el estado del arte, con el fin de compararlas para tomar aquella que produzca los mejores resultados. Por lo tanto, una vez estudiadas las técnicas de detección de puntos de Harris y SIFT en un trabajo anterior, se requiere el estudio de dos de las otras técnicas más avanzadas en esa área.

1.2 Justificación

De acuerdo a estudios anteriores sobre métodos de identificación y comparación de puntos de interés por parte de estudiantes del semillero Lún de la Universidad de Ibagué en relación con las técnicas Harris y SIFT, se llevó a cabo el análisis de nuevas técnicas que han surgido producto del avance tecnológico, el cual, busca siempre el mejoramiento de la eficiencia de cada método. Las técnicas SURF y ORB, pretenden ser más rápidas, precisas y robustas a alteraciones y perturbaciones presente en imágenes, por lo tanto se realizó un estudio comparativo con el fin de obtener la técnica más eficiente. Las imágenes analizadas en este estudio son ‘fotografías estándar empleadas para el estudio de técnicas de procesamiento de

imágenes de cultivos de pasto adquiridas en los espectros visible e infrarrojo cercano. Estas fotografías son empleadas para encontrar índices de vegetación que permitan evaluar el crecimiento de los pastos y encontrar las mejores variedades.

1.3 Estado del arte

El procesamiento digital de imágenes, desarrolla algoritmos que permiten mejorar la calidad de una imagen, para luego obtener tanta información como sea posible de esta misma. El avance en las técnicas utilizadas en procesamiento de imágenes permite extraer información más precisa y exacta aun cuando la imagen presente perturbaciones. La identificación de puntos de interés sirve para encontrar sitios de referencia que permitan hacer una correspondencia entre una imagen y otra con el fin de fusionarlas o reconocer objetos en ellas. Así, las aplicaciones de las técnicas de detección, descripción y comparación de puntos de interés son amplias, como por ejemplo: en el seguimiento de objetos, reconocimiento facial, visión robótica, exploración espacial, reconocimiento de objetos, calibración de registro de fotografías, entre otros. A lo largo de los años estas técnicas han ido mejorando con el fin de hacerlas más precisas y eficientes.

Las primeras técnicas se basaron en la detección de puntos de interés. En 1988 Harris introdujo un detector de esquinas basado en cambios de intensidad en distintas direcciones presentes en una imagen [11]. El algoritmo de Harris guarda en una matriz, de tamaño de 2x2, los gradientes de un punto y determina si corresponde a una esquina examinando sus componentes. Una desventaja de este método es que no es invariante al cambio en escala de los objetos. Diez años más tarde, Tony Lindeberg, profesor de KTH Royal Institute of Technology, desarrolla un método que incluye la selección de escala automática en la identificación de puntos de interés en una imagen [13], lo cual permite detectar un mismo objeto independiente de su distancia al observador. Para ello, Lindeberg trabajó con el determinante de la matriz Hessiana y con el Laplaciano para detectar estructuras. Un año más tarde, el científico David Lowe introdujo en su artículo “Object recognition from local scale-invariant features” [13] un método para reconocer objetos de características invariantes a escala local, en el cual propuso aproximar el laplaciano del gausiano (LoG) por el filtro de diferencias de gaussianos (DoG). En el mismo año, Paul L. Rosin, integrante del departamento de Sistemas de Información e Informática de la Universidad de Brunel en Middlesex, Reino Unido, desarrolló la técnica “Measuring corner properties”[15], la

cual mide las propiedades de las esquinas de nivel de gris, como el ángulo subtendido, la orientación, el contraste, la brusquedad o redondeo del vértice y curvatura del límite para cúspides. En el 2001 Krystian Mikolajczyk y Cordelia Schmid en su artículo “Indexing based on scale invariant” [16] presentaron un método basado en detectores de características robustos e invariantes en escala con alta repetitividad, que denominaron Harris-Laplace y Hessian-Laplace. Básicamente utilizaron el algoritmo del Harris, adaptado a la escala, y el determinante de la matriz Hessiana para seleccionar la ubicación, y el Laplaciano para seleccionar la escala, buscando una mayor velocidad. Otros algoritmos que fueron surgiendo, se centraron en la detección de puntos de interés en secuencia de videos. En el año 2003, Josef Sivic y Andrew Zisserman, del grupo de investigación en robótica, del departamento de ciencias de la Ingeniería en la Universidad de Oxford, Reino Unido, publicó el método “Object Matching in Videos” [17], el cual se enfoca en la recuperación de objetos y escenas. Este método busca y localiza todas las ocurrencias de un objeto en un video. Un año más tarde David Lowe en su artículo “Distinctive Image Features from Scale-Invariant Keypoints” [18] propone una técnica que extrae características invariantes distintivas de las imágenes y puede usarse para realizar una comparación confiable entre diferentes vistas de un objeto o escena. En el mismo año, David Lowe presentó una de las técnicas más relevantes en el área de identificación de puntos de interés: el método SIFT presentado en su artículo ”Scale invariant feature transform” [19], el cual es un método que incluye un detector de puntos y un descriptor. El detector se basa en reducir la escala de la imagen original y obtener la diferencia de Gaussianos (DoG) para encontrar los puntos de interés y el descriptor está dado por un vector de 128 elementos que representan la ubicación y sentido de los puntos de interés, el cual, permite la comparación de imágenes, conocida en inglés como “matching”. El SIFT es un algoritmo muy usado actualmente gracias a sus características de invariancia a la rotación, traslación y cambio de escala. En el año 2005 K. Mikolajczyk y C. Schmid presentaron el artículo “A performance Evaluation of Local Descriptors” [20], en el cual se comparan una serie de descriptores calculados para las regiones de interés local, extraídos con el detector Harris-Affine. Un año después, Edward Rosten and Tom Drummond de la universidad de Cambridge publicaron la técnica “high-speed corner detection”[21], la cual es empleada como primer paso en muchas tareas de visión como seguimiento de objetos, SLAM (localización y mapeo simultáneos), coincidencia de imágenes y reconocimiento aplicados a escenas 3D.

En el semillero de procesamiento de imágenes, visión por computador y reconocimiento de

patrones LÚN de la Universidad de Ibagué, Juan Guillermo Moreno y Andrea Liseth Acosta realizaron bajo la dirección del Doctor Ingeniero Manuel Forero un estudio comparativo de los algoritmos SIFT y Harris. Como continuación de ese trabajo, en este proyecto de grado se realiza el análisis comparativo de dos de las técnicas más recientes en esta área con el fin de encontrar el método más eficiente para realizar un registro muy preciso de las imágenes enviadas por el CIAT “Centro Internacional de Agricultura Tropical”.

1.4 Objetivo general

Realizar un estudio práctico y analítico de los métodos SURF y ORB para la detección y descripción de puntos de interés en fotografías aéreas de cultivos en el espectro infrarrojo cercano y visible.

1.4.1 Objetivos específicos

- Recopilar el estado del arte sobre el desarrollo de métodos basados en las técnicas SURF y ORB.
- Analizar e implementar los métodos de descriptores BRIEF y detector de puntos Fast, que conforman la técnica ORB.
- Analizar e implementar el método SURF para detección de puntos de interés.
- Realizar un estudio comparativo de los métodos SURF Y ORB.

2. Marco Teórico

2.1 Procesamiento digital de imágenes

El Procesamiento digital de imágenes es un área de investigación, en la cual se estudian y desarrollan técnicas aplicadas a las imágenes digitales con el objetivo de obtener como resultado imágenes de mejor calidad y obtener información de interés contenida en ellas. Las técnicas estudiadas por el procesamiento de imágenes permiten: eliminar señales no deseadas en la imagen como el ruido producido en el proceso de adquisición de la imagen, suavizar la imagen con el fin de reducir variaciones entre los pixeles cuando la imagen es muy saturada, resaltar rasgos de la imagen como bordes, eliminar características de la imagen como objetos no deseados, manchas, información irrelevante (barro, arena, polvo), resaltar un objeto de interés dentro de una imagen, obtener características o puntos de interés de la imagen para poder compararla con otras imágenes, etc [1]. El procesamiento digital de imágenes se compone de varias etapas:

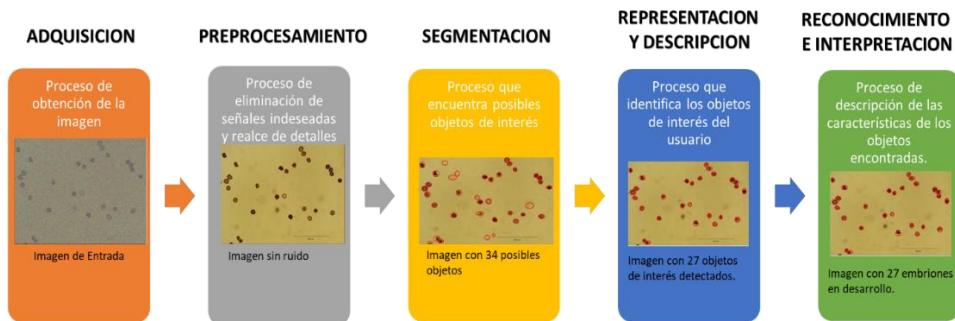


Figura 1. Procesamiento digital de imágenes

2.2 Punto de interés

Una imagen está conformada por determinadas áreas locales que contienen información relevante para el procesamiento digital. Estas áreas pueden ser identificadas por medio de patrones clave llamados puntos de interés; los cuales, facilitan la detección de zonas esenciales en una imagen aun cuando contienen alteraciones o perturbaciones como ruido. Un punto de interés en una imagen es un píxel que tiene una posición bien definida y puede ser detectado de manera robusta. Los pixeles situados en un borde presentan un mayor interés. Sin embargo, no son útiles, puesto que la región donde se encuentra un pixel en un contorno se parece mucho a los demás pixeles de este borde en la misma región. Por el contrario, un punto de una esquina es mucho más informativo, puesto que permite delimitar de mejor manera un objeto y no son puntos tan repetitivos, por lo cual son utilizados como puntos de interés [22].

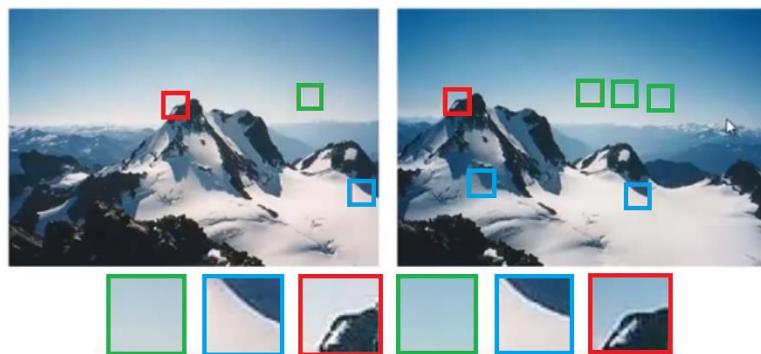


Figura 2. Puntos de interés en dos imágenes

En la figura 2, se presentan dos imágenes donde los recuadros en color rojo son coincidencias de esquinas, recuadros azules son bordes y recuadros verdes zonas planas entre dos imágenes [22].

2.3 Algoritmo SURF

El método conocido por la sigla SURF, del inglés Speeded-Up Robust Features, es un algoritmo que permite extraer, describir y comparar características propias de una imagen. El algoritmo fue presentado en Austria en el año 2006 por Herbert Bay, Andreas Ess, Tinne Tuytelaars y Luc Van Gool en la Conferencia Internacional de visión por computador. Este algoritmo es más rápido y robusto a transformaciones como cambio de escala, orientación, rotación, desenfoque, cambios de iluminación y ruido que el SIFT, considerado como estado del arte en ese momento. El algoritmo SURF posee tres etapas compuestas por el detector, el descriptor y el comparador. El

detector se encarga de encontrar los puntos de interés característicos de una imagen mediante la aproximación de diferencia de Gaussianos (DoG) y la matriz hessiana. En el descriptor se construye un vector de 64 características que describen la posición y orientación de los puntos de interés utilizando las sumas de las respuestas del wavelet Haar. Finalmente, en la etapa de comparación (matching) se identifican los mismos puntos en dos imágenes, comparando la distancia euclídea entre los vectores descriptores de los puntos de interés. Cada una de las etapas se explica con mayor profundidad a continuación. La figura 3 presenta el diagrama de bloques del método SURF [1].



Figura 3.Algoritmo SURF

En color amarillo aparecen los bloques que hacen parte del detector, en verde los del descriptor y el bloque azul es el comparador.

2.3.1 Detector

El detector está constituido por varias sub-etapas descritas a continuación.

- Imágenes integrales

Después de adquirir la imagen de entrada, el algoritmo SURF crea una nueva imagen llamada integral. Es una imagen donde el valor de cada pixel está dado por la suma de las intensidades de los pixeles en los ejes horizontal y vertical. La figura 4 muestra un ejemplo de una imagen y su imagen integral resultante. El uso de la imagen integral permite reducir el número de operaciones de $n \times n$ sumas enteras por pixel a sólo 4 operaciones por pixel, siendo n el tamaño

del lado de la máscara, lo cual disminuye notablemente el costo computacional y aumenta la velocidad del método [1].

Imagen Original				Imagen Integral			
1	3			1	4	6	10
2	1	5	9	3	7	14	27
5	2	2	1	8	14	23	37
1	1	2	2	9	16	27	43
10	12	15	14	19	38	64	94

Figura 4.Imagen Integral

La ecuación que describe la imagen integral está dada por:

$$U_{(x,y)} = \sum_{0 \leq i < M} \sum_{0 \leq j < N} q_{(i,j)} \quad ECU\ 1$$

Donde $q_{(i,j)}$ es valor del pixel en la posición (i,j) .

La mayor ventaja del uso de imágenes integrales es que para encontrar la intensidad de un pixel o de una región solo se deben hacer 4 operaciones.

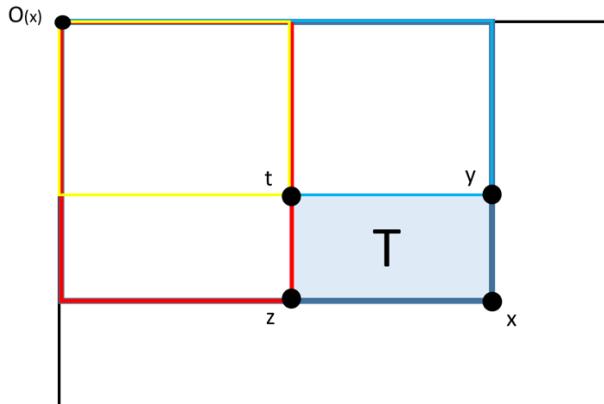


Figura 5.Operaciones Imagen integral

Así, para hallar la intensidad de la zona (T):

$$T = O_{(x)} - O_{(z)} - O_{(y)} + O_{(t)} \quad ECU\ 2$$

- Matriz Hessiana

Después de crear la imagen integral se proceden a hallar los puntos de interés. Debido a que los filtros de derivación de una imagen permiten encontrar los cambios abruptos de intensidad en una dirección dada, el algoritmo SURF utiliza una matriz cuadrada, denominada Hessiana, que contiene las segundas derivadas parciales de la imagen para detectar cambios de todas las direcciones, tal como ilustra la figura 6.

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} \end{bmatrix} \quad ECU 3$$

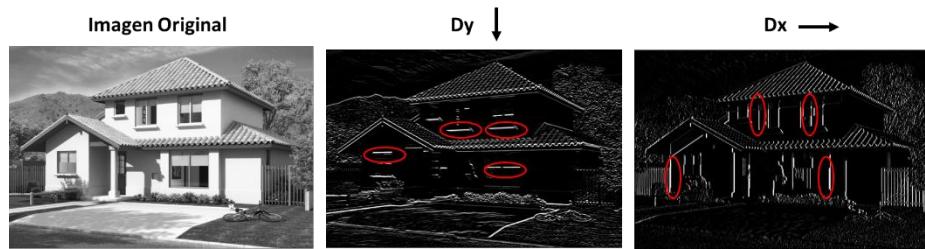


Figura 6. Izquierda a derecha: imagen original, derivada en y, derivada en x

La derivada es un buen mecanismo para detectar cambios de intensidad de la imagen pero su principal inconveniente es su sensibilidad al ruido. Por esta razón el SURF realiza primero un suavizado mediante un filtro gaussiano expresado por la ecuación (4) con el fin de evitar falsos puntos de interés debidos a ruido [1].

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad ECU 4$$

Donde la desviación estándar σ determina el grado de suavización de la imagen.

Luego se obtienen la segundas derivadas de la imagen a lo largo de los ejes x, y y la derivada xy. Puesto que las imágenes son bidimensionales, la matriz hessiana es de tamaño 2x2, dada por:

$$H(x, \sigma) = \begin{bmatrix} Lxx_{(x, \sigma)} & Lxy_{(x, \sigma)} \\ Lxy_{(x, \sigma)} & Lyy_{(x, \sigma)} \end{bmatrix} \quad ECU\ 5$$

Donde cada componente de la matriz, corresponde a la imagen ($I_{(x)}$) filtrada por la segunda derivada de una gaussiana de tamaño (σ), matemáticamente denotada por:

$$Lxy_{(x, \sigma)} = I_{(x)} * \frac{\partial^2}{\partial xy} g(\sigma) \quad ECU\ 6$$

El determinante de la matriz Hessiana ($\Delta H(x, \sigma)$) expresa la magnitud del cambio local alrededor de un punto. Por lo tanto, si el valor del determinante está por encima de un umbral este punto sería un posible punto de interés.

$$\text{Posible punto de interés} = \Delta H(x, \sigma) > \text{umbral} \quad ECU\ 7$$

Detección de puntos de interés en una señal unidimensional

Para demostrar el comportamiento del proceso de detección de puntos de interés se utiliza una señal unidimensional, a la cual se le aplican los filtros del método SURF para encontrar los puntos de interés, tal como ilustra la figura 7.

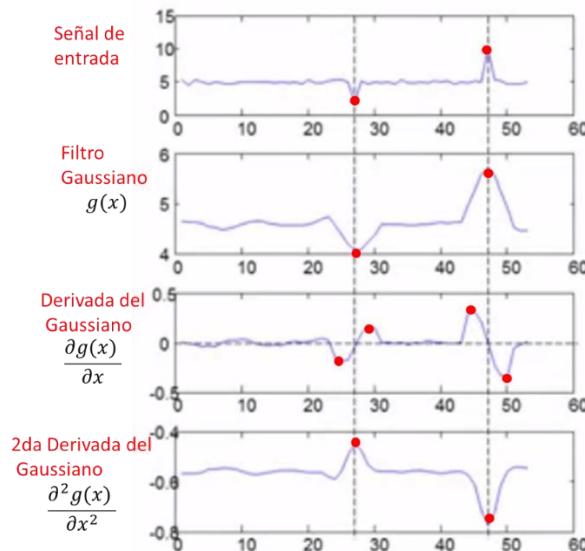


Figura 7.Comportamiento de la señal a los filtros.

La señal de entrada presenta algunos puntos relevantes que podrían ser puntos de interés pero también debidos a ruido. Para eliminar el ruido se utiliza un filtro gaussiano que también suaviza la imagen. Para hallar los puntos relevantes como bordes se aplica la derivada del Gaussiano o gradiente que resalta más los puntos de interés. Con el fin de localizar los cambios locales más significativos se calcula la segunda derivada o Laplaciano (LoG) que determina los verdaderos puntos máximos y mínimos de la señal [1]. La expresión matemática que define el LoG está dada por:

$$LoG(x, y) = -\frac{1}{\pi\sigma^4} \left[1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2+y^2}{2\sigma^2}} \quad ECU\ 8$$

- Aproximación de los filtros

Una de las principales características del Algoritmo SURF es su rapidez, la cual se logra gracias a que los filtros mencionados antes son binarizados para obtener cálculos más rápidos.

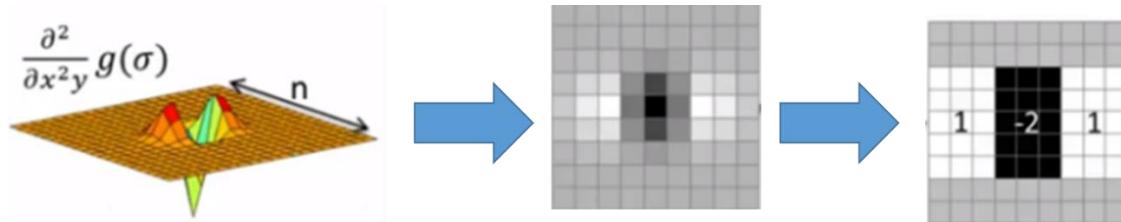


Figura 8.Filtro Binarizado

Para ello se aproxima cada filtro, intentando conservar su forma básica, asignando pesos enteros a la máscara en lugar de valores decimales. Así, se logra un filtro computacionalmente menos costoso debido a que deja de hacer $n \times n$ multiplicaciones entre números decimales, siendo n la dimensión de lado del filtro, a $n \times n$ sumas con números enteros [1]. La figura 9, muestra cada uno de los filtros aproximados de la matriz Hessiana.

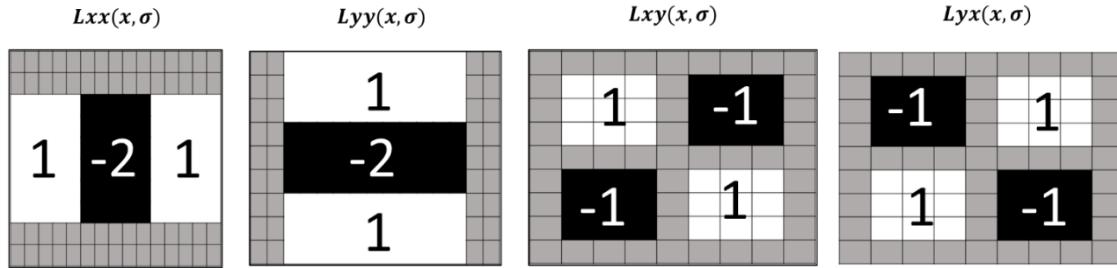


Figura 9.Filtros aproximados de la matriz Hessiana

- Detección del punto máximo entre escalas

La matriz Hessiana es aplicada en cada escala determinada por la desviación estándar σ , donde encuentra los puntos de interés. Para terminar el proceso de detección, el último paso consiste en encontrar el punto de interés con mayor magnitud entre escalas, tal como ilustra la figura 10.

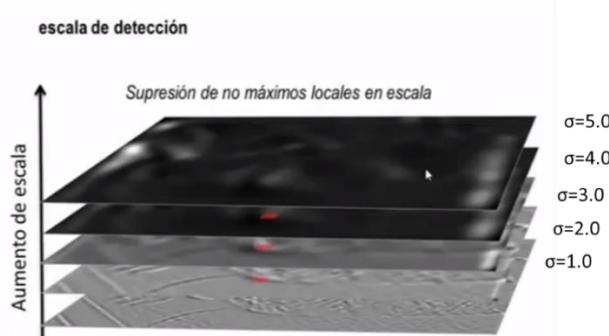


Figura 10.Escala de detección

Así, se localizan los máximos locales en una vecindad $3 \times 3 \times 3$, incluyendo de esta manera, las escalas inmediatamente superior e inferior (figura 10). Si un pixel es un máximo local será considerado un punto de interés y su importancia estará dada por su magnitud de gradiente

2.3.2 Descriptor

Una vez encontrados los puntos de interés, se procede a encontrar sus características con el fin de describirlos y compararlos con puntos en otra imagen. En la fase de descripción se obtiene la orientación de los puntos de interés y se emplean filtros wavelet de Haar, llamados así debido al matemático Alfred Haar, quien los propuso en 1909. El filtro Haar permite encontrar cambios de

intensidad en una dirección[1]. La función escalar que describe el wavelet de Haar está dada por:

$$\varphi(t) = \begin{cases} 1 & 0 \leq t < 1, \\ 0 & \text{De otra forma.} \end{cases} \quad ECU\ 9$$

Los filtros Haar se comportan como operadores de gradiente. Para encontrar la orientación, se obtiene la respuesta a los filtros Haar en la dirección “x” y “y” en los puntos de interés. Las máscaras de Haar están dados por unos y menos unos, con el fin de emplear nuevamente imágenes integrales, lo cual permite acelerar los cálculos. De esta manera las máscaras empleadas son las siguientes:

$$\begin{bmatrix} -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 \\ -1 & -1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Figura 11.Filtros de Haar

- Orientación

En la fase de detección del punto de interés se realizó una vecindad de 3x3x3 para determinar el punto más significativo entre las tres escalas. La escala donde se encontró el punto de mayor magnitud, es la escala de detección S.

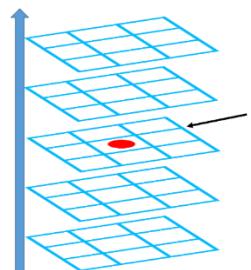


Figura 12.Escala de detección S

La región donde se encuentra el punto de interés debe ser analizada para poder determinar el vector descriptor del punto de interés. Para esto se debe ubicar un vecindaje circular de tamaño 6S donde el centro sea el punto de interés.



Figura 13. Vecindaje circular de tamaño 6S

Después de centrar la región circular en el punto de interés, el algoritmo divide la circunferencia en $\pi/3$ y en cada sección se aplican los filtros de wavelet de Haar en "x" y en "y". Donde las máscaras del filtro (figura 11) tienen un tamaño de 4S.

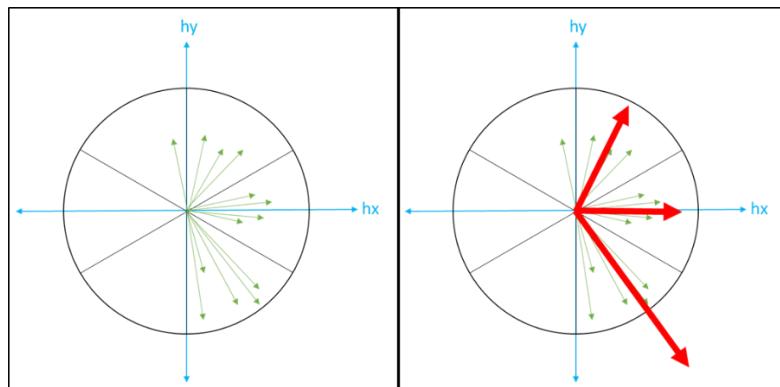


Figura 14. Suma de respuestas de Haar

Después de tener cada uno de los vectores de la respuesta del filtro de Haar en cada sección de $\pi/3$ se suman los vectores para hallar un vector total en cada sección angular. El vector sumatoria que tenga mayor magnitud será el que dará la orientación del punto de interés como se muestra en la figura 15.

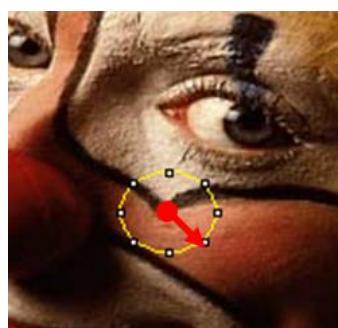


Figura 15. Dirección del punto de interés

- Descripción

La descripción es la etapa del método SURF donde se halla un vector de 64 valores que describen la ubicación del punto de interés. El algoritmo ubica una región cuadrada en el punto de interés de un tamaño 20S [1]. La región cuadrada se subdivide a su vez en 16 cuadriculas como se observa en la figura 16.



Figura 16.Cuadricula 4x4 en punto de interes

Posteriormente el algoritmo toma cada subregión de la cuadricula de 4x4 y la divide de nuevo en una cuadricula de 5x5(figura 17). A cada región de la matriz de 5x5 se le aplica el filtro de Haar en dirección “x” y “y”. El tamaño de este último filtro de Haar es de 2S.

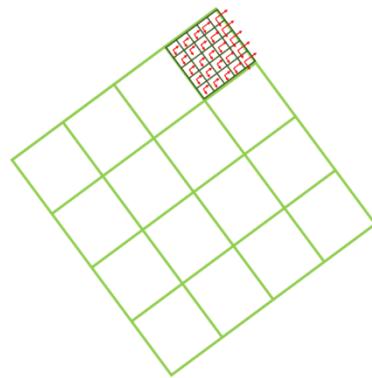


Figura 17.Subcuadricula de 5x5

Después de tener la respuesta de Haar, para cada una de las 16 subregiones se suman las respuestas en “x” y en “y” por separado. El mismo proceso se realiza para la sumatoria de

respuestas absolutas en “x” y “y”.

$$\text{en 1 de las 16 subregiones } hx = \sum hx \quad ECU 10$$

$$\text{en 1 de las 16 subregiones } hy = \sum hy \quad ECU 11$$

$$\text{en 1 de las 16 subregiones } |hx| = \sum |hx| \quad ECU 12$$

$$\text{en 1 de las 16 subregiones } |hy| = \sum |hy| \quad ECU 13$$

El paso siguiente del algoritmo es concatenar cada valor en un vector de 4 valores de la forma:

$$V = (\sum hx, \sum hy, \sum |hx|, \sum |hy|) \quad ECU 14$$

Finalmente se concatena cada vector de cada una de las 16 regiones obteniendo un vector de 64 valores, que se normaliza a suma 1 para mejor interpretación. Este vector será el que describirá la posición del punto de interés y que servirá de ayuda para la última fase de comparación o emparejamiento.

$$Vd = (\sum hx1, \sum hy1, \sum |hx1|, \sum |hy1|, \sum hx2, \sum hy2, \sum |hx2|, \dots, \sum hx16, \sum hy16, \sum |hx16|, \sum |hy16|) \quad ECU 15$$

2.3.3 Emparejamiento

En la etapa de emparejamiento se buscan los puntos de interés que son comunes en dos imágenes de entrada. Para este paso se utiliza una interesante característica del SURF desarrollada en la etapa de detección que fue el signo de la Matriz Hessiana. Si el signo de la matriz hessiana era positivo representaba un punto brillante en un fondo oscuro, y si el signo era negativo representaba un punto oscuro en fondo brillante.

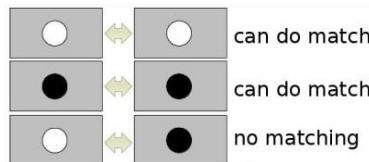


Figura 18. Emparejamiento de puntos de interés

Esta clasificación del método (figura 18) hace que el método sea más rápido y menos costoso computacionalmente pues en el último paso el algoritmo compara el vector descriptor de una imagen con los vectores descriptores de la otra imagen y esto evita realizar operaciones computacionales innecesarias [1].

2.4 Algoritmo ORB

El algoritmo ORB, conocido así por las iniciales de su nombre en inglés: Oriented FAST and Rotated BRIEF, resulta como una fusión del detector de puntos de interés FAST y el descriptor BRIEF. Según los autores, los resultados del algoritmo ORB, muestran que no es afectado significativamente por el ruido de la imagen. El método ORB puede utilizarse para sistemas en tiempo real para habilitar dispositivos de baja potencia sin aceleración GPU, realizar fotografías panorámicas mediante la fusión de varias imágenes y seguimiento de parches en teléfonos móviles [2], y reducir el tiempo de detección de objetos basados en características PC estándar.

2.4.1. Detector FAST

El detector conocido por la sigla FAST, por su nombre en inglés: Features from accelerated segment test, es una técnica que localiza bordes y esquinas. Este método recibe como parámetro el umbral de la diferencia de intensidad entre el píxel central y aquellos situados en un círculo alrededor del centro. Este detector no produce una medida que cuantifique que tanto un punto puede ser considerado como una esquina, y de esta manera también produce altas respuestas a lo largo de bordes y esquinas. Por esta razón, se emplea una medida empleada en la técnica de detección de esquinas de Harris para ordenar los puntos de acuerdo a su importancia, resaltando aquellos correspondientes a esquinas y obtener un número determinado de puntos de interés. Así, en el algoritmo debe establecerse un umbral suficientemente bajo para obtener más de los N puntos de interés deseados, para, una vez hallados, ordenarlos según la medida de Harris y seleccionar únicamente los N puntos más significativos [2].

El método FAST no produce características multiescala. Por esta razón, se emplea una escala piramidal de donde se extraen características FAST, filtradas por la medida de Harris, en cada nivel de la pirámide.

2.4.2. Orientación por intensidad del centro

Para medir la orientación de una esquina, correspondiente a un punto de interés, en el método ORB se emplea como medida la intensidad del centro, en la cual se asume que la intensidad de la esquina está desplazada de su centro y este vector puede usarse para asignarle una orientación. Los momentos de un parche fueron definidos por Rosin mediante la expresión:

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y) \quad ECU 16$$

El centro está expresado por la ecuación:

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad ECU 17$$

Podemos construir un vector desde el centro de las esquinas, O, hasta el centro \vec{OC} . La orientación del parche es:

$$\theta = \text{atan2}(m_{01}, m_{10}) \quad ECU 18$$

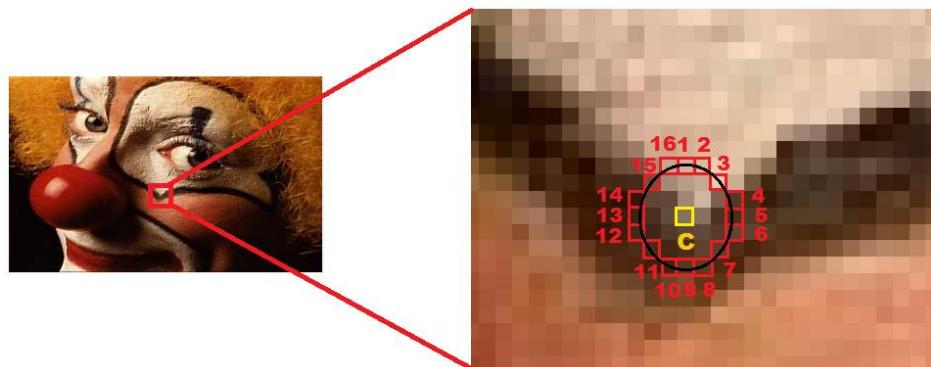


Figura 19. Detección de características mediante círculo de 16 pixeles.

Donde atan2 es la versión compatible con cuadrantes de arctan. La colofonia se debe tener en cuenta si el rincón es oscuro o claro; sin embargo, para nuestros propósitos, podemos ignorar esto ya que las medidas de los ángulos son consistentes independientemente del tipo de esquina. Para mejorar la invariancia de rotación de esta medida, nos aseguramos de que los momentos se calculen con X y Y, que permanecen dentro de una región circular de radio r.

Empíricamente, se escoge que r sea el tamaño del parche, de modo que X y Y se ejecuten desde $[-r, r]$. Como $|C|$ se acerca a 0, la medida se vuelve inestable; Con las esquinas FAST, se ha visto que esto rara vez sucede.

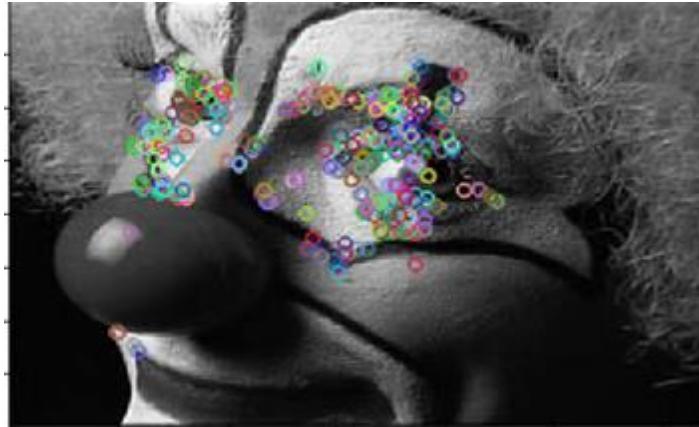


Figura 20. Puntos de interés detectados mediante FAST.

2.4.3. Descriptor BRIEF

El descriptor conocido por la sigla BRIEF, por su nombre en inglés: Binary Robust Independent Elementary Features, es una descripción de la cadena de bits de un parche de imagen, el cual, fue construido a partir de un conjunto de pruebas de intensidad binaria. Considere un parche de imagen suavizado, p . Una prueba binaria τ se define por:

$$\tau(p: x, y) := \begin{cases} 1 & : p(x) < p(y) \\ 0 & : p(x) \geq p(y) \end{cases} \quad ECU\ 19$$

Donde $p(x)$ es la intensidad de p en un punto x . La característica se define como un vector de n pruebas binarias:

$$f_n(p) := \sum_{1 \leq i \leq n} 2^{i-1} \tau(p: x_i, y_i) \quad ECU\ 20$$

También proporciona una alta tasa de emparejamiento entre puntos, pero falla fuertemente cuando una imagen es levemente rotada con respecto a otra. Con el fin de hacer el descriptor BRIEF invariante a la rotación, los autores sugieren orientar el BRIEF en la misma dirección de los puntos clave [2]. Para ello, sea cualquier conjunto de características de n pruebas binarias en la ubicación (x_i, y_i) , se define la matriz $2 \times n$ como:

$$S = \begin{pmatrix} X_1 & \dots & X_n \\ Y_1 & \dots & Y_n \end{pmatrix} \quad ECU\ 21$$

Usando la orientación de parche θ , y la matriz de rotación R_θ correspondiente, construimos una versión S_θ de S:

$$S_\theta = R_\theta S \quad ECU\ 22$$

Ahora el operador BRIEF dirigido se convierte en

$$g_n(P, \theta) := f_n(P)|_{(x_i, y_i) \in S_\theta} \quad ECU\ 23$$

El descriptor BRIEF, presenta una gran variación y una media cercana a 0.5 por cada característica de bit, pero una vez que está orientada a lo largo de la dirección del punto clave, pierde esta propiedad y se distribuye más. La alta varianza hace que una característica sea más discriminativa, ya que responde de manera diferente a las entradas.

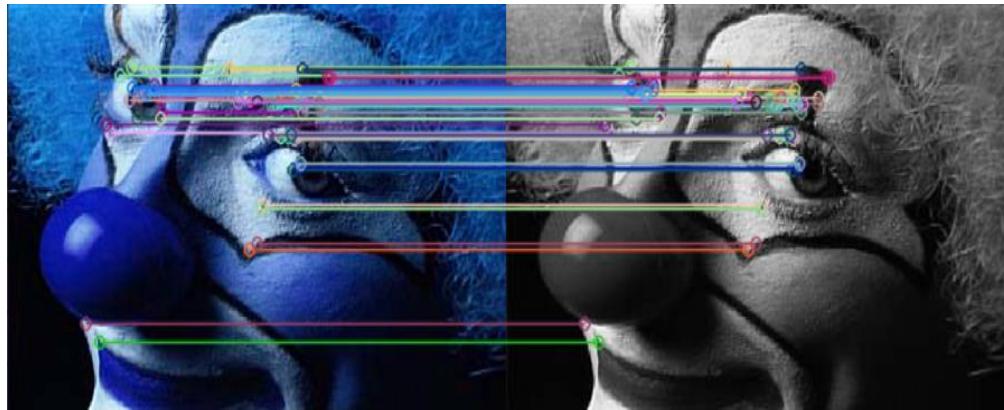


Figura 21. Identificación de puntos de interés con el método ORB

2.5 Filtro de Gabor

El filtro de Gabor toma su nombre del investigador húngaro Dennis Gabor quien propuso, en 1946, un operador lineal que combina una función Gaussiana con una función sinusoidal, tal como ilustra la figura 22. Este filtro es utilizado en análisis de texturas debido a que permite analizar si hay algún contenido de frecuencia específico en la imagen. El filtro permite examinar la frecuencia en direcciones específicas de una región alrededor de un punto. Una de las aplicaciones más comunes del filtro gaussiano es para la detección y procesamiento de huellas

dactilares. La expresión matemática que define la función de Gabor se presenta en la ecuación 24:

$$h(x, y, \theta_k, f, \sigma_x, \sigma_y) = e^{-\frac{1}{2}\left(\frac{x_{\theta_k}^2}{\sigma_x^2} + \frac{y_{\theta_k}^2}{\sigma_y^2}\right)} e^{(i2\pi f x_{\theta_k})} \quad \text{con } k = 1, \dots, m. \quad ECU 24$$

Donde:

$$x_{\theta_k} = x \cdot \cos(\theta_k) + y \cdot \sin(\theta_k)$$

$$Y_{\theta_k} = -x \cdot \sin(\theta_k) + y \cdot \cos(\theta_k)$$

Y f es la frecuencia de onda sinusoidal, m es la cantidad de orientaciones del Gabor, θ_k hace referencia a la enésima orientación del filtro Gabor, σ_x es la desviación estándar de la gaussiana a lo largo del eje x, y σ_y es la desviación estándar de la gaussiana a lo largo del eje y.

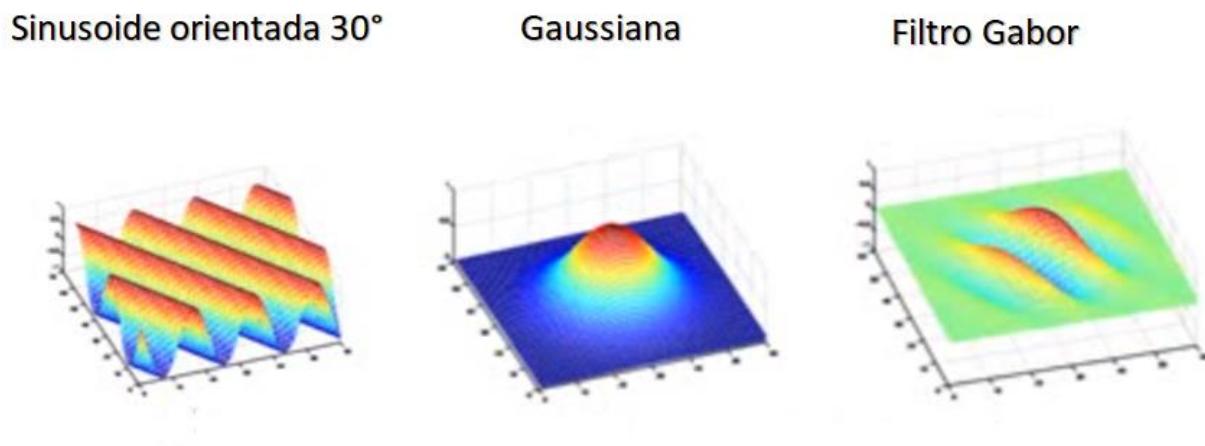


Figura 22. Forma de la Función de Gabor

El filtro de Gabor es direccional y convoluciona con la imagen en diferentes ángulos determinados por el usuario. De esta manera, se obtiene un resultado en cada ángulo escogido, realizando las señales que aparecen en la misma dirección y frecuencia a la del filtro. Dado que el filtro es simétrico, basta con aplicar el filtro únicamente entre 0 y 180 grados. La figura 23 presenta los resultados obtenidos al convolucionar tres imágenes con el filtro Gabor en diferentes ángulos y frecuencias.

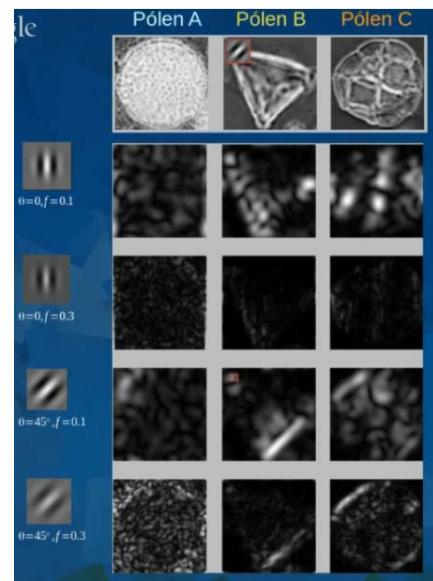


Figura 23. Resultados de la aplicación del método Gabor en distintos ángulos.

Dos ejemplos de máscaras de aproximación al filtro Gabor son:

$$\begin{bmatrix} -1 & -0.5 & 1 \\ -0.5 & 0.8 & -0.5 \\ 1 & -0.5 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix}$$

3. Desarrollo

Después de realizar el análisis teórico de las técnicas SURF y ORB se realizó una búsqueda exhaustiva de códigos con el fin de evaluarlos y verificar si podían ser usados directamente en esta investigación. Se encontraron códigos en java y python, los cuales fueron analizados, mejorados o modificados para permitir el estudio comparativo de estas técnicas. En este capítulo se describen en detalle los pasos seguidos para el desarrollo de los métodos y el proceso de evaluación.

3.1 Método SURF

Un código fuente del algoritmo SURF, desarrollado por Eugen Labun, fue empleado como base para el análisis de este método. El código fue ampliamente modificado, haciéndolo mucho más eficiente y permitiendo la comparación directa entre dos imágenes. El método está implementado en lenguaje Java como plugin del programa de libre acceso ImageJ, empleado en las investigaciones del semillero Lún. La implementación sigue todos los pasos del método original propuesto por los autores Hebert Bay, Tinne Tuytelaars y Luc Van Gool [1]. El plugin recibe como parámetros los valores presentados en la tabla1, los cuales como puede observarse aparecen en inglés, puesto que permite que este código tenga mayor divulgación.

Tabla 1. Valores de los parámetros de entrada del método SURF

Parámetro	Descripción	Valor
Octaves	Número de octavas.	4
Layers Per Octaves	Número de escalas por octava.	4
Hessian threshold	Umbral empleado para evaluar el determinante de la matriz Hessiana. Si éste es superior al umbral, éste es considerado de interés.	0,001
Initial Sampling Step	Valor inicial de muestreo, el cual, duplica las octavas.	2

El algoritmo SURF encontrado no permitía observar los puntos de interés emparejados entre las dos imágenes áreas donde se establecían los puntos de interés, ya que se detectaban en gran cantidad, por esta razón, se realizaron cambios en el algoritmo para una mejor visualización en la identificación y emparejamiento de puntos de interés. De esta manera, el plugin traza una línea entre un punto en una imagen y el punto más cercano en la otra.



a). Descriptor

b) Emparejamiento de objetos

Figura 24. Detección y descripción del algoritmo SURF

3.2 Método ORB

La implementación del algoritmo ORB es bastante compleja debido a que se requiere el uso de librerías externas que presentaron errores de compatibilidad con ImageJ. Por lo tanto, sólo se completó el detector FAST y se empleó el método ORB de la librería openCV sobre Python 2.7 para la evaluación. Se realizaron modificaciones en el código original con el fin de medir el tiempo de ejecución del algoritmo y manipular la cantidad de puntos de interés graficados, dado que en la versión original eran demasiado difícil de distinguir, lo que impedía la interpretación de los resultados.

Las variables empleadas para analizar el rendimiento del algoritmo fueron: cantidad y calidad de los puntos de interés detectados y el tiempo de ejecución del método. Los parámetros de los métodos SURF y ORB fueron los mismos sugeridos por los autores. Los parámetros del método ORB utilizados se presentan en la tabla 2.

Tabla 2. Parámetros de entrada del método ORB

PARAMETRO	DESCRIPCIÓN	VALOR
nfeatures	Número máximo de características encontradas por el algoritmo.	500
 scaleFactor	índice en que se reduce la escala en cada nivel de las pirámide	1.2
nlevels	número de niveles de la pirámide del ORB	8
edgeThreshold	umbral para la detección de esquinas del filtro Harris	31
WTA_K	número de pixeles a evaluar en la etapa de detección	2
patchSize	número de pixeles de la máscara	31

3.3 Protocolo de evaluación

El proceso de evaluación se realizó con una imagen estándar, clown, escogida debido a que contiene un número importante de posibles puntos de interés, falsos (en el cabello) y verdaderos, y a su tamaño, de 320 x 200 pixeles, lo cual permite una fácil visualización de los resultados. Además, se utilizaron 49 pares de imágenes aéreas de cultivos tomadas en infrarrojo cercano y en espectro visible, para estudiar la capacidad de estos métodos para responder correctamente en este tipo de tarea, tema de esta investigación.



Figura 25. De Izquierda a derecha: imagen estándar (Clown), imagen NIR, imagen VIS

En la primera sección, las pruebas consistían en someter las imágenes a cambios rotacionales y alteraciones de sus propiedades para aplicar el método y observar su robustez a los cambios. Las alteraciones a las que se sometió la imagen fueron: cambio a escala de grises (ByN), cambio de escala, espejo horizontal, Rotación, Traslación y Ruido Gaussiano.

El procedimiento consistió aplicar el método sobre dos imágenes: la imagen estándar y una copia de esta pero con una de las alteraciones mencionadas anteriormente. Con el objetivo de que el matching nos diera como resultado el número de puntos de interés congruentes en ambas imágenes. Las variables que se utilizaron para analizar la robustez y el rendimiento del algoritmo fueron: cantidad de puntos de interés detectados y el tiempo de ejecución del método.

Los parámetros de los métodos SUR y ORB fueron los mismos que plantearon los autores para una comparación exacta de los métodos estudiados. Se utilizó la misma imagen de entrada y el mismo computador para que las pruebas fueran lo más equivalentes posibles. La segunda sección consistió en aplicar las técnicas SURF y ORB en 98 imágenes de cultivos VIS y NIR. Los datos de las variables mencionadas anteriormente fueron posteriormente analizados para después comparar los resultados de ambos métodos.

4. Resultados

A continuación, se encuentra los resultados de la ejecución de los algoritmos SURF y ORB en la imagen Clown con los valores de parámetros predeterminados anteriormente. Los datos a evaluar son: cantidad de puntos de interés y tiempo de ejecución del algoritmo.

4.1 Análisis De Resultados Método SURF



Figura 26. Líneas de congruencia entre la imagen original a la imagen blanco y negro.

Cabe resaltar, que el algoritmo SURF recibe imágenes de cualquier tipo, sea a color o blanco y negro, por tal motivo, no muestra error alguno en la comparación, ya que se puede apreciar en la figura 26, que los puntos de la imagen original concuerdan con la imagen ByN, también se logra visualizar un gran número de puntos de interés. Los resultados se encuentran en la tabla 3.

Tabla 3. Resultados obtenidos del algoritmo SURF

Parametro	Cantidad de Congruencias de Puntos de Interés	Tiempo de respuesta (ms)
Blanco y Negro	86	266

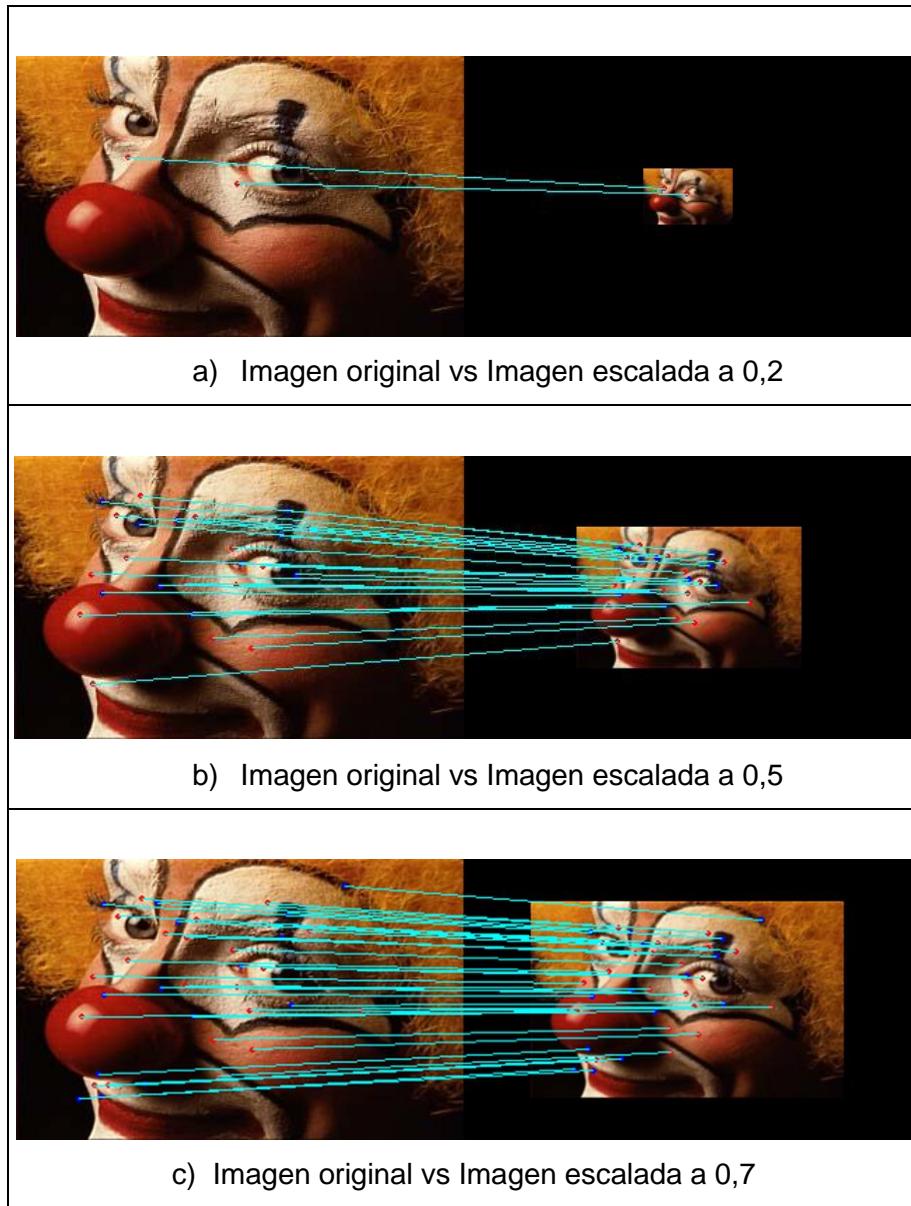


Figura 27. Líneas de congruencia entre la imagen original y la imagen escalada a un tamaño con porcentaje menor.

De acuerdo a la figura 27, cuando se escala una imagen a un tamaño menor, disminuye los números de puntos de interés, esto se debe a que la imagen escalada reduce su tamaño obteniendo menos espacio para desplazar las octavas, por tal motivo, encuentra menos puntos de interés. Los datos se registran en la siguiente tabla.

Tabla 4. Resultados obtenidos del algoritmo SURF en cambio de escala

Parametro (Imagen Clown)	Cantidad de Congruencias de Puntos de Interés	Tiempo de respuesta (ms)
Escalada 20%	2	35
Escalada 50%	37	50
Escalada 70%	25	42



Figura 28. Líneas de congruencia entre la imagen original y la imagen en espejo horizontal

En la figura 28 se denota un cambio drástico en la detección de puntos de interés, que a comparación de las demás pruebas, esta presenta un único punto clave, esto se debe al corrimiento inverso de pixeles, que forman un efecto de espejo, por tal razón, las escalas de grises no son congruentes, a pesar de esto, el algoritmo SURF alcanza a detectar un punto de interés, y se debe a que en esa zona el pixel resalta su valor, logrando así una única coincidencia, esta prueba demuestra la robustez que otorga el método SURF.

Tabla 5. Resultados obtenidos del algoritmo SURF al cambio de espejo horizontal

Parametro (Imagen Clown)	Cantidad de Congruencias de Puntos de Interés	Tiempo de respuesta (ms)
Espejo	1	46

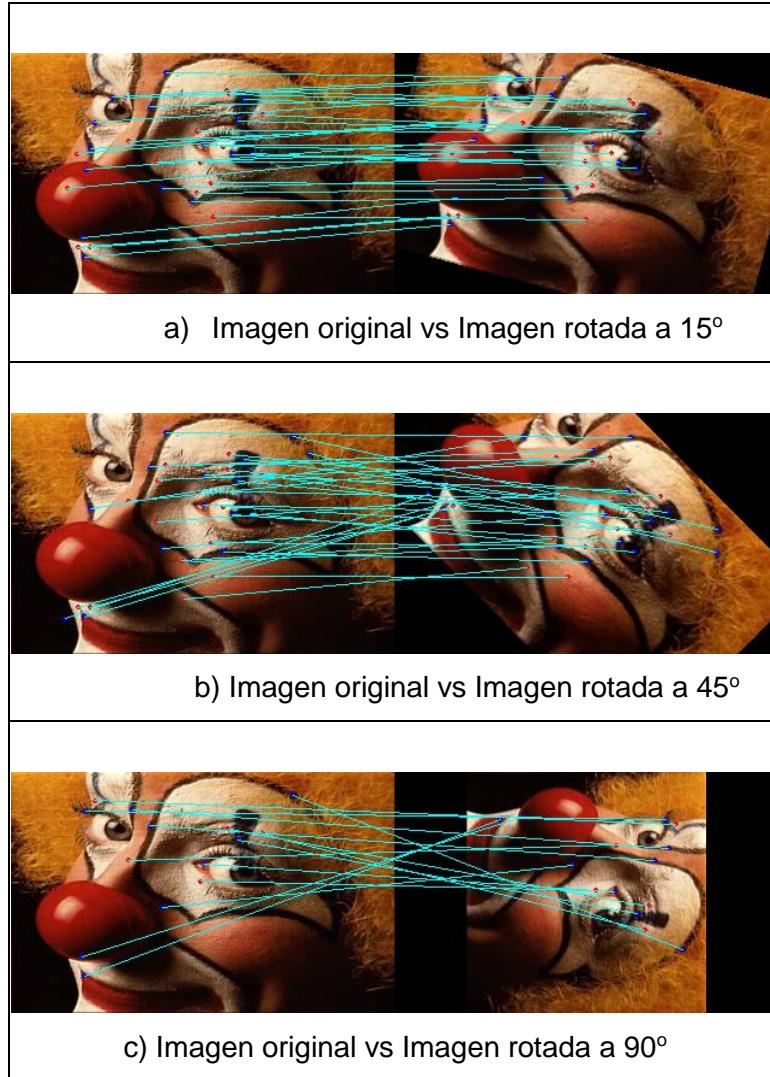


Figura 29. Líneas de congruencia entre la imagen original y la imagen con cambios de rotación

En la figura 29, se puede notar que cuando se rota la imagen de 15° y 45°, los puntos de interés no disminuyen drásticamente, tan solo cuando la imagen se establece en los 90°, se pierden 17 puntos de interés, esto se debe al cambio de posición entre píxeles vecinos, donde la imagen detecta puntos de interés en otras zonas de la imagen. Los datos se registran en la siguiente tabla.

Tabla 6. Resultados obtenidos del algoritmo SURF al cambio de espejo horizontal

Parametro (Imagen Clown)	Cantidad de Congruencias de Puntos de Interés	Tiempo de respuesta (ms)
Rotación 15°	31	52
Rotación 45°	31	85
Rotación 90°	14	56

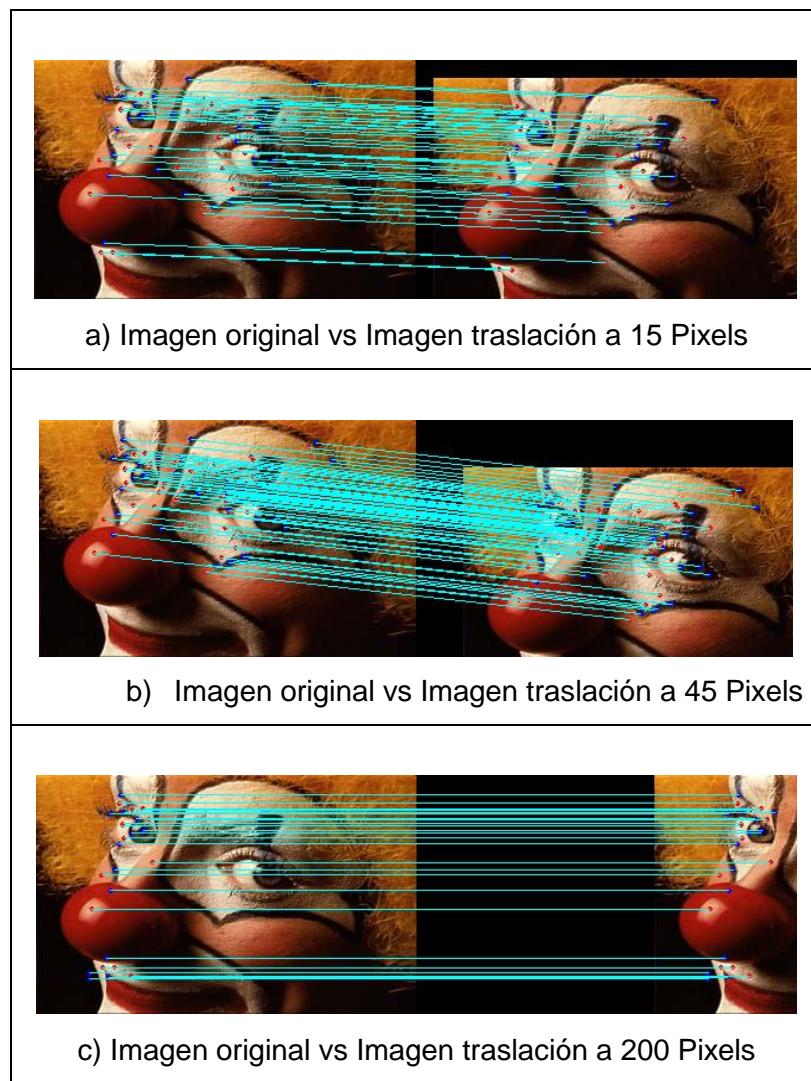
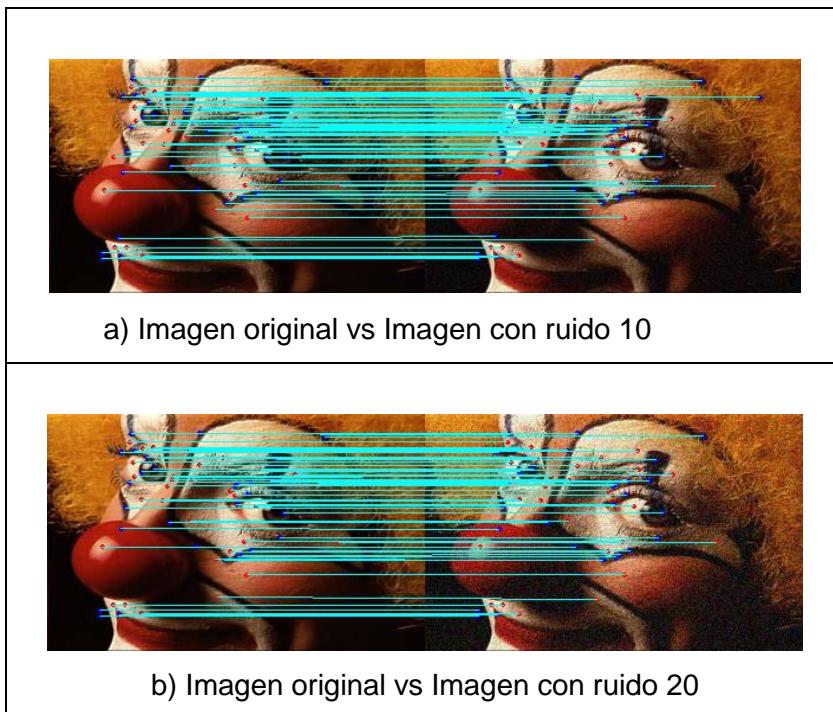


Figura 30. Líneas de congruencia entre la imagen original y la imagen con cambios de translación en pixeles

En la figura 30, la cantidad de puntos de interés son de gran proporción, y cuando la imagen se traslada por pixels, al algoritmo realiza de nuevo una búsqueda para actualizar los puntos de interés y buscar nuevas zonas donde se establecen los puntos de interés para luego marcar su congruencia. Los datos se encuentran en la siguiente tabla.

Tabla 7. Resultados obtenidos del algoritmo SURF al cambio de translación de la imagen

Parametros (Imagen Clown)	Cantidad de Congruencias de Puntos de Interés	Tiempo de respuesta (ms)
Traslación 15	49	58
Traslación 45	71	57
Traslación 200	28	41



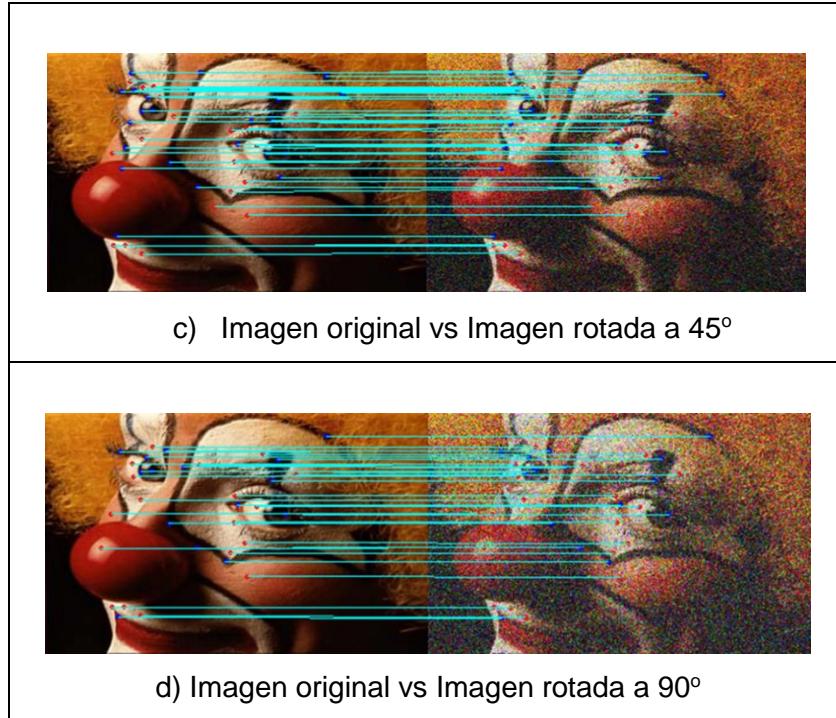


Figura 31. Líneas de congruencia entre la imagen original y la imagen con ruido gaussiano.

En la figura 31, se observa una gran cantidad de puntos de interés, que a pesar de las perturbaciones, las líneas de congruencia no presentan falencia. Los datos se encuentran en la siguiente tabla.

Tabla 8. Resultados obtenidos del algoritmo SURF en respuesta a una imagen con perturbación de ruido gaussiano.

Parametro (Imagen Clown)	Cantidad de Congruencias de Puntos de Interés	Tiempo de respuesta (ms)
Ruido 10%	76	54
Ruido 20%	62	56
Ruido 40%	44	57
Ruido 60%	33	52

Durante el desarrollo de las pruebas de rendimiento, se logró observar que el tiempo de respuesta por cada parámetro no tuvo mayor variación, tan solo en la primera prueba blanco y negro, esto se debe a que el algoritmo requiere de más tiempo para procesar la imagen de un

tipo diferente y convertirlo a ByN. En la siguiente gráfica, se observa la distribución de puntos de interés de cada parámetro calculado.

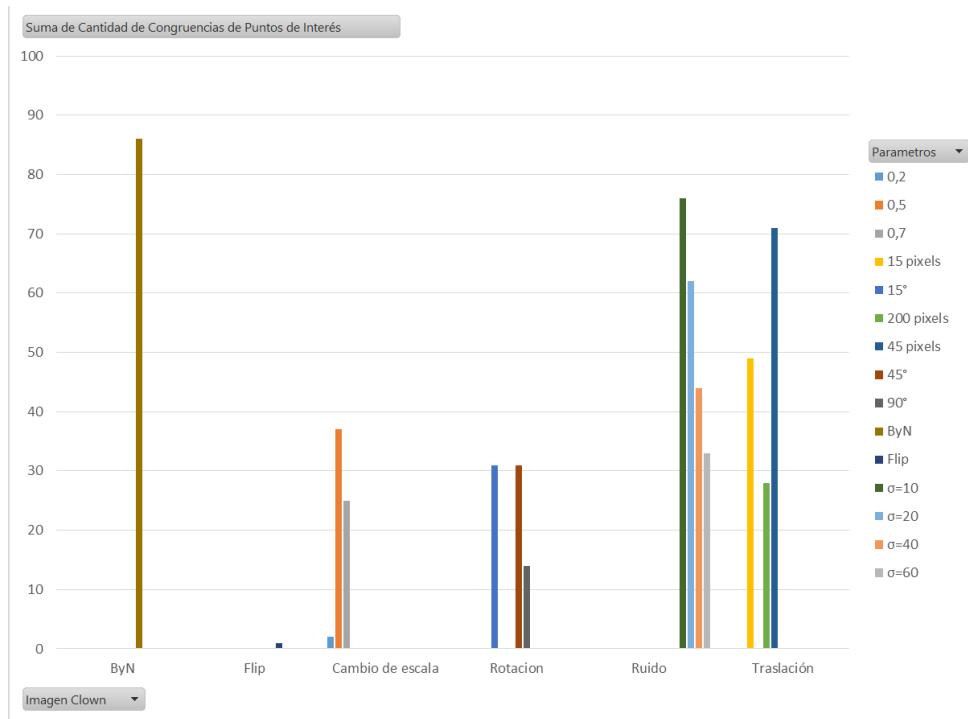


Figura 32. Grafica resultante de la cantidad de puntos de interés del algoritmo SURF

En la figura 32, la imagen con cambio de escala de grises ByN presenta la mayor cantidad de puntos de interés detectados, esto se debe a que el algoritmo SURF trabaja con esquinas y bordes, que representan un cambio de intensidad entre pixeles, mas no se basa en el tipo de niveles de gris de la imagen. Por otro lado, el dato espejo horizontal presenta un único punto, siendo el parámetro con la menor cantidad detectada de puntos de interés.

4.2 Análisis De Resultados Método ORB

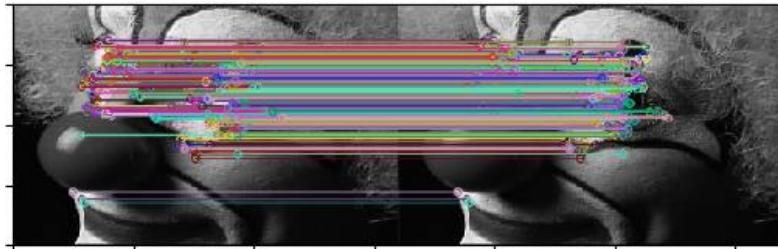


Figura 33. Líneas de congruencia detectadas por el algoritmo ORB en blanco y negro.

La prueba de utilizar la imagen de entrada en formato escala de grises (figura 33) en vez de RGB para el método ORB no tiene mucha repercusión pues aunque este método recibe todo tipo de imágenes, la primera parte del algoritmo convierte la imagen RGB a escala de grises para que las operaciones sean más rápidas y sencillas. Por lo tanto el emparejamiento equivaldría a comparar la misma imagen consigo misma como se observa en la figura 33. La cantidad de puntos de interés para este caso fue de 409 sin ningún falso positivo y el tiempo de ejecución fue de 1.24 segundos.

La prueba de cambio de escala se realizó para establecer el comportamiento del algoritmo cuando se reduce el tamaño de la imagen. Para esto se utilizó la copia exacta de la imagen de entrada pero escalada al 70%, 50% y 20% de su tamaño original.

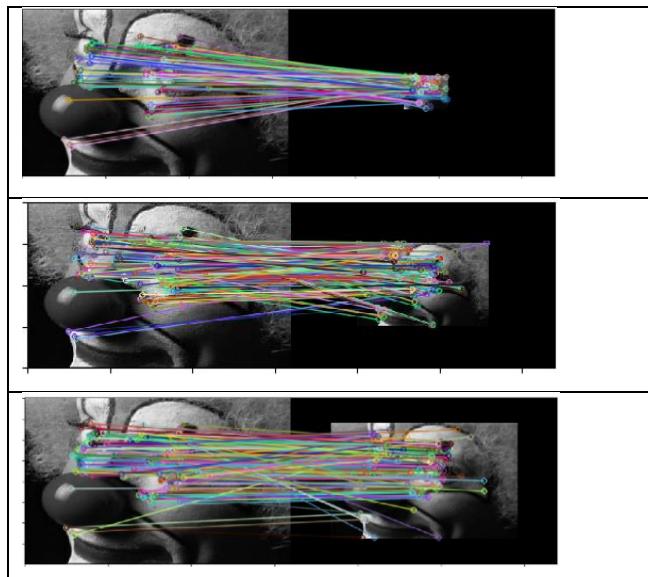


Figura 34. Similitud de puntos de interés entre imágenes escaldadas por el algoritmo ORB.

En la figura 34, se puede evidenciar la congruencia de los puntos de interés para cada una de las imágenes escaladas, donde se observa que los puntos de interés disminuyeron su cantidad. Además se puede observar que el algoritmo detecta falsos positivos en mayor proporción a medida que la escala va disminuyendo. También se observa una disminución sustancial del tiempo de ejecución del algoritmo a medida que la imagen reducía su tamaño.

Tabla 9. Resultados Obtenidos del Algoritmo ORB al cambio de niveles de grises ByN

Parámetro	Puntos de interés emparejados	Falsos Positivos	Tiempo(s)
Escalada al 20%	90	22	0.98
Escalada al 50%	187	15	1
Escalada al 70%	210	12	1.02

La prueba siguiente fue el espejo horizontal de la imagen pues varios métodos de identificación de puntos de interés tienen problemas de congruencia en esta prueba como se pudo evidenciar en el SURF donde solo se encontró un punto.

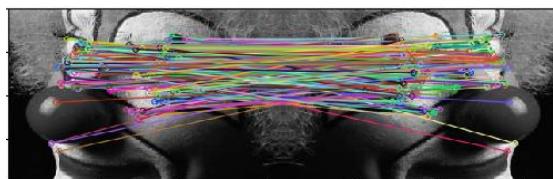
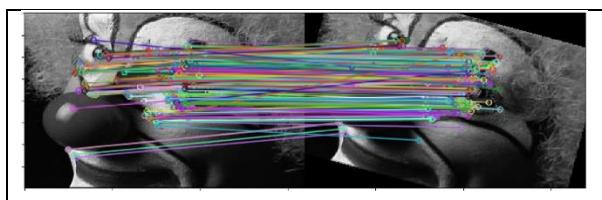


Figura 35. Similitud de puntos de interés en una imagen con espejo horizontal por el algoritmo ORB.

El método ORB arroja resultados muy buenos en esta prueba (figura 35) pues detectó una gran cantidad de puntos de interés. Se obtuvo un total de 198 puntos emparejados correctamente y 18 falsos positivos en un tiempo de 0.99 s.

Para la prueba de rotación, una de las más interesantes pues los autores aseguran que el método es invariante a la rotación, se utilizaron 3 ángulos ($15^\circ, 45^\circ, 90^\circ$).



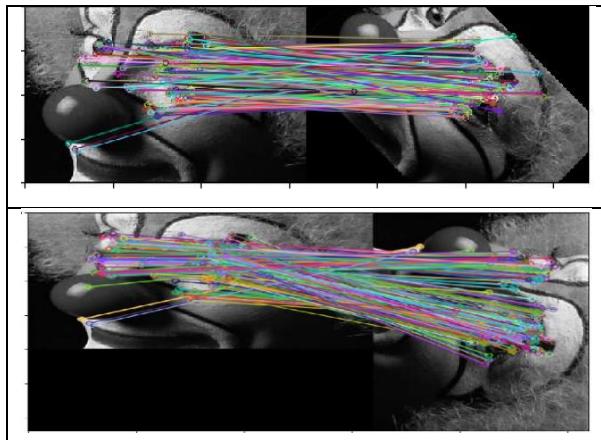


Figura 36. Similitud de puntos de interés en imágenes con diferentes ángulos de rotación por el algoritmo ORB

La figura 36 muestra el comportamiento de el algoritmo a cada una de las rotaciones establecidas en las imagenes de prueba. Los resultados estadísticos muestran que el método detecto gran cantidad de puntos de interés verdaderos en los tres ángulos, comprobando su eficacia a las rotaciones. Otra característica destacada fue que el método encuentra menor cantidad de falsos positivos cuando la imagen esta sobre los ejes cartesianos.

Tabla 10.Resultados Obtenidos del algoritmo ORB a cambios de rotación de la imagen

Parámetro	Puntos de interés emparejados	Falsos Positivos	Tiempo(s)
Rotación 15°	282	7	1.02
Rotación 45°	236	5	1.09
Rotación 90°	398	3	0.949

Para la prueba de translación (figura 37) se desplazó la imagen tres distancias distintas, la primera se desplazó 15 pixeles a la derecha y 15 hacia abajo, la segunda 45 pixeles en ambos sentidos y la ultima 200 pixeles hacia la derecha.



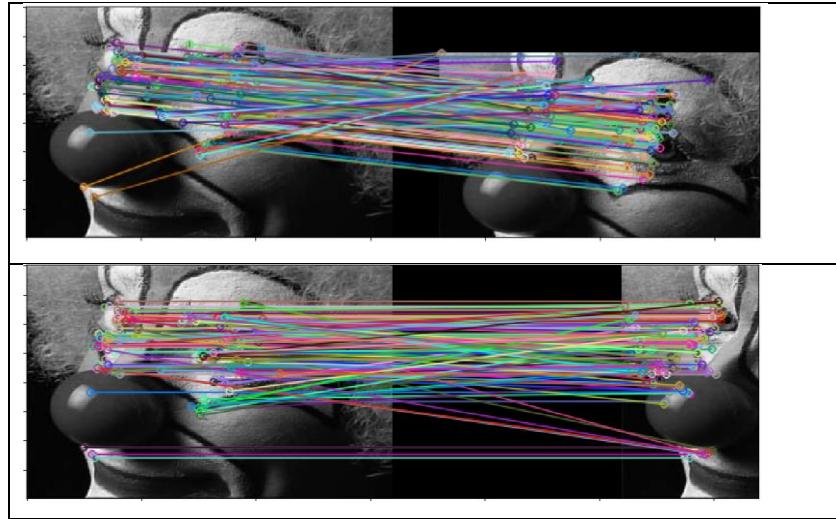


Figura 37. Similitud de puntos de interés en imágenes con traslación de pixeles por el algoritmo ORB

La respuesta a la translación también fue óptima pues se encontraron cantidades muy grandes de puntos de interés verdaderos como se observa en la tabla 11. Hay que resaltar que uno de las consecuencias de la translación es que se pierde información original de la imagen por lo tanto el algoritmo debe buscar nuevas zonas con puntos de interés y correlacionar la mayor cantidad posible por lo tanto a medida que la imagen se desplazaba más aumento el número de falsos positivos pues ubicaba la mayoría de puntos en el cabello del payaso que tiene bordes y esquinas muy similares.

Tabla 11. Resultados Obtenidos del algoritmo ORB a cambios de rotación de la imagen

Parámetro	Puntos de interés emparejados	Falsos Positivos	Tiempo(s)
Traslación 15 pixeles	312	7	1.02
Traslación 45 pixeles	292	8	1.09
Traslación 200 pixeles	147	14	0.949

Para finalizar se evaluó la incidencia del ruido en la congruencia de imágenes. Para esta última prueba se contamino la imagen con ruido gaussiano variando la desviación estándar (σ) del ruido en 10, 20, 40 y 60 unidades.

Los resultados demostraron que el ruido disminuye considerablemente la cantidad de puntos de

interés (figura 37) verdaderos más sin embargo la disminución se comienza a presentar en presencia de un alto índice de ruido. También se encontró que la cantidad de falsos positivos durante las pruebas fue muy pequeño lo que nos lleva inferir que el aumento de ruido no ayuda a la aparición desmedida de falsos positivos. Si se tuviera estas cantidades tan altas de ruido primero la imagen debería ser filtrada con un filtro de ruido gaussiano.

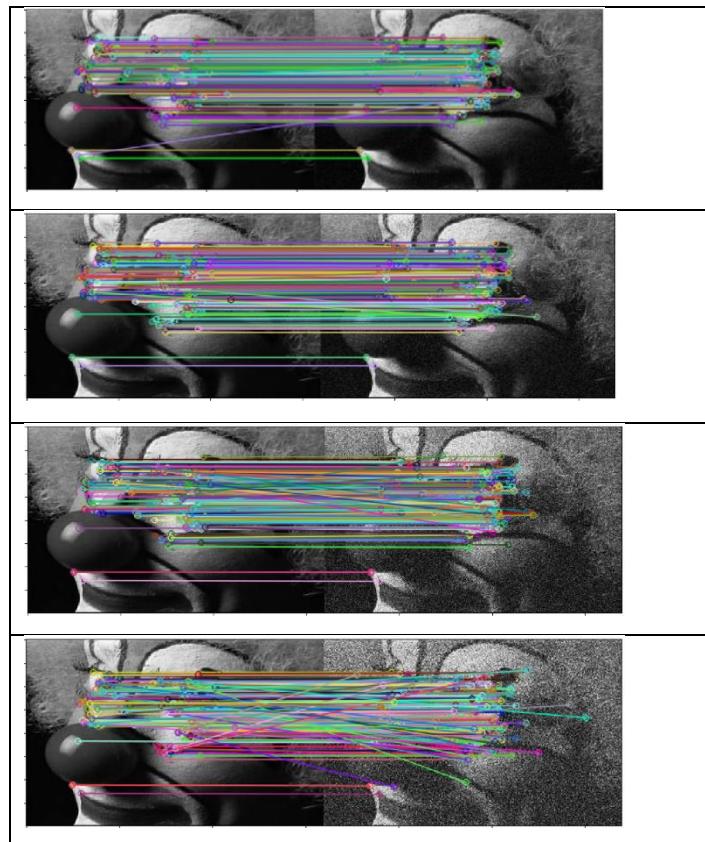


Figura 38. Similitud de puntos de interés en imágenes con perturbaciones de ruido gaussiano por parte del algoritmo ORB

Tabla 12. Resultados Obtenidos del algoritmo ORB a perturbaciones en la imagen como el ruido gaussiano

Parámetro	Puntos de interés emparejados	Falsos Positivos	Tiempo(s)
Ruido con $\sigma=10$	362	1	1.085
Ruido con $\sigma=20$	323	1	0.979
Ruido con $\sigma=40$	280	2	1.013
Ruido con $\sigma=60$	260	6	1.092

Se pudo observar también a lo largo de todas las pruebas de rendimiento que el tiempo de ejecución del algoritmo no tuvieron variaciones exageradas pues los tiempos se movieron en un rango entre los 0.9 y 1.2 segundos.

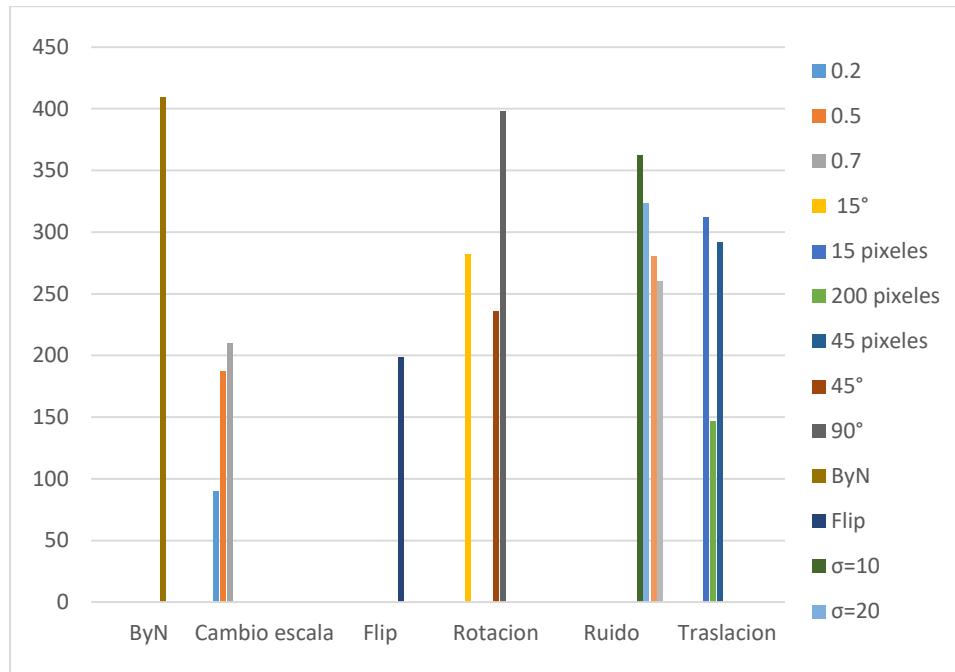


Figura 39. Grafica resultante de la cantidad de puntos de interés del algoritmo ORB

La figura 39 muestra el comportamiento de los puntos de interés en cada una de las pruebas realizadas. Las pruebas donde mayor número de puntos interés verdadero se detectó fueron en Rotación y ByN. La figura 38 comprueba la resistencia del método a la rotación, debido a que fue una de las pruebas que mayor índice de puntos de interés verdaderos presento en cada parámetro experimentado.

Tabla 13. Resultados de las pruebas del algoritmo ORB

Prueba	Puntos interés verdaderos	Falsos Positivos
ByN	409	0
Cambio escala	162.3	16.3
Espejo Horizontal	198	18
Rotación	305.3	5
Ruido	306.2	3.3
Traslación	250.3	9.6

Además, se pudo observar que a una mayor inclinación los puntos de interés verdaderos sufren un decremento, aunque cuando la imagen está totalmente ubicada sobre los ejes cartesianos ($0^\circ, 90^\circ, 180^\circ$ y 270°) el índice de puntos de interés verdaderos aumenta considerablemente con respecto a rotaciones angulares intermedias. De la figura 39 también podemos interpretar que la mayor falencia del método ORB se presentó en el cambio de escala. Al reducir el tamaño en un 20% del tamaño original de la imagen los puntos de interés se redujeron en un 78%. Las pruebas con imágenes ruidosas mostraron una reducción exponencial a medida que la cantidad de ruido gaussiano aumentaba. También cabe resaltar que la prueba de cambio de escala fue la que más falsos positivos introduce.

4.3 Comparación de los métodos en pruebas de rendimiento

Las pruebas de rendimiento nos dan una interpretación del rendimiento individual de los algoritmos. El siguiente paso fue realizar un análisis comparativo de los resultados obtenidos por el método ORB y SURF conjuntamente.

Para realizar la comparación primero se sacó el promedio de los puntos de interés por cada prueba realizada, y de igual forma se promedió el tiempo de cada prueba para realizar una posterior comparación de promedios de puntos de interés y tiempos de ejecución e SURF y ORB.

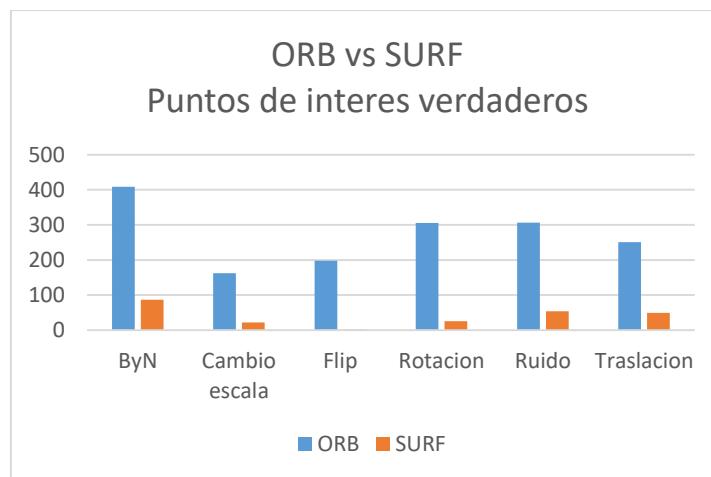


Figura 40. Cantidad de puntos de interés detectados en cada parámetro por los algoritmos

SURF y ORB

La figura 40, representa la cantidad promedio de puntos de interés verdaderos detectados por ambos algoritmos. El método ORB obtuvo una gran cantidad de puntos de interés verdaderos en comparación con la técnica SURF, sin embargo, el algoritmo ORB presento mayor indicie de falsos positivos. Siguiendo con las observaciones, el rendimiento del algoritmo SURF obtuvo malos resultados en tanto a la prueba de Espejo Horizontal, ya que este método detecto un único punto de interés verdadero, el cual se redujo en un 99.5% en comparación con el algoritmo ORB. Los resultados demuestran que el método SURF es más óptimo en cuanto a eficiencia se refiere pues aunque devuelve menor número de puntos verdaderos, en ninguna de las pruebas entrega falsos positivos.

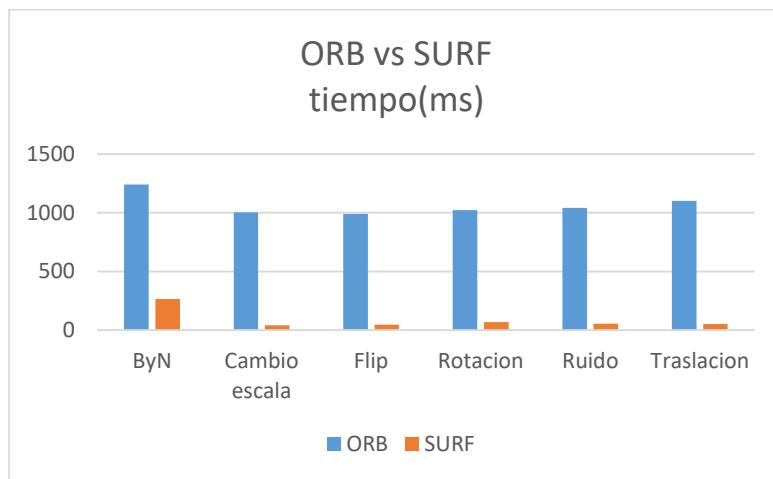


Figura 41. Tiempo de respuesta de los algoritmos SURF y ORB por cada parámetro

En la figura 41, se registra el tiempo de respuesta de las técnicas SURF y ORB al momento de ejecutar su algoritmo para cada parámetro, dando como resultado que el algoritmo SURF tarda menos tiempo en ejecutarse que el algoritmo ORB.

4.4 Experimentación con imágenes aéreas

La experimentación de los métodos ORB y SURF en fotografías aéreas tomadas en el espectro visible e infrarrojo cercano adquirido por parte del CIAT, “Centro Internacional de Agricultura Tropical”, permitió determinar el rendimiento de estos métodos para esos tipos singulares de imágenes. Las pruebas se realizaron en 49 imágenes en espectro visible (VIS) y 49 imágenes

en espectro infrarrojo cercano (NIR). El tamaño de las imágenes es de 774 de ancho por 518 de alto en pixeles, estas imágenes se tomaron desde un dron con ángulo oblicuo. El algoritmo ORB fue aplicado inicialmente a las fotografías aéreas con parámetros de entrada propuestos por los autores para que la técnica fuera la más cercana a la original, por tal motivo, no se realizó pruebas de matriz [R/G/B] independiente Vs NIR, ya que este algoritmo trabaja con imágenes B/N.

4.4.1. Análisis De Resultados Del Método SURF con fotografías aéreas

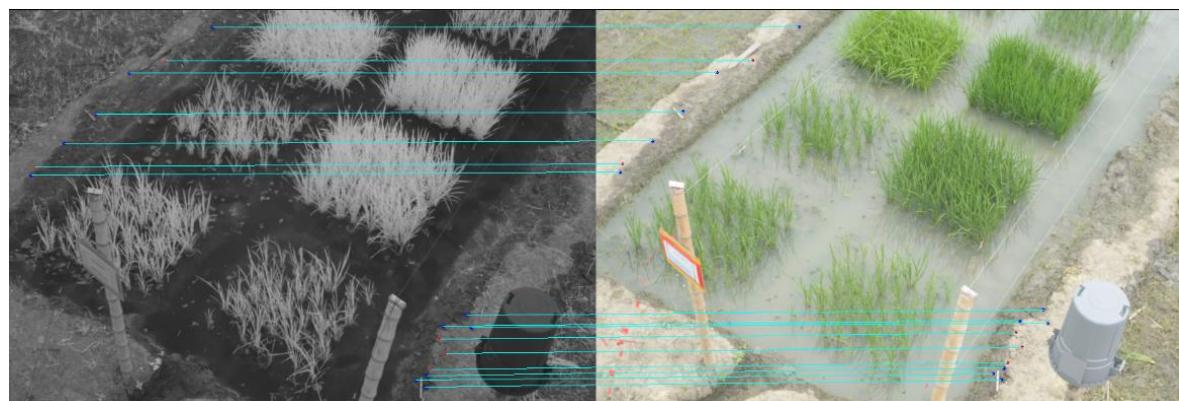
Las imágenes presentes en este apartado, son productos del proceso algorítmico del método SURF. Esta técnica toma dos imágenes en espectros distintos y calcula la mayor cantidad de puntos de interés para después encontrar la congruencia en cada objeto, ya sea pasto, tierra, postes o canecas presentes en las fotografías. Para llevar a cabo estas pruebas, se redujo la escala de las imágenes a un tamaño del 20% en comparación con la imagen original, con el objetivo de facilitar el desarrollo algorítmico de este método y lograr visualizar de forma detallada las líneas de congruencia entre objetos. La cantidad de puntos de interés por parámetro y el tiempo de respuesta se encuentran registrados en tablas.



Figura 42. Resultados de las fotografías de cultivo de pasto con los parámetros: Octaves: 3, Layers Per Octaves: 3, Hessian Thresholded: 0.00010, Initial Sampling Step: 1.

Tabla 14. Resultados de la imagen NIR_A_T1 y VIS_A_T1 con parámetros: Octaves: 3, Layers Per Octaves: 3, Hessian Thresholded: 0.00010, Initial Sampling Step: 1

Resultados de NIR_A_T1							
Octave	Layer	IPCandidates	Thresholded	Suppressed	InterpolationNotSucceed	BadInterpolationResult	Final IPs
0	1	365540	292145	71691	0	0	1704
1	1	83520	63592	19392	0	0	536
2	1	17201	11596	5479	0	0	126
							TOTAL IP
							2366
Resultados de VIS_A_T1							
Octave	Layer	IPCandidates	Thresholded	Suppressed	InterpolationNotSucceed	BadInterpolationResult	Final IPs
0	1	365540	324624	39675	0	1	1240
1	1	83520	72774	10379	0	0	367
2	1	17201	13678	3446	0	0	77
							TOTAL IP
							1684
Emparejamiento de puntos de interés							
38							
Tiempo De Respuesta (s)							
1,809							



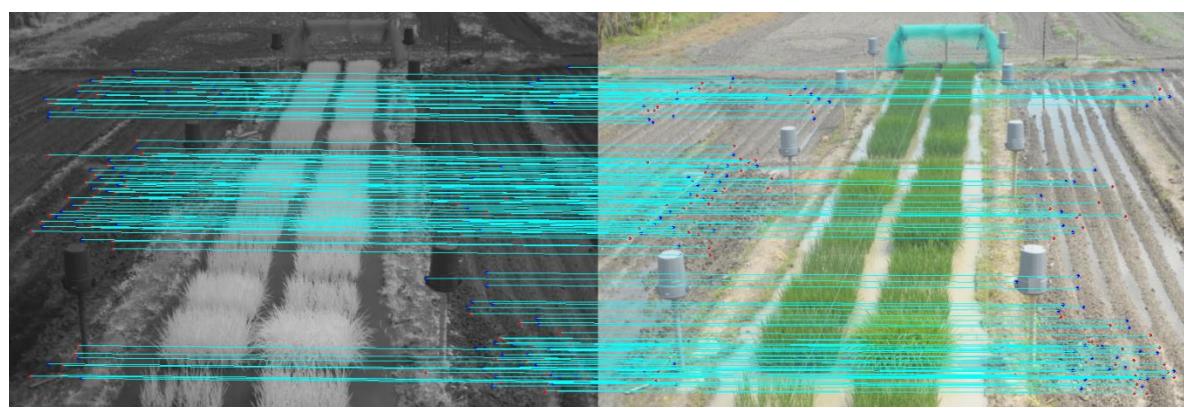
a) NIR_A_T1

b) VIS_A_T1

Figura 43. Resultados de las fotografías de cultivo de pasto con los parámetros: Octaves: 4, Layers Per Octaves: 4, Hessian Thresholded: 0.00010, Initial Sampling Step: 2.

Tabla 15. Resultados de la imagen NIR_A_T1 y VIS_A_T1 con parámetros: Octaves: 4, Layers Per Octaves: 4, Hessian Thresholded: 0.00010, Initial Sampling Step: 2.

Resultados De NIR_A_T1							
Octave	Layer	IP Candidates	Thresholded	Suppressed	InterpolationNotSucceed	BadInterpolationResult	Final IPs
0	1	90768	72166	17588		0	1 1013
0	2	90768	70442	19822		0	0 504
1	1	20585	15618	4666		0	0 301
1	2	20585	14747	5717		0	0 121
2	1	4233	2828	1348		0	0 57
2	2	4233	2492	1718		0	0 23
3	1	665	361	292		0	0 12
3	2	665	320	340		0	0 5
							TOTAL IP 2036
Resultados De VIS_A_T1							
Octave	Layer	IP Candidates	Thresholded	Suppressed	InterpolationNotSucceed	BadInterpolationResult	Final IPs
0	1	90768	80184	9724		0	0 860
0	2	90768	79361	11024		0	0 383
1	1	20585	17867	2490		0	0 228
1	2	20585	17118	3378		0	0 89
2	1	4233	3367	823		0	0 43
2	2	4233	2889	1316		0	0 28
3	1	665	393	264		0	0 8
3	2	665	382	279		0	0 4
							TOTAL IP 1643
Emparejamiento de puntos de interés							
16							
Tiempo De Respuesta (s)							
1,45							



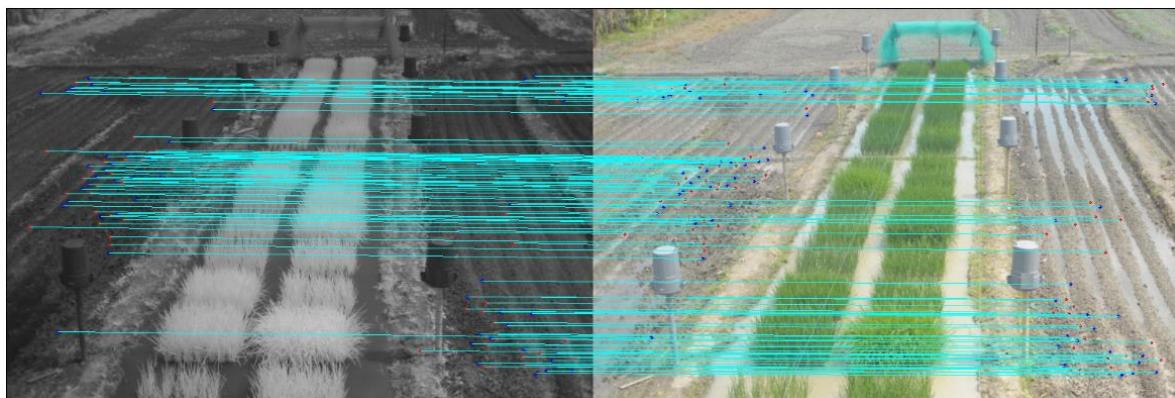
a) NIR_C_T2

b) VIS_C_T2

Figura 44. Resultados de las fotografías de cultivo de pasto con los parámetros: Octaves: 3, Layers Per Octaves: 3, Hessian Thresholded: 0.00010, Initial Sampling Step: 1.

Tabla 16. Resultados de la imagen NIR_C_T2 y VIS_C_T2 con parámetros: Octaves: 3, Layers Per Octaves: 3, Hessian Thresholded: 0.00010, Initial Sampling Step: 1.

Resultados De NIR_C_T2							
Octave	Layer	IPCandidates	Thresholded	Suppressed	InterpolationNotSucceed	BadInterpolationResult	Final IPs
0	1	365540	316133	48245	0	0	1162
1	1	83520	68301	14798	0	0	421
2	1	17201	13020	4086	0	0	95
							Total IP
							1678
Resultados De VIS_C_T2							
Octave	Layer	IPCandidates	Thresholded	Suppressed	InterpolationNotSucceed	BadInterpolationResult	Final IPs
0	1	365540	295082	68874	0	2	1582
1	1	83520	63272	19690	0	0	558
2	1	17201	12536	4534	0	0	131
							Total IP
							2271
Emparejamiento de puntos de interés							
153							
Tiempo (s)							
1,903							



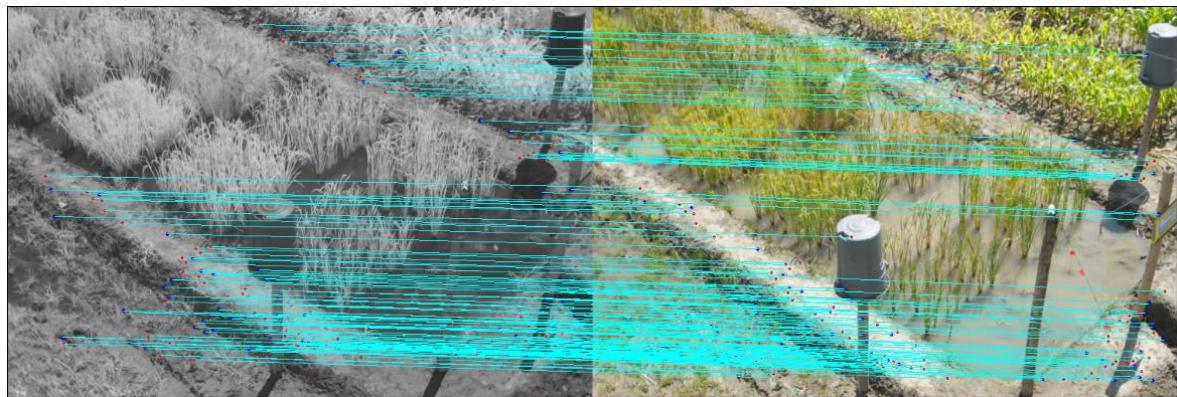
a) NIR_C_T2

b) VIS_C_T2

Figura 45. Resultados de las fotografías de cultivo de pasto con los parámetros: Octaves: 4, Layers Per Octaves: 4, Hessian Thresholded: 0.00010, Initial Sampling Step: 2

Tabla 17. Resultados de la imagen NIR_C_T2 y VIS_C_T2 con parámetros: Octaves: 4, Layers Per Octaves: 4, Hessian Thresholded: 0.00010, Initial Sampling Step: 2.

Resultado De NIR_C_T2								
Octave	Layer	IPCandidates	Thresholded	Suppressed	InterpolationNotSucceed	BadInterpolationResult	Final IPs	
0	1	90768	78509	11503	0	0	756	
0	2	90768	76282	14072	0	1	413	
1	1	20585	16789	3536	0	0	260	
1	2	20585	16192	4267	0	0	126	
2	1	4233	3211	961	0	0	61	
2	2	4233	3056	1147	0	0	30	
3	1	665	444	211	0	0	10	
3	2	665	445	214	0	0	6	
							Total IP	
								1662
Resultado De VIS_C_T2								
Octave	Layer	IPCandidates	Thresholded	Suppressed	InterpolationNotSucceed	BadInterpolationResult	Final IPs	
0	1	90768	73209	16567	0	0	992	
0	2	90768	70951	19324	0	0	493	
1	1	20585	15602	4664	0	0	319	
1	2	20585	15177	5252	0	0	156	
2	1	4233	3083	1076	0	0	74	
2	2	4233	3043	1157	0	0	33	
3	1	665	439	212	0	0	14	
3	2	665	433	225	0	0	7	
							Total IP	
								2088
Emparejamiento de puntos de interés								
111								
Tiempo De Respuesta (s)								
1,31								



a) NIR_H_T6

b) VIS_H_T6

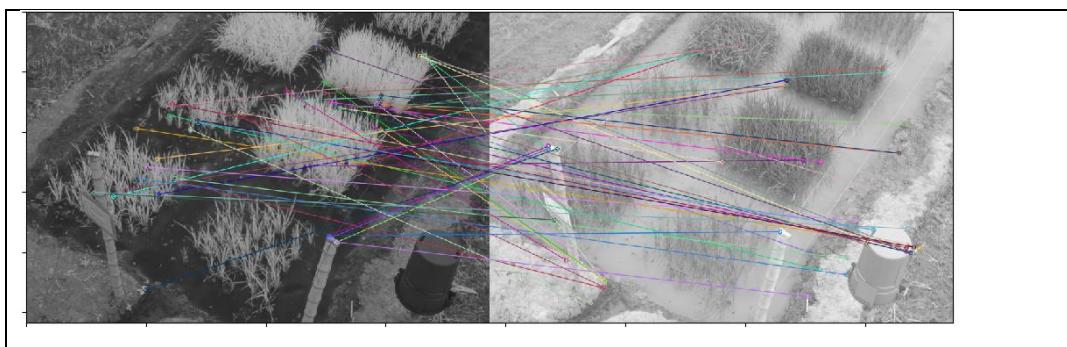
Figura 46. Resultados de las fotografías de cultivo de pasto con los parámetros: Octaves: 3, Layers Per Octaves: 3, Hessian Thresholded: 0.00010, Initial Sampling Step: 1

Tabla 18. Resultados de la imagen NIR_H_T6 y VIS_H_T6 con parámetros: Octaves: 3, Layers Per Octaves: 3, Hessian Thresholded: 0.00010, Initial Sampling Step: 1.

Resultado NIR_H_T6							
Octave	Layer	IP Candidates	Thresholded	Suppressed	InterpolationNotSucceed	BadInterpolationResult	Final IPs
0	1	365540	276699	86900	0	2	1939
1	1	83520	60582	22382	0	1	555
2	1	17201	11264	5822	0	0	115
							Total IP
							2609
Resultado VIS_H_T6							
Octave	Layer	IP Candidates	Thresholded	Suppressed	InterpolationNotSucceed	BadInterpolationResult	Final IPs
0	1	365540	279929	83789	0	1	1821
1	1	83520	62048	20938	0	0	534
2	1	17201	11977	5111	0	0	113
							Total IP
							2468
Emparejamiento de puntos interés							
150							
Tiempo De Respuesta (s)							
2,076							

4.4.2. Análisis De Resultados Del Método ORB

Después de aplicar el método ORB en las imágenes aéreas de los cultivos, los resultados no fueron los esperados. En las pruebas de rendimiento el algoritmo mostro un alto índice de falsos positivos, pero en las imágenes de los cultivos los puntos de interés hallados eran erróneos. La mayoría de los puntos de interés emparejados eran falsos positivos.



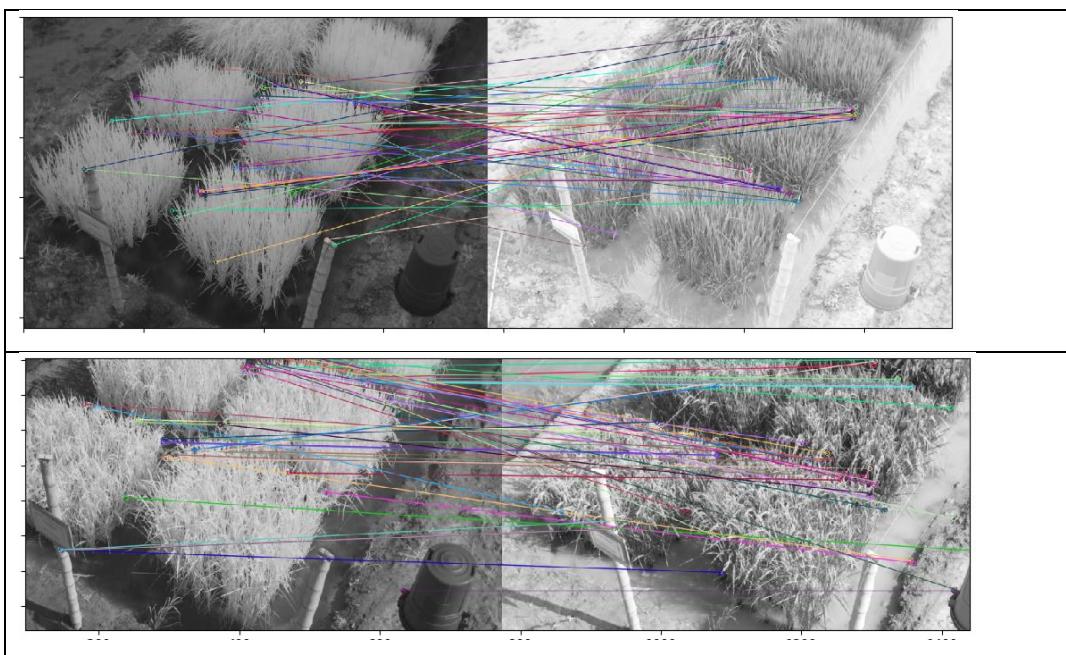


Figura 47. Resultados de las fotografías de cultivo de pasto con los parámetros: nfeatures: 500, scaleFactor: 1.2, nlevels: 8, edgeThreshold: 31, WTA_K: 2, patchSize: 31

Debido a los pésimos resultados del ORB con las imágenes NIR y VIS se modificó los parámetros de entrada del algoritmo para lograr un mejor emparejamiento de los puntos de interés. El parámetro después de la experimentación que mejoró las características de detección de puntos verdaderos fue nlevels, este parámetro determina la cantidad de niveles que debe tener la pirámide. Se redujo el número de escalas a un 20%, esto permitió que el algoritmo filtrara la imagen solo en dos ocasiones. Como consecuencia el método solo puede distinguir los objetos de mayor dimensión de la imagen como palos y canecas que presentan esquinas muy bien definidas y de gran tamaño. Sin embargo solo se lograron 6 coincidencias con el nuevo parámetro por lo que el método siguió siendo ineficiente para la correlación en imágenes VIS y NIR.

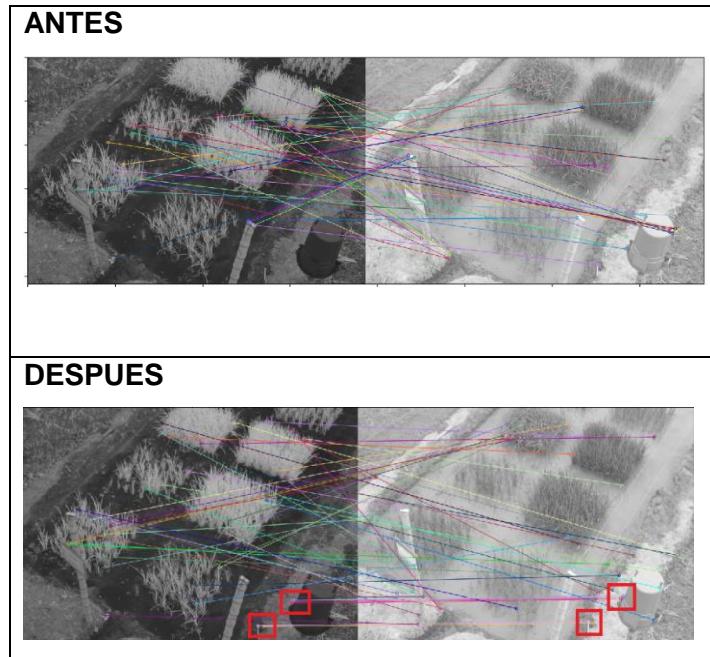


Figura 48. Resultados de las fotografías de cultivo de pasto con los parámetros: nfeatures: 500, scaleFactor: 1.2, nlevels: 2, edgeThreshold: 31, WTA_K: 2, patchSize: 31

5. Descriptor de Gabor

Debido a que los resultados obtenidos con los descriptores del método ORB no fueron apropiados dado que los descriptores de los puntos de interés encontrados en una imagen VIS no coinciden con los hallados en la NIR, se decidió utilizar el descriptor de textura Gabor, con el fin de determinar si la textura si permitía encontrar una similitud entre los puntos de interés que permita su emparejamiento. La textura en imágenes es una región que presenta un patrón común en los pixeles. Este patrón es repetitivo en la microregion de la imagen. Dado que lo que se desea es que poder hallar congruencia entre las imágenes en VIS y NIR se tomaron pequeñas muestras que representarían las regiones u objetos más importantes de las fotografías aéreas. Los objetos a los que se les analizó su textura fueron canecas, palos, agua, pasto y tierra.



Figura 49. Objetos de referencia para la detección de textura Gabor.

Se utilizaron 25 muestras de cada objeto tanto en espectro visible como en infrarrojo cercano. Después se utilizó el filtro de Gabor, el cual arroja como resultado una pila de imágenes donde cada imagen es el resultado del filtro en los ángulos escogidos. En este estudio se tomaron seis ángulos: 30, 60, 90, 120, 150, 180.

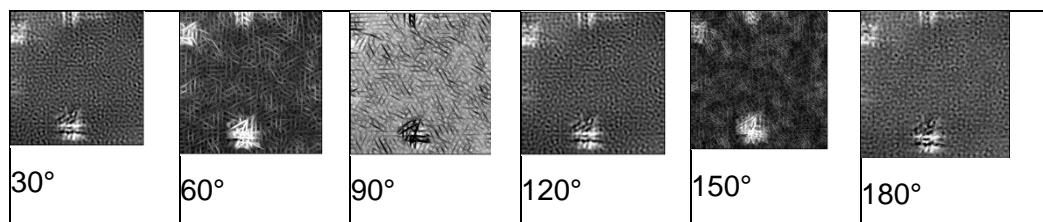


Figura 50. Resultados del algoritmo Gabor en distintos ángulos.

Después de obtener las imágenes resultantes del filtro de Gabor, se desarrolló un código que determinara la media, la desviación estándar y la entropía de todos los pixeles de las 6 imágenes

resultantes. El análisis de los datos entregados por el código, fue realizado en el software de acceso libre dedicado a computación estadística y gráficos RStudio.

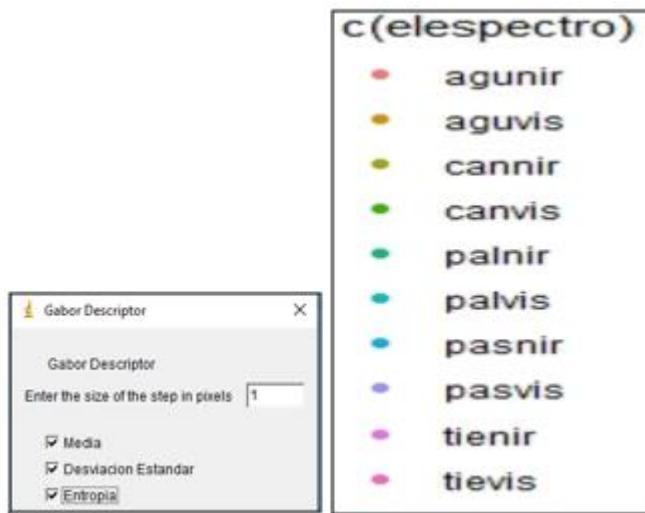


Figura 51. Parámetros del descriptor Gabor

Los datos entregados por el código desarrollado en imageJ se pueden interpretar de una forma más adecuada mediante una diagrama de funciones gaussianas. Se realizó una gráfica para cada ángulo y para cada técnica (media, desviación estándar y entropía) con los resultados de cada objeto (Caneca, Agua, Tierra, Palo, Pasto) tanto en VIS como en NIR.

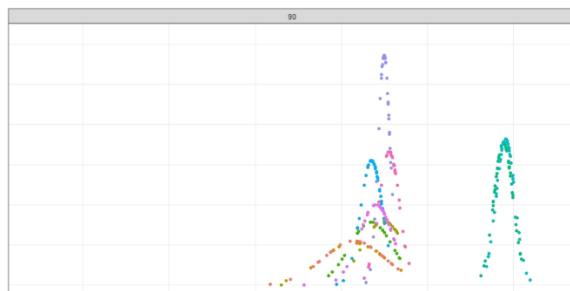


Figura 52. Resultado obtenido de la entropía en las imágenes NIR y VIS

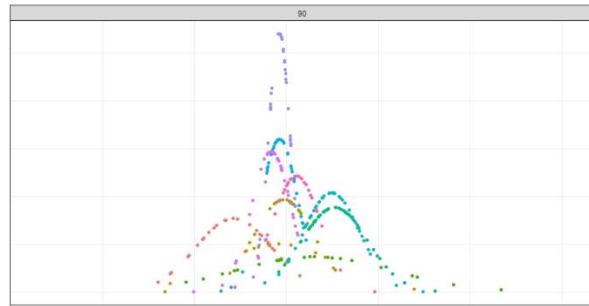


Figura 53. Resultado obtenido de la desviación estándar en cada imagen aérea NIR y VIS

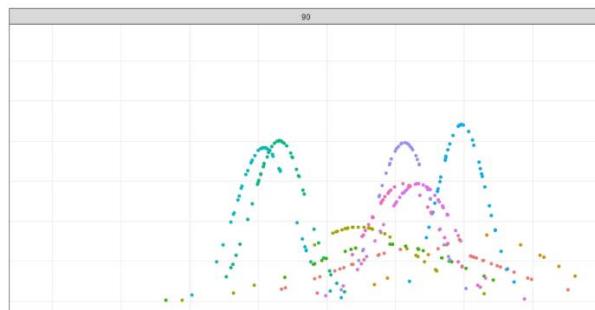


Figura 54. Resultado obtenido de la media en las imágenes aéreas NIR y VIS

Las figuras 52, 53 y 54, presentan los resultados de cada técnica para el ángulo de 90°. Pudimos encontrar que la mejor técnica fue la entropía puesto que permitió distinguir claramente el palo de los demás objetos, y además su textura fue aproximadamente igual tanto en espectro visible como en infrarrojo cercano.

6. Conclusiones y recomendaciones

6.1. Conclusiones

A partir del trabajo desarrollado se obtuvieron las siguientes conclusiones:

- El algoritmo ORB no es adecuado para realizar el registro de imágenes de diferentes modalidades, como fotografías adquiridas en los espectros infrarrojo cercano y visible, pues no logran identificar los mismos puntos de interés en ambas imágenes. Esto debido a que emplea un descriptor de baja calidad a cambio de aumentar su velocidad. Ventaja que ha hecho que el ORB sea recomendado para su uso en tiempo real. Así, el descriptor ORB es de tamaño 32, mientras que el SURF de tamaño 64 y SIFT de 128. Dado que este algoritmo no es apropiado para el estudio propuesto, no se implementó en lenguaje java y se prefirió en su lugar explorar descriptores de textura con el fin de determinar si resultaban más apropiados para emparejar las imágenes de fotografía aérea.
- La comparación de una imagen con su espejo horizontal permitió encontrar que ninguno de los descriptores es invariante a esta transformación. Aun así, el algoritmo SURF logró encontrar una coincidencia entre un punto reflejado y su par, lo cual se debe a su mayor robustez con respecto al ORB, y al parecer debido a que la región detectada tiene una alta simetría, lo que hace que allá una alta coincidencia entre el punto de interés y su reflexión.
- El método ORB es muy sensible al cambio de escala. De manera que al reducir la escala en un 20% de una imagen representó una reducción del 78% en el número de puntos de interés encontrados.
- El método ORB emplea una mayor cantidad de puntos de interés entre imágenes de un mismo tipo, pero con un alto índice de falsos positivos. Por el contrario, el SURF entrega una menor cantidad de puntos, pero más significativos. Esto permite demostrar que al algoritmo SURF es más eficiente y robusto que el ORB.
- Para llevar a cabo las evaluaciones de los métodos SURF y ORB con fotografías aéreas, fue necesario disminuir el tamaño de estas imágenes para lograr una mejor visualización de las líneas de congruencia producidas por los algoritmos. Esto permitió también reducir

notablemente el tiempo de respuesta de los algoritmos. Sin embargo, disminuye también la cantidad de puntos de interés encontrados.

- En imágenes pequeñas el algoritmo SURF es más rápido que el ORB. Sin embargo, en imágenes de gran tamaño el SURF se vuelve notablemente más lento que el ORB.
- La entropía de las imágenes de textura, obtenidas mediante filtros de Gabor, es el descriptor más apropiado, en comparación con la media y la desviación estándar, para caracterizar los postes empleados como puntos de referencia en cultivos en imágenes de los espectros visible e infrarrojo cercano, siendo el detector de puntos SURF el mejor método para la detección de puntos de interés en estas fotografías.

6.2. Recomendaciones

Es necesario implementar mecanismos de identificación que otorgue un orden para cada punto de interés detectado, con el fin de entender visualmente la congruencia entre cada imagen, destacando su área local y cantidad de puntos claves presentes.

Se recomienda utilizar el algoritmo ORB para aplicaciones de video donde el mayor requerimiento sea la velocidad de ejecución del algoritmo.

Se recomienda utilizar descriptores de textura para el tipo de imágenes de cultivos estudiado debido a que logran identificar los mismos postes de referencia en las dos imágenes, por lo cual resulta un poderoso descriptor para este propósito.

Referencias bibliográficas

- [1] Bay, Herbert. Tuytelaars, Tinne. Van Gool, Luc. "SURF: Speeded Up Robust Features".(2008)
- [2] Rublee, Ethan. Rabaud, Vincent. Konolige, Kurt. Bradski, Gary." ORB: an efficient alternative to SIFT or SURF". (2013)
- [3] Heinly, Jared. Dunn, Enrique. Frahm, Jan-Michael. "Comparative Evaluation of Binary Features", The University of North Carolina at Chapel Hill. (2012)
- [4] Kulkarni, A.V. Jagtap, J.S. Harpale, V.K. "Object recognition with ORB and its Implementation on FPGA", International Journal of Advanced Computer, Volume (3). (2013)
- [5] Miksik, Ondrej. Krystian Mikolajczyk." Evaluation of Local Detectors and Descriptors for Fast Feature Matching".(2014)
- [6] Luo, Juan. Oubong, Gwun. "A Comparison of SIFT, PCA-SIFT and SURF, International Journal of Image Processing (IJIP)" Volume(3). (2010).
- [7] Bekele, Dagmawi. Teutsch, Michael. Schuchert, Tobias."Evaluation Of Binary Keypoint Descriptors", Karlsruhe Institute of Technology, Karlsruhe, Germany. (2014)
- [8] Xin, Yang. Kwang, Ting. "Accelerating SURF Detector on Mobile Devices, Dept. of Electrical and Computer Engineering", University of California, Santa Barbara, USA. (2012)
- [9] Karami, Ebrahim. Prasad, Siva. Shehata, Mohamed. "Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images", Faculty of Engineering and Applied Sciences, Memorial University, Canada.(2015)
- [10] Azad, P., Asfour, T., & Dillmann, R. "Combining Harris interest points and the SIFT descriptor for fast scale-invariant object recognition". *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*, 4275–4280. (2009)
- [11] Harris, C., & Stephens, M. "A Combined Corner and Edge Detector". *Proceedings of the Alvey Vision Conference 1988*, 23.1-23.6. (1988).
- [12] Moravec, H."Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover", Tech Report CMU-RI-TR-3, Carnegie-Mellon University, Robotics Institute.(1980)
- [13] T. Lindeberg."Feature detection with automatic scale selection". *IJCV*, 30(2):79 – 116 .(1998)
- [14] D. Lowe."Object recognition from local scale-invariant features". In *ICCV*. (1999)
- [15] P. L. Rosin. "Measuring corner properties". *Computer Vision and Image Understanding*,

73(2):291 – 307, (1999).

- [16] K. Mikolajczyk and C. Schmid. “Indexing based on scale invariant interest points”. In ICCV, volume 1, pages 525 – 531, 2001.
- [17] J. Sivic and A. Zisserman. “Video google: A text retrieval approach to object matching in videos. International Conference on Computer Vision”, page 1470, 2003.
- [18]] D. G. Lowe. “Distinctive image features from scale-invariant keypoints”. International Journal of Computer Vision, 60(2):91–110, 2004.
- [19]] D. Lowe. “Object recognition from local scale-invariant features”. In ICCV, 1999.
- [20]] K. Mikolajczyk and C. Schmid. “A performance evaluation of local descriptors”. PAMI, 27(10):1615–1630, 2005.
- [21] E. Rosten and T. Drummond.”Machine learning for highspeed corner detection”. In European Conference on Computer Vision, volume 1, 2006.
- [22] J. Moreno and A. Acosta. “Análisis Comparativo De Las Técnicas De Harris Y Sift Para La Detección De Puntos De Interés En Fotografías Aéreas”. Ibagué – Colombia. Universidad de Ibagué – Facultad de ingeniería. 2018.