



UNIVERSIDAD DEL PAPALOAPAN

Campus Loma Bonita

INGENIERÍA EN COMPUTACIÓN

DETECCIÓN DE PUNTOS CARACTERÍSTICOS EN IMÁGENES
DIGITALES

TESIS

QUE PARA OBTENER EL TÍTULO DE:
INGENIERA EN COMPUTACIÓN

PRESENTA:

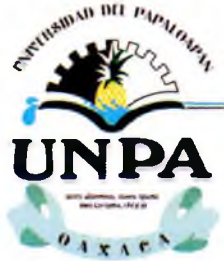
MARÍA DEL CARMEN DÍAZ CHÁVEZ

ASESOR DE TESIS:

M.C. ALMA ALHELÍ PEDRO PÉREZ

LOMA BONITA, OAXACA.

2016



UNIVERSIDAD DEL PAPALOAPAN

INGENIERÍA EN COMPUTACIÓN

LA PRESENTE TESIS TITULADA “**DETECCIÓN DE PUNTOS CARACTERÍSTICOS EN IMÁGENES DIGITALES**” PRESENTADA POR LA SUSTENTANTE DE LICENCIATURA **C. MARÍA DEL CARMEN DÍAZ CHÁVEZ** BAJO LA DIRECCIÓN DE LA **M.C. ALMA ALHELÍ PEDRO PÉREZ**, HA SIDO REVISADA Y ACEPTADA POR EL COMITÉ EXAMINADOR PARA SER DEFENDIDA EN EL EXAMEN PROFESIONAL Y OBTENER EL TÍTULO DE INGENIERA EN COMPUTACIÓN.

**M.C. ALMA ALHELÍ PEDRO
PÉREZ
ASESOR**

**M.C. RAFAEL F. GONZÁLEZ
ZÁRATE
PRESIDENTE**

**M.C. EDUARDO ORTIZ
HERNÁNDEZ
SECRETARIO**

**DR. EDUARDO SÁNCHEZ
SOTO
VOCAL**

Dedicatoria

A mi madre:

Con la mayor gratitud por todos tus esfuerzos y sacrificios para que yo pudiera terminar mi carrera profesional, por haberme dado todo y enseñarme a luchar por lo que se quiere. Gracias por guiarme en mi camino y por estar para mí a pesar de la enorme distancia que nos separa, porque gracias a ti, dios me ha dado la oportunidad de vivir, gracias por haberme impulsado y por la dicha de ser tu hija. Gracias por nunca dejar de creer en mí, todo esto te lo debo a ti.

Te Amo

Agradecimientos

En especial gracias a mi madre Verónica Chávez por ser la persona más hermosa y maravillosa que he conocido, gracias por amarme, cuidarme y guiarme, por ser mi inspiración y mi motivo para nunca rendirme, por ser el mayor ejemplo de fortaleza y dulzura, por celebrar mis triunfos y apoyarme en mis fracasos, por enseñarme lo que significa el verdadero amor, aquel que es incondicional y que aún en la distancia siempre te alienta a ser la mejor versión de ti. Te amo infinitamente mamá.

Agradezco también de forma muy especial a mamá Carmen la mujer más fuerte y valiente, gracias por ser mi guía, por enseñarme lo que es el esfuerzo, por ser un ejemplo en todos los aspectos, por ser mi segunda madre, por cuidar y seguir cuidando de mi como el primer día, por todas sus enseñanzas, sus consejos, gracias por todo su amor y por animarme a nunca renunciar ni rendirme, la quiero mucho.

A ti Dios que me diste la oportunidad de vivir y prestarme a una familia maravillosa. Gracias a mi familia en general que estando cerca o lejos, siempre me brinda su apoyo, su cariño, sus buenos deseos y sobre todo por el sentimiento de unidad que existe entre todos nosotros. Gracias por creer en mi, los quiero.

A mi asesor de tesis, la M.C. Alma Alhelí Pedro Pérez por tener siempre la disposición de guiarme para la elaboración de este trabajo. Le agradezco todas las enseñanzas, las pláticas y las llamadas de atención; por sus consejos, su motivación y su amistad. Gracias por ayudarme a concluir un logro más.

A mis revisores, el Dr. Eduardo Sánchez Soto, al M.C. Eduardo Ortiz Hernández y al M.C. Rafael F. González Zarate, gracias por sus enseñanzas, por sus observaciones y por el tiempo dedicado a la lectura y corrección de esta tesis.

A mis compañeros y amigos por los excelentes y divertidos ratos, por los conocimientos compartidos, por todos los desvelos, el estrés y las carnes asadas. Gracias a todos: Leydy, Zule, Marleny, José Luis, Alex, Gencas, Kikin, Winbugs, Ezequiel, Arturo y Nieto. A mi amiga Rosa, por dejarme confiar en ella, por las extensas pláticas, por los ratos divertidos y por levantarme el ánimo cuando lo necesite, gracias amiga por enseñarme que la amistad existe.

A José Luis por ser la persona que siempre está para mí, por tu compañía, amistad y apoyo inmenso, pero sobre todo, por amarme como lo haces. Gracias por compartir nuevos e inolvidables momentos en mi vida.

A mi tía Juani por haber cuidado de mi en mi infancia, por preocuparte por mi y por todo tu cariño, sabes que te veo como a otra mamá, te quiero mucho. A mi hermano Juan Manuel gracias por su apoyo, por escucharme hablar de mi trabajo e interesarse y por presionarme para que terminara mi tesis.

Por último quiero agradecer a mi abuelo David Díaz que aunque ya no se encuentres con nosotros físicamente, siempre estará presente en mi corazón. Gracias por su amor.

Resumen

El trabajo que se describe en esta tesis aborda el tema de la detección de puntos característicos en imágenes digitales. Los puntos característicos, también conocidos como puntos de interés o puntos clave son características espaciales que capturan información útil para tareas especializadas como el reconocimiento de objetos.

El método que se utilizó para la detección de los puntos característicos es la transformada SIFT (*Scale Invariant Features Transform*). Funciona como un detector y descriptor de puntos característicos. En este trabajo se utilizó el método detector, el cual se centra en buscar puntos característicos en el denominado espacio-escala. El detector se divide en 3 etapas: etapa 1 descomposición de la imagen en octavas y escalas, etapa 2 diferencia Gaussiana y etapa 3 detección de máximos y mínimos. La implementación del algoritmo se realizó en la herramienta de software Matlab en su versión 8.4 R2014b.

En la etapa experimental se emplearon tres categorías de imágenes: Instituto de Agroingeniería, Laboratorio de Mecatrónica y Auditorio, estas imágenes se capturaron en un ambiente no controlable. Se utilizaron 5 ejemplos de cada categoría y se realizó un conteo del total de puntos característicos en cada imagen. De acuerdo con los resultados obtenidos se observó que la categoría donde se detectó un mayor número de puntos característicos fue el Laboratorio de Mecatrónica. Esto se debe a que en esta categoría interviene una mayor cantidad de elementos (como nubes, aves, vegetación y sombras).

Abstract

The work described in this thesis addresses the issue of detection of characteristic points in digital images. The characteristic points, also known as points of interest or keypoints, capturing spatial characteristics are useful for specialized tasks such as object recognition information.

The method used to detect the characteristic points is the transform SIFT (Scala Invariant Features Transform), it functions as a detector and descriptor of characteristic points. The detector was used in this work, which focuses on finding characteristic points in the so called space-scale. The detector is divided into three stages: stage one decomposition image in octaves and scales, stage 2 difference of Gaussians and stage 3 maximum and minimum detection. The implementation of the algorithm was realized in the software tool Matlab R2014 version 8.4.

In the experimental stage three categories of images were used: The Agricultural Engineering Institute, the Mechatronics Laboratory and the Auditorium, these images were captured on a non-controllable environment. 5 examples of each category were used and a count of total characteristic points was detected in each of the images. According to the results obtained the category in which the greater number of characteristic points detected was the Mechatronics Laboratory. The reason for this is in this category a greater number of elements (such as clouds, birds, vegetation and shadows) was involved.

Índice general

1. Planteamiento del problema	1
1.1. Introducción	1
1.2. Descripción del problema	2
1.3. Objetivo general	3
1.3.1. Objetivos específicos	3
1.3.2. Alcances y Limitaciones	3
1.4. Justificación	4
1.5. Hipótesis	4
1.6. Organización de la tesis	4
2. Estado del Arte	6
2.1. <i>A Vector SIFT Detector for Interest Point Detection in Hyperspectral Imagery</i>	6
2.2. <i>Using Gradient Features From Scale-Invariant Keypoints on Face Recognition</i>	7
2.3. <i>Image Recognition by Scale Invariant Feature Transformation (SIFT)</i>	8
2.4. <i>SAR-SIFT: A SIFT-Like Algorithm for SAR Images</i>	8
2.5. Reconocimiento e identificación de Logotipos en imágenes con transformada SIFT	9
2.6. Detección y asociación automática de puntos característicos para diferentes aplicaciones	9
2.7. Clasificación de objetos en imágenes usando SIFT	10
2.8. Análisis del problema	12

<i>ÍNDICE GENERAL</i>	VII
2.8.1. Problemas relacionados con las imágenes	12
2.8.2. Categorías de las imágenes a analizar	12
2.9. Propuesta de solución	13
2.10. Discusión	13
3. Marco Teórico	14
3.1. Imagen Digital	14
3.2. Punto característico	15
3.3. Transformada de características invariantes a escala SIFT	15
3.3.1. Detección de extremos en el espacio escala	16
3.3.1.1. Detección de máximos y mínimos	19
3.3.2. Localización de puntos clave	19
3.3.2.1. Eliminación de puntos de bajo contraste	20
3.3.2.2. Eliminación de puntos localizados en los bordes	21
3.3.3. Asignación de la orientación	22
3.3.4. Descriptor de puntos característicos	23
3.4. Discusión	25
4. Análisis, diseño e implementación del sistema	26
4.1. Análisis del Sistema	26
4.2. Diseño del sistema	27
4.2.1. Arquitectura del sistema	27
4.2.1.1. Etapa 1: Descomposición de la imagen en octavas y escalas	28
4.2.1.2. Etapa 2: Aplicación de la Diferencia Gaussiana	32
4.2.1.3. Etapa 3: Extracción de puntos característicos	35
4.3. Diseño e implementación del sistema	39
4.3.1. Descripción del ambiente de implementación	39
4.4. Discusión	39

<i>ÍNDICE GENERAL</i>	VIII
5. Experimentos y resultados	41
5.1. Ambiente de pruebas	41
5.2. Pruebas	42
5.2.1. Prueba 1: Detección de puntos característicos en imágenes de la categoría Agroingeniería	42
5.2.2. Prueba 2: Detección de puntos característicos en imágenes de la categoría Mecatrónica	46
5.2.3. Prueba 3: Detección de puntos característicos en imágenes de la categoría Auditorio	50
5.2.4. Gráfica del total de puntos detectados	54
5.3. Discusión	54
6. Conclusiones generales	56
6.1. Conclusiones	56
6.2. Aportaciones	57
6.3. Trabajos futuros	57

Índice de figuras

2.1. Puntos de interés detectados utilizando la transformada de SIFT y el Vector SIFT propuesto en [3].	7
3.1. Representación de la imagen digital	15
3.2. Pirámide Gaussiana y Diferencia Gaussiana. Fuente David Lowe [13]	18
3.3. Localización de máximos y mínimos locales. Fuente David Lowe [13].	19
3.4. Etapas de selección de puntos característicos. Fuente David Lowe [13].	20
3.5. Descriptor de puntos característicos. Fuente David Lowe [13].	23
4.1. Elementos que intervienen en el análisis del sistema.	27
4.2. Arquitectura del sistema.	28
4.3. Pasos de la etapa 1.	29
4.4. Categoría de imágenes. (a) Instituto de Agroingeniería; (b) Auditorio; (c) Laboratorio de Mecatrónica.	30
4.5. Conversión de la imagen original a escala de grises. (a) Imagen original; (b) Imagen en escala de grises.	30
4.6. Imágenes con aplicación de desenfoque Gaussiano	31
4.7. Pasos de la etapa 2	33
4.8. Proceso para obtener las imágenes de Diferencia Gaussiana $D(x, y, \sigma)$	34
4.9. Pirámide de Diferencia Gaussiana $D(x, y, \sigma)$	35

<i>ÍNDICE DE FIGURAS</i>	X
4.10. Pasos de la etapa 3	36
4.11. En rojo: píxel de estudio. En amarillo: 26 vecinos (8 de la escala actual, 9 de la escala anterior y 9 de la escala posterior).	37
4.12. Detección de máximo local.	37
4.13. Detección de mínimo local.	37
4.14. Detección de puntos. En verde puntos máximos, en amarillo puntos mínimos.	38
4.15. Esquema de interacción del usuario con el sistema.	39
5.1. Categoría 1. Instituto de Agroingeniería.	42
5.2. Detección de puntos característicos de la categoría Agroingeniería.	43
5.3. Valores promedio de los puntos máximos detectados en las imágenes de la Categoría Agroingeniería.	45
5.4. Valores promedio de los puntos mínimos detectados en las imágenes de la Categoría Agroingeniería.	45
5.5. Categoría 2. Laboratorio de Mecatrónica.	46
5.6. Detección de puntos característicos de la categoría Mecatrónica.	47
5.7. Valores promedio de los puntos máximos detectados en las imágenes de la Categoría Mecatrónica.	49
5.8. Valores promedio de los puntos mínimos detectados en las imágenes de la Categoría Mecatrónica.	49
5.9. Categoría 3. Auditorio.	50
5.10. Detección de puntos característicos de la categoría Auditorio.	51
5.11. Valores promedio de los puntos máximos detectados en las imágenes de la Categoría Auditorio.	53
5.12. Valores promedio de los puntos mínimos detectados en las imágenes de la Categoría Auditorio.	53
5.13. Número de puntos característicos detectados por cada categoría.	54

Índice de tablas

2.1. Resumen de trabajos relacionados	11
2.2. Métodos para detección de puntos característicos	12
2.3. Categoría y características de las imágenes a analizar para esta investigación	12
4.1. Descomposición de la imagen en octavas.	32
5.1. Número de puntos detectados en la Figura 5.1(a).	43
5.2. Número de puntos detectados en la Figura 5.1(b).	43
5.3. Número de puntos detectados en la Figura 5.1(c).	44
5.4. Número de puntos detectados en la Figura 5.1(d).	44
5.5. Número de puntos detectados en la Figura 5.1(e).	44
5.6. Número total de puntos característicos de la categoría Agroingeniería.	44
5.7. Número de puntos detectados en la Figura 5.5(a).	47
5.8. Número de puntos detectados en la Figura 5.5(b).	47
5.9. Número de puntos detectados en la Figura 5.5(c).	48
5.10. Número de puntos detectados en la Figura 5.5(d).	48
5.11. Número de puntos detectados en la Figura 5.5(e).	48
5.12. Número total de puntos característicos de la categoría Mecatrónica.	48
5.13. Número de puntos detectados en la Figura 5.9(a).	51
5.14. Número de puntos detectados en la Figura 5.9(b).	51

ÍNDICE DE TABLAS

XII

5.15. Número de puntos detectados en la Figura 5.9(c).	52
5.16. Número de puntos detectados en la Figura 5.9(d).	52
5.17. Número de puntos detectados en la Figura 5.9(e).	52
5.18. Número total de puntos característicos de la categoría Auditorio.	52

Capítulo 1

Planteamiento del problema

En este capítulo inicial se describe el problema de esta investigación, el cual consiste en detectar puntos característicos en imágenes digitales. Así mismo, se expone la propuesta de solución que comprende el objetivo general y los objetivos específicos. También se precisan los alcances y limitaciones de la presente investigación. Para finalizar se describe la organización de los capítulos en los cuales se explica cómo se desarrolló esta tesis.

1.1. Introducción

Hoy en día la visión artificial es una disciplina utilizada en distintas áreas, con la finalidad de extraer información del mundo físico a partir de imágenes, por medio de una computadora [10]. La visión artificial es un proceso que produce una descripción útil con información relevante, a partir de imágenes del mundo exterior.

La descripción de una imagen se puede realizar aplicando diversos métodos, entre los cuales se encuentra la detección de puntos característicos. Estos puntos son la base de numerosas tareas en la visión por computadora.

La detección de los puntos característicos, dentro del procesamiento de imágenes tiene como objetivo ubicar en la imagen, determinados puntos que aporten información relevante, como la localización y la escala. Los puntos característicos de una imagen, por lo regular, están localizados donde se produce un cambio en una propiedad de la imagen. Estos puntos por lo tanto, aportan mayor cantidad de información que otros puntos de su entorno.

En la literatura existen investigaciones que abordan la temática de las técnicas para la detección de puntos característicos [19], [3], [4], [8]. Cabe mencionar que estos trabajos se emplean para realizar tareas más sofisticadas como el reconocimiento de rostros, de objetos y la detección de puntos característicos en imágenes hiperespectrales, tareas en las cuales es de suma importancia la detección de puntos característicos.

En este trabajo de investigación se realiza la detección de puntos característicos de edificios de la Universidad del Papaloapan campus Loma Bonita, para ello es necesario llevar a cabo una serie de etapas como la transformada SIFT durante la detección, y que será detallada más adelante.

1.2. Descripción del problema

La detección de puntos característicos en imágenes digitales es una tarea difícil, debido a que la mayoría de las veces, las imágenes se adquieren bajo ambientes no controlados como cambios de iluminación, movimiento de los objetos, entre otros. Aunado a lo anterior si las imágenes corresponden a escenas naturales incluyen distintas clases de elementos como: árboles, flores, pasto, cielo, nubes, entre otros.

1.3. Objetivo general

El objetivo principal de este trabajo de investigación consiste en implementar un algoritmo para detectar puntos característicos en imágenes de edificios de la Universidad del Papaloapan.

1.3.1. Objetivos específicos

Los objetivos específicos que se presentan en esta tesis son los siguientes:

1. Adquirir las imágenes de tres edificios de la Universidad del Papaloapan campus Loma Bonita.
2. Estudiar y analizar los métodos para la detección de puntos de característicos en las imágenes.
3. Elegir un método para la detección de puntos característicos.
4. Implementar un método para la detección de puntos característicos.
5. Realizar pruebas con imágenes de edificios de la Universidad del Papaloapan.
6. Evaluar los resultados obtenidos.

1.3.2. Alcances y Limitaciones

Los alcances y limitaciones que presenta esta tesis son los siguientes:

1. Las imágenes serán solo de los edificios de la Universidad del Papaloapan campus Loma Bonita.
2. Se usarán al menos tres categorías de edificios.
3. Se aplicará un solo método para la detección de los puntos característicos de imágenes.
4. La implementación del algoritmo se realizará en Matlab.

1.4. Justificación

Este trabajo de tesis se origina a partir de la necesidad de obtener información de imágenes digitales, con la finalidad de utilizar esta información en tareas especializadas como: el reconocimiento de objetos, la reconstrucción de escenas en ambientes virtuales, la recuperación de información con base en el contenido, entre otras aplicaciones.

En las últimas décadas se han desarrollado numerosas investigaciones relacionadas con la detección de puntos característicos. Sin embargo, en la mayoría de los casos se trabaja con ambientes controlados en los cuales se cuida el contexto de adquisición de las imágenes.

La finalidad es experimentar con imágenes adquiridas en un ambiente no controlado y que integren diferentes clases de objetos como: vegetación, cielo y edificios. La información obtenida servirá para realizar un análisis de la información y determinar como se comporta el algoritmo bajo éstas condiciones de aplicaciones.

1.5. Hipótesis

El algoritmo a implementar detectará los puntos característicos de forma óptima en imágenes de edificios de la Universidad del Papaloapan, obtenidas en un ambiente no controlado y con la inclusión de diversas clases de objetos.

1.6. Organización de la tesis

El presente documento de tesis se compone de seis capítulos y una sección de referencias. A continuación se describen de forma general los capítulos que integran el contenido de la tesis.

Capítulo 2. Estado del arte, análisis del problema y propuesta de solución.

En este capítulo se describen los trabajos relacionados con la presente investigación, también se expone el análisis del problema y la propuesta de solución.

Capítulo 3. Marco teórico.

En este capítulo se describen los conceptos principales, así como la teoría que permite comprender el desarrollo del trabajo de esta tesis.

Capítulo 4. Análisis, diseño e implementación del sistema.

En este capítulo se realiza la implementación de las técnicas, el desarrollo y funcionamiento del algoritmo para la detección de puntos característicos.

Capítulo 5. Experimentos y resultados.

En este capítulo se llevan a cabo las pruebas pertinentes y se analizan los resultados obtenidos para conocer el desempeño del algoritmo que se implemento.

Capítulo 6. Conclusiones generales.

En este capítulo se muestran las conclusiones que se derivan de todo el proceso de investigación y de los experimentos realizados.

Capítulo 2

Estado del Arte

En este capítulo se realiza una revisión de investigaciones relacionadas al tema de reconocimiento de imágenes a partir de sus puntos característicos. Con base en la información recopilada se desarrolla un análisis del problema y se diseña una propuesta de solución.

2.1. *A Vector SIFT Detector for Interest Point Detection in Hyperspectral Imagery*

En este trabajo [3] se presenta un algoritmo para la extracción de puntos de interés en imágenes hiperespectrales aplicando una extensión a la transformada SIFT (*Scale Invariant Feature Transform*), las imágenes fueron adquiridas con los sensores del Sistema de Imágenes Hiperespectrales Aéreas (AISA) Eagle y Hyperion.

La transformada SIFT funciona bien con imágenes en escala de grises o de una sola banda, pero, no con imágenes hiperespectrales, las cuales se dividen en varias bandas, debido a que, es difícil mapear con precisión una imagen vectorial en una imagen escalar. En la extensión que realizaron capturan todas las bandas espectrales de la imagen y en lugar de utilizar el desenfoque gaussiano propuesto por Lowe, lo hacen por difusión no lineal, lo que conduce a una mejor detección de los puntos de interés, debido a que, conserva mejor los bordes de la imagen.

En la fase experimental se aplican dos pruebas para probar el algoritmo, en la primera se evalúa si la difusión no lineal mejora la detección de los puntos de interés, mientras que, en la segunda se estudia la invariancia de escala de los puntos de interés detectados en las imágenes AISA en diferentes resoluciones.

Los resultados mostraron que el vector SIFT propuesto detecto más puntos de interés en comparación con la transformada SIFT propuesta por Lowe. En la Figura 2.1 se observa que, los puntos de interés detectados por el vector SIFT están ubicados en los bordes, mientras que los detectados por la transformada de SIFT están mas dispersos en la imagen.

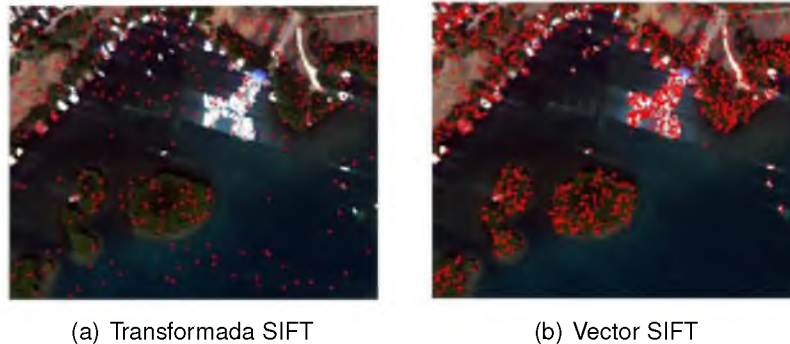


Figura 2.1: Puntos de interés detectados utilizando la transformada de SIFT y el Vector SIFT propuesto en [3].

2.2. *Using Gradient Features From Scale-Invariant Keypoints on Face Recognition*

En este artículo [19] los autores proponen la aplicación del algoritmo *PCA-SIFT* para el reconocimiento de rostros, que combina el *Án*alisis de Componentes Principales (*PCA*), la transformada SIFT y las características de gradiente. El algoritmo es apropiado para resolver algunos de los problemas que se presentan en el reconocimiento facial, como los accesorios, las expresiones faciales y la boca, la cual es la parte que más cambia del rostro.

El proceso total se compone de tres etapas: pre-proceso, extracción de características y clasificación. En el pre-proceso se reduce el ruido de las imágenes y las variaciones de iluminación, se descarta el cabello y la ropa, obteniendo únicamente el rostro. Para la extracción de los puntos característicos aplicaron las primeras tres etapas del algoritmo SIFT. En la tercer etapa se empleo el algoritmo *K-means* para la agrupación de los puntos característicos. Por último, las similitudes se combinan para clasificar la imagen de la cara.

En la etapa experimental se realizaron comparaciones con otros métodos como *LBP*, *SIFT_CLUSTER* y *SIFT_GRID*. El método *PCA-SIFT* obtuvo los mejores resultados ante variaciones de accesorios y expresiones con un 94.9% y 95.9% respectivamente.

Con base en los resultados los investigadores mencionan que los descriptores *PCA-SIFT* son más robustos y eficientes que los descriptores locales SIFT ante las variaciones de escala, rotación, accesorios y expresiones del rostro.

2.3. *Image Recognition by Scale Invariant Feature Transformation (SIFT)*

En este artículo [7] se describe el desarrollo de un software capaz de realizar el reconocimiento de objetos en diferentes imágenes almacenadas en una base de datos.

El software fue desarrollado en el entorno Java con una base de datos en PostgreSQL. Para realizar el reconocimiento de los objetos se aplicó el algoritmo SIFT, el cual se encarga de extraer los puntos característicos de las imágenes, que se utilizan para reconocer a un objeto presente en dos imágenes con diferentes perspectivas.

En los experimentos realizados el software reconoce los objetos sin importar las variaciones que presentaran, como rotación, tamaño e incluso problemas de iluminación. Además, mediante una interfaz permite al usuario modificar los parámetros como, el número de escalas para la imagen, el mínimo de octavas y la dimensión de las imágenes. El resultado del reconocimiento depende de los parámetros del algoritmo SIFT, así como también de las características propias de la imagen.

2.4. *SAR-SIFT: A SIFT-Like Algorithm for SAR Images*

En este trabajo [2] describen un nuevo algoritmo para la extracción de descriptores locales llamado SAR-SIFT, desarrollado para imágenes SAR (Radar de apertura sintética), esto debido a que SIFT no funciona correctamente con este tipo de imágenes, porque, suelen estar dañadas por un fuerte ruido multiplicativo, llamado speckle y detectar falsos puntos característicos.

El algoritmo SIFT fue adaptado a las imágenes SAR mediante un nuevo método llamado gradiente de relación (GR). Con base en los resultados obtenidos se observó que el algoritmo obtuvo mejores resultados que SIFT con las imágenes SAR. Para la detección de los puntos aplican el detector LoG-Harris. Esto permitió reducir el ruido y el número de falsos puntos detectados.

2.5. Reconocimiento e identificación de Logotipos en imágenes con transformada SIFT

El objetivo principal de este trabajo [8] es la creación de un sistema de reconocimiento de logotipos, utiliza la transformada *SIFT* para la extracción de los puntos característicos de las imágenes.

El sistema tiene una fase de entrenamiento y una de prueba, fue desarrollado en Matlab y utiliza la herramienta *VLFEAT*. De un total de 300 imágenes de 100 logotipos diferentes, 200 imágenes fueron para la fase de entrenamiento y 100 para la de prueba. En la primera fase se extrajeron las características de los logotipos que se iban a identificar, mientras que en la segunda se realizó el reconocimiento de los logotipos que aparecen en las imágenes que el usuario introduzca en el sistema.

En las pruebas, las 100 imágenes fueron clasificadas en 8 categorías: individuales, con cambios de escala, giros mayores a 30 grados, giros menores a 30 grados, cambios de iluminación, oclusiones parciales, múltiples logotipos en una imagen y en contexto. Al probar el sistema con las diferentes categorías de imágenes se concluyó que el sistema trabaja mejor con algunos logotipos y tiene una mejor respuesta para imágenes sin ningún tipo de cambios, además entre más características tenga una imagen más sencillo es realizar el reconocimiento de los logotipos. Se realizaron 5 pruebas distintas y se calcularon los valores de tres variables, precisión (P), cobertura (R) y medida (F1). El resultado final de las pruebas arrojó un 62.68 % para la localización de los logotipos existentes en una imagen.

2.6. Detección y asociación automática de puntos característicos para diferentes aplicaciones

En este proyecto Enebral [5] realizó la búsqueda de puntos característicos y la asociación de correspondencias entre imágenes. Los puntos característicos obtenidos son estables frente a va-

riaciones, como cambios de orientación, traslaciones, cambios de escala y perspectiva.

Los métodos implementados para la detección de puntos característicos, son el detector de esquinas Harris [6] y la transformada SIFT desarrollada por Lowe en 2004 [13]. Como descriptor de Harris escogió el estudio de color a partir de matrices RGB, en una ventana de tamaño concreto alrededor del punto, y para SIFT, histogramas de los gradientes locales e imágenes en escala de grises. El algoritmo fue desarrollado en MATLAB.

Se realizaron pruebas con imágenes de figuras geométricas básicas y fotografías reales, para el detector Harris se utilizaron dos descriptores diferentes: matrices RGB y cálculo de transformación espacial, dando mejores resultados el segundo. En la prueba con fotografías reales no se obtuvieron buenos resultados ante las variaciones de rotación y cambio de escala. En cambio, con SIFT el porcentaje de aciertos estuvo por encima del 80% y hasta el 90% dependiendo de la iluminación y el grado de semejanza entre los objetos. Por lo tanto, se concluyó que el detector de esquinas Harris obtuvo buenos resultados ante variaciones de traslaciones o rotación, pero no de iluminación y escala. En cambio, SIFT resultó ser robusto ante dichas variaciones, pero más caro computacionalmente.

2.7. Clasificación de objetos en imágenes usando SIFT

En este trabajo [4] se realizó el reconocimiento de clases de objetos en imágenes. Utilizó el método SIFT para la extracción de puntos característicos y para la clasificación de clases de objetos aplicó una Red Bayesiana de tres niveles. En el primer nivel se tiene los descriptores extraídos para el objeto que se desea clasificar, en segundo nivel se tienen los clústers obtenidos en la fase de entrenamiento y en el tercer nivel se tienen las clases de los objetos entrenados.

Las pruebas fueron realizadas con una base de datos perteneciente a la comisión europea PASCAL, dicha base de datos contiene 102 clases y distinto número de imágenes por cada clase, el tamaño de las imágenes es de 300 x 300 píxeles y son formato JPG. Se realizaron dos experimentos, en el primer experimento se utilizaron 100 imágenes por cada clase, y en el segundo experimento se obtuvieron 10 imágenes aleatoriamente, las primeras tienen poca variación en escala, oclusión, iluminación y rotación, mientras que las segundas sufren agresiones en dichas invarianzas. El primer experimento tiene una precisión en promedio del 93%, y el segundo experimento tiene una pequeña mejora, dando como resultado 93.3% de precisión.

Ambos mantienen la invarianza ante iluminación, escala, rotación y oclusión hasta con un 40%. En la tabla 2.1 se muestra un resumen con descripciones de otros trabajos relacionados al tema de esta tesis.

Artículo o Tesis	Método empleado	Categoría de imágenes	Invarianza Abordada / Tipo de ambiente	Métodos con mejor desempeño
[Rister, 2013] [18]	-SIFT -GPU móvil	-Escenas de edificios -Texturas naturales	Escala Ambiente no controlado	Procesamiento casi en tiempo real
[Zhenwei, 2013] [21]	-SIFT	-Rostros -Texturas naturales y artificiales -Edificios	Escala Ambiente controlado y no controlado	No se menciona
[Mesa, 2008] [14]	-RNA -MLP -RBF -Kohonen -Plessey -Susan -Sojka	Imágenes en escala de grises	No se menciona	La detección por medio de RNA fue de menor calidad, pero MLP y RBF fueron más eficaces.
[Constanzo, 2014] [1]	-SIFT -Detectores forenses	-Texturas naturales y artificiales -Piezas arqueológicas -Escenas naturales	Áreas Clonadas Ambientes no controlables	No se menciona

Tabla 2.1: Resumen de trabajos relacionados

En la tabla 2.2 se muestran algunos de los métodos que han sido utilizados para la detección de puntos característicos. Éstos comprenden la detección de esquinas, bordes y puntos característicos.

Autor / Año	Método	Tipo de detección
Moravec (1981)	Detector de Moravec	Detector de esquinas
Harris y Stephens (1988)	Detector de Harris y Stephens	Detector de bordes y esquinas
Harris (1992)	Detector de Harris	Detector de esquinas
David Lowe (2004)	SIFT (<i>Scale Invariant Feature Transform</i>)	Detector y descriptor de puntos característicos
Herbert Bay (2006)	SURF (<i>Speeded-Up Robust Features</i>)	Detector y descriptor de puntos característicos

Tabla 2.2: Métodos para detección de puntos característicos

2.8. Análisis del problema

En este apartado se presentan los problemas relacionados con esta investigación y se describe de forma general la solución propuesta para realizar la detección de puntos característicos en imágenes de los edificios de Universidad del Papaloapan.

2.8.1. Problemas relacionados con las imágenes

Las imágenes utilizadas en esta tesis presentan diversos problemas:

- Existe variabilidad en la iluminación de las imágenes adquiridas, debido a que se tomaron en un ambiente no controlable.
- Las imágenes presentan cambios de escala debido a que en la obtención de las imágenes no se considero un riguroso proceso para su obtención por lo tanto, las imágenes obtenidas se adquirieron a diferentes distancias y con diferentes ángulos.

2.8.2. Categorías de las imágenes a analizar

Categoría de imágenes	Características
Instituto de AgroIngeniería	<ul style="list-style-type: none"> ■ Variedad de elementos en una misma imagen por ejemplo: árboles, césped y cielo. ■ Variedad en la tonalidad de los elementos.
Laboratorio de Mecatrónica	
Auditorio	

Tabla 2.3: Categoría y características de las imágenes a analizar para esta investigación

2.9. Propuesta de solución

En esta sección se plantea una propuesta para resolver el problema con base en la revisión del estado del arte. La metodología se compone de tres pasos los cuales son:

1. **Descomposición de la imagen en octavas y escalas.** En esta primer etapa la imagen de referencia será descompuesta en octavas y escalas.
2. **Diferencia Gaussiana.** En esta etapa se calculará la Diferencia Gaussiana de las imágenes resultantes de la etapa anterior.
3. **Detección de máximos y mínimos locales.** Esta etapa consiste en detectar los puntos característicos correspondientes a cada octava y escala de la imagen de referencia. De cada punto detectado se deberá guardar la posición, la octava y la escala a la cual pertenece.

2.10. Discusión

En este capítulo se presentó el estado del arte de esta investigación, donde se describen los métodos y sus explicaciones para la detección de puntos característicos. Éstos métodos son utilizados para realizar tareas importantes como el reconocimiento de objetos y de rostros, así como la clasificación de objetos y la obtención de puntos característicos en imágenes hiperespectrales, por mencionar algunos.

- En los trabajos relacionados se observó que se emplean distintas categorías de imágenes por ejemplo: imágenes de rostros, hiperespectrales, edificios, logotipos, objetos como motocicletas, billetes y cámaras digitales.
- De acuerdo al análisis del problema se determinó aplicar una metodología en tres etapas aplicando el algoritmo SIFT en su fase de detector. Este proceso se implementó en imágenes de edificios en ambientes no controlables.

Capítulo 3

Marco Teórico

En este capítulo se presentan los conceptos y definiciones que permitirán al lector, de una forma sencilla, leer y comprender el contenido de esta tesis. Se describen los conceptos que se relacionan con el tema de esta tesis.

3.1. Imagen Digital

La imagen digital [17] es una representación bidimensional (2D) de datos, donde cada punto o dato se denomina píxel (del inglés *picture element*). Se presenta en forma de matriz con una resolución de $M \times N$ elementos. Cada elemento de la matriz o *píxel* tendrá un valor asignado que corresponde al nivel de luminosidad del punto correspondiente en la escena captada; dicho valor es el resultado de la cuantización de *intensidad* o *nivel de gris*.

La imagen digital puede definirse como una función discreta $f(x, y) = z$, donde x e y son las coordenadas espaciales y z es el valor que toma la imagen en esas coordenadas. El valor de z indica la intensidad de la luz o nivel de gris de la imagen en ese punto, que es cuantificada dentro de un rango de valores (0 a 255).

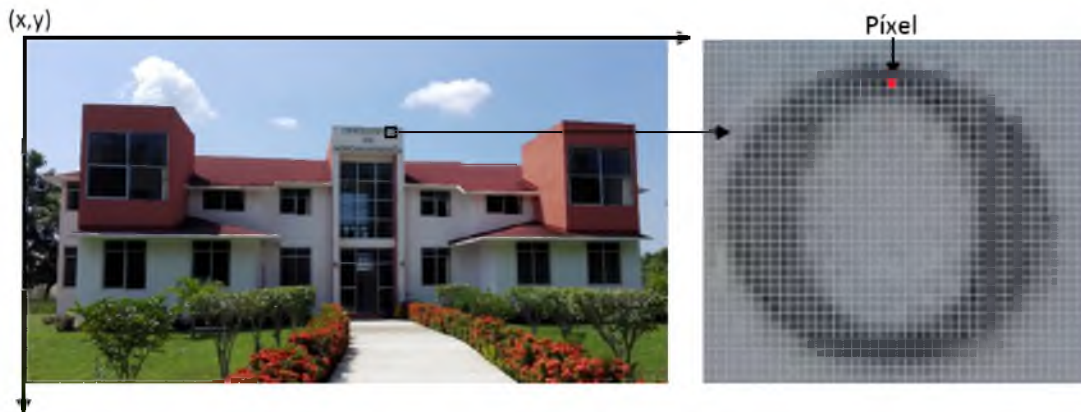


Figura 3.1: Representación de la imagen digital

3.2. Punto característico

Se entiende por punto característico o de interés a cualquier punto de la imagen en el cual la señal cambia bidimensionalmente [16]. Los puntos característicos son características espaciales de la imagen que capturan información de sus vecinos, son distintivos y estables bajo transformaciones de traslación y rotación [3]. Éstos puntos son aplicados a problemas de visión por computadora como el reconocimiento de imágenes, la comparación, la reconstrucción en 3D y la detección de cambios.

3.3. Transformada de características invariantes a escala SIFT

El algoritmo SIFT (Scale Invariant Feature Transform) fue desarrollado por David Lowe y los pasos para extraer los puntos característicos invariantes a la escala, orientación e iluminación, se describen en uno de sus documentos de investigación escrito en 2004, al cual cita como "*Distinctive image feature from Scale-Invariant Keypoints*" [13].

El objetivo principal de la transformada SIFT es la extracción de los puntos característicos que describen los objetos presentes en distintas imágenes. Los puntos ayudan a encontrar las similitudes existentes entre las imágenes almacenadas en una base de datos y una imagen nueva. La transformada SIFT se realiza mediante cuatro etapas que se describen a continuación:

1. **Detección de extremos espacio-escala (*Scale-Space*):** En esta etapa se representa a la imagen en diferentes escalas y tamaños para detectar los puntos característicos invariantes a escala y orientación. Se lleva a cabo de manera eficiente mediante el uso de la función de Diferencia Guassiana (DoG).
2. **Localización de puntos clave:** Los puntos característicos son seleccionados con base a las medidas de su estabilidad para descartar aquellos que no sean adecuados.
3. **Asignación de la orientación:** Se asignan una o más orientaciones a cada punto característico extraído, basándose en las direcciones locales de la imagen gradiente, es decir, en las variaciones de píxeles cercanos en las dos dimensiones.
4. **Descriptor de puntos clave:** Los gradientes locales de la imagen se miden en la región que rodea al punto característico. Éstos son transformados mediante una representación que permitirá medir niveles de distorsión y cambios en la iluminación de forma local.

3.3.1. Detección de extremos en el espacio escala

Es la primera de las etapas, en la que se representa a la imagen en diferentes escalas y tamaños para detectar los puntos característicos invariantes a escala y orientación. Se lleva a cabo de manera eficiente mediante el uso de la función de Diferencia Guassiana (DoG). A continuación se definen dos términos que ayudarán a comprender la detección de extremos en el espacio-escala:

- Octava: Conjunto de imágenes con el mismo tamaño pero con diferente nivel de borrosidad.
- Escala: Conjunto de imágenes con el mismo nivel de borrosidad pero con diferente tamaño.

El objetivo es obtener puntos característicos de la imagen que puedan ser identificados de forma repetida bajo diferentes vistas del mismo objeto. La detección de la ubicación de los puntos característicos invariantes a los cambios de escala se realiza mediante la función continua conocida como espacio-escala (*Scale-Space*), desarrollada por Witkin y descrita en su trabajo de investigación en 1986 [20].

El trabajo de Koenderik (1984) [9] y Lindeberg (1994) [11] encontró que el único filtro apropiado para estos efectos son los filtros Guassianos de paso bajo, debido a que la función Gaussiana es invariante a escala en el espacio, para la detección de puntos característicos, además elimina el ruido de la imagen. Por lo tanto, el "Espacio-Escala" de una imagen se define como una función

$L(x, y, \sigma)$. Esta función se obtiene a partir de la convolución de la Gaussiana $G(x, y, \sigma)$ y una imagen de entrada $I(x, y)$.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (3.1)$$

Donde (x, y) son las coordenadas espaciales, σ el factor de escala y $*$ es la operación de convolución entre x e y , y:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (3.2)$$

Lowe [12] propuso el uso de la función DoG (Diferencia de Gaussianas) convolucionada con la imagen $D(x, y, \sigma)$ para detectar de manera más eficiente las ubicaciones de los puntos característicos en el espacio escala. Esta función DoG $D(x, y, \sigma)$ se calcula a partir de la diferencia de dos escalas cercanas en cada octava, donde k es una constante multiplicativa del factor escala.

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (3.3)$$

Lowe explica que existen varias razones para escoger esta función. En primer lugar, es una función particularmente eficiente para calcular. Las imágenes suavizadas L , son calculadas para la descripción de características en el espacio-escala, y por lo tanto, D puede calcularse por una simple resta de la imagen. Además, la función de Diferencia Gaussiana, proporciona una aproximación cercana a la Laplaciana normalizada en escala de la Gaussiana $\sigma^2 \nabla^2 G$, tal y como estudió Linderbeg en 1994 [11]. Lindenberberg demostró que la normalización de la Laplaciana mediante el factor σ^2 es necesaria para obtener una escala invariante.

En comparaciones experimentales, Mikolajczyk en 2002 [15] encontró que el máximo y mínimo de $\sigma^2 \nabla^2 G$ producen características más estables de la imagen, en comparación con otras funciones como, el gradiente, el Hessiano o el detector de esquinas Harris. La relación entre $D(x, y, \sigma)$ y $\sigma^2 \nabla^2 G$, puede entenderse mediante la ecuación de difusión por calor (parametrizada en términos de σ , en lugar de $t = \sigma^2$).

$$\frac{\partial G}{\partial \sigma} = \sigma \nabla^2 G \quad (3.4)$$

Esto permite observar que $\nabla^2 G$ puede ser calculada por la aproximación de diferencia finita de $\partial G / \partial \sigma$ usando la diferencia de escalas cercanas de $k\sigma$ y σ

$$\sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma} \tag{3.5}$$

y por lo tanto:

$$G(x, y, k\sigma) - G(x, y, \sigma) \approx (k - 1)\sigma^2 \nabla^2 G \tag{3.6}$$

Con esto se demuestra que cuando la función de Diferencia Gaussiana tiene escalas diferentes por un factor constante, ya incorpora la normalización de la escala σ^2 , necesaria para la Laplaciana invariante en escala. El factor $(k - 1)$ en la ecuación es una constante en todas las escalas y por lo tanto no influye en el proceso de la ubicación extrema. El error de aproximación no tiene casi impacto en la estabilidad de la detección de extremos o en la localización para diferencias significativas en la escala, tal como $k \approx 2^{1/2}$.

Una forma eficaz de construir $D(x, y, \sigma)$ se muestra en la Figura 3.2. La imagen inicial se convoluciona con Gaussianas para producir distintas imágenes separadas por un factor constante de espacio escala k que se muestran en la columna izquierda. Se dividirá cada octava del espacio-escala (es decir, doblando σ) en un número entero de intervalos, s , tal que $k = 2^{1/s}$. De esta manera se deberán producir $s + 3$ imágenes borrosas para cada octava, y de acuerdo a las investigaciones de Lowe, el valor que mejores resultados consigue es el de $s = 3$. La diferencia Gaussiana se obtiene al restar dos imágenes vecinas.

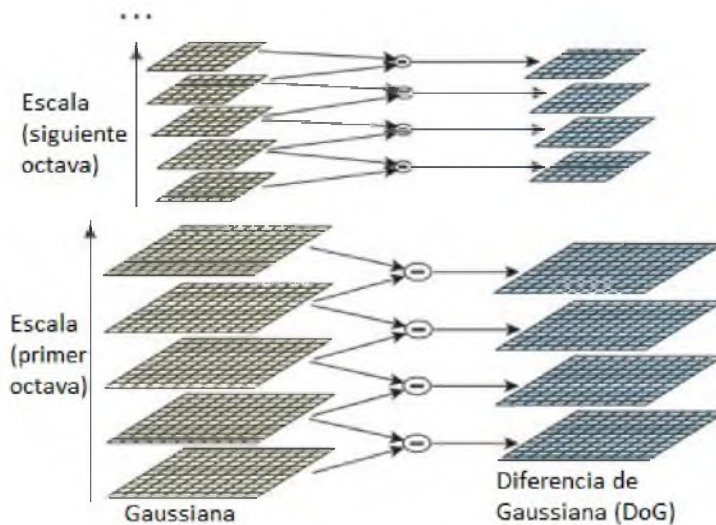


Figura 3.2: Pirámide Gaussiana y Diferencia Gaussiana. Fuente David Lowe [13]

3.3.1.1. Detección de máximos y mínimos

A partir de los cálculos anteriores, se obtendrán los máximos y mínimos locales del espacio $D(x, y, \sigma)$, para ello, cada punto se compara con sus ocho vecinos de la imagen local y los nueve vecinos de la escala anterior y posterior, en total se compara con 26 píxeles vecinos. La Figura 3.3 indica este procedimiento. Se seleccionará este punto solo si el valor resulta ser mayor o menor al de todos sus vecinos, de esta manera se identifica el punto como máximo o mínimo local respectivamente. El costo de este proceso es bajo debido a que la mayoría de los puntos de muestreo serán eliminados después de los primeros controles. Un aspecto importante es determinar la frecuencia de muestreo en la imagen y la escala, necesaria para detectar de forma fiable los picos (máximos y mínimos).

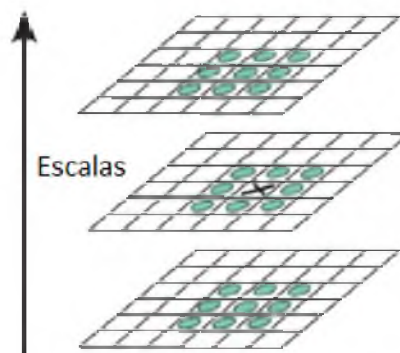


Figura 3.3: Localización de máximos y mínimos locales. Fuente David Lowe [13].

3.3.2. Localización de puntos clave

Los puntos característicos son seleccionados con base a las medidas de su estabilidad para descartar aquellos que no sean adecuados. Una vez que un punto característico candidato ha sido encontrado tras haberlo comparado con sus 26 vecinos, se realiza un estudio de su estabilidad. Para ello se debe almacenar toda la información de cada punto característico. Para cada punto encontrado se guardará a que octava y escala pertenece, así como su posición dentro de la imagen correspondiente.

Los puntos con bajo contraste o que estén localizados a lo largo de bordes serán descartados, debido a que, son más sensibles al ruido y por lo tanto, no podrían ser detectados bajo pequeños cambios de iluminación o variaciones en el punto de vista.

3.3.2.1. Eliminación de puntos de bajo contraste

Muchos de los puntos tienen un bajo contraste, por lo tanto, no serían estables a los cambios de iluminación. Para descartar estos puntos se utiliza una función evaluada en el extremo $D(\hat{x})$.

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x} \quad (3.7)$$

Donde D y sus derivadas son evaluadas en el punto de la muestra y $x = (x, y, \sigma)^T$ es el desplazamiento de este punto. Finalmente, si el valor de $D(\hat{x})$ es menor a 0.03, el punto es eliminado, suponiendo que D toma valores entre 0 y 1.

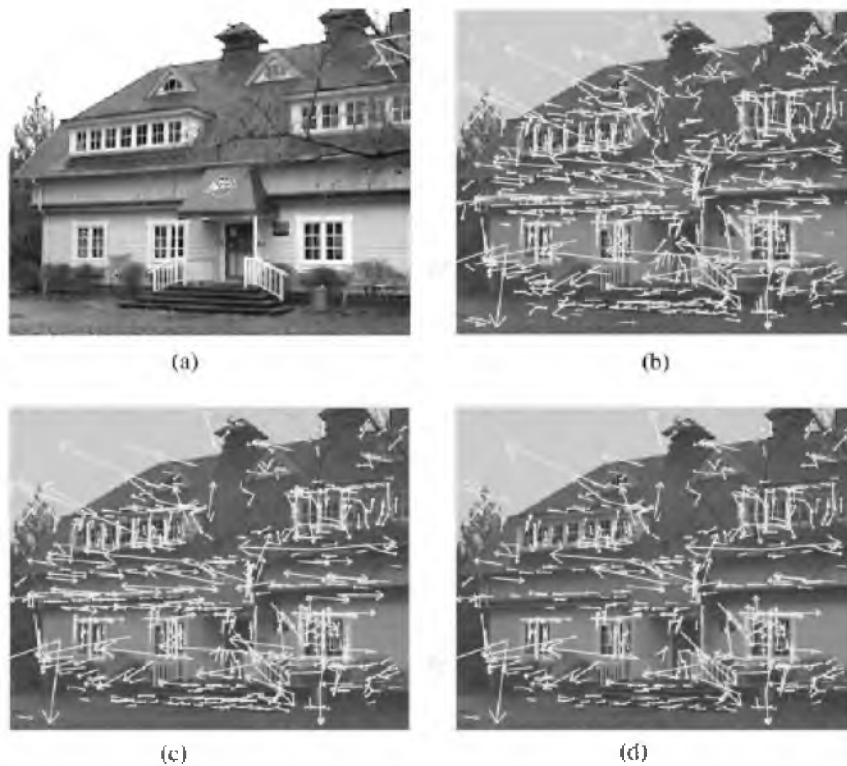


Figura 3.4: Etapas de selección de puntos característicos. Fuente David Lowe [13].

En la Figura 3.4 se muestran las etapas de selección de puntos característicos. (a) La imagen original 233 x 189 píxeles. (b) Los 832 puntos característicos obtenidos (máximos y mínimos) obtenidos mediante la función de Diferencia Gaussiana. (c) Muestra los 729 puntos característicos que permanecen después de aplicar un umbral mínimo de contraste ($D(\hat{x}) < 0,03$) a la imagen anterior. (d) Muestra los últimos 536 puntos característicos que permanecen después de un umbral adicional en relación de las curvaturas principales.

3.3.2.2. Eliminación de puntos localizados en los bordes

Además de quitar aquellos puntos con bajo contraste, se deben eliminar los puntos característicos que estén en los bordes. La función de Diferencia Gaussiana $D(x, y, \sigma)$ tiene una gran respuesta ante puntos situados sobre bordes, aunque la localización en estas zonas no está bien determinada y por lo tanto inestable a pequeñas cantidades de ruido. Un pico mal definido en la función de Diferencia Gaussiana tendrá una gran curvatura a través del borde, pero una pequeña en la dirección perpendicular. Las curvaturas principales se pueden calcular mediante una matriz Hessiana (3.8) de 2x2 sobre la localización y escala del punto en estudio:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (3.8)$$

Las derivadas son estimadas tomando las diferencias de puntos de muestreo vecinos. Los valores de H son proporcionales a las curvaturas principales de D . Siendo α el valor de mayor longitud y β el menor. Entonces, podemos calcular la suma de los valores propios de la traza de H (3.9) y su producto del determinante (3.10).

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (3.9)$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \quad (3.10)$$

En caso de que el determinante sea negativo, las curvaturas tienen un signo diferente al de los puntos descartados por no ser extremos. Sea r la relación entre las magnitudes mayor y menor de los valores propios, tal que $\alpha = r\beta$:

$$\frac{Tr(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(r + 1)^2}{r} \quad (3.11)$$

Dependerá de la relación entre los valores propios en lugar de sus valores individuales. La cantidad $(r + 1)^2/r$ es un mínimo cuando los dos valores propios son iguales y aumentan con

r . Por lo tanto, para comprobar que la relación de las curvaturas principales está por debajo del umbral r , sólo tenemos que comprobar si:

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \quad (3.12)$$

Lo anterior es muy eficiente de calcular, ya que se requieren menos de 20 operaciones de punto flotante para probar cada punto característico. Luego de varios experimentos, Lowe propone un umbral de $r = 10$.

3.3.3. Asignación de la orientación

En la asignación de la orientación se asignan una o más orientaciones a cada punto característico extraído, basándose en las direcciones locales de la imagen gradiente, es decir, en las variaciones de píxeles cercanos en las dos dimensiones.

La descripción de los puntos característicos se puede representar mediante la asignación de una orientación a cada punto característico de la imagen basado en las propiedades locales de la misma. De esta manera se obtienen características invariantes a la rotación de la imagen.

La escala de los puntos característicos se utiliza para seleccionar la imagen Gaussiana suavizada (L), con la escala más cercana, de modo que todos los cálculos se realicen de una manera invariante a la escala. Para cada punto de la imagen $L(x, y)$, a esta escala, se calcula la magnitud del gradiente, $m(x, y)$ y su orientación $\theta(x, y)$, mediante las siguientes ecuaciones:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (3.13)$$

$$\theta(x, y) = \tan^{-1} \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \quad (3.14)$$

El histograma de direcciones se genera a partir de las orientaciones del gradiente de los puntos de muestreo dentro de una región que rodea al punto característico. Este histograma, tiene 36 contenedores que cubren el rango de 360 grados de las orientaciones, es decir, cada contenedor tiene 10 grados. Cada muestra añadida al histograma, se pondera por la magnitud del gradiente y por una ventana circular de ponderación Gaussiana con un σ que es 1.5 veces mayor que la escala del punto característico.

Los picos en el histograma de orientación corresponden a las direcciones dominantes de los gradientes locales. El pico más alto se detecta en el histograma, junto con los picos locales que estén dentro del 80% del pico más alto, éstos se utilizan para crear un punto clave en esa dirección. Por lo tanto, para localizaciones con muchos picos de magnitud similar, habrá múltiples puntos significativos creados en la misma posición y escala, pero con diferentes orientaciones. Sólo al 15% de los puntos se les asigna más de una orientación, pero éstas contribuyen significativamente a la estabilidad de las posteriores búsquedas.

3.3.4. Descriptor de puntos característicos

Los gradientes locales de la imagen se miden en la región que rodea al punto característico. Éstos son transformados mediante una representación que permitirá medir niveles de distorsión y cambios en la iluminación de forma local.

En las etapas anteriores del algoritmo SIFT se ha asignado a cada punto característico una escala, una ubicación y una orientación. Los resultados de estas etapas proporcionan invariabilidad a la localización en la imagen, la escala y a la rotación. Esta última etapa consiste en calcular un descriptor para cada punto clave. En la Figura 3.5 se muestra el cálculo del descriptor para características.

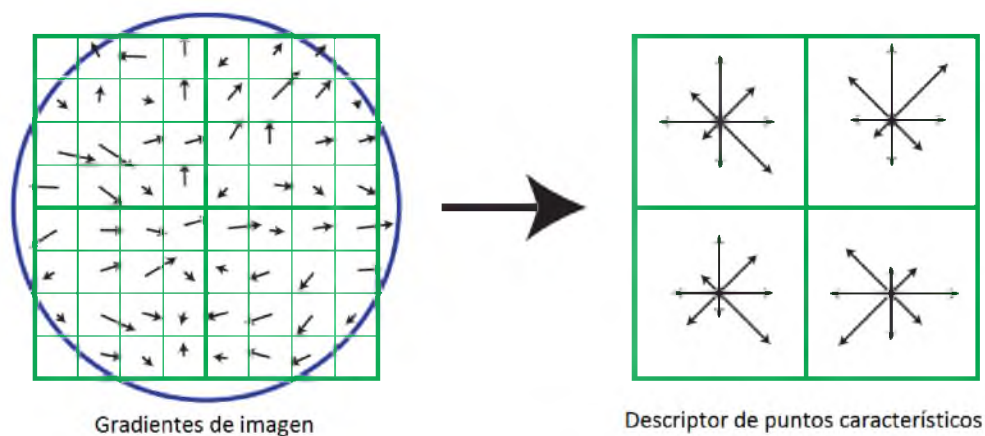


Figura 3.5: Descriptor de puntos característicos. Fuente David Lowe [13].

El primer paso consiste en calcular las magnitudes del gradiente y las orientaciones de la imagen alrededor del punto característico, utilizando la escala del punto para seleccionar el nivel de desenfoque Gaussiano para la imagen. Con la finalidad de lograr invarianza en la orientación, las coordenadas del descriptor y las orientaciones del gradiente son rotadas respecto a la orientación del punto característico. Para obtener una mayor eficiencia, los gradientes son calculados para todos los niveles de la pirámide. Éstos se ilustran con flechas pequeñas en el lado izquierdo de la imagen.

Una función de ponderación de Gauss con σ igual a la mitad del ancho de banda de la ventana del descriptor se utiliza para asignar un peso a la magnitud de cada punto de muestreo. Esto se ilustra con una ventana circular en la parte izquierdo de la Figura 3.5. El propósito de esta ventana gaussiana es evitar cambios repentinos en el descriptor con pequeños cambios en la posición de la ventana y dar menos énfasis a los gradientes que están lejos del centro del descriptor, ya que éstos son los más afectados por errores de registro erróneo.

El descriptor del punto característico se mostró en la parte derecha de la Figura 3.5. Permite un cambio significativo en las posiciones del gradiente mediante la creación de un histograma de orientación sobre una región de 4×4 de la muestra. Una muestra de gradiente a la izquierda de la Figura 3.5 puede cambiar hasta cuatro posiciones de la muestra mientras siga contribuyendo al mismo histograma a la derecha, logrando así el objetivo de permitir grandes cambios en la posición local.

El descriptor de cada punto característico se forma a partir de un vector que contiene los valores de las 8 orientaciones de los 4×4 histogramas componiendo un vector de características de $4 \times 4 \times 8 = 128$ elementos para cada punto. En el ejemplo de la la Figura 3.5 utilizan únicamente $2 \times 2 = 4$ subregiones en vez de $4 \times 4 = 16$. De acuerdo a las investigaciones de Lowe con la matriz de 4×4 se obtienen mejores resultados.

Finalmente, el vector de características es modificado para reducir los efectos de los cambios de iluminación. El objetivo del descriptor es ser invariante a tres tipos de variación, que son el contraste, el brillo y la iluminación no lineal.

En primer lugar, el vector es normalizado por una unidad de longitud. Un cambio en el contraste de la imagen en la que cada valor de píxel y gradiente son multiplicados por una constante, por lo que el cambio de contraste será cancelado por el vector de normalización.

El descriptor es invariante a los cambios de iluminación, debido a que el añadir una constante a cada píxel de la imagen, no afecta los valores de gradiente, ya que se calcula a partir de las diferencias entre píxeles vecinos. Sin embargo los cambios no lineales de iluminación son los más difíciles de controlar, y se pueden producir por la saturación de la cámara a cambios de iluminación sobre superficies 3D con distintas orientaciones y cantidades. Éstos problemas pueden afectar las magnitudes de algunos gradientes. Por lo tanto, se reduce la influencia de grandes magnitudes del gradiente mediante una umbralización de los valores en el vector de características a cada unidad que no sea mayor que 0.2 y posteriormente se renormalizará de nuevo la unidad de longitud.

Una vez realizadas estas modificaciones, se puede dar por finalizado el proceso de construcción de los descriptores. El algoritmo SIFT termina en este punto.

3.4. Discusión

En este apartado se presentaron los fundamentos teóricos y los conceptos relacionados con esta tesis, además se describió el método a utilizar para la detección de puntos característicos en imágenes. De acuerdo con la información presentada en este capítulo se concluye lo siguiente:

- El algoritmo SIFT es un detector y descriptor de puntos característicos.
- SIFT es uno de los métodos más eficientes para la detección de los puntos característicos.
- SIFT es un algoritmo que se utiliza ampliamente en el campo de la visión por computadora, sin embargo, tiene un alto costo computacional.

Capítulo 4

Análisis, diseño e implementación del sistema

En este capítulo se describe de manera general el análisis y diseño del proceso necesario para la detección de puntos característicos en imágenes digitales. Se presentan las etapas que lo componen, así como el diseño y la implementación del sistema, con base en el capítulo 3, en el cual se muestra la metodología propuesta para la detección de puntos característicos.

4.1. Análisis del Sistema

El análisis del sistema se define con base en los requerimientos necesarios para cumplir con el objetivo planteado en esta tesis. En el análisis se genera un modelo de representación del problema, de las etapas necesarias y de los métodos asociados para la detección de puntos característicos. De acuerdo a la propuesta de solución, presentada en el capítulo 2, en el análisis del sistema intervienen los siguientes elementos: la base de imágenes de edificios de la Universidad, la descomposición de la imagen en diferentes octavas y escalas, la diferencia Gaussiana y la extracción de mínimos y máximos locales, estos elementos se relacionan entre sí (como se muestra en la Figura 4.1) con la finalidad de desarrollar una arquitectura, el diseño y la implementación del sistema.



Figura 4.1: Elementos que intervienen en el análisis del sistema.

4.2. Diseño del sistema

Una vez que se analizó el problema, el siguiente paso es el diseño del sistema, en esta etapa se define su arquitectura, la cual cumple con los puntos establecidos durante la fase de análisis. En la siguiente sección se presenta la arquitectura general propuesta para el sistema de detección de puntos característicos en imágenes digitales.

4.2.1. Arquitectura del sistema

El esquema que se presenta a continuación muestra la descripción general del sistema, en éste se definen sus componentes, los procesos y las relaciones entre éstos. El esquema propuesto se divide en 3 etapas: la etapa 1 es la descomposición de la imagen en octavas y escalas, la etapa 2 aplicación de la diferencia gaussiana para las escalas de la imagen y la etapa 3 extracción de puntos característicos de la imagen. La Figura 4.2 muestra de manera general la arquitectura del sistema con las 3 etapas y los elementos que intervienen.

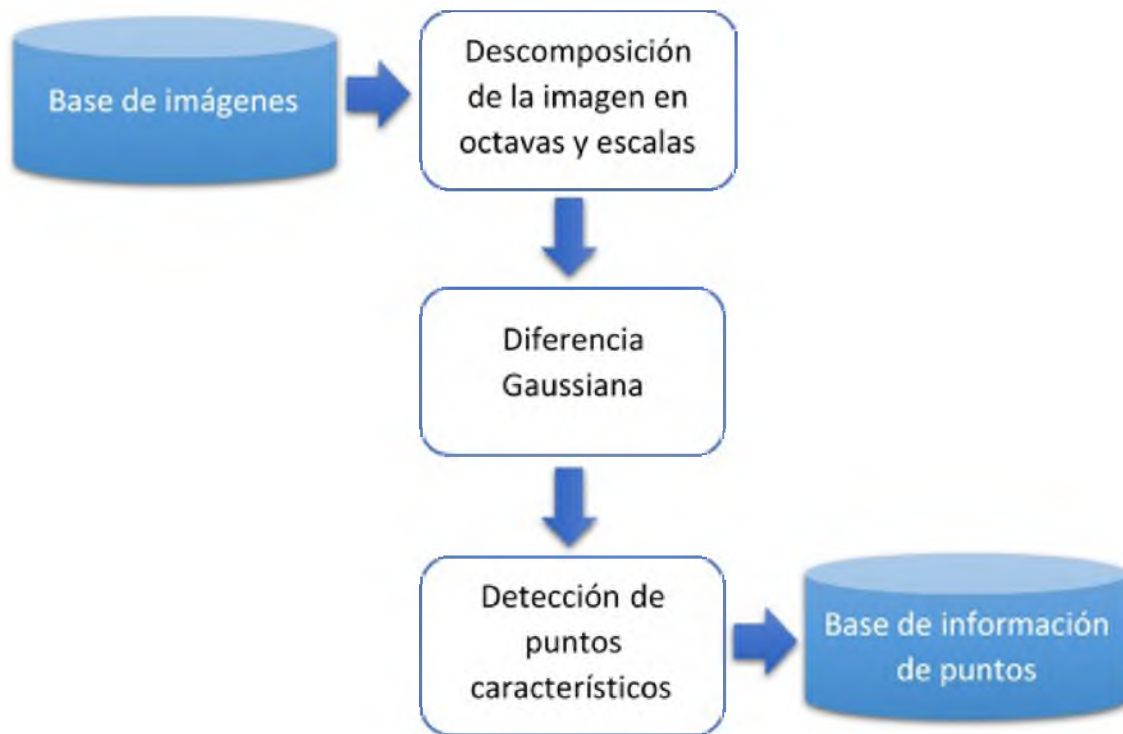


Figura 4.2: Arquitectura del sistema.

4.2.1.1. Etapa 1: Descomposición de la imagen en octavas y escalas

La etapa de descomposición de la imagen en octavas y escalas es un proceso que conlleva una serie de pasos descritos a continuación. En la Figura 4.3 se muestra de forma gráfica los pasos de esta etapa y posteriormente se detalla la secuencia de pasos a realizar.

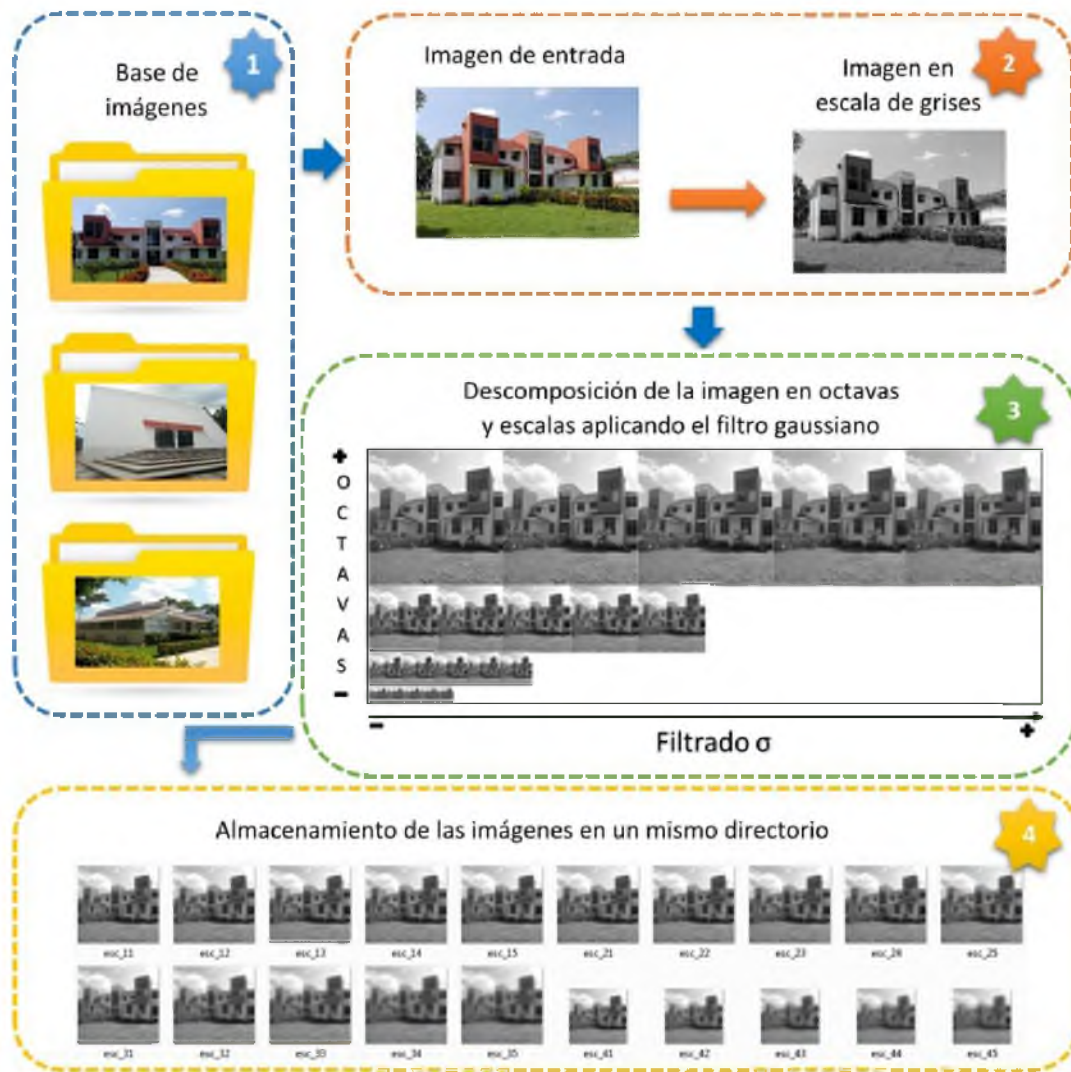


Figura 4.3: Pasos de la etapa 1.

Paso 1: Adquisición de las imágenes para la detección de puntos característicos. Las imágenes fueron capturadas con una cámara digital marca Kodak easysshare C1530 de 14 MP y cumplen con las siguientes especificaciones: imágenes del Instituto de Agroingeniería, Auditorio y Laboratorio de Mecatrónica con diferentes cambios de escala, perspectiva, con iluminación natural y variaciones en el fondo (variaciones en el color del cielo, iluminación y nubes). Se cuenta con una base de imágenes de edificios de 90 elementos, es decir 30 imágenes por categoría. En la Figura 4.4 se muestran 3 ejemplos de las categorías de imágenes.



Figura 4.4: Categoría de imágenes. (a) Instituto de Agroingeniería; (b) Auditorio; (c) Laboratorio de Mecatrónica.

Paso 2: Se selecciona una imagen del conjunto de imágenes y se convierte a escala de grises, es decir, la imagen de trabajo tendrá únicamente un canal con valores entre 0 y 255 (Figura 4.5).



Figura 4.5: Conversión de la imagen original a escala de grises. (a) Imagen original; (b) Imagen en escala de grises.

Paso 3: Descomposición de la imagen en escalas y octavas. En este paso la imagen seleccionada en el paso anterior es descompuesta en un conjunto de octavas y escalas. En este caso se obtendrán 4 octavas y 5 escalas, lo anterior es recomendado por Lowe [13] citado en el capítulo 3, donde se explica que es un método para la detección y descripción de puntos característicos. Para efectos de esta tesis, solo se aplicará la etapa de detección de puntos característicos.

La imagen original tiene una dimensión de 512x512 píxeles, esta es la primer octava que se tiene. Una vez que se tenga la primer octava, se procede a obtener las escalas de la imagen, para esto se utilizó el filtro Gaussiano paso bajo. Este filtro difumina las imágenes y elimina el ruido.

La imagen filtrada se obtiene al convolucionar la imagen en escala de grises con la máscara de 3x3 (4.1). Este proceso se repite hasta obtener 5 imágenes con diferente nivel de borrosidad, es decir, con diferente valor de σ .

$$g = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (4.1)$$

En la Figura 4.6 se observan las 5 escalas de la octava, esto también se conoce como desenfoque Gaussiano y se puede apreciar que los bordes del edificio se atenúan en cada escala.



Figura 4.6: Imgenes con aplicación de desenfoque Gaussiano

La siguiente octava se genera al reducir la octava anterior a la mitad de tamaño, es decir, si la primer octava tiene un tamaño de 512x512, la segunda tendrá un tamaño de 256x256, la tercera de 128x128 y la cuarta de 64x64. En la Tabla 4.2.1.1 se muestra la descomposición de la imagen para obtener las 4 octavas de la imagen y las 5 escalas de cada octava.


















				
Octava 3 con un tamaño de 128x128				
				
Octava 4 con un tamaño de 64x64				
				

Tabla 4.1: Descomposición de la imagen en octavas.

Paso 4: Almacenamiento de las imágenes. En este último paso se guardan todas las imágenes generadas por las 4 octavas y las 5 escalas, es decir, de una sola imagen en escala de grises se obtendrán 20 imágenes. Estas imágenes serán utilizadas en la siguiente etapa.

4.2.1.2. Etapa 2: Aplicación de la Diferencia Gaussiana

Esta etapa se realiza a partir de las imágenes que fueron generadas en la etapa uno. Con base en las imágenes se realiza una diferencia Gaussiana para cada octava de la imagen y se guardan todas las imágenes generadas por la diferencia Gaussiana en un mismo directorio para su uso en la siguiente etapa. En la Figura 4.8 se muestra de forma gráfica los pasos de esta etapa y posteriormente se detalla la secuencia de pasos a realizar.

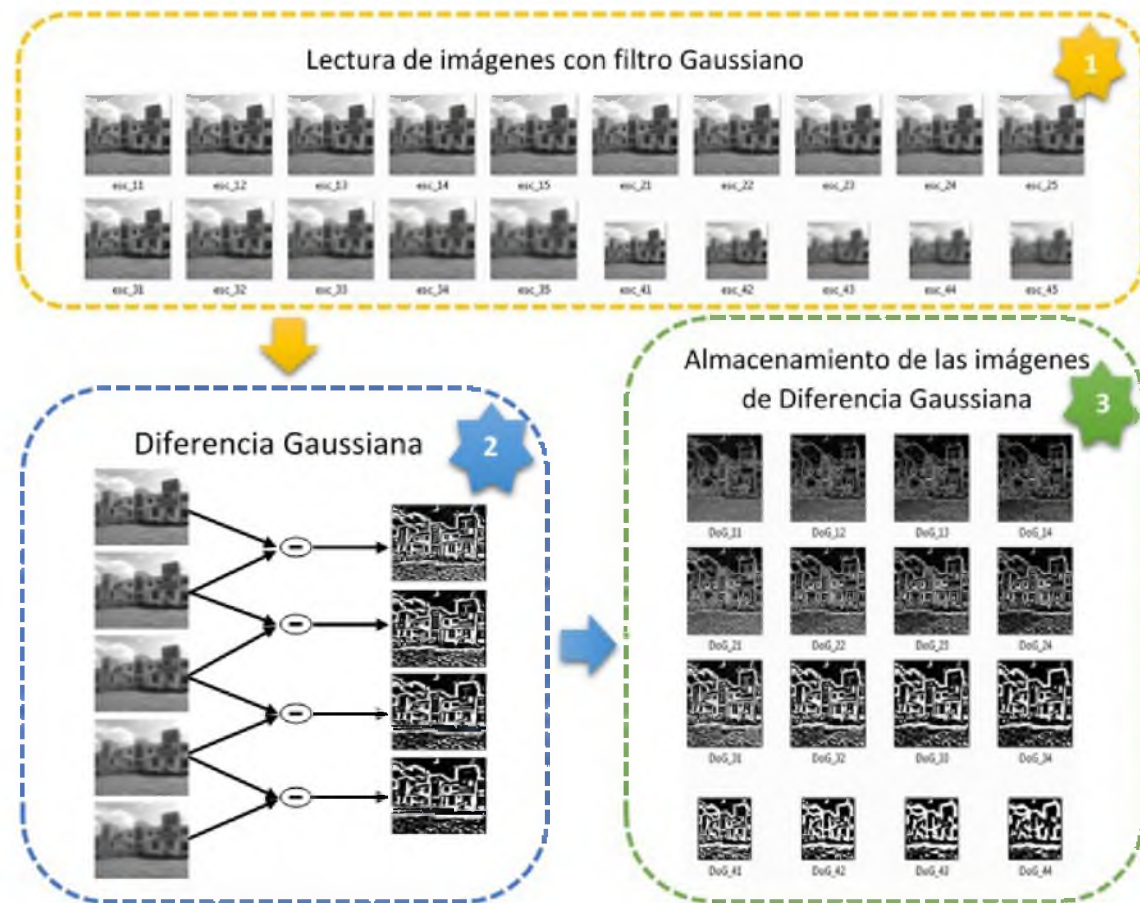


Figura 4.7: Pasos de la etapa 2

Paso 1: Lectura de las imágenes. En este paso se realiza la lectura de las imágenes con filtro Gaussiano que fueron obtenidas en la etapa 1.

Paso 2: Diferencia Gaussiana. En este paso se aplica la Diferencia Gaussiana $D(x, y, \sigma)$ para detectar de manera más eficiente la ubicación de los puntos característicos en el espacio escala. La función DoG no implica un alto coste computacional ya que se calcula restando dos imágenes vecinas que pertenezcan a una misma octava. En la Figura 4.8 se puede ver gráficamente lo explicado anteriormente, en este caso para una sola octava. El mismo proceso se realiza las octavas restantes.

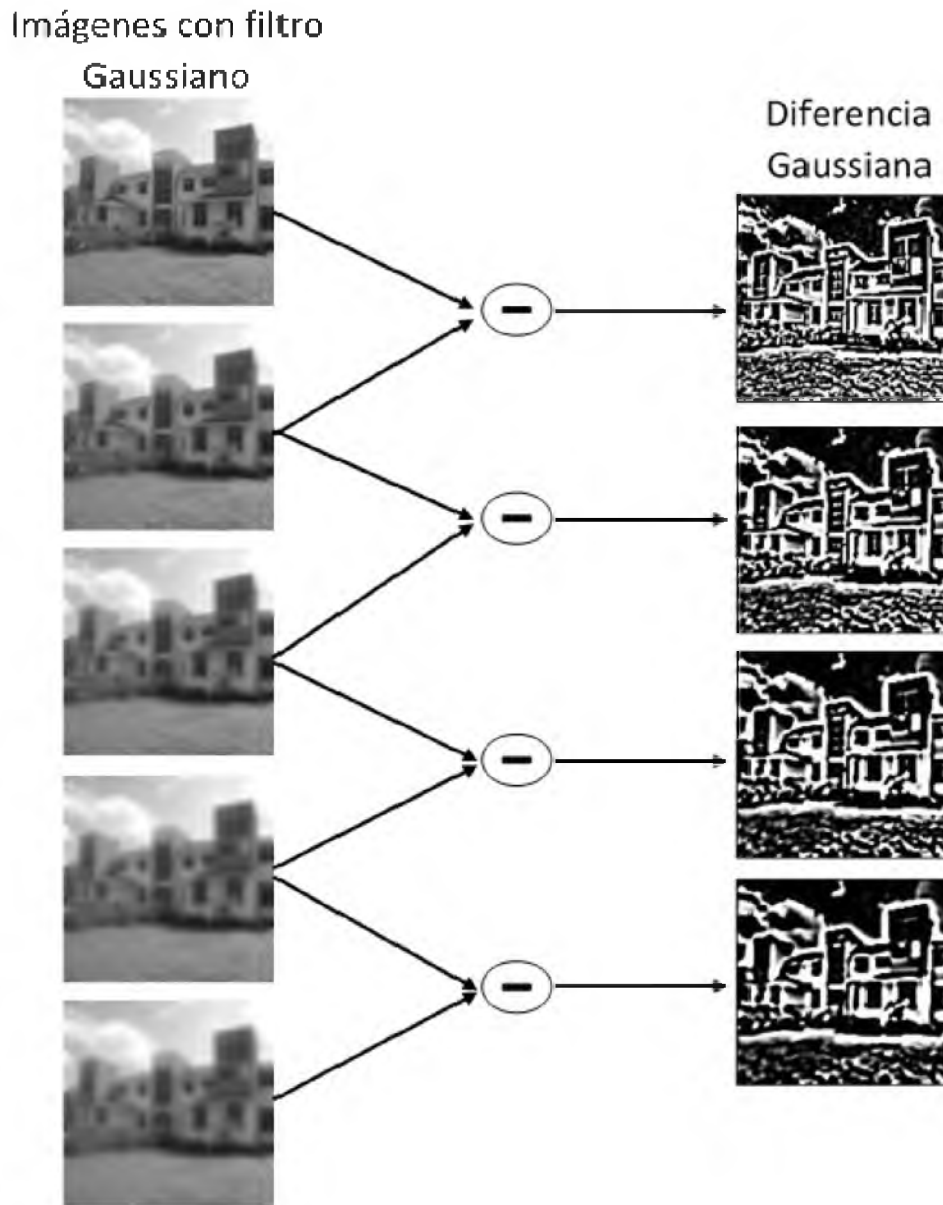


Figura 4.8: Proceso para obtener las imágenes de Diferencia Gaussiana $D(x, y, \sigma)$

En la Figura 4.9 se pueden observar las 16 imágenes resultantes de aplicar la Diferencia Gaussiana a las imágenes con filtro Gaussiano. A cada octava corresponden 4 imágenes.



Figura 4.9: Pirámide de Diferencia Gaussiana $D(x, y, \sigma)$

Paso 3: Almacenar las imágenes DoG. En este paso se almacenan en un mismo directorio las imágenes de Diferencia Gaussiana generadas en el paso anterior. Estas imágenes serán utilizadas en la siguiente etapa donde finalmente se extraerán los puntos característicos.

4.2.1.3. Etapa 3: Extracción de puntos característicos

Esta etapa se realiza a partir de las imágenes de Diferencia Gaussiana que fueron generadas en la etapa 2. Con base en esas imágenes se realiza la detección de puntos característicos máximos y mínimos. Los puntos detectados en esta etapa son almacenados en una matriz junto con la escala y la octava a la cual pertenecen. En la Figura 4.10 se muestra de forma gráfica los pasos de esta etapa y posteriormente se detalla la secuencia de pasos a realizar.

Paso 1: Lectura de imágenes DoG. Se leen las imágenes de Diferencia Gaussiana que fueron generadas en la etapa anterior.

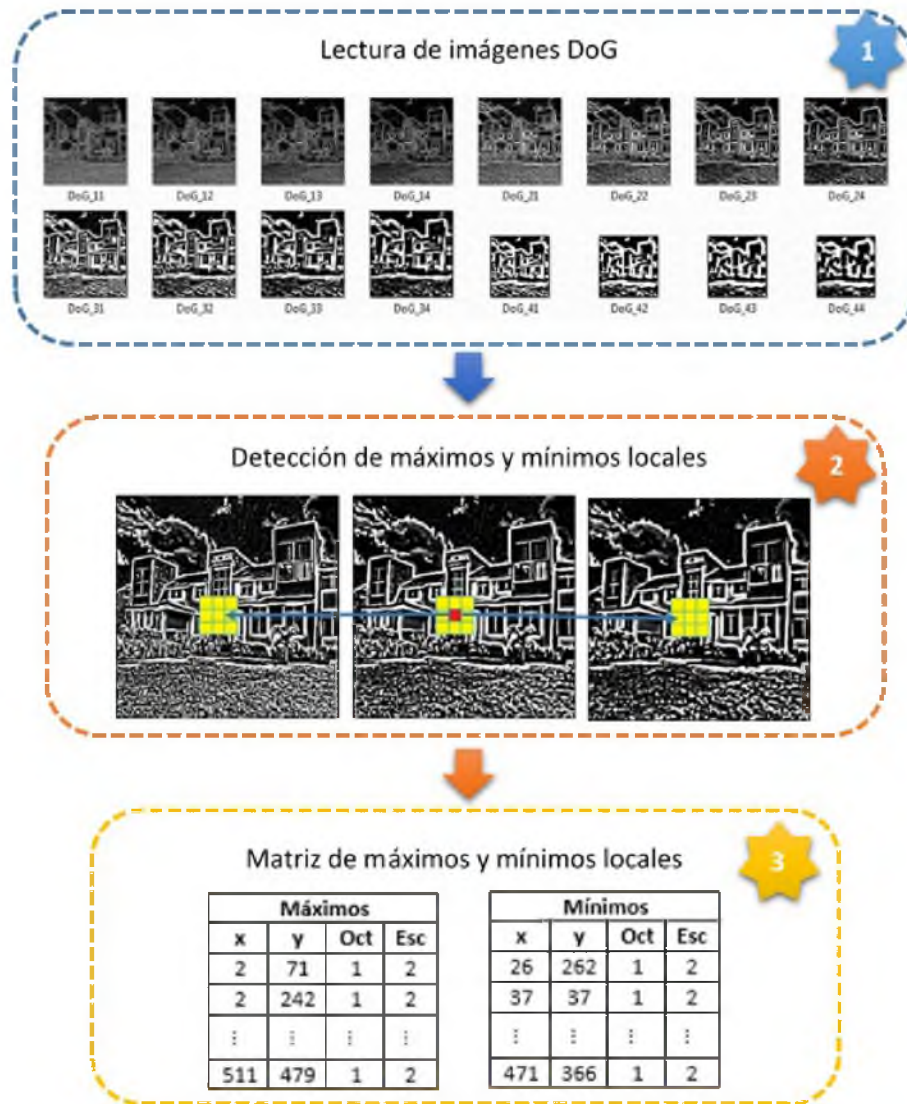


Figura 4.10: Pasos de la etapa 3

Paso 2: Detección de extremos en el espacio escala. A partir de las imágenes DoG generadas en la etapa anterior, se detectan los máximos y mínimos locales, para obtenerlos, se utilizan las imágenes centrales de cada octava, para cada uno de sus píxeles se considera su vecindad-8. Cada vecindad es comparada con las vecindades vecinas, una de la escala anterior y otra posterior. El píxel central de la vecindad-8 en total es comparado con 26 píxeles vecinos (8 vecinos de la misma escala, 9 de la escala anterior y 9 de la escala posterior). En la Figura 4.11 se muestra gráficamente lo explicado, se consideró a las primeras 3 escalas de la primer octava.



Figura 4.11: En rojo: píxel de estudio. En amarillo: 26 vecinos (8 de la escala actual, 9 de la escala anterior y 9 de la escala posterior).

Un píxel será seleccionado como punto característicos solo si es mayor que sus 26 vecinos o menor que todos ellos. En la Figura 4.12 se ejemplifica la detección de un máximo local y en la 4.13 la de un mínimo local. El píxel de estudio en ambos casos es el de color amarillo, y el resto son los píxeles vecinos con los que será comparado. Para el primer ejemplo el píxel de estudio tiene un valor de 14 mientras que, sus vecinos tiene un valor menor a este, lo que indica que es un punto máximo. En el siguiente ejemplo el valor del píxel de estudio es de 236 mientras que sus vecinos tienen un valor mayor a este, por lo tanto, se considera como un punto mínimo.

DoG1			
	56	57	58
50	0	0	0
51	5	9	0
52	0	0	4

DoG2			
	56	57	58
50	11	0	5
51	0	14	0
52	11	0	0

DoG3			
	56	57	58
50	0	5	6
51	3	1	3
52	0	0	0

Figura 4.12: Detección de máximo local.

DoG1			
	57	58	59
27	255	255	239
28	253	255	249
29	252	255	254

DoG2			
	57	58	59
27	254	255	253
28	255	236	244
29	251	255	240

DoG3			
	57	58	59
27	253	255	255
28	255	247	239
29	255	237	255

Figura 4.13: Detección de mínimo local.

El proceso descrito anteriormente se repite para cada octava de la imagen, al tener solo 4 escalas de la imagen, solo las imágenes centrales podrán ser comparadas con la anterior y posterior, de lo contrario la imagen de la escala 1 no tendría una imagen anterior y la imagen de la escala 4 no tendría una posterior. De lo anterior resultan solo 8 imágenes DoG con puntos característicos y no las 16 que se tenían al iniciar la tercer etapa.

Paso 3: Matriz de puntos máximos y mínimos locales. Se tendrá una matriz para máximos y una para mínimos, donde se guardarán las coordenadas del punto, además de la octava y la escala a la cual pertenecen. En la Figura 4.14 se muestran las imágenes junto con los puntos detectados en ellas.

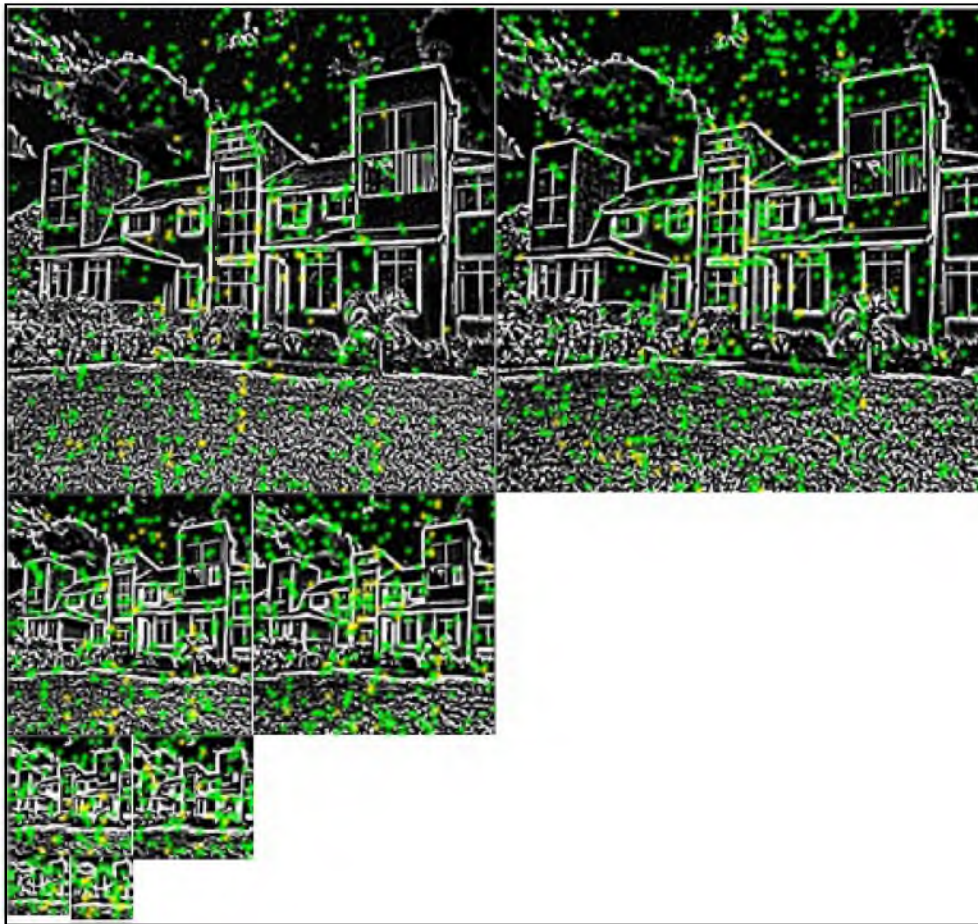


Figura 4.14: Detección de puntos. En verde puntos máximos, en amarillo puntos mínimos.

4.3. Diseño e implementación del sistema

En este apartado se describe el ambiente, el diseño y la implementación de las etapas de descomposición de la imagen, diferencia Gaussiana y la detección de puntos característicos descritos en el capítulo 4.2

4.3.1. Descripción del ambiente de implementación

El algoritmo para la detección de puntos característicos, desarrollado en esta tesis, fue implementado en *Matlab* 2014. Para su implementación se utilizó solo una función para convertir una imagen a escala de grises. El resto del código se implementó siguiendo los pasos del método propuesto para la solución del problema. La interacción del usuario con el algoritmo descrito se realiza de forma directa, introduciendo el nombre del directorio de la imagen a procesar y seleccionando la ejecución del programa. El esquema de interacción del usuario con el sistema en esta etapa se muestra en la Figura 4.15. Las 3 etapas del programa fueron desarrolladas en *Matlab*.



Figura 4.15: Esquema de interacción del usuario con el sistema.

4.4. Discusión

El análisis y diseño del sistema determinó los elementos que intervienen para la solución del problema y las etapas necesarias para la implementación del sistema, éstas son: descomposición de la imagen en octavas y escalas, aplicación de la diferencia Gaussiana y la detección de puntos característicos, las 3 etapas fueron implementadas en el entorno de trabajo *Matlab*.

Además de presentar la arquitectura del sistema, en la sección 4.3.1 se muestra la forma en que el usuario puede interactuar con el sistema y realizar las actividades que se incluyen en la arquitectura del sistema.

Capítulo 5

Experimentos y resultados

En este capítulo se presentan las pruebas realizadas al sistema de detección de puntos característicos en imágenes digitales. Estas pruebas tienen por objetivo saber cual es la mejor octava y escala para la detección de puntos característicos en imágenes de un ambiente natural.

Además de presentantar las pruebas, al final del capítulo se realiza un análisis de los resultados obtenidos.

5.1. Ambiente de pruebas

En esta sección se muestran las características del equipo de cómputo utilizado para realizar los casos de prueba del sistema y la base de imágenes que se utilizó en las pruebas.

La base de imágenes que se utilizó para las pruebas del sistema tiene un formato JPG y son imágenes de tres categorías: Instituto de Agroingeniería, Laboratorio de Mecatrónica y Auditorio, fueron obtenidas mediante captura personal. El número de imágenes a probar por cada categoría es de 10.

El equipo de cómputo en el que se realizaron las pruebas tiene las siguientes características:

- Procesador: Intel(R) Core(TM) i5-3210M, 2.50 GHz.
- Memoria RAM: 4.00 GB
- Sistema Operativo de 64 bits, Windows 7.

5.2. Pruebas

5.2.1. Prueba 1: Detección de puntos característicos en imágenes de la categoría Agroingeniería

Objetivo: Evaluar el desempeño del algoritmo que se aplicó para la detección de puntos característicos en imágenes de la categoría Agroingeniería. En la Figura 5.1 se muestran las imágenes que se utilizaron en esta prueba.



Figura 5.1: Categoría 1. Instituto de Agroingeniería.

Procedimiento: En esta prueba se utilizaron 5 imágenes, a cada una se le aplicó una descomposición en octavas y escalas, después se realizó una diferencia Gaussiana y por último se detectaron los puntos máximos y mínimos que caracterizan a la imagen. El número de subimágenes por cada imagen de referencia es de 20. Al aplicar la diferencia Gaussiana el número de subimágenes se reduce a 16. Una vez realizada la detección de puntos característicos, se obtienen 8 subimágenes.

Resultados: En las pruebas realizadas se obtuvieron los puntos característicos por cada imagen de referencia. A partir del número de puntos se realiza un promedio de las 4 octavas de la escala 2 y 3. En las Tablas 5.1, 5.2, 5.3, 5.4 y 5.5 se muestran el número de puntos máximos y mínimos detectados en cada imagen. En La Figura 5.2 se muestra el resultado de la detección de puntos característicos de la Figura 5.1(a).

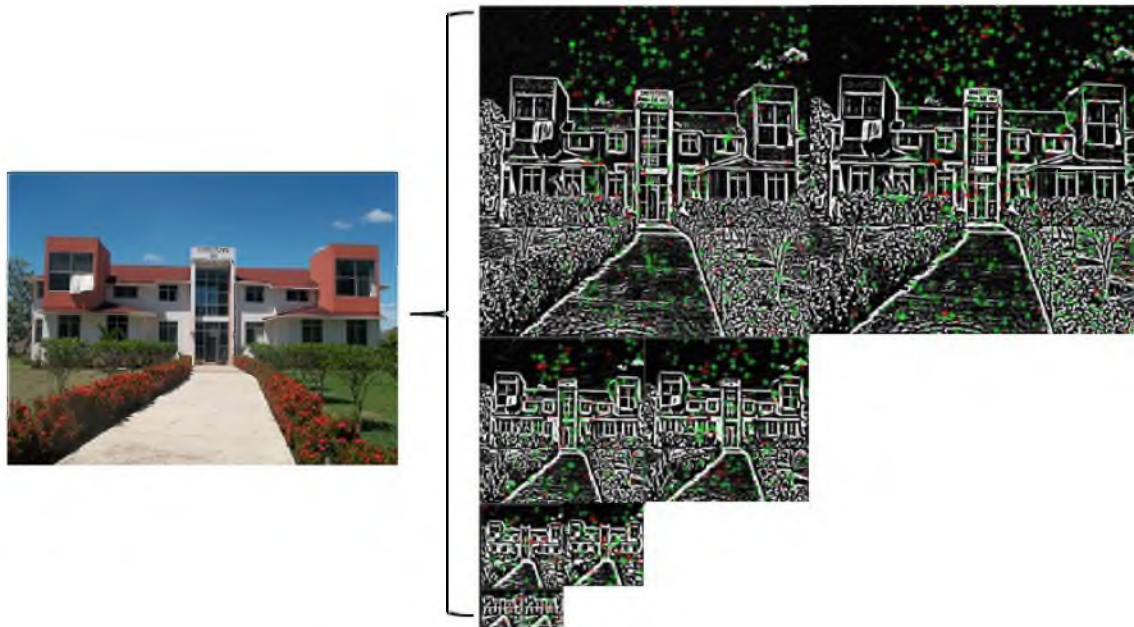


Figura 5.2: Detección de puntos característicos de la categoría Agroingeniería.

A continuación se muestran las tablas del número de puntos característicos detectados en las imágenes de referencia utilizadas para esta prueba.

Figura 5.1(a)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	452	145	30	4	631	157.75
	Escala 3	690	189	59	17	955	238.75
Mínimos	Escala 2	78	35	11	5	129	32.25
	Escala 3	98	38	17	5	158	39.5

Tabla 5.1: Número de puntos detectados en la Figura 5.1(a).

Figura 5.1(b)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	414	124	39	6	583	145.75
	Escala 3	721	206	65	14	1006	251.5
Mínimos	Escala 2	82	27	21	2	132	33
	Escala 3	108	38	16	4	166	41.5

Tabla 5.2: Número de puntos detectados en la Figura 5.1(b).

Figura 5.1(c)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	430	147	41	14	632	158
	Escala 3	712	234	60	15	1021	255.25
Mínimos	Escala 2	91	40	12	5	148	37
	Escala 3	93	35	11	8	147	36.75

Tabla 5.3: Número de puntos detectados en la Figura 5.1(c).

Figura 5.1(d)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	407	150	31	8	596	149
	Escala 3	658	208	50	17	933	233.25
Mínimos	Escala 2	87	32	16	3	138	34.5
	Escala 3	108	44	15	8	175	43.75

Tabla 5.4: Número de puntos detectados en la Figura 5.1(d).

Figura 5.1(e)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	419	117	39	6	581	145.25
	Escala 3	743	210	55	11	1019	254.75
Mínimos	Escala 2	78	26	6	3	113	28.25
	Escala 3	87	27	7	5	126	31.5

Tabla 5.5: Número de puntos detectados en la Figura 5.1(e).

En la Tabla 5.6 se muestra el número total de los puntos característicos que se obtuvieron para las imágenes de la categoría Agroingeniería utilizadas en esta prueba. El valor incluye el número total de puntos máximos y mínimos de la escala 2 y 3.

Imagen	Número total de puntos
5.1(a)	1873
5.1(b)	1887
5.1(c)	1949
5.1(d)	1842
5.1(e)	1839

Tabla 5.6: Número total de puntos característicos de la categoría Agroingeniería.

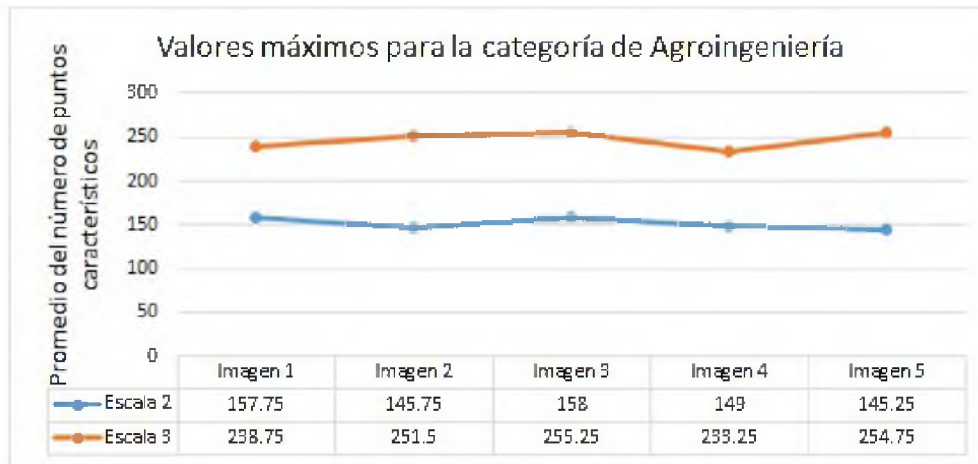


Figura 5.3: Valores promedio de los puntos máximos detectados en las imágenes de la Categoría Agroingeniería.

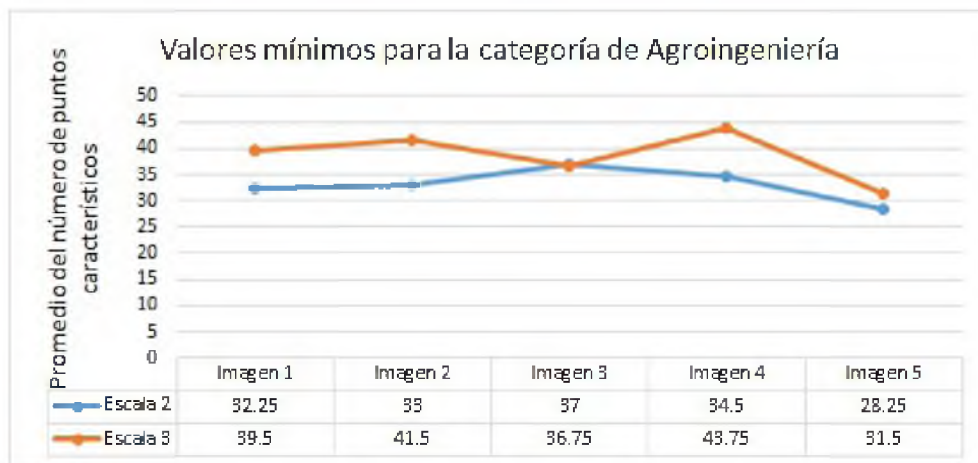


Figura 5.4: Valores promedio de los puntos mínimos detectados en las imágenes de la Categoría Agroingeniería.

5.2.2. Prueba 2: Detección de puntos característicos en imágenes de la categoría Mecatrónica

Objetivo: Evaluar el desempeño del algoritmo que se aplicó para la detección de puntos característicos en imágenes de la categoría Mecatrónica. En la Figura 5.5 se muestran las imágenes que se utilizaron en esta prueba.



Figura 5.5: Categoría 2. Laboratorio de Mecatrónica.

Procedimiento: En esta prueba se utilizaron 5 imágenes, a cada una se le aplicó una descomposición en octavas y escalas, después se realizó una diferencia Gaussiana y por último se detectaron los puntos máximos y mínimos que caracterizan a la imagen. El número de subimágenes por cada imagen de referencia es de 20. Al aplicar la diferencia Gaussiana el número de subimágenes se reduce a 16. Una vez realizada la detección de puntos característicos, se obtienen 8 subimágenes.

Resultados: En las pruebas realizadas se obtuvieron los puntos característicos por cada imagen de referencia. A partir del número de puntos se realiza un promedio de las 4 octavas de la escala 2 y 3. En las Tablas 5.7, 5.8, 5.9, 5.10 y 5.11 se muestran el número de puntos máximos y mínimos detectados para cada imagen de referencia. En La Figura 5.6 se muestra el resultado de la detección de puntos característicos de la Figura 5.5(a).

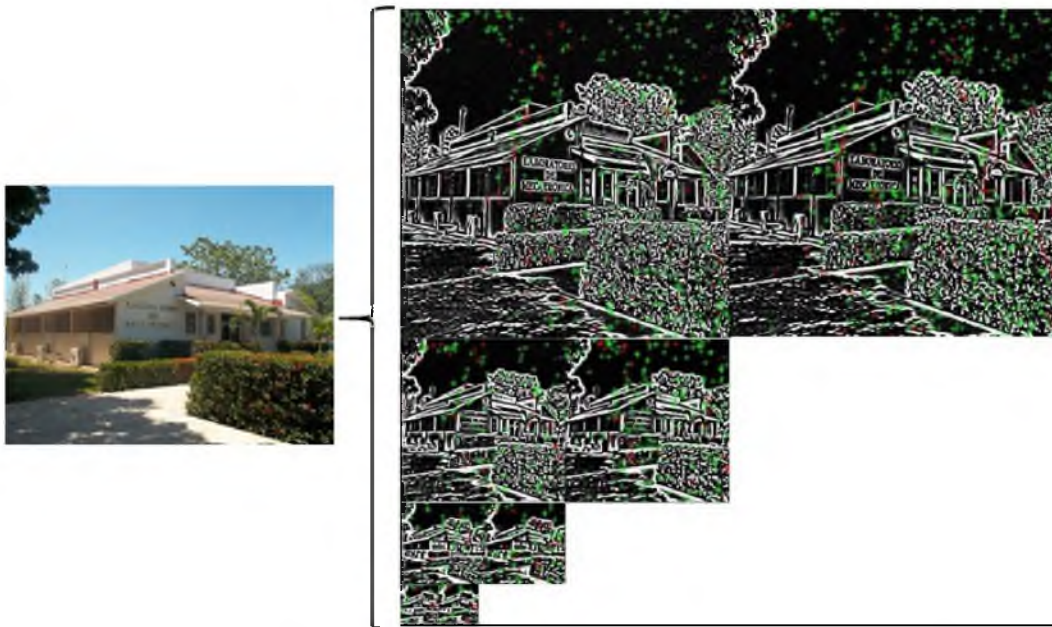


Figura 5.6: Detección de puntos característicos de la categoría Mecatrónica.

A continuación se muestran las tablas del número de puntos característicos detectados en las imágenes de referencia utilizadas para esta prueba.

Figura 5.5(a)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	506	126	42	11	685	171.25
	Escala 3	776	226	62	18	1082	270.5
Mínimos	Escala 2	92	24	6	3	125	31.25
	Escala 3	136	41	15	4	196	49

Tabla 5.7: Número de puntos detectados en la Figura 5.5(a).

Figura 5.5(b)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	485	141	29	10	665	166.25
	Escala 3	796	261	65	22	1144	286
Mínimos	Escala 2	106	33	12	4	155	38.75
	Escala 3	134	52	11	9	206	51.5

Tabla 5.8: Número de puntos detectados en la Figura 5.5(b).

Figura 5.5(c)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	558	155	39	13	765	191.25
	Escala 3	864	251	77	21	1213	303.25
Mínimos	Escala 2	115	40	19	4	178	44.5
	Escala 3	125	51	26	3	205	51.25

Tabla 5.9: Número de puntos detectados en la Figura 5.5(c).

Figura 5.5(d)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	498	159	38	18	713	178.25
	Escala 3	844	262	75	28	1209	302.25
Mínimos	Escala 2	105	22	6	2	135	33.75
	Escala 3	146	45	15	6	212	53

Tabla 5.10: Número de puntos detectados en la Figura 5.5(d).

Figura 5.5(e)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	427	105	38	11	581	145.25
	Escala 3	747	221	54	18	1040	260
Mínimos	Escala 2	69	20	4	2	95	23.75
	Escala 3	106	35	14	3	158	39.5

Tabla 5.11: Número de puntos detectados en la Figura 5.5(e).

En la Tabla 5.12 se muestra el número total de los puntos característicos que se obtuvieron para las imágenes de la categoría Mecatrónica utilizadas en esta prueba. El valor incluye el número total de puntos máximos y mínimos de la escala 2 y 3.

Imagen	Número total de puntos
5.1(a)	2088
5.1(b)	2170
5.1(c)	2361
5.1(d)	2269
5.1(e)	1874

Tabla 5.12: Número total de puntos característicos de la categoría Mecatrónica.

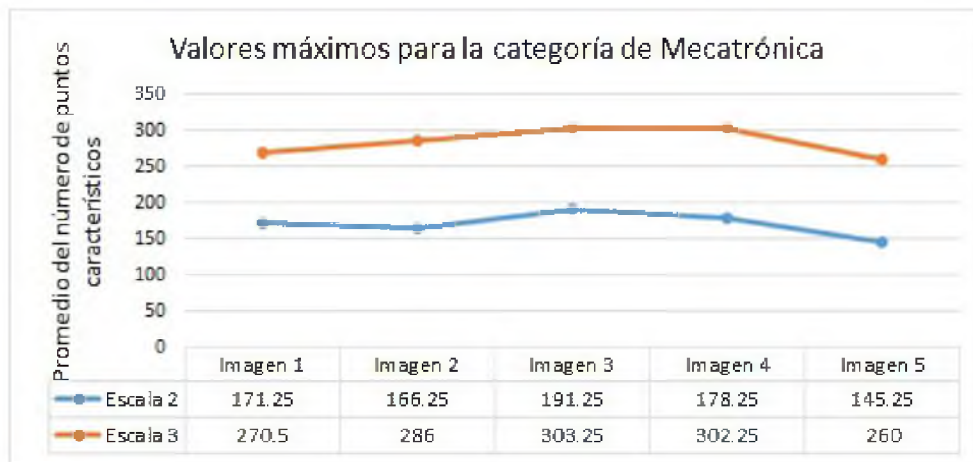


Figura 5.7: Valores promedio de los puntos máximos detectados en las imágenes de la Categoría Mecatrónica.

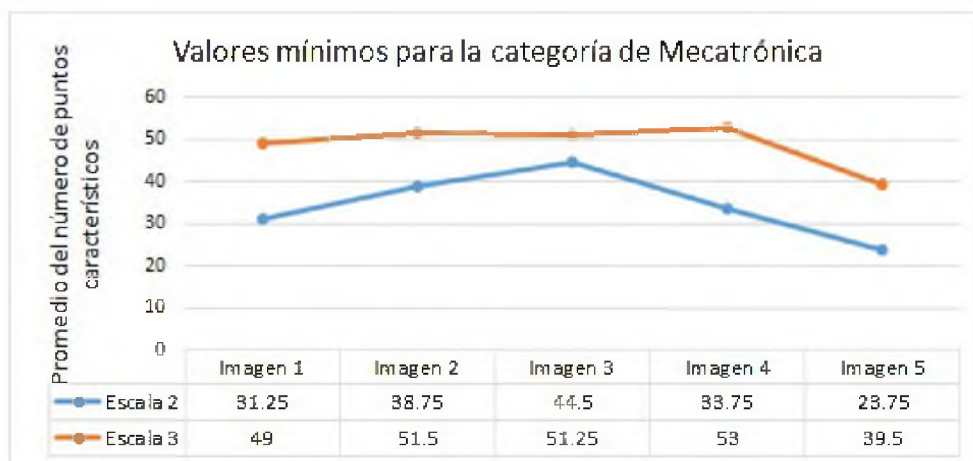


Figura 5.8: Valores promedio de los puntos mínimos detectados en las imágenes de la Categoría Mecatrónica.

5.2.3. Prueba 3: Detección de puntos característicos en imágenes de la categoría Auditorio

Objetivo: Evaluar el desempeño del algoritmo que se aplicó para la detección de puntos característicos en imágenes de la categoría Auditorio. En la Figura 5.9 se muestran las imágenes que se utilizaron en esta prueba.



Figura 5.9: Categoría 3. Auditorio.

Procedimiento: En esta prueba se utilizaron 5 imágenes, a cada una se le aplicó una descomposición en octavas y escalas, después se realizó una diferencia Gaussiana y por último se detectaron los puntos máximos y mínimos que caracterizan a la imagen. El número de subimágenes por cada imagen de referencia es de 20. Al aplicar la diferencia Gaussiana el número de subimágenes se reduce a 16. Una vez realizada la detección de puntos característicos, se obtienen 8 subimágenes.

Resultados: En las pruebas realizadas se obtuvieron los puntos característicos por cada imagen de referencia. A partir del número de puntos se realiza un promedio de las 4 octavas de la escala 2 y 3. En las Tablas (5.13, 5.14, 5.15, 5.16 y 5.17) se muestran el número de puntos máximos y mínimos detectados para cada imagen de referencia. En La Figura 5.10 se muestra el resultado de la detección de puntos característicos de la Figura 5.9(a).

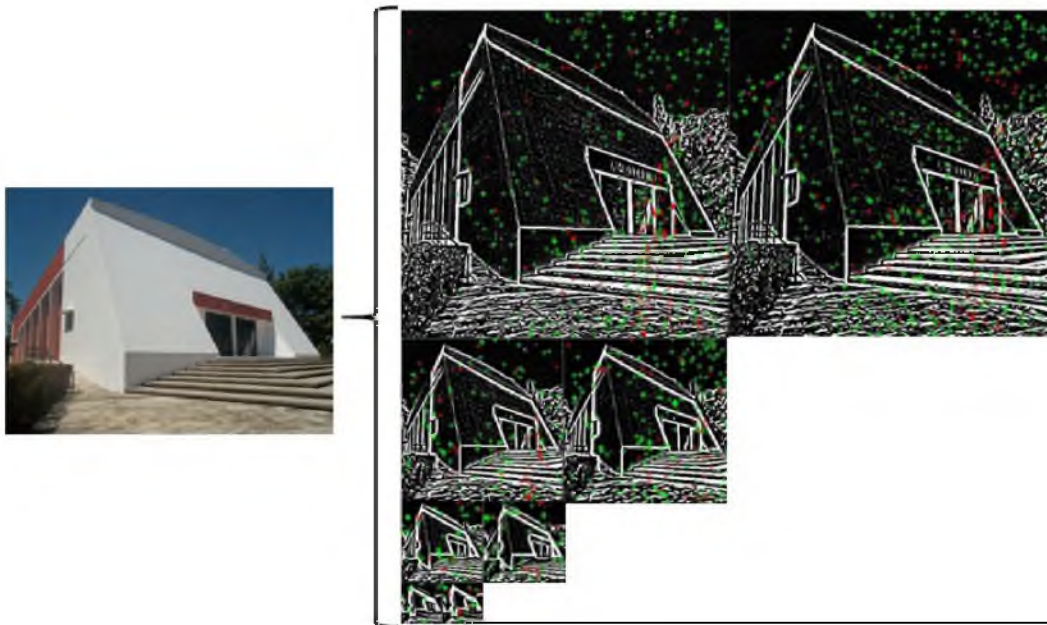


Figura 5.10: Detección de puntos característicos de la categoría Auditorio.

Figura 5.9(a)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	404	116	47	12	579	144.75
	Escala 3	649	194	53	14	910	227.5
Mínimos	Escala 2	127	40	10	0	177	44.25
	Escala 3	154	34	11	7	206	51.5

Tabla 5.13: Número de puntos detectados en la Figura 5.9(a).

Figura 5.9(b)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	378	108	40	12	538	134.5
	Escala 3	640	200	50	22	912	228
Mínimos	Escala 2	105	49	6	1	161	40.25
	Escala 3	166	71	10	4	251	62.75

Tabla 5.14: Número de puntos detectados en la Figura 5.9(b).

Figura 5.9(c)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	398	117	26	11	552	138
	Escala 3	632	186	59	18	895	223.75
Mínimos	Escala 2	142	58	8	2	210	52.5
	Escala 3	186	61	8	9	264	66

Tabla 5.15: Número de puntos detectados en la Figura 5.9(c).

Figura 5.9(d)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	427	104	40	16	587	146.75
	Escala 3	642	186	51	18	897	224.25
Mínimos	Escala 2	116	48	11	5	180	45
	Escala 3	158	67	17	2	244	61

Tabla 5.16: Número de puntos detectados en la Figura 5.9(d).

Figura 5.9(e)		Octava 1	Octava 2	Octava 3	Octava 4	Total	Promedio
Máximos	Escala 2	428	106	30	10	574	143.5
	Escala 3	672	182	50	13	917	229.25
Mínimos	Escala 2	124	66	10	2	202	50.5
	Escala 3	157	58	23	3	241	60.25

Tabla 5.17: Número de puntos detectados en la Figura 5.9(e).

En la Tabla 5.18 se muestra el número total de los puntos característicos que se obtuvieron para las imágenes de la categoría Auditorio utilizadas en esta prueba. El valor incluye el número total de puntos máximos y mínimos de la escala 2 y 3.

Imagen	Número total de puntos
5.1(a)	1872
5.1(b)	1862
5.1(c)	1921
5.1(d)	1908
5.1(e)	1934

Tabla 5.18: Número total de puntos característicos de la categoría Auditorio.

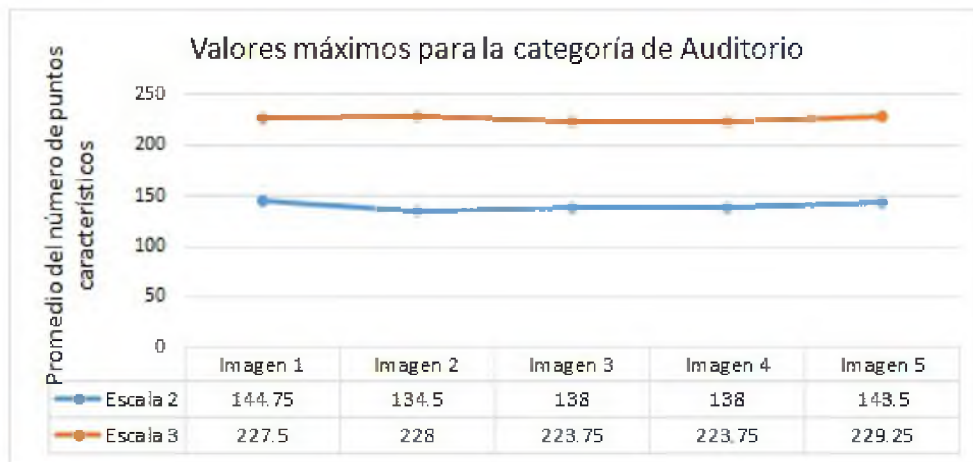


Figura 5.11: Valores promedio de los puntos máximos detectados en las imágenes de la Categoría Auditorio.

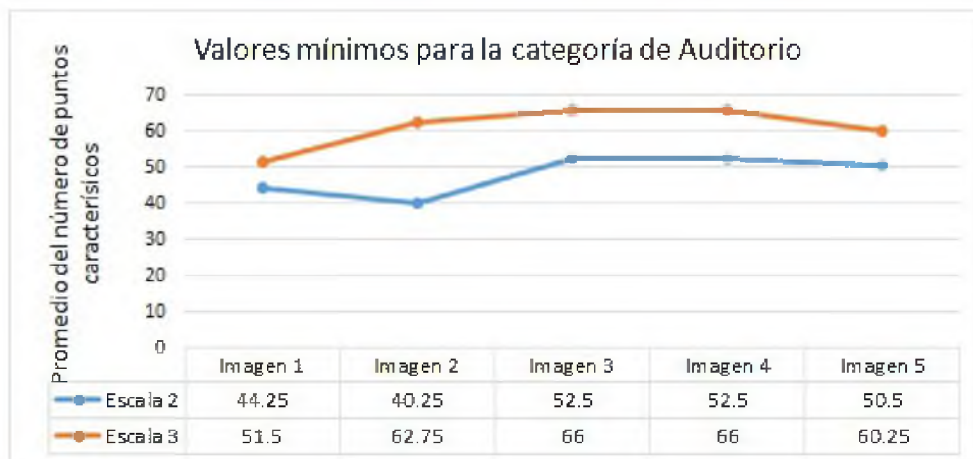


Figura 5.12: Valores promedio de los puntos mínimos detectados en las imágenes de la Categoría Auditorio.

5.2.4. Gráfica del total de puntos detectados

En la Figura 5.13 se muestra la gráfica del total de puntos característicos que fueron detectados en las imágenes de referencia de las 3 categorías con las que se trabajó en esta tesis. En la gráfica se puede observar que en la categoría de Mecatrónica existe una mayor variación en el número de puntos detectados a diferencia de la categoría Agroingeniería y Mecatrónica. Esto se debe a que las imágenes de Mecatrónica varían con respecto al contenido de cada una.

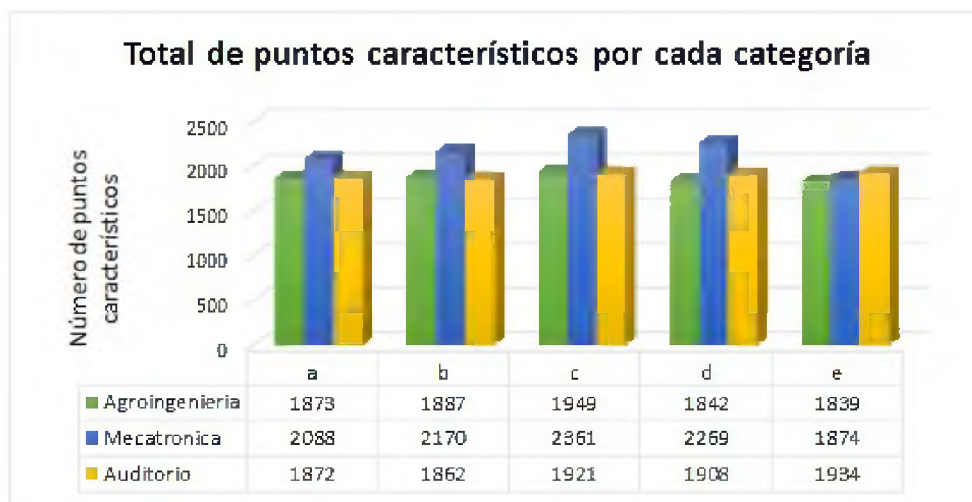


Figura 5.13: Número de puntos característicos detectados por cada categoría.

5.3. Discusión

En este capítulo se realizó la descripción de los casos de prueba realizados y la comparación de los resultados obtenidos. Se realizó una prueba aplicada a cada categoría.

En la prueba aplicada a la categoría Agroingeniería se tiene un total de 1873 puntos para la imagen 5.1(a), 1887 para la imagen 5.1(b), 1949 para la imagen 5.1(c), 1842 para la imagen 5.1(d) y 1839 para la imagen 5.1(e). También se observa que las escalas donde se detecta una mayor cantidad de puntos característicos es en la escala 3, tanto para los valores máximos, como para los valores mínimos. Esto se debe a que la escala número 3 toma las escalas 2 y 4 para realizar los cálculos de puntos máximos y mínimos. En estas imágenes los bordes se acentúan más, en comparación con la primer escala.

En los resultados de la prueba aplicada a la categoría de Mecatrónica se obtuvo un total de 2088 puntos para la imagen 5.5(a), 2170 para la imagen 5.5(b), 2361 para la imagen 5.5(c), 2269 para la imagen 5.5(d) y 1874 para la imagen 5.5(e). En esta prueba se observó que el mayor número de puntos máximos y mínimos también se obtuvo en la escala 2 y 3. Se realizó una comparación y se observó que en la categoría mecatrónica se obtuvo el mayor número de puntos característicos en comparación con las otras categorías.

La última prueba se aplicó a la categoría Auditorio, en esta se obtuvo un total de 1872 puntos característicos para la imagen 5.9(a), 1862 para la imagen 5.9(b), 1921 para la imagen 5.9(c), 1908 para la imagen 5.9(d) y 1934 para la imagen 5.9(e). Al igual que con las otras dos pruebas, se obtuvo un mayor número de puntos característicos en la escala 3 de cada octava.

Capítulo 6

Conclusiones generales

En este capítulo se presentan las conclusiones finales, las aportaciones y los trabajos futuros que se derivan de este trabajo de tesis.

6.1. Conclusiones

Con base en las pruebas realizadas se determinan las siguientes conclusiones:

- Se observó que al aplicar el algoritmo para la detección de puntos característicos, las imágenes que contienen varios elementos (áreas de texturas como vegetación, nubes y otros elementos como aves y sombras) presentaron una mayor cantidad de puntos característicos detectados.
- El tamaño de la imagen es un factor a considerar en el total de puntos característicos detectados. Cuanto más grande es el tamaño de la imagen, mayor será el número de puntos característicos y el tiempo de ejecución del algoritmo.
- En los experimentos que se aplicaron a la categoría Agroingeniería, se utilizaron 5 imágenes, las cuales reportan los siguientes valores totales de puntos característicos de 1873 Figura 5.1(a), 1887 Figura 5.1(b), 1949 Figura 5.1(c), 1842 Figura 5.1(d) y 1839 Figura 5.1(e). Éstos valores determinan que existe variación en el contenido de las imágenes a pesar de que pertenecen a la misma categoría. Esto se debe a que las imágenes fueron capturadas en un ambiente no controlado y por lo tanto existen elementos (como las nubes, las sombras y el viento) que cambian de posición o alteran el contenido de la imagen. Las imágenes que reportan los valores más altos de puntos característicos, mantienen una uniformidad de los elementos que contienen.

- En los experimentos que se realizaron con imágenes de la categoría Mecatrónica, se utilizaron 5 imágenes, las cuales reportan los siguientes valores totales de puntos característicos de 2088 Figura 5.5(a), 2170 Figura 5.5(b), 2361 Figura 5.5(c), 2269 Figura 5.5(d) y 1874 Figura 5.5(e). De acuerdo a los resultados se observó que en la Figura 5.5(c) se detectó un mayor número de puntos característicos. Este resultado se debe a que en la imagen interviene una mayor cantidad de elementos (como aves, nubes, sombras, vegetación y cambios de iluminación) por esta razón es la categoría con más puntos detectados.
- En los resultados que se obtuvieron de las pruebas aplicadas a la categoría Auditorio, se utilizaron 5 imágenes, las cuales reportan los siguientes valores totales de puntos característicos de 1872 Figura 5.9(a), 1862 Figura 5.9(b), 1921 Figura 5.9(c), 1908 Figura 5.9(d) y 1934 Figura 5.9(e). En esta categoría, se detectaron más puntos característicos en la última imagen con la que se experimentó. En esta imagen existe una mayor concentración de puntos característicos en el cielo y en algunas sombras que se presentan. Esto se debe a que los niveles de azul del cielo varían en la imagen y la cantidad de sombra de los elementos aumenta en comparación con las demás imágenes.
- Con base en los experimentos que se realizaron, se observó que los resultados coinciden con el comportamiento que se reporta en los trabajos que se revisaron en el estado del arte.

6.2. Aportaciones

- La principal aportación del presente trabajo es la implementación de un algoritmo que detecta puntos característicos en imágenes de escenas naturales.
- Creación de una base de datos de 3 categorías de imágenes de la Universidad del Papaloapan, campus Loma Bonita, las cuales fueron adquiridas mediante captura personal.

6.3. Trabajos futuros

- Implementar el descriptor de puntos característicos SIFT a partir del algoritmo que se implementó en esta tesis.
- Hacer pruebas de detección de puntos característicos que incluyan imágenes con cambios de escala y rotación.

- Implementar otros métodos de detección de puntos característicos para comparar los resultados obtenidos.
- Implementar un sistema de reconocimiento de patrones que incluya un número mayor de imágenes de ejemplo para realizar el entrenamiento del sistema y el algoritmo de detección de puntos característicos que se aplicó en este trabajo de tesis.
- Utilizar un mayor número de imágenes de ejemplos, para realizar el entrenamiento del sistema, con el objetivo de implementar un sistema de reconocimiento de imágenes.

Bibliografía

- [1] A. Constanzo, I. Amerini, R. Caldelli, and M. Barni. Forensic Analysis of SIFT Keypoint Removal and Injection. *IEEE Transactions on Information Forensics and Security*, 2014.
- [2] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin. SAR-SIFT: A Sift-Like Algorithm for Sar Images. *IEEE Transactions on Geoscience and Remote Sensing*, 2015.
- [3] L. Dorado and M. Vélez. A Vector Sift Detector for Interest Point Detection in Hyperspectral Imagery. 2012.
- [4] M. Duarte and L. Chang. Clasificación de Objetos en Imágenes Usando SIFT. Master's thesis, INAOE.
- [5] J. G. Enebral. Detección y Asociación Automática de Puntos Característicos para Diferentes Aplicaciones., 2009.
- [6] C. Haris and M. Sthephens. A Combined Corner and Edge Detector. *Fourth Alvey Vision Conference*, 1988.
- [7] M. Herigert, N. Bouchet, and I. Pianetti. Reconocimiento de Imágenes Mediante Scale Invariant Feature Transformation (sift).
- [8] M. C. Juárez. Reconocimiento e Identificación de Logotipos en Imágenes con Transformada SIFT, 2011.
- [9] J.J. Koenderink and A.J. Van Doorn. Spatiotemporal integration in the detection of coherent motion. *Vision Res*, 1984.
- [10] L. Larcher, E. Biasoni, C. Catteno, A. Ruggeri, and C. Herrera. Algoritmo para Detección de Bordos y Ulterior Determinación de Objetos en Imágenes Digitales. *Asociación Argentina de Mecánica Computacional*, 2011.

- [11] T. Lindeberg. Scale-Space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, 1994.
- [12] D.G. Lowe. Object recognition from local scale-invariant features. *International Conference on Computer Vision*, 1999.
- [13] D.G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *Internacional Journal of Computer Vision*, (2):91–110, 2004.
- [14] H.J. Mesa, J.W. Branch, and C. López. Sistema de detección de esquinas en imágenes digitales en tonos de gris, basado en redes neuronales artificiales. *Avances en Sistemas e Informática*, 2008.
- [15] K. Mikolajczyk and C. Schmid. An affine invariant interest point detector. *In European Conference on Computer Vision (ECCV)*, 2002.
- [16] G. Pajares and J.M. De la Cruz. *Visión por computador: Imágenes digitales y aplicaciones*. 2 edition, 2008.
- [17] G. Pajares, J.M. De la Cruz, J.M. Molina, J. Cuadrado, and A. López. *Imágenes Digitales Procesamiento Práctico con Java*. 1 edition, 2004.
- [18] B. Rister, C. Wang, M. Wu, and J. R. Callardo. A fast and efficient sift detector using the mobile gpu. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013.
- [19] L. D. Shinfeng, L. Jia, and C. Cheng. Using Gradient Features from Scale-Invariant Keypoints on Face Recognition. *International Journal of Innovative Computing, Information and Control*, 2011.
- [20] A.P. Witkin. Scale-space filtering. *International Joint Conference on Artificial Intelligence*, 1983.
- [21] M. Zhenwei and J. Xudong. Interest point detection using rank order LoG filter. *Elsevier Patern Recognition*, 2013.