



UNIVERSIDAD DEL PAPALOAPAN

Campus Loma Bonita

INGENIERÍA EN COMPUTACIÓN

GENERACIÓN DE NUEVAS IMÁGENES CON DIFERENTES
PUNTOS DE VISTA A PARTIR DE UNA IMAGEN DE REFERENCIA

TESIS

QUE PARA OBTENER EL TÍTULO DE:
INGENIERO EN COMPUTACIÓN

PRESENTA:

JOSÉ LUIS SÁNCHEZ DEL VALLE

ASESOR DE TESIS:

M.C. ALMA ALHELÍ PEDRO PÉREZ

LOMA BONITA, OAXACA.

2016



UNIVERSIDAD DEL PAPALOAPAN

INGENIERÍA EN COMPUTACIÓN

LA PRESENTE TESIS TITULADA “GENERACIÓN DE NUEVAS IMÁGENES CON DIFERENTES PUNTOS DE VISTA A PARTIR DE UNA IMAGEN DE REFERENCIA” PRESENTADA POR EL SUSTENTANTE DE LICENCIATURA **C. JOSÉ LUIS SÁNCHEZ DEL VALLE** 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 INGENIERO 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 Norma E. Del Valle, por darme la vida, por haberme apoyado en todo momento, por sus consejos, por sus valores, por la motivación constante que me ha permitido ser una persona de bien, pero más que nada, por su amor y cariño. Te quiero mucho mamá.

A mi segunda madre María Luisa Rodríguez, que con la sabiduría de Dios me has enseñado a ser quien soy hoy, gracias por tus consejos, por el amor que me has dado y por el apoyo incondicional en mi vida. Te quiero mamá licha.

A mi familia quienes por ellos soy lo que soy. Los quiero.

Agradecimientos

Gracias a Dios por haberme permitido llegar hasta este punto y haberme dado salud para lograr mis objetivos, además de su infinita bondad y amor.

Quiero agradecer de todo corazón a mi mamá Norma, a mi hermana Silvia, a mi abuela mamá Licha, a mi papá Arturo, a mi tía Alicia, a mi tía Bertha, a mi tía Elva, a mi tío Abel, a mi tío Agustín, a mi tía Josefa, a mi primo Fabio, a mi primo Edgar, a mi prima Saraí, a mi prima Luz, a mi sobrino Jorge Arturo, a mi sobrina Gisselle Guadalupe, a mi sobrino Fabio Maximiliano, por su amor, por su cariño, por su apoyo y comprensión. Es una alegría inmensa ser parte de esta familia.

Agradezco profundamente a mi asesor, la M.C. Alma Alhelí Pedro Pérez, por la oportunidad que me dio de trabajar a su lado y haberme guiado a la culminación de esta tesis. Gracias Maestra por confiar en mí, por tenerme paciencia, así como ayudarme y apoyarme en los momentos más difíciles. Gracias por ser una amiga para mí. Siempre le estaré agradecido.

Gracias a los miembros del comité revisor de esta tesis, el M.C. Eduardo Ortiz Hernández, el M.C. Rafael F. González Zárate y el Dr. Eduardo Sánchez Soto, por todo el tiempo dedicado a la corrección de esta tesis, así como sus invaluable comentarios y sus valiosos consejos para el desarrollo de este trabajo.

Gracias a mis compañeros y amigos, con el cual compartí gran parte de mi vida durante estos 5 años de carrera, en donde vivimos todo tipo de experiencias. Gracias por su apoyo y comprensión: Carmen(mari), Leydy (pilin), Zuleima (la Sra. de latin), Marleny, Alejandro (abuelo), Genaro (Gepe), Ezequiel (cheke), Arturo (latin), Christian (winbux), Luis Enrique (kike), Nieto (akatsuki). Gracias a Carmen por estar a mi lado, por darme su apoyo y comprensión incondicional, gracias por tu amor y hacerme pasar agradables momentos en mi vida.

Resumen

En esta tesis se aborda el problema que consiste en generar nuevas imágenes con diferentes puntos de vista a partir de una imagen de referencia. La técnica empleada para generar el cambio de vista es la geometría proyectiva. Esta técnica conlleva una descomposición de matrices, que se aplica a una imagen de referencia.

Para aplicar el cambio de vista en una imagen se inicia desde la captura de imágenes de los edificios de Agroingeniería, Biblioteca y Auditorio de la Universidad del Papaloapan. Como segundo paso se elige una imagen y se seleccionan cuatro puntos en la imagen. Estos puntos van en el sentido de las manecillas del reloj para los cambios de vista hacia la izquierda y en el sentido contrario para los cambios de vista hacia la derecha, formando un rectángulo. Por cada punto se obtienen sus coordenadas (x, y) , con estos puntos se genera una matriz de 8×8 . Esta matriz se multiplica por la matriz de traslación y de rotación. A partir de esta multiplicación se obtiene una matriz proyectiva. En el paso tres, la imagen se descompone en tres canales RGB, por cada canal se le aplica la matriz proyectiva. Finalmente se regenera la imagen y se obtiene el cambio de vista.

En la etapa experimental se emplearon 30 imágenes de las categorías de Agroingeniería, Biblioteca y Auditorio. Se realizaron 10 pruebas, seis son pruebas básicas en las que se aplicaron dos cambios de vista a una imagen por cada una de las tres categorías. También se realizaron dos pruebas más, en las que se integraron imágenes de las tres categorías, una con un área de selección grande y otra con un área de selección pequeña.

En las seis pruebas básicas, los resultados muestran que al hacer el cambio de vista hacia la izquierda se observó un mejor desempeño que al hacer el cambio hacia la derecha. En las cuatro pruebas restantes que integraron imágenes de las tres categorías, se observó que al seleccionar una área pequeña el cambio de vista es más notorio que al hacerlo con un área grande.

Con base en las pruebas realizadas se determinó que el cambio de vista de una imagen hacia la derecha es más recomendable. También se recomienda utilizar un área de selección pequeña para realizar el cambio de vista. Con base en el trabajo que se realizó en esta tesis, se propone una metodología la cual se compone de las siguientes etapas: construcción de una base de imágenes, selección de cuatro puntos de referencia y la aplicación de una transformación proyectiva.

Abstract

This thesis addresses the problem of generating new images with different viewpoints from a reference image. The technique used to generate the change of view is projective geometry. This technique involves a matrix decomposition, which is applied to a reference image.

In order to change the view on an image, in the first step we started with images captured of the Agricultural Engineering buildings, the Library and Auditorium of the University of the Papaloapan. In the second step after an image was selected four points were selected of that image. The points were going in a clockwise position and changed to a counter clockwise position, forming a rectangle. For each point its coordinates (x,y) were obtained, with these points an 8x8 matrix was generated. This matrix was multiplied by the matrix of translation and rotation. From this multiplication a projective matrix was obtained. In step three, the image was broken down into three RGB channels; each channel was applied to a projective matrix. Finally the image was restored and the change of view was obtained.

In the experimental stage 30 images 10 of each the Agricultural Engineering building, the Library and Auditorium were used. A total of 10 tests were performed to the images of the three buildings, six basic tests in which two changes of view were applied to each of the images. Two tests were performed, in which images of each of the three were integrated, one with a large selection area and one with a small selection area.

In the six basic tests, the results showed that by changing the view to the left a better performance was observed than that of the right. In the remaining four tests it was observed that by selecting a small area the change of view was more noticeable than in doing so with a large area.

Based on the tests it was determined that the change of view of an image to the right is recommended. It is also recommended to select a small area rather than a large area to change the view of an image. Based on the work done in this thesis, a methodology which consists of the following steps is proposed: building a base of images, selection of four reference points and the application of a projective transformation.

Índice general

1. Planteamiento del problema	1
1.1. Introducción	1
1.1.1. Descripción del problema	2
1.1.2. Propuesta de solución	2
1.1.2.1. Objetivo general	3
1.1.2.2. Objetivos específicos	3
1.1.2.3. Alcances y limitaciones	3
1.2. Hipótesis	4
1.3. Justificación	4
1.4. Organización de la tesis	4
2. Estado del Arte	6
2.1. <i>Composition of novel views through an efficient image warping</i>	6
2.2. <i>A Hybrid Method of image Synthesis in IBR for Novel Viewpoints</i>	7
2.3. <i>3D Object Reconstruction with Calibrated Images</i>	8
2.4. <i>Changing Perspective in Stereoscopic Images</i>	8
2.5. <i>Superview 3D image warping for visibility gap errors</i>	9
2.6. <i>A novel multi-view image coding scheme based on view-warping and 3D-DCT</i>	10
2.7. Análisis del problema	12

2.7.1. Problemas relacionados con imágenes de edificios de la Universidad del Papaloapan	12
2.7.2. Propuesta de solución	13
2.7.3. Discusión	13
3. Marco Teórico	14
3.1. Modelo RGB	14
3.2. Representación de imágenes	16
3.3. Plano proyectivo	17
3.3.1. Puntos y líneas	17
3.3.1.1. Representación homogénea de líneas	17
3.3.1.2. Representación homogénea de puntos	17
3.3.2. Grados de libertad (dof)	18
3.4. Modelo de una cámara	18
3.4.1. Cámara CCD	19
3.4.2. Cámara Pinhole	19
3.4.3. Modelo de proyección central	20
3.5. Transformaciones Proyectivas en 2D	21
3.5.1. Transformaciones de líneas	23
3.6. Categorización de las Transformaciones Proyectivas	23
3.6.1. Transformación Isométrica (Euclidiana)	23
3.6.2. Transformación de similitud	24
3.6.3. Transformación Afín	25
3.6.4. Transformación Proyectiva	27
3.7. Descomposición de una transformación proyectiva	27
3.8. Discusión	28

<i>ÍNDICE GENERAL</i>	X
4. Análisis, Diseño e implementación del sistema	30
4.1. Análisis del sistema	30
4.2. Esquema del sistema	31
4.3. Proceso para generar imágenes con diferentes puntos de vista	32
4.4. Diseño e implementación del sistema	39
4.4.1. Descripción del ambiente de implementación	39
4.4.2. Diseño general del sistema	39
4.5. Discusión	41
5. Experimentos y resultados	42
5.1. Ambiente de pruebas	42
5.1.1. Hardware	42
5.1.2. Software	43
5.1.3. Base de imágenes	43
5.2. Pruebas con cambio hacia la derecha	43
5.2.1. Prueba para la categoría Agroingeniería con cambio hacia la derecha	43
5.2.2. Prueba para la categoría Biblioteca con cambio hacia la derecha	44
5.2.3. Prueba para la categoría Auditorio con cambio hacia la derecha	45
5.3. Pruebas con cambio hacia la izquierda	46
5.3.1. Prueba para la categoría Agroingeniería con cambio hacia la izquierda	47
5.3.2. Prueba para la categoría Biblioteca con cambio hacia la izquierda	47
5.3.3. Prueba para la categoría Auditorio con cambio hacia la izquierda	48
5.4. Pruebas para las tres categorías con los mismos puntos de referencia	49
5.4.1. Prueba 1: Con cambio hacia la derecha con un área de selección de 27,676.18 pixeles	49
5.4.2. Prueba 2: Con cambio hacia la izquierda con un área de selección de 29,220.27 pixeles	51

<i>ÍNDICE GENERAL</i>	XI
5.4.3. Prueba 3: Con cambio hacia la derecha con un área de selección de 59,098.27 pixeles	52
5.4.4. Prueba 4: Con cambio hacia la izquierda con un área de selección de 58,403.30 pixeles	54
5.5. Discusión	56
6. Conclusiones	58
6.1. Aportaciones	60
6.2. Trabajos futuros	60

Índice de figuras

3.1. Tetraedro de color RGB. Figura tomada de [2].	15
3.2. Imagen RGB.	15
3.3. Convención de ejes utilizadas para la representación de imágenes digitales.	16
3.4. Modelo geométrico de cámara pinhole. Figura tomada de [15].	19
3.5. Distorsiones que surgen de una transformación planar afín. (a) Rotación por $R(\theta)$. (b) Una deformación $R(-\phi)DR(\phi)$. Figura tomada de [15].	26
4.1. Elementos que intervienen en el análisis del sistema.	31
4.2. Esquema general del sistema.	32
4.3. Proceso general para generar nuevas imágenes.	33
4.4. Imagen de la biblioteca.	34
4.5. Descomposición de la imagen en canales RGB. (a) Matriz $mCanal_{Rojo}$. (b) Matriz $mCanal_{Verde}$. (c) Matriz $mCanal_{Azul}$	37
4.6. Transformación proyectiva por cada canal RGB. (a) Canal rojo en escala de grises. (b) Canal verde en escala de grises. (c) Canal azul en escala de grises.	38
4.7. Cambio de vista de la imagen de referencia.	38
4.8. Esquema de interacción del usuario con el sistema.	39
4.9. Diseño general del sistema: Proceso para generar imágenes con diferentes vistas.	40

5.1. Resultado al aplicar un cambio de vista hacia la derecha de la categoría Agroingeniería.	44
5.2. Resultado al aplicar un cambio de vista hacia la derecha de la categoría Biblioteca.	45
5.3. Resultado al aplicar un cambio de vista hacia la derecha de la categoría Auditorio. .	46
5.4. Resultado al aplicar un cambio de vista hacia la izquierda de la categoría Agroingeniería.	47
5.5. Resultado al aplicar un cambio de vista hacia la izquierda de la categoría Biblioteca.	48
5.6. Resultado al aplicar un cambio de vista hacia la izquierda de la categoría Auditorio.	49
5.7. Resultado al aplicar un cambio de vista hacia la derecha en las tres categorías. . .	50
5.8. Resultado al aplicar un cambio de vista hacia la izquierda en las tres categorías. . .	52
5.9. Resultado al aplicar un cambio de vista hacia la derecha en las tres categorías. . .	53
5.10. Resultado al aplicar un cambio de vista hacia la izquierda en las tres categorías. . .	55

Índice de tablas

2.1. Resumen de los trabajos revisados.	11
2.2. Categoría de imágenes de edificios usados para esta investigación sus características.	13
3.1. Resumen de Transformaciones Proyectivas 2D	28
5.1. Resultados de los puntos que se extraen en la imagen de referencia de la categoría Agroingeniería.	44
5.2. Resultados de los puntos que se extraen en la imagen de referencia de la categoría Biblioteca.	45
5.3. Resultados de los puntos que se extraen en la imagen de referencia de la categoría Auditorio.	46
5.4. Resultados de los puntos que se extraen en la imagen de referencia de la categoría Agroingeniería	47
5.5. Resultados de los puntos que se extraen en la imagen de referencia de la categoría Biblioteca.	48
5.6. Resultados de los puntos que se extraen en la imagen de referencia de la categoría Auditorio.	49
5.7. Resultados de los puntos que se extraen en la imagen de referencia de las tres categorías.	51

5.8. Resultados de los puntos que se extraen en la imagen de referencia de las tres categorías.	52
5.9. Resultados de los puntos que se extraen en la imagen de referencia de las tres categorías.	54
5.10. Resultados de los puntos que se extraen en la imagen de referencia de las tres categorías.	55

Capítulo 1

Planteamiento del problema

En este primer capítulo se describe el problema que da lugar a la presente investigación, el cual consiste en generar una imagen con un punto de vista diferente a partir de una imagen de referencia. También se expone la propuesta que comprende el objetivo general, los objetivos específicos, los alcances y limitaciones con los cuales se plantea una solución al problema. Finalmente, se presenta la organización de los capítulos que se desarrollaron en esta tesis.

1.1. Introducción

Las imágenes son representaciones visuales de un objeto real y se consideran elementos principales de la visión por computadora [5]. En el campo de la visión artificial existen aplicaciones que requieren del procesamiento digital de imágenes para realizar tareas de reconstrucción 3D, construcción y navegación de entornos virtuales. Estas aplicaciones utilizan métodos para generar diferentes puntos de vista a partir de una imagen de referencia. Estos métodos tienen como objetivo consumir un ilimitado número de imágenes a partir de las cuales se generan nuevos puntos de vistas [15] [14]. En el caso de la construcción y navegación de un entorno virtual, el número de imágenes es mayor que el número de imágenes empleado en tareas de reconstrucción 3D de objetos.

En la literatura existen investigaciones que abordan el tema de la generación de nuevas imágenes con diferentes puntos de vista, aplicando diferentes técnicas como son: líneas epipolares, deformación de la imagen, reconstrucción de puntos, deformación 2D y transformaciones proyectivas [4] [12] [9]. Estos trabajos se concretan en la reconstrucción 3D de estatuas y figuras prehispánicas, recorridos virtuales de interiores de edificios y exteriores.

Contar con métodos para generar nuevas vista ayudará a realizar tareas más especializadas, como el desarrollo de entornos virtuales. Por lo anterior el objetivo de esta investigación es: implementar un algoritmo para generar una imagen con un punto de vista diferente a partir de imágenes de referencia. Las imágenes con las que se trabajará son edificios de la Universidad del Papaloapan Campus Loma Bonita.

1.1.1. Descripción del problema

Una de las problemáticas que se presentan para generar nuevos puntos de vista en imágenes, es que se requiere un número extenso de imágenes y un equipo de cómputo con capacidad de proceso y memoria para almacenar todas las imágenes. Otro problema que se presenta es que al trabajar con imágenes de dos dimensiones, no se cuenta con información de profundidad que permita generar un nuevo cambio de vista de forma eficiente. El objetivo que se persigue en este trabajo es encontrar un proceso que permita generar un nuevo punto de vista a partir de una imagen digital de dos dimensiones.

1.1.2. Propuesta de solución

El problema que se pretende resolver consiste en implementar un algoritmo para la generación de nuevos puntos de vista, mediante una imagen de referencia. El algoritmo permitirá que el procesamiento de las imágenes sea más eficiente y rápido sin la necesidad de un número extenso de imágenes de referencia y un exceso en el procesamiento de cómputo.

1.1.2.1. Objetivo general

Implementar un algoritmo para generar una imagen con un punto de vista diferente a partir de una imagen de referencia.

1.1.2.2. Objetivos específicos

Los objetivos específicos que permitirán lograr el objetivo general se describen a continuación:

1. Revisar el estado del arte.
2. Estudiar los métodos para generar nuevas vistas de imágenes.
3. Seleccionar un método para generar el cambio de vista en la imagen.
4. Aplicar un método para generar el nuevo punto de vista.
5. Realizar pruebas y analizar resultados de los métodos implementados.

1.1.2.3. Alcances y limitaciones

Los alcances y limitaciones que presentan este proyecto de tesis son:

1. Se trabajará con un número limitado de imágenes.
2. Se trabajará con una sola cámara para la adquisición de las imágenes.
3. Se utilizarán imágenes de los edificios de la Universidad del Papaloapan campus Loma Bonita.
4. Se trabajará con al menos tres categorías de edificios de la Universidad del Papaloapan.
5. Se aplicará un método para generar un nuevo punto de vista en la imagen.

1.2. Hipótesis

El algoritmo realizará un cambio de vista de una imagen sin la necesidad de un número extenso de imágenes.

1.3. Justificación

El procedimiento para obtener diferentes puntos de vistas de una imagen, es contar con un número extenso de imágenes de referencia. Esto implica un mayor costo computacional. En este trabajo se requiere una imagen para generar una vista sin la necesidad de recurrir a un número extenso de imágenes. Además se obtiene una aplicación para generar nuevas vistas a partir de una imagen de referencia sin utilizar una aplicación comercial.

1.4. 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 antecedentes, así como los trabajos relacionados con la presente investigación. Finalmente se expone un análisis del problema y propuesta de solución.

Capítulo 3. Marco Teórico

En este capítulo se describen los conceptos principales en que se apoya el presente estudio, así como la teoría que permite la comprensión del desarrollo del trabajo de esta tesis.

Capítulo 4. Análisis, diseño y aplicación del sistema.

En este capítulo se analiza, diseña y aplica el sistema para la generación de nuevas imágenes con diferentes puntos de vistas, a partir de una imagen de referencias.

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 al generar un cambio de vista de una imagen de referencia.

Capítulo 6. Conclusiones

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

Finalmente se presenta a detalle cada una de las referencias bibliográficas que han servido como base para el desarrollo del presente trabajo de tesis.

Capítulo 2

Estado del Arte

En este capítulo se revisan investigaciones relacionadas al problema de la generación de nuevas imágenes con diferentes puntos de vista a partir de imágenes de referencias. Con base en la información recopilada se analiza el problema y se diseña una propuesta de solución.

2.1. Composition of novel views through an efficient image warping

En este trabajo de investigación se presenta un algoritmo de deformación-inversa eficiente, para la generación de nuevos puntos de vista mediante la combinación de múltiples imágenes de referencia [4].

El método aplicado procede en tres etapas. En la primera etapa, las imágenes de referencia son preprocesadas para extraer los píxeles del borde que se utilizarán para acelerar el proceso de una deformación. En la segunda etapa, con base en las características de las líneas epipolares, se realiza una deformación inversa a la imagen deseada a partir de una imagen de referencia. Solo los pares de los píxeles del borde de la imagen de referencia son necesarios para calcular u obtener una disparidad generalizada de todos los puntos en la imagen deseada. Se realiza un proceso de las pruebas de visibilidad para garantizar la exactitud de la imagen derivada en esta etapa. Por

último, en la tercer etapa los huecos en la imagen deseada se llenan mediante la búsqueda en los puntos correspondientes en otras imágenes de referencia. En esta etapa, se aplicó dos técnicas de aceleración, las cuales son: relleno de huecos por segmentos y posible recorte de disparidad.

El algoritmo fue implementado en C ++ en un sistema operativo Windows NT 4.0. Procesador Pentium II con memoria de 128 MB en RAM. Las imágenes se obtuvieron con una resolución de 400x400 pixeles, el algoritmo puede alcanzar una velocidad acerca de 2-3 fotogramas por segundos. Con base en los experimentos se observó que la complejidad algorítmica se redujo mediante el aprovechamiento de las características de la línea epipolar.

2.2. A Hybrid Method of image Synthesis in IBR for Novel Viewpoints

En este artículo presenta una solución eficiente para generar una nueva vista de objetos reales y sintéticos a partir de imágenes de referencia. Se basa en imágenes de modelado y renderizado. IBM indica un grupo de técnicas que usan imágenes como fuente de datos para la síntesis de una imagen [6].

El método propuesto combina los métodos de deformación hacia adelante y hacia atrás, para cumplir con las tareas de la composición de nuevas imágenes con un nuevo punto de vista. Se utilizaron diferentes métodos heurísticos para elegir una imagen principal de referencia y otra imagen para el relleno de huecos.

Un problema de la deformación 3D de una imagen son los huecos generados por los cambios de visibilidad cuando la imagen cambia a un nuevo punto de vista. Una solución a este tipo de problemas es utilizar varias imágenes de entrada con diferentes puntos de vista. El algoritmo ha sido implementado en C en una Índigo con memoria RAM de 64 MB. Las imágenes usadas están a una resolución de 400x400 pixeles.

2.3. 3D Object Reconstruction with Calibrated Images

Este artículo propone una arquitectura de solución al problema de reconstrucción de objetos 3D, así como la organización y selección de los métodos para obtener el modelo final. Para cumplir con el objetivo de este trabajo se propone un sistema de reconstrucción de objeto 3D a partir de un conjunto de imágenes con diferentes ángulos de visión del objeto [12].

El sistema está compuesto por dos módulos principales; el primero realiza el procesamiento de dos vistas, su objetivo es determinar el mapa de profundidad en un par de vistas, cada par de vistas tiene una secuencia de fases: detección de puntos de interés, correspondencia de puntos y reconstrucción de puntos; en el proceso de reconstrucción se determinan los parámetros que describen el movimiento (matriz de rotación R y el vector de traslación T) entre las dos vistas. El segundo módulo determina el modelo 3D incorporando los puntos de los pares de vistas de todos los puntos generados; una vez obtenido el mapa de profundidad total, se genera la malla 3D, aplicando el método de triangulación de Delaunay.

Para validar el sistema se emplearon imágenes de 640x480 píxeles captadas por una cámara en un ambiente donde el objeto se encuentra sobre una plataforma giratoria con un fondo negro. La cantidad de puntos característicos detectados influye directamente en la cantidad de puntos 3D reconstruidos y la calidad de la reconstrucción, es decir, mayores detalles del objeto. Por lo tanto, las áreas con más detalles están definidas por la densidad de puntos en el mapa.

2.4. Changing Perspective in Stereoscopic Images

Este artículo se propone una técnica para llevar a cabo la manipulación perspectiva en un par de imágenes estereoscópica, en donde el usuario especifica los parámetros de la cámara para la nueva vista [14].

En este trabajo, como primer paso se calcula una deformación 2D para cada imagen con un marco de optimización. En segundo lugar, se utilizaron dos tipos de correspondencias: puntos fuertes y segmentos de línea recta. Estos por lo general se pueden determinar con una confianza relativamente alta en un par estéreo. Por último utilizan la deformación a un par de imágenes estereoscópicas, a partir de técnicas de deformación 2D de una imagen. Estos métodos permiten el uso escaso de correspondencias a través del par de imágenes estereoscópicas; pueden ser más fácilmente y de forma fiable encontrado que un mapa de profundidad denso.

El algoritmo se ha probado con una variedad de imágenes estereoscópicas, produciendo resultados en menos de 10 segundos, utilizando una PC con 3.0 GHz, cuatro núcleos y 4 GB de RAM. El método produce resultados con disparidad casi horizontal, proporcionando una experiencia de visualización cómoda.

2.5. *Superview 3D image warping for visibility gap errors*

En este trabajo, se analiza y propone un esquema de deformación 3D basado en un campo de visión más extenso, para eliminar los errores de visibilidad que se producen con frecuencia con el método IBR. Estos errores son causados por el limitado campo de visión de una imagen 2D normal [16].

En el método de deformación 3D, la imagen de referencia es más grande que la imagen principal, esto se realiza para reducir o eliminar errores de visibilidad, sin embargo al deformar una imagen de mayor tamaño aumenta el tiempo de renderizado. Con el fin de abordar el problema de la velocidad de la deformación de la imagen, se presentan métodos de aceleración como lo son el sub-muestreo de la imagen, el promedio y recorte de los píxeles.

Se aplicaron los métodos de aceleración a una imagen de 240x240 píxeles y una imagen de referencia al 25% 50% y 100% de aumentos. Esto corresponde en tamaños de la imagen en 300x300, 360x360, 480x480 píxeles, los resultados se obtuvieron en una PC Pentium 200 MHz, sin ningún hardware de aceleración de gráficos o texturas. El uso de una imagen de referencia de gran tamaño junto con un sub-muestreo de la imagen de referencia, produce una imagen de mejor calidad.

2.6. A novel multi-view image coding scheme based on view-warping and 3D-DCT

En este artículo se presenta un nuevo enfoque, para explotar al máximo la redundancia entre los diferentes puntos de vista de la misma escena, considerando tanto la textura. El objetivo es lograr un mejor rendimiento en la compresión de la imagen [17].

El proceso aplicado se realiza en tres pasos: En el primero la deformación 3D se utiliza en lugar de la etapa de predicción de movimiento con el fin de mejorar la precisión y evitar las limitaciones dadas por los tamaños de bloque fijos en la imagen. En el segundo, el algoritmo deforma todos los puntos de vistas disponibles con respecto a una imagen de referencia y almacena todos los puntos de vistas en una pila. Por último se ocupan de las oclusiones que se presentan en las regiones que no son visibles en las imágenes de referencia.

Los resultados experimentales en este artículo, muestra como el algoritmo propuesto superan al método H.264 AVC. En la Tabla 2.1 se muestra un resumen con descripciones de los trabajos relacionados.

Artículo ó Tesis	Técnicas ó Descriptores empleados	Categoría de imágenes	Técnicas ó descriptor con mejor rendimiento
[Graig M, 1995]	EREW,CREW, PRAM	Personas	EREW-PRAM
[Xuehui Liu, 2001]	Geometría Epipolar, Relleno de huecos	Interiores de Edificios	No se menciona
[Enhua, 2003]	Líneas epipolares, discontinuidad de profundidad, deformación de la imagen, relleno de huecos	Interiores de Edificios	Líneas epipolares, discontinuidad de profundidad, deformación de la imagen, relleno de huecos
[Paul Bao, 2003]	Sub-sampling, pixel averaging, clipping	Edificios y objetos	Sub-sampling y clipping
[Natividad, 2007]	Detección de puntos de interés, correspondencia de puntos y reconstrucción de puntos, triangulación de Delaunay	Estatuas y figuras prehispánicas	No se menciona
[Zamarin, 2010]	3D-DCT, H.264 MVC	Videos de personas	H.264 MVC
[Song-Pei, 2013]	Deformaciones 2D de la imagen, transformaciones proyectivas, manipulación perspectiva.	Edificios y objetos	Deformaciones 2D de la imagen, transformaciones proyectivas, manipulación perspectiva.

Tabla 2.1: Resumen de los trabajos revisados.

2.7. Análisis del problema

A continuación se presentan los problemas relacionados con la presente tesis y se describen de forma general la solución propuesta para generar nuevas vistas a partir de una imagen de referencia. En el análisis se pueden definir problemas relacionados con las imágenes, los cuales se explican a continuación.

2.7.1. Problemas relacionados con imágenes de edificios de la Universidad del Papaloapan

Las imágenes digitales presentan diversas características y estas pueden considerarse problemas dependiendo del enfoque de la investigación. En este trabajo las imágenes que se utilizarán fueron tomadas de los edificios de la universidad y sus características, se citan a continuación:

- Existe variabilidad entre imágenes de una misma categoría, debido a que las imágenes se adquirieron en un ambiente no controlado. Esto genera que existan diferencias entre una imagen y otra, por ejemplo los cambios se pueden observar en la iluminación, contraste e intensidad, entre otros.
- Las imágenes cuentan con diferentes categorías de elementos, entre estos se encuentran el edificio, cielo, césped y árboles. Estos elementos se distorsionan de diferente forma cuando se aplican una transformación.
- La imagen de referencia es una imagen cuadrada antes de aplicarse la transformación, una vez aplicada la transformación, la imagen de resultado presenta una apariencia rectangular.

En la Tabla 2.2 Se presenta un resumen de las características de las imágenes que se emplean en esta tesis.

Categoría de imágenes	Características
Edificio de Agroingeniería, Auditorio, Biblioteca	<ul style="list-style-type: none"> ■ Imágenes adquiridas en un ambiente no controlado. ■ Variedad de elementos contenidos en una misma imagen.

Tabla 2.2: Categoría de imágenes de edificios usados para esta investigación sus características.

2.7.2. Propuesta de solución

Para solucionar el problema planteado en esta tesis, se presenta una metodología para realizar el proceso de generar un cambio de vista de una imagen. Como primer paso se define las clases o categoría de edificios de la Universidad que se utilizan. El segundo paso consiste en tomar la imagen de referencia. Por ejemplo la categoría 1: es el edificio de Agroingeniería, en esta etapa se redimensiona la imagen a una resolución de 640x480 pixeles, por último se obtiene cuatro puntos de referencia de la imagen y se calcula la matriz de proyección *matriz_{Proyectiva}* la cual sirve para generar una nueva imagen con un punto de vista diferente.

2.7.3. Discusión

En este capítulo se expusieron los trabajos relacionados al problema abordado en esta tesis, se realiza el planteamiento del problema y una descripción general de la propuesta de solución. Con base en la información revisada, se citan los siguientes planteamientos.

- En los trabajos relacionados se abordó el tema de la generación de nuevas imágenes empleando técnicas en el dominio espacial. Las técnicas empleadas en estos trabajos son: líneas epipolares, triangulación, deformaciones hacia adelante y hacia atrás, transformaciones proyectivas.
- De acuerdo al análisis del problema se determinó emplear el método de transformación proyectiva para generar imágenes con nuevas vistas, dichas transformaciones se implementan en el dominio espacial de la imagen.

Capítulo 3

Marco Teórico

En este capítulo se describen los fundamentos teóricos de la visión estéreo, necesaria para comprender el problema que se aborda en esta tesis. Actualmente la visión estéreo se basa en geometría proyectiva, por lo que se realiza una revisión de ésta, desde la representación de una imagen, coordenadas homogéneas hasta proyecciones.

3.1. Modelo RGB

El propósito de un modelo de color es facilitar la especificación de los colores de alguna forma estándar. En esencia, un modelo de color es una especificación de un sistema de coordenadas $3D$ y un subespacio dentro de dicho sistema donde cada color se representa por un punto [5].

En procesamiento de imágenes se utiliza el modelo de color RGB (del inglés Red, Green, Blue). En este modelo está basado en el sistema de coordenadas cartesianas. El subespacio de color de interés es el tetraedro mostrado por la Figura 3.1. En la cual los colores RGB están en tres vértices y los colores cyan, magenta y amarillo se sitúan en otros tres vértices [2], el negro corresponde al origen y el blanco se encuentra en el vértice más alejado del origen.

En este modelo, la escala de grises se extiende desde el negro al blanco a lo largo de la diagonal que une esos dos puntos, y los colores son puntos dentro del tetraedro, definidos por vectores desde el origen [5]. Los valores de R, G y B están normalizados en el rango $[0,1]$.

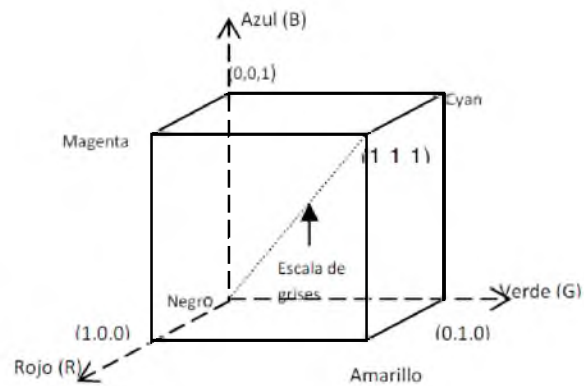


Figura 3.1: Tetraedro de color RGB. Figura tomada de [2].

Todos los colores que se visualizan en las imágenes, utilizan el modelo de color RGB. La cantidad de cada color se especifica con un valor entre 0 y 255. La figura 3.2, muestra una imagen con el modelo RGB.



Figura 3.2: Imagen RGB.

3.2. Representación de imágenes

La formación de la imagen ocurre cuando un sensor registra la radiación que ha interactuado con ciertos objetos físicos. La imagen obtenida por el sensor es una función bidimensional $f(x, y)$, el valor de la función f corresponde a la intensidad o brillantez en cada punto de la imagen (x, y) [5]. En la Figura 3.3 se muestra la representación de una imagen digital.

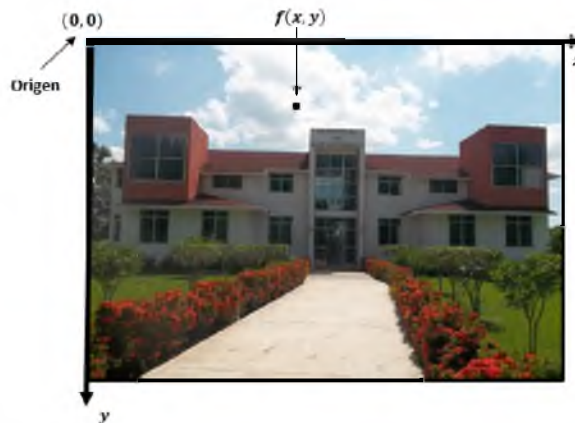


Figura 3.3: Convención de ejes utilizadas para la representación de imágenes digitales.

Una imagen digital en escala de grises ha sido discretizada tanto en valor de intensidad como espacialmente, una imagen digital se representa por una matriz de $N \times M$, cada valor es un número que representa el nivel de intensidad del punto correspondiente de la imagen [7]. Los elementos de la matriz (3.1) se conocen como píxeles. Los píxeles están representados en un plano proyectivo por las coordenadas (x, y) . A continuación se describe el plano proyectivo de una imagen.

$$imagen = \begin{bmatrix} 0 & 1 & 8 & 3 \\ 5 & 4 & 1 & 9 \\ 3 & 2 & 0 & 1 \\ 9 & 4 & 7 & 1 \end{bmatrix} \quad (3.1)$$

3.3. Plano proyectivo

Un punto en un plano se representa por un par de coordenadas (x, y) en R^2 . Comúnmente R^2 se identifica como un plano. Se puede considerar entonces a R^2 como un espacio vectorial, en el que el par de coordenadas (x, y) es un vector [15][1]. En esta sección se introducirá la representación de coordenadas homogéneas para líneas y puntos sobre un plano.

3.3.1. Puntos y líneas

3.3.1.1. Representación homogénea de líneas

Una línea en el plano se representa por una ecuación de la siguiente forma:

$$ax + by + c = 0 \quad (3.2)$$

donde los coeficientes (a, b, c) toman diferentes valores para representar así diferentes líneas [1]. De esta manera, una línea es representada por un vector columna $(a, b, c)^T$. La correspondencia entre las líneas y vectores $(a, b, c)^T$ no es uno a uno, ya que las líneas $ax + by + c = 0$ y $(ka)x + (kb)y + (kc) = 0$ son las mismas para toda constante $k \neq 0$. Dos vectores relacionados entre sí por un factor de escala son considerados vectores equivalentes o vectores homogéneos. Cualquier vector particular $(a, b, c)^T$ es uno de la clase equivalente. Una serie de vectores de clases equivalentes en R^3 , a excepción del vector $(0, 0, 0)^T$ forma el espacio proyectivo P^2 .

3.3.1.2. Representación homogénea de puntos

Un punto (x, y) pertenece a la $l = (a, b, c)^T$ si y solo si la ecuación (3.2) es verdadera. Esta ecuación se escribe utilizando el producto escalar de vectores; al expresar a $(x, y) \in R^2$ en coordenadas homogéneas [8].

$$0 = \mathbf{x}^T \cdot \mathbf{l} = [x, y, 1]^T \cdot [a, b, c] \quad (3.3)$$

La ecuación (3.3) es equivalente a la ecuación (3.2). De esta forma, al igual que con las líneas, los puntos (x, y) se representan por un vector $(x, y, 1)^T$. Se dice entonces que los puntos en un plano bidimensional se representan por vectores homogéneos de tres dimensiones, cuyos dos primeros elementos son las coordenadas del punto en el plano y el tercer elemento es 1.

En términos generales, si se tiene un vector homogéneo de tres dimensiones dados por $(x_1, x_2, x_3)^T$ que representa un punto en el plano, las coordenadas de este punto están definidas como:

$$(x, y) = (x_1/x_3, x_2/x_3) \in R^2 \quad (3.4)$$

3.3.2. Grados de libertad (dof)

Para representar un punto en el espacio R^2 se requiere proporcionar dos valores, las coordenadas x y y , así también, para representar una recta de la forma $(ax + by + c = 0)$ se define un par de parámetros independientes dados por las relaciones entre los coeficientes $\{a : b : c\}$ [15]. Entonces los grados de libertad son el número mínimo de parámetros que se requiere para especificar un punto, una recta o cualquier otra entidad geométrica sin considerar los valores de sus coeficientes.

3.4. Modelo de una cámara

La cámara de un sistema de visión por computadora, es el dispositivo que recibe la luz reflejada por la escena y la utiliza para generar imágenes. La imagen obtenida por el sensor es una función bidimensional, donde el valor de la función corresponde a la intensidad de cada punto de la imagen [9]. Es importante conocer el concepto básico de la cámara CCD y de la cámara Pinhole.

3.4.1. Cámara CCD

La cámara CCD (*charge-coupled-device*) es un sensor de imagen que utiliza elementos semiconductores fotosensibles en forma de arreglos matriciales [1]. Un sensor CCD transforma la energía luminosa incidente en energía eléctrica. Las cámaras CCD son consideradas como dispositivos que poseen una muy baja deformación geométrica de la imagen, una buena sensibilidad a la luz y una resolución aceptable [10].

3.4.2. Cámara Pinhole

El modelo matemático de la cámara Pinhole. Se considera un sistema de coordenadas Euclideo cuyo origen está en el centro de la cámara, llamado *centro de proyección* o *centro óptico* [15]. El eje z es considerado el eje principal, y a una distancia f sobre este eje se encuentra un plano paralelo al plano xy ; la distancia f se conoce como longitud focal, el plano como plano focal o plano de la imagen y el punto localizado donde el eje z que se encuentra con este plano es llamado punto principal de la imagen. Ver Figura(3.4).

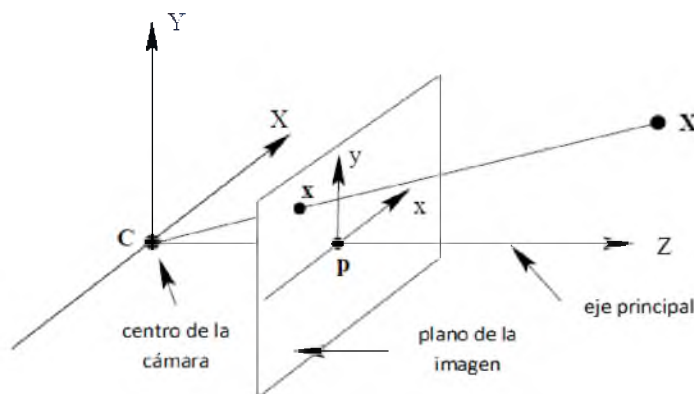


Figura 3.4: Modelo geométrico de cámara pinhole. Figura tomada de [15].

Considerando este modelo de la cámara Pinhole, un punto en el espacio con coordenadas $P = (X; Y, Z)^T$ es proyectado a un punto en el plano de la imagen donde una línea une el punto P con el centro de proyección a través del plano de la imagen. Con base a la semejanza de

triángulos, se puede observar que el punto $(X, Y, Z)^T$ es proyectado al punto $(fX/Z, fY/Z)^T$ sobre el plano de la imagen.

3.4.3. Modelo de proyección central

En la práctica el mapeo del espacio Euclideo R^3 al espacio Euclideo R^2 resulta en

$$(X, Y, Z)^T \mapsto (fX/Z + p_x, fY/Z + p_y)^T \quad (3.5)$$

donde (p_x, p_y) son las coordenada del punto principal con respecto del sistema de coordenadas del plano de la imagen, ya que el origen de éste regularmente no se encuentra en el punto principal [15]. La proyección central se puede expresar como un mapeo entre coordenadas homogéneas y a su vez en términos de la multiplicación de matrices como en la ecuación (3.6).

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} f & & 0 \\ & f & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.6)$$

La expresión anterior se puede reescribir de la siguiente forma:

$$x = K[I|0]Xc; \quad (3.7)$$

donde I es la matriz identidad de 3×3 , 0 es un vector columna de ceros de 3×1 y la matriz K , en la ecuación (3.8), es llamada *matriz de calibración de la cámara*, definida como:

$$K = \begin{bmatrix} f & p_x \\ & f & p_y \\ & & 1 \end{bmatrix} \quad (3.8)$$

3.5. Transformaciones Projectivas en 2D

En esta sección se estudian las transformaciones proyectivas en $2D$, así también, se abordan los modelos de transformación para las diferentes entidades geométricas. Dichas transformaciones son estudiadas y modeladas con una matriz de transformación H .

La geometría proyectiva en 2D es el estudio de las propiedades del plano proyectivo P^2 que son invariantes bajo un grupo de transformaciones conocidas como proyectividades [15].

Una proyectividad es un mapeo invertible h desde P^2 así mismo tres puntos cualesquiera x_1, x_2 y x_3 pertenecen a la misma línea si y solo si $h(x_1), h(x_2)$ y $h(x_3)$ también pertenecen a la misma línea [15]. Una proyectividad es llamada también una colineación, una transformación proyectiva o una homografía. Todos estos términos son sinónimos.

Una transformación $h: P^2 \rightarrow P^2$ es una proyectividad si y solo si existe una matriz de 3×3 no singular H tal que para cualquier punto en P^2 representado por un vector x se cumple que $h(x) = Hx$ [15]. Para interpretar este teorema, cualquier punto en P^2 es representado como x , un vector homogéneo de 3×3 , y Hx como una transformación o función lineal de coordenadas homogéneas.

Una transformación proyectiva es una transformación lineal de vectores homogéneos de 3 coordenadas representada por una matriz no singular de 3×3 elementos, como se ve en la ecuación (3.9) o de forma compacta como $x' = Hx$.

$$\begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (3.9)$$

Se puede escribir la transformación proyectiva de la ecuación (3.9) en la forma inhomogénea como:

$$x'_i = \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + 1} \quad (3.10)$$

$$y'_i = \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + 1} \quad (3.11)$$

Cada correspondencia de puntos genera 2 ecuaciones para los elementos de H , los cuales luego de multiplicarlos se obtiene:

$$x'_i(h_{31}x_i + h_{32}y_i + 1) = h_{11}x_i + h_{12}y_i + h_{13} \quad (3.12)$$

$$y'_i(h_{31}x_i + h_{32}y_i + 1) = h_{21}x_i + h_{22}y_i + h_{23} \quad (3.13)$$

Para determinar H se da un conjunto de correspondencias de cuatro puntos en 2D. Este vector puede ser expresado en términos de un producto cruz como $x'_i \times H_{x_i} = 0$. Donde x'_i es una matriz de $2N \times 8$ y H_{x_i} es un vector de 8×1 .

$$\begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x_i * x' & -y_i * x' \\ 0 & 0 & 0 & x_i & y_i & 1 & -x_i * y' & -y_i * y' \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} = 0 \quad (3.14)$$

La matriz H puede modificarse multiplicándola por un factor de escala distinto de cero, sin alterar la transformación proyectiva. Por lo que se dice que H es una matriz homogénea. La transformación proyectiva posee 8 grados de libertad.

3.5.1. Transformaciones de líneas

Si los puntos x_i están situados sobre la línea l , entonces la transformación de los puntos $x'_i = Hx_i$ bajo una transformación proyectiva, se encuentran sobre la línea $l' = H^{-T}l$ [15]. De esta manera los puntos que se encuentran sobre la línea se mantienen, dado que $l'^T x'_i = l^T H^{-1} Hx_i = 0$. Esto enuncia la regla de transformación para líneas: bajo la transformación de puntos $x' = Hx$, una línea es transformada como $l' = H^{-T}l$.

3.6. Categorización de las Transformaciones Proyectivas

En esta sección se describen características importantes de una transformación proyectiva y algunas de sus propiedades geométricas. Se introducirán estas transformaciones comenzando por las más especializadas, las isometrías, y progresivamente generalizándolas hasta que se alcancen las transformaciones proyectivas.

3.6.1. Transformación Isométrica (Euclidiana)

Las isometrías son transformaciones del plano R^2 que preservan la distancia euclidiana, es decir la distancia entre dos puntos es igual a la distancia entre los puntos transformados [15]. Una transformación euclideana también se conoce como un desplazamiento. Una isometría está representada como:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \epsilon \cos(\theta) & -\text{sen}(\theta) & t_x \\ \epsilon \text{sen}(\theta) & \cos(\theta) & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \quad (3.15)$$

donde $\epsilon = \pm 1$. Si $\epsilon = 1$ la isometría preserva la orientación y es una transformación Euclidiana (una composición de traslaciones y rotaciones). Si $\epsilon = -1$ entonces la isometría tiene una orientación invertida. La transformación euclidiana modela el movimiento de un objeto rígido. Una transformación euclidiana puede ser escrita en una forma compacta como:

$$H_E = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \quad (3.16)$$

donde R es una matriz de rotación de 2×2 , t un vector de traslación de dos dimensiones, y 0 un vector nulo de dos dimensiones.

La transformación euclideana posee tres grados de libertad, uno para la rotación y dos para la traslación [15]. Por lo tanto, se deben especificar tres parámetros para definir la transformación. Esta transformación se puede calcular desde la correspondencia de dos puntos (4 coordenadas). Las características invariantes de esta transformada, es decir, las propiedades que se mantienen después de aplicar la transformación isométrica son: la distancia entre dos puntos, el ángulo entre dos líneas y el área.

3.6.2. Transformación de similitud

Una transformación de similitud es una isometría compuesta con un factor de escala isotrópico (aplican un mismo cambio de escala en todas las dimensiones) [15]. Las transformaciones euclidianas compuestas con escalamientos son similitudes y tiene la siguiente representación matricial y puede ser escrita de forma compacta como la ecuación 3.18

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} s \cos(\theta) & -s \sin(\theta) & t_x \\ s \sin(\theta) & s \cos(\theta) & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \quad (3.17)$$

$$H_S = \begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix} \quad (3.18)$$

El escalar s representa el escalado isotrópico. Una transformación de semejanza también se le conoce como una transformación de igual-forma, debido a que preserva la forma. Una transformación de semejanza planar posee cuatro grados de libertad, los tres grados de libertad de la transformación euclídeana y el escalado. Al igual que una transformación euclídeana, la transformación de semejanza se puede especificar a partir de la correspondencia de dos puntos (4 coordenadas). Las características invariantes de esta transformada son: los ángulos entre líneas, las líneas paralelas, la relación entre dos distancias y la relación entre áreas.

3.6.3. Transformación Afín

Una transformación afín o una afinidad, es una transformación lineal no singular seguida por una traslación [15]. Esta transformación posee una representación matricial de la forma como:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \quad (3.19)$$

que puede ser escrita de forma compacta como:

$$H_A = \begin{bmatrix} A & t \\ 0^T & 1 \end{bmatrix} \quad (3.20)$$

donde A es una matriz no singular de 2×2 . Una transformación afín posee seis grados de libertad correspondientes a los seis elementos de la matriz. Esta transformación puede especificarse a partir de la correspondencia de tres puntos (6 coordenadas). Las características invariantes de esta transformación son: líneas paralelas y razón entre dos áreas.

Una forma de entender los efectos geométricos de la componente lineal A de una transformación afín es la composición de dos transformaciones fundamentales, normalmente rotaciones y escalado no isotrópico. La matriz afín A siempre se puede descomponer como:

$$A = R(\theta)R(-\phi)DR(\phi) \quad (3.21)$$

donde $R(\theta)$ y $R(\phi)$ son rotaciones por θ y ϕ respectivamente, y D es una matriz diagonal de la forma:

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad (3.22)$$

Esta descomposición se obtiene directamente del método SVD (descomposición en valor singular), escribiendo, $A = UDV^T = (UV^T)(UDV^T) = R(\theta)R(-\phi)DR(\phi)$ debido a que U y V son matrices ortogonales [15].

Por lo tanto, la matriz de afinidad A es el producto de una rotación (por ϕ), un escalado por λ_1 y λ_2 respectivamente en las direcciones rotadas x e y , una rotación hacia atrás (por $-\phi$) y finalmente otra rotación (por θ). La única diferencia entre la afinidad y una similaridad, es el escalamiento no isotrópico. La esencia de una afinidad es el escalado en direcciones ortogonales, orientadas en un ángulo particular [15], en la Figura 3.5, se ve un ejemplo.

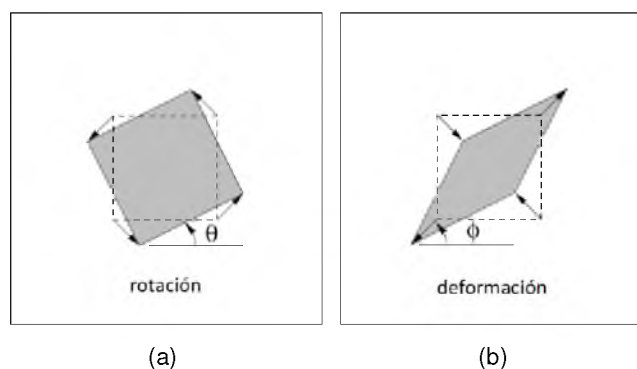


Figura 3.5: Distorsiones que surgen de una transformación planar afín. (a) Rotación por $R(\theta)$. (b) Una deformación $R(-\phi)DR(\phi)$. Figura tomada de [15].

3.6.4. Transformación Projectiva

Una transformación proyectiva es una transformación lineal no singular de coordenadas homogéneas. Esta transformación generaliza una transformación afín, la cual es la composición de una transformación lineal no singular general de coordenadas no homogéneas y una traslación [15]. La forma de una transformación proyectiva viene dada por la siguiente matriz:

$$H_P = \begin{bmatrix} A & t \\ V^T & u \end{bmatrix} \quad (3.23)$$

donde el vector $V = (v_1, v_2)^T$, es un factor de escalamiento para x y y respectivamente. La matriz posee 9 elementos donde solo sus relaciones son significativas, así la transformación se puede especificar por medio de 8 parámetros.

La característica invariante más importante de la transformación proyectiva es la relación de cruz de cuatro puntos colineales: una relación de longitudes sobre una línea es invariante bajo afinidad, pero no bajo una proyectividad. Sin embargo, una relación de relaciones o relación cruz de longitudes sobre una línea es un invariante proyectivo [15].

3.7. Descomposición de una transformación proyectiva

Una transformación proyectiva se descompone en una cadena de transformaciones, donde cada matriz en la cadena representa una transformación más alta en la jerarquía que la anterior, como se expresa a continuación:

$$H = H_S H_A H_P = \begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} I & 0 \\ V^T & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0^T & u \end{bmatrix} = \begin{bmatrix} A & t \\ V^T & u \end{bmatrix} \quad (3.24)$$

donde A es una matriz no singular dada por $A = sRK + tv^T$, y K una matriz triangular superior normalizada cuyo $\det(K) = 1$. Esta descomposición es válida siempre y cuando $u \neq 0$, y es única si s se elige positivo [15]. En la Tabla 3.1 Se presenta un resumen de las transformaciones proyectivas en 2D.

Grupo	Matriz H	Propiedades Invariantes
Proyectivo (8 dof)	$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$	razón de cruz.
Afín (6 dof)	$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$	líneas paralelas, razón entre dos áreas.
Similitud (4 dof)	$\begin{bmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$	ángulos entre rectas, razón entre dos distancias.
Euclídea (2 dof)	$\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$	longitud entre puntos.

Tabla 3.1: Resumen de Transformaciones Proyectivas 2D

3.8. Discusión

En este capítulo se presentaron los fundamentos teóricos y los conceptos que involucran la generación de nuevas imágenes con diferentes puntos de vista a partir de una imagen de referencia. De acuerdo con la información presentada en este capítulo se concluye lo siguiente:

- Las técnicas de visión por computadora permiten obtener información tridimensional a partir de imágenes bidimensionales adquiridas en una escena. Esta información es necesaria cuando se pretende generar nuevas imágenes a partir de una imagen de referencia.
- Las técnicas para generar nuevos puntos de vistas se basan en una geometría proyectiva que utiliza una matriz y transformaciones para obtener así una matriz de proyección H_P , con la cual se logra el cambio de vista en la imagen.

Capítulo 4

Análisis, Diseño e implementación del sistema

En este capítulo se describe el análisis y diseño del proceso necesario para generar nuevas imágenes con diferentes puntos de vista a partir de una imagen de referencia. Este proceso incluye la extracción de puntos de referencias y la aplicación de una transformación proyectiva 2D. También se realiza la implementación del sistema para la generación de nuevas imágenes.

4.1. Análisis del sistema

En el análisis del sistema intervienen los siguientes elementos: las imágenes de edificios, la descomposición en canales RGB de la imagen y una transformación proyectiva. Estos elementos se relacionan entre sí, como se muestra en la Figura 4.1.

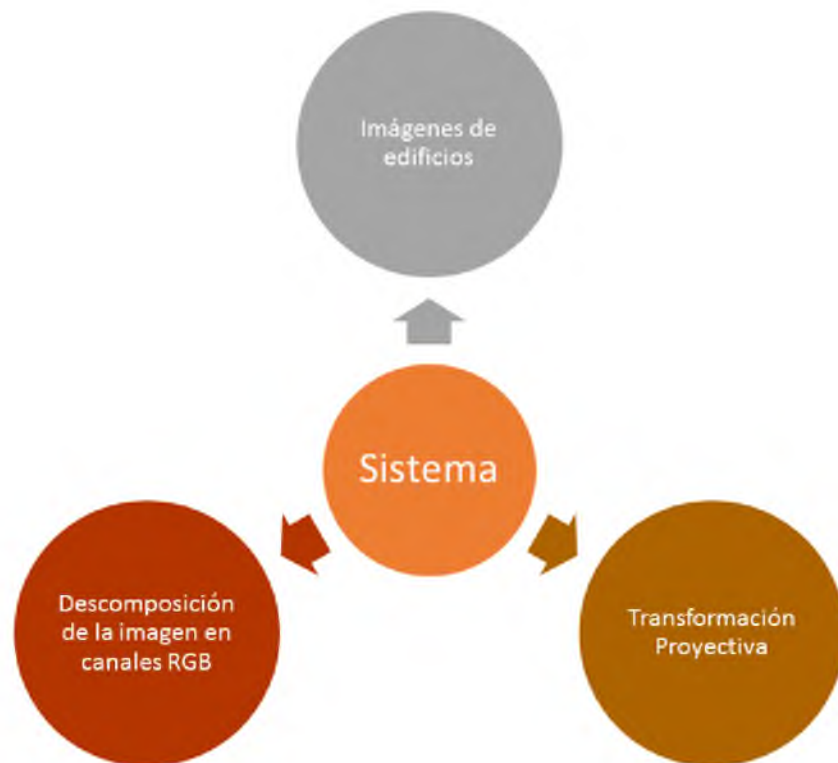


Figura 4.1: Elementos que intervienen en el análisis del sistema.

4.2. Esquema del sistema

El esquema que se presenta a continuación muestra la descripción general del sistema, en éste se define sus componentes, los procesos y las relaciones entre éstos. El esquema propuesto empieza a partir de la construcción de una base de datos de imágenes tomadas de los edificios de la Universidad, se selecciona una imagen de la base de datos, se seleccionan cuatro puntos en la imagen, de éstos puntos se obtiene sus coordenadas (x, y) , las cuales nos sirve para generar una matriz de 8×8 y obtener así la matriz proyectiva de 3×3 . En la sección 4.3 se muestra el proceso detalladamente, posteriormente se obtienen los canales en RGB, a cada canal de la imagen se le aplica la transformación proyectiva, la cual da como resultado un nuevo punto de vista de la imagen seleccionada originalmente. En la figura 4.2 se detalla el esquema general del sistema.



Figura 4.2: Esquema general del sistema.

4.3. Proceso para generar imágenes con diferentes puntos de vista

En esta sección se describe el proceso para generar nuevas imágenes con diferentes puntos de vista. Este proceso consta de 3 pasos como se muestra en la Figura 4.3. A continuación se detallan los pasos a seguir.

En el paso 1: se obtiene las imágenes de la Universidad del Papaloapan, éstas se clasifican en tres categorías: 1) Edificio de Agroingeniería, 2) Biblioteca, 3) Auditorio. Con las imágenes se crea una base de imágenes. La dimensión de las imágenes es de 640 por 480 píxeles. No existe restricción en el tamaño de las imágenes, sin embargo, con base en el estado del arte expuesto en el capítulo dos se optó por usar la dimensión mencionada.

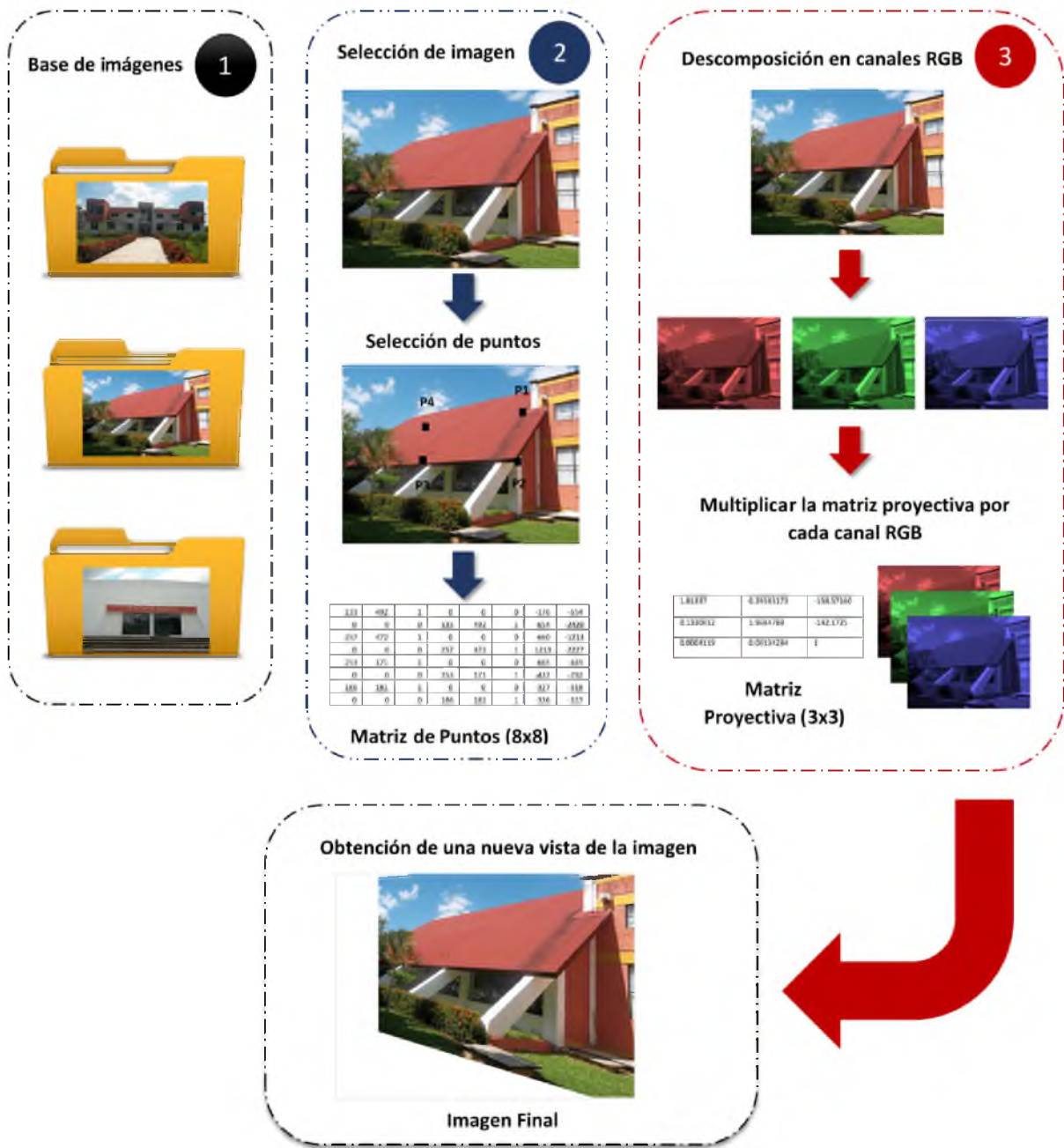


Figura 4.3: Proceso general para generar nuevas imágenes.

En el paso 2: se elige una imagen de referencia de la base de imágenes, posteriormente se seleccionan cuatro puntos en la imagen. Estos puntos van en el sentido de las manecillas del reloj para el cambio de vista hacia la izquierda y en el sentido contrario para el cambio de vista hacia la derecha, formando un rectángulo. En la Figura 4.4 se muestra los puntos seleccionados.



Figura 4.4: Imagen de la biblioteca.

Tomando como referencia la Figura 4.4, se obtiene las coordenadas (x, y) de cada uno de los puntos seleccionados, las coordenadas obtenidas para este ejemplo son las siguientes $Coordendas_X = [139, 249, 266, 172]$, $Coordendas_Y = [488, 470, 217, 216]$, para cada uno de los puntos se genera una matriz P_N de (2×8) con sus respectivas coordenadas (x, y) como se muestra a continuación:

$$P_N = \begin{bmatrix} punto_X & punto_Y & 1 & 0 & 0 & 0 & -coordenada_X * punto_X & -coordenada_X * punto_Y \\ 0 & 0 & 0 & punto_X & punto_Y & 1 & -coordenada_Y * punto_X & -coordenada_Y * punto_Y \end{bmatrix}$$

La matriz P_N se obtuvo de la ecuación 3.14 descrita en el capítulo 3, donde el $punto_X$ se obtiene de la primera posición del vector $Coordendas_X$ más un valor nuevo, para este ejemplo el valor es 0, al igual que el $punto_Y$, el $punto_Y$ se obtiene de la primera posición del vector $Coordendas_Y$, para los puntos restantes se hace el mismo procedimiento, la matriz para el punto 1 es la siguiente:

$$P_1 = \begin{bmatrix} 139 & 488 & 1 & 0 & 0 & 0 & -19321 & -67832 \\ 0 & 0 & 0 & 139 & 488 & 1 & -67832 & -238144 \end{bmatrix}$$

La matriz $puntos_P$ de $2N \times 8$ se obtuvo de la concatenación de la matriz P_N para cada uno de los cuatro puntos seleccionados en la imagen, como se muestra continuación:

$$puntos_P = \begin{bmatrix} 139 & 488 & 1 & 0 & 0 & 0 & -19321 & -67832 \\ 0 & 0 & 0 & 139 & 488 & 1 & -67832 & -238144 \\ 249 & 470 & 1 & 0 & 0 & 0 & -62001 & -117030 \\ 0 & 0 & 0 & 249 & 470 & 1 & -117030 & -220900 \\ 256 & 217 & 1 & 0 & 0 & 0 & -68096 & -57722 \\ 0 & 0 & 0 & 256 & 217 & 1 & -55552 & -47089 \\ 182 & 216 & 1 & 0 & 0 & 0 & -31304 & -37152 \\ 0 & 0 & 0 & 182 & 216 & 1 & -39312 & -46656 \end{bmatrix} \quad (4.1)$$

$$coordenadas = \begin{bmatrix} 139 \\ 488 \\ 249 \\ 470 \\ 266 \\ 217 \\ 172 \\ 216 \end{bmatrix} \quad (4.2)$$

La matriz m_H se obtuvo de la siguiente ecuación $m_H = puntos_P^{-1} * coordenadas$, donde $puntos_P^{-1}$ es la inversa de la matriz $puntos_P$ y donde $coordenadas$ es la matriz de los puntos seleccionados con sus respectivas coordenadas (x, y) , dando como resultado una matriz de 8×1 , como se muestra a continuación:

$$m_H = \begin{bmatrix} 1,76425424091740 \\ 0,307474467015981 \\ -156,400821912046 \\ 0,0410834415380199 \\ 2,02275382584927 \\ -154,163443446616 \\ 0,000212535590650272 \\ 0,00141189493649482 \end{bmatrix} \quad (4.3)$$

Para obtener la matriz $matriz_H$, se tomó cada uno de los elementos de la matriz m_H , empezando en la posición (1, 1) hasta la (1, 3) para formar la primera fila de la matriz $matriz_H$, la segunda fila se obtuvo de la posición (1, 4) hasta la posición (1, 6), la tercera fila se obtuvo de la posición (1, 7) hasta la posición (1, 8) que dando como último valor 1 por la ecuación 3.9 descrita en el capítulo 3, la matriz $matriz_H$ se muestra a continuación:

$$matriz_H = \begin{bmatrix} 1,76425424091740 & 0,307474467015982 & -156,400821912046 \\ 0,0410834415380190 & 2,02275382584927 & -154,163443446615 \\ 0,000212535590650270 & 0,00141189493649483 & 1 \end{bmatrix} \quad (4.4)$$

$$m_{Rotacion} = \begin{bmatrix} 0,996917333733128 & -0,0784590957278449 & 0 \\ 0,0784590957278449 & 0,996917333733128 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.5)$$

$$m_{Traslacion} = \begin{bmatrix} 1 & 0 & -240 \\ 0 & 1 & -960 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.6)$$

La matriz proyectiva $m_{Proyectiva}$ se obtiene al multiplicar la matriz $m_{matrizH}$, por la inversa de la matriz $m_{Traslacion}$, por la matriz $m_{Rotacion}$, por último se multiplica por la matriz $m_{Traslacion}$, dando como resultado la siguiente ecuación $m_{Proyectiva} = (m_{matrizH} * m_{Traslacion}^{-1} * m_{Rotacion} * m_{Traslacion})$, la matriz proyectiva $m_{Proyectiva}$ resultante de esta ecuación se muestra a continuación:

$$m_{Proyectiva} = \begin{bmatrix} 1,78293980252421 & 0,168104833472193 & -27,0905084956435 \\ 0,199660231054835 & 2,01329498119167 & -183,141382059362 \\ 0,000322656414334570 & 0,00139086718534932 & 0,993757643415451 \end{bmatrix} \quad (4.7)$$

En el paso 3: La imagen de referencia tomada, se descompone en sus tres canales RGB (Figura 4.5) y se obtiene una matriz por cada canal (rojo, verde, azul). Cada una de las matrices se multiplica por la matriz proyectiva $m_{Proyectiva}$, dando como resultado una imagen por cada canal (Figura 4.6). Posteriormente se regenera la imagen dando como resultado el nuevo cambio de vista de la imagen seleccionada (Figura 4.7).



Figura 4.5: Descomposición de la imagen en canales RGB. (a) Matriz $m_{CanalRojo}$. (b) Matriz $m_{CanalVerde}$. (c) Matriz $m_{CanalAzul}$.

$$img_{roja} = \begin{bmatrix} 1,7829398025 & 0,16810483347 & -27,090508495 \\ 0,19966023105 & 2,0132949811 & -183,14138205 \\ 0,00032265641433 & 0,0013908671853 & 0,99375764341 \end{bmatrix} \begin{bmatrix} 66 & 55 & 55 \\ 31 & 22 & 41 \\ 99 & 66 & 53 \end{bmatrix} \quad (4.8)$$

$$img_{verde} = \begin{bmatrix} 1,7829398025 & 0,16810483347 & -27,090508495 \\ 0,19966023105 & 2,0132949811 & -183,14138205 \\ 0,00032265641433 & 0,0013908671853 & 0,99375764341 \end{bmatrix} \begin{bmatrix} 192 & 187 & 188 \\ 203 & 194 & 186 \\ 217 & 224 & 228 \end{bmatrix} \quad (4.9)$$

$$img_{azul} = \begin{bmatrix} 1,7829398025 & 0,16810483347 & -27,090508495 \\ 0,19966023105 & 2,0132949811 & -183,14138205 \\ 0,00032265641433 & 0,0013908671853 & 0,99375764341 \end{bmatrix} \begin{bmatrix} 207 & 206 & 208 \\ 206 & 205 & 207 \\ 202 & 202 & 202 \end{bmatrix} \quad (4.10)$$

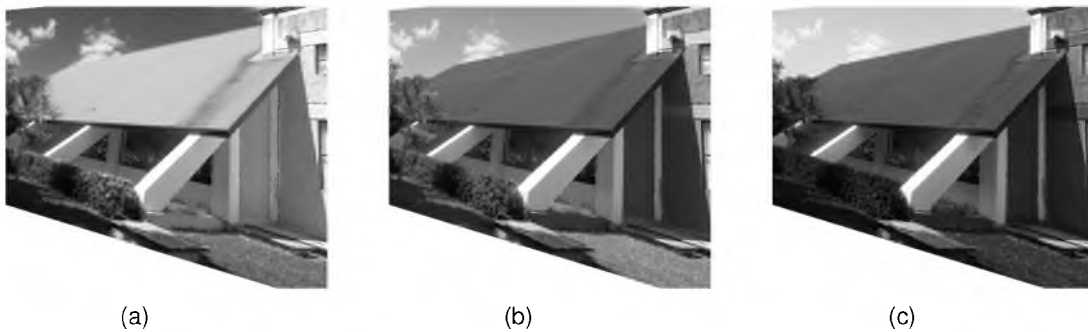


Figura 4.6: Transformación proyectiva por cada canal RGB. (a) Canal rojo en escala de grises. (b) Canal verde en escala de grises. (c) Canal azul en escala de grises.



Figura 4.7: Cambio de vista de la imagen de referencia.

4.4. Diseño e implementación del sistema

En este apartado se describe el ambiente de implementación, el diseño del sistema y la implementación del proceso para generar nuevas imágenes con diferentes puntos de vista, descritas en la sección 4.3.

4.4.1. Descripción del ambiente de implementación

El programa para generar nuevos puntos de vista se desarrolló en Matlab empleando herramientas de procesamiento de imágenes y desarrollo de interfaces gráficas GUIDE. La interacción del usuario con la herramienta descrita se realiza de forma directa. El usuario selecciona una imagen de las tres categorías de la Universidad, después se seleccionan los cuatro puntos dentro de la imagen y ejecuta el programa. El esquema de interacción del usuario con el sistema se muestra en la Figura 4.8.

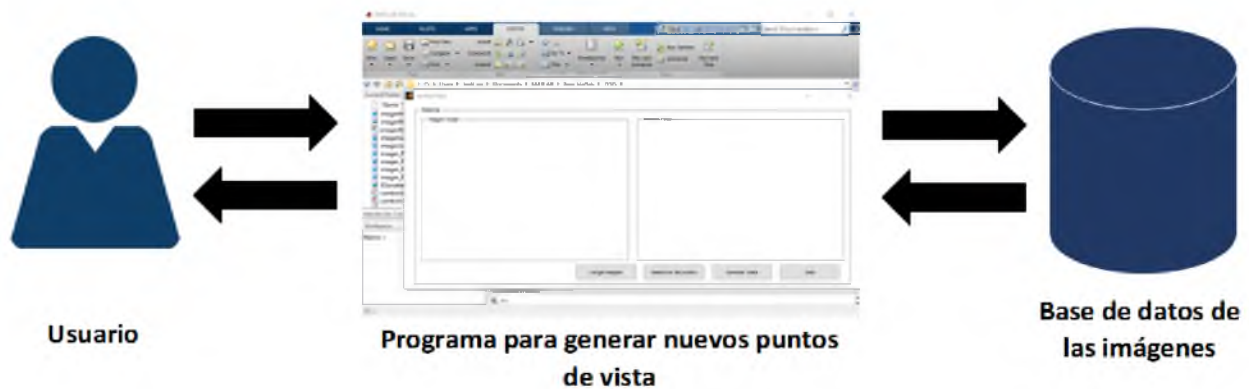


Figura 4.8: Esquema de interacción del usuario con el sistema.

4.4.2. Diseño general del sistema

En esta sección se define el diseño del proceso para generar nuevas imágenes con diferentes puntos de vista, en estos se describen las entradas, salidas y la secuencia de los procesos, los cuales se observan a detalle en la Figura 4.9.

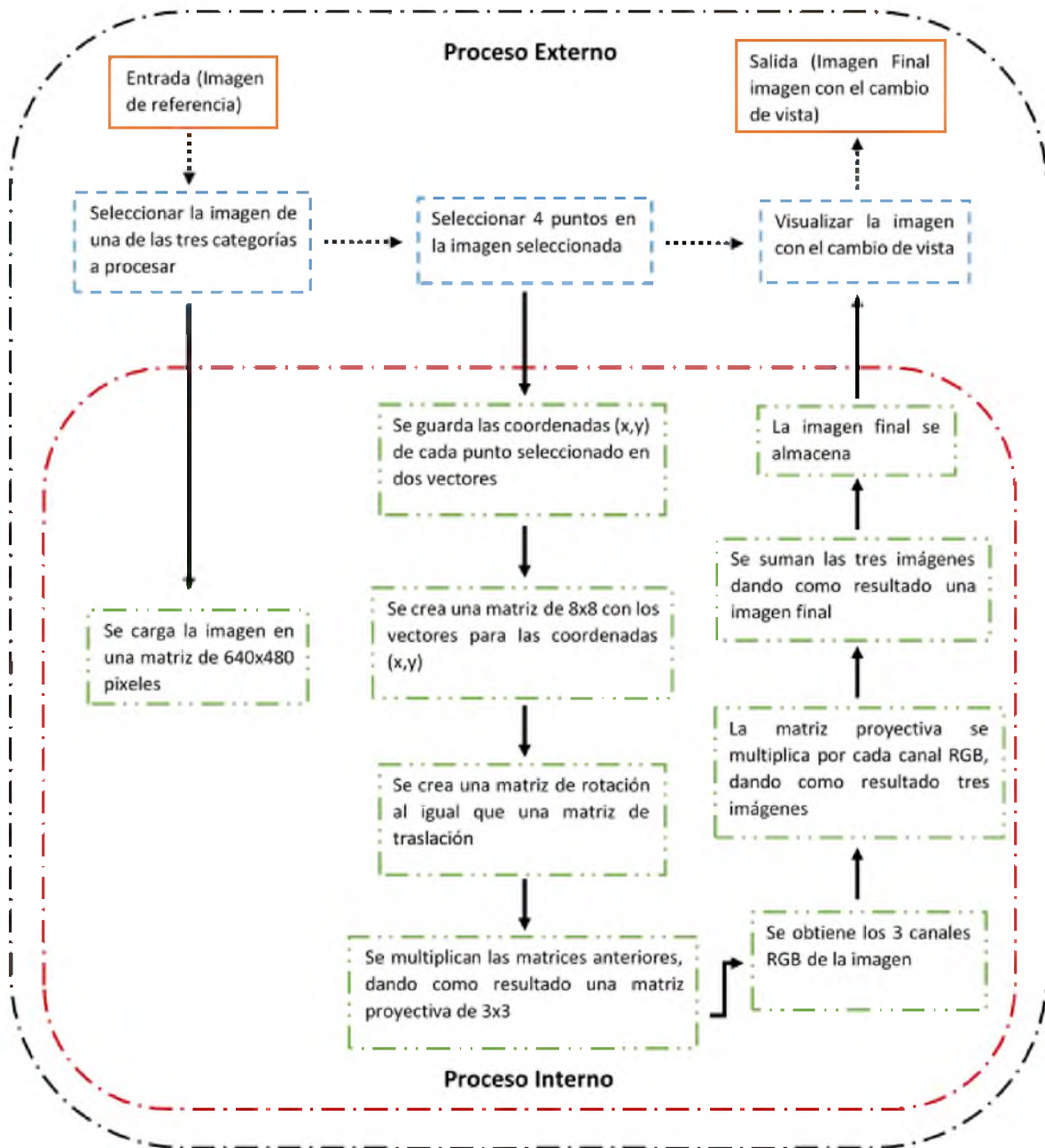


Figura 4.9: Diseño general del sistema: Proceso para generar imágenes con diferentes vistas.

4.5. Discusión

El análisis del sistema determinó los elementos que intervienen para la solución del problema y los procesos necesarios para la implementación. Las etapas que componen el proceso para generar nuevos puntos de vista de una imagen son: crear la base de imágenes obtenidas de la universidad, seleccionar cuatro puntos en la imagen de referencia, obteniendo así una matriz proyectiva, la cual se multiplica por cada canal RGB y finalmente se obtiene el cambio de vista de la imagen seleccionada.

El sistema se implementó en el entorno de programación *Matlab*. El sistema genera nuevos puntos de vista con dirección a la derecha o con dirección a la izquierda, de acuerdo a los cuatro puntos seleccionados en la imagen.

Capítulo 5

Experimentos y resultados

En este capítulo se presentan las pruebas realizadas y los resultados que se obtuvieron. El algoritmo para la generación de nuevas imágenes con diferentes puntos de vistas, se implementó en *Matlab*. Estas pruebas tienen por objetivo generar un cambio de vista por cada imagen seleccionada de la base de imágenes.

5.1. Ambiente de pruebas

En esta sección se muestran las características del equipo de cómputo que se utilizó para realizar cada una de las pruebas, al igual que el software y las imágenes que se utilizaron en dichas pruebas.

5.1.1. Hardware

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

- Procesador: Intel(R) Core(TM) i7-4510U CPU 260GHz.
- Memoria RAM: 8 GB.
- Sistema Operativo de 64 bits, Windows 10.

5.1.2. Software

El programa se elaboró en *Matlab R2014a* utilizando las herramientas de procesamiento de imágenes y el desarrollo de interfaces gráficas GUIDE, que tiene como finalidad la selección del modo de operación del sistema y su visualización.

5.1.3. Base de imágenes

La base de imágenes se integró de imágenes de los edificios de la Universidad Papaloapan, formando tres categorías: Edificio de Agroingeniería, Biblioteca y Auditorio, (obtenidas mediante captura personal), el número de imágenes por cada categoría es de 10, estas imágenes tienen una dimensión de 640 x 640 píxeles.

5.2. Pruebas con cambio hacia la derecha

Objetivo: Observar el desempeño del algoritmo planteado para obtener un cambio de vista hacia la derecha con una imagen de referencia de las categorías Agroingeniería, Biblioteca y Auditorio. En esta prueba se espera que las imágenes presenten un cambio significativo de vista.

Procedimiento: En esta prueba se utilizó una imagen de referencia, en esta imagen se seleccionan cuatro puntos. Estos puntos van en el sentido contrario a las manecillas del reloj y forman un rectángulo. A partir de estos puntos se genera la matriz proyectiva, la cual se usa para generar el cambio de vista en la imagen.

5.2.1. Prueba para la categoría Agroingeniería con cambio hacia la derecha

Resultado: Al realizar la prueba se observa un cambio de vista, como se muestra en la Figura 5.1, donde el edificio se aproxima a la vista del usuario. El acercamiento es más notable en el extremo izquierdo de la imagen. En la Tabla 5.1 se muestran los puntos de referencia que se utilizaron.

Imagen con orientación hacia la derecha		
Punto	Coordenada X	Coordenada Y
punto 1	231	143
punto 2	231	385
punto 3	388	381
punto 4	376	162

Tabla 5.1: Resultados de los puntos que se extraen en la imagen de referencia de la categoría Agroingeniería.



Figura 5.1: Resultado al aplicar un cambio de vista hacia la derecha de la categoría Agroingeniería.

5.2.2. Prueba para la categoría Biblioteca con cambio hacia la derecha

Resultado: Al realizar la prueba se observa un cambio significativo, como se muestra en la Figura 5.2. El edificio de la biblioteca se aproxima a la vista del usuario. El acercamiento es más notable en el extremo izquierdo de la imagen y del lado derecho los objetos en la imagen se alejan. En la Tabla 5.2 se muestran los puntos de referencia que se utilizaron.

Imagen con orientación hacia la derecha		
Punto	Coordenada X	Coordenada Y
punto 1	507	119
punto 2	504	264
punto 3	240	268
punto 4	265	142

Tabla 5.2: Resultados de los puntos que se extraen en la imagen de referencia de la categoría Biblioteca.



Figura 5.2: Resultado al aplicar un cambio de vista hacia la derecha de la categoría Biblioteca.

5.2.3. Prueba para la categoría Auditorio con cambio hacia la derecha

Resultado: Al realizar la prueba se observa un cambio significativo, como se muestra en la Figura 5.3. El auditorio se aproxima a la vista del usuario. El acercamiento es más notable en el extremo izquierdo de la imagen y del lado derecho los objetos dentro de la imagen se alejan. En la Tabla 5.3 se muestran los puntos de referencia que se utilizaron.

Imagen con orientación hacia la derecha		
Punto	Coordenada X	Coordenada Y
punto 1	214	115
punto 2	232	326
punto 3	454	328
punto 4	447	136

Tabla 5.3: Resultados de los puntos que se extraen en la imagen de referencia de la categoría Auditorio.



Figura 5.3: Resultado al aplicar un cambio de vista hacia la derecha de la categoría Auditorio.

5.3. Pruebas con cambio hacia la izquierda

Objetivo: Observar el desempeño del algoritmo planteado para obtener un cambio de vista hacia la izquierda con una imagen de referencia de la categoría Agroingeniería, Biblioteca y Auditorio. En esta prueba se espera que la imagen presente un cambio significativo de vista.

Procedimiento: En esta prueba se utilizó una imagen de referencia, en la que se seleccionan cuatro puntos. Estos puntos van en el sentido a las manecillas del reloj y forman un rectángulo. A partir de estos puntos se genera la matriz proyectiva, la cual genera el cambio de vista en la imagen.

5.3.1. Prueba para la categoría Agroingeniería con cambio hacia la izquierda

Resultado: Al realizar la prueba se observa un cambio significativo, como se muestra en la Figura 5.4. El edificio se aproxima a la vista del usuario. El acercamiento es más notable en el extremo derecho de la imagen. En la Tabla 5.4 se muestran los puntos de referencia que se utilizaron.

Imagen con orientación hacia la izquierda		
Punto	Coordenada X	Coordenada Y
punto 1	422	168
punto 2	421	380
punto 3	218	388
punto 4	230	161

Tabla 5.4: Resultados de los puntos que se extraen en la imagen de referencia de la categoría Agroingeniería



Figura 5.4: Resultado al aplicar un cambio de vista hacia la izquierda de la categoría Agroingeniería.

5.3.2. Prueba para la categoría Biblioteca con cambio hacia la izquierda

Resultado: Al realizar la prueba se observa un cambio significativo, como se muestra en la Figura 5.5, donde la biblioteca se aproxima a la vista del usuario. El acercamiento es más notable en el extremo derecho de la imagen y del lado izquierdo los objetos en la imagen se alejan. En la Tabla 5.5 se muestran los puntos de referencia que se utilizaron para generar el cambio de vista.

Imagen con orientación hacia la izquierda		
Punto	Coordenada X	Coordenada Y
punto 1	175	143
punto 2	163	271
punto 3	434	265
punto 4	445	123

Tabla 5.5: Resultados de los puntos que se extraen en la imagen de referencia de la categoría Biblioteca.



Figura 5.5: Resultado al aplicar un cambio de vista hacia la izquierda de la categoría Biblioteca.

5.3.3. Prueba para la categoría Auditorio con cambio hacia la izquierda

Resultado: Al realizar la prueba se observa un cambio significativo, como se muestra en la Figura 5.6, donde el auditorio se aproxima a la vista del usuario. El acercamiento es más notable en el extremo derecho de la imagen. En la Tabla 5.6 se muestran los puntos de referencia que se utilizaron.

Imagen con orientación hacia la izquierda		
Punto	Coordenada X	Coordenada Y
punto 1	186	219
punto 2	195	323
punto 3	51	325
punto 4	46	180

Tabla 5.6: Resultados de los puntos que se extraen en la imagen de referencia de la categoría Auditorio.



Figura 5.6: Resultado al aplicar un cambio de vista hacia la izquierda de la categoría Auditorio.

5.4. Pruebas para las tres categorías con los mismos puntos de referencia

5.4.1. Prueba 1: Con cambio hacia la derecha con un área de selección de 27,676.18 píxeles

Objetivo: Observar el desempeño del algoritmo que se implementó para obtener un cambio de vista hacia la derecha con una imagen de referencia de las tres categorías (Agroingeniería, Biblioteca, Auditorio). En esta prueba se espera que las imágenes presenten un cambio de vista.

Procedimiento: En esta prueba se utilizaron imágenes de referencia de las tres categorías, a cada

una de las imágenes se le seleccionó la misma área (27,676.18 píxeles) para los cuatro puntos (Tabla 5.7). A partir de estos puntos se generó una matriz proyectiva y se genera el cambio de vista en las imágenes.

Resultado: Al realizar la prueba se observa un cambio significativo en las imágenes, como se muestra en la Figura 5.7. En la categoría Agroingeniería, Biblioteca, se observa el mismo resultado donde los elementos del lado izquierdo se acercan más a la vista del usuario y los elementos del lado derecho se alejan, dando como resultado el cambio de vista en la imagen. En la categoría Auditorio el cambio es un acercamiento en el lado izquierdo en la imagen.



Figura 5.7: Resultado al aplicar un cambio de vista hacia la derecha en las tres categorías.

Prueba 1: imagen con orientación hacia la derecha		
Punto	Coordenada X	Coordenada Y
punto 1	166	172
punto 2	159	283
punto 3	405	290
punto 4	401	171

Tabla 5.7: Resultados de los puntos que se extraen en la imagen de referencia de las tres categorías.

5.4.2. Prueba 2: Con cambio hacia la izquierda con un área de selección de 29,220.27 pixeles

Objetivo: Observar el desempeño del algoritmo planteado para obtener un cambio de vista hacia la izquierda con una imagen de referencia de las tres categorías (Agroingeniería, Biblioteca, Auditorio). En esta prueba se espera que las imágenes presente un cambio significativo de vista.

Procedimiento: En esta prueba se utilizaron imágenes de referencia de las tres categorías, a cada una de las imágenes se le seleccionó la misma área (29,220.27 pixeles) para los cuatro puntos (Tabla 5.8). A partir de estos puntos se generó una matriz proyectiva, para generar el cambio de vista en las imágenes.

Resultado: Al realizar la prueba se observa un cambio significativo en las imágenes, como se muestra en la Figura 5.8. En la categoría Agroingeniería y Biblioteca los elementos del lado derecho se aproximan a la vista del usuario y los elementos del lado izquierdo se alejan. En la categoría Auditorio se observa un acercamiento al edificio.



Figura 5.8: Resultado al aplicar un cambio de vista hacia la izquierda en las tres categorías.

Prueba 2: imagen con orientación hacia la izquierda		
Punto	Coordenada X	Coordenada Y
punto 1	382	171
punto 2	391	330
punto 3	199	332
punto 4	207	173

Tabla 5.8: Resultados de los puntos que se extraen en la imagen de referencia de las tres categorías.

5.4.3. Prueba 3: Con cambio hacia la derecha con un área de selección de 59,098.27 pixeles

Objetivo: Observar el desempeño del algoritmo que se implementó para obtener un cambio de vista hacia la derecha con una imagen de referencia de las tres categorías (Agroingeniería, Biblioteca, Auditorio). En esta prueba se espera que las imágenes presente un cambio de vista.

Procedimiento: En esta prueba se utilizaron imágenes de referencia de las tres categorías, a cada una de las imágenes se le seleccionó la misma área (59,098.27 píxeles) para los cuatro puntos (Tabla 5.9). A partir de estos puntos se generó una matriz proyectiva, la cual se usa para generar el cambio de vista en las imágenes.

Resultado: Al realizar la prueba se observa un cambio significativo en las imágenes, como se muestra en la Figura 5.9. En la categoría Agroingeniería, la información del lado izquierdo se elimina para hacer el acercamiento de frente al edificio. En la categoría Biblioteca el edificio tiene un pequeño cambio donde el extremo izquierdo se acerca más a la vista del usuario. En la categoría Auditorio los elementos del lado derecho se alejan y en el lado izquierdo el acercamiento es más notorio.



Figura 5.9: Resultado al aplicar un cambio de vista hacia la derecha en las tres categorías.

Prueba 3: imagen con orientación hacia la derecha		
Punto	Coordenada X	Coordenada Y
punto 1	271	172
punto 2	267	348
punto 3	581	355
punto 4	585	155

Tabla 5.9: Resultados de los puntos que se extraen en la imagen de referencia de las tres categorías.

5.4.4. Prueba 4: Con cambio hacia la izquierda con un área de selección de 58,403.30 pixeles

Objetivo: Observar el desempeño del algoritmo que se implementó para obtener un cambio de vista hacia la izquierda con una imagen de referencia de las tres categorías (Agroingeniería, Biblioteca, Auditorio). En esta prueba se espera que las imágenes presente un cambio significativo de vista.

Procedimiento: En esta prueba se utilizaron imágenes de referencia de las tres categorías, a cada una de las imágenes se le seleccionó la misma área (58,403.30 pixeles) para los cuatro puntos (Tabla 5.10). A partir de estos puntos se generó una matriz proyectiva, para generar el cambio de vista en las imágenes.

Resultado: Al realizar la prueba se observa un cambio significativo en las imágenes, como se muestra en la Figura 5.10. En la categoría Agroingeniería y Biblioteca, los elementos del lado izquierdo se alejan y los elementos del lado derecho se acercan más a la vista del usuario dando como resultado el cambio en la imagen. En la categoría Auditorio el cambio es un acercamiento en el lado izquierdo.



Figura 5.10: Resultado al aplicar un cambio de vista hacia la izquierda en las tres categorías.

Prueba 4: imagen con orientación hacia la izquierda		
Punto	Coordenada X	Coordenada Y
punto 1	373	61
punto 2	377	297
punto 3	117	289
punto 4	127	64

Tabla 5.10: Resultados de los puntos que se extraen en la imagen de referencia de las tres categorías.

5.5. Discusión

En este capítulo se realiza la descripción de los casos de prueba y la forma en la que se comparan los resultados. Las pruebas integraron imágenes de las tres categorías, también se hicieron pruebas con una ventana pequeña y una ventana grande para la selección de puntos. De acuerdo a las pruebas realizadas con cada categoría se determina que el algoritmo aplicado reporta un buen desempeño para generar cambios de vistas a partir de imágenes de referencias.

En las pruebas para la categoría Agroingeniería con cambios hacia la derecha, las imágenes presentaron un cambio significativo, en el lado izquierdo se obtuvo un acercamiento de los objetos de la imagen a la vista del usuario.

En cuanto a las pruebas aplicadas a la categoría Biblioteca con cambios hacia la izquierda, se observa que las imágenes de resultado presentaron un cambio notable. Los objetos del lado izquierdo de la imagen se alejan más de la vista del usuario.

En las pruebas para la categoría Auditorio con cambios hacia la izquierda, se obtuvo un mejor resultado que al hacer el cambio hacia la derecha. En la imagen de resultado se observa un acercamiento hacia el lado derecho con una vista de frente al edificio.

En las pruebas en las que se integraron imágenes de las tres categorías con un área de selección de 29,220.27 píxeles para el cambio de vista hacia la izquierda, las imágenes presentaron un cambio significativo. Los objetos dentro de la imagen del lado derecho tuvieron un acercamiento notorio para las tres categorías. En cuanto al área de selección de 27,676.18 píxeles para el cambio de vista hacia la derecha, se observa que los elementos dentro de las imágenes de resultados para el lado izquierdo se alejan a la vista del usuario.

En cuanto a las pruebas aplicadas a las tres categorías para el cambio de vista hacia la derecha con un área de selección de 59,098.27 píxeles, se observa que las imágenes obtenidas presentaron un pequeño cambio en comparación con las pruebas anteriores. En cuanto al cambio de vista hacia la izquierda con el área de selección de 58,403.30 píxeles se observó que el resultado final fue menor en comparación de los resultados anteriores.

Con base en los resultados obtenidos, se determina que el desempeño del algoritmo, cumplió con el objetivo que se planteó en esta tesis; el cual fue generar un cambio de vista para cada una de las imágenes de las tres categorías: Agroingeniería, Biblioteca y Auditorio.

Capítulo 6

Conclusiones

A continuación se presentan las conclusiones finales, así como las aportaciones y los trabajos futuros de esta investigación. Con base en la información recopilada en el estado del arte, se determinan los siguientes.

Las técnicas más comunes que se utilizan para generar imágenes con cambio de vista son: líneas epipolares, deformación de la imagen, detección de puntos de interés, correspondencia de puntos, reconstrucción de puntos, deformaciones 2D y transformaciones proyectivas.

Las imágenes que se utilizan en los trabajos que se revisaron corresponden a las siguientes categorías: imágenes de personas, interiores de edificios, estatuas, objetos, videos de personas y edificios.

Las tareas que se tratan de resolver al aplicar una técnica para generar imágenes con nuevos puntos de vistas son las siguientes: construcción de entornos virtuales, generar vistas de objetos reales y sintéticos, reconstrucción de objetos 3D y la comprensión de la imagen.

Con base en los resultados obtenidos en las pruebas realizadas se determinan las siguientes conclusiones.

En los experimentos con imágenes de la categoría Agroingeniería con cambios hacia la izquierda, presentó un mejor desempeño en la imagen resultado, que al hacer el cambio de vista hacia la derecha, la imagen muestra un cambio notable en el acercamiento del edificio hacia la vista del usuario.

En los experimentos con las imágenes de la categoría Biblioteca para generar nuevos cambios de vistas, el cambio que presentó un mejor desempeño en la imagen, fue al hacer el cambio de vista hacia la derecha, el resultado muestra un cambio significativo, donde el acercamiento fue más notorio del lado izquierdo y del lado derecho los objetos dentro de la imagen se alejan a la vista del usuario.

En los experimentos para la categoría Auditorio, el cambio que presentó un mejor desempeño en la imagen, fue el cambio de vista hacia la izquierda, la imagen muestra un cambio notable, donde el acercamiento fue hacia el lado derecho con una vista de frente al usuario.

En los experimentos en los que se integraron imágenes de las tres categorías con cambios de vista hacia la derecha. Las imágenes presentaron un mejor desempeño con un área de selección pequeña de 27,676.18 píxeles. El cambio de vista hacia la izquierda presentó el mismo desempeño. Las imágenes muestran que con un área pequeña para la selección de puntos, el cambio de vista es más notable que con un área mayor. Esto se debe a que el área que se ve afectada por las transformaciones al hacer el cambio de vista en la imagen, se encuentra al exterior del área seleccionada, por esta razón, si se selecciona un área mayor el espacio de afectación será menor.

6.1. Aportaciones

Las principales aportaciones que se derivan al realizar este trabajo de investigación.

- La principal aportación de esta investigación fue aplicar un algoritmo para realizar el cambio de vista de una imagen de referencia.
- Se generó una base de imágenes de las categorías: Agroingeniería, Biblioteca y Auditorio. Esta base de imágenes se formó a partir de las imágenes capturadas en edificios de la Universidad del Papaloapan campus Loma Bonita.
- Se presenta un estudio sobre diversas técnicas revisadas en el estado del arte para generar nuevas vistas de una imagen de referencia.

6.2. Trabajos futuros

A continuación se presentan los trabajos futuros que se derivan al realizar este trabajo de investigación.

- La optimización de los códigos hechos en *Matlab*, para minimizar el tiempo de procesamiento de las imágenes.
- Desarrollar una aplicación para generar nuevas vistas a partir de n número de imágenes de referencia en tiempo real.
- Desarrollar un algoritmo que permita rellenar las áreas de los laterales, que resultan al generar una nueva imagen con el cambio de vista.

Bibliografía

- [1] Ángel Ernesto Gómez Sánchez. Visión Estereoscópica y Estimación de Pose para el Posicionamiento de un Brazo Robótico. *Centro Nacional de Investigación y Desarrollo Tecnológico*, 2008.
- [2] Diana Avellaneda Avellaneda. Categorización de Texturas Naturales. *Centro Nacional de Investigación y Desarrollo Tecnológico*, 2009.
- [3] Craig M. Wittenbrink y Arun K. Somanti. 2D and 3D Optimal Parallel Image Warping. *Elsevier*, 1995.
- [4] Enhua Wu y Xin Zheng. Composition of Novel Views Through an Efficient Image Warping. *The Visual Computer Springer*, 2003.
- [5] Gonzalo Pajares y Jesús M. de la Cruz García. Visión por Computador: Imágenes Digitales y Aplicaciones. *RA - MA*, 2008.
- [6] Hanqiu Sun y Enhua Wu. A Hybrid Method of Image Synthesis in IBR for Novel Viewpoints. *The Visual Computer Springer*, 2001.
- [7] Hearn D y Baker M. P. Gráficos por Computadora con OpenGL. *Pearson Educación S.A*, 2006.
- [8] Israel Vite Silva. Reconstrucción Tridimensional de Objetos Mediante Técnicas Evolutivas. *Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional.*, 2008.
- [9] Julio Molleda Meré. Técnicas de Visión por Computador para la Reconstrucción en Tiempo Real de la Forma 3D de Productos Laminado. *Universidad de Oviedo*, 2009.
- [10] Kenneth R. Castleman. Digital Image Processing. *Prentice-Hall*, 1997.

- [11] Lawrence Zitnick y Sing Bing Kang. Stereo for Image-Based Rendering Using Image Over-Segmentation. *International Journal of Computer Vision*, 2006.
- [12] Natividad Grandón Pastén y Diego Aracena Pizarro. 3D Object Reconstruction with Calibrated Images. *Ingeniare*, 2007.
- [13] Robert W. Mercado. Optimizing an Inverse Warper. *Department of Electrical Engineering and Computer Science, Cambridge, MA*, 1998.
- [14] Song-Pei Du y Ralph R. Martin. Changing Perspective in Stereoscopic Images. *Transactions on Visualization and Computer Graphics, IEEE*, 2013.
- [15] R. I. Hartley y A. Zisserman. Multiple View Geometry in Computer Vision. *Cambridge University Press*, 2004.
- [16] Paul Bao y Douglas Gourlay. Superview 3D Image Warping For Visibility Gap Errors. *Transactions on Consumer Electronics, IEEE*, 2003.
- [17] M. Zamarin y S. Milani. A Novel Multi-View Image Coding Scheme Based on View-Warping and 3D-DCT. *ElSevier*, 2009.