

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS E INGENIERÍA



**Estudio del diseño de algoritmo de correspondencia a través de
transformada Census mediante uso de FPGA**
**TRABAJO DE INVESTIGACIÓN PARA LA OBTENCIÓN DEL
GRADO DE BACHILLER EN CIENCIAS CON MENCIÓN EN
INGENIERÍA ELECTRÓNICA**

AUTOR

Dylan Hugo Mayor Vega

ASESOR:

Donato Andrés Flores Espinoza

Lima, diciembre, 2020

Resumen

La visión estereoscópica es la capacidad de adquirir dos imágenes en simultáneo cuando sus cámaras están separadas a una corta distancia; para que mediante algoritmos de procesamientos de imágenes se pueda determinar la correspondencia entre el pixel de la imagen original y su par estéreo.

Existen diferentes algoritmos de correspondencia, siendo uno de ellos la transformada Census, la particularidad de este algoritmo es el uso de la operación lógica XOR, por lo que su desarrollo es a través de dicha compuerta digital; asimismo, la transformada Census permitirá obtener la distancia de Hamming y próximamente desarrollar la función costo de correspondencia para determinar los pixeles correspondientes dentro del par estéreo.

Dicho algoritmo puede ser desarrollado en diferentes dispositivos electrónicos como la tarjeta de desarrollo FPGA, estas tarjetas cuentan con elementos adicionales al chip FPGA como una memoria externa SDRAM de 64 MB que puede ser usado para el almacenamiento de información y un procesador NIOS II para enviar la información de las imágenes estereoscópicas al módulo digital de transformada Census.

En este trabajo se busca realizar el modelo de solución en base a la teoría del algoritmo de correspondencia en imágenes estereoscópicas a través de transformada Census y realizar un modelo de solución para su desarrollo en un FPGA.

Palabras claves: Transformada Census, FPGA, función costo de correspondencia, distancia de Hamming, procesador NIOS II.

Índice

Introducción	1
Capítulo 1: Diseño del hardware para un vehículo aéreo no tripulado	3
1.1. Motivación de problemática	3
1.2. Estado del arte	5
1.2.1. Sistema basado en FPGA para visión estéreo usando transformada Census para robots autónomos móviles	5
1.2.2. Sistema de alta calidad de visión estéreo en tiempo real para FPGA	5
1.2.3. Marco novedoso para estimación de disparidad en FPGA	6
1.2.4. Diseño de un sistema embebido en tiempo real con FPGA para visión estéreo	6
1.2.5. Sistema de doble canal de adquisición de imágenes basado en FPGA	7
1.3. Justificación	7
1.4. Objetivos	8
1.4.1. Objetivo general	8
1.4.2. Objetivos específicos	8
Capítulo 2: Teoría	9
2.1. Visión general	9
2.2. Marco teórico	10
2.2.1. Vehículos aéreos no tripulados	10
2.2.2. Visión estereoscópica	11
a) Geometría epipolar	11

b)	Rectificación de imagen	11
c)	Disparidad	12
2.2.3.	Tabla de FPGA	12
2.2.4.	FPGA Cyclone V (5CSEMA5F31C6N)	13
2.2.5.	Diseño de la plataforma de integración de sistema	13
2.2.6.	Memoria SDRAM	13
2.2.7.	Función de correspondencia a partir de la transformada Census	14
2.2.8.	Transformada rápida de Census	16
2.3.	Modelo	17
2.3.1.	Diagrama de bloques	17
2.3.2.	Modelo teórico	18
2.4.	Propuesta de diseño	19
2.5.	Modelo de diseño utilizando Matlab	21
2.5.1.	Desarrollo de modelo transformada Census mediante el software Matlab ..	21
2.5.2.	Resultados del algoritmo desarrollado en Matlab	23
Conclusiones		25
Recomendaciones y trabajos futuros		26
Bibliografía		27

Índice de figuras

Figura 1: Descripción de la transformada Census. “FPGA- based stereo vision system using census transform for autonomous mobile robot”	14
Figura 2: Obtención de la distancia de Hamming. “FPGA- based stereo vision system using census transform for autonomous mobile robot”	15
Figura 3: Función costo de correspondencia generado en Matlab.	15
Figura 4: Análisis de transformada rápida de Census. “Semi-global matching-based disparity estimate using fast census transform”	16
Figura 5: Modelo del problema planteado.....	17
Figura 6: Mapa conceptual.	18
Figura 7.a y 7.b: Imágenes estereoscópicas de prueba (Middlebury Benchmark dataset).	21
Figura 8: Mapa de disparidad con las distancias de Hamming generado en Matlab.....	23
Figura 9: Mapa de disparidad en mapa de colores generado en Matlab.....	24

Índice de tablas

Tabla 1: Información de FPGA Cyclone V. Intel: “Cyclone V FPGA features”..... 12



Introducción

La visión estereoscópica es la obtención de 2 imágenes adquiridas por 2 cámaras separadas a una corta distancia. Con las imágenes obtenidas, a través de algoritmos de procesamiento de imágenes se puede determinar la correspondencia entre el pixel de la imagen original y su par estéreo. Esto contribuye en la construcción de mapas de disparidad o de profundidad, para identificar la posición en la que se encuentran los objetos respecto a las cámaras.

Estos mapas de disparidad tienen diferentes usos como la navegación autónoma para brindar a los robots una navegación independiente y así contar con un sistema de evasión de obstáculos en ambientes cerrados tales como el socavón de una minería, los campos de vides, entre otros.

Para la creación de dicho mapa de disparidad se utilizará el algoritmo de correspondencia de transformada Census, y este será desarrollada en una tarjeta FPGA de Intel. De mismo modo, se presentará trabajos de investigación en el análisis de correspondencia mediante FPGA; sin embargo, en este presente trabajo se desea desarrollar un modelo utilizando elementos adicionales al chip FPGA Cyclone V como su procesador interno NIOS II y su memoria externa SDRAM.

El objetivo principal de este trabajo será obtener el modelo de solución para realizar el algoritmo de correspondencia en imágenes estereoscópicas a través de transformada Census mediante el uso de tarjeta FPGA para determinar la distancia en la que se encuentran los objetos respecto a la fuente de adquisición de imágenes.

En el primer capítulo, se presentará la motivación de la navegación autónoma y sus beneficios de implementación en el área de agricultura, así como el estado del arte de las tecnologías empleadas para la visión estereoscópica usando un FPGA; para luego presentar los

objetivos del trabajo. En el segundo capítulo, se presentará la teoría de los vehículos aéreos no tripulados y el algoritmo de correspondencia de transformada Census, de igual manera, se mostrará las características de la tarjeta FPGA Cyclone V y se desarrollará el diseño de su modelo en el software Matlab que servirá como punto de comparación con el diseño para el hardware planteado. Finalmente se terminará con las conclusiones y trabajos futuros de dicha investigación, así como sus recomendaciones al momento de la implementación.



Capítulo 1: Diseño del hardware para un vehículo aéreo no tripulado

1.1. Motivación de problemática

En la agroindustria, hay trabajadores que realizan la actividad de conteo de frutos en los campos. Sin embargo, estas agroindustrias al tener grandes áreas de cultivo, aumenta la probabilidad de un mal conteo estimado respecto al real para la producción de la industria. Esto puede ser por diferentes factores, tales como el agotamiento de la vista del trabajador, debido a que el enfoque visual cercano y lejano de manera repetitiva, conocido como “reflejo acomodativo”, causa la fatiga visual [1].

Por ello, se ha planteado la necesidad de automatizar esta actividad mediante el uso de un vehículo aéreo no tripulado, también conocido como dron, para reducir la fatiga visual del trabajador en dicha actividad.

En adición, la automatización en la agricultura ayudaría a disminuir los accidentes laborales en los campos de cultivo. Según la SUNAFIL (Superintendencia Nacional de Fiscalización Laboral), los accidentes más frecuentes en el sector agrícola son los problemas respiratorios. Esto es consecuente debido a la exposición del trabajador a agentes tóxicos como los plaguicidas y fertilizantes durante largas jornadas laborales.

Asimismo, la prolongada jornada laboral expuesto al calor intenso, trae consigo consecuencias como edemas en las extremidades, calambres musculares y deshidratación. Siendo el más peligroso el denominado golpe de calor, debido a la superación del calor corporal a los 40.6 °C, siendo mortal entre un 15% y 25% de los casos registrados. [2]

Otro problema a identificar son las posiciones no ergonómicas para el conteo de frutas como es el caso de la vid. Como consecuencia de ello, ocasionaría en el agricultor problemas lumbares, lo que conllevaría a un descanso médico para el trabajador y por ende pérdidas en horas laborales para la empresa [3].

Por esta razón, se propone la navegación autónoma en drones para apoyar a los agricultores en la actividad del conteo de frutas y reducir las enfermedades o accidentes laborales descritas anteriormente.

Para la navegación autónoma de un vehículo aéreo no tripulado, se puede realizar mediante el uso del GPS que tiene incorporado para su navegación. Sin embargo, es necesario considerar que la separación entre cada tallo de fruto tiene un aproximado de dos metros de longitud. Por lo tanto, el sistema de navegación por GPS de un dron al tener un margen de error de ± 5 metros, ocasionaría la colisión del dron dentro del área de cultivo, impidiendo su navegación en espacios reducidos.

Es por esta razón, que se propone otro método de navegación autónoma, como es el caso de procesamiento de imágenes; específicamente en el empleo de visión estereoscópica, el cual consiste en la adquisición de dos imágenes a través de dos cámaras separadas a una corta distancia, para luego emplear algoritmos de correspondencia y así obtener los pixeles correspondientes en el par estéreo para finalmente generar un mapa de profundidad o disparidad que permita conocer la proximidad de los objetos hacia la fuente de adquisición de imágenes.

En adición, esta propuesta de navegación autónoma es un sistema de evasión de obstáculos para el vehículo aéreo no tripulado, para lo cual evitaría las colisiones con objetos cercanos y permitiría el ingreso a lugares reducidos para los agricultores, aumentando la eficacia en la actividad de conteo de frutas.

Finalmente, se debe mencionar que el diseño del algoritmo de correspondencia para imágenes estereoscópicas será desarrollado en una tarjeta FPGA, debido a su capacidad de tener una arquitectura programable y por ende permite tener modificaciones a futuro, así como su bajo consumo de potencia al momento de la operación.

1.2. Estado del arte

1.2.1. Sistema basado en FPGA para visión estéreo usando transformada Census para robots autónomos móviles

En el año 2015 Kaichiro Nakazato et.al, llegó a demostrar que la producción de imágenes estéreo se puede realizar con un FPGA mediante el cálculo de bordes con los filtros de Sobel y próximamente aplicando la transformada Census, realizando un análisis de correspondencia entre el pixel de la imagen de la izquierda con respecto a sus pixeles vecinos de la imagen de la derecha, este método tiene la ventaja de que toma pocos ciclos de reloj [4].

Además, la información se obtiene de manera directa a través del CMOS de las cámaras OmniVision OV5642, a una tasa de un byte por ciclo de reloj. Para la recepción de imágenes son en diferentes formatos, como el RGB, HSV o YUYV; este último es recomendado por el paper ya que primero deberá pasarse a escala de grises para realizar el cálculo más rápido a través de sus 3 variables importantes, la luminosidad, cromancia U y cromancia V.

Luego los patrones de intensidad mediante la transformada Census se representan por un vector y el costo computacional entre los dos vectores Census es definido por la distancia de hamming [5]. Por lo cual se procede a calcular dicha distancia para así conocer la disparidad de los objetos y su profundidad en la que se encuentran.

En este trabajo se demostró que el sistema de visión estereoscópica es 4.3 veces más rápido que el procedimiento hecho por software; además, presentó un error de emparejamiento de 4.17%. Se debe de considerar que en el paper recomiendan que las cámaras estén distanciadas entre 1000 mm y 10000 mm debido su resolución.

1.2.2. Sistema de alta calidad de visión estéreo en tiempo real para FPGA

En el año 2015, Wenqiang Wang et al., propuso otro algoritmo para la visión estereoscópica, usando un Altera 5SGSMD5H2 FPGA, y por el método de optimización

semiglobal, como la Transformada Census de diferencia absoluta (AD-Census) le permite hallar su emparejamiento en la otra imagen de la otra cámara, dando una precisión de pérdida de información de 0.3%. Se demostró que el algoritmo propuesto alcanza buenos resultados en imágenes de 1600 x 1200 a una velocidad de 42.61 FPS, con un margen de error porcentual de 5.61; además, el algoritmo Adaptive Census (ADCensus) demostró ser efectivo, pero pierde su precisión conforme aumenta la resolución de las cámaras. [6]

1.2.3. Marco novedoso para estimación de disparidad en FPGA

En el año 2016, Vasundhara et. al, propone que se puede usar el algoritmo de suma y diferencia absoluta y un FPGA descrito en un lenguaje de descripción de hardware como VHDL, para luego ser probado en un Zed board zynq XC7Z2020. Este método consiste en utilizar las imágenes en escala de grises y mediante el análisis por métodos locales, se puede calcular la correspondencia entre ambas imágenes [7]. Para ello se emplea la ecuación de SAD (sum and absolute difference) que se presenta en la ecuación 1,

$$SAD = \sum_{(i,j) \in W} |I_l(i,j) - I_r(i+d,j)|^2, \quad (1)$$

donde el análisis de pixeles por ambas cámaras debe de encontrar el menor valor de la distancia 'd' en la función de costo de disparidad. Para el diseño del hardware, se debe de contar con un buffer de 24 bits para almacenar la cadena de bits que compone el pixel de la imagen de la cámara OV7670 que realiza en formato RGB565. Como resultado se presentó una serie de optimización de recursos, siendo SAD una alternativa para procesamiento rápido para hallar la profundidad de los objetos.

1.2.4. Diseño de un sistema embebido en tiempo real con FPGA para visión estéreo

En el año 2018, Stefania Perri, propuso un sistema embebido "system on chip" para el FPGA, con la finalidad de implementar el Support Local Binary Pattern (SLBP) que consiste en el mismo método de cálculo de distancia de disparidad, que incluye el procesamiento del

sistema, así como la lógica programable. Esto optimiza el tiempo de ejecución debido a que usa HSL pipeline.

Se comprobó la funcionalidad de este dispositivo con otros estados del arte del Middlebury benchmark. Como resultado se demostró que el sistema embebido propuesto tiene buenos resultados con un fps de 101 para la visión estereoscópica. [8]

1.2.5. Sistema de doble canal de adquisición de imágenes basado en FPGA

En 2019, Shi Haitao et. al, diseñó una arquitectura en FPGA para la adquisición de imágenes con 2 cámaras CMOS OV 5640. En ello, expone las ventajas en la ejecución de procesamiento, el método de pipeline y el procesamiento paralelo.

Para la adquisición de imágenes de las cámaras plantea el protocolo de comunicación SCCB (serial camera control bus), asimismo, realiza el diseño de una máquina de estados para realizar la escritura y lectura de los datos, al igual que el uso de dos FIFO asíncronos para la lectura y escritura de información.

Finalmente, el autor desarrolla el módulo de operación almacenada en la memoria DDR3 SDRAM para ser mostrada en un monitor vía cable HDMI. Dando como resultado la muestra de imágenes duales en tiempo real a una resolución de 640x 720 píxeles. [9]

1.3. Justificación

La agricultura es una actividad importante en la economía del Perú; sin embargo, es una actividad que trae la mayor tasa de accidentes laborales según un estudio realizado por la fundación de prevención de riesgos laborales. El estudio indica, que el 67% de los trabajadores son sometidos a movimiento repetitivos y 26% de ellos permanecen en el campo todo el día e incluso en superficies irregulares.

Además, el 48% de los trabajadores realizan posturas forzadas para sus actividades como el conteo de vid o análisis de campo. Es por ello, que se considera necesario un sistema

automático para realizar dicha actividad, por ello, se empleará el uso de drones para transitar en dicha zona [10].

Además, el sistema de visión estereoscópica es necesario para poder conocer la profundidad de los objetos y respecto a ello tomar acciones de movimiento para realizar la evasión del obstáculo. En cuanto al uso del FPGA se debe por su principal ventaja, siendo este el diseño de su arquitectura, lo cual permitirá al usuario realizar diferentes modificaciones a futuro.

Y finalmente, un sistema automático de monitoreo permitirá llevar un mejor análisis de los cultivos, ya que puede ser lanzado a cualquier hora para analizar los frutos, y si un fruto presenta una enfermedad, este puede ser guardado en un registro para avisar a los trabajadores y darle su control requerido.

1.4. Objetivos

1.4.1. Objetivo general

Diseñar el modelo de solución para el algoritmo de correspondencia entre imágenes estereoscópicas utilizando transformada Census mediante un FPGA.

1.4.2. Objetivos específicos

- Estudiar la plataforma de integración de sistema para desarrollar un algoritmo de correspondencia en FPGA.
- Desarrollar y estudiar los algoritmos de correspondencia por transformada Census.
- Diseñar el modelo de solución del algoritmo de correspondencia en el software Matlab.

Capítulo 2: Teoría

2.1. Visión general

La visión estereoscópica es utilizada para obtener la disparidad de los objetos con la finalidad de calcular la distancia en la que se encuentra los objetos con la imagen proyectada por las cámaras, esta teoría tiene diferentes aplicaciones, siendo uno de ellos en vehículos aéreos no tripulados o drones, esto servirá como sistema de evasión de obstáculos, para conocer la proximidad de los obstáculos y así evitar colisiones.

Para el desarrollo de la visión estereoscópica, se basa en algoritmos de correspondencia; es por ello que se analizará el algoritmo de la transformada Census, que consiste en seleccionar dentro de una ventana sobre la imagen un conjunto de píxeles para luego realizar el análisis entre píxel central con el píxel dentro de la ventana, obteniendo así una cadena de bits. Luego de repetir dicho análisis tanto para la imagen original como para el par estéreo, se realiza una operación XOR con ambas cadenas de bits para determinar lo que se conoce como distancia de Hamming.

Al repetir dicho proceso a lo largo del par estéreo, se obtienen diferentes distancias de Hamming para cada caso, esto permite construir la función costo de correspondencia. Para luego, determinar el menor valor de dicha función y así conocer la correspondencia de los píxeles como relación de disparidad o profundidad en la que se encuentran los objetos respecto a la fuente receptora de imágenes. Dado que los valores enteros representan una imagen a escala de grises, un mapa de colores facilitará la visión de la profundidad de los objetos.

Para el análisis de algoritmo de correspondencia por transformada Census se plantea un modelo de simulación en un software de procesamiento de imágenes como Matlab para determinar el análisis de las funciones costo de correspondencia. Y en base a ello, se puede definir un modelo de solución para su desarrollo en una tarjeta FPGA.

2.2. Marco teórico

2.2.1. Vehículos aéreos no tripulados

Un vehículo aéreo no tripulado, o también conocido como drones, tiene la particular característica de ser manipulado por el usuario desde tierra con un control remoto. Los drones fueron principalmente creados con fines militares. Sin embargo, hoy en día se usa en diferentes áreas como comercialización, entretenimiento y en casos particulares como la agricultura. Se debe de mencionar que existen dos tipos de drones en el mercado: [11]

- **Multi-rotor:** Tienen hélices en la parte superior como un helicóptero, se caracteriza por su simpleza y su estabilidad debido a que sus motores se ubican a la misma distancia de su centro de gravedad. Este tipo de dron no tiene una mayor autonomía de 20 minutos de vuelo.
- **Alas fijas:** Con alas y un motor en la parte trasera como un avión, dado por su habilidad de planeamiento pueden ahorrar el consumo de energía. El planeamiento en este tipo de dron causa que el aterrizaje sea más seguro, pero es necesario que el área sea mayor al no poder ascender o descender verticalmente.

La ventaja de los primeros es la capacidad de mantenerse en un mismo punto de vuelo, por lo que mediante el acelerómetro y giroscopio mantiene su equilibrio durante el vuelo.

Los drones pueden llevar consigo una computadora embebida, estos dan la posibilidad del dron a llevar consigo periféricos de entrada y salida. Tales como sensores, cámaras, entre otros. Asimismo, cuentan con una memoria interna donde almacena la información de los periféricos, o inclusive son guardados a una memoria externa.

Para la manipulación del dron, es a través de un sistema llamado Pixhawk, es un hardware de código abierto con la finalidad de brindar al usuario manipular los motores que controlan las hélices del dron.

Para el sistema de navegación del dron existen dos tipos de navegación, mediante el uso de un control automático para direccionar al dron, o también mediante ‘waypoints’, que consiste en una configuración predeterminada del usuario para indicar la ruta del dron mediante una serie de puntos en el mapa. Es necesario aclarar que ambos métodos de navegación utilizan GPS (global positioning system), por lo que es diferente al modelo de solución, que busca una navegación autónoma por visión estereoscópica.

2.2.2. Visión estereoscópica

La visión estereoscópica consiste en el análisis de dos imágenes adquiridas por cámaras separadas a una corta distancia, y a través del análisis de correspondencia permite determinar la proximidad de los objetos hacia la fuente de recepción de imágenes, análogamente, el ser humano realiza la misma operación a través de los ojos [12]

Para el desarrollo del modelo de solución, se analizará dos imágenes que están separadas a una corta distancia y serán evaluados por los siguientes criterios en la visión estereoscópica.

a) Geometría epipolar

Es la geometría de visión estereoscópica que permite determinar la relación de un punto y su proyección en cada cámara. La unión del centro de las cámaras corresponde a una línea base y la unión del punto proyectado a dicha línea base se conoce como líneas epipolares; en base a esta información se permite construir matrices extrínsecas o esenciales para conocer el comportamiento de la cámara en cuanto a translación y rotación, y de esta manera poder calcular la correspondencia de los píxeles de una imagen y su par estéreo.

b) Rectificación de imagen

Es el proceso de transformación de imágenes estereoscópica para que el par estéreo sean paralelas. Para conseguir la rectificación de imágenes es necesario las matrices esenciales de ambas cámaras y asumir que no existe una rotación relativa entre ellas.

Luego, se debe de determinar que las líneas epipolares de ambas imágenes adquiridas sean paralelas, lo cual resumiría el cálculo de correspondencia de píxeles a una sola dimensión en el eje horizontal, permitiendo que el procesamiento sea más rápido.

c) **Disparidad**

Es la distancia que determina la correspondencia del píxel de la imagen izquierda con su par estéreo a través de la línea epipolar de dicha imagen.

Mediante algoritmos de correspondencia, el valor resultante ayuda a determinar que mientras este valor se hace más grande indica que el objeto está más cerca al par de cámaras, siendo esto un indicador fundamental para la aproximación de objetos. [13]

2.2.3. **Tabla de FPGA**

A continuación, en la tabla 1 se presenta la siguiente tabla de información de los FPGA de la serie Cyclone V de la compañía Intel. Para el desarrollo de este trabajo de investigación, se está utilizando la tarjeta FPGA 5CSEMA5F31C6N, siendo la tercera columna de la tabla.

Tabla 1: Información de FPGA Cyclone V. Intel: “Cyclone V FPGA features”. [14]

Producto	Cyclone V SE SoC			
	5CSEA2	5CSE4	5CSA5	5CSA6
LE (K)	25	40	85	110
ALMS	9,434	15,094	32,075	41,509
Registros	37,736	60,376	128,300	166,036
Variable Precisión DSP	36	84	87	112
18 x 18 multiplicados	72	168	174	224
Núcleos de procesador	Solo o dual			
Reloj de CPU (MHz)	925			
PLL (FPGA)	5	5	6	6
PLL (HPS)	3			
GPIO (FPGA)	145	145	288	288
GPIO (HPS)	181			
Memoria	DDR3, DDR2, LPDDR2			

2.2.4. FPGA Cyclone V (5CSEMA5F31C6N)

El hardware para el desarrollo del trabajo de investigación será un FPGA cuya descripción es en Verilog HDL y tiene la ventaja de realizar su procesamiento de forma paralela, así como el uso de pipeline. Cuenta con 85 000 elementos lógicos programables para realizar las tareas que se describirán durante su desarrollo.

Además, se debe de mencionar que el chip Cyclone V es un SoC (system on chip) que permite tener elementos adicionales como el procesador incorporado, tales como el procesador NIOS II y otro procesador modelo ARM Cortex A9, con un reloj interno es de 800 MHz. Estos procesadores se pueden utilizar en la tarea de distribución de información de las imágenes estereoscópicas

Otro aspecto a considerar del FPGA es la memoria externa en la tarjeta. Cuenta con una memoria de 64MB SDRAM que transmite información por una canal de bus de 16-bits. Se utilizará para almacenar la información de los pixeles de las imágenes estereoscópicas.

2.2.5. Diseño de la plataforma de integración de sistema

La plataforma de integración de sistema permite la interacción de las compuertas lógicas del chip FPGA con su procesador incorporado NIOSII del empaquetado Cyclone V, para realizar la interacción se planea usar la herramienta Qsys del software Quartus. En adición, se define las direcciones bases de los puertos de entradas y salidas para luego ser manipulado por el procesador NIOSII y realizar los algoritmos de correspondencias a través de la transformada Census.

2.2.6. Memoria SDRAM

La tarjeta FPGA Cyclone V cuenta con una memoria externa SDRAM de 64 MB que está organizada como una memoria de 8Mx16bitsx4 bancos, este dispositivo electrónico será utilizado para almacenar las imágenes estereoscópicas que cuentan con una resolución de

640x480. Asimismo, luego de procesados las imágenes por el bloque de transformada Census, los datos procesados deberán de ser guardados en la misma memoria para visualizarlos.

2.2.7. Función de correspondencia a partir de la transformada Census

Es una medida no paramétrica que es robusta frente a cambios de ganancia y de ganancia (offset), este reemplaza el valor de un pixel mediante el análisis del pixel central con sus pixeles vecinos.

La transformada Census es una variación de la transformada rank que mantiene la distribución espacial de los vectores. Para cada valor central se obtiene una cadena de bits correspondientes que se leerá de izquierda a derecha. Para ello, se necesita tener la imagen a escala de grises para tener el valor de su intensidad, para luego realizar el siguiente algoritmo:

Tomar el valor central de una ventana de análisis de 5x5 como se puede observar en la figura 1; luego, si el valor del pixel central es menor al valor cercano y dentro de la ventana, se asigna el valor de '1', caso contrario se asigna el valor de '0'.

Con los números obtenidos, se crea una cadena de bits correspondiente a la transformada Census del primer pixel de la imagen de la izquierda.

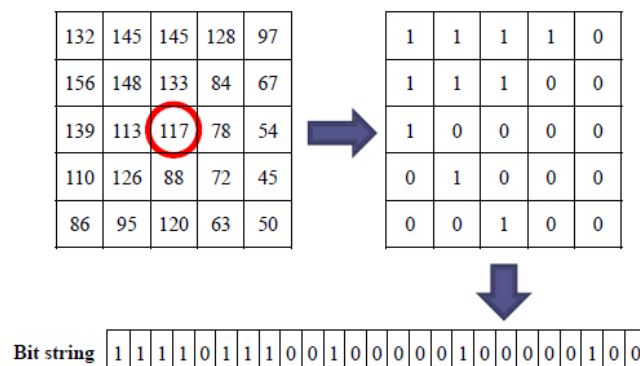


Figura 1: Descripción de la transformada Census. “FPGA- based stereo vision system using census transform for autonomous mobile robot” [15]

Este mismo procedimiento se realiza para el primer pixel del par estéreo, por lo que se obtienen 2 cadenas de bits, estos deben de pasar por la operación XOR y a la cadena de bits resultante se debe de sumar los valores de ‘1’ como se muestra en la figura 2, para obtener el denominado distancia de Hamming.

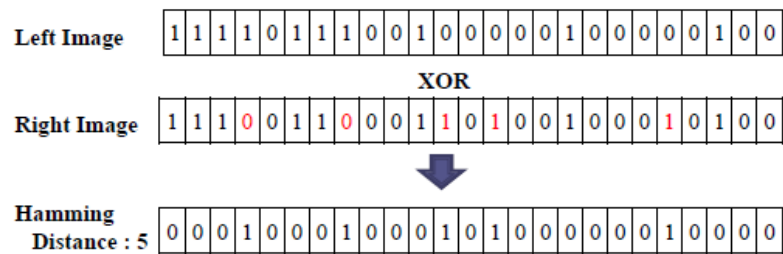


Figura 2: Obtención de la distancia de Hamming. “FPGA- based stereo vision system using census transform for autonomous mobile robot” [16]

Al realizar dicha operación continuamente con los próximos pixeles en la imagen estéreo se puede construir una función costo de correspondencia, para finalmente determinar que la posición del menor valor de dicha función indica el pixel correspondiente de la imagen original con su par estéreo como se muestra en la figura 3.

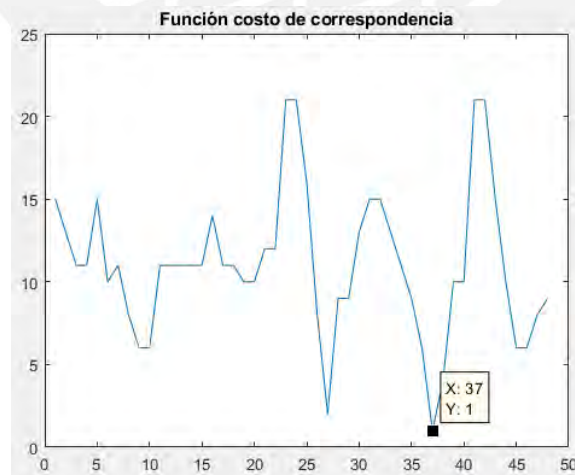


Figura 3: Función costo de correspondencia generado en Matlab.

La transformada Census también puede ser procesado por el método de métodos locales, pero tiene la desventaja de agregar un costo computacional al procesado. También, se puede

realizar por métodos globales que optimiza el cálculo de la disparidad inicial y tienen mayor precisión [17].

2.2.8. Transformada rápida de Census

La transformada rápida de Census permite realizar una operación más rápida al analizar menor cantidad de píxeles y en una ventana mayor, como es el caso de una ventana de 7×7 , donde se presenta el siguiente algoritmo. [18]

Analizar el valor central con su fila, columna y diagonal en la que se encuentra, teniendo un total de 8 caminos por analizar, partiendo del bit central a analizar. En el caso de las filas el valor central ubicado en $I(i,j)$ es analizado con el número de $I(i+2,j)$, mientras que para el valor de $I(i+1,j)$ se debe de analizar con el valor de $I(i+3,j)$ tal como se muestra en la figura 4.

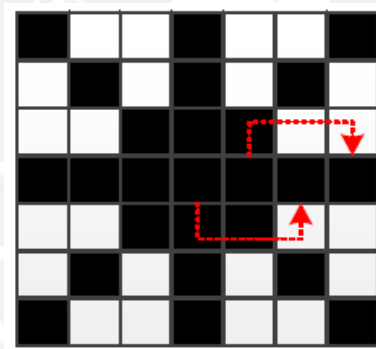


Figura 4: Análisis de transformada rápida de Census. “Semi-global matching-based disparity estimate using fast census transform” [19]

El resto de los valores que no forman parte de los caminos definidos anteriormente, no serán analizados por tener una relación con sus píxeles cercanos, haciendo que el desarrollo de este algoritmo no pierda información vital durante el procesamiento de la imagen.

Luego, se realiza el mismo método expresado en la transformada Census para obtener las distancias de Hamming y así construir la función costo de correspondencia. Este método permite ahorrar ciclos de reloj en el sistema implementado.

2.3. Modelo

A continuación, se presenta el modelo de solución para el trabajo de investigación, el desarrollo de los bloques se explicó con anterioridad, así como las razones de su implementación.

2.3.1. Diagrama de bloques

En la figura 5 se presenta un diagrama de bloques que permite observar los pasos a seguir para la obtención de la disparidad de los objetos en las imágenes estereoscópicas.

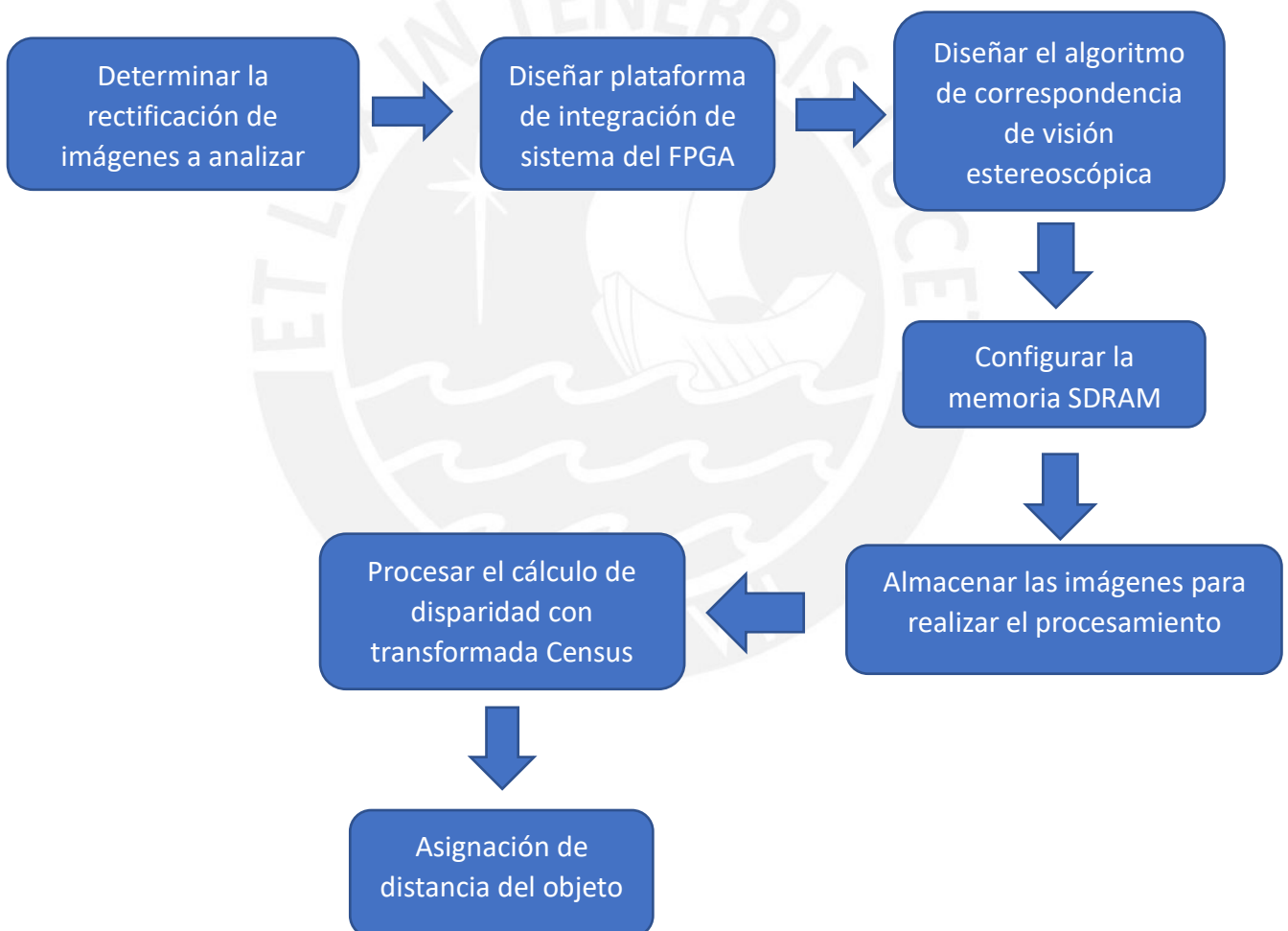


Figura 5: Modelo del problema planteado.

Aquí se muestra el modelo de solución para el trabajo de investigación, es de carácter lineal y sin necesidad de retroalimentación.

2.3.2. Modelo teórico

Para el desarrollo del trabajo de investigación se plantea el siguiente mapa conceptual con los requerimientos necesarios para su implementación como se muestra en la figura 6.

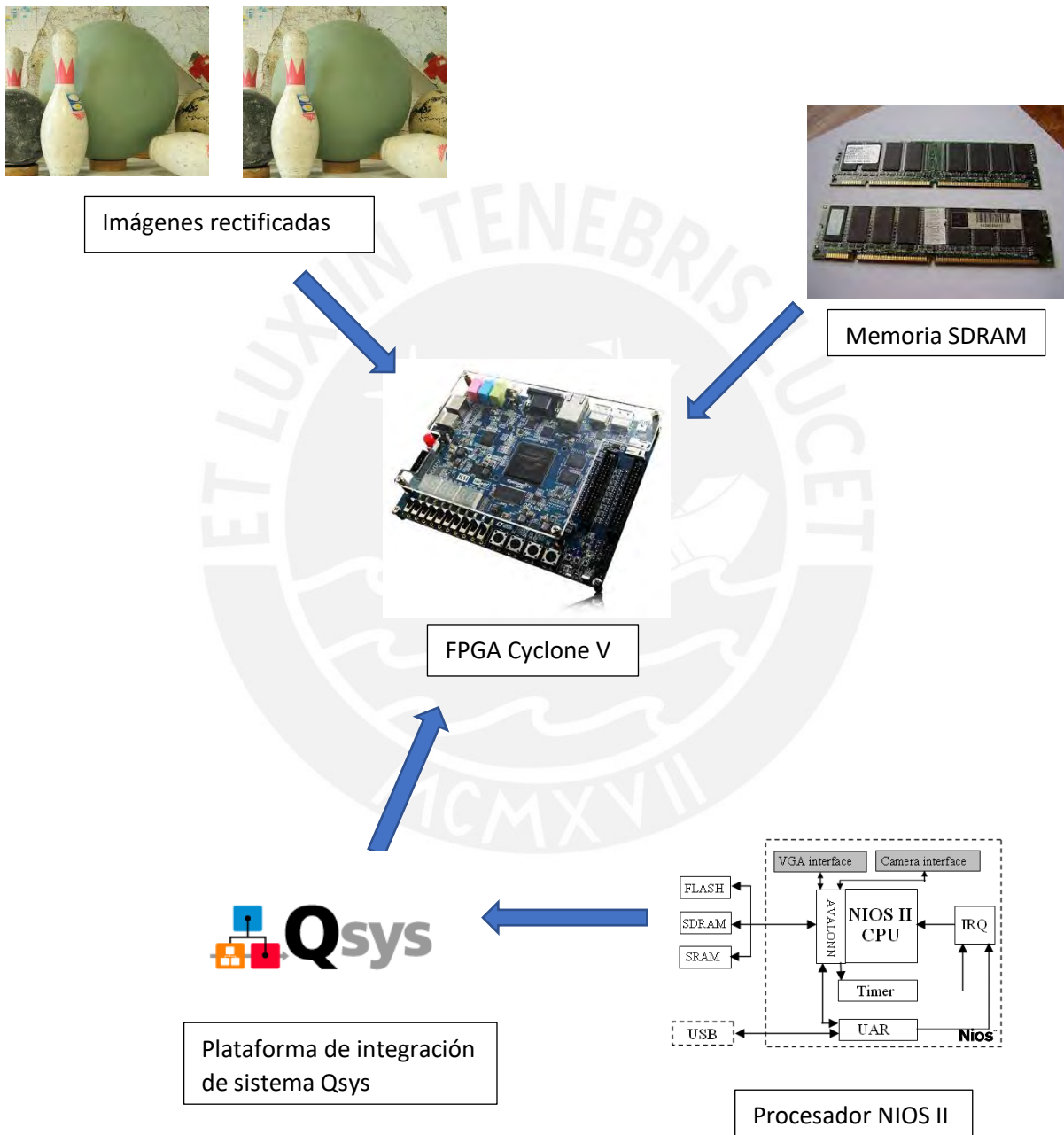


Figura 6: Mapa conceptual.

2.4. Propuesta de diseño

Luego de explicado la teoría, para el diseño de sistema de visión estereoscópica mediante un FPGA, se hará uso de las compuertas lógicas presentes en el chip FPGA de la tarjeta DE1 Cyclone V de Intel; sin embargo, el chip Cyclone V es un “system on chip” (SoC), por lo que es un circuito integrado con módulos adicionales a los de un computador, siendo uno de ellos el procesador integrado NIOS II.

De esta manera, se realiza el diseño de un sistema embebido para generar una combinación entre los elementos de software presente en el procesador y las compuertas lógicas hardware ubicados en el chip FPGA; para ello, es necesario generar una plataforma de integración de sistema, para que mediante uso de algoritmos descritos en lenguaje de programación C o lenguaje ensamblador, el procesador NIOS II pueda interactuar con los módulos digitales y los componentes externos al chip FPGA Cyclone V.

Para el diseño digital de la plataforma de integración de sistema se hará uso de la herramienta Qsys del software Quartus Prime; posteriormente, el diseño será conectado con los módulos digitales ubicados en el chip FPGA diseñados en lenguaje de descripción de hardware Verilog HDL; finalmente, para uso del procesador NIOS II se hará uso de la herramienta NIOS II Software Build Tool for Eclipse del software Quartus Prime.

Las ventajas de esta propuesta de diseño es disminuir las complicaciones de operación por sus algoritmos secuenciales en el procesador NIOS II, y en el ahorro de ciclos de reloj debido a la operación paralela que puede realizar las compuertas lógicas en el chip FPGA.

Con el hardware mencionado, se procede a la descripción del modelo de diseño, en base a la teoría descrita, es necesario contar con 2 imágenes, la imagen izquierda y derecha, también llamado imagen original y su par estéreo, para el análisis de pixeles de correspondencia mediante la transformada Census. Las imágenes estéreo de prueba que se utilizarán serán las

imágenes del dataset de Middlebury Benchmark del año 2014, debido a que principalmente las imágenes ya están rectificadas como parte del preprocesamiento; además, son las imágenes de prueba para la evaluación de algoritmos de visión estereoscópica, así como su comparación con otros algoritmos posicionados en un ranking a nivel de investigación.

De esta manera, las imágenes deben de ser guardadas en una memoria ubicada en la tarjeta FPGA; a continuación, se describe las 2 memorias que cuenta dicha tarjeta:

- Primero está la “On-Chip memory” que está incorporada en el SoC del chip Cyclone V, esta memoria puede ser configurada como memoria RAM o memoria ROM; sin embargo, tiene una capacidad máxima de 4096 bytes, por lo que no podría almacenar la información de una imagen.
- En segundo lugar, se tiene la memoria SDRAM, es una memoria externa al chip Cyclone V y puede ser accedida mediante un módulo digital a través del chip FPGA o utilizando el procesador NIOS II, y su capacidad de almacenamiento es de 64 MB, por lo que sería ideal para guardar la información de píxeles de la imagen original y su par estéreo.

De esta manera, se determina que es necesario utilizar la memoria SDRAM para guardar las imágenes del dataset de Middlebury Benchmark y realizar la operación de correspondencia mediante transformada Census; sin embargo, la memoria SDRAM es una memoria volátil, por lo que será necesario cargar las imágenes antes del procesamiento, es por ello que se utilizará el procesador NIOS II para generar un algoritmo de inicio de memoria antes de enviar la información al módulo digital de transformada Census.

Finalmente, los bloques digitales de transformada Census realizarán las operaciones de correspondencia, y los datos de salidas serán almacenado en un arreglo de registros denominado mapa de disparidad para ser mostrado en un monitor por puerto VGA a una

resolución de 640x480 a 60 Hz, para ello será necesario diseñar dicho módulo; asimismo, estos datos del mapa de disparidad deberán ser retornados al procesador NIOS II y enviados al computador a través de un módulo digital para ser visualizados por una terminal.

2.5. Modelo de diseño utilizando Matlab

Como parte del desarrollo del trabajo de investigación, se desea implementar el algoritmo de la transformada Census para determinar la correspondencia entre 2 imágenes estereoscópicas en una tarjeta FPGA; sin embargo, se ha desarrollado ese mismo algoritmo en el software Matlab para ser utilizado como modelo de referencia.

2.5.1. Desarrollo de modelo transformada Census mediante el software Matlab

Para el desarrollo del algoritmo descrito, es necesario cargar las imágenes estereoscópicas, para ello se utiliza la función `imread` (`image`). Tal como se muestra en la figura 7.a y 7.b, se puede visualizar el par estéreo de nombre `Adinrodack.png` que será redimensionado a la resolución de 640x480 píxeles a través de la función `imresize` (`image`, [`width height`]) para la operación Census.



(a)

(b)

Figura 7.a y 7.b: Imágenes estereoscópicas de prueba (Middlebury Benchmark dataset).

Luego, la imagen debe de pasar de formato RGB a escala de grises para su procesamiento, para ello se utiliza la función `rgb2gray`(`image`). Con las imágenes obtenidas, se procede a realizar el siguiente algoritmo en Matlab.

- a) Se crea dos iterativas para ir por todo el ancho y largo de ambas imágenes.
- b) A una variable se le asigna un arreglo que representa la ventana Census de 9x9 para la imagen de la izquierda; luego, se crea un segundo arreglo de ventana Census 9x9 con un desplazamiento de 48 niveles de disparidad en la imagen de la derecha, dando de resultados una matriz de 9x56.
- c) A ambos arreglos, sus elementos cuentan con un valor inicial de '1' mediante la función `ones(n1, n2)`.
- d) Se procede a hacer el análisis Census para determinar que elemento se le asigna el valor de '0' o el valor de '1', esto se realiza tanto para el arreglo 9x9 de la imagen de la izquierda y para el primer arreglo 9x9 de la imagen derecha.
- e) Se transforma ambos arreglos 9x9 a vectores 81x1 que son denominados vectores Census.
- f) Se convierte el tipo de variable `double` a valor lógico mediante la función `logical(vector)`.
- g) Se realiza la operación XOR entre ambos vectores Census mediante la función `bitxor(vector1, vector2)`.
- h) Se retorna el tipo de variable resultante de lógico a `double` mediante la función `double(vector)`.
- i) Se realiza el conteo de '1' presente en el vector final mediante la función `sum(vector_double)` y se almacena dicho valor en un vector denominado vector de Hamming de 48 elementos.
- j) Se repite la operación (b) con el siguiente arreglo 9x9 de la imagen de la derecha, hasta completar el vector de Hamming con sus 48 elementos.
- k) Se calcula el menor valor y posición del vector de Hamming mediante la función `[valor_min,pos_min] = min(vector_Hamming)`.

- l) La variable `pos_min` es almacenado en un arreglo de mapa de disparidad.
- m) Se presenta los resultados con `imshow(mapa_disparidad)`.

2.5.2. Resultados del algoritmo desarrollado en Matlab

Luego de realizar el algoritmo en el software Matlab, se ejecuta el algoritmo obteniendo los siguientes resultados. En la figura 8 se puede observar en el mapa de disparidad los valores en un rango de 1 a 48, indicando la posición de la menor distancia de Hamming para cada caso. Esto se ve representado en una imagen a escala de grises al aplicar la función `imshow(mapa_disparidad)`.



Figura 8: Mapa de disparidad con las distancias de Hamming generado en Matlab.

Para tener una mejor visión de los resultados se planteó realizar una barra de colores para segmentar en un rango de distancias de Hamming un color específico. Para generar la barra de colores se creó una matriz de 48x3 elementos (las 3 columnas representan los colores RGB) y se inicializó todos los elementos con el valor de cero a través de la función `zeros(48,3)`; luego, en cada rango se asigna el valor de '1' en la columna correspondiente para los colores RGB y sus combinaciones de colores secundarios, dando como resultado la figura 9.

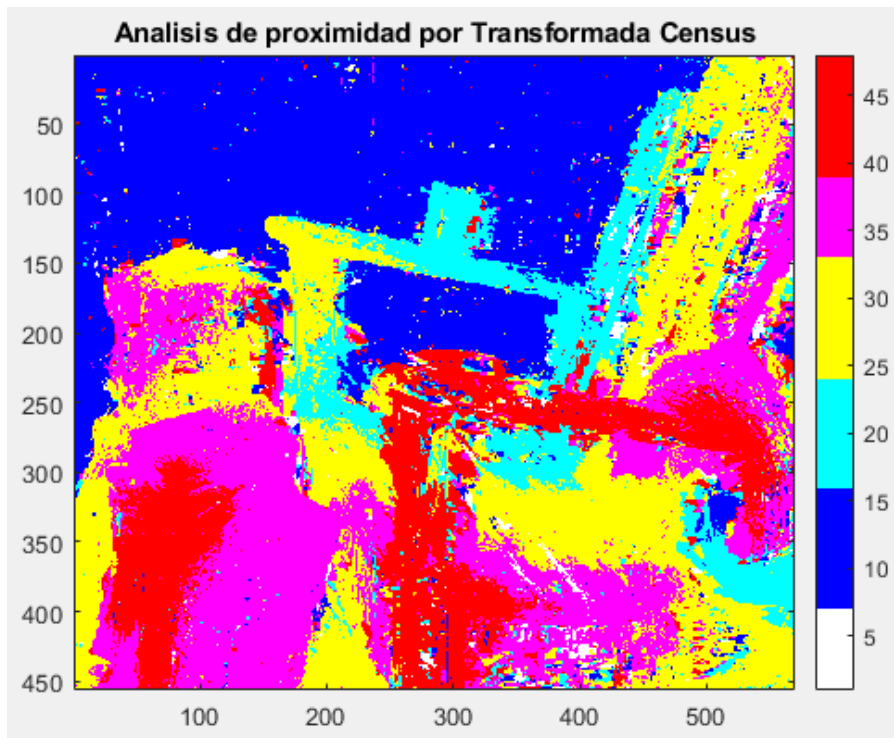


Figura 9: Mapa de disparidad en mapa de colores generado en Matlab.

Los resultados obtenidos como el mapa de disparidad con las distancias de Hamming y la imagen obtenida en formato de colores servirán como punto de comparación para el modelo de solución para el hardware implementado.

Conclusiones

- En el trabajo de investigación se ha planteado un modelo de solución para el desarrollo de algoritmos de correspondencia a través de transformada Census y su implementación en FPGA considerando las características del hardware mencionado.
- Se ha analizado la teoría de la transformada Census, así como la obtención de la distancia de Hamming, y la construcción de la función costo de correspondencia entre los píxeles de la imagen original y su par estéreo.
- En base a lo expuesto, la transformada Census permite un análisis de correspondencia con la operación lógica digital XOR, por lo que sería ideal implementarlo en un sistema FPGA para aprovechar las compuertas programables en el chip.
- Se propone como modelo de solución desarrollar una plataforma de integración de sistema que permita unir el procesador NIOS II de la tarjeta Cyclone V y el chip FPGA.
- Se ha desarrollado un algoritmo en el software Matlab para realizar el algoritmo de correspondencia a través de transformada Census para imágenes estereoscópicas.

Recomendaciones y trabajos futuros

- Como trabajo futuro se propone desarrollar el módulo digital de la transformada Census en el software Quartus Prime en el lenguaje de descripción de hardware Verilog HDL.
- Dicho módulo debe determinar la distancia de Hamming y construir la función costo de correspondencia para determinar la posición del menor valor de dicha función.
- Se diseñará la plataforma de integración de sistema para la conexión entre el procesador NIOS II con el chip FPGA a través de la herramienta Qsys y ejecutar los algoritmos por el procesador desarrollados en lenguaje C en el software NIOS II Software Build Tool for Eclipse.
- Se configurará la memoria SDRAM de la tarjeta FPGA para almacenar la información de las imágenes estereoscópicas.
- Se recomienda verificar que los recursos de la tarjeta Cyclone V 5CSEMA5F31C6N sean necesario para el diseño del módulo de la transformada Census.
- Se recomienda realizar una comparación entre los resultados obtenidos por la tarjeta FPGA y su modelo en software desarrollado por el programa Matlab para verificar la proximidad de los datos obtenidos en hardware y software.

Bibliografía

- [1] Varilux, “¿Cómo identificar la fatiga visual?”, 2019. [En línea]. Disponible en: <https://varilux.es/salud-visual/fatiga-visual/> [Accedido: 05-mayo-2019]
- [2] Revista colombiana de salud ocupacional, “Implicaciones en la salud de los trabajadores originadas por la exposición térmica en cultivos”, Artículo de investigación científica o tecnológica, Mayo 2016.
- [3] Sunafil, “Seguridad y Salud en el Trabajo en el Sector Agroindustrial”, 2016 [En línea]. Disponible en: <https://www.sunafil.gob.pe/noticias/item/3833-seguridad-y-salud-en-el-trabajo-en-el-sector-agroindustrial.html#ii-comit%C3%A9-de-seguridad-y-salud-en-el-trabajo> [Accedido: 05-mayo-2019]
- [4] Intel Software, “Census transform algorithm overview”, 2017 [En línea]. Disponible en: <https://software.intel.com/en-us/sample-census-transform-census-transform-algorithm-overview> [Accedido: 05-mayo-2019]
- [5] K. Nakazato et. al, “FPGA- based stereo vision system using census transform for autonomous mobile robot”, 2015 International Conference on Informatics, Electronics & Vision (ICIEV), June 2015
- [6] W. Wang et. al, “Real-Time High-quality stereo vision system in FPGA”, IEEE transactions on circuits and systems for video technology, vol.25, October 2015.
- [7] G. Vasundhara et. al, “A novel framework for disparity estimation in FPGA”, 2016 International Conference on Automatic Control and Dynamic Optimization Techniques, September 2016.
- [8] S. Perri et.al, “Design of real-time FPGA-based embedded system for stereo vision”, 2018 IEEE International Symposium on Circuits and Systems (ISCAS), May 2018.

- [9] Shi Haitao et.al, “Dual-channel image acquisition system based on FPGA”, 2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), p. 421-424, 2019.
- [10] Portal multimedia para la promoción de la ergonomía en el sector agrario, “Riesgos ergonómicos en el sector”, 2015 [En línea]. Disponible en: http://agrario.ibv.org/index.php?option=com_content&view=article&id=50&Itemid=141 [Accedido: 22-mayo-2019]
- [11] Ingeoexpert, “Tipos de dron: formas y control”, 2018. [En línea] Disponible en: <https://ingeoexpert.com/tipos-de-drones/?v=3acf83834396> [Accedido: 09-junio-2019].
- [12] E. Cárdenas et. al, “La estereoscopia, métodos y aplicaciones en diferentes áreas de conocimiento”, Revista científica General José María Córdova, vol.13, pp 201-219, Diciembre 2015.
- [13] H. Kenji et. al, “CS231A Course Notes 3: Epipolar Geometry”, Standford University. 2017.
- [14] Cyclone V producto table, “Cyclone V FPGA features”, 2019. [En línea] Disponible en: <https://www.intel.la/content/dam/www/programmable/us/en/pdfs/literature/pt/cyclone-v-product-table.pdf> [Accedido: 22-junio-2019].
- [15], [16] y [17] K. Nakazato et. al, “FPGA- based stereo vision system using census transform for autonomous mobile robot”, 2015 International Conference on Informatics, Electronics & Vision (ICIEV), June 2015
- [18] y [19] P. Xu et. al, “Semi-global matching-based disparity estimate using fast census transform”, 2016 International Congress on Image and Signal Processing, BioMedical Engineering and Infomatics (CISP-BMEI 2016), Octubre 2016.