



PROTOTIPO DE SISTEMA DE CAPTURA DE MOVIMIENTO DE LA
EXTREMIDAD SUPERIOR DEL CUERPO HUMANO BASADO EN VIDEO,
UTILIZANDO OPENCV

PRESENTADO POR:

JAVIER DAVID LARA PEREZ
ELKIN MANUEL TORRES MEDRANO

UNIVERSIDAD DEL SINÚ ELÍAS BECHARÁ ZAINÚM SECCIONAL CARTAGENA

FACULTAD DE CIENCIAS EXACTAS E INGENIERÍA

ESCUELA DE INGENIERÍA DE SISTEMAS

CARTAGENA - COLOMBIA

JUNIO DE 2019



UNIVERSIDAD DEL SINÚ
Elías Bechara Zainúm
Seccional Cartagena

PROTOTIPO DE SISTEMA DE CAPTURA DE MOVIMIENTO DE LA
EXTREMIDAD SUPERIOR DEL CUERPO HUMANO BASADO EN VIDEO,
UTILIZANDO OPENCV

Trabajo de grado presentado como requisito para optar el título de
INGENIERO DE SISTEMAS

Asesor disciplinar

LUIS FERNANDO MURILLO FERNÁNDEZ

Asesor metodológico

EUGENIA ARRIETA RODRÍGUEZ

UNIVERSIDAD DEL SINÚ ELÍAS BECHARÁ ZAINÚM SECCIONAL CARTAGENA

FACULTAD DE CIENCIAS EXACTAS E INGENIERÍA

ESCUELA DE INGENIERÍA DE SISTEMAS

CARTAGENA-COLOMBIA

JUNIO DE 2019

ACTA DE CALIFICACIÓN Y APROBACIÓN

Nota de aceptación:

Director de Escuela

Director de Investigaciones

Firma del jurado

Firma del jurado

Cartagena de Indias, junio de 2019

 <p>UNIVERSIDAD DEL SINÚ Eliás Bechara Zainúm Seccional Cartagena</p>	<p>PROCESO: INVESTIGACIÓN TÍTULO: ACTA PROYECTO DE INVESTIGACIÓN CÓDIGO: R-INVE-021</p>
--	---

**EL DIRECTOR DE INVESTIGACIONES DE LA UNIVERSIDAD DEL SINU
“ELIAS BECHARA ZAINUM” SECCIONAL CARTAGENA**

HACE CONSTAR QUE:

En Cartagena, a los 4 días del mes de junio del 2019, en la Oficina de la Dirección de Investigaciones de la Universidad, se aprobó por el jurado y se realizó la sustentación del Trabajo de Grado titulado “PROTOTIPO DE SISTEMA DE CAPTURA DE MOVIMIENTO DE LA EXTREMIDAD SUPERIOR DEL CUERPO HUMANO BASADO EN VIDEO, UTILIZANDO OPENCV “, que se desarrolló bajo la dirección de la Ingeniera EUGENIA ARRIETA RODRÍGUEZ y presentado por los egresados JAVIER DAVID LARA PÉREZ, ELKIN MANUEL TORRES MEDRANO.

Los jurados designados fueron los ingenieros AISNER JOSÉ MARRUGO JULIAO, MARÍA MERCEDES SUÁREZ SÁNCHEZ

Teniendo en cuenta la aprobación emitida, se encuentra que los egresados han cumplido con los requisitos de presentación y sustentación del trabajo de investigación, exigidos por el programa de INGENIERÍA DE SISTEMAS, Resolución 0178 de 15 de marzo de 2010.

Se expide esta constancia a los 7 del mes de junio de 2019

DIRECCIÓN DE INVESTIGACIONES
Universidad del Sinú

COORDINADOR DE INVESTIGACIONES
Escuela de Ingeniería de Sistemas

TABLA DE CONTENIDO

Agradecimientos	8
RESUMEN	9
INTRODUCCIÓN	10
1. DISEÑO METODOLÓGICO	12
1.1. PLANTEAMIENTO DEL PROBLEMA	12
1.1.1. Descripción del problema	12
1.1.2. Justificación	13
1.1.3. Formulación del problema	15
1.1.4. Alcance	15
1.1.5. Objetivos	16
1.2. ESTADO DEL ARTE	17
1.3. MARCOS DE REFERENCIA	21
1.3.1. Marco teórico	21
1.3.2. Marco conceptual	24
□ Ruido:	24
□ Cobot:	24
□ Internet de las cosas (internet of things IoT):	25
□ Esqueletización:	25
□ Chroma Key:	26
1.3.3. Marco legal y aspectos éticos	26
1.4. METODOLOGÍA	27
1.4.1. Línea de Investigación	27
1.4.2. Tipo de Investigación	27
1.4.3. Definición De La Metodología	27
2. DESARROLLO	30
2.1. Investigación, selección y pruebas de las técnicas de captura de movimiento de forma independiente	30
2.2. Combinación de diferentes técnicas de captura de movimiento, para el desarrollo del algoritmo de esqueletización	34
3. RESULTADOS	46
4. CONCLUSIONES	47
Trabajos futuros	48
ANEXOS	52
Bibliografía	49

Tabla de ilustración

Ilustración 1 Cámara web	16
Ilustración 2 Articulaciones del cuerpo humano utilizadas en captura de movimiento. Briseño Cerón, Domínguez Ramírez, & Saucedo Ugalde, 2012.	22
Ilustración 3 Erosión de imagen en tiempo real	32
Ilustración 4 Dilatación de imagen en tiempo real	32
Ilustración 5 Captura con ajuste de erosión y dilatación	33
Ilustración 6 Captura con sombra	33
Ilustración 7 Filtro negativo	34
Ilustración 8 Filtro HSV	34
Ilustración 9 Contorno de imagen en tiempo real	35
Ilustración 10 Entrega de datos de la capture en fps	35
Ilustración 11 Contorno de la imagen	37
Ilustración 12 Suavización de la imagen.....	37
Ilustración 13 Captura en tiempo real	38
Ilustración 14 Captura de contorno y marcado de puntos.....	38
Ilustración 15 Detección de puntos.....	39
Ilustración 16 Cámara con sensor infrarrojo	40
Ilustración 17 Perlas reflectivas para infrarrojo	41
Ilustración 18 CMU Graphics Lab Motion Capture Database	41
Ilustración 19 Método de Haar Cascade. Tomado de (online: [77])	42
Ilustración 20 Diagrama de flujo algoritmo de esqueletización	43
Ilustración 21 Captura y esqueletización de la extremidad superior	45
Ilustración 22 Inclinómetro con algoritmo.....	46

Tabla de figuras

Gráfico 1 Técnica de gradiente morfológico	31
Gráfico 2 Detección y seguimiento de objeto con Python y OpenCv	34

Agradecimientos

Primeramente, le damos gracias a Dios y a todas esas personas que nos brindaron su ánimo y apoyo, porque sin ello no habría sido posible este trabajo. También le queremos dar las gracias al tutor de este proyecto, Luis Fernando Murillo Fernández, por su apoyo y ayuda. Gracias por todos sus consejos a la hora de desarrollar este proyecto, como por su esfuerzo en corregir y perfeccionar el trabajo. En segundo lugar, dar las gracias a todos esos compañeros con los que hemos convivido durante estos años en la carrera de ingeniería de sistemas y que nos han enseñado cada uno algo que nos ha hecho mejorar tanto estudiantil como personalmente y por último y no menos importante, dar las gracias a nuestros familiares, que siempre están ahí y que sin ellos nada de esto sería posible.

RESUMEN

En este proyecto se presenta el desarrollo de un prototipo de captura de movimiento de la extremidad superior derecha del cuerpo humano, con herramientas de bajo costo; con el cual se puede detectar, marcar y realizar un trazo de las articulaciones universales del brazo, obteniendo un esqueleto que sigue el movimiento de la extremidad superior en tiempo real, como resultado se muestran las posiciones de esos puntos en un plano cartesiano, es decir, los valores de una articulación marcada en un eje (x, y) y el ángulo que estas forman respecto a la horizontal de dicha articulación.

Todo este proyecto va enmarcado a un objetivo específico de un macro-proyecto de la universidad del Sinú seccional Cartagena, con el cual se quiere desarrollar un algoritmo de fusión de datos para prototipo de sistema de captura de movimiento de miembro superior mediante vídeo y sensores inerciales en sistemas de seguridad para Robótica Colaborativa.

INTRODUCCIÓN

En la actualidad se observa un gran avance de la industria cinematográfica respecto a la producción de mejores animaciones y efectos de video. También se escucha hablar de robótica colaborativa e industria 4.0, dichas tecnologías están ligadas a la visión por computadora, la cual permite obtener información de los objetos que existen en el entorno, saber dónde se encuentran, si están en movimiento o no y que trayecto siguen; se puede determinar un camino libre de obstáculos, es decir, se logra una descripción de los objetos físicos que son captados por la cámara. Por esta razón la visión por computadora es el eje principal al que se apunta con el proyecto de desarrollo de un prototipo de captura de movimiento del cuerpo humano “extremidad superior” basado en video utilizando librerías de uso libre (open source) y dispositivos de bajo costo. Es importante resaltar que lo relacionado con la captura de movimiento del cuerpo humano mediante cámara, apunta al termino Mocap (Motion capture o captura del movimiento) que es usado para referirse a aquellas técnicas de grabación de movimiento del cuerpo humano, que captura datos espaciotemporales y representados digitalmente. Dentro de los procedimientos estandarizados para la captura detallada del movimiento del cuerpo humano encontramos los sistemas ópticos estacionarios y portables, los cuales consisten en una red de cámaras, usualmente infrarrojas, video cámaras, sensores y marcadores (algunos sistemas no requieren marcadores) que convierten la información real en datos digitales para su tratamiento en entornos confinados o laboratorio.

Cuando se avanza con la investigación referente a métodos y técnicas de captura de movimiento del cuerpo humano mediante video, se encuentran muchos métodos y técnicas que contribuyen al desarrollo de un algoritmo que permita la aprehensión de movimiento mediante cámara web y herramientas de bajo costo. La investigación respecto a las diferentes técnicas y algoritmos que existen para la captura de movimiento mediante cámara permiten adentrarse en el mundo de la cinematografía, televisión, videojuegos y robótica, puesto de muchas de estas industrias ya utilizan procedimientos estandarizados, que colaboran con el desarrollo de este proyecto

aportando nuevos conocimientos en el área de la investigación y desarrollo de tecnologías de visión artificial que actualmente tienen un costo elevado y que en Colombia son muy poco conocidas.

1. DISEÑO METODOLÓGICO

1.1. PLANTEAMIENTO DEL PROBLEMA

1.1.1. Descripción del problema

La robótica colaborativa se ha introducido a la industria, apuntando a la necesidad de las empresas (más específicamente pequeñas y medianas “pymes”), de ganar competitividad en el mercado y buscando optimizar los procesos de producción. La combinación de tecnologías como visión artificial, Big Data, reconstrucción 3D, realidad aumentada y otras, dan lugar a una nueva revolución industrial conocida como industria 4.0.

La industria 4.0 “se refiere a la denominada cuarta revolución industrial, impulsada por la transformación digital, y significa un salto cualitativo en la organización y gestión de la cadena de valor del sector.”, esto implica que la industria como la conocíamos se transformará a partir de la aplicación de las nuevas tecnologías, y hará obsoleta gran parte de los procesos productivos actuales [1].

“Dentro de la iniciativa de industria 4.0 se menciona un conjunto disperso de tecnologías necesarias para conseguir sus planteamientos. Todas ellas convergen en el internet de las cosas (Internet of Things), como eje fundamental, pero también se presentan una serie de tecnologías claves necesarias como el Big Data, la automatización industrial, la gestión del ciclo de vida de producto, los dispositivos inteligentes, la ciberseguridad y las tecnologías semánticas. Las tecnologías que se mencionan, en la mayoría de los casos, usan computación visual (Visual Computing) como medio integrador” [2].

En la robótica colaborativa los robots son caracterizados por ser ligeros, flexibles y fáciles de instalar, y han sido diseñados especialmente para interactuar con humanos en un espacio de trabajo compartido sin necesidad de instalar vallas de seguridad. Pero la idea de personas laborando junto a máquinas cuya fuerza en

ocasiones suele ser muy superior, trae consigo la consecuencia de riesgos en el espacio de trabajo, debido a la probabilidad de accidentes que puedan causar lesiones, muertes, pérdida de producción y tiempo.

Una solución a esto es lograr que las máquinas identifiquen la posición en el espacio de trabajo, de las personas; con el fin de poder anticiparse a los movimientos de la de ella, y de esta manera el dispositivo tome acciones que eviten una colisión o impacto, cuando exista la aproximación de una parte del cuerpo de la persona a la máquina. Es aquí, donde los sistemas de captura de movimiento ejercen un papel importante, y entre estos, los sistemas que utilizan visión artificial son de los más utilizados.

Estos sistemas son ofrecidos actualmente en el mercado, los cual constan de sensores, cámaras que integran con un software, y en general, hardware de muy altas especificaciones que son necesarios para la captura de movimiento, pero la mayoría de estos sistemas suele costar miles de dólares, lo cual significa una inversión inmensa si se quiere instalar a cada máquina. Es importante mencionar que los sistemas profesionales utilizan trajes especiales y fondos de colores, los cuales contrastan con la vestimenta, implicando así una inversión extra para que la maquina logre identificar al ser humano.

1.1.2. Justificación

La “Visión artificial” como un campo de la Inteligencia Artificial (IA), permite la obtención, el procesamiento y análisis de información de cualquier tipo, captada a través de imágenes digitales, es así como nacen los sistemas de captura de movimiento, capaces de detectar diversas características, tales como la distancia, posición con respecto al espacio, velocidad y dirección del movimiento y más, de una persona u objeto. En las aplicaciones de industria 4.0 y especialmente en la robótica colaborativa, la visión artificial permite implementar sistemas de seguridad, telemando y HMI (Human Machine Interface).

Estos sistemas son ofrecidos actualmente en el mercado, el cual consta de una cámara que integra el software y el hardware necesario para la captura de movimiento, pero la mayoría de estos sistemas suele costar miles de dólares, lo cual significa una inversión inmensa si se quiere instalar a cada máquina. Se observa como algunos sistemas para la captura de movimiento desarrollados y ofrecidos por diferentes empresas, entre los que se encuentran NANSENSE, la cual cuenta con una serie de sensores que permiten la captación de los movimientos realizados por la persona, dichos sensores tienen un costo que va desde los 1.000 dólares hasta 6.000 dólares [3]. También se encuentran algunos como lo es BIOMECH estos sistemas de captura de movimiento que tienen un costo que van desde los 76 millones de pesos hasta más de 300 millones de pesos colombianos. [4]

En 1999 surgió una biblioteca o librería de visión artificial conocida como OpenCv (Open Source Computer Vision) desarrollada por Intel [5], caracterizada por ser de código abierto y que en el transcurso de los años ha tenido una evolución sorprendente para múltiples aplicaciones como en sistemas de seguridad con detección de rostro, sistemas de seguimiento de objetos y otros sistemas que requieren de visión artificial.

Hoy en día, el desarrollo de sensores de bajo costo y de buenas prestaciones, así como la popularización de sistemas de desarrollo de SoC (System on Chip) empotrados de costo razonable, genera una tendencia a desarrollar sistemas de captura de movimiento utilizando este hardware de bajo costo, y aprovechando la disponibilidad de herramientas software de código abierto.

Con este proyecto se desarrolló un prototipo del sistema de captura de movimiento basado en video capaz de detectar el movimiento del cuerpo humano “extremidad superior”, utilizando recursos hardware de bajo costo, tales como sistemas SoC (sistemas en chip), cámaras web y PC (Computador Personal) y software libre, como lo es OpenCV con la finalidad de brindar apoyo a las pequeñas y medianas empresas de la industria, tanto en la implementación y pruebas de robótica

colaborativa, y así como el desarrollo de este proyecto brinda las bases para el desarrollo de un macro-proyecto enfocado en el desarrollo de un algoritmo de fusión de datos para prototipo de sistema de captura de movimiento de miembro superior mediante vídeo y sensores inerciales en sistemas de seguridad para Robótica Colaborativa. Se espera que en futuras fases del proyecto (Macroproyecto) se pueda implementar este sistema en la interacción de una persona con un robot, utilizando técnicas de robótica colaborativa.

1.1.3. Formulación del problema

¿Cómo desarrollar un prototipo que basado en video sea capaz de detectar el movimiento de una extremidad superior del cuerpo humano, utilizando recursos de hardware y tecnologías de bajo costo?

1.1.4. Alcance

El desarrollo de un prototipo de captura de movimiento “extremidad superior” del cuerpo humano basado en video y utilizando Python, librerías OpenCV y dispositivos de bajo costo, permitirá cimentar las bases del macro-proyecto, dando cumplimiento al primer objetivo específico con el desarrollo de un prototipo de captura de movimiento en un entorno controlado (laboratorio de electrónica de la Universidad del Sinú – Cartagena), dicho entorno dispone de un fondo color verde, una cámara web con resolución 1.3 MPX (ver Ilustración 1), la cual deberá estar ubicada aproximadamente a una distancia de 1.5 metros del objeto que se desea captar, con una inclinación no superior a 95° ni inferior a 85° y debe contar con un soporte fijo donde la cámara este ubicada verticalmente, también se dispone de un equipo (computador personal) al cual tiene cargado el algoritmo que realiza la captura de video, esqueletización de la extremidad superior y la obtención de los datos.



Ilustración 1 Cámara web

Cabe resaltar que este algoritmo no cuenta con una interfaz gráfica de usuario y la verificación de su funcionamiento se lleva a cabo mediante la interfaz ofrecida por la cámara web y las librerías OpenCv. La obtención y entrega de los datos está en un formato decimal y expresado en ángulos y coordenadas en (x, y) , para que en proyectos futuros se continúe con el desarrollo. Es importante mencionar que, para iniciar captura de la extremidad superior es necesario que el brazo este en una posición de martillo y con el puño totalmente cerrado, con los señaladores correctamente ubicados, que en este caso son círculos de colores en las 2 articulaciones universales del brazo (codo y hombro) y de colores azul y amarillo respectivamente, pero no hay que dejar a un lado que el brazo, la cabeza y toda la extremidad superior derecha, deben estar completamente dentro del fondo de color verde para evitar inconsistencia y distorsión en la captura del video.

1.1.5. Objetivos

Objetivo general

Desarrollar un prototipo de sistema de captura de movimiento de una extremidad superior del cuerpo humano basado en video, utilizando recursos de hardware y tecnologías de bajo costo.

Objetivos específicos

- Investigar y seleccionar las técnicas y/o algoritmos de captura de movimiento mediante video que más se adecuan a las necesidades del proyecto.
- Probar las técnicas de captura de movimiento previamente seleccionada, mediante un conjunto de pruebas independientes.
- Construir el algoritmo de esqueletización de la extremidad superior, usando la combinación de técnicas previamente seleccionadas.
- Realizar las pruebas necesarias y pertinentes para validar el correcto funcionamiento del prototipo en lo que respecta a captura de movimiento

1.2. ESTADO DEL ARTE

Los sistemas de captura de movimiento basados en video tienen un campo amplio en la industria del cine, robótica y video juegos, ya que en la actualidad se ha incrementado el interés y el uso de estos sistemas, debido a la disponibilidad de métodos de procesamiento de imágenes de buen prestigio, así mismo como el uso de sensores y cámaras de bajo costo. También se ha aumentado el número de proyectos enfocados a la visión por computador y el uso práctico de estos métodos en la vida cotidiana para el desarrollo e implementación de esta tecnología en robótica colaborativa, por lo que se mencionan los siguientes proyectos y estudios enfocados a la visión por computadora o visión artificial.

El estudio denominado: “Navegación de robot móvil usando Kinect, OpenCV y Arduino” [6], cuyo objetivo fue el desarrollo de un sistema de apoyo a la navegación de un robot móvil por medio de imágenes de profundidad y el reconocimiento de objetos por sus canales primarios. Permitted ver como se utiliza el sensor Kinect de Microsoft para realizar la captura de la imagen RGB y la imagen de profundidad,

este sensor consta de una cámara RGB y un emisor de infrarrojos que proyecta un patrón irregular de haces con una intensidad variable. Se logra ver como el sensor de profundidad reconstruye una imagen a partir de la distorsión del patrón, buscando todos los puntos rojos de la escena, contándolos, calculando el centroide, el diámetro, la posición y la distancia a la que se encuentra el objeto que se desea captar respecto a la ubicación del Kinect. Por último, se analiza como procesan los datos y como se realiza la toma de decisión de movimiento para ser enviada al Arduino, donde se controlan los motores. se estudian los resultados obtenidos en la investigación, los cuales indican que las imágenes de profundidad capturadas por el Kinect requieren de escenarios con iluminación controlada.

En el estudio denominado “Captura de movimiento utilizando el Kinect para el control de una plataforma robótica controlada de forma remota por medio de seguimiento de los puntos de articulación del cuerpo” [7]. La apertura de mercados y la visualización de una calidad de vida más amena y accesible a todos, hace que cada día crezcan las líneas de tecnología con miras a convertirse en aliadas del ser humano para toda actividad y más si esta coloca en riesgo la preservación de la especie humana como tal. Se analiza como en áreas donde se liberen gases, espacios confinados, terrenos minados y desastres naturales; se hace necesaria la interacción natural entre el hombre y la máquina; convirtiéndose el robot en una aplicación moderna para interactuar en diversas circunstancias con las personas. También se observa como a lo largo del tiempo se han desarrollado métodos de captura de movimiento, los cuales han sido empleados en áreas como la medicina, la educación, la seguridad entre otros y vemos como de esta manera se presenta en el mercado el sensor Kinect, dispositivo diseñado con tal propósito inicialmente, lo cual se fue diversificando a lo largo que fueron presentándose códigos abiertos para acceder y focalizar más aplicaciones con él. Vemos como en este proyecto se hace la integración del Kinect con una línea bastante robusta como es la robótica, utilizando el software ROS, haciendo uso de controladores y paquetes para obtener la esqueletización y visualización de joints. El paquete Skeleton_Marker que tiene como característica reconocer un usuario y realizar la esqueletización generando

un solo ID, también se hizo uso del controlador Nite para obtener la imagen de profundidad que garantiza la estabilidad de los joints. Se observa que una vez obtienen la imagen, pasan a la etapa de creación de un nuevo paquete para recibir la información generada por los puntos censados; para esta tarea implementan Transform Frame de ROS (TF), que hace relación a una transformación de los datos generados para ser enviados; en esta secuencia se utilizó los mensajes de ROS (listener) para enviar y recibir la información de un paquete a otro. y para que logaran una visión del comportamiento del joint se implementan por código la toma de 100 muestras de cada punto, vistas por terminal en los tres ejes coordenados; y 10 muestras reales las cuales se hicieron con un medidor laser, con el fin de relacionar las medidas generadas por el Kinect y un instrumento de medida. Por último, se analiza como con la obtención de todos los datos, realizan el tratamiento estadístico para presentar el grado de error a nivel de joints y global tomando el esqueleto.

En la tesis: “Desarrollo de una aplicación de detección de movimiento basada en comparativa estructural de imágenes” [8], se observa el desarrollo de una aplicación de detección de movimiento cuyo funcionamiento está basado en la comparativa estructural de imágenes, y que puede aplicarse a nuevos dispositivos móviles de pequeño tamaño. El método que se ha empleado para la detección de esta tesis consiste en la aplicación de una medida de calidad de imagen, la similitud estructural, que permite determinar las diferencias existentes entre dos imágenes. La creciente necesidad de seguridad ha propiciado el desarrollo en los últimos años de muy variados sistemas de vigilancia. Es importante resaltar todos los avances que se han llevado a cabo en el campo del procesado de señales para hacer más eficiente a los sistemas de seguridad para esta tesis.

El estudio referente a este proyecto permitió ver aplicación de detección de movimiento cuyo funcionamiento está basado en la comparativa estructural de imágenes, y que puede aplicarse a nuevos dispositivos móviles de pequeño tamaño. Se observa como aplicaron métodos para la detección, que consisten en

la aplicación de una medida de calidad de imagen, la similitud estructural, que permite determinar las diferencias existentes entre dos imágenes. Se revisa el análisis realizado a las diversas pruebas para estudiar el buen funcionamiento del método desarrollado para el proyecto mencionado.

En el Documento de grado “Diseño y desarrollo de un dispositivo objeto controlado a través del sensor Kinect y la plataforma Arduino orientado al uso ludico-pedagógico de niños en la primera infancia” [9], el estudio de este proyecto, nos permitió observar como se hizo el diseño y desarrollo de un títere que detecta los movimientos corporales de las extremidades superiores e inferiores de las personas que se ubiquen en frente de él controlado a través de un sensor Kinect y la placa Arduino orientado en el uso lúdico - pedagógico de niños de la primera infancia. Se han analizado las diferentes teorías del aprendizaje que intentan dotar la enseñanza de aspectos útiles con el objetivo de tener una mejor perspectiva. Las teorías más influyentes estudiadas han sido el conductismo, el cognitivismo, el constructivismo y el conectivismo. Se estudia el sistema de capas que, con los planes de estudio actuales, partiendo desde primaria hasta los ciclos formativos de grado superior, permita emplear la robótica en todos los niveles educativos. Con estas premisas y enfocado en el construccionismo, una variante del constructivismo se el desarrollador de este proyecto logra estudiar el comportamiento tanto de la plataforma robótica, como la electrónica y programación a partir de unos patrones usando material asequible y normalizado. También se analiza el uso de los patrones tiene como objetivo facilitar la versatilidad, el modularidad, la interconexión entre módulos y reducir las variables que intervienen en el sistema. Se estudia la implementación de la tarea principal, que se encarga de detectar y seguir una línea negra sobre fondo blanco de modo autónomo. Cabe resaltar que el sistema microcontrolador implementado en el proyecto mencionado puede utilizarse en otras aplicaciones, aunque no sean robóticas, para conseguir el aprendizaje necesario en la tecnología de los microcontroladores y en su programación.

Con el desarrollo de un prototipo de captura de movimiento de una extremidad superior del cuerpo humano y utilizando herramientas de bajo costo, se obtiene la esqueletización de una extremidad superior con los ángulos de las articulaciones detectadas y su posicionamiento en un plano cartesiano (x, y), que permitiera realizar la predicción del movimiento y con capacidad de trabajar en conjunto con la robótica colaborativa, para evitar colisiones y accidentes de trabajo entre máquina y hombre en el macro-proyecto de algoritmo de fusión de datos para prototipo de sistema de captura de movimiento.

1.3. MARCOS DE REFERENCIA

1.3.1. Marco teórico

A través del tiempo, el campo de robótica enfoca precisamente a lo que uno conoce como Inteligencia artificial (IA) Controles difusos, Cobots (Robot Colaborativo), Domótica, IoT (internet of things), Sistemas Embebidos, Sistemas de captura, entre otros. Con base a esto, se busca como soporte de acercamiento en nuestro proyecto señalar los conceptos y estrategias respecto a nuestro tema de investigación.

De acuerdo con la revista internacional de tecnología, ciencia y sociedad: el sistema óseo del ser humano es complejo y para realizar el análisis, modelado y representación de los movimientos es una tarea ardua, sin embargo, con el desarrollo de las tecnologías de la información que permiten capturar las trayectorias de movimiento, ahorrando tiempo y esfuerzo para el diseño y desarrollo de herramientas de apoyo para áreas como la fisioterapia.

Captura de movimiento: Los sistemas de captura de movimiento son, por definición, sistemas que generan para un computador información que representan medidas físicas de movimiento capturado [10]. Consiste en vestir a un actor con un traje especial donde serán posicionados reflectores (Sistemas Ópticos) o transmisores (Sistemas Magnéticos) o básicamente sincronizar los movimientos

físicos de algún miembro. Estos marcadores son generalmente posicionados en las llamadas “Articulaciones Universales”, que son 19 posiciones que ofrecen un mínimo de precisión para representación de un movimiento humano. En general los sistemas de captura tienen una estructura como muestra la Ilustración 2.

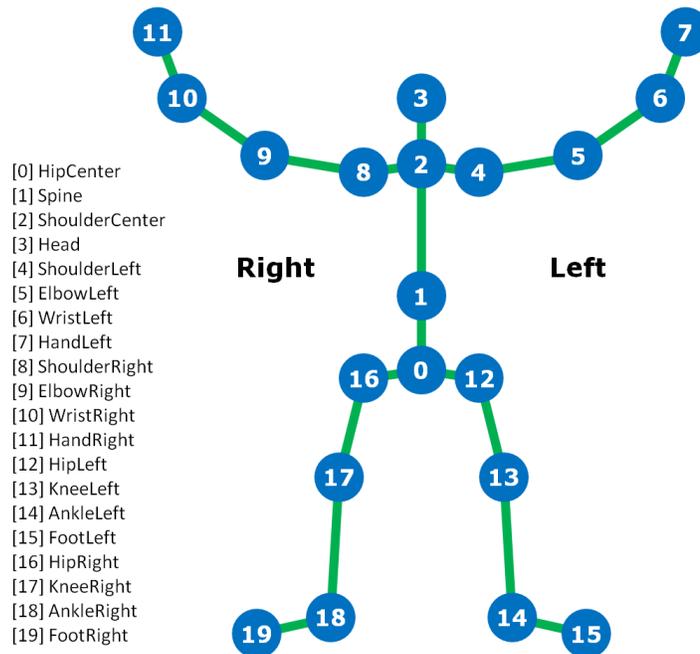


Ilustración 2 Articulaciones del cuerpo humano utilizadas en captura de movimiento. Briseño Cerón, Domínguez Ramírez, & Saucedo Ugalde, 2012.

Visión Artificial: La visión artificial abarca en todas las aplicaciones industriales y no industriales en las que una combinación de hardware y software brinda un guiado operativo a los dispositivos en la ejecución de sus funciones de acuerdo con la captación y procesamiento de imágenes. cuentan con sensores digitales protegidos en el interior de cámaras industriales con ópticas especializadas para adquirir imágenes, de forma que el hardware y software informático pueden procesar, analizar y medir diversas características a la hora de tomar decisiones [11].

La visión artificial la componen un conjunto de procesos destinados para realizar el análisis de imágenes que permita la obtención, procesamiento y análisis de cualquier tipo de información especial obtenida a través de imágenes digitales.

Estos procesos son: captación de imágenes, memorización de la información, procesado e interpretación de resultados. También tiene aplicaciones en la industria actual como:

- Identificación e inspección de objetos.
- Determinación de la posición de objetos en el espacio.
- Mediciones Tridimensionales.
- Establecimiento de relaciones espaciales entre varios objetos (Guiado de robots)

Identificación: Un sistema de reconocimiento óptico de caracteres (OCR) lee caracteres alfanuméricos sin conocimiento previo, mientras que un sistema de comprobación óptica de caracteres (OCV) confirma la presencia de una cadena de caracteres [12]. Además, los sistemas de visión artificial pueden identificar piezas mediante la localización de un patrón único o identificar elementos según su color, forma o tamaño. Estas aplicaciones como DPM marcan un código o una cadena de caracteres directamente en la pieza. Fabricantes de todas las industrias suelen usar esta técnica para detectar errores, habilitar estrategias eficaces de contención y supervisar el control de procesos y las métricas de control de calidad.

Librerías OpenCV: Se lanza bajo una licencia BSD, es gratis para los académicos y para uso comercial. La licencia BSD permite el uso exclusivo, y para el software lanzado bajo la licencia para ser incorporado en productos propietarios. Los trabajos basados en el material pueden ser publicados bajo una licencia de propiedad o como software de código cerrado. Esta es la razón para el uso generalizado de la Código BSD en productos propietarios, que van desde Juniper hasta redes enrutadores para Mac OS X [13].

Opencv es una colección de funciones C y algunas clases de C ++, que implementar un poco de procesamiento de imágenes populares y algoritmos de visión por computadora y cuenta con funciones gratis para usos comerciales y no comerciales.

- cv: Principales funciones de OpenCV, realiza el procesamiento algoritmos de imágenes y visión.
- cvaux: Funciones auxiliares (experimentales) de OpenCV.
- cxcore: estructuras de datos, soporte de álgebra lineal, soporte XML, funciones de dibujo y otros algoritmos.
- highgui: funciones GUI, imagen y video I/O

Método de captura de movimiento: En la actualidad son numerosos los sistemas de captura de movimientos, y las tecnologías relacionadas con el mocap, las cuales consisten en el análisis y grabación de los datos de un movimiento. En las sesiones de captura lo que se analiza son las posiciones, las fuerzas, las velocidades y los impulsos de los movimientos del actor, tomando como referencia las variables sacadas en un tiempo determinado (por ejemplo, tomando datos cada 2 Ms). Para la captura hay diferentes equipos basados en tecnologías diferentes, las más conocidas son las ópticas de marcado, que se reconocen por usar deflectores en puntos estratégicos del cuerpo y las inerciales, como el mando Wiimote de las videoconsolas Wii que constituye un ejemplo muy básico de mocap de este tipo [14].

1.3.2. Marco conceptual

- **Ruido:** Usualmente se encuentran dos tipos de ruido luminancia, causado por la falta de luz y el ruido de color, causado por el calentamiento de sensores. El ruido aparece cuando el color y la luz de algunos de los pixeles se ven alterados [15].
- **Cobot:** Un "cobot" es un dispositivo robótico que manipula objetos en colaboración con un operador humano [16]. Un cobot proporciona asistencia al operador humano mediante la configuración de superficies virtuales que

se pueden utilizar para restringir y guiar el movimiento. Si bien las pantallas hápticas servoactuadas convencionales también se pueden usar de esta manera, una distinción importante es que, mientras que las pantallas hápticas son dispositivos activos que pueden suministrar energía al operador humano, los cobots son intrínsecamente pasivos. Esto se debe a que los cobots no usan servos para implementar restricciones, sino que emplean juntas no holonómicas "dirigibles". Como consecuencia de su pasividad, los cobots están potencialmente bien equipados para tareas críticas para la seguridad (por ejemplo, cirugía) o aquellas que involucran grandes fuerzas de interacción (por ejemplo, ensamblaje de automóviles).

- **Internet de las cosas (internet of things IoT):** El Internet de las cosas está impulsado por una expansión de Internet a través de la inclusión de objetos físicos combinados con la capacidad de proporcionar servicios más inteligentes al entorno a medida que haya más datos disponibles. Varios dominios de aplicaciones que van desde Green-IT y eficiencia energética a logística ya están empezando a beneficiarse de los conceptos de Internet de las cosas. Existen desafíos asociados con la Internet de las cosas, más explícitamente en las áreas de confianza y seguridad, la estandarización y la gobernabilidad requeridas para garantizar una Internet abierta y justa y confiable de las cosas que proporcione valor a toda la sociedad. El Internet de las cosas ocupa un lugar destacado en la agenda de investigación de varias multinacionales, así como de la Comisión Europea y países como China [17].
- **Esqueletización:** La esqueletización busca representar una imagen binarizada (formada por pixels blancos y negros) por un grafo de grosor fino (típicamente 1 pixel de grosor) cuyos puntos cumplen que la distancia local respecto a los bordes de la imagen binarizada es máxima. Los algoritmos de esqueletización se basan en la ejecución de un conjunto de iteraciones,

donde en cada una se realiza el borrado de los pixels pertenecientes a los bordes de la imagen, hasta que solamente queda el esqueleto [18].

- **Chroma Key:** Es un método para el recorte de una filmación, donde el fondo es de color azul o verde, se utilizan desde el inicio de la Televisión en aplicaciones simples hasta los sofisticados sistemas de múltiples capas, básicamente consisten en insertar letras o imágenes en un vídeo, reemplazando un color por el video a insertar [19].

1.3.3. Marco legal y aspectos éticos

Las pruebas son aplicadas en los autores principales de este proyecto, dando consentimiento y aprobación para la utilización de marcadores en las articulaciones, los cuales son círculos de foamy y no son invasivos para el cuerpo humano, ni generan ningún tipo de repercusión. Es importante mencionar que este proyecto no contamina el ambiente ya que no genera desechos tóxicos, ni es necesario utilizar materia prima donde este se vea afectado.

El desarrollo del algoritmo para la captura de movimiento está enmarcado a la norma ISO / TS 15066: 2016, que especifica los requisitos de seguridad para los sistemas de robot industrial de colaboración y el entorno de trabajo, y complementa los requisitos y la guía sobre la operación de robot industrial de colaboración que figuran en ISO 10218-1 (Robots y dispositivos robóticos. Requisitos de seguridad para robots industriales - parte 1) e ISO 10218-2 (Robots y dispositivos robóticos. Requisitos de seguridad para robots industriales - parte 2). Se aplica a los sistemas de robots industriales como se describe en ISO 10218-1 e ISO 10218-2. No se aplica a robots no industriales, aunque los principios de seguridad presentados pueden ser útiles para otras áreas de la robótica [20].

1.4. METODOLOGÍA

1.4.1. Línea de Investigación

El proyecto se enfoca en dos líneas de investigación de la escuela de ingeniería de sistemas de la universidad del Sinú Seccional Cartagena, las cuales son: la línea automatización y robótica y la línea de inteligencia artificial del semillero de investigación DEARTICA.

1.4.2. Tipo de Investigación

Este proyecto esta enfocados a investigación aplicada, dado que cuenta con una problemática establecida, pero es necesario realizar una investigación para darle solución. Por este motivo con la adquisición de nuevos conocimientos respecto a la programación y desarrollo de algoritmos en Python y utilizando la librería OpenCv para la captura de movimiento, utilizando técnicas de visión por computador y demás temas que se desprenden, se logra establecer la esqueletización de la extremidad superior del cuerpo humano basado en video.

1.4.3. Definición De La Metodología

Este proyecto, está basado en el modelo de Martin y McClure y del SEI [21], que consiste en tres grandes fases:

- Investigación y desarrollo inicial
- Investigación aplicada
- Transferencia.

Cada una de ellas involucrando el desarrollo de varios proyectos de investigación, posiblemente cada uno de ellos con métodos y técnicas diferentes.



Ilustración 3 Esquema de la estrategia de investigación (Tomado de: <http://web.imt-atlantique.fr/x-info/harbol07/MIFISIS2004.pdf>)

A. Investigación y desarrollo inicial

Es necesario conocer a fondo los métodos, funciones y técnicas con los que dispone OpenCV para la captura de movimiento. Por esta razón, es necesario dirigirse a la página oficial de OpenCV y consultar todas las herramientas que brinda esta librería. Primero nos encontramos con las técnicas de filtrado, dilatación, erosión y contorno, revisando su documentación y que brinda cada una de ellas, también se encontraron métodos de espacio de color, como lo son el filtro negativo y HSV; cabe resaltar que cada uno de las técnicas y métodos mencionados anteriormente fueron probados de forma independiente, con la finalidad de llevar a cabo el desarrollo de un algoritmo que permita la captura de tres articulaciones universales de la extremidad superior, señalizándolas y siguiendo su movimiento y así mismo realizar la esqueletización de la extremidad.

B. Investigación aplicada

Una vez interiorizados los conceptos fundamentales de la visión por computador, se realizan pruebas con diferentes algoritmos, dentro de los cuales se destacan la técnica de contorno, detección de objetos, reconocimiento facial, el método de Lucas kanade y reconocimiento de puño.

Posteriormente se procede al desarrollo del algoritmo final, el cual es capaz de marcar la extremidad superior y sus articulaciones en forma de esqueleto. Para ello es necesario un entorno controlado (en este caso se utiliza el laboratorio de electrónica de la universidad del Sinú seccional Cartagena), el cual dispone con un chroma key, que no es más que un fondo de color verde, buena luminosidad y cámara(s). El algoritmo empieza importando las librerías numpy, cv2 y math, luego se implementa el método de Haar cascade para detección de objetos y rostro, se usa el método Lucas Kanade que sigue el flujo de movimiento de un objeto marcado, se inicia el ciclo de la captura de frame () por frame en un ciclo infinito para la detección de las articulaciones del brazo derecho, en este caso la articulación del hombro será marcada con un círculo de color amarillo y la del codo con uno de color azul, se realiza la detección de esos colores respectivamente y se traza una línea para simular el esqueleto y se aplica la arco-tangente para calcular los ángulos de mano a codo y de codo a hombro, todo esto respecto a la horizontal.

C. Transferencia

Se realiza la sustentación del proyecto en la Universidad del Sinú seccional Cartagena por parte de los autores principales, con la finalidad de ser evaluado y aprobado como proyecto de grado para la obtención del título de ingeniero de sistemas y se entregara el documento, junto a un artículo científico referente al proyecto.

2. DESARROLLO

En este capítulo se describe el desarrollo del prototipo y todas sus fases, también se describen las técnicas, métodos y algoritmos que permiten lograr el propósito de este trabajo.

2.1. Investigación, selección y pruebas de las técnicas de captura de movimiento de forma independiente

Para iniciar con el desarrollo de este trabajo, es necesario realizar la instalación de Python 3, un editor de código “sublime text”, las librerías necesarias (OpenCv, numpy y maht) y el complemento de Python para la ejecución de algoritmos desde el editor de código; también se hace impredecible investigar acerca de OpenCV, el cual es una librería de visión por computador de código abierto, que ha sido diseñado para ser eficiente en cuanto a gasto de recursos computacionales y con un enfoque hacia las aplicaciones de tiempo real. Uno de los objetivos de OpenCV es proveer una infraestructura de visión por computador fácil de utilizar que ayuda a los programadores a desarrollar aplicaciones “sofisticadas” de visión por computadora (Computer Vision “CV”).

Es necesario investigar a fondo cuales son las técnicas que usa OpenCV y la definición correspondiente al funcionamiento de cada una de ellas. Primeramente, nos encontramos con la técnica de filtrado, que implementa métodos para resaltar o suprimir, de forma selectiva, información contenida en una imagen a diferentes escalas espaciales, para destacar algunos elementos de la imagen, o también para ocultar valores anómalos. El proceso de filtrado consiste en la aplicación a cada uno de los pixeles de la imagen de una matriz de filtrado de tamaño NxN (generalmente de 3x3 aunque puede ser mayor) compuesta por números enteros y que genera un nuevo valor mediante una función del valor original y los de los pixeles circundantes. El resultado final se divide entre un escalar, generalmente la suma de los coeficientes de ponderación. Los filtros se pueden expresar mediante la siguiente ecuación [22].

$$P_{smooth} = \frac{P1 + P2 + P1 + P2 + P1 + P2 + P1 + P2 + P_{centro}}{9}$$

Donde p es igual a pixel, y (x, y) son los ejes. El Kernel entonces queda definido como:

$$P_{smooth}(x,y) = \begin{bmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{bmatrix}$$

$$P_{smooth}(x,y) = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Como primera instancia para el desarrollo del algoritmo de captura de movimiento del cuerpo humano (extremidad superior), se empieza con la aplicación de dos técnicas que permiten captar el objeto que se desea visualizar. La técnica de erosión, que erosiona los límites del objeto de primer plano (siempre tratando de mantener el primer plano en blanco) y la técnica de dilatación que aumenta la región blanca en la imagen o el tamaño del objeto en primer plano aumenta. Para la obtención de una mejor captura del objeto se utiliza la técnica gradiente morfológico (Morphological Gradient), que no es más que la diferencia entre la erosión y la dilatación, y que ayuda a encontrar la silueta de los objetos o el contorno. La implementación de esta técnica esta descrita por el siguiente diagrama:



Gráfico 1 Técnica de gradiente morfológico (fuente propia)

El desarrollo de un algoritmo que utilice las técnicas de erosión y dilatación arroja los siguientes resultados de forma independiente:

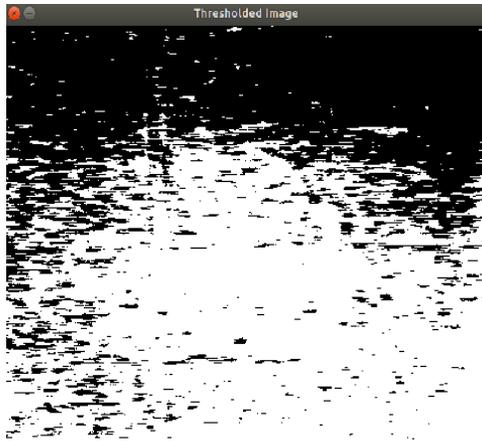


Ilustración 3 Erosión de imagen en tiempo real

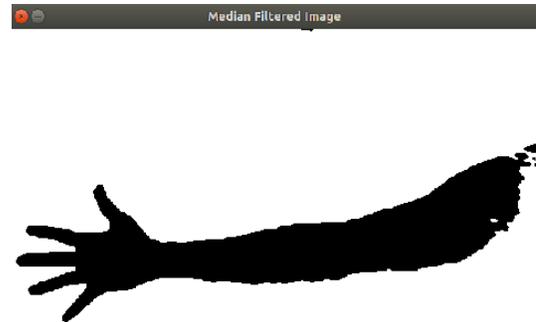


Ilustración 4 Dilatación de imagen en tiempo real

Se puede observar en la Ilustración 3, una gran cantidad de ruido, evitando la captura de una buena imagen, esto se debe a las condiciones del momento, es decir, al realizar una detección sin objeto previo dentro de un primer plano, no se muestra la imagen correcta del objeto que se desea detectar, mientras que en la Ilustración 4, al tener un objeto en el plano, muestra la aplicación correcta de esta técnica.

Para tener un mejor control del ruido presente en la captura, se investiga como mejorar el algoritmo y la técnica implementada; por ello se decide utilizar Chroma key o mal llamado fondo verde, que es una técnica comúnmente usada en la industria cinematográfica, de la televisión y los videojuegos para extraer el color de una imagen o video y reemplazarla en el área que ocupaba ese color por otra imagen o video. Se usa el color verde, puesto que no tiene incidencia en la piel, es decir, que es muy poco probable que la cámara relacione este color con alguna parte de nuestro cuerpo. Con el fondo ya en su lugar, se pasa a probar un algoritmo que permite calibrar valores de la dilatación, evitando así el ruido y realizando una mejor detección del objeto en primer plano. Pero se observa que aun así la técnica suele fallar y a la hora de realizar movimientos de la extremidad genera sombras

como se ve en el Ilustración 6 y al tener presente este inconveniente, se piensa en alternativas para solucionarlo.

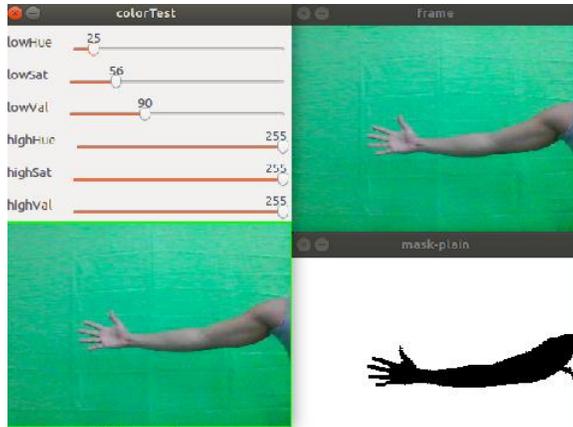


Ilustración 5 Captura con ajuste de erosión y dilatación



Ilustración 6 Captura con sombra

A medida que se avanza con la investigación, se decide usar el filtro negativo, que cumple la función de invertir los colores buscando generar un mejor contraste, pero al aplicar esta técnica se observó que es menos eficiente a la hora de realizar la detección de contorno y la sombra se hace mucho más evidente. Ver Ilustración 7. Por esta razón se descartó y se continuó trabajando e investigando acerca de una técnica más eficiente, dentro de la cual se encontró el espacio de color HSV que consta de los siguientes tres componentes:

- H - Tono (longitud de onda dominante).
- S - Saturación (Pureza / tonos del color).
- V - Valor (Intensidad).

La separación de componentes de color y de la intensidad es fundamental para la detección de objetos, por esta razón se tienen en cuenta HSV, que es uno de los muchos espacios de color que separan el color de la intensidad.

La detección y seguimiento de objeto es una técnica que permite la obtención de información de un objeto de color presente en una imagen o video en tiempo real y

para realizar este proceso, es necesario empezar con la detección natural de la cámara (captura imagen en RGB) y luego se convierte a HSV, para ello se utiliza la siguiente línea de código `cv2.COLOR_BGR2HSV`. Es importante mencionar que para HSV el rango de tono es $[0,179]$, el rango de saturación es $[0,255]$ y el rango de valor es $[0,255]$, la detección y seguimiento del objeto de la siguiente manera:



Gráfico 2 Detección y seguimiento de objeto con Python y OpenCv



Ilustración 7 Filtro negativo

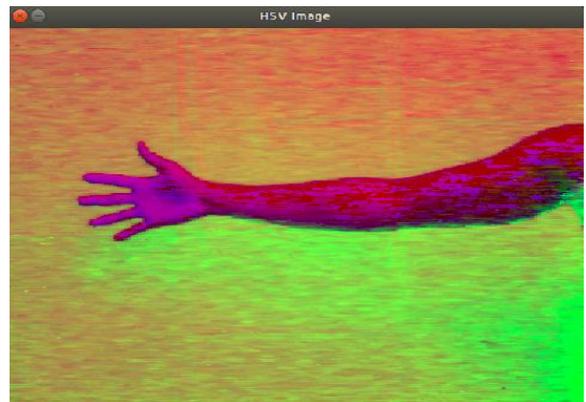


Ilustración 8 Filtro HSV

2.2. Combinación de diferentes técnicas de captura de movimiento, para el desarrollo del algoritmo de esqueletización.

Con el uso de las técnicas de erosión, dilatación y Morphological Gradient mencionadas anteriormente y la aplicación de Chroma key, se procederá con el desarrollo de un algoritmo para el cálculo de contorno, con el cual se espera llegar a la esqueletización de la extremidad superior. La ejecución de este algoritmo permitirá obtener los valores del contorno y representarlos, dichos valores son entregados por la terminación de ejecución en formato de FPS, estos datos se verán modificados o alterados con el movimiento de la extremidad superior dentro del

fondo. La aplicación de esta técnica resulto ser limitante para la esqueletización, puesto que hay poca documentación respecto a cómo contraer al máximo el contorno, es por esta razón que se decide implementar una técnica de marcado de puntos y contorno.



Ilustración 9 Contorno de imagen en tiempo real

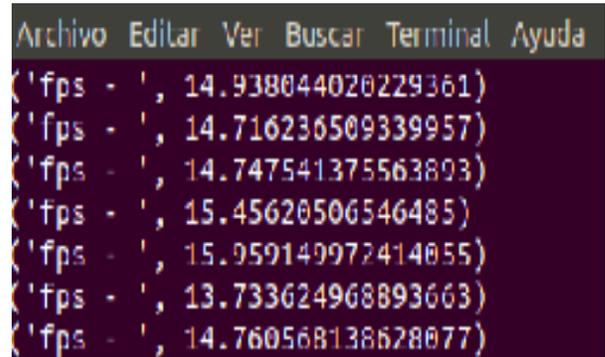


Ilustración 10 Entrega de datos de la capture en fps

Se procede al desarrollo del algoritmo capaz de detectar un objeto de color azul y para ello, fue necesario seguir los siguientes pasos: importar las librerías de openCV y numpy, iniciar la captura de video mediante cámara web, implementar un condicional que tome cada cuadro de la captura y los convierta de RGB a HSV, definir el rango de color en que detectara, en este caso sera el color azul, establecer el umbral de imagen HSV para obtener solo los colores azules y finalmente se enmasacara la imagen o video para mostrar el color establecido e imprimir los resultado por pantalla en tiempo real.

Se continúa trabajando en un algoritmo que permita detectar y realizar contornos de un imagen o video en tiempo real, para llevar a cabo el proceso se usa la técnica de umbral simple, que es un grupo de algoritmo cuya finalidad es segmentar en gráficos rasterizados o imagen de mapa de bits, es decir, separa los objetos de una imagen que nos interese del resto. Cuando el valor del pixel es mayor que el umbral se le asigna un valor de 100, que es el que permite resaltar con mayor precisión el brazo. Cabe mencionar que esta técnica arroja mejores resultados si se realiza con

fondos en la escala de grises. Lo primero que realiza este método es captar la imagen de video mediante la cámara web dimensionando la ventana a un ancho de 200px y un alto de 10px y posteriormente se le asigna un nombre a la ventana. Para llevar a cabo el umbral simple, se captura la imagen y se pasa a una escala de gris, con la finalidad de obtener una mejor captura en binario de los fotogramas y realizar el contorno de forma eficiente. El siguiente fragmento de código realiza el proceso de entrega de la imagen en binario.

```
camera = cv2.VideoCapture(1)
camera.set(10,200)
cv2.namedWindow('trackbar')
cv2.createTrackbar('trh1', 'trackbar', threshold, 100, printThreshold)
```

Se busca mejorar el algoritmo de contorno, para lo cual se toma la imagen original y se le aplica la técnica de suavizado de imagen, dentro de la cual se encuentran los siguientes métodos: Averaging, Gaussian Blurring, Median Blurring. Para llevar a cabo este proceso se realizaron pruebas para escoger la que mejor se adapte a las necesidades de este proyecto.

- **Averaging:** Esta técnica gira la imagen con un filtro de caja normalizado y toma el promedio de todos los píxeles en el área del Kernel reemplazando el elemento central. Cabe resaltar que es necesario especificar el ancho y la altura del Kernel.
- **Gaussian Blurring:** Esta técnica en lugar del filtro de caja, utiliza el Kernel gaussiano e implementa la función `cv2.GaussianBlur()`, con la cual se especifica el ancho y la altura del Kernel, que deben ser positivos e impar. También se debe especificar la desviación estándar en la dirección x, y, w, `sigmaX` y `sigmaY` respectivamente, si solo se especifica `sigmaX` y `sigmaY` se toma igual que el valor de `sigmaX`, ambos se dejan como ceros y se calcula a partir del tamaño del Kernel.
- **Median Blurring:** Esta técnica toma la mediana de todos los píxeles en el área del núcleo y el elemento central se reemplaza con el valor de la

mediana, en el caso de los filtros anterior el elemento central es un valor recién calculado que puede ser un valor de pixel en la imagen o un nuevo valor.



Ilustración 11 Contorno de la imagen



Ilustración 12 Suavización de la imagen

La implementación de estas técnicas permitió observar y analizar cuáles son las ventajas y desventajas que ofrecen cada una de ellas, con la finalidad de utilizar la que más se acople al proyecto. La técnica Averaging es útil para eliminar ruidos en los bordes de la imagen, pero esta operación difumina los difumina muy poco, la técnica Median Blurring, realiza un desenfoque para eliminar el ruido de sal y pimienta de la imagen, mientras que la técnica Gaussian Blurring arroja mejores resultados en comparación de las anteriores, ya que al implantar el filtro a la imagen original se aprecia que esta suavización y mejora el contorno de la extremidad superior.

El algoritmo de marcado de puntos y contorno resulta ser bastante eficiente para la captura de una mano, puesto que integra todas las técnicas y métodos mencionados anteriormente, pero al momento de ingresar toda la extremidad, presenta limitaciones, puesto que no identifica la extremidad muy bien y la señalación de puntos es muy difusa. En la Ilustración 13 se observa la detección de

la imagen en tiempo real y en la Ilustración 14 la captura ya procesada con el contorno y los puntos marcados.

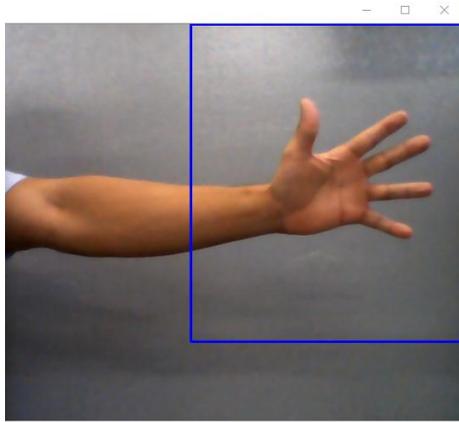


Ilustración 13 Captura en tiempo real

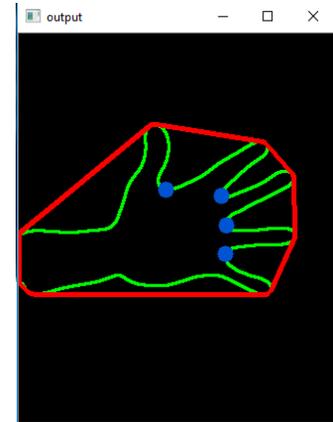


Ilustración 14 Captura de contorno y marcado de puntos

Al ver que la técnica de marcado de puntos en los errores presentados en la Ilustración 14 fue eficiente, se decide desarrollar un algoritmo que sea capaz de detectar un objeto y seguir su movimiento. Para el desarrollo de este algoritmo se empieza realización la captura del video en tiempo real, procesando los fotogramas y convirtiendo la imagen de RGB a HSV, se establece el rango de colores en el cual va a detectar la cámara, es decir, se dan los valores corresponden al rango de colores que se desea calcular para detectar el objeto, se realiza el contorno al objeto de color establecido que aparezca en la imagen, se crea un condicional para ubicar el punto central del objeto que se está detectando y seguir su movimiento. se hace entrega de los datos que se están calculando, en este caso los pixeles en los que se está moviendo el objeto, arrojando por pantalla el video en tiempo real, la marcación de puntos en el objeto y el seguimiento del mismo.

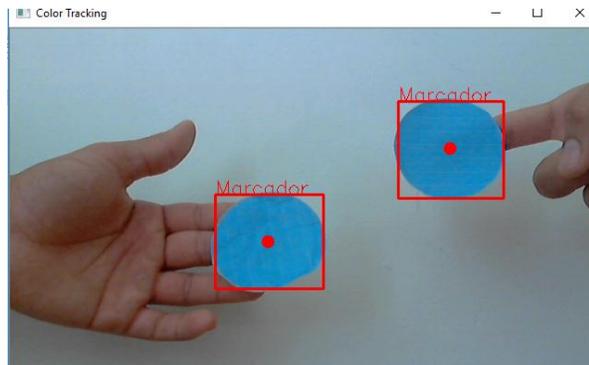


Ilustración 15 Detección de puntos

Para la optimización del algoritmo de marcado de puntos, se buscan alternativas para la detección del contorno de la imagen y para ello se utiliza una cámara con infrarrojo, que usa luz infrarroja, imperceptible para el usuario, para iluminar unos marcadores de material retro-reflectante. Al trabajar con este tipo de luz, las cámaras infrarrojas sólo observan los marcadores. Esto supone una ventaja frente a los sistemas basados en vídeo, puesto que aquí no es necesario aplicar técnicas de visión por computador para diferenciar los marcadores del resto de la imagen, alcanzando velocidades superiores a 90 capturas por segundo. Puesto que la cámara con la que se cuanta actualmente no ve en infrarrojo, se decide adaptarla para lo cual se realiza un raspado del componente para ver infrarrojo y se coloca un sensor infrarrojo, una vez colocado el infrarrojo en la cámara (ver Ilustración 16) se procede a realizar pruebas en el laboratorio de electrónica de la universidad del Sinú seccional Cartagena, donde se hace necesario que este con todas las luces apagadas, puesto que el sensor detecta en ambientes totalmente oscuros, dado que al adaptar la cámara queda más sensible a las luz, debe calibrarse el enfoque y para ello se empieza detectando un objeto cualquiera a una distancia de 70 cm y posteriormente la extremidad es enfocada correctamente a una distancia máxima de un metro.

Una vez se logra enfocar la extremidad superior, se realizan las pruebas para corroborar el correcto funcionamiento del sensor. Se inició apagando las luces del laboratorio, para que este estuviese totalmente oscuro como es recomendado,

obteniendo excelentes resultados por parte del sensor infrarrojo, pero muy pocos por parte de la cámara, también se probó enciéndiendo y apagando algunas luces en la parte más alejada del área de trabajo, viendo mejores resultados en cuanto a cámara, pero un poco más bajos con el sensor. Cabe resaltar que estas pruebas se hicieron sin ningún tipo de algoritmo cargado.



Ilustración 16 Cámara con sensor infrarrojo

Luego se continúa trabajando con perlas que son comúnmente usadas en vestidos de ropa común, estas perlas se asemejan a los sensores reflectivos utilizadas en los trajes de cine, que son sensible a la luz infrarroja, estas perlas se adaptan a un suéter para intentar similar la silueta que se usa en la industria del cine con marcadores en las articulaciones universales para hacer la esqueletización del cuerpo humano. Los vestidos profesionales para grabar películas en cine se asemejan a los mostrados en la Ilustración 18, los cuales permiten realizar una esqueletización bastante detalla.



Ilustración 17 Perlas reflectivas para infrarrojo



Ilustración 18 CMU Graphics Lab Motion Capture Database

Se usan las perlas junto con el sensor infrarrojo y se llega a la conclusión de que no son lo suficientemente reflectivas para poder ser detectadas con el sensor, y esto hace que se pierda fácilmente en el umbral, y a la hora de realizar la detección no sirven como punto de referencia de las articulaciones universales de la extremidad superior, por ende, no se pueden trazar las líneas que formarían el esqueleto del brazo y hallar los ángulos y coordenadas en el plano cartesiano.

Se procede a utilizar el método Gaussian Blurring con el sensor infrarrojo, pero se observa que no se obtienen resultados eficientes, puesto que los colores en los que ve este sensor son diferentes a los que se ve una cámara normal (RGB), por esto se decide ajustar el algoritmo para captar en infrarrojo, pero no es eficiente, puesto que genera una sombra y no deja realizar la detección en debida forma de la extremidad, por esto es dejado a un lado, pero sin olvidar las técnicas usadas anteriormente.

Se llega a un punto clave de la investigación, donde se cuenta con un algoritmo que cumple muchas de las funciones y métodos para la captura de movimiento, el cual es desarrollado por Ronald Torres, Amaury Arnedo y Sergio Puello como proyecto para obtener el grado de la corporación universitaria Rafael Núñez y que será implementado en este proyecto junto con su set de datos para la captura facial. Se consulta y prueba, el funcionamiento y es necesario un entorno controlado (en este caso se utiliza el laboratorio de electrónica de la universidad del Sinú seccional

Cartagena), el cual dispone con un chroma key, que no es más que un fondo de color verde, buena luminosidad y cámara(s). El algoritmo empieza importando las librerías numpy, cv2 y math, luego se implementa el método de Haar cascade para detección de objetos y rostro, el cual permite la detección de rostros, caras y ojos. Este es un método se basa en la detección de objetos mediante clasificadores en cascada basados en funciones de Haar. Este método usa aprendizaje automático en función cascada y se entrena a partir de muchas imágenes positivas (imágenes de caras) y negativas (imágenes sin caras). Este algoritmo extrae los rasgos como se puede apreciar en la Ilustración 19, y en él nos basamos para hallar el punto o articulación principal (ver Ilustración 20), se usa el método Lucas Kanade que sigue el flujo de movimiento de un objeto marcado, se inicia el ciclo de la captura de frame por frame en un ciclo infinito para la detección de las articulaciones del brazo derecho, en este caso la articulación del hombro será marcada con un círculo de color amarillo y la del codo con uno de color azul, se realiza la detección de esos colores respectivamente y se traza una línea para simular el esqueleto.

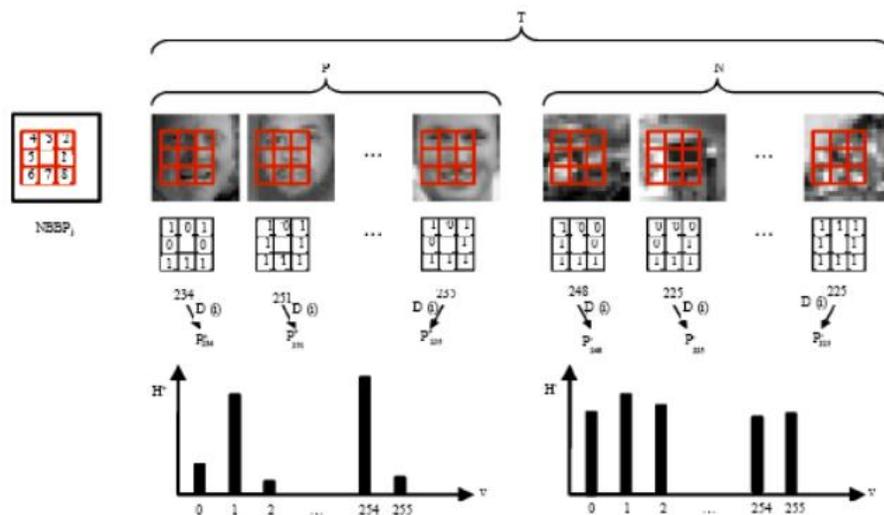


Ilustración 19 Método de Haar Cascade. Tomado de (online: https://docs.opencv.org/3.4.1/d7/d8b/tutorial_py_face_detection.html)

Con esta información se continua con el desarrollo del algoritmo de esqueletización, donde teniendo un punto y realizando la unión de dos articulaciones, se aplica la

arco-tangente para calcular los ángulos de la línea trazada entre la mano y el codo y la del codo al hombro con respecto a la horizontal. Se hace necesario, con toda esta información realizar un diagrama de flujo de las tareas a realizar por el algoritmo para así tener un control para saber de qué manera se implementarán las deferentes faces, técnicas, métodos y funciones.

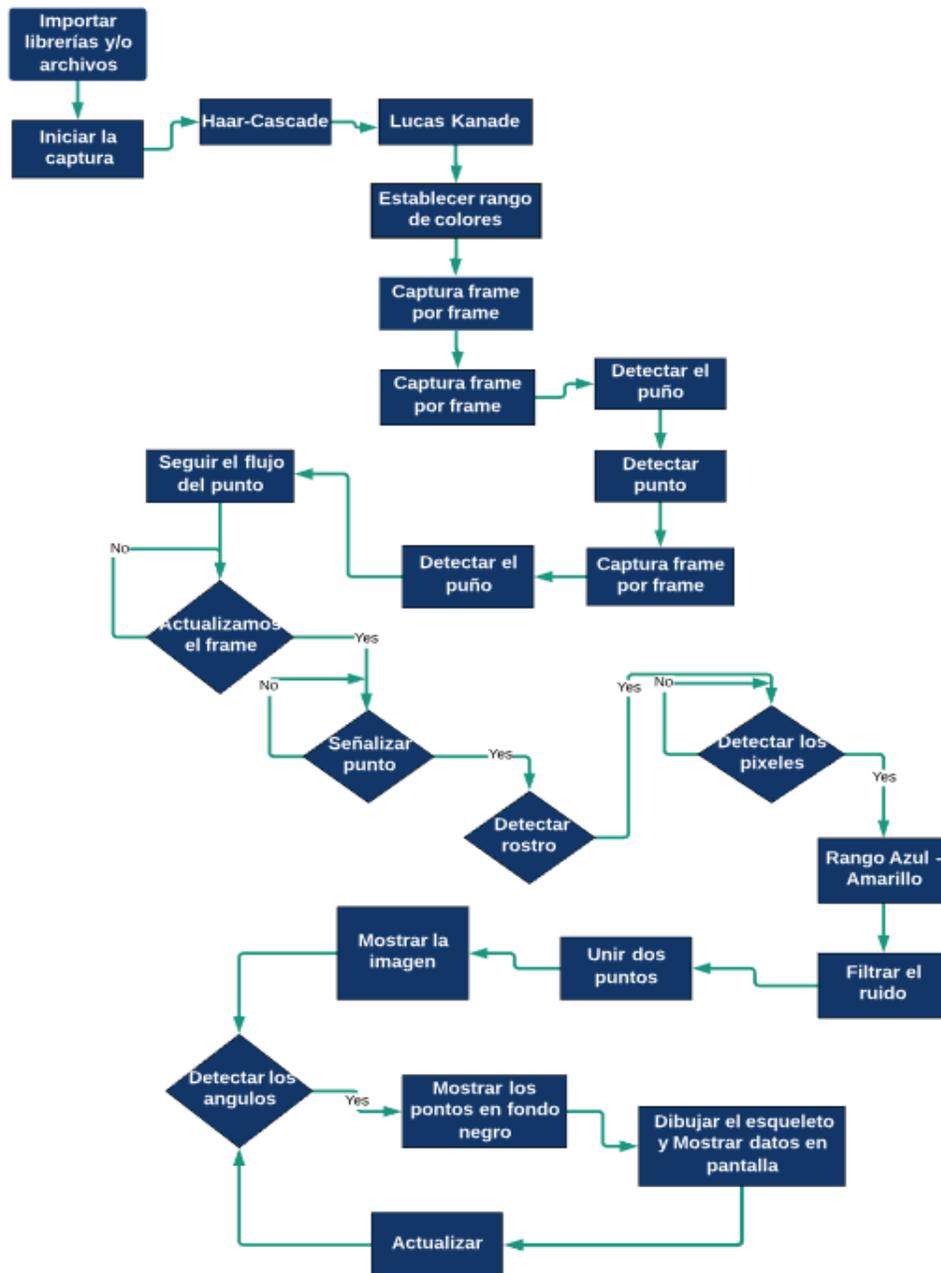
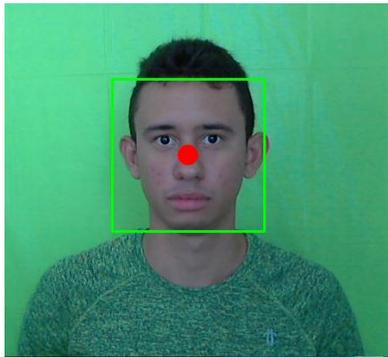
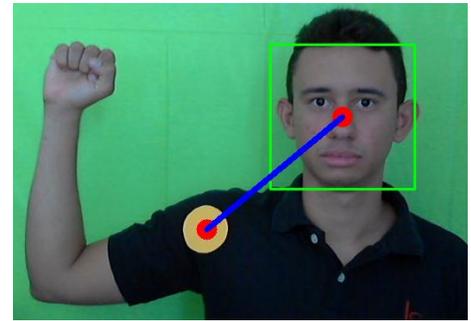


Ilustración 20 Diagrama de flujo algoritmo de esqueletización

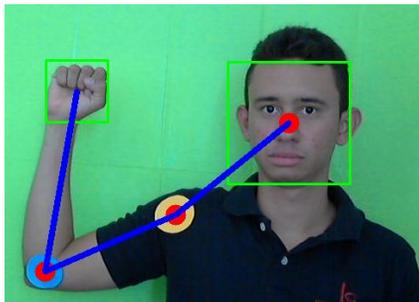
A continuación se muestra el proceso para realizar la captura de movimiento y esqueletización de la extremidad superior del cuerpo humano.



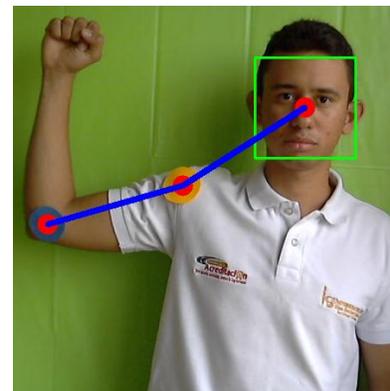
Detección facial y marcado de la articulación



Detección de la articulación (hombro)



Esqueletización completa de la extremidad



Detección de dos articulaciones (hombro y codo)

Como se puede observar en la secuencia de ilustraciones anteriores, se realiza la captura de la extremidad, empezando por el rostro como articulación principal para la detección y tomando el puño como articulación final.

Posteriormente se pasa la captura del esqueleto detectado y es mostrado en un nuevo frame de color negro, el cual tendrá marcado las articulaciones, el trazo de la línea, formando el esqueleto, los datos de las posiciones respecto a (x,y) y los ángulos de las articulaciones, como se muestra en la siguiente Ilustración.

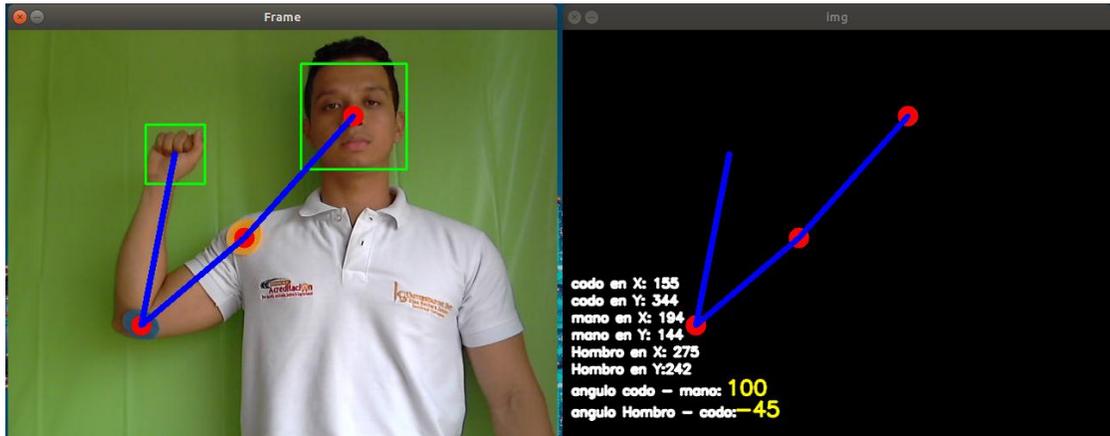


Ilustración 21 Captura y esqueletización de la extremidad superior

3. RESULTADOS

La puesta en ejecución del algoritmo arroja datos de posicionamiento en un plano cartesiano (x, y) de los puntos hallados en las articulaciones, así mismo arroja el ángulo que se forma con la horizontal, tomando como referencia dos puntos, la mano y el codo, para hallar el primer ángulo, y el codo y el hombro para hallar el segundo. Los datos obtenidos reflejan información de un objeto en el plano, sabiendo así su posicionamiento.

Es importante resaltar que la veracidad de los datos arrojados por el algoritmo es corroborada con un inclinómetro, que es colocado en las articulaciones marcadas, para validar los ángulos. Los valores obtenidos de las pruebas del inclinómetro son comparados con los datos arrojados por el algoritmo en las diferentes pruebas. Y cabe resaltar que la detección de esos ángulos fue tomada de forma independiente con el inclinómetro y con el algoritmo, puesto que al colocar el inclinómetro junto a la captura del algoritmo presenta error, ya que capta ese dispositivo como un punto de la articulación (ver Ilustración 22).

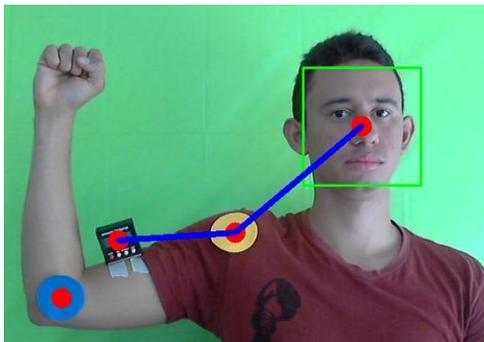


Ilustración 22 Inclinómetro con algoritmo

Es importante mencionar que por la forma del musculo biset, se presentara un margen de error con el inclinómetro, puesto que no está colocado en un área plana.

En la tabla que se encuentra en el anexo 2, se puede observar una comparación de los datos obtenidos, con los datos reales.

4. CONCLUSIONES

Con la investigación de este proyecto de grado, se logra ampliar conocimientos en diferentes áreas, como lo es la cinematografía, video juegos, inteligencia artificial, robótica y en la industria, conociendo así muchas de las técnicas aplicadas en estas áreas y cuáles son las limitaciones y ventajas de cada una de ellas.

También se logra cumplir con todos los objetivos de este proyecto, por lo cual se desarrolló un algoritmo de captura de movimiento de la extremidad superior con herramientas de bajo costo, cimentando las bases para el desarrollo del macro-proyecto de algoritmo de fusión de datos para prototipo de sistema de captura de movimiento de miembro superior mediante vídeo y sensores inerciales en sistemas de seguridad para Robótica Colaborativa, al cual está ligado este trabajo de forma directa y que da cumpliendo al primero objetivo específico de ese macro-proyecto.

Trabajos futuros

Como continuación de este proyecto de investigación, y cualquier otro proyecto, existen diversas líneas de investigación que quedan abiertas y en las que es posible continuar trabajando. Durante el desarrollo de este proyecto has surgida algunas líneas futuras en las que se espera abordar trabajos, los cuales puedan estar directamente relacionadas con este proyecto y son el resultado de cuestiones que han ido surgiendo durante la realización de la misma. Estas líneas pueden servir para retomarlas posteriormente o como opción a trabajos futuros para otros investigadores.

A continuación, se presentan trabajos futuros que pueden desarrollarse como resultado de este proyecto o que, por exceder el alcance de esta tesis, no han podido ser tratados con la suficiente profundidad. Además, se sugieren algunos desarrollos específicos para apoyar y mejorar el modelo y metodología propuestos. Entre los posibles trabajos futuros se destacan:

- Desarrollo de un algoritmo de fusión de datos para prototipo de sistemas de captura de movimiento de miembro superior mediante video y sensores inerciales en sistemas de seguridad para Robótica Colaborativa. El desarrollo de este algoritmo permitirá obtener una data que será inicialmente utilizada como parte de un sistema de seguridad en entornos de robótica colaborativa permitiendo la interacción hombre y robot en tareas complementarias, en una misma celda o puesto de trabajo. Este desarrollo tiene aplicación también en sistemas de análisis de movimiento como herramienta de apoyo diagnóstico, mejoramiento deportivo, sistemas de biometría, entre otros. Más adelante, a partir de este desarrollo inicial se pretende desarrollar un sistema de captura de movimiento completo y que funcione con múltiples personas en una misma celda.

Bibliografía

- [1] M. Buisan, «Industria 4.0,» 2017. [En línea]. Available: <http://www.expansion.com/economia-digital/2017/12/30/5a47d73de5fdea115e8b458d.html>.
- [2] C. Ynzunza, «El Entorno de la Industria 4.0: Implicaciones y Perspectivas Futuras,» [En línea]. Available: <https://www.redalyc.org/jatsRepo/944/94454631006/html/index.html>.
- [3] «nansense,» [En línea]. Available: <https://www.nansense.com/>.
- [4] «Biomec,» [En línea]. Available: <http://www.biomec.com.co/index.html>.
- [5] E. Caballero, «APLICACIÓN PRÁCTICA DE LA VISIÓN ARTIFICIAL,» [En línea]. Available: <http://repository.udistrital.edu.co/bitstream/11349/6104/1/CaballeroBarrigaEdisonRene2017.pdf>.
- [6] C. D. y. C. Romero, «Navegación de robot móvil usando Kinect, OpenCV y Arduino Universidad de los Llanos en Villavicencio,» 2012. [En línea]. Available: <http://ojs.uac.edu.co/index.php/prospectiva/article/view/398>.
- [7] O. P. y. F. Leandro, «Captura de movimiento utilizando el Kinect para el control de una plataforma robótica controlada de forma remota por medio de seguimiento de los puntos de articulación del cuerpo de la Universidad Tecnológica de Pereira.,» 2015. [En línea]. Available: <http://repositorio.utp.edu.co/dspace/handle/11059/5136>.
- [8] R. Ormaechea, «Desarrollo de una aplicación de detección de movimiento basada en comparativa estructural de imágenes de la Universidad de Valladolid.,» 2012. [En línea]. Available: https://www.lpi.tel.uva.es/~santi/personal/PFCs/PFC_R_Ormaechea_2012.pdf.
- [9] M. Jerez, «Diseño y desarrollo de un dispositivo objeto controlado a través del sensor Kinect y la plataforma Arduino orientado al uso ludico-pedagogico de niños en la primera infancia, de la Corporación Universitaria Rafael Núñez,» 2016. [En línea]. Available: http://scienti.colciencias.gov.co:8081/cvlac/visualizador/generarCurriculoCv.do?cod_rh=0001351334.

- [10] «Revista mexicana de ingeniería biomédica,» [En línea]. Available: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0188-95322016000200149.
- [11] Cognex, «Vision artificial,» [En línea]. Available: <https://www.cognex.com/es-co/what-is/machine-vision/what-is-machine-vision..>
- [12] Signaturit, «Blog de Signaturit,» [En línea]. Available: <https://blog.signaturit.com/es/ocr-que-es-el-reconocimiento-optico-de-caracteres>.
- [13] OpenCV, «librería de visión por computador,» [En línea]. Available: <https://www.um.es/geograf/sigmur/teledet/tema06.pdf>.
- [14] L. Gómez, «Sistemas de captura y análisis de movimiento cinemático humano,» Una revisión sistemática, 2018. [En línea]. Available: <http://www.scielo.org.co/pdf/prosp/v16n2/1692-8261-prosp-16-02-00024.pdf>.
- [15] C. e. e. ruido. [En línea]. Available: <https://www.casanovafoto.com/blog/2014/06/como-evitar-ruido-digital/>.
- [16] F. d. robotica. [En línea]. Available: www.Cap_Muest_Barrientos_8448156366.pdf.
- [17] J. & Books, «Internet of Things (IoT): A vision, architectural elements, and future directions,» [En línea]. Available: <https://www.sciencedirect.com/science/article/pii/S0167739X13000241>.
- [18] C. p. I. E. d. H. Dactilares. [En línea]. Available: <http://petrus.upc.es/emsy/jcra05.pdf>.
- [19] P. STUDIO, «Chroma key,» [En línea]. Available: <https://www.popcornstudio.es/chroma-key>.
- [20] O. i. d. normalización, «International Organization for Standardization,» [En línea]. Available: <https://www.iso.org/standard/62996.html> .
- [21] A. H. C. Jaime, «Modelo de Investigación en Ingeniería del Software: Una propuesta de investigación tecnológica,» 2004. [En línea]. Available: <http://web.imt-atlantique.fr/x-info/harbol07/MIFISIS2004.pdf>.

- [22] T. filytrado. [En línea]. Available:
<https://www.um.es/geograf/sigmur/teledet/tema06.pdf>.
- [23] G. Gipuzkoa, «Robótica Colaborativa & Visual Computing,» *Gorain Gipuzkoa*, p. 11.
- [24] M. Buisan, «Expansión - Economía digital,» 2017. [En línea]. Available:
<http://www.expansion.com/economia-digital/2017/12/30/5a47d73de5fdea115e8b458d.html>.
- [25] A. Briseño Cerón, O. A. Dominguez Ramirez y I. Saucedo Ugalde, «El uso de captura de movimiento corporal para el análisis de discapacidades en miembros superior inferior,» *Revista Internacional de Tecnología, Ciencia y Sociedad*, p. 11, 2012.

ANEXOS

Anexo 1.

Fragmentos del algoritmo con las principales funciones y métodos para cumplir con el objetivo de esta investigación.

Detección de rostro

```
fcx=0
fcy=0
face = face_cascade.detectMultiScale(framegray, 1.3, 5)
for region_face in face:
    (x, y, w, h) = region_face
    frame = cv2.rectangle(frame, (x, y), (x + w, y + h), (0, 255, 0), 2)
    fcx=int(x+(w/2))
    fcy=int(y+(h/2))
    cv2.circle(frame, (fcx,fcy), 12, (0, 0, 255), -1)
    cv2.circle(img, (fcx,fcy), 12, (0, 0, 255), -1)
```

Detección de ángulos

```
font=cv2.FONT_HERSHEY_SIMPLEX#tipo de fuente para escribir sobre la imagen
#angulo codo a puño
if fx != 0 and azx != 0 and fx != 0 and fy != 0:
    if azy != fy and azx != fx:
        distCPY=azy-fy #distancia-codo puño en y
        distCPX=azx-fx #distancia-codo puño en x
        anguloCP=math.atan(distCPY/distCPX)
        if math.degrees(anguloCP) <= 0 :
            cv2.putText(img, str(int(math.degrees(anguloCP))+180), (190,425), font, 0.8, (0,255,255), 2, cv2.LINE_AA)
        else:
            cv2.putText(img, str(int(math.degrees(anguloCP))), (190,425), font, 0.8, (0,255,255), 2, cv2.LINE_AA)

#angulo codo hombro
if azx != amx and azy != amy:
    distCHY=amy-azy #distancia codo hombro en Y
    if azx != 0 and amx != 0 and azy != 0 and amy != 0:
        distCHX=amx-azx #distancia codo hombro en X
        anguloCH=math.atan(distCHY/distCHX)
        cv2.putText(img, str(int(math.degrees(anguloCH))), (200,450), font, 0.8, (0,255,255), 2, cv2.LINE_AA)
```

Anexo 2.

Pruebas para validar ángulos de la extremidad superior

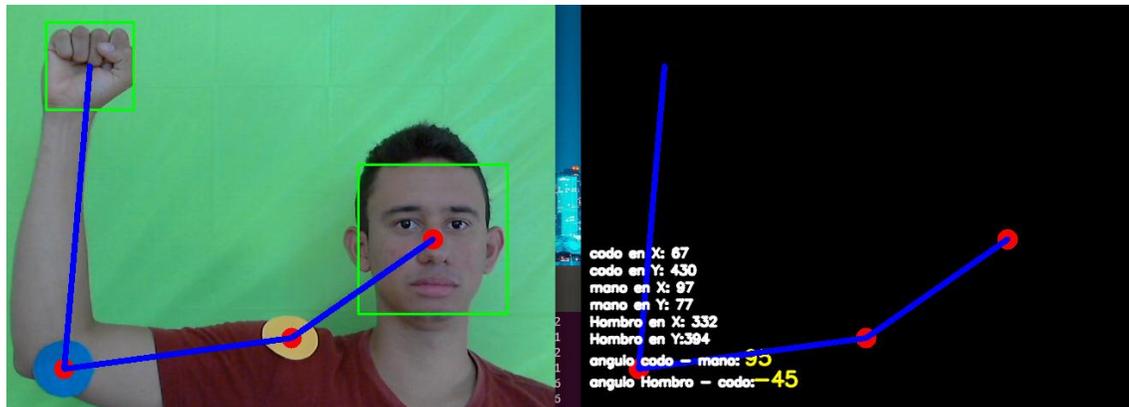
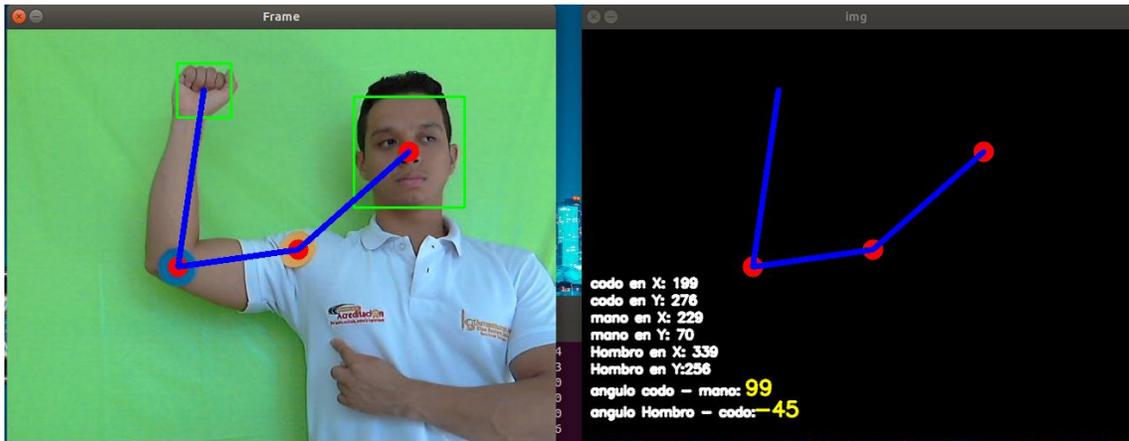


Tabla de resultados ángulos

	Ángulo codo/hombro algoritmo	Inclinómetro	Ángulo codo/mano algoritmo	Inclinómetro
Sujeto 1	-45	33	92	15
	-45	33	99	17
	-43	30	95	16
	-41	31	97	18
	-45	32	87	10
	-46	34	81	6
Sujeto 2	-45	32	90	11
	-42	32	95	18
	-46	34	95	17
	-42	31	83	9
	-48	36	85	10
	-50	39	93	16

Es importante resaltar que el inclinómetro presenta un error aproximado de 7° a 9° y el valor de cero, en el algoritmo es directamente proporcional a 90°