

**SISTEMA DE CONTROL DEL *MOUSE* PARA LA ASISTENCIA A PERSONAS  
CON LIMITACIONES EN LOS MIEMBROS SUPERIORES MEDIANTE EL  
ANÁLISIS DE IMÁGENES DEL MOVIMIENTO DE LA CABEZA Y GESTOS  
DEL ROSTRO**

**OSCAR DARIO MOLINA QUISOBONI**

**UNIVERSIDAD DEL CAUCA  
FACULTAD DE CIENCIAS NATURALES EXACTAS Y DE LA EDUCACIÓN  
DEPARTAMENTO DE FÍSICA  
INGENIERÍA FÍSICA  
POPAYÁN, 2009**

**SISTEMA DE CONTROL DEL *MOUSE* PARA LA ASISTENCIA A PERSONAS  
CON LIMITACIONES EN LOS MIEMBROS SUPERIORES MEDIANTE EL  
ANÁLISIS DE IMÁGENES DEL MOVIMIENTO DE LA CABEZA Y GESTOS  
DEL ROSTRO**

**OSCAR DARIO MOLINA QUISOBONI**

**Trabajo de grado presentado para optar al título de Ingeniero Físico**

**Director**

**Ing. LEONAIRO PENCUE FIERRO**

**UNIVERSIDAD DEL CAUCA  
FACULTAD DE CIENCIAS NATURALES EXACTAS Y DE LA EDUCACIÓN  
DEPARTAMENTO DE FÍSICA  
INGENIERÍA FÍSICA  
POPAYÁN, 2009**

**NOTA DE ACEPTACIÓN**

---

---

---

---

---

**Director**  
**Ing. Leonairo Pencue Fierro**

---

**Jurado**  
**MSc. Jorge W. Coronel García**

---

**Jurado**  
**Ing. Mario Andrés Córdoba G.**

Fecha de sustentación: Popayán, Diciembre 4 de 2009.

## **AGRADECIMIENTOS**

Agradecimientos a Dios por ser mi guía espiritual y darme la fortaleza para afrontar positivamente las dificultades y triunfos que se presentan en mi vida, a mis padres Víctor Manuel y Aura Tulia por su apoyo incondicional, por darme los valores y enseñanzas necesarios para mi formación personal y por permitirme siempre contar con ellos, a mis hermanos Faiver y Dina Marely por su confianza, motivación y apoyo absoluto en la realización de mis metas.

Gracias a toda mi familia por creer en mis capacidades y valorar mi objetivo de obtener la formación profesional y a todos mis amigos que de una u otra forma han contribuido y dado su confianza a lo largo del pregrado para que éste sea posible.

Agradecimientos a Leonairo Pencue quien fue el director de trabajo de grado y al Grupo de Óptica y Láser de la Universidad del Cauca, ya que con sus enseñanzas logre incursionar en el maravilloso mundo de la visión artificial y a todos los profesores del Departamento de Física de la Universidad del Cauca, quienes con sus conocimientos contribuyeron con mi formación profesional.

***Oscar Darío Molina Quisoboni***  
***Pitalito, Octubre de 2009***

## TABLA DE CONTENIDO

<b>ÍNDICE DE FIGURAS</b> .....	V
<b>ÍNDICE DE TABLAS</b> .....	VIII
<b>INTRODUCCIÓN</b> .....	1
<b>CAPÍTULO 1. DISCAPACIDAD</b> .....	4
1.1. DISCAPACIDAD DE LOS MIEMBROS SUPERIORES. . . . .	5
1.2. ESTADÍSTICA DE LA DISCAPACIDAD EN COLOMBIA. . . . .	5
1.3. DIFICULTAD SOCIAL DE LOS DISCAPACITADOS. . . . .	7
1.4. SISTEMAS DE ASISTENCIA PARA DISCAPACITADOS. . . . .	8
1.4.1. <i>MOUSE</i> DEL COMPUTADOR CONTROLADO POR VOZ. . . . .	9
1.4.2. <i>EYE MOUSE</i> . . . . .	9
1.4.3. EMULADOR DE TECLADO Y <i>MOUSE</i> PARA DISCAPACITADOS MOTRICES SEVEROS UTILIZANDO DISPOSITIVOS LÓGICOS PROGRAMABLES. . . . .	9
1.4.4. OCZ NIA - NEURAL IMPULSE ACTUATOR. . . . .	9
1.4.5. HEADMOUSE ® EXTREME. . . . .	10
1.4.6. <i>MOUSE</i> OCULAR. . . . .	10
1.4.7. SMARTNAV 4 AT. . . . .	10
1.4.8. ORGANIZACIÓN MILBA”T. . . . .	10
1.4.9. HEADMOUSE. . . . .	11
<b>CAPÍTULO 2. VISIÓN POR COMPUTADOR</b> .....	12
2.1. LA VISIÓN POR COMPUTADOR EN APLICACIONES PARA LOS DISCAPACITADOS. . . . .	13
2.2. MÉTODOS PARA LA DETECCIÓN DE MOVIMIENTO. . . . .	14
2.2.1 SUSTRACCIÓN DE FONDO. . . . .	15
2.2.2 FLUJO ÓPTICO. . . . .	16
2.2.2.1. MÉTODO DE HORN & SCHUNK. . . . .	17
2.2.2.2. MÉTODO DE LUCAS & KANADE. . . . .	19
2.2.3. MÉTODO POR COLOR. . . . .	21

2.2.3.1. ESPACIO HSV. ....	21
2.2.3.2. ESPACIO YCBCR. ....	22
<b>CAPÍTULO 3. DETECCIÓN DE LA CABEZA. ....</b>	<b>24</b>
3.1. FILTROS. ....	24
3.1.1. FILTRO GAUSSIANO. ....	25
3.1. DETECCIÓN DE BORDES. ....	26
3.2. SUSTRACCIÓN DE FONDO. ....	27
3.3. INTEGRALES PROYECTIVAS. ....	29
3.4. IMPLEMENTACIÓN DEL ALGORITMO. ....	31
<b>CAPÍTULO 4. MOVIMIENTO DE LA CABEZA. ....</b>	<b>34</b>
4.1. SISTEMA DE TRES PUNTOS. ....	36
4.1.1. CONSTRUCCIÓN DEL SISTEMA DE TRES PUNTOS. ....	37
4.1.1.1. MATERIALES. ....	37
4.1.1.2. PASOS DE CONSTRUCCIÓN DEL SISTEMA DE TRES PUNTOS. ....	38
4.1.1.2.1. BANDA. ....	38
4.1.1.2.2. ELEMENTO L. ....	40
4.1.2. MODO DE UTILIZAR EL SISTEMA DE TRES PUNTOS. ....	43
4.2. SEGMENTACIÓN. ....	44
4.2.1. DETECTOR DE CANNY. ....	45
4.2.2. CONTORNOS CERRADOS. ....	46
4.3. DETECCIÓN DE LOS TRES PUNTOS. ....	48
4.3.1. INTEGRALES PROYECTIVAS. ....	48
4.3.2. SEGUIMIENTO DE PUNTOS. ....	50
4.4. DISTANCIAS ENTRE LOS PUNTOS. ....	51
4.5. IMPLEMENTACIÓN DE ALGORITMO. ....	54
<b>CAPÍTULO 5. DETECCIÓN DE LOS GESTOS DEL ROSTRO. ....</b>	<b>55</b>
5.1. GESTOS DE LOS OJOS. ....	55
5.1.1. DETECCIÓN DEL MOVIMIENTO DE LOS OJOS: FLUJO ÓPTICO. ....	56

5.1.2. IMPLEMENTACIÓN DEL ALGORITMO. . . . .	57
5.2. GESTOS DE LAS CEJAS. . . . .	58
5.2.1. DETECCIÓN DEL MOVIMIENTO DE LAS CEJAS: FLUJO ÓPTICO. . . . .	58
5.2.2. IMPLEMENTACIÓN DEL ALGORITMO. . . . .	59
5.3. GESTOS DE LA BOCA. . . . .	60
5.3.1. DETECCIÓN DEL MOVIMIENTO DE LA BOCA: FLUJO ÓPTICO. . . . .	60
5.3.2. IMPLEMENTACIÓN DEL ALGORITMO. . . . .	61
<b>CAPÍTULO 6. INTERFAZ ENTRE MOVIMIENTOS, GESTOS Y LOS CONTROLES DEL <i>MOUSE</i>. . . . .</b>	<b>63</b>
6.1. MOVIMIENTOS DEL CURSOR DEL <i>MOUSE</i> . . . . .	63
6.2. CONTROLES DEL <i>MOUSE</i> . . . . .	64
6.2.1. “ <i>CLICK</i> ”. . . . .	65
6.2.2. DOBLE “ <i>CLICK</i> ”. . . . .	65
6.2.3. “ <i>CLICK</i> ” SOSTENIDO. . . . .	66
6.2.4. “ <i>CLICK</i> ” DERECHO. . . . .	66
<b>CAPÍTULO 7. INTERFAZ GRAFICA. . . . .</b>	<b>67</b>
7.1. INTERFAZ GRAFICA DE USUARIO. . . . .	67
7.2. MANUAL DE USUARIO. . . . .	69
7.3. CONDICIONES DE FUNCIONAMIENTO. . . . .	70
<b>CAPÍTULO 8. RESULTADOS Y DISCUSIÓN. . . . .</b>	<b>71</b>
8.1. DETECCIÓN DE LA CABEZA. . . . .	71
8.2. MOVIMIENTOS DE LA CABEZA. . . . .	72
8.3. DETECCIÓN DE LOS GESTOS DEL ROSTRO. . . . .	74
8.4. SCODEM 2009. . . . .	75
8.4.1. PRUEBAS CON USUARIOS. . . . .	77
<b>CONCLUSIONES. . . . .</b>	<b>81</b>
<b>TRABAJO A FUTURO. . . . .</b>	<b>83</b>
<b>REFERENCIAS. . . . .</b>	<b>84</b>
<b>ANEXOS. . . . .</b>	<b>88</b>

**ANEXO A: MANUAL DE USUARIO..... 88**  
**ANEXO B: ENCUESTA DE EVALUACIÓN DEL SISTEMA..... 95**

## ÍNDICE DE FIGURAS

<b>Figura 1.</b> Diagrama de la resta de imágenes. . . . .	15
<b>Figura 2.</b> Detección falsa de flujo óptico por movimiento e iluminación, Imagen tomada de [6]. . . . .	20
<b>Figura 3.</b> Espacio de color HSV. . . . .	22
<b>Figura 4.</b> Imagen en escala de grises con el filtro gaussiano. . . . .	26
<b>Figura 5.</b> Operador de Sobel. (a) Dirección horizontal. (b) Dirección vertical. . . . .	27
<b>Figura 6.</b> Sustracción de fondo de la cabeza del usuario. (a) Aumento de movimiento. (b) Disminución de movimiento. . . . .	28
<b>Figura 7.</b> Proyecciones. (a) Imagen a obtener las proyecciones. (b) Proyección vertical (c) Proyección horizontal. . . . .	30
<b>Figura 8.</b> Suma de los filtros de Sobel en la dirección horizontal y vertical. . . . .	32
<b>Figura 9.</b> Detección de la cabeza del usuario. . . . .	33
<b>Figura 10.</b> Movimiento de giro de la cabeza. (a) Horizontal. (b) Vertical. . . . .	35
<b>Figura 11.</b> Movimiento de la cabeza. (a) Posición inicial. (b) Horizontal. (c) Vertical. . . . .	35
<b>Figura 12.</b> Sistema de tres puntos. . . . .	36
<b>Figura 13.</b> Diagrama tridimensional del sistema de tres puntos. . . . .	37
<b>Figura 14.</b> Dimensiones de la cinta de velcro en los extremos de la banda del sistema de tres puntos. . . . .	39
<b>Figura 15.</b> Dimensiones de la banda del sistema de tres puntos. . . . .	39
<b>Figura 16.</b> Elemento L. (a) Diagrama con las dimensiones. (b) Cubierto con cinta negra. . . . .	40

<b>Figura 17.</b> Frente en papel del elemento L. (a) Diagrama con las dimensiones. (b) Imagen Real. ....	41
<b>Figura 18.</b> Frente del elemento L. ....	41
<b>Figura 19.</b> Diagrama de dimensiones de la tela que cubre el frente del elemento L. ....	42
<b>Figura 20.</b> Elemento L. ....	42
<b>Figura 21.</b> Posición de la banda en la cabeza. ....	43
<b>Figura 22.</b> Posición del elemento L en la cabeza. ....	43
<b>Figura 23.</b> Zona de análisis del sistema de tres puntos. ....	44
<b>Figura 24.</b> Detector de Canny aplicado a la zona del sistema de tres puntos. ....	46
<b>Figura 25.</b> Detección de contornos cerrados. ....	47
<b>Figura 26.</b> Segmentación del sistema de tres puntos. ....	48
<b>Figura 27.</b> Regiones para la identificación de puntos en el sistema de tres puntos. ....	49
<b>Figura 28.</b> Detección del sistema de tres puntos por el método de integrales proyectivas. ....	50
<b>Figura 29.</b> Secuencia de imágenes del seguimiento del sistema de puntos. ....	51
<b>Figura 30.</b> Diagrama de distancias del sistema de tres puntos. ....	52
<b>Figura 31.</b> Diagrama del sistema de tres puntos con inclinación. ....	53
<b>Figura 32.</b> Región de análisis de los ojos. ....	56
<b>Figura 33.</b> Flujo óptico en la región de los ojos. (a) cerrándolos. (b) abriéndolos. ....	57
<b>Figura 34.</b> Región de análisis de las cejas. ....	59

<b>Figura 35.</b> Flujo óptico en la región de las cejas. ....	59
<b>Figura 36.</b> Región de análisis de la boca. ....	61
<b>Figura 37.</b> Flujo óptico en la región de la boca. ....	61
<b>Figura 38.</b> <i>Mouse</i> del computador. ....	64
<b>Figura 39.</b> Interfaz grafica de usuario. ....	68

## ÍNDICE DE TABLAS

<b>Tabla 1.</b> Departamentos de Colombia con los porcentajes más altos de personas con limitaciones permanentes. . . . .	6
<b>Tabla 2.</b> Departamentos de Colombia con los porcentajes más altos de las personas en condiciones de discapacidad con limitaciones permanentes para usar los brazos y manos. . . . .	7
<b>Tabla 3.</b> Materiales utilizados para la fabricación del sistema de tres puntos. . . . .	38
<b>Tabla 4.</b> Respuesta de los parámetros de distancia al movimiento de giro de la cabeza del usuario. . . . .	53
<b>Tabla 4.</b> Resultados de la encuesta realizada a los usuario para evaluar SCODEM 2009. . . . .	78

---

---

## INTRODUCCIÓN

---

---

Por un lado los sistemas de visión artificial, se han convertido actualmente en una herramienta muy útil para la innovación tecnológica y solución de problemas de forma eficiente, por lo cual es viable realizar aplicaciones de este tipo. Por otro lado las personas con discapacidades en los miembros superiores tienen dificultades al intentar acceder a un computador, herramienta vital en el mundo moderno para un adecuado desarrollo y alcance a la información, para quienes no es suficiente adquirir computadores convencionales.

Se ha desarrollado un sistema de control del *mouse* en propiedades específicas como el desplazamiento horizontal, vertical, *click*, doble *click*, *click* derecho y *click* sostenido, mediante los movimientos de la cabeza y gestos del rostro, orientado para las de personas con limitaciones en los miembros superiores, los cuales pueden utilizar el sistema como una asistencia para el manejo del computador.

La adquisición de video se realiza con una cámara de bajo costo, y el sistema de tres puntos utilizado para detectar el movimiento de la cabeza es de bajo costo y fácil fabricación. Con la utilización del procesamiento de imagen se obtuvieron buenos resultados y se crean nuevas opciones para nuevas aplicaciones basadas en estos métodos.

En el capítulo 1 se enfatiza en los discapacitados, su población en Colombia y las diferentes aplicaciones que se ha realizado para los limitados en los miembros superiores.

En el capítulo 2 se hace un análisis sobre la visión por computador y su viabilidad para este proyecto, además se analizan los diferentes métodos que se estudiaron para la detección de movimientos.

La detección de la cabeza en la imagen se estudia en el capítulo 3, este proceso es vital para las etapas posteriores que dependen de una buena ubicación de la cabeza para su óptimo funcionamiento.

En el capítulo 4 se analiza y explica el sistema de tres puntos que se implemento para la detección del movimiento de la cabeza, enfatizando el movimiento vertical (arriba y abajo) y el movimiento horizontal (izquierda y derecha), utilizados para el desplazamiento del cursor del *mouse* en las cuatro direcciones respectivas.

Luego de ubicada la cabeza se hace necesario determinar los gestos de la cara basados específicamente en la boca, ojos y cejas. Los métodos y características de los gestos detectados se analizan en el capítulo 5.

En el capítulo 6 se estudian las opciones para hacer la interfaz entre los movimientos de la cabeza y los gestos de la cara con los controles del *mouse*, como el movimiento izquierda, derecha, arriba, abajo y *click*, doble *click*, *click* derecho, *click* sostenido, respectivamente.

En el capítulo 7 se presenta la interfaz grafica de usuario, el manual de usuario y se describen las condiciones de funcionamiento de “**SCODEM 2009**” nombre asignado a la aplicación final.

En el capítulo 8 se analizan y discuten los resultados obtenidos en todo el proyecto.

Finalmente se encuentran las conclusiones y el trabajo a futuro de la aplicación, basado en la importancia de la detección de los grados de libertad, obtenidos en el movimiento de la cabeza.

# CAPÍTULO 1

---

---

## DISCAPACIDAD

---

---

La concepción de la discapacidad ha tenido cambios positivos en las últimas décadas debido a la búsqueda de la humanidad por la dignificación del hombre. Las personas con discapacidades tienen una incidencia social, cultural, jurídica y económica, que hace necesario tenerlas en cuenta para un adecuado desarrollo social [1].

La discapacidad en Colombia se trabaja bajo el enfoque el cual la considera como un problema social, resultado de un complejo conjunto de condiciones de entorno, muchas de ellas creadas por el ambiente social; desde esta perspectiva, la discapacidad es un hecho multidimensional [1].

Las acciones que se realicen entorno a mejorar las condiciones de los discapacitados deben ser integrales, donde se involucren diferentes sectores con entidades tanto públicas como privadas.

Las condiciones sociales, económicas, políticas e incluso demográficas de nuestro país, incluidas la guerra, la violencia, los niveles de pobreza, el desamparo, el desplazamiento y

el envejecimiento de la población, hacen que la discapacidad se convierta cada vez más en un fenómeno complejo que afecta a un grupo importante de personas para las que es necesario garantizar un acceso equitativo a la atención y a los servicios de salud, educación y recreación, igualdad de condiciones y oportunidades laborales y una plena participación en la sociedad que les permita contribuir a su propio desarrollo y al desarrollo socioeconómico de su comunidad [2].

### **1.1. DISCAPACIDAD DE LOS MIEMBROS SUPERIORES**

Entre las diferentes discapacidades se encuentra la de los miembros superiores consistente en tener problemas para tocar, coger y manipular objetos con las manos, mover y utilizar brazos, hombros, manos, dedos [3], donde algunas de estas limitaciones son causadas por amputación, parálisis de los miembros superiores, enfermedades reumáticas, enfermedades congénitas, entre otras.

Este tipo de discapacidad con frecuencia limita la capacidad para aprender de la experiencia, resolver problemas y superar obstáculos de la vida cotidiana, así como para ajustarse con flexibilidad a nuevas situaciones. Estas personas requieren de una adecuada educación y un entorno que facilite su normal desarrollo en las diferentes áreas [3].

### **1.2. ESTADÍSTICA DE LA DISCAPACIDAD EN COLOMBIA**

A nivel mundial ha habido un interés por conocer los datos de la población discapacitada, lo cual ha sido difícil por diferentes circunstancias como del mismo concepto de discapacidad que establece cada país, sin embargo organismos como Naciones Unidas han contribuido para que organismos especializados cooperen con los países en desarrollo en la recopilación de datos estadísticos.

En Colombia el “DANE” (Departamento Administrativo Nacional De Estadística), en el 2005 realizo un censo de la población con discapacidad estableciendo que en Colombia el 6.3% del total de la población son personas con alguna clase de discapacidad [4].

Algunos departamentos muestran un porcentaje mayor de discapacitados como se puede ver en la tabla 1.

**Tabla 1.** Departamentos de Colombia con los porcentajes más altos de personas con limitaciones permanentes.

Fuente: Departamento Administrativo Nacional De Estadística, DANE, censo general 2005. Discapacidad personas con limitaciones permanentes, 2005.

<b>DEPARTAMENTOS</b>	<b>PORCENTAJE DE LIMITACIONES PERMANENTES</b>
Cauca	9.5
Nariño	9.0
Boyacá	8.8
Huila	8.5
Quindío	8.0
Tolima	7.8
Putumayo	7.4

De acuerdo al “DANE” del 6.3% de la población con discapacidad el 14.6% tiene limitaciones físicas en los miembros superiores [4], de tal forma hay 400.000 personas aproximadamente con alguna clase de discapacidad en los miembros superiores.

**Tabla 2.** Departamentos de Colombia con los porcentajes más altos de las personas en condiciones de discapacidad con limitaciones permanentes para usar los brazos y manos.

Fuente: Departamento Administrativo Nacional De Estadística, DANE, censo general 2005. Discapacidad personas con limitaciones permanentes, 2005.

<b>DEPARTAMENTOS</b>	<b>PORCENTAJE DE LIMITACIONES PERMANENTES PARA USAR LOS BRAZOS Y MANOS</b>
Vichada	55.1
Vaupés	30.2
Guaviare	23.7
Guainía	18.3
Cesar	16.3
Amazonas	16.3
Magdalena	16.1
Meta	16.1
Boyacá	15.9
Antioquia	15.8
Huila	15.6
Atlántico	15.5
Cundinamarca	15.4
Bogotá D.C	15.0
Putumayo	15.0
Caldas	14.8
Tolima	14.7
Santander	14.7
Cauca	14.5

### **1.3. DIFICULTAD SOCIAL DE LOS DISCAPACITADOS**

Las personas discapacitadas requieren de entornos adecuados para su desarrollo integral, por lo cual, estas personas tienen que luchar por los derechos que tienen a su educación, salud, actividad física, actividad laboral, entre otros. En Colombia se han creado leyes para la protección de los discapacitados siguiendo las políticas internacionales, convirtiéndose en un país que busca mejorar cada vez más la calidad de vida de los discapacitados, sin embargo es difícil que toda la legislación sea puesta en práctica sin un compromiso serio, por parte de toda la sociedad [5].

Las personas con discapacidades actualmente son discriminadas en muchos sectores de la sociedad a pesar de la existencia de leyes que protegen sus intereses y el libre desarrollo. La tecnología, las entidades públicas y privadas, los espacios comunes, centros deportivos, los proyectos de vivienda social, entre otros, muchas veces no tienen en cuenta las necesidades de las personas discapacitadas y no tienen las condiciones que ellos requieren, violando así, sus derechos constitucionales.

En la actualidad el computador se ha convertido en una herramienta muy útil, sobre todo con la llegada de la internet, donde se permite interaccionar con el mundo en diferentes áreas, como en el conocimiento, el entretenimiento, o la comunicación [6]. Además, el computador se ha incorporado rápidamente en diferentes ámbitos como en la educación, la ciencia, la medicina, la industria, entre otros. Así, se hace necesario que todas las personas, sin tener en cuenta su condición física, tengan facilidad para acceder a estos dispositivos, pero es evidente que las personas con limitaciones en los miembros superiores tienen inconvenientes al intentar acceder a computadores convencionales, debido a que éstos no son fabricados con periféricos adecuados para éstas limitaciones.

Por lo tanto, surge la necesidad de crear sistemas que sean alternativos a los periféricos tradicionales del computador y así éstos se adecuen a los requerimientos de las personas con discapacidades en los miembros superiores.

#### **1.4. SISTEMAS DE ASISTENCIA PARA DISCAPACITADOS**

Actualmente se han desarrollado sistemas para ayudar a los discapacitados en el manejo de los periféricos del computador. Estos sistemas han sido diseñados con métodos y técnicas que aprovechan las capacidades del discapacitado.

#### **1.4.1. *MOUSE* DEL COMPUTADOR CONTROLADO POR VOZ**

Es un software desarrollado en la Universidad de Washington que proporciona una alternativa al periférico tradicional del *mouse* usando la voz humana. El Joystick Vocal capta 100 veces por segundo los sonidos e instantáneamente los convierte en movimientos en la pantalla. Los diferentes sonidos de las vocales dictan la dirección, y otros sonidos mueven el cursor en una de ocho direcciones. Los usuarios pueden realizar una transición suave de una vocal a otra, y los sonidos más fuertes hacen que el cursor se mueva más rápido. Dos sonidos simulan el hacer clic y también los botones del ratón se manejan con sonidos [7].

#### **1.4.2. *EYE MOUSE***

La aplicación fue desarrollada por estudiantes argentinos, donde se efectúan los movimientos del mouse con los movimientos de los ojos. Este sistema consta de un software y un hardware [8].

#### **1.4.3. EMULADOR DE TECLADO Y *MOUSE* PARA DISCAPACITADOS MOTRICES SEVEROS UTILIZANDO DISPOSITIVOS LÓGICOS PROGRAMABLES**

El sistema está destinado a personas con discapacidades motoras severas que sólo pueden controlar un tipo de dispositivo equivalente a un pulsador. El sistema emula al teclado y al mouse de una computadora con un diseño en hardware basado en dispositivos lógicos programables que hacen que el proceso se visualice de forma externa al monitor [9].

#### **1.4.4. OCZ NIA - NEURAL IMPULSE ACTUATOR**

El sistema tiene tres tipos de sensores ubicados sobre la frente del usuario que miden impulsos eléctricos enviados por el sistema nervioso; con este sistema se controla el

computador y se controlan funciones de los juegos que permiten al usuario una mayor autonomía en esta clase de aplicaciones [10].

#### **1.4.5. HEADMOUSE ® EXTREME**

El sistema utiliza los movimientos de la cabeza del usuario para los movimientos del puntero del *mouse*. Este dispositivo no necesita de un software y está diseñado para personas con limitación en las manos [11].

#### **1.4.6. MOUSE OCULAR**

El creador del sistema fue el profesor Manuel Cardoso, de la Universidad Federal de Amazonas, que trabajó apoyado por investigadores de la Fundación Feitoza. El mouse se compone de cinco sensores que se colocan en la frente y en las sienes. Los sensores captan los movimientos del globo ocular y los transmiten al computador a través de un pequeño módem. El "ENTER" es ordenado a través de un parpadeo [12].

#### **1.4.7. SMARTNAV 4 AT**

El sistema consta de un hardware y software que se utilizan como una alternativa al *mouse* convencional utilizando los movimientos de la cabeza. Este sistema está diseñado para las personas con esclerosis lateral, lesiones de la médula espinal, tales como tetraplejia, distrofia muscular, y otras necesidades especiales [13].

#### **1.4.8. ORGANIZACIÓN MILBA”T**

Es una organización liderada por el ingeniero Dany Barak que contribuye con soluciones a personas discapacitadas especialmente a soldados heridos en combate, mejorando las condiciones de vida [14].

#### **1.4.9. HEADMOUSE**

Este sistema fue diseñado por el grupo de robótica de la Universidad de Lleida, para personas con discapacidades motrices. El *HeadMouse* utiliza los movimientos de la cabeza y los gestos del rostro para controlar el *mouse* del computador. Cuenta con una versión absoluta y una versión relativa para el movimiento del *mouse* [15].

# CAPÍTULO 2

---

---

## VISIÓN POR COMPUTADOR

---

---

La visión por computador es una rama de la inteligencia artificial que tiene por objetivo modelar matemáticamente los procesos de percepción visual en los seres vivos y generar programas que permitan simular estas capacidades visuales por computador [16].

Es preciso reconocer que hoy por hoy la visión por computador a veces no es la mejor solución a un problema. Existen muchas ocasiones en las que el problema es tan complejo que la solución humana es lo mejor. El sistema de visión humana puede describir automáticamente una textura en detalle, un borde, un color, una representación bidimensional de una tridimensional, ya que puede diferenciar entre imágenes de diferentes personas, firmas, colores, etc. Pero a veces, las soluciones humanas tienden a ser inexactas o subjetivas y en ocasiones lentas, las personas se fatigan, se distraen causando inconvenientes en los procesos [17].

La visión por computador es una herramienta que ha tenido cada vez más importancia por su gran versatilidad y adaptabilidad, para solucionar problemas que se presentan en

diferentes sectores, teniendo así un gran campo de acción en donde se pueden desarrollar distintas aplicaciones, especialmente cuando se requiere buena precisión, alta velocidad en los procesos y bajos costos.

A continuación se presentan algunos sectores en los que se utiliza la visión por computador.

Navegación en robótica: se puede utilizar como un elemento de un sistema multisensorial. Se extrae la información del entorno del robot y se integra con el resto de la información proporcionada por otro tipo de sensores. El resultado es la reconstrucción de la escena 3-D, que permite la navegación autónoma del sistema [17].

Biología: se pueden realizar análisis microscópico y macroscópico, ya sea de microorganismos, células o texturas en vegetales respectivamente.

Medicina: éste es un campo muy amplio, donde se realizan muchas aplicaciones en las diferentes áreas de la medicina. Se hacen diagnósticos de dolencias o enfermedades, en base a imágenes obtenidas de radiografías, resonancias magnéticas, tomografías, etc [17].

Sistemas de calidad: la visión por computador es muy útil en estos sistemas debido a la velocidad de los procesos, reducción de costos, funcionamiento continuo, entre otros.

## **2.1. LA VISIÓN POR COMPUTADOR EN APLICACIONES PARA LOS DISCAPACITADOS**

Dentro de las muchas aplicaciones que pueden realizarse para los discapacitados está la asistencia para el control del *mouse* para los limitados de los miembros superiores, para quienes ya se han creado varios sistemas basados en diferentes técnicas que se mencionaron en el capítulo 1. Algunos de los sistemas utilizan la visión por computador como

*headmouse*, *eye mouse*, entre otros; demostrando que esta herramienta es robusta para la solución de problemas para los discapacitados.

En este proyecto se utiliza técnicas de la visión por computador para implementar el sistema de control del *mouse* para la asistencia a personas con limitaciones en los miembros superiores, debido a su versatilidad, adaptación a diferentes entornos, bajo costo, ejecución de procesos en tiempo real, entre otros.

Las diversas herramientas de la visión por computador como la detección de movimiento, filtros, segmentación, etc; permiten el análisis de los movimientos de la cabeza y la detección de los gestos del rostro, los cuales se utilizan para controlar el *mouse* del computador.

## **2.2. MÉTODOS PARA LA DETECCIÓN DE MOVIMIENTO**

El ser humano posee capacidades visuales como la de seguir objetos que están en el campo de visión, y aunque muchas se hagan de forma inconsciente, estas capacidades permiten realizar actividades cotidianas, como leer, caminar, conducir, entre otras [18]. La visión humana es capaz de detectar rápidamente un objeto en movimiento y analizar su trayectoria, hay que tener en cuenta que con objetos a altas velocidades se pueden tener problemas.

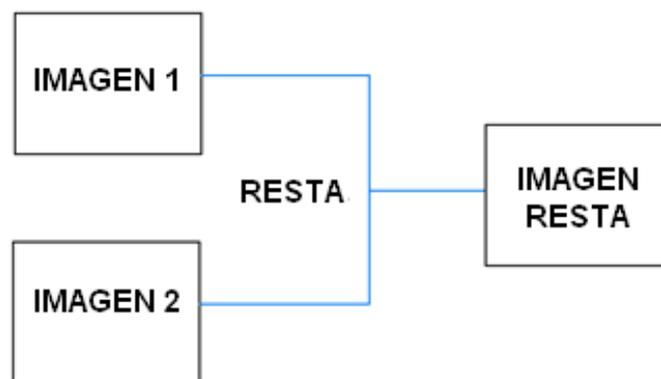
La introducción de las diferentes capacidades visuales en los sistemas de visión artificial, es una de las áreas de investigación en la visión por computador y la robótica de la actualidad [18]. Se han realizado estudios en la detección de movimiento, porque hay muchas aplicaciones tecnológicas las cuales requieren incorporar este tipo de detecciones, para lo cual hay que superar varios retos, debidos a factores como las características del objeto a detectar, la velocidad de movimiento, la iluminación y el entorno en el que se mueve el objeto.

Para este proyecto es muy importante la detección de movimiento y se hace necesario hacer un análisis de diferentes métodos con el fin de encontrar el que mejor se adecue a las condiciones en el cual se desarrolla la aplicación.

Para la aplicación la detección debe hacerse en tiempo real, y debe funcionar en un entorno variable y con iluminación no constante.

### 2.2.1. SUSTRACCIÓN DE FONDO

El método de sustracción de fondo tiene como objetivo eliminar el fondo, entiéndase como fondo al conjunto de píxeles inmóviles que no pertenecen a ningún objeto [19], para esto se realiza una resta de dos imágenes consecutivas (ver Figura 1.), con el fin de obtener la traza del objeto que cambió de posición.



**Figura 1.** Diagrama de la resta de imágenes.

La sustracción de fondo es adecuada para la detección de movimiento porque no es tan sensible al ruido, aunque hay que tener en cuenta las condiciones de luz, de tal forma que no sea variable entre las imágenes a restar, lo cual se puede controlar con la alta velocidad de captura de las imágenes. Hay que evitar los materiales que reflejen la luz, debido a que pueden incorporar ruido en la detección del objeto en movimiento.

El método tiene la ventaja de ser rápido, lo cual es necesario en aplicaciones que funcionan en tiempo real.

### **2.2.2. FLUJO ÓPTICO**

El flujo óptico se puede definir como el movimiento aparente captado en una secuencia de imágenes mediante el análisis de los cambios que sufre la intensidad en los niveles de gris de los puntos de las imágenes, el cual describe la dirección y velocidad del movimiento de dichos puntos [20].

Debido a que los objetos se mueven en el espacio tridimensional y la imagen es una proyección bidimensional, se obtiene la velocidad real transformada en un campo de movimiento, con la pérdida de información consiguiente, que hace necesario utilizar secuencias temporales de imágenes para obtener los parámetros [20].

Para encontrar el flujo óptico se puede utilizar un método basado en los gradientes espacio-temporales, que determina los cambios espaciales y temporales del patrón de grises de la imagen [21]. Los métodos varían pero se deben tener en cuenta algunas etapas para obtener datos confiables.

La primera etapa es la de filtrado previo o suavizado, con filtros paso-bajo, para extraer la estructura de interés de la señal y para aumentar la relación señal/ruido [20].

La segunda etapa es la de extracción de medidas básicas, como derivadas espacio-temporales, para medir las componentes normales de la velocidad, o búsqueda de zonas mediante técnicas de correlación [20].

La tercera etapa es la de integración de las medidas realizadas en el punto anterior para producir un campo bidimensional de flujo, lo que generalmente supone, realizar restricciones sobre la suavidad del campo de flujo fundamental que se ha medido [20].

Si la intensidad de imagen de cada punto visible de la escena no cambia respecto al tiempo, se tiene [20]:

$$I(x, y, t) \approx I(x + \delta x, y + \delta y, t + \delta t) \quad (1)$$

Donde  $I(x, y, t)$  es la función de la intensidad de la imagen,  $(\delta x, \delta y)$  es el desplazamiento de la región en  $(x, y, t)$  después de un tiempo  $\delta t$ . A la ecuación (1) se le realiza un desarrollo en serie de Taylor y queda de la siguiente forma:

$$I(x + \delta x, y + \delta y, t + \delta t) = I(x, y, t) + \nabla I \cdot (\delta x + \delta y) + \nabla^2 I_1 + R^n \quad (2)$$

En la ecuación (2),  $\nabla I = (I_x, I_y)$  e  $I_1$  son derivadas parciales de primer orden y  $R^n$  representa los términos de segundo orden y superiores, los cuales se suponen despreciables. Por lo tanto se tiene:

$$(I_x, I_y) \cdot (u, v) + I_1 = 0 \quad (3)$$

En donde  $(I_x, I_y)$  corresponde al gradiente espacial de la intensidad y  $(u, v)$  corresponde al flujo óptico, siendo  $u = \frac{dx}{dt}$  y  $v = \frac{dy}{dt}$ . En la ecuación (3) se tienen dos incógnitas,  $(u, v)$ , por lo cual se hace necesarias otras restricciones para calcular el flujo óptico.

### 2.2.2.1. MÉTODO DE HORN & SCHUNK

El método de Horn & Schunck necesita de restricciones como una variación lenta del movimiento en la imagen, para lo cual se supone que las velocidades de objetos vecinos son

prácticamente idénticas. Se realiza una restricción de suavidad a la ecuación (3), de tal forma se minimiza una suma ponderada del término de suavidad (factor de la izquierda de la ecuación (4)) y de la expresión de la restricción del cambio de la luminancia (factor de la derecha de la ecuación (4)) [20,22].

$$\iint_{\text{sobre imagen}} \left( \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 + \left( \frac{\partial v}{\partial x} \right)^2 + \left( \frac{\partial v}{\partial y} \right)^2 \right) dx dy + \beta \iint_{\text{sobre imagen}} (\nabla I \cdot \vec{u} + I_t) dx dy \quad (4)$$

$\beta$ : El factor de peso de la importancia relativa del suavizado para adecuarse a los datos.

La realización física de esta técnica usa los valores de estas variables en un punto (i,j) y sus vecinos, aproximando la desviación de la suavidad ecuación (5). El error de la restricción de flujo óptico está dado por la ecuación (6).

$$S_{ij} = \frac{1}{4} \left[ (u_{i+1,j} - u_{i,j})^2 + (u_{i,j+1} - u_{i,j})^2 + (v_{i+1,j} - v_{i,j})^2 + (v_{i,j+1} - v_{i,j})^2 \right] \quad (5)$$

$$C_{ij} = (I_x u_{i,j} + I_y v_{i,j} + I_t)^2 \quad (6)$$

Los valores del conjunto  $u_{i,j}, v_{i,j}$  deben minimizar por lo tanto se tiene:

$$E = \sum_i \sum_j S_{ij} + \gamma C_{ij} \quad (7)$$

$\gamma$ : Constante de regularización.

$$\frac{\partial E}{\partial u_{kl}} = 2(u_{kl} - \bar{u}_{kl}) + 2\gamma(I_x u_{kl} + I_y v_{kl} + I_t)I_x \quad (8)$$

$$\frac{\partial E}{\partial v_{kl}} = 2(v_{kl} - \bar{v}_{kl}) + 2\gamma(I_x u_{kl} + I_y v_{kl} + I_t)I_y \quad (9)$$

$\bar{u}, \bar{v}$ : Medias locales de  $u, v$ .

$$(1 + \gamma I_x^2)u_{kl} + \gamma I_x I_y v_{kl} = \bar{u}_{kl} - \gamma I_x I_t() \quad (10)$$

$$(1 + \gamma I_y^2)v_{kl} + \gamma I_x I_y v_{kl} = \bar{v}_{kl} - \gamma I_y I_t() \quad (11)$$

Las ecuaciones (10) y (11) resultan de igualar las ecuaciones (8) y (9) a cero.

$$u_{kl}^{n+1} = \bar{u}_{kl}^n - \frac{I_x \bar{u}_{kl}^n + I_y \bar{v}_{kl}^n + I_t}{1 + \gamma(I_x^2 + I_y^2)} I_x \quad (12)$$

$$v_{kl}^{n+1} = \bar{v}_{kl}^n - \frac{I_x \bar{u}_{kl}^n + I_y \bar{v}_{kl}^n + I_t}{1 + \gamma(I_x^2 + I_y^2)} I_y \quad (13)$$

La ecuación (12) y (13) hace referencia a los nuevos valores de  $(u, v)$ . El primer termino de la derecha son los valores que rodean a  $(u, v)$ , y el segundo termino es un valor de ajuste en la dirección del gradiente de la luminosidad. Los valores de  $I_x, I_y$  e  $I_t$ , requiere de una estimación, calculada a partir de medias locales en vecindades alrededor de cada punto de la imagen [20].

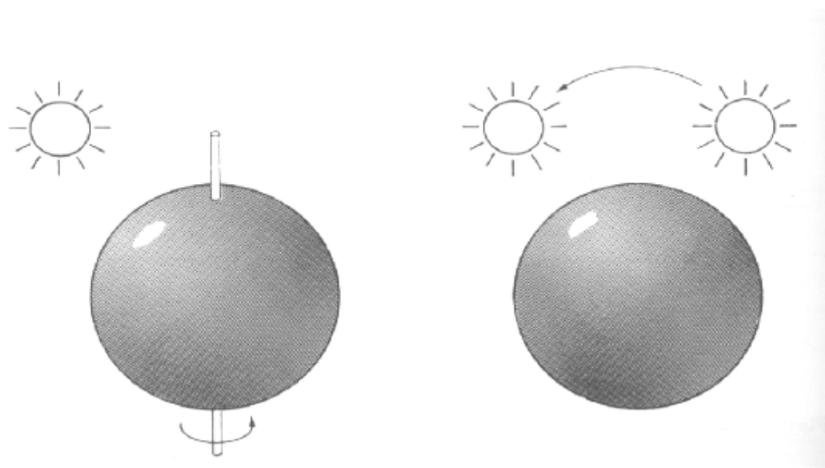
Horn & Schunck propusieron un algoritmo iterativo, que encuentra los valores de flujo óptico que satisfacen la variación suave del campo de flujo y el movimiento de los bordes debe mantenerse compatible con los gradientes espaciales y temporales del nivel de gris [20].

#### 2.2.2.2. MÉTODO DE LUCAS & KANADE

El método de Lucas & Kanade considera al flujo localmente constante y lo determina con una solución de mínimos cuadrados con ponderación. Extrae los dos componentes de la velocidad utilizando una función ventana, la cual obtiene los valores de los gradientes espaciales y temporal del entorno del punto en estudio [19,20].

El flujo óptico es un buen método para determinar el movimiento, es robusto si se minimizan las condiciones en donde pueden haber problemas, y tiene gran importancia porque puede determinar con mucha exactitud la dirección del movimiento, lo cual es indispensable para esta aplicación.

La detección del flujo óptico tiene una dependencia de la intensidad de los niveles de gris de la imagen, esto crea problemas en la determinación del movimiento [23]. A continuación se describen dos casos que generan error en la detección del movimiento por iluminación y por el movimiento sin cambios en los niveles de gris de la imagen:



**Figura 2.** Detección falsa de flujo óptico por movimiento e iluminación, Imagen

Fuente: Molina Rafael. Campo de movimiento en el plano de la imagen: Flujo óptico y correspondencia, 2003.

En el primer caso hay movimiento rotativo de una esfera con iluminación constante (ver Figura 2. Izquierda), esto hace que no exista variación en la intensidad de los niveles de gris en las imágenes consecutivas, lo cual implica un falso positivo, debido a que hay movimiento pero el flujo óptico no lo detecta [23].

En el segundo caso hay una esfera sin movimiento pero se cambia la iluminación sobre la esfera (ver Figura 2. Derecha), esto hace que se genere un cambio en la intensidad de los

niveles de gris en las imágenes consecutivas y se de un positivo falso, ya que no hay movimiento pero por flujo óptico si detecta movimiento en la esfera [23].

El algoritmo de flujo óptico tiene una velocidad de procesamiento bajo y tiene dependencia del tamaño de la imagen lo que genera problemas en la implementación de aplicaciones en tiempo real.

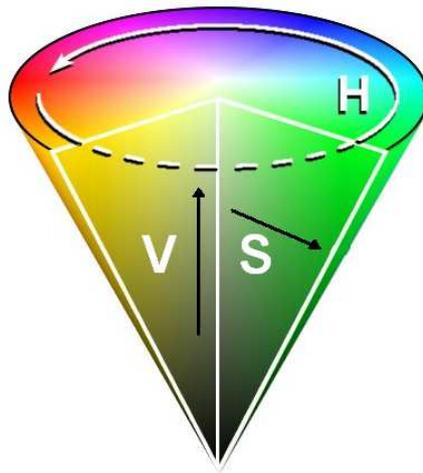
### **2.2.3. MÉTODO POR COLOR**

El análisis del color es importante para muchas aplicaciones entre las cuales está la detección de movimiento por color, es indispensable la caracterización del objeto a detectar y el entorno o fondo en el que se mueve, es así como se puede analizar la viabilidad del método por color. En este proyecto se busca la detección de la cabeza del usuario sus movimientos y el reconocimiento de gestos del rostro, por lo tanto es adecuado analizar la detección de movimiento por color de piel.

Dependiendo del tipo de sensor y aplicación se han desarrollado diversos espacios de color basados en propiedades tales como brillo, tono, coloración, luminosidad, croma, entre otros [24].

#### **2.2.3.1. ESPACIO HSV**

El espacio de color HSV corresponde por sus siglas H, S y V a tono (Hue), saturación (Saturation) y valor (Value) respectivamente. El espacio de color se puede representar en un cono de color como se ve en la Figura 3, donde el tono se mide como el ángulo alrededor del eje S, situándose el color rojo a 0°, el color verde a 120° y el color azul a 240°, la saturación tiene el rango de valores desde 0 hasta 1, al igual que el valor [25]. La dirección de aumento de los rangos de los canales se ve en la Figura 3.



**Figura 3.** Espacio de color HSV.

El espacio HSV es buen método para reconocer la piel, porque el tinte (H) de la piel para la mayoría de las razas se sitúa alrededor del naranja, la saturación (S) suele ser baja, el valor (V) es poco adecuado por la dependencia que tiene con la iluminación y los ajustes de la cámara [26].

### **2.2.3.2. ESPACIO YCbCr**

El espacio YCbCr es usada comúnmente por los estudios de televisión europeos y en la compresión de imágenes. El YCbCr es importante porque el color está representado por la luminancia (Y) y por (Cb y Cr) que son características colorimétricas del color [25].

La separación de las componentes de luminancia y de crominancia del color hace que este método sea viable para la detección del color de la piel, con este método se pueden hacer arreglos para modelar el color de partes específicas del rostro.

Luego de detectar la piel humana diferenciándola del fondo se procede a identificar la posición en la cual se encuentra dentro de la imagen, esto se realiza en todas las imágenes del video para que se realice el seguimiento del bloque de piel que se desee.

El método de color tiene inconvenientes por la variación que presenta el color por la iluminación, las sombras, la variación de color de piel de acuerdo a la raza, la barba, etc. y para ésta aplicación el método tiene que soportar variaciones de iluminación y es preferible que soporte a un usuario con barba, lo que disminuye el nivel de viabilidad. En los algoritmos que se implementaron con esta técnica se lograron buenos resultados cuando la iluminación era muy constante y cuando el color de piel de los usuarios no tenía una variación notable.

# CAPÍTULO 3

---

---

## DETECCIÓN DE LA CABEZA

---

---

La detección de la cabeza es importante porque en este proceso se empieza por reducir información, delimitando el área de análisis para los siguientes procesos y obteniendo más rápidamente los parámetros que posteriormente controlan el *mouse*. La detección de la cabeza del usuario se realiza mediante la utilización de una serie de filtros y mediante el método de sustracción de fondo, logrando separar la cabeza del usuario del fondo de la imagen.

### 3.1. FILTROS

El sistema de filtrado en una imagen tiene varios propósitos entre los que están la eliminación de ruido y el realce de patrones específicos. Para la utilización de filtros en el procesamiento digital de imágenes se debe tener en cuenta la aplicación que se va a desarrollar, debido a que la imagen filtrada tiene menos información que la original, pero adquiere mayor importancia porque los datos que quedan, contribuyen con el propósito específico de la aplicación [27].

Los métodos para mejorar la imagen se pueden dividir en dos campos diferentes: métodos en el dominio de la frecuencia y métodos en el dominio espacial. En el primero se utilizan mecanismos matemáticos para pasar del espacio de la imagen al espacio de las frecuencias, donde se analiza las frecuencias altas y bajas para eliminar la que se desee, pero los costos computacionales de estos filtros pueden ser altos, disminuyendo la velocidad de computo. En el segundo método se realizan operaciones sobre los píxeles de la imagen modificando sus valores de acuerdo a los píxeles cercanos y al proceso específico que se realice sobre ellos [28].

En el filtrado espacial está el concepto de frecuencia espacial, que especifica la magnitud de cambios de los datos por unidad de distancia, en una región de la imagen. Los valores de los píxeles de una región de la imagen con pequeños cambios o transiciones graduales, se definen como regiones de bajas frecuencias. Regiones con grandes cambios o rápidas transiciones se determinan como regiones de altas frecuencias [28].

### 3.1.1. FILTRO GAUSSIANO

Una distribución gaussiana con desviación típica  $\sigma$  y media  $\mu$  se describe en la ecuación (14), ahora si hay una convolución de la función  $g_{\sigma}(x)$  con una señal  $f(x)$  se obtiene una nueva señal suavizada  $h(x)$ , donde el valor en cada punto es el resultado de promediar con distintos pesos los valores vecinos a ambos lados de dicho punto [29].

$$g_{\sigma}(x) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \quad (14)$$

En la ecuación (15) se ha extendido la función a dos dimensiones y se ha supuesto que la campana de gauss se centra sobre la media. En la ecuación (16) se tiene en cuenta el caso discreto con la imagen  $f(i, j)$  [29].

$$g_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2} \left( \frac{x^2+y^2}{\sigma^2} \right)} \quad (15)$$

$$F(i,j) = \frac{1}{2\pi\sigma^2} \sum_k \sum_l e^{-\frac{1}{2}\left(\frac{k^2+l^2}{\sigma^2}\right)} f(i+k, j+l) \quad (16)$$

El filtro gaussiano reduce el ruido en la imagen y además se obtiene el efecto de suavizado, características muy vitales en el proyecto. En la Figura 4. Se puede apreciar una imagen real, la cual se ha convertido a escala de grises y posteriormente se ha sometido a un filtro gaussiano.



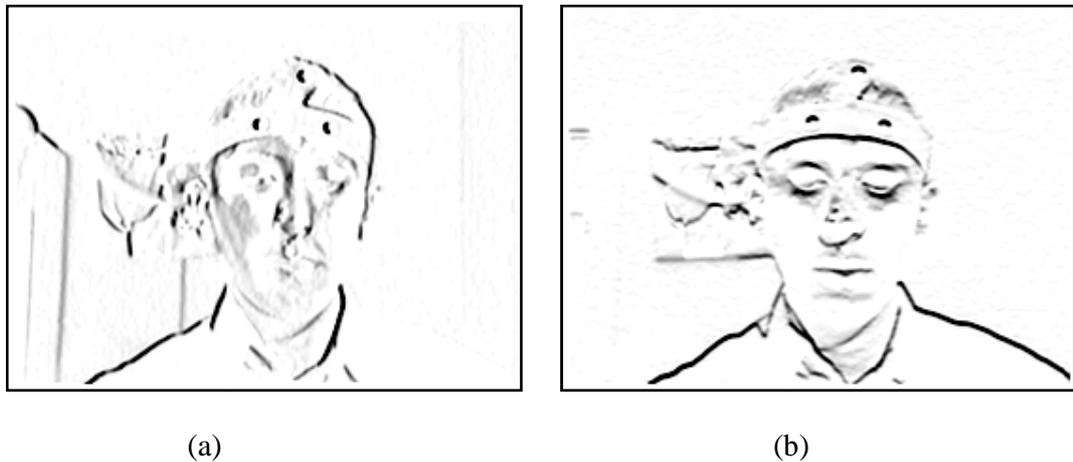
**Figura 4.** Imagen en escala de grises con el filtro gaussiano.

### 3.2. DETECCIÓN DE BORDES

Las técnicas de detección de bordes extraen información de los límites entre los objetos y el fondo de la imagen o los límites entre los mismos objetos, basándose en la localización de los puntos en los que se produce una variación de intensidad. Los métodos para encontrar los bordes se basan en operadores de derivada, utilizando específicamente la primera derivada (gradiente) y la segunda derivada (laplaciana). En el primer método se buscan grandes picos y en el segundo se buscan pasos de respuesta positiva a negativa o viceversa [24].

El operador de Sobel es una técnica basada en el gradiente, calculando la aproximación al gradiente de la función de intensidad de la imagen. Otra propiedad es que la diferencia entre píxeles vecinos de acuerdo a la relación puede dar lugar a derivadas unidimensionales o bidimensionales y se puede aplicar en alguna dirección determinada de la imagen [24].

En la Figura 5. Se puede ver una imagen real después de aplicarle el operador de Sobel en dos direcciones distintas. En las imágenes se disminuye la información del fondo uniforme, para facilitar el procesamiento en las siguientes etapas y se empieza a caracterizar la proyección de la cabeza del usuario que es finalmente el objetivo.



**Figura 5.** Operador de Sobel. (a) Dirección horizontal. (b) Dirección vertical.

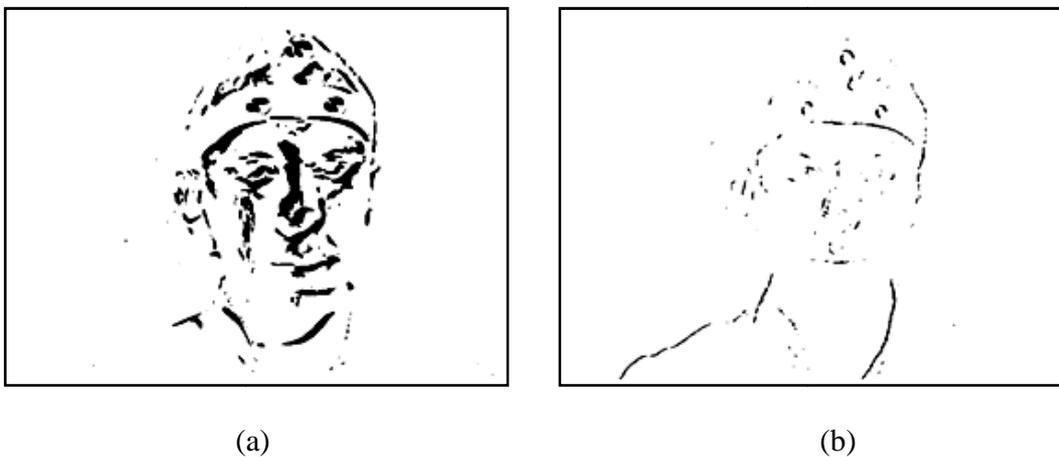
### 3.3. SUSTRACCIÓN DE FONDO

Como se explico anteriormente la sustracción de fondo es una técnica para eliminar el fondo de un objeto en movimiento, a pesar de que el método tiene problemas con el cambio de iluminación en imágenes seguidas, se ha buscado la forma de minimizar estos errores y aumentar la viabilidad. Su bajo costo computacional es una ventaja importante en aplicaciones en tiempo real como la que se hace en este proyecto.

Se ha creado un buffer que almacena tres imágenes seguidas, para luego aplicar un proceso

de sumas y restas entre las imágenes, hasta lograr potencializar la cabeza del usuario con un movimiento leve. Las imágenes a cargar en el buffer han sido filtradas y se les aplicó el proceso de realce de bordes para facilitar aun más la detección de la cabeza.

Con este método se ha logrado disminuir el ruido por brillo de zonas del fondo por materiales sensibles a la reflexión de la luz. Se tiene dificultades si se genera un movimiento detrás del usuario, debido a que este puede ser detectado y generar una falsa ubicación de la cabeza del usuario.



**Figura 6.** Sustracción de fondo de la cabeza del usuario. (a) Aumento de movimiento. (b) Disminución de movimiento.

En la Figura 6. Se tiene dos imágenes reales luego de la sustracción de fondo, en donde se puede apreciar la diferencia cuando el usuario genera un movimiento rápido con la cabeza (ver Figura 6. (a)) y cuando el movimiento es lento (ver Figura 6. (b)). Aunque la cabeza no va a estar en constante movimiento y en ocasiones estará totalmente quieta perdiendo información sobre su ubicación, esta técnica es viable porque en la pruebas se comprobó que movimientos muy leves logran la detección y se hicieron arreglos en los algoritmos para que el sistema detecte la cabeza cuando no se mueve, basado en las coordenadas anteriores.

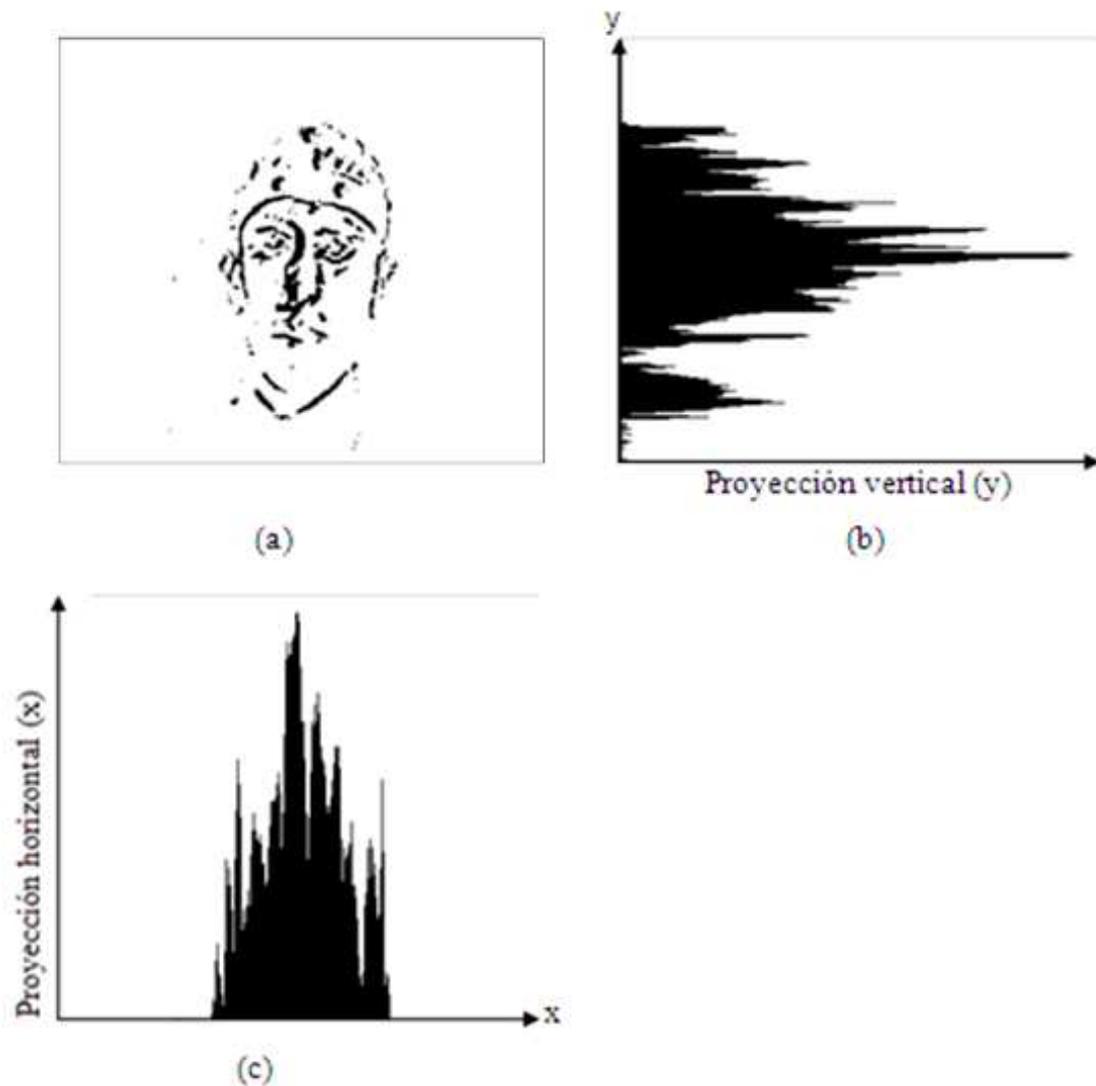
### 3.4. INTEGRALES PROYECTIVAS

En la visión artificial en algunas aplicaciones se dificulta el análisis de imágenes debido al manejo de espacios de elevada dimensionalidad [30]. Los procesos con grandes cantidades de valores tienen un elevado costo computacional y se pierde capacidad de análisis en tiempo real. A través de todo el procesamiento de imágenes se han implementado técnicas para reducir el volumen de datos sin perder la información vital.

Es indispensable la reducción a sub-espacios de menor dimensionalidad para optimizar el procesamiento y en particular el de caras humanas. Son importantes los sub-espacios lineales, los cuales se obtienen a través de transformaciones lineales sobre las imágenes, donde cada valor de salida es una combinación lineal de todos los píxeles de entrada. Algunos ejemplos son: las autocaras, las fishercaras, el análisis de componentes independientes, los filtros *wavelets* de Haar y Gabor, las integrales proyectivas, etc [30].

En las integrales proyectivas cada valor de salida es la media aritmética de una fila o columna de píxeles de la entrada, con lo cual es evidente la reducción de datos en la imagen, además las aplicaciones que se han realizado con esta técnica demuestra la robustez frente al ruido. Suponiendo que el ruido es aditivo y aleatorio, las proyecciones tienen el efecto de disminuir la varianza del ruido sin afectar a la forma de las señales [30].

La técnica de integrales proyectiva es vital en este proyecto para determinar las coordenadas de la cabeza del usuario en base a la imagen, es por esto que la integral proyectiva se aplica a una imagen previamente procesada y binaria, pero con información del movimiento de la cabeza como se ve en la imagen real de la Figura 7. (a), y las integrales proyectivas vertical y horizontal en la Figura 7. (b) y (c) respectivamente.



**Figura 7.** Proyecciones. (a) Imagen a obtener las proyecciones. (b) Proyección vertical. (c) Proyección horizontal.

En las integrales proyectivas aplicadas para la detección de la cabeza se obtienen las coordenadas de un sub-espacio rectangular donde se encuentra la cabeza del usuario. Es necesario implementar el umbral que determina a partir de que rango del número de píxeles con una intensidad de 1, se encuentra la cabeza del usuario, disminuyendo el efecto del ruido.

### 3.5. IMPLEMENTACIÓN DEL ALGORITMO

Para la implementación del algoritmo se tuvieron en cuenta las ventajas y desventajas de los diferentes métodos y los resultados de pruebas realizadas con imágenes reales, de tal forma que las decisiones tomadas para la implementación cumplieran con la detección de la cabeza del usuario.

Para iniciar la aplicación es necesaria la captura de las imágenes provenientes de la cámara webcam y garantizar el tamaño de 640x480 píxeles de las imágenes, ya que es preciso para mantener un bajo costo computacional y garantizar la calidad suficiente para diferenciar los detalles de la cabeza y específicamente el rostro.

Se implementó un filtro gaussiano que minimiza el ruido de la imagen y contribuye con el suavizado siendo este muy conveniente en los procesos siguientes. Como en la detección de bordes basado en el operador Sobel, al obtener buenos resultados con el ajuste del operador en dos direcciones, horizontal y vertical como se puede ver en la Figura 5. (a) y (b) respectivamente. Para aumentar la definición de los bordes de la cabeza se sumaron las imágenes del operador Sobel en las dos direcciones horizontal y vertical (ver Figura 8.).

Aunque los métodos de color permiten ubicar la cabeza del usuario sin movimiento conllevando a muchas ventajas para la aplicación, tienen muchas dificultades con las condiciones de iluminación, puede tener problemas con los fondos de cierto color que son muy parecidos al color característico de la piel y presentan problemas de robustez con parámetros particulares de algunas personas como la barba. Otro método es el de flujo óptico el cual tiene un costo computacional muy alto con imágenes grandes y presenta algunos problemas con brillos del fondo dando positivos falsos. Así estos dos métodos no son óptimos para la implementación de algoritmos para la detección de la cabeza.



**Figura 8.** Suma de los filtros de Sobel en la dirección horizontal y vertical.

Para la detección de la cabeza del usuario se implemento un arreglo de la sustracción de fondo, basado en un buffer que almacena tres imágenes y luego realiza sumas y restas entre ellas, obteniendo una mayor caracterización del movimiento del usuario como se puede ver en la Figura 6, que cuando se aplica una sola resta de imágenes. Inicialmente el usuario deberá mover la cabeza para que el sistema detecte su posición, luego no es necesario, debido a que la aplicación se basa en las coordenadas anteriores para ubicar la cabeza cuando el usuario se queda quieto.

Finalmente se realizan integrales proyectivas para la identificación de las coordenadas del rectángulo que rodea la cabeza del usuario (ver Figura 7). De tal forma se obtiene el sub-espacio de la cabeza (ver Figura 9.). Lo cual es vital para minimizar la cantidad de datos a utilizar en los procesos siguientes debido a que no es necesario tener en cuenta toda la imagen.



**Figura 9.** Detección de la cabeza del usuario.

# CAPÍTULO 4

---

---

## MOVIMIENTOS DE LA CABEZA

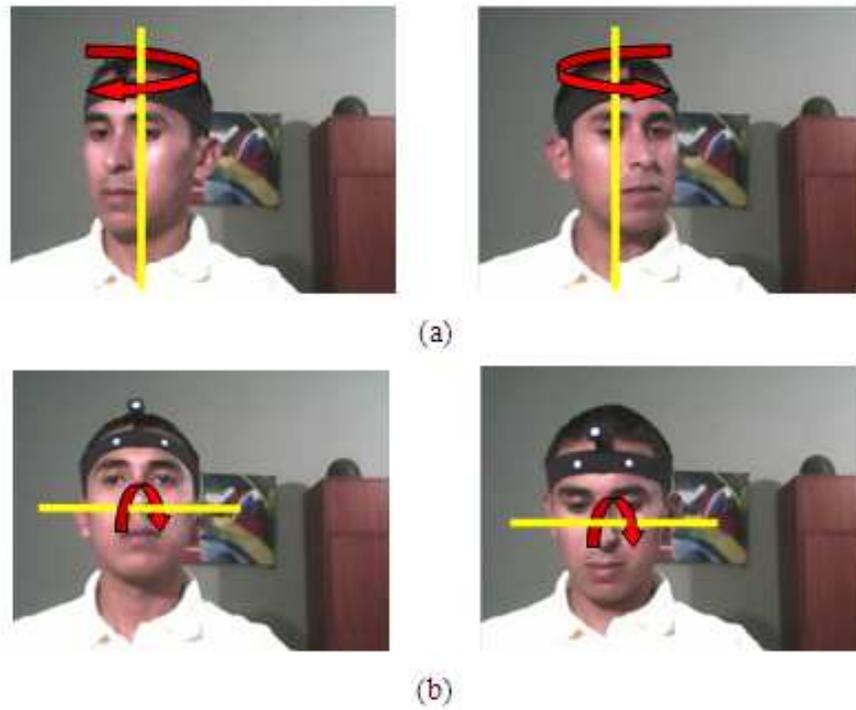
---

---

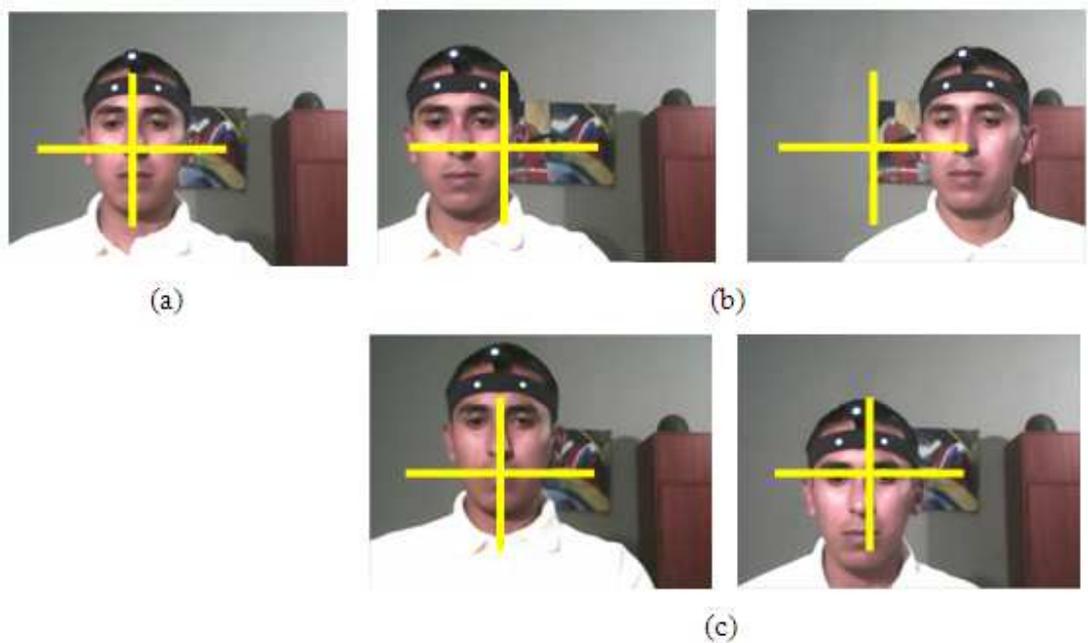
La cabeza humana realiza movimientos en diferentes direcciones lo cual hace que sea un sistema de movimiento con varios grados de libertad, pero en este proyecto se busca la detección del movimiento del giro horizontal, específicamente izquierda, derecha y vertical, específicamente arriba, abajo (ver Figura 10.). Además hay que tener en cuenta que es diferente el movimiento que puede realizar toda la cabeza hacia alguna de las direcciones horizontal y vertical (ver Figura 11.).

Para facilitar la correcta detección del movimiento de la cabeza se diseñó un sistema de tres puntos con el cual se obtuvieron buenos resultados, incluyendo tres grados de libertad más de los requeridos. Este sistema se explica en el capítulo incluyendo su construcción, modo de uso, y toda la caracterización necesaria para su correcto funcionamiento dentro del proyecto.

Finalmente se obtienen los parámetros que caracterizan el movimiento de giro por parte de la cabeza, los cuales son necesarios para el correspondiente movimiento del *mouse*.



**Figura 10.** Movimiento de giro de la cabeza. (a) Horizontal. (b) Vertical.



**Figura 11.** Movimiento de la cabeza. (a) Posición inicial. (b) Horizontal. (c) Vertical.

#### 4.1. SISTEMA DE TRES PUNTOS

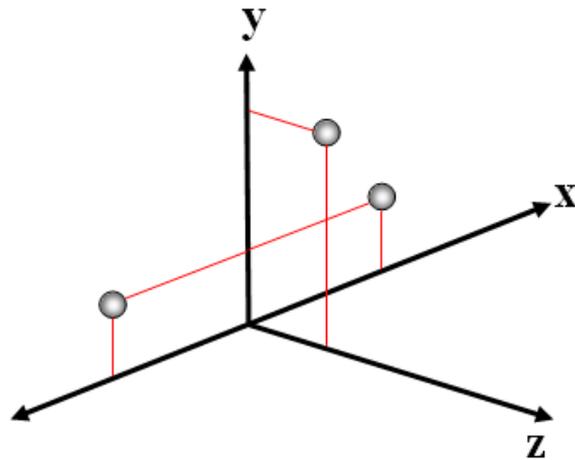
Se diseñó un sistema de tres puntos (ver Figura 12.), que permite la detección del movimiento de la cabeza y ayuda como parámetro guía para la ubicación de las coordenadas de partes del rostro como los ojos, cejas y boca. El sistema es vital para diferenciar el movimiento de giro (ver Figura 10.), del movimiento de toda la cabeza en una dirección determinada (ver Figura 11.). El sistema lo conforman dos elementos, uno es una banda y el otro es un elemento en L, los cuales en conjunto logran un sistema de tres puntos blancos sobre fondo negro, donde la variación de las distancias entre ellos caracteriza diferentes movimientos de la cabeza.



**Figura 12.** Sistema de tres puntos.

La banda de tela negra con dos puntos blancos de referencia, funciona como un sistema de dos puntos en el mismo plano del rostro, mientras que el elemento L tiene un punto blanco sobre un cuadro cubierto con tela negra, que está ubicado a tres centímetros del rostro y en un plano paralelo a éste. Por lo tanto, los tres puntos conforman un sistema tridimensional (ver Figura 13.), pero cuando, se captura la imagen éste se vuelve un sistema bidimensional, con la proyección de los tres puntos en un mismo plano.

El elemento con el punto a tres centímetros del rostro contribuye a la sensibilidad en la detección del movimiento de la cabeza, debido a que con movimientos pequeños se puede obtener variaciones medibles de las distancias entre los puntos, éstas medidas se realizan en la imagen bidimensional capturada por la webcam.



**Figura 13.** Diagrama tridimensional del sistema de tres puntos.

#### **4.1.1. CONSTRUCCIÓN DEL SISTEMA DE TRES PUNTOS**

La utilización de un sistema de tres puntos para el proyecto hace necesario que a continuación se especifiquen los materiales requeridos y se explique paso a paso la construcción del sistema. Los materiales y el método de construcción pueden variar de acuerdo a la creatividad del usuario, aquí solamente se dan las medidas que se deben respetar y se presenta un método de construcción con el cual se obtuvieron los resultados favorables para la aplicación.

##### **4.1.1.1. MATERIALES**

Los materiales utilizados para la construcción del sistema de tres puntos son fáciles de adquirir y tienen un bajo costo, facilitando la fabricación del sistema el cual es vital para la

aplicación. En la Tabla 3. Se especifican los materiales, las características, las dimensiones y cantidad requerida de cada uno de ellos para la construcción del sistema de tres puntos.

**Tabla 3.** Materiales utilizados para la fabricación del sistema de tres puntos.

<b>MATERIAL</b>	<b>CARACTERÍSTICAS</b>	<b>DIMENSIONES</b>	<b>CANTIDAD</b>
TELA NEGRA	La tela debe tener propiedades que eviten reflejar la luz (deberá ser opaca).	60.0 x 7.0 cm	1
		7.5 x 7.5 cm	1
CINTA DE VELCRO	La cinta de velcro debe ser negra y puede tener una medida de ancho inferior a 3.0 cm.	6.0 cm	1
LAMINA METÁLICA	La lamina metálica debe tener propiedades que permitan doblarla fácilmente en L.	6.0 x 0.8 x 0.1 cm	1
PAPEL GRUESO	El color debe ser negro.	7.0 x 2.0 cm	1
CINTA NEGRA		Rollo	1
CORRECTOR PARA TINTA			1

#### **4.1.1.2. PASOS DE CONSTRUCCIÓN DEL SISTEMA DE TRES PUNTOS**

La construcción del sistema se divide en dos partes, debido al número de componentes que lo conforman, empezando por la banda y luego por el elemento en L. La construcción no tiene procesos complicados con el fin de que el sistema sea asequible al usuario y con el fin de disminuir al máximo el tiempo de fabricación.

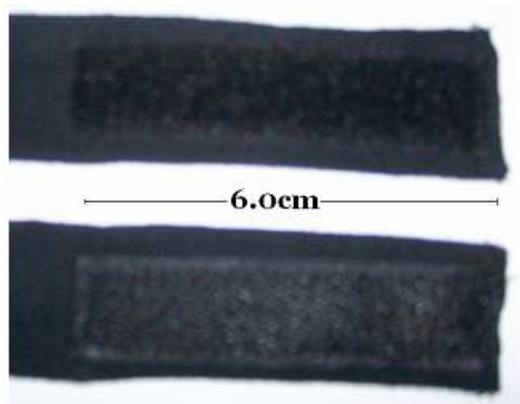
##### **4.1.1.2.1. BANDA**

Pasos para la construcción de la banda:

- La tela negra de dimensiones 60.0 x 7.0 cm se dobla para darle mayor consistencia, quedando con medidas de 60.0 x 3.0 cm, debido a que 1.0 cm se deja en superposición

para realizar una costura sobre la banda y así fijarla de forma permanente.

- La cinta de velcro se ubica en los extremos de la banda y se fija mediante una costura (ver Figura 14.). La cinta de velcro sirve para que la banda se sujete y adapte fácilmente a la cabeza del usuario.



**Figura 14.** Dimensiones de la cinta de velcro en los extremos de la banda del sistema de tres puntos.



**Figura 15.** Dimensiones de la banda del sistema de tres puntos.

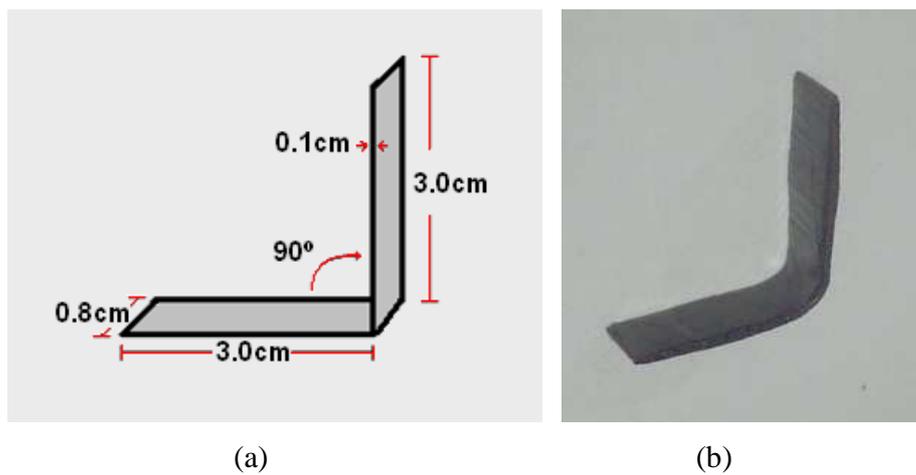
- En la parte central del a banda se dibujan dos puntos blancos con el corrector para tinta,

los puntos tienen que tener un diámetro de 0.6 cm cada uno y con una separación entre ellos de 6.5 cm. La banda terminada debe verse como en la Figura 15.

#### 4.1.1.2.2. ELEMENTO L

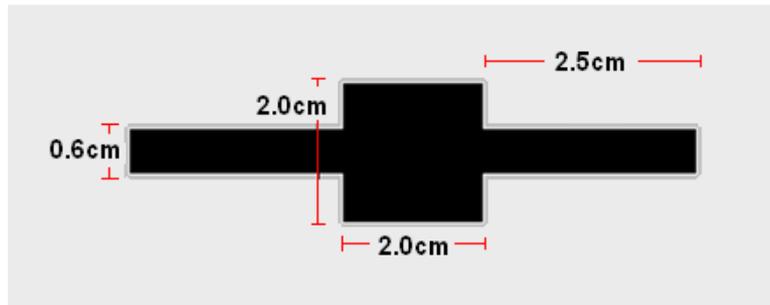
Pasos para la construcción del elemento en L:

- La lámina metálica se dobla por la mitad formando un ángulo de 90 grados, de tal forma que tenga las dimensiones de la Figura 16. (a).
- La lámina metálica se cubre con la cinta negra, para evitar el reflejo que puede producir el metal. (ver Figura 16. (b)).



**Figura 16.** Elemento L. (a) Diagrama con las dimensiones. (b) Cubierto con cinta negra.

- El papel grueso se recorta con las dimensiones de la Figura 17. (a) y debe verse como en la Figura 17. (b).



(a)



(b)

**Figura 17.** Frente en papel del elemento L. (a) Diagrama con las dimensiones.

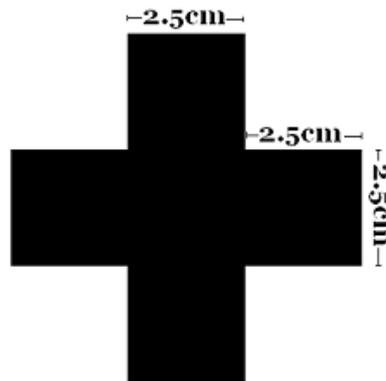
(b) Imagen Real.

- En uno de los extremos de la lámina se ubica el recorte de papel grueso, de tal forma que se cree un tablero en el frente como en la Figura 18.



**Figura 18.** Frente del elemento L

- El tablero del frente se cubre con la tela negra de dimensiones 7.5 x 7.5 cm y recortada de la forma que se muestra en la Figura 19.



**Figura 19.** Diagrama de dimensiones de la tela que cubre el Frente del elemento L.

- En la parte central del tablero y sobre la tela se pinta un punto blanco con el corrector de tinta con un diámetro de 0.6 cm. El elemento en L final debe verse como en la Figura 20.



**Figura 20.** Elemento L.

#### **4.1.2. MODO DE UTILIZAR EL SISTEMA DE TRES PUNTOS**

El sistema de tres puntos está diseñado para adherirse sobre la parte superior de la cabeza del usuario. La banda se ubica en la frente y se asegura con la cinta de velcro en la parte posterior de la cabeza, garantizando que los puntos estén centrados como se muestra en la Figura 21. El elemento L se asegura con la banda garantizando que la proyección del punto blanco quede, en la mitad de la distancia entre los dos puntos de la banda, como se muestra en la Figura 22.



**Figura 21.** Posición de la banda en la cabeza.

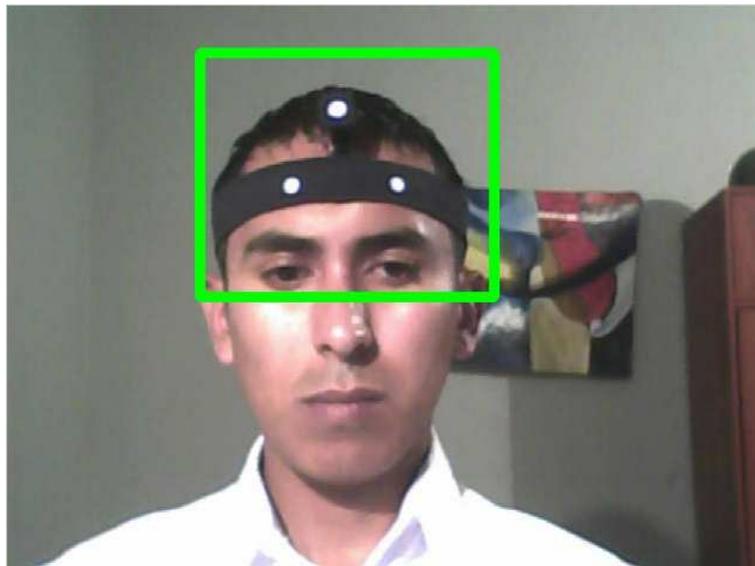


**Figura 22.** Posición del elemento L en la cabeza.

## 4.2. SEGMENTACIÓN

El objetivo de la segmentación es separar los objetos del fondo de la imagen y que se puedan distinguir entre si [24], en este caso es necesario separar los tres puntos blancos del resto de la imagen.

El proceso de la segmentación se puede basar en las siguientes tres propiedades: En primer lugar la similitud, donde cada uno de los pixeles de un elemento tiene valores parecidos para alguna propiedad, en segundo lugar la discontinuidad, donde los objetos se destacan del fondo de la imagen y tienen unos contornos definidos y en tercer lugar la conectividad, de tal forma que los pixeles pertenecientes al mismo objeto tienen que ser contiguos [24].



**Figura 23.** Zona de análisis del sistema de tres puntos.

En los procesos anteriores se determinó el rectángulo que delimita la cabeza del usuario, y si se tiene en cuenta que el sistema de tres puntos está ubicado en la parte superior de la cabeza del usuario, es posible determinar que el sistema siempre va a estar en la mitad superior del rectángulo. Así, es conveniente analizar solamente esta parte de la imagen, de

tal forma que se minimiza la cantidad de datos a ser procesados y se aumenta la velocidad de ejecución de los algoritmos. En la Figura 23. Se puede observar la subregión de la imagen que es tomada en cuenta para el análisis del sistema de tres puntos.

#### **4.2.1 DETECTOR DE CANNY**

El detector de Canny hace parte de los métodos de detección de bordes basados en las derivadas gaussianas, éste detector se basa en tres criterios los cuales son [24]:

**Error:** se deben detectar todos los bordes, evitando suministrar falsos bordes.

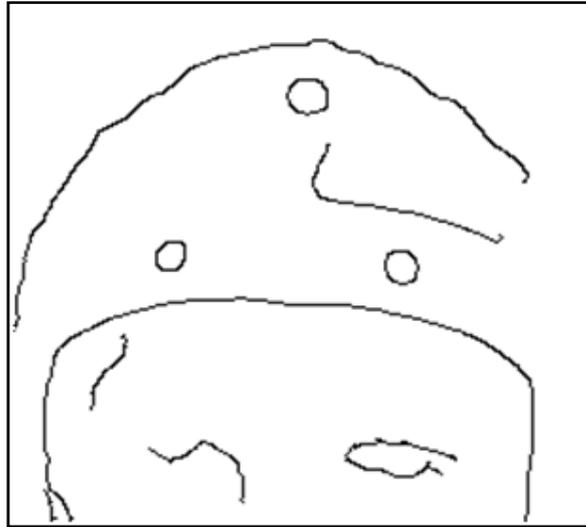
**Localización:** la distancia entre el pixel señalado como borde y el borde real se debe minimizar lo más que se pueda.

**Respuesta:** no se debe identificar varios pixeles como bordes cuando solo exista uno.

Para aplicar el detector es conveniente aplicar un filtro gaussiano para minimizar el ruido y la detección de falsos bordes, pero no se debe exceder en el suavizado debido a que puede perderse el detalle de los bordes.

El detector de Canny busca la detección de los bordes de los puntos blancos del sistema de tres puntos y como se puede observar en la Figura 24, a parte de los tres puntos se detecta otros bordes que suelen cambiar con cada imagen dependiendo de la iluminación, del movimiento del usuario etc. y para esta aplicación estos bordes se clasifican como ruido. En los procesos siguientes se busca la eliminación de los bordes que no pertenecen a los tres puntos del sistema.

Con el detector de Canny se obtuvieron buenos resultados en la detección de los bordes del sistema de tres puntos, el cual es vital para la caracterización del movimiento de la cabeza del usuario y el posterior movimiento del *mouse*.



**Figura 24.** Detector de Canny aplicado a la zona del sistema de tres puntos.

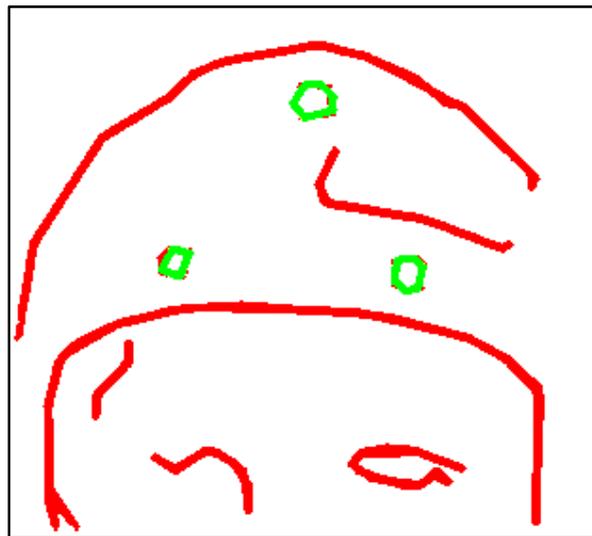
#### **4.2.2. CONTORNOS CERRADOS**

El proceso del detector de Canny deja muchos contornos que no pertenecen al sistema de tres puntos, por lo que se analizaron los contornos clasificados como ruido para esta aplicación en específico, y se encontraron características importantes para determinar el proceso a seguir en su respectiva eliminación.

La propiedad más evidente y constante es donde los contornos de los tres puntos son líneas cerradas (los extremos de la línea se unen), mientras que los contornos determinados como ruido, se presentan en la gran mayoría como líneas abiertas (los extremos de la línea no se unen) y rara vez se cierran.

Se hizo un análisis de las líneas de ruido que se presentan de forma cerrada originando falsos positivos y se logró determinar que la gran mayoría difieren en tamaño de los tres puntos del sistema, por lo cual se implemento un filtro de tamaño para garantizar la eliminación de contornos no deseados.

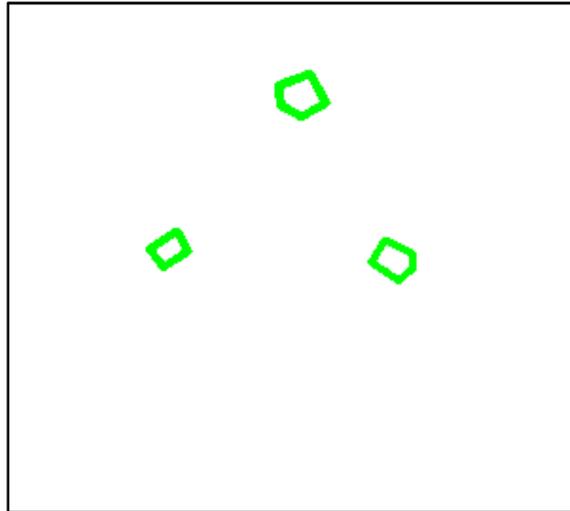
En la Figura 25. Se puede ver las líneas de color verde que identifican los contornos cerrados detectados de la imagen en la Figura 24. Correspondientes a los contornos de los tres puntos del sistema y los contornos del ruido en líneas de color rojo los cuales se proceden a eliminar.



**Figura 25.** Detección de contornos cerrados.

Los procesos de segmentación tienen características que hacen que sea dependiente a los cambios bruscos de iluminación, debido a que puede generarse ruido que afecte la ubicación de los tres puntos de sistema. Se debe evitar la utilización de objetos por parte del usuario en el área de análisis, que puedan producir contornos cerrados causando falsos positivos para el sistema.

Los procesos de la segmentación tienen un bajo costo computacional y se adaptaron bien a los requerimientos del sistema. En la Figura 26. Se observa la imagen final del sector de análisis para el sistema de tres puntos, después de la segmentación.



**Figura 26.** Segmentación del sistema de tres puntos.

### **4.3. DETECCIÓN DE LOS TRES PUNTOS**

Luego del proceso de segmentación es necesario determinar las coordenadas de la posición de los tres puntos del sistema en la imagen. Para este proceso se probaron dos métodos diferentes con el objetivo de identificar el más adecuado y robusto para implementarlo en la aplicación.

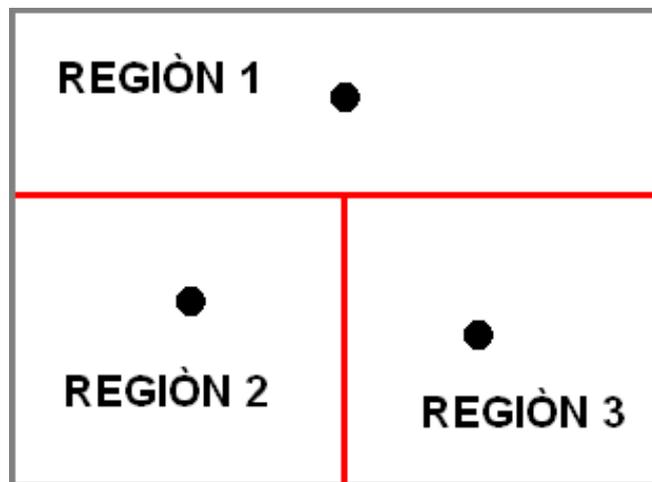
#### **4.3.1. INTEGRALES PROYECTIVAS**

La imagen luego de la segmentación tiene tres puntos bien definidos y separados entre sí, para lo cual, se puede aplicar el método de integrales proyectivas con el fin de identificar las coordenadas de los puntos, ya que anteriormente se obtuvieron buenos resultados con la detección de las coordenadas del rectángulo que delimita la cabeza del usuario con este método.

Para que el sistema de tres puntos sea viable se debe identificar las coordenadas y la posición de un punto con respecto a los otros dos. Así, se procede a realizar las integrales proyectivas en las direcciones horizontal y vertical de la imagen segmentada, con el fin de

conocer las coordenadas de los puntos que están de primeros en las direcciones: izquierda-derecha, derecha-izquierda, inferior-superior y superior-inferior, para determinar tres zonas de análisis.

Cuando están definidas las tres regiones que contienen los tres puntos del sistema (ver Figura 27.) se procede a realizar las integrales proyectivas horizontal y vertical de cada región, determinando las coordenadas y la posición de un punto con respecto a los otros dos.



**Figura 27.** Regiones para la identificación de puntos en el sistema de tres puntos.

Para la detección del sistema de tres puntos el método de integrales proyectivas tiene que aplicarse varias veces causando demoras en el proceso, pero las demoras se minimizan debido a que las regiones son pequeñas y es más rápido procesarlas.

En la Figura 28. Se tiene una imagen con la ubicación de los tres puntos con el método de integrales proyectivas, obteniendo buenos resultados, sin embargo el método presenta deficiencias si el proceso de segmentación deja pasar más de tres objetos, ya que el sistema puede arrojar datos erróneos.



**Figura 28.** Detección del sistema de tres puntos por el método de integrales proyectivas.

#### **4.3.2. SEGUIMIENTO DE PUNTOS**

Este método de seguimientos de puntos se puede dividir en dos etapas, una que extrae información sobre los objetos de la imagen y la otra que realiza el seguimiento.

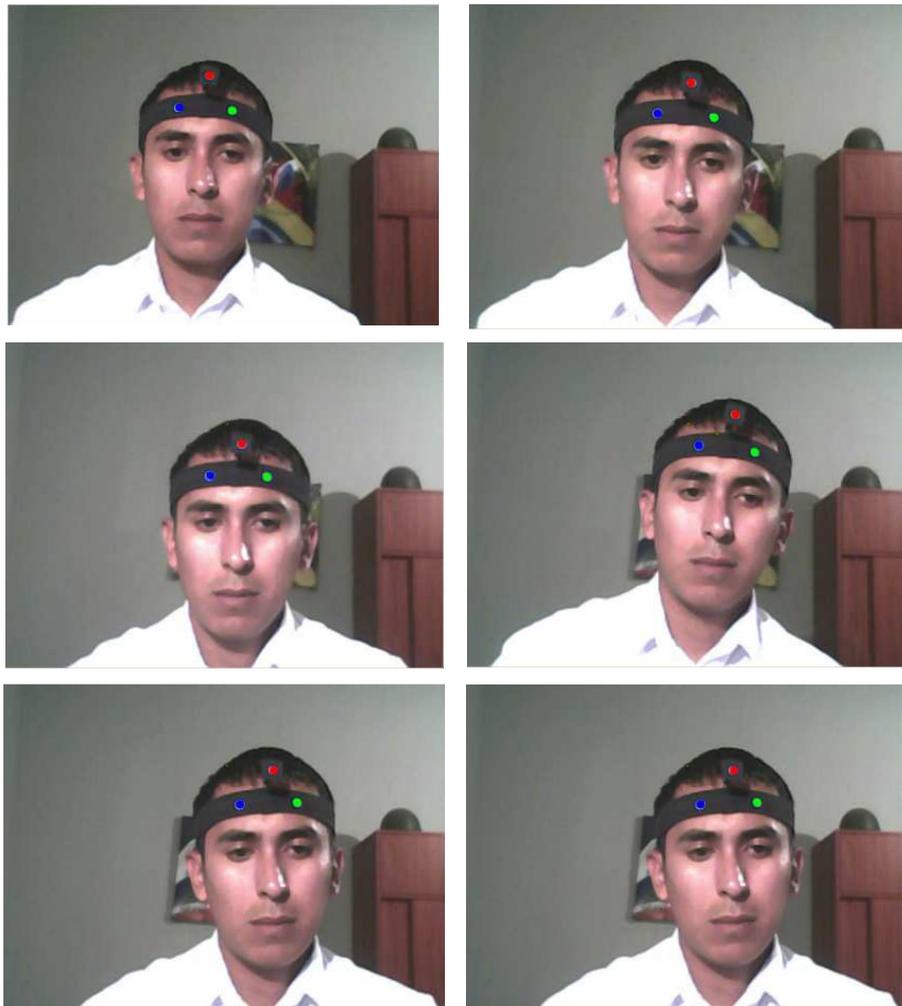
En la primera etapa se utiliza una imagen inicial en la cual se ubican todos los objetos de la imagen, y se obtienen características, como el área, perímetro y las coordenadas del rectángulo que delimita a cada objeto. También se etiquetan los objetos con el fin de realizar el seguimiento de cada uno.

En la segunda etapa se utiliza la imagen siguiente a la inicial, donde se obtienen las mismas características de la primera etapa, con la diferencia que para realizar el etiquetado se tienen en cuenta condiciones que garantizan el reconocimiento de los mismos objetos, luego del movimiento y cambio de posición en la segunda imagen con respecto a la primera.

El método es robusto debido a que si el proceso de segmentación deja pasar más de tres puntos por efecto de ruido el seguimiento de puntos no tiene problemas y si se llega a perder la ubicación de uno de los puntos el sistema reinicia el seguimiento. La velocidad

del proceso es buena y no presenta inconvenientes en tiempo real.

En la Figura 29. Se tiene una secuencia de seis imágenes reales en las cuales se detecto los tres puntos del sistema y se le realizo el respectivo seguimiento. El color de cada punto permite observar que se sigue a cada objeto sin equivocación.

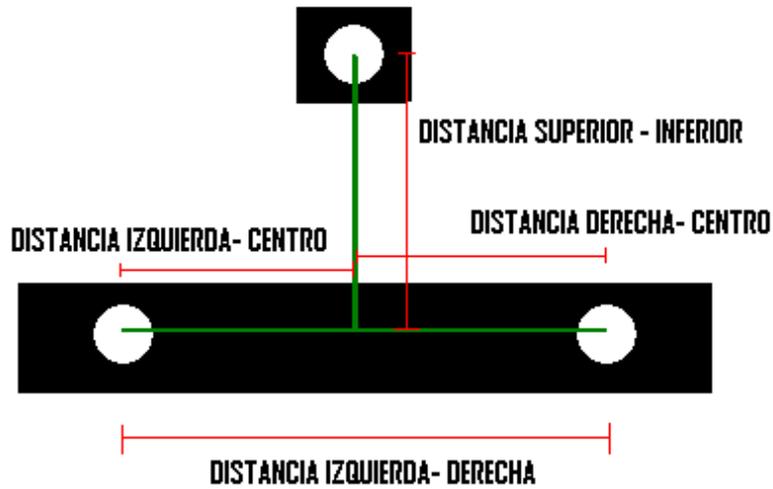


**Figura 29.** Secuencia de imágenes del seguimiento del sistema de puntos.

#### 4.4. DISTANCIAS ENTRE LOS PUNTOS

La distancia entre los puntos es vital para determinar la dirección y la cantidad de

movimiento efectuado por parte del usuario. Las distancias adecuadas para analizar se describen en la Figura 30. Donde se pueden identificar cuatro parámetros de distancia, suficientes para definir siete movimientos diferentes de la cabeza del usuario.



**Figura 30.** Diagrama de distancias del sistema de tres puntos.

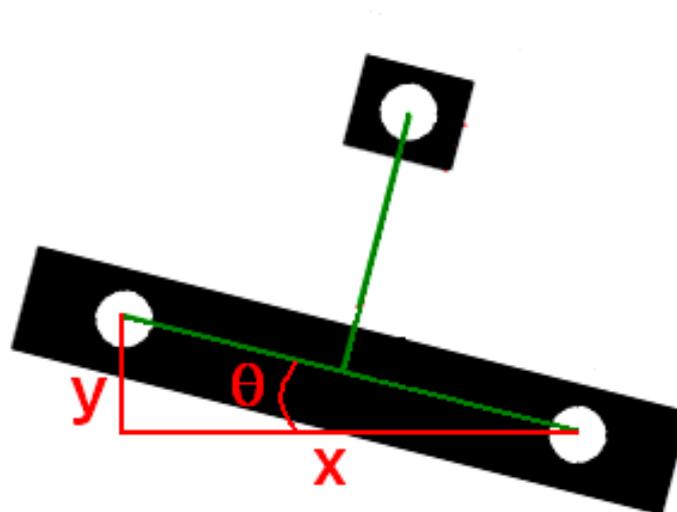
Cuando se inicializa el sistema se toman los datos de las distancias como los parámetros iniciales y a partir de éstos se realizan todas las comparaciones posteriores. De cada imagen se obtienen los parámetros de distancia y se comparan con los datos obtenidos inicialmente; en la Tabla 4. Se describe la dirección del movimiento de giro de acuerdo a la variación de distancia.

Cuando el usuario se acerca o aleja de la cámara también puede ser detectado por el sistema, de tal forma que si la distancia superior-inferior e izquierda-derecha aumentan proporcionalmente, es la respuesta al movimiento del usuario acercando la cabeza a la cámara y si por el contrario disminuyen proporcionalmente es debido a que el usuario se aleja de la cámara.

**Tabla 4.** Respuesta de los parámetros de distancia al movimiento de giro de la cabeza del usuario.

DISTANCIA	Izquierda-Centro:	Derecha-Centro:	Superior-Inferior:
<b>MOVIMIENTO DE GIRO DE LA CABEZA DEL USUARIO</b>			
Derecha	Aumenta	Disminuye	x
Izquierda	Disminuye	Aumenta	x
Arriba	x	x	Aumenta
Abajo	x	x	Disminuye

En el caso de inclinación de la cabeza por parte del usuario, el sistema logra determinar el cambio con respecto a la inclinación inicial y por otra parte obtiene las distancias entre los puntos como se grafica en líneas verdes en la Figura 31.



**Figura 31.** Diagrama del sistema de tres puntos con inclinación.

La inclinación de la cabeza se determina por el ángulo  $\theta$  señalado en la Figura 31 y su valor se establece mediante la ecuación (17). Parámetro importante para determinar alguna opción de control, sobre el programa final.

$$\theta = \tan^{-1} y/x \quad (17)$$

Donde  $y$  está dado por el valor absoluto de la resta entre las coordenadas  $y$  de los puntos inferiores y  $x$  de las coordenadas en  $x$ .

#### **4.5. IMPLEMENTACIÓN DEL ALGORITMO**

La implementación del algoritmo inicia con la determinación de la región de análisis de acuerdo a rectángulo que delimita la cabeza del usuario, por lo cual se escoge la mitad superior ya que el sistema de tres puntos siempre se va a encontrar en este sector.

Para el proceso de segmentación se realiza un suavizado bajo los criterios de no afectar los contornos de interés y que exista una disminución del ruido en la imagen. Posteriormente se hace un arreglo con los algoritmos del detector de Canny y el de contornos cerrados con el fin de lograr la segmentación requerida (los tres puntos del sistema) para los otros procesos.

Luego de probar y analizar los dos métodos para encontrar las coordenadas de los tres puntos, se optó por implementar el seguimiento de puntos, debido a que demostró ser más robusto que las integrales proyectivas, ya que requiere menos condiciones específicas de iluminación.

Para obtener las distancias se implementó un arreglo trigonométrico de tal forma que las variaciones de distancia no se afectaran por la inclinación de la cabeza, además se hizo un algoritmo que no compara las distancias como tal, sino porcentajes entre éstas, con lo cual se elimina el error que puede presentarse cuando el usuario se aleja o acerca de la cámara webcam.

# CAPÍTULO 5

---

---

## DETECCIÓN DE LOS GESTOS DEL ROSTRO

---

---

Los gestos del rostro forman parte de los métodos de comunicación del ser humano, con ellos se puede expresar mucha información. Es una forma natural de expresión con los cuales se dan a conocer estados de ánimo, se representan palabras concretas o simplemente acompañan a la comunicación verbal.

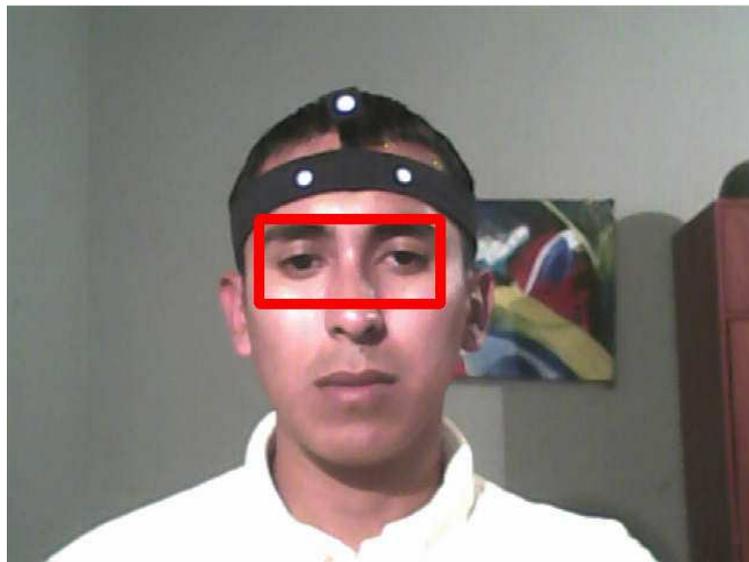
En este proyecto se ha implementado la detección de gestos que sean fáciles de realizar por parte del usuario, como los gestos que se basan en los movimientos de los ojos, cejas y boca.

### 5.1. GESTOS DE LOS OJOS

El movimiento de los ojos de cierre y apertura, es muy natural para el ser humano debido al parpadeo (mecanismo involuntario pero que varía la frecuencia en cada persona). El gesto en los ojos a detectar es el de cierre y apertura, teniendo en cuenta que el parpadeo es muy rápido, por lo cual se puede diferenciar del movimiento consiente y con tiempo de cierre más prolongado.

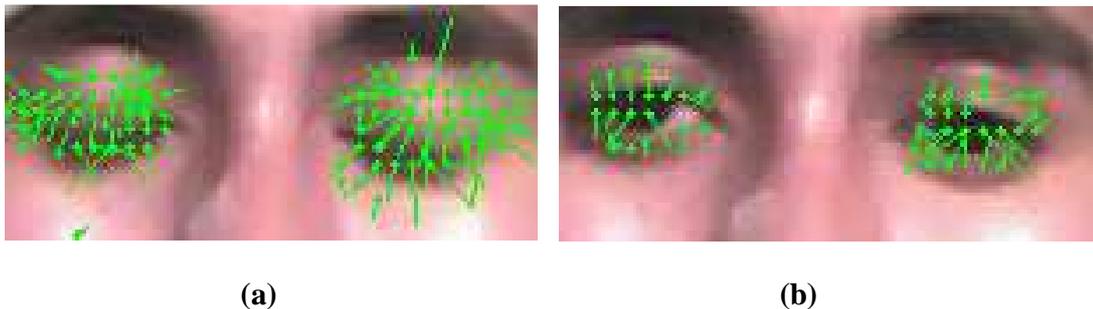
### 5.1.1. DETECCIÓN DEL MOVIMIENTO DE LOS OJOS: FLUJO ÓPTICO

Hay que detectar el movimiento de los ojos, para lo cual se utiliza el método de flujo óptico específicamente el algoritmo de Horn & Schunck con el cual se obtuvieron mejores resultados. El alto costo computacional que presenta el algoritmo es minimizado con el tamaño de la imagen a procesar, ya que solo se procesa la región de los ojos como se muestra en la Figura 32. Ésta región es obtenida basados en las coordenadas de la detección de la cabeza y de los puntos inferiores del sistema de tres puntos, de los cuales se obtiene el espacio en donde se encuentran los ojos del usuario.



**Figura 32.** Región de análisis de los ojos.

El flujo óptico con el método de Horn & Schunck, obtiene los datos de las componentes de la velocidad y dirección de movimiento, luego se promedian para obtener la dirección y magnitud de la velocidad del movimiento de toda la región analizada. En la Figura 33. Se dibujan las componentes del flujo óptico sobre la región de los ojos, en (a) el usuario está cerrando los ojos y en (b) los está abriendo.



**Figura 33.** Flujo óptico en la región de los ojos. (a) cerrándolos. (b) abriéndolos.

### 5.1.3. IMPLEMENTACIÓN DEL ALGORITMO

En la implementación de un algoritmo que permita la detección del movimiento de los ojos, es necesario detectar cuando los ojos se cierran y cuando se abren para calcular el tiempo que duran cerrados y diferenciar éste movimiento de un parpadeo. Entonces con los datos de dirección y magnitud de la componente de velocidad en la región de análisis, se determina si el movimiento ocurre hacia abajo (cerrar los ojos) o hacia arriba (abrir los ojos).

Cuando se detecta el movimiento de cierre de los ojos el algoritmo da un tiempo de espera mínimo de tres imágenes, en donde no tienen en cuenta los datos del flujo óptico y de este modo se termine el movimiento de cerrar los ojos, luego se da un tiempo de seis imágenes para detectar el movimiento causado al abrir los ojos. De esta forma se controla que no exista detección del parpadeo ya que este suele ser muy rápido y se presenta en un número inferior de imágenes a las tenidas en cuenta por el algoritmo.

Se implementó un algoritmo que permite saber cuándo se detectan dos gestos de los ojos en menos de 14 imágenes seguidas. Esto se realizó con el fin de obtener un nuevo gesto el cual es diferente al movimiento de los ojos hecho solamente una vez, y así poderlo utilizar en la interfaz con otro comando o control.

El algoritmo antes de notificar el movimiento de los ojos se apoya en el análisis de flujo óptico realizado en la región de la boca, de tal forma que si ésta se mueve en la misma dirección que la región de los ojos, es indicativo del movimiento de toda la cabeza y no existe el gesto por parte del usuario.

El algoritmo determina que existe movimiento en los ojos cuando el vector de velocidad promedio de flujo óptico de la región de análisis supera un umbral, por lo cual se recomienda que el usuario cuando realice el movimiento de abrir los ojos, lo haga de forma muy marcada, rápida y que finalice abriendo bien los ojos.

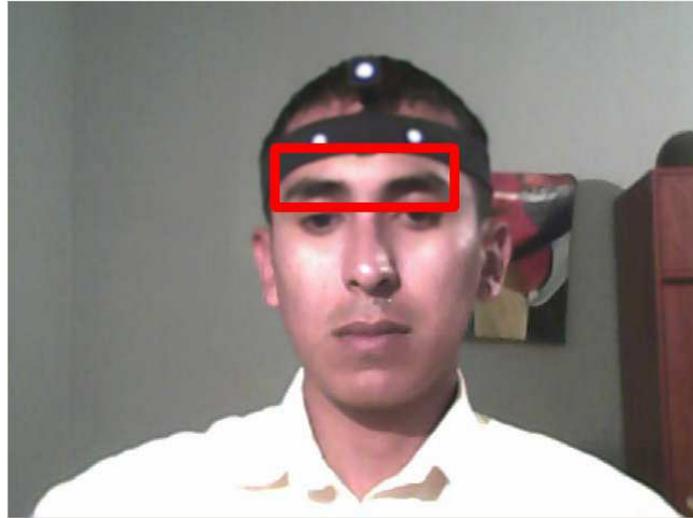
## **5.2. GESTOS DE LAS CEJAS**

El gesto de las cejas consiste en moverlas hacia arriba y regresarlas a la posición inicial, es muy común que muchas personas acostumbren a realizar este gesto durante la interacción con otras personas. Para la aplicación se considero muy útil la detección del gesto de las cejas debido a la facilidad con que las personas lo realizan.

### **5.2.1. DETECCIÓN DEL MOVIMIENTO DE LAS CEJAS: FLUJO ÓPTICO**

Para la detección del movimiento de la cejas se utiliza el método de flujo óptico específicamente el algoritmo de Horn & Schunck. Al igual que la región de los ojos, la región de las cejas se obtiene en base a las coordenadas de la detección de la cabeza y del sistema de tres puntos. El tiempo de ejecución del algoritmo del flujo óptico disminuye con el tamaño de la región, por lo cual el área de esta región es pequeña comparada con la imagen completa, como se puede ver en la Figura 34.

En la Figura 35. Se puede observar el flujo óptico del movimiento de las cejas cuando éstas ejecutan un movimiento hacia arriba.



**Figura 34.** Región de análisis de las cejas.



**Figura 35.** Flujo óptico en la región de las cejas.

### **5.2.3. IMPLEMENTACIÓN DEL ALGORITMO**

El algoritmo permite calcular la dirección y magnitud de la componente de velocidad en la región donde se encuentran las cejas, para determinar cuando existe un movimiento hacia arriba por parte del usuario. El movimiento tiene un tiempo de duración por lo cual se da un espacio de tres imágenes antes de verificar un nuevo movimiento de las cejas.

Se implementó un algoritmo que determina cuando se presentan dos gestos de los ojos en menos de 10 imágenes, de tal forma que se establece un tiempo de tres imágenes para el primer movimiento de cejas y luego el sistema espera 7 imágenes consecutivas para que se

detecte el segundo movimiento de cejas. Esta doble detección del gesto con las cejas es vital para realizar la interfaz con otro comando del *mouse*, diferente al controlado por un solo movimiento de las cejas.

El algoritmo se apoya en los datos del flujo óptico en la boca, porque si las direcciones del movimiento coinciden no se activa la detección del gesto.

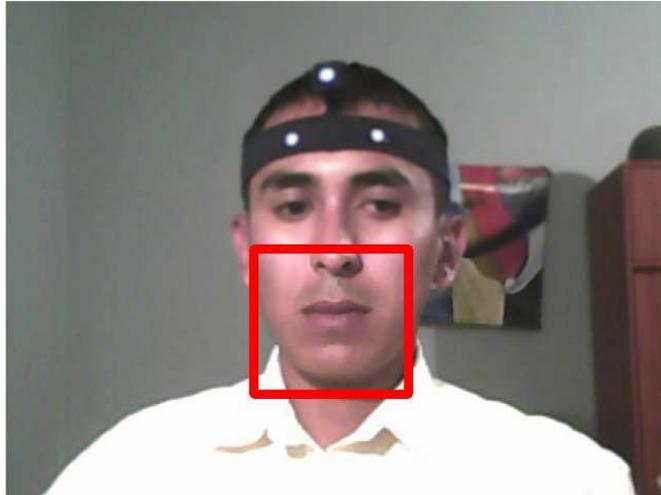
### **5.3. GESTOS DE LA BOCA**

El gesto de la boca consiste en abrirla y cerrarla, siendo muy natural y fácil de realizar. El usuario cuando se encuentra manipulando un computador es muy probable que hable, entonces es fundamental que se diferencie el movimiento realizado con la boca al hablar, del movimiento de abrir y cerrar la boca como gesto.

#### **5.3.1. DETECCIÓN DEL MOVIMIENTO DE LA BOCA: FLUJO ÓPTICO**

La región de análisis de la boca (ver Figura 36.), también se apoyo en las coordenadas del rectángulo que delimita la cabeza y del sistema de tres puntos, donde no solo se ubico la boca sino la parte del mentón, el cual contribuye con información del movimiento de la boca al abrirla.

El método de flujo óptico utilizado para el movimiento de la boca fue el de Horn & Schunck y los resultados se pueden ver en la Figura 37, donde las componentes de velocidad del movimiento se graficaron en color verde.



**Figura 36.** Región de análisis de la boca.



**Figura 37.** Flujo óptico en la región de la boca.

### **5.3.3. IMPLEMENTACIÓN DEL ALGORITMO**

El algoritmo implementado analiza una región que aparte de encontrarse la boca, se encuentra el mentón del usuario. Esta región es importante porque cuando se abre la boca se mueve toda la región entre los labios y el mentón, marcando aun más el movimiento de la boca como gesto.

El algoritmo determina cuando se abre la boca lo suficiente, como para superar un umbral

de magnitud de la velocidad del movimiento, que diferencia este movimiento del causado por el efecto de hablar.

El algoritmo se basa en los datos del movimiento realizado en el sector de las cejas para evitar que se detecte un gesto de la boca, cuando se mueve toda la cabeza del usuario.

# CAPÍTULO 6

---

---

## INTERFAZ ENTRE MOVIMIENTOS, GESTOS Y LOS CONTROLES DEL *MOUSE*

---

---

En los capítulos anteriores, se analizó y explicó cómo se obtuvieron los parámetros que describen el movimiento de la cabeza del usuario y los gestos específicos de los ojos, cejas y boca. En este capítulo se describirá la interfaz entre los parámetros obtenidos anteriormente y los respectivos controles del *mouse*.

### 6.1. MOVIMIENTOS DEL CURSOR DEL “*MOUSE*”

El movimiento del cursor del *mouse* en el computador se basa en direcciones específicas siguiendo los ejes vertical arriba, abajo y horizontal derecha, izquierda y los movimientos en diagonal se realizan haciendo combinaciones de movimiento en el eje vertical y el eje horizontal al mismo tiempo. El usuario realiza movimientos de giro de la cabeza basado en el sistema de tres puntos, que corresponden a las direcciones en el eje vertical arriba, abajo y el eje horizontal derecha, izquierda, de tal forma que al girar la cabeza en alguna dirección mencionada el cursor del *mouse* se mueve en la misma dirección.

El movimiento del cursor del *mouse* reflejado en el computador es proporcional al movimiento realizado por el usuario, lo que quiere decir que a mayor giro en las direcciones vertical, horizontal de la cabeza, mayor será la velocidad de respuesta del movimiento del *mouse* en la dirección correspondiente. El movimiento del cursor del *mouse* se detiene cuando la cabeza regresa al punto inicial, donde no hay giro de la cabeza por parte del usuario.

Los movimientos del cursor del *mouse*, responden al movimiento leve de la cabeza del usuario.

Así queda completa la interfaz entre los movimientos realizados por la cabeza del usuario y los movimientos del cursor del *mouse* en el computador.

## 6.2. CONTROLES DEL “MOUSE”

Los controles del *mouse* se refieren a las acciones efectuadas por los botones izquierdo y derecho (ver Figura 38.), los cuales son muy importantes para la asistencia del control del computador de las personas con discapacidad en los miembros superiores.



**Figura 38.** *Mouse* del computador.

Para esta aplicación se ha logrado generar el control de cuatro comandos específicos como: *click*, doble *click*, *click* sostenido y *click* derecho, estos comandos son vitales para un fácil manejo y un mejor control del computador.

### **6.2.1. “CLICK”**

El *click* es fundamental en la utilización de un computador, con este comando se puede seleccionar una ventana, aplicación, etc. En los menús es vital para la selección de un control específico.

Siendo el *click* uno de los comandos más necesarios en el uso del computador, este requiere de un fácil acceso, así que el usuario puede acceder a este comando con el movimiento de las cejas, debido a la facilidad con la que se puede hacer este gesto. El usuario debe subir y bajar las cejas para que se ejecute la acción de *click*.

Para ejecutar un nuevo *click* se deberá esperar aproximadamente tres segundos, tiempo en el cual se restringe el ingreso de un nuevo *click*, esto se hace por si se realiza una orden de doble *click* por parte del usuario.

### **6.2.2. DOBLE “CLICK”**

El doble *click* tiene como función dar la orden para ejecutar una acción determinada.

Este comando en un *mouse* convencional se activa pulsando rápidamente dos veces el botón izquierdo y en este proyecto se ha permitido que se active con la ejecución del gesto utilizado para el *click* efectuándolo dos veces. Se realiza por el movimiento de cejas, específicamente subiendo y bajando las cejas dos veces en menos de 3 segundos aproximadamente, ya que de realizarse en un tiempo mayor se activa el *click*.

### **6.2.3. “CLICK” SOSTENIDO**

El *click* sostenido o arrastre permite seleccionar o arrastrar diferente tipo de información. Este comando se ejecuta en dos partes, cuando el *click* queda sostenido y cuando se libera por parte del usuario.

Para activar el *click* sostenido se debe abrir y cerrar la boca, luego se libera la ejecución repitiendo el mismo movimiento de la boca. El movimiento que realiza el usuario para hablar no activa el *click* sostenido.

### **6.2.4. “CLICK” DERECHO**

El click derecho muestra un menú de las acciones que se pueden realizar sobre dicho punto en donde se está ejecutando. El usuario activa este comando con el movimiento de los ojos, se deben cerrar y abrir los ojos esperando de 1 a 2 segundos aproximadamente con los ojos cerrados, para diferenciar el movimiento de un parpadeo.

# CAPÍTULO 7

---

---

## INTERFAZ GRAFICA

---

---

La guía de usuario explica todo lo referente a la aplicación final a la cual se le dio el nombre de “**SCODEM 2009**”. En el capítulo se incluye la interfaz grafica de usuario donde se explica las respectivas opciones que se tienen para la manipulación de la aplicación, luego el manual de usuario expone los pasos que se deben seguir para la correcta utilización de la aplicación, y por último se dan a conocer las condiciones de funcionamiento de “**SCODEM 2009**”.

### 7.1. INTERFAZ GRAFICA DE USUARIO

La interacción persona-computador, refleja la necesidad de un espacio que media la relación de un sujeto y un computador surgiendo la interfaz de usuario. Surgen conceptos como, manipulación directa, iconos controles y texto, vitales en la creación de la interfaz [31].

La interfaz grafica de usuario debe ser funcional, simple, fácil de operar. En este caso se ha optado por crear la interfaz grafica (ver Figura 39.) donde la información de la imagen

tomada por la cámara webcam, se muestra reducida en tamaño y se enmarca la región donde el sistema ha ubicado la cabeza del usuario, además de la ubicación del sistema de tres puntos.



**Figura 39.** Interfaz grafica de usuario.

En la ventana no se encuentra ningún botón de control, pero se tienen los siguientes comandos implementados con gestos y movimientos detectados anteriormente, para que la persona con discapacidad pueda controlar con mayor eficiencia el software.

El comando “INICIO” se activa con el movimiento de la boca (abrir y cerrar), tiene la función de iniciar la interacción usuario-computador, proporcionándole al usuario el control total de las funciones y el movimiento del *mouse*. Este inicio no se realiza cuando se ejecuta la aplicación, con el fin de darle espacio para que el usuario esté listo y se ubique correctamente en el campo de visión de la cámara webcam. El comando “INICIO” también sirve para quitar la aplicación del modo de pause.

El comando “PAUSA” se ejecuta abriendo y cerrando los ojos dos veces seguidas en menos de cinco segundos, éste da la orden con la cual se cancela la interacción entre el usuario y el

computador temporalmente, también es útil cuando el sistema necesita reiniciar valores, debido a una mala ubicación por parte del usuario.

El comando “SALIR” se activa inclinando la cabeza 20° hacia la izquierda o la derecha, el cual cancela la aplicación y sale completamente.

Cuando se ejecuta el software “**SCODEM 2009**” se da una espera máxima de un minuto para la ubicación del usuario y la activación del comando inicio, de lo contrario el software se cierra automáticamente. Éste es el mismo tiempo que puede permanecer el sistema en modo de pause, de lo contrario el software también se cerrara automáticamente.

Si el usuario mueve el periférico del *mouse*, la aplicación inactiva los controles ejercidos por el movimiento de la cabeza y los gestos del rostro del usuario, por un lapso de siete segundos.

## **7.2. MANUAL DE USUARIO**

El manual de usuario es una herramienta vital para explicar el funcionamiento de la aplicación final. El usuario podrá utilizar “**SCODEM 2009**” sin inconvenientes, porque en él se explican los requisitos de funcionamiento, la configuración de la cámara webcam, la utilización del sistema de tres puntos, las condiciones de funcionamiento, la interfaz grafica y lo más importante la explicación de la conexión entre los movimientos y los gestos con los controles del *mouse*.

El manual de usuario (ver ANEXO A) se ha realizado de una forma sencilla para que el usuario no tenga dificultades y por el contrario pueda acceder fácilmente a “**SCODEM 2009**”.

### **7.3. CONDICIONES DE FUNCIONAMIENTO**

La aplicación final deberá cumplir con unos requisitos de funcionamiento debidos al entorno y a las características de los algoritmos que la constituyen.

La iluminación es importante en los sistemas de visión artificial, debido a que simplifican el análisis de una escena. Es un factor que afecta la complejidad de los algoritmos, por lo cual es recomendable tener buenos sistemas de iluminación de acuerdo a la aplicación [24]. En éste caso se recomienda que las condiciones de luz sean óptimas en el rostro del usuario para evitar sombras no deseadas, bajo contraste, saturación y brillos que confunden al sistema.

Hay un parámetro importante para tener en cuenta en el correcto funcionamiento y es el que tiene que ver con la distancia entre el usuario y la cámara webcam, debido a que esta variación cambia el valor de algunos parámetros vitales en el procesamiento de imágenes. Si bien, no es necesario que el usuario este centrado en la imagen, es necesario que la cabeza este en su totalidad dentro de la imagen, pero que no se aleje el usuario una distancia donde la cabeza ocupe un espacio, de menos del 30% aproximadamente de la imagen.

Se deben evitar en el fondo de la imagen objetos brillantes, focos de luz y objetos en movimiento, debido a que pueden confundir al sistema.

La utilización de gafas por parte del usuario causa brillos indeseables para la detección del sistema de tres puntos y evita la detección de movimientos de los ojos, por lo cual se recomienda no utilizarlas.

# CAPÍTULO 8

---

---

## RESULTADOS Y DISCUSIÓN

---

---

### 8.1. DETECCIÓN DE LA CABEZA

Los resultados en la detección de la cabeza son adecuados para la aplicación, debido a que con mínimos movimientos de giro o desplazamiento de la cabeza por parte del usuario se logra la ubicación. Luego de detectada la cabeza el usuario puede dejar de moverla y el algoritmo obtiene su posición de acuerdo a las últimas coordenadas detectadas. La ejecución del algoritmo se hace en tiempo real contribuyendo con los objetivos de la aplicación.

Se determinó que para la ubicación de la cabeza por el método de resta de imágenes, es suficiente un buffer de tres imágenes para realizar todo el arreglo que establece las coordenadas del rectángulo que delimita la cabeza.

La iluminación sobre el rostro del usuario debe ser óptima, debido a que si es muy baja no se puede garantizar la calidad de la imagen necesaria para su procesamiento y si es demasiado alta puede causar brillos y saturaciones que causan errores en el procesamiento.

En el fondo de la imagen los objetos que producen brillos con la luz pueden presentar problemas de ruido al igual que focos de iluminación por lo cual es adecuado evitar este tipo de objetos mientras se utiliza el sistema, también es importante que no existan objetos en movimiento en el fondo de la imagen.

Cuando el usuario se ubica muy lejos de la cámara, ocupando aproximadamente menos del 30% de la imagen, el sistema determina que el usuario no está en el rango de visión de la cámara, lo que requiere que el usuario se ubique más cerca a la cámara.

## **8.2. MOVIMIENTO DE LA CABEZA**

El usuario tiene que utilizar el sistema de tres puntos para que la aplicación funcione correctamente, este es un condicionamiento que puede crear dificultad al usuario, puesto que este implemento se debe fabricar, pero que se diseño para aumentar las ventajas en los controles del *mouse* y para que en un trabajo futuro se puedan utilizar los movimientos, que hasta el momento no se han aplicado a ninguna interfaz, en otras aplicaciones, ya que la implementación de nuevos controles contribuye eficazmente en la asistencia a personas con limitaciones en los miembros superiores.

Se logro la reducción de la región de análisis del sistema de tres puntos en comparación con el tamaño de la imagen, basado en el rectángulo que delimita la cabeza del usuario. Esta reducción aumenta la velocidad de proceso del algoritmo y contribuye en el objetivo de realizar una aplicación que funcione en tiempo real.

Se logro determinar la ubicación de cada punto del sistema de tres puntos e identificar la posición de cada uno con respecto a los otros dos, por dos métodos diferentes, el de integrales proyectivas y el de seguimiento de puntos. Luego de probar en diferentes condiciones los dos métodos, el de seguimiento de puntos demostró se mas robusto de tal forma que este método fue el que se implemento para la aplicación final.

El usuario al momento de ubicar el sistema de tres puntos en la cabeza debe de alinear los puntos de la banda y centrar el punto del elemento L con respecto a los otros dos puntos para un correcto funcionamiento. Pero si el usuario no logra la alineación y se presenta algún grado de inclinación o el punto superior no queda completamente centrado con respecto a los otros dos, el sistema inicialmente detecta el grado de inclinación y la posición del punto superior con respecto a los puntos inferiores, para determinar los porcentajes de distancia entre los puntos y realizar la posterior comparación con los parámetros de las nuevas imágenes.

De acuerdo a lo mencionado anteriormente, el sistema se vuelve más robusto ya que funciona cuando el alineamiento del sistema de tres puntos no está como se sugiere permitiéndole al usuario cierta libertad y facilidad para ubicar el sistema de tres puntos. Sin embargo el desalineamiento no debe ser demasiado ya que puede presentarse pérdida de sensibilidad del movimiento en alguna de las direcciones.

El algoritmo detecta el sistema de tres puntos aproximadamente en un ángulo de giro de la cabeza horizontal inferior a  $30^\circ$ , tanto a la derecha como hacia la izquierda, este valor cambia sutilmente dependiendo de la iluminación. Si la iluminación es baja el valor disminuye y si aumenta la iluminación el ángulo puede aumentar.

El giro de la cabeza vertical hacia abajo puede ser detectado por el algoritmo hasta que los puntos del sistema de tres puntos se alinean horizontalmente en la imagen, ya que si el punto superior queda en la parte inferior de la imagen con respecto a los otros puntos, el sistema determina que no hay localización de los puntos y no produce ningún movimiento del cursor del *mouse*.

El usuario debe tener en cuenta que el movimiento de giro de la cabeza vertical hacia arriba, mueve el cursor del *mouse* cuando el punto superior del sistema de tres puntos no se sale del área del rectángulo que delimita la detección de la cabeza.

### **8.3. DETECCIÓN DE LOS GESTOS DEL ROSTRO**

La detección de los tres gestos originados por los ojos, cejas y boca se hizo utilizando el algoritmo de flujo óptico obteniéndose buenos resultados. Una de las desventajas del método de flujo óptico es la baja velocidad de proceso en imágenes de tamaño de 640x480 píxeles. Para la viabilidad del método de flujo óptico en la detección de los gestos se analizaron zonas pequeñas correspondientes a la ubicación de cada parte del rostro, con lo cual se disminuyó considerablemente el costo computacional.

La detección del gesto de las cejas consiste en el movimiento de las cejas hacia arriba, sin que exista gran movimiento en el sector de la boca, esto se hace con el fin de que no se presente la detección del gesto de las cejas cuando se está moviendo toda la cabeza.

Cuando se realizan dos movimiento de las cejas en un periodo de tiempo aproximadamente inferior a tres segundos, el sistema toma este proceso como un gesto independiente al de subir las cejas una sola vez.

El gesto de los ojos se detecta cuando se cierran y se abren los ojos, pero cuando se cierran se deben dejar así por un periodo de dos segundos aproximadamente con el fin de diferenciar el gesto de un parpadeo. Para este gesto también se verifica que no exista un alto grado de movimiento en el sector de la boca, para evitar una falsa detección.

Cuando se abren y cierran los ojos dos veces seguidas, en un periodo de tiempo aproximadamente inferior a cuatro segundos, el sistema determina este proceso como un nuevo gesto diferente al de abrir y cerrar los ojos una sola vez.

El gesto de los ojos se ejecuta con mayor precisión cuando el movimiento de abrir los ojos se realiza de forma contundente, rápida y es importante que el usuario en este proceso abra bien los ojos.

La detección del gesto de la boca se presenta cuando se abre y cierra la boca de forma extensa con el fin de diferenciar este gesto del proceso de abrir y cerrar la boca cuando se habla. Para este caso se verifica que no exista gran movimiento en el sector de las cejas para evitar que se detecte el gesto cuando se mueve toda la cabeza.

Para la detección de cada gesto los algoritmos utilizan pocas imágenes, por lo que se hace viable la implementación de interfaces con los controles del *mouse*, para que se realicen con una doble detección de un gesto, como se hizo para el comando de doble *click* (doble movimiento de las cejas) y pausar la aplicación (abrir y cerrar los ojos dos veces seguidas).

#### **8.4. SCODEM 2009**

“SCODEM 2009” es la aplicación final, del proceso de realizar un sistema de control del *mouse* para la asistencia a personas con limitaciones en los miembros superiores. Se utilizaron diferentes herramientas de la visión por computador como filtros, segmentación, detección de movimiento, entre otros; para implementar un sistema que se basa en los movimientos leves de la cabeza para mover el cursor del *mouse* y en los gestos del rostro realizados por los ojos, cejas y boca para el control de los comandos del *mouse* como *click* (subir las cejas), doble *click* (subir las cejas dos veces seguidas), *click* derecho (cerrar y abrir los ojos) y *click* sostenido (abrir y cerrar la boca).

Se lograron implementar comandos para el control de la aplicación “SCODEM 2009”, como pause (cerrar los ojos dos veces seguidas), donde el usuario pierde el control del *mouse* mediante los movimientos de la cabeza y los gestos del rostro, des pause (abrir la boca), donde el usuario retoma el control del *mouse* y salir de la aplicación (inclinación de 20° de la cabeza).

Se logro detectar cuando el usuario se aleja o acerca de la cámara a partir de la posición inicial. Estos parámetros no se utilizaron en ninguna interfaz de la aplicación pero se dejan para posteriores aplicaciones como interfaces con controles de juegos o simuladores.

Cuando el usuario gira la cabeza un ángulo en el cual se pierde la ubicación del sistema de tres puntos o se retira del ángulo de visión de la cámara webcam, el sistema deja el cursor del *mouse* en la última posición determinada por el usuario y sigue analizando las imágenes hasta obtener la ubicación de la cabeza y del sistema de tres puntos nuevamente. Al ubicar el sistema de tres puntos el algoritmo no inicializa variables, continua con los datos con los que se inicializo el sistema antes de perder la ubicación.

El sistema necesita de una cámara webcam de bajo costo y un sistema de tres puntos el cual se diseño con el fin de identificar y caracterizar los movimientos de la cabeza. Los resultados obtenidos fueron buenos ya que se detectaron siete movimientos diferentes, tales como: giro de la cabeza horizontal hacia la izquierda y derecha, giro de la cabeza vertical hacia arriba y abajo, inclinación de la cabeza y el acercamiento o alejamiento de la cabeza con la cámara.

Los movimientos del cursor del *mouse* se realizan con movimientos muy leves del movimiento de la cabeza del usuario. Si los movimientos de la cabeza son demasiado bruscos puede causar que el sistema no funcione correctamente.

Para la ubicación inicial se recomienda que el usuario tenga la cabeza levemente inclinada hacia abajo, ya que de esta forma se contribuye en la sensibilidad del movimiento del giro vertical hacia arriba haciendo que el movimiento del cursor del *mouse* en esta misma dirección sea más rápido.

Para que se detecten los gestos del rostro es necesario que cuando el usuario los realice no exista un movimiento fuerte de la cabeza para evitar que el sistema determine que no es un gesto lo que se hizo sino un movimiento de la cabeza.

Cuando la iluminación sobre el rostro es baja es posible que no se detecten los gestos del rostro, con lo cual se pierde el control de los comandos de *mouse*, mientras que la detección

del sistema de tres puntos es más robusta a baja iluminación, aunque empieza a disminuir el rango de detección en los giros de la cabeza. La iluminación demasiado alta puede saturar la imagen con brillos no deseados perdiendo la posición del sistema de tres puntos y por lo tanto se pierde el control de todo el sistema.

Los algoritmos se implementaron en C++ debido a que permite realizar aplicaciones en tiempo real y permite el uso de librerías para el procesamiento de imágenes. En este proyecto se utilizó la librería de *OpenCV (Open Source Computer Vision)* la cual tiene funciones para el procesamiento de imágenes y los tiempos de los procesos son bajos, además esta librería tiene licencia gratuita.

El sistema implementado en el proyecto controla los movimientos y comandos del *mouse*, y para controlar de forma más autónoma el computador se requiere de un teclado virtual que permita escribir en aplicaciones de texto, navegar en internet, entre otras. Estos teclados se activan por medio del *mouse*, por lo que son adecuados para complementar el sistema. Se recomiendan los teclados virtuales de licencia gratuita como *Virtual KeyBoard 1.0*, *Click-N-Type*, *On-Screen KeyBoard Portable*, etc. Sin embargo se deja a consideración del usuario la implementación y el tipo de teclado a instalar.

#### **8.4.1. PRUEBAS CON USUARIOS**

Se hicieron pruebas, las cuales consistían en que 10 personas utilizaran el sistema “SCODEM 2009”, realizaran algunas tareas y ejecutaran aplicaciones apoyados de un teclado virtual de licencia gratuita para complementar el control del *mouse* y poder acceder a las aplicaciones de forma más completa. Luego se les pidió que contestaran una encuesta (ver Anexo B) para que evaluaran el sistema desde su propia experiencia. Los resultados se pueden ver en la Tabla 5.

**Tabla 5.** Resultados de la encuesta realizada a los usuarios para evaluar SCODEM 2009.

La calificación se realiza en una escala de 1 a 10, siendo 1 el valor mínimo y 10 el valor máximo												
PREGUNTAS	USUARIOS	1	2	3	4	5	6	7	8	9	10	PROMEDIO
	EDAD (AÑOS)	26	19	23	25	26	12	21	18	32	30	
Establezca el grado de dificultad que usted encontró para lograr la detección de la cabeza.		1	3	1	2	1	1	1	1	1	1	1.3
Establezca el grado de dificultad que usted encontró para coordinar los movimientos de la cabeza con los movimientos del cursor del <i>mouse</i> .		2	4	1	2	3	4	1	2	3	2	2.4
Establezca el grado de dificultad que usted encontró para activar los comandos del <i>mouse</i> por medio de los gestos.	Click	1	3	1	2	2	2	1	1	1	1	1.5
	Doble Click	2	3	1	2	2	2	1	1	1	1	1.6
	Click derecho	1	3	1	2	2	2	1	1	2	2	1.7
	Click sostenido	3	5	2	3	3	4	1	2	2	2	2.7
Establezca el grado de dificultad que usted encontró para realizar tareas y ejecutar aplicaciones en el computador.	Abrir una carpeta de archivos	2	3	1	1	2	2	1	1	1	2	1.6
	Crear una carpeta	2	3	1	2	2	3	1	1	2	2	1.9
	Abrir y cerrar un archivo de texto	2	3	1	1	2	2	1	1	1	2	1.6
	Arrastrar una carpeta de un punto a otro en el escritorio	2	5	2	3	3	4	1	2	2	2	2.6
	Navegar en internet con la ayuda de un teclado virtual activado con SCODEM 2009	2	4	2	2	4	4	2	2	3	2	2.7
¿Cómo se sintió físicamente utilizando el sistema?	Dolor de cabeza	1	1	1	1	1	1	1	1	1	1	1
	Dolor de cuello	1	4	1	2	2	2	1	1	1	1	1.6
	Molestia en los ojos	1	2	1	1	1	1	1	1	1	1	1.1
	Dolor de espalda	1	2	1	1	1	1	1	1	1	1	1.1
Califique en términos generales el sistema.		9	7	10	9	9	8	10	10	8	9	8.9

A partir de los resultados mostrados en la Tabla 5. Se puede concluir lo siguiente:

El sistema no presenta ningún grado de dificultad con la detección de la cabeza, para la mayoría de los usuarios, solamente dos personas presentaron algún tipo de dificultad, que evaluaron en la escala de rangos como 2 y 3, siendo 10 el valor más alto de dificultad y 1 el más bajo.

Según los usuarios encuestados el promedio obtenido para medir el grado de dificultad que se tiene en coordinar los movimientos de la cabeza con los movimientos del cursor del *mouse*, el cual consiste en llevar el cursor a puntos específicos y moverlo por toda la pantalla, fue de 2.4, mostrando así, que los usuarios presentan algo de dificultad y es mas evidente cuando el usuario inicia la aplicación por primera vez, ya que con el tiempo se mejora la coordinación de los movimientos.

Para activar los comandos del *mouse* los usuarios presentaron mayor facilidad para el *click*, activado con el movimiento de las cejas (subir y bajar), y el que mayor presento dificultad fue el *click* sostenido, activado con el gesto de abrir y cerrar la boca, esto se debe a que el gesto de la boca tiene que diferenciarse del movimiento causado al hablar ocasionando que los umbrales de detección sean más altos que para los otros gestos. Los comandos de doble *click* (doble movimiento de las cejas) y *click* derecho (cerrar y abrir los ojos) se facilitan a los usuarios quienes los evaluaron en promedio con 1.6 y 1.7 respectivamente en la escala de dificultad.

Los usuarios debieron realizar varias tareas y ejecutar algunas aplicaciones para probar el funcionamiento del sistema. Los usuarios presentaron mayor facilidad para abrir una carpeta de archivos al igual que al abrir y cerrar un archivo de texto presentando un valor promedio de dificultad de 1.6; los valores promedio del grado de dificultad mas altos fueron los obtenidos cuando los usuarios tenían que arrastrar una carpeta de un lugar a otro para lo cual se tenía que utilizar el *click* sostenido y los movimientos de la cabeza y navegar

en internet con la ayuda de un teclado virtual, en este caso la dificultad se debió en gran parte al proceso de escribir con el teclado virtual ya que este proceso no es muy rápido y necesita de práctica para aumentar la velocidad de escritura. La creación de una carpeta obtuvo una calificación promedio de dificultad por parte de los usuarios de 1.9, para lo cual el usuario tiene que utilizar el gesto de cerrar y abrir los ojos para activar el *click* derecho, los movimientos de la cabeza y el movimiento de las cejas para dar *click*.

A los usuarios se les preguntó por dolores físicos causados por el uso del sistema. Ninguno presentó dolor de cabeza, solo un usuario presentó molestia en los ojos y dolor de espalda muy leve y cuatro personas presentaron algún grado de dolor de cuello, esto puede deberse a que el usuario debe estar pendiente de no mover libremente la cabeza ya que con estos movimientos controla el cursor del *mouse*.

La calificación del sistema por parte de los usuarios en general fue en promedio de 8.9 en una escala de 1 a 10 siendo 1 la calificación más baja y 10 la más alta. Todos los usuarios presentaron un buen grado de aceptación del sistema y dieron un concepto favorable al proyecto por ser una herramienta útil para las personas con limitaciones en los miembros superiores.

---

---

## CONCLUSIONES

---

---

- Se ha diseñado e implementado un sistema de control del *mouse* para la asistencia a personas con limitaciones en los miembros superiores mediante el análisis de imágenes del movimiento de la cabeza y gestos del rostro.
- Se construyó e implementó un sistema de tres puntos para ser utilizado por el usuario, con el cual se detectaron siete movimientos diferentes de la cabeza, tres más de los requeridos para el desplazamiento del *mouse*, y que pueden ser utilizados para otras aplicaciones futuras.
- Se logró identificar el movimiento de giro de la cabeza e implementar la interfaz con los movimientos del cursor del *mouse*.
- Se analizaron varios métodos para la detección de movimiento, determinando la viabilidad de cada uno en los diferentes procesos de la aplicación.
- Se logró la detección de los gestos del rostro realizado por los ojos, cejas y boca, correspondiente a los movimientos como los de: abrir y cerrar los ojos, mover las cejas hacia arriba, abrir y cerrar la boca.

- Se implemento la interfaz entre los gestos del rostro con los comandos del *mouse* correspondientes a *click* (movimiento de cejas), doble *click* (doble movimiento de cejas), *click* derecho (abrir y cerrar los ojos), *click* sostenido (abrir y cerrar la boca).
- Se logro implementar controles para la aplicación “SCODEM 2009” correspondientes a pausar la aplicación (abrir y cerrar los ojos dos veces seguidas), des pausar la aplicación (abrir y cerrar la boca) y salir (inclinarse la cabeza un ángulo de 20°).
- La librería OpenCV demostró ser óptima en la implementación de los algoritmos para el procesamiento de imágenes en tiempo real.

---

---

## **TRABAJO A FUTURO**

---

---

En el trabajo a futuro se recomienda implementar la interfaz para controlar programas específicos de texto, reproductores, juegos, etc. Utilizando los movimientos y gestos que se detectaron.

Se recomienda realizar la interfaz con simuladores y juegos, para que sean controlados con todos los movimientos y gestos detectados en esta aplicación.

---

---

## REFERENCIAS

---

---

- [1] Departamento Administrativo Nacional De Estadística, DANE. Información estadística de la discapacidad, 2004.
- [2] Departamento Administrativo Nacional De Estadística, DANE. Resultados preliminares de la implementación del registro para la localización y caracterización de las personas con discapacidad, 2007.
- [3] Departamento Administrativo Nacional De Estadística, DANE, censo general 2005. Conceptos generales, 2005.
- [4] Departamento Administrativo Nacional De Estadística, DANE, censo general 2005. Discapacidad personas con limitaciones permanentes, 2005.
- [5] Ministerio de la protección social. Política pública en discapacidad, fortalecimiento y movilización social, 2005.
- [6] Rivera, Alma Yadira. Proyecto Maseducacion. Uso didáctico de la red internet. 2005.

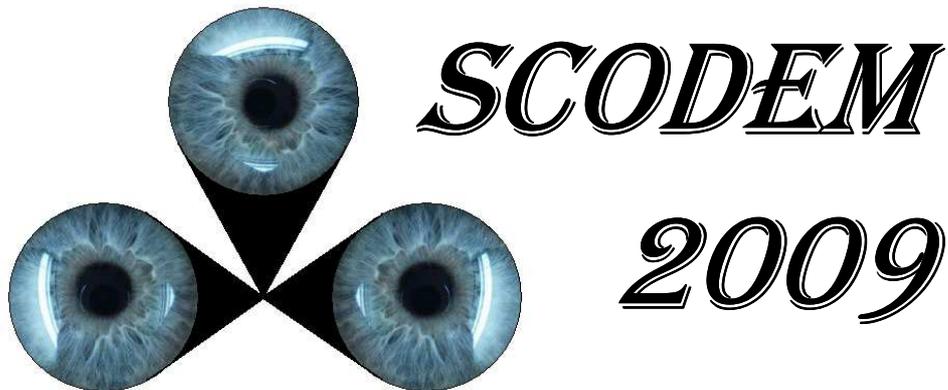
- [7] [http://www.xombra.com/go\\_news.php?articulo=3155](http://www.xombra.com/go_news.php?articulo=3155), consultado, 29 de Octubre de 2009.
- [8] <http://www.correvedile.com/informatica/mouse-visual.html>, consultado, 29 de Octubre de 2009.
- [9] González Claudio, Gayoso Carlos, Arnone Leonardo y García Juan. Emulador de teclado y mouse para discapacitados motrices severos utilizando dispositivos lógicos programables. Artículo, Universidad Nacional de Mar del Plata, 2007.
- [10] <http://www.hispazone.com/Articulo/381/OCZ-NIA---Neural-Impulse-Actuator-Primeras-impresiones-.html>, consultado, 29 de Octubre de 2009.
- [11] [http://shop.orin.com/shop/index.php?main\\_page=product\\_info&products\\_id=1&zenid=584a6659c351bac41e5c545e93598d91](http://shop.orin.com/shop/index.php?main_page=product_info&products_id=1&zenid=584a6659c351bac41e5c545e93598d91), consultado, 29 de Octubre de 2009.
- [12] <http://weblog.mendoza.edu.ar/informatica/archives/021145.html>, consultado, 30 de Octubre de 2009.
- [13] <http://www.enablemart.com/Head-Eye-Controlled-Input/SmartNAV3-AT>, consultado, 30 de Octubre de 2009.
- [14] [http://www.delacole.com/cgi\\_perl/medios/vernota.cgi?medio=comunidades&numero=465&nota=465-15](http://www.delacole.com/cgi_perl/medios/vernota.cgi?medio=comunidades&numero=465&nota=465-15), consultado, 30 de Octubre de 2009.
- [15] <http://robotica.udl.es/>, consultado, 30 de Octubre de 2009.
- [16] <http://www.depi.itch.edu.mx/apacheco/expo/html/ai11/vision.html>, consultado, 30 de Octubre de 2009.

- [17] Pajares Gonzalo. Aplicaciones industriales de la visión por computador. Universidad Complutense de Madrid 2009.
- [18] Cayetano Guerra, Artal. Seguimiento visual artificial. Artículo, Universidad de las Palmas, 2003.
- [19] Pérez Alfonso y Solís Hugo. Métodos para seguimiento de dedo en tiempo real. Artículo. 2005.
- [20] Nieto M. Fernando. Implementación de un sistema en tiempo real para la detección de desplazamiento basado en la optimización de un conjunto de puntos de control para el estudio del flujo óptico, 2008.
- [21] Martín José, Zuloaga Aitzol y Cruzado Josu. Banco de pruebas para la estimación del movimiento basado en el cálculo del flujo óptico. Artículo, Universidad del País Vasco, 2001.
- [22] Martín José, Zuloaga Aitzol y Cruzado Josu. FPGA implementation of the Horn & Shunk optical flow algorithm for motion detection in real time Images, Artículo, Universidad Politécnica de Madrid, 2001.
- [23] Molina Rafael. Campo de movimiento en el plano de la imagen: Flujo óptico y correspondencia, 2003.
- [24] Escalera, Arturo. Visión por computador. Madrid: Prentice Hall, 2001.
- [25] Benito Darío. Detección automática del color de la piel en imágenes bidimensionales basado en el análisis de regiones. Proyecto fin de carrera, Universidad Rey Juan Carlos, 2005.

- [26] Conde Patricia y Rodríguez Fernando. Detección de caras en imágenes capturadas por cámaras cam. Artículo, Universidad de Vigo, 2005.
  
- [27] Chávez, Jorge. Introducción al tratamiento digital de imágenes, México, 2002.
  
- [28] Trujillo, Isaías. Estudio sobre viabilidad de utilización de un sistema de visión artificial para la realización de control de calidad en la fabricación de adhesivos sólidos en H.B Fuller Colombia. Trabajo de grado, Universidad del Cauca, 2007.
  
- [29] Monllor, Fernando. Sistemas Informáticos: Análisis, diseño e implementación de un dispositivo gestual. 2004.
  
- [30] García Mateos, Ginés. Procesamiento de caras humanas mediante integrales proyectivas. Tesis doctoral, Universidad de Murcia, 2007.
  
- [31] Marrero Expósito Carlos. Interfaz gráfica de usuario: Aproximación semiótica y cognitiva. Proyecto de investigación, Universidad de la Laguna, 2006.

## ANEXO A: MANUAL DE USUARIO

# *Manual de Usuario*



## *Sistema De Control Del Mouse*

### *Requisitos*

- Computador con requisitos mínimos de: Procesador de 1.8GHz y memoria RAM de 512MB
- Cámara webcam.
- Software “SCODEM 2009”
- Sistema de tres puntos.

## *Webcam*

- La cámara *webcam* tiene que soportar una resolución de 640x480.
- La configuración de la cámara *webcam* deberá ser la siguiente:
  - El control de parpadeo tiene que corresponder a la frecuencia de la red eléctrica (50Hz ó 60Hz).
  - El control automático de ganancia tiene que estar activo.

## *Sistema de tres puntos*

- El sistema de tres puntos está conformado por una banda y un elemento L.
- La banda se ubica sobre la frente, rodeando la cabeza y asegurándose en la parte posterior, verificando que los dos puntos blancos queden centrados en la parte frontal de la cabeza.
- El elemento L se ubica en la parte frontal de la cabeza del usuario, asegurándose con la banda y verificando que la proyección del punto blanco sobre la banda, quede centrado en el medio de los dos puntos.



## *Condiciones del entorno*

- La aplicación “**SCODEM 2009**” para su correcto funcionamiento requiere de un entorno con optimas condiciones de iluminación, específicamente sobre el rostro del usuario.
- La imagen debe contener por completo la cabeza del usuario y no debe estar muy alejada de la cámara webcam. No es necesario que la cabeza este centrada en la imagen.

## *Opciones de SCODEM 2009*

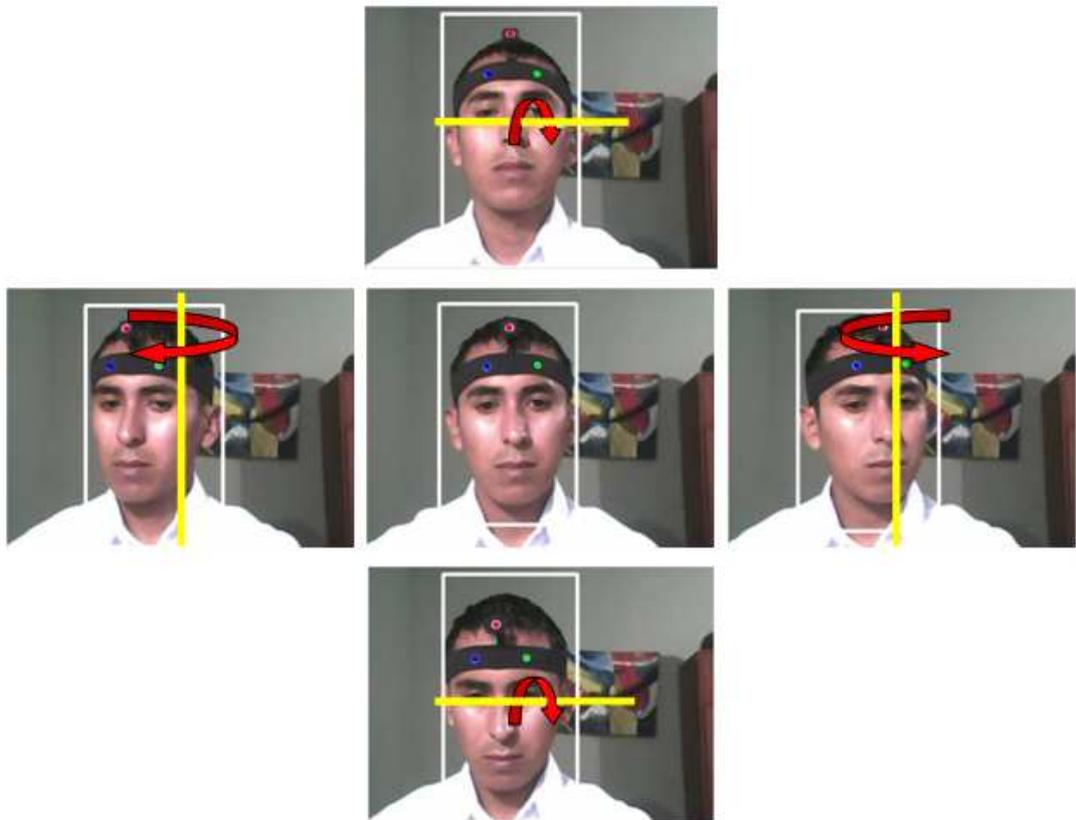
- El comando “**INICIO**” se activa con el movimiento de la boca (abrir y cerrar), y éste permite el inicio de la interacción usuario-computador, otorgando el control total del *mouse* al usuario. También quita la aplicación del modo de pause. Para inicializar se recomienda inclinar la cabeza de forma muy sutil hacia abajo, para mejorar la sensibilidad en los movimientos del cursor del *mouse*.
- El comando “**PAUSA**” se ejecuta con el movimiento de cerrar y abrir los ojos dos veces en menos de cinco segundos, con el cual se cancela la interacción entre el usuario y el computador hasta nueva orden por parte del usuario.
- El comando “**SALIR**” se activa con la inclinación de 20° de la cabeza hacia la izquierda o la derecha, cancelando la aplicación y saliendo completamente.
- Cuando el software “**SCODEM 2009**” se ejecuta o está en modo de pause, se da una espera máxima de un minuto para activar el comando de inicio, de lo contrario el software se cierra automáticamente.
- Si se mueve el periférico del *mouse* el sistema inactiva los controles ejercidos por parte del usuario por siete segundos.



## *Controles del mouse*

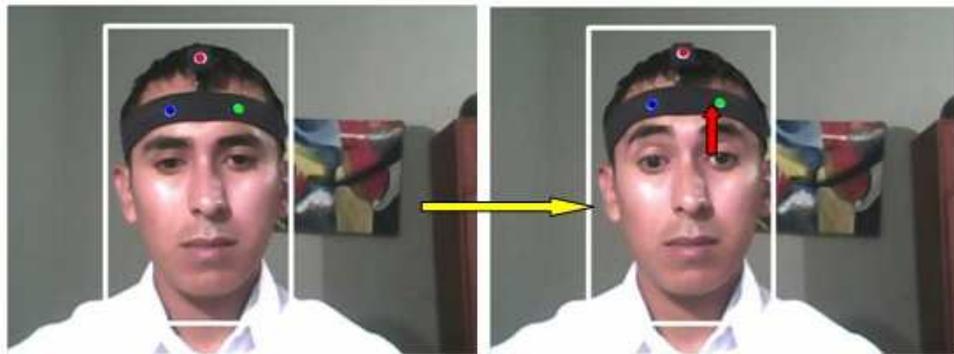
### *Movimiento del cursor del mouse*

- Los movimientos del cursor del mouse se realizan con los movimientos de giro de la cabeza. Si la cabeza gira hacia arriba, abajo, derecha, izquierda, el cursor se mueve en las mismas direcciones respectivamente.
- La velocidad del movimiento del cursor *mouse* es proporcional al giro de la cabeza; así, a mayor giro, mayor velocidad y a menor giro, menor velocidad del puntero del *mouse*.
- Para detener el movimiento del *mouse* es necesario regresar a la zona inicial, donde no hay giro de la cabeza.



## *Click*

- El *click* del *mouse* se realiza subiendo y bajando las cejas a la posición inicial.



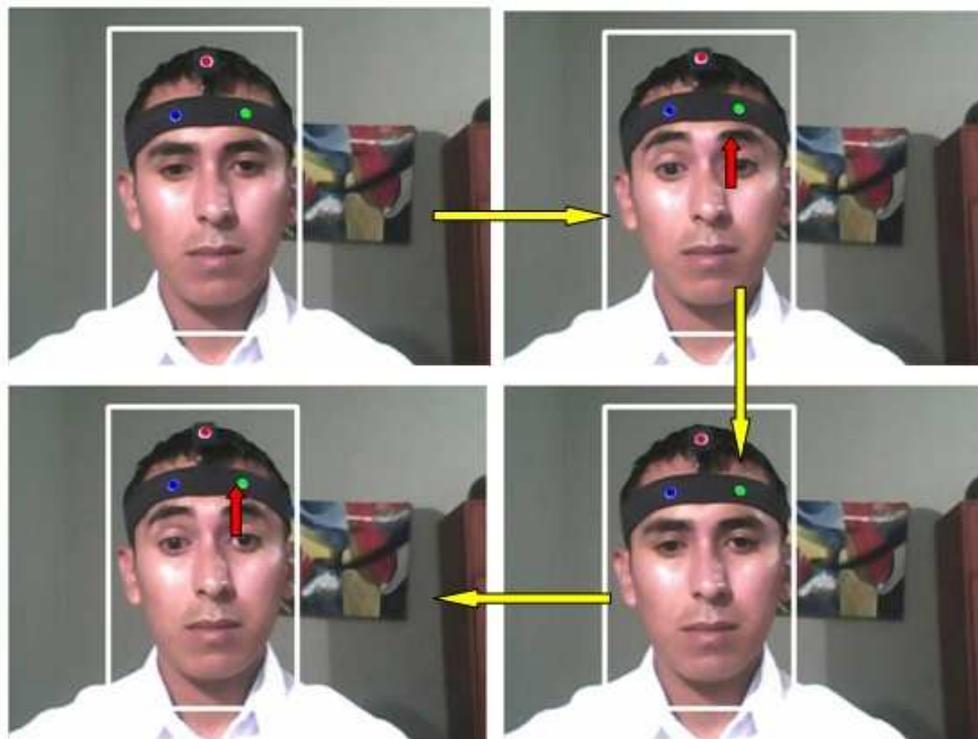
## *Click derecho*

- El click derecho del *mouse* se realiza cerrando los ojos durante un tiempo de 1 a 2 segundos.



## *Doble click*

- El doble *click* requiere de realizar dos veces seguidas el *click* (subir y bajar las cejas), pero condicionado a un lapso de tiempo de uno a tres segundos entre los *clicks*.



## *Click sostenido*

- El *click* sostenido del *mouse* se realiza de la siguiente forma:
  - Se abre y cierra la boca para que se ejecute el *click* sostenido, luego de hacer el proceso que se deseaba, se abre y cierra la boca para liberar el *click* sostenido.



## ANEXO B: ENCUESTA DE EVALUACIÓN DEL SISTEMA

### EVALUACIÓN DE SCODEM 2009



Fecha: \_\_\_\_\_

Nombre : \_\_\_\_\_

Edad: \_\_\_\_\_

Califique las preguntas de 1 a 10, siendo 1 la calificación más baja y 10 la más alta.

1. Establezca el grado de dificultad que usted encontró para lograr la detección de la cabeza.

1    2    3    4    5    6    7    8    9    10  
                          

2. Establezca el grado de dificultad que usted encontró para coordinar los movimientos de la cabeza con los movimientos del cursor del *mouse*.

1    2    3    4    5    6    7    8    9    10  
                          

3. Establezca el grado de dificultad que usted encontró para activar los comandos del *mouse* por medio de los gestos.

a. *Click*.

1    2    3    4    5    6    7    8    9    10  
                          

b. *Doble Click*.

1    2    3    4    5    6    7    8    9    10  
                          

c. *Click derecho*.

1    2    3    4    5    6    7    8    9    10

d. *Click sostenido.*

1 2 3 4 5 6 7 8 9 10

4. Establezca el grado de dificultad que usted encontró para realizar tareas y ejecutar aplicaciones en el computador.

a. Abrir una carpeta de archivos.

1 2 3 4 5 6 7 8 9 10

b. Crear una carpeta.

1 2 3 4 5 6 7 8 9 10

c. Abrir y cerrar un archivo de texto.

1 2 3 4 5 6 7 8 9 10

d. Arrastrar una carpeta de un punto a otro en el escritorio.

1 2 3 4 5 6 7 8 9 10

e. Navegar en internet con la ayuda de un teclado virtual activado con SCODEM 2009.

1 2 3 4 5 6 7 8 9 10

5. ¿Cómo se sintió físicamente utilizando el sistema?

a. Dolor de cabeza.

1 2 3 4 5 6 7 8 9 10

b. Dolor de cuello.

1 2 3 4 5 6 7 8 9 10

c. Molestia en los ojos.

1    2    3    4    5    6    7    8    9    10  
                          

d. Dolor de espalda.

1    2    3    4    5    6    7    8    9    10  
                          

6. Califique en términos generales el sistema.

1    2    3    4    5    6    7    8    9    10