

DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA MULTITOUCH BASADO EN  
ILUMINACIÓN DIFUSA PARA SU APLICACIÓN EN OBRAS DE ARTE  
MULTIMEDIA



PABLO ANDRES PRADO DIAZ  
RICHARD HAMILTON GONZALEZ LUNA

Universidad  
del Cauca

UNIVERSIDAD DEL CAUCA  
FACULTAD DE CIENCIAS NATURALES, EXACTAS Y DE LA EDUCACIÓN  
DEPARTAMENTO DE FÍSICA  
POPAYÁN  
2014

DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA MULTITOUCH BASADO EN  
ILUMINACIÓN DIFUSA PARA SU APLICACIÓN EN OBRAS DE ARTE  
MULTIMEDIA

PABLO ANDRES PRADO DIAZ  
RICHARD HAMILTON GONZALEZ LUNA

Trabajo de investigación presentado como requisito parcial 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  
POPAYÁN  
2014

## **AGRADECIMIENTOS**

Los autores expresamos nuestros agradecimientos a:

Nuestros familiares por todo su trabajo, esfuerzo, valores y sacrificio, nuestros amigos por creer en esta noble causa y nuestros profesores por la excelente formación académica.

Nota de Aceptación

---

---

---

---

Director \_\_\_\_\_  
Ing. Edgar Leonairo Pencue Fierro

Evaluador \_\_\_\_\_  
Ing. Carlos Alberto Ospina Caicedo.

Evaluador \_\_\_\_\_  
MSc. Luis Fernando Echeverri.

Fecha de sustentación: Popayán, 21 de noviembre de 2014.

# CONTENIDO

<b>INTRODUCCIÓN .....</b>	<b>1</b>
<b>1. ARTE Y NUEVAS TECNOLOGIAS .....</b>	<b>2</b>
1.1. ARTE MULTIMEDIA .....	2
1.1.1. Arte Interactivo .....	2
1.2. INTERFAZ DE USUARIO. ....	3
1.2.1. Interfaz gráfica de usuario (gui).....	4
1.2.2. Interfaz natural de usuario (nui).....	4
<b>2. SUPERFICIES MULTITOUCH.....</b>	<b>5</b>
2.1. SUPERFICIES MULTITOUCH EN LA ACTUALIDAD.....	6
2.2. SUPERFICIES MULTITOUCH UTILIZANDO SISTEMAS ÓPTICOS.....	7
2.2.1. Esquema general .....	8
2.2.2. Comunicación usuario-computador .....	9
2.2.3. Comunicación computador-usuario.....	10
<b>3. FUNDAMENTOS TEORICOS DEL HARDWARE. ....</b>	<b>12</b>
3.1. SISTEMA DE VISUALIZACIÓN .....	13
3.1.1. Proyector .....	13
3.1.2. Pantalla de proyección .....	14
3.2. SISTEMA DE DETECCIÓN .....	14
3.2.1. Cámara infrarroja.....	15
3.2.2. Detección de toques por iluminación difusa.....	15
<b>4. FUNDAMENTOS TEORICOS DEL SOFTWARE .....</b>	<b>17</b>
4.1. SOFTWARE DE VISIÓN ARTIFICIAL .....	18

4.2. PROCESAMIENTO DIGITAL DE IMÁGENES.....	19
4.2.1. Espacios de color.....	20
4.2.2. Operaciones sobre vecindades.....	20
4.2.3. Operaciones en los niveles de gris .....	21
4.2.4. Operaciones morfológicas .....	22
4.3 PROTOCOLO TUIO.....	23
4.3.1. Transferencia de datos.....	23
<b>5. DESARROLLO DEL DISPOSITIVO.....</b>	<b>25</b>
5.1. DESARROLLO DEL HARDWARE .....	26
5.1.1. Elementos utilizados en este proyecto.....	26
5.1.2. Etapa 1, sistema de visualización .....	29
5.1.3. Etapa 2, sistema de detección .....	35
5.2. DESARROLLO DEL SOFTWARE DE VISIÓN ARTIFICIAL .....	36
5.2.1. Análisis general.....	37
5.2.2. Pre-procesamiento.....	38
5.2.3. Etapa 1, adecuación de la imagen.....	41
5.2.4. Etapa 2, filtro de niveles de gris .....	42
5.2.5. Etapa 3, filtro morfológico .....	44
<b>6. RESULTADOS Y DISCUSIÓN.....</b>	<b>47</b>
6.1. EVALUACIÓN DE CADA SISTEMA .....	47
6.1.1. Evaluación de la visualización .....	47
6.1.2. Evaluación de la detección .....	51
6.2. FUNCIONAMIENTO COMO UN DISPOSITIVO INTEGRADO .....	56
6.2.1. Puesta en funcionamiento .....	57
6.2.2. Utilidad del dispositivo .....	59
<b>7. CONCLUSIONES .....</b>	<b>61</b>

<b>TRABAJOD FUTUROS .....</b>	<b>62</b>
<b>REFERENCIAS.....</b>	<b>63</b>
<b>BIBLIOGRAFÍA .....</b>	<b>66</b>

## LISTA DE FIGURAS

Figura 2.1. Superficies multitouch con áreas de interacción grande .....	6
Figura 2.2. Esquema general.....	8
Figura 2.3. Filtrado de la señal.....	10
Figura 2.4. Gestos interactivos.....	11
Figura 3.1. Esquema detallado del hardware.....	12
Figura 3.2. Parámetro del proyector.....	14
Figura 3.3. Detección de toques.....	16
Figura 4.1. Esquema detallado del software .....	17
Figura 4.2. Ruidos de resplandor .....	18
Figura 4.3. Filtrado de imágenes.....	19
Figura 4.4. Análisis según los primeros vecinos.....	22
Figura 5.1. Componentes del hardware .....	26
Figura 5.2. Elementos del sistema de visualización .....	27
Figura 5.3. Iluminadores infrarrojos.....	28
Figura 5.4. Disposición electrónica iluminador tipo 2 .....	28
Figura 5.5. <i>Genius widecam 1050</i> .....	28
Figura 5.6. Diseño estructural del dispositivo .....	30
Figura 5.7. Componentes del área de interacción.....	30
Figura 5.8. Diagrama de rayos para el montaje de espejos .....	33
Figura 5.9. Implementación de la etapa 1 .....	34
Figura 5.10. Revestimiento del sistema de visualización .....	34
Figura 5.11. Disposición de iluminadores infrarrojos.....	36
Figura 5.12. Área de interacción en el rango del infrarrojo.....	36
Figura 5.13. Escena a recortar .....	38
Figura 5.14. Método integrales proyectivas.....	39
Figura 5.15. Región de interés (ROI) .....	40
Figura 5.16 Imagen pre-procesada. a) Imagen detectada, b) Resta de <i>background</i> .....	41
Figura 5.17. Pasos de la etapa 1.....	42
Figura 5.18. Máscara de enfoque, media.....	42
Figura 5.19. Ejemplo con ruido de nivel alto .....	43
Figura 5.20. Atenuación de niveles de ruido .....	44
Figura 5.21. Umbralización.....	44
Figura 5.22. Errores en las imágenes binarias .....	45
Figura 5.23. Eliminación de errores debidos a resplandores .....	46
Figura 6.1. Ubicación de fuente de iluminación exterior .....	48
Figura 6.2. Análisis de luminiscencia por zonas.....	48
Figura 6.3. Plantillas de calibración.....	49

Figura 6.4. Resultados de la calibración de luminiscencia .....	51
Figura 6.5. Disminución de reflejos de iluminadores infrarrojos .....	52
Figura 6.6. Evaluación de toques por zonas .....	53
Figura 6.7. Evaluación del filtro morfológico.....	53
Figura 6.8. Evaluación de resplandores cercanos.....	54
Figura 6.9. Detección del CCV .....	55
Figura 6.10. Estableciendo comunicación local.....	57
Figura 6.11. Prueba de identificación y seguimiento de toques .....	58
Figura 6.12. Dispositivo integrado .....	58
Figura 6.13. Continuidad de interacciones .....	59
Figura 6.14. Capacidades inmersivas del dispositivo.....	60
Figura 6.15. Capacidades colaborativas del dispositivo .....	60

## RESUMEN

En este documento se explica el desarrollo de una mesa interactiva, cuya interfaz de usuario se controla por medio de toques sobre su superficie, y está conformada por dos sistemas ópticos que trabajan en conjunto. Este proyecto es realizado con el objetivo de lograr una mayor fluidez en el manejo del computador, de tal forma que sea útil en obras de arte interactivo; por lo que resulta imperativo alcanzar ciertos estándares de calidad, para dar una buena impresión visual al usuario.

Primero se desarrolla un sistema de visualización donde se generan imágenes de la interfaz gráfica del computador, por medio de un proyector cuyo objetivo es una pantalla horizontal.

Por otro lado, se elabora un sistema de detección de toques funcional en el infrarrojo cercano, usando una cámara que apunta hacia la pantalla del sistema de visualización; iluminada en el infrarrojo de forma cuasi homogénea, para poder obtener un reflejo notable de los objetos que entran en contacto con esta.

Las imágenes digitales capturadas por la cámara, son tratadas para resaltar los reflejos de los toques del usuario, sustrayendo el fondo que contiene información no relevante. Los datos obtenidos después del procesamiento de la imagen, son utilizados para generar interacciones en la interfaz gráfica gracias a un protocolo de comunicación.

Finalmente, se verifica que los sistemas trabajan en conjunto de una forma óptima; para pasar a la evaluación llevada a cabo por medio de aplicaciones interactivas, desarrolladas para validar la respuesta del dispositivo ante los toques del usuario.

## OBJETIVOS

### Objetivo general

Diseñar e implementar un dispositivo *multitouch*, en donde la detección de toques se dará por medio de un sistema basado en iluminación difusa, proyectando a la vez una interfaz de usuario interactiva.

### Objetivos específicos

- Elaborar el sistema de visualización que sirva para proyectar imágenes en el área de interacción.
- Construir el sistema de detección de toques del usuario utilizando una cámara web e iluminadores infrarrojos.
- Desarrollar el software necesario que permita la detección de toques para su posterior implementación en otras aplicaciones.
- Diseñar y probar aplicaciones que sirvan para evaluar el funcionamiento del sistema final.

## INTRODUCCIÓN

Un dispositivo *multitouch* es aquel que despliega una interfaz gráfica permitiendo el accionar de varios puntos de interacción táctil, sobre una superficie plana de forma simultánea; sin la necesidad de utilizar controles periféricos. Este tipo tecnologías se han venido utilizando desde hace varios años en el arte, el entretenimiento y la educación; gracias a sus características inmersivas que brindan la posibilidad de internar al usuario en un tipo de interfaz cuyo rango de acción es mucho mayor, que en las interfaces gráficas comúnmente conocidas.

Las superficies *multitouch* se perfilan como una tecnología de gran potencial para el arte multimedia, donde ciertas tendencias apuntan a que se demande la acción del espectador para adquirir sentido y completar obras de arte; que dependen de la oportunidad que trae cada persona de forma única dentro de estas. Por esta razón, un dispositivo de este tipo puede jugar un rol activo para permitir al usuario interpretar una obra sintiéndose partícipe y creador; es decir, de forma interactiva.

Debido al auge de las superficies *multitouch*, está floreciendo con mucho éxito el uso de interfaces con soporte para la ejecución de aplicaciones colaborativas, que cuentan con la ventaja de incluir la interacción de múltiples usuarios a la vez; lo cual se traduce en la necesidad de implementar dispositivos con una superficie de operación relativamente grande.

En la actualidad hay diferentes técnicas para la detección de toques sobre una superficie *multitouch*. Refiriéndose a la fabricación de mesas interactivas, resulta más conveniente implementar sistemas basados en visión artificial; ya que al tratarse de superficies que se consideran de gran tamaño en este tipo de tecnologías, puede representar un problema económico y logístico usar otras técnicas de detección en un área de gran magnitud.

Este proyecto se enfoca en el desarrollo una mesa interactiva con superficie *multitouch*, donde las interacciones se obtuvieron por medio de un sistema de detección basado en visión artificial; mientras despliega su interfaz gracias a un montaje de espejos, realizado para satisfacer los estándares estéticos y de comodidad demandados por el usuario en general para este tipo de dispositivo.

## **1. ARTE Y NUEVAS TECNOLOGÍAS**

Al incorporar nuevas tecnologías como engranaje principal en el proceso de creación artístico, se cuenta con la posibilidad de abrir una nueva ventana de oportunidades en los modos de exhibición y circulación de las obras de arte.

El uso de nuevas tecnologías en el arte debe estar pensado para superar las limitaciones explicativas de los modelos deterministas y mecanicistas, ayudando a entregar la visión del artista al espectador de forma más dialógica [1]. Visto desde un enfoque técnico, es necesario atender múltiples procesos interconectándolos de forma sistemática, para así brindar una nueva gama de herramientas que ayuden a establecer un vínculo más estrecho en la relación que busque lograr el artista entre la obra y el espectador.

Gracias a la revolución digital, se ha acelerado el desarrollo de nuevas tecnologías que pueden entrar a jugar un rol de herramientas comunicativas para los artistas; cambiando así la forma en que puede concebirse una obra de arte, llegando en algunos casos a renovar el concepto de arte actual al establecer el objetivo de lograr una “estética digital”, donde se busca dar un toque humanístico a las nuevas tecnologías que entrega la ciencia.

### **1.1. ARTE MULTIMEDIA**

El arte multimedia es aquel en el que se coordinan varias formas de estimulación en el espectador como texto, sonido e imágenes fijas o en movimiento. El ejemplo más representativo de esta forma de arte es el cine, que desde sus inicios combinó más de un medio para cautivar al espectador; uno de estos fue el texto que seguía a las imágenes en movimiento en el cine mudo, en muchas ocasiones acompañado por música que ambientaba cada escena; esto impulsó el séptimo arte al punto de lograr con los años un gran impacto en el público al combinar secuencias de video y bandas sonoras [2].

#### **1.1.1. Arte interactivo**

Desde una vista contemporánea, el arte multimedia ya no se focaliza solamente en combinar varios tipos estimulación para presentar una obra al espectador. Debido a que el principal objetivo de los espacios de arte es enriquecer la relación entre la obra y el público, hoy en día se busca que se permita el control de la exhibición; haciendo partícipe al espectador, por medio del uso de nuevas tecnologías en instalaciones interactivas que parten de un

estado potencial que no cambia hasta que alguien lo activa, manipula o interfiere [3].

En el arte interactivo, una obra debe ser pensada como un sistema de relaciones siempre cambiante. Sin las acciones concretas del participante, la obra permanece en un estado de latencia, en un estado germinal incapaz de completar las posibilidades de las que la ha dotado su creador [4].

Durante el modelado de una obra interactiva, el artista define un conjunto de interacciones iniciales; define qué se vinculará con qué y de qué modo, es decir, define reglas de juego. El vínculo, la conexión misma toma estatuto de obra dejando al participante el propio desarrollo de tales reglas de juego [5].

La interpretación de una obra de arte exhibida en una instalación interactiva, requiere además de un compromiso intelectual del participante, un compromiso físico; debido a que la obra ya no se basa solo en la contemplación. Por tal motivo, hay que hablar de espectadores que son usuarios, ya que los participantes de una obra interactiva deben hacer uso de ella, operarla, estimularla; de otra manera carece de todo sentido y función [6].

Los desarrolladores de instalaciones interactivas tienen que tener en cuenta que la necesidad primordial a satisfacer en sus diseños, es minimizar la percepción de alguna línea divisoria entre la realidad física del usuario y la realidad virtual del dispositivo, que se encarga de los procesos sistemáticos en la instalación en general, sin caer en el error de que este mismo dispositivo sea más llamativo para el público que la obra de arte en sí. En este punto se puede afirmar entonces, que el vínculo buscado por el artista entre una obra interactiva y el espectador partícipe de esta, depende de la interfaz de usuario implementada por el desarrollador de la instalación interactiva.

Por esta razón, los diseños deben enfocarse en que el usuario tenga un control más natural del dispositivo que sirve como herramienta comunicativa en este tipo de arte, haciendo que el artista logre que su obra entre en un contacto dinámico con los sentidos del público; por lo tanto, las interacciones necesarias tienen que ir más allá de apretar un simple botón o mover un elemento en una pantalla con un dispositivo periférico, si se quiere envolver la percepción del usuario de una manera más inmersiva [1].

## **1.2. INTERFAZ DE USUARIO**

La interfaz de usuario es el medio que sirve como puente de comunicación entre las personas y la funcionalidad del computador, recibe órdenes e incorpora datos por medio de dispositivos como el teclado, el ratón, los controladores de juego y toda una variedad de dispositivos de entrada [7].

Debido a la creciente proliferación de la información y el auge de aplicaciones cuyo software es más complicado comparado con los desarrollos logrados hace algunos años; siempre ha existido un amplio interés por parte de una gran cantidad de desarrolladores hacia el tipo de tecnologías que permitan mejorar el modo en que el hombre interactúa con el computador, buscando con esto que la experiencia de uso sea más agradable y se logre de esta forma aprovechar las aplicaciones de software eficientemente.

A lo largo de los años se han ido mejorando las interfaces con el objetivo de orientar de una forma efectiva al usuario en su ejecución, haciendo que las funciones requeridas sean más perceptibles y apoyando un aprendizaje eficaz del contenido informacional de la máquina; lo que conlleva a que cada vez se logre llegar a un público más amplio. Por esta razón, el mundo ha sido testigo de la evolución de los tipos de interfaces de usuario; desde las líneas de comandos (CLI), a la interfaz gráfica de usuario (GUI), llegando a la interfaz natural de usuario (NUI); la cual es catalogada como la siguiente generación de las interfaces [8].

### **1.2.1. Interfaz gráfica de usuario (GUI)**

En este tipo de interfaz se permite la interacción del usuario con el computador por medio de la utilización de imágenes y objetos pictóricos (ventanas o íconos) además de texto [9]. Surgió como la evolución de las líneas de comando de los primeros sistemas operativos.

### **1.2.2. Interfaz natural de usuario (NUI)**

Es aquella en la cual se interactúa con un sistema haciendo uso de las indicaciones gestuales de las personas, aumentando así las modalidades de dispositivos de entrada, como seguimiento de movimiento, reconocimiento de voz, entre otras [8].

Debido a que los sistemas que utilizan NUI brindan una comunicación más agradable y sensible a las necesidades de las personas, se logra explotar en gran medida la capacidad de aprendizaje y el interés por parte de los usuarios hacia las nuevas tecnologías, lo que trae como consecuencia que las aplicaciones lleguen a un público mayor y puedan ser utilizadas en una variedad más amplia de disciplinas; explorando las posibilidades culturales y estéticas de los desarrollos informáticos de una forma más eficaz en comparación a los sistemas con los que se interactúa por medio de mandos periféricos.

## 2. SUPERFICIES *MULTITOUCH*

Las superficies *multitouch* son dispositivos de entrada y de salida, que tienen la capacidad de detectar varias interacciones táctiles del usuario, sobre un área en la cual se despliega la interfaz gráfica de un computador.

A diferencia de los dispositivos de entrada tradicionales que envían la señal de una interacción a la vez hacia el computador (teclado, ratón, entre otros); en las superficies *multitouch* se cuenta con una serie de gestos identificables en el comportamiento de los toques del usuario, sobre una superficie física que sirve como “área de interacción”, en la cual se presenta una comunicación del usuario con el computador y viceversa. Estos gestos pueden establecer respuestas más inmediatas por parte del contenido digital del computador, que para estos casos se le conoce como “aplicación interactiva”, logrando un vínculo más inmersivo entre el usuario el computador [10].

Como en los dispositivos de este tipo se interviene sin la necesidad de controles periféricos, gracias a que el área de interacción del usuario con la interfaz es la misma área de detección de toques; se ubican las superficies *multitouch* en la categoría de interfaz natural de usuario, aunque se siga valiendo de una interfaz gráfica para su funcionamiento.

En este tipo de tecnologías, el impacto que se quiere lograr depende primordialmente del tamaño del área de interacción en la que el usuario tiene posibilidad de participar en las aplicaciones interactivas desplegadas en la interfaz gráfica, más que de cualquier otra característica [11]. Siguiendo este orden de ideas, si se quiere que el usuario perciba su comunicación con el dispositivo de una forma más privada, entonces es necesaria un área de interacción pequeña, conociéndose este vínculo entre el usuario y el computador como de tipo personal.

Por otro lado, cuando el dispositivo tiene un área de interacción grande, las aplicaciones desplegadas en la interfaz gráfica pueden llamar fácilmente la atención de un público de múltiples personas que cuentan con la posibilidad de ser partícipes en la comunicación con el dispositivo a la vez, debido a que se puede desplegar más información al mismo tiempo comparado con un dispositivo de uso personal; lo que puede impulsar las superficies *multitouch* de tamaño grande como una herramienta idónea para aplicaciones que busquen la integración social de varios usuarios, siendo el vínculo entre el usuario y el computador de tipo cooperativo [12].

## 2.1. SUPERFICIES *MULTITOUCH* EN LA ACTUALIDAD

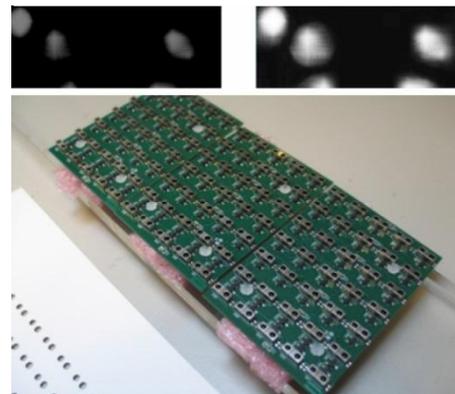
Hoy en día son de uso muy común los dispositivos que requieren de las interacciones táctiles del usuario para su funcionamiento; como son las tabletas, algunos modelos de celulares y computadores, entre otros. Para la detección de toques existen varias técnicas cuya utilización en el diseño del dispositivo varía en conveniencia económica y logística en su implementación, dependiendo del rango de tamaño del área de interacción del mismo y del problema que se hayan propuesto a solucionar sus desarrolladores.

En general, se destacan dos tipos de tecnologías empleadas para detectar las interacciones del usuario con las superficies *multitouch*; están las que se basan en situar una capa de sensores bajo la superficie del área de interacción para medir los cambios eléctricos inducidos por los contactos del usuario [13], y aquellas donde las interacciones son detectadas gracias a sistemas ópticos.

El motivo de estudio de esta sección son las superficies *multitouch* de tamaño grande, que tienen el potencial de desplegar aplicaciones cooperativas en la interfaz gráfica y pueden ser utilizados como mesas interactivas.



(a)



(b)



(c)



(d)

Figura 2.1. Superficies *multitouch* con área de interacción grande, a) *Diamond Touch* [14], b) *Thin Sight* [15], c) *Microsoft Surface* [16], d) *Reactable* [17].

Los dispositivos *multitouch* más conocidos que cuentan con las características mencionadas anteriormente son: *Diamond Touch*, *Thin Sight*, *Microsoft Surface* y *Reactable*, observados en la figura 2.1; en los cuales se utilizan técnicas de detección acordes a la amplia área de interacción.

En la *Diamond Touch* el área de interacción tiene integrado un arreglo matricial de “antenas” capacitivas, en el cual la detección se hace por medio de la descarga provocada por los toques del usuario [14]. Por otro lado, la *Thin Sight* detecta los toques por medio de una serie de sensores fotoeléctricos [15].

La *Microsoft Surface* y la *Reactable* funcionan gracias a sistemas ópticos, aprovechando los brillos que genera el usuario al entrar en contacto con el área de interacción debidamente iluminada [18].

En la *Reactable* se cuenta con la ventaja identificar piezas que tengan grabadas en su estructura figuras conocidas como “*fiducials*”, para establecer una interacción específica por medio de estas. Como está inspirada en los sintetizadores de los años setenta, las interacciones son respuestas musicales que pueden ser controladas por el usuario, ganando fama como una herramienta novedosa para *DJ's* desde que la artista islandesa *Björk* la utilizó en su gira *Volta* en el año 2007 [19].

## 2.2. SUPERFICIES MULTITOUCH UTILIZANDO SISTEMAS ÓPTICOS

Las superficies *multitouch* que funcionan gracias a sistemas ópticos son, en cierta medida, la opción más rentable para desarrollar un dispositivo con un área de interacción lo suficientemente grande que logre el óptimo funcionamiento de aplicaciones colaborativas entre varios usuarios.

Por esta razón, han sido reproducidos dispositivos con este tipo de sistemas en diferentes grupos de investigación, como también por desarrolladores independientes que siguen la filosofía *DIY (Do It Yourself)*. Todo esto gracias al aumento del interés del público en general en este tipo de tecnologías y al acceso a los elementos de hardware y herramientas computacionales necesarias para su desarrollo, que hace una década no se tenía.

En 2006 fue fundada por Christian Moore la comunidad *nui group* [20], cuyo objetivo es mantener un registro de los avances relacionados con interfaces naturales de usuario a través de su página web; que cuenta con foros donde se discute y comparte información acerca del progreso de este tipo de tecnologías alrededor del mundo.

En cuanto a superficies *multitouch* se refiere, se puede encontrar documentación de varios proyectos elaborados por medio de sistemas ópticos,

que sirven como referente para aquellos que busquen desarrollar un dispositivo de esta clase.

### 2.2.1. Esquema general

Las superficies *multitouch* que funcionan gracias a sistemas ópticos se componen de varios elementos cuya configuración en general se muestra en la figura 2.2.

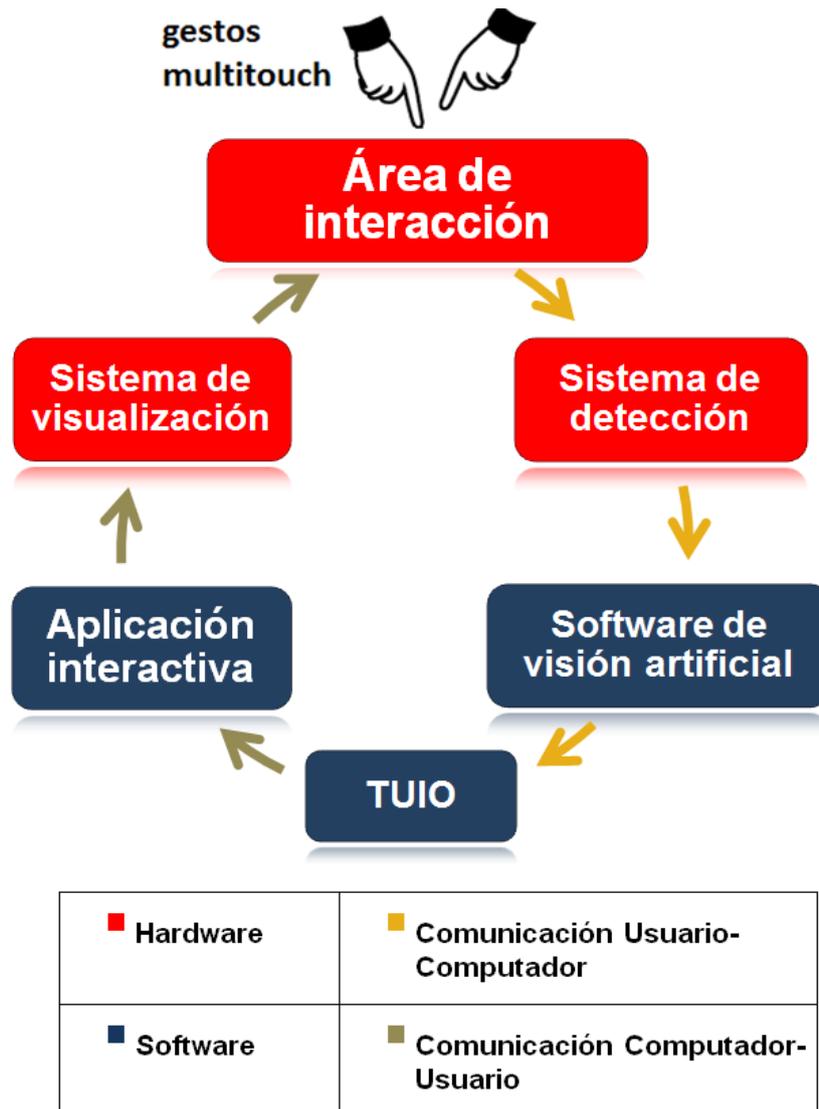


Figura 2.2. Esquema general.

Según este esquema, el hardware cumple con dos etapas; la primera es un “sistema de detección” de toques, que sirve para obtener la información de entrada, y la segunda para entregar la salida, que es el “sistema de visualización” de la interfaz gráfica.

El sistema de visualización consta primordialmente de un proyector y una pantalla de proyección que se constituye como un elemento primordial para el área de interacción. El diseño de todo el dispositivo está orientado en gran parte por las propiedades del proyector; por otro lado, el conocimiento de las propiedades del material de la pantalla de proyección, ayuda a tener una noción de la apreciación cualitativa de la calidad visual percibida por el usuario en la interfaz gráfica.

En el sistema de detección la señal de entrada es video capturado por medio de una cámara apuntada hacia el área de interacción, en el cual la utilidad de la información que de este se obtenga depende de las condiciones de iluminación. Debido a esto, es necesario utilizar elementos iluminadores acordes a las técnicas establecidas para conseguir una distribución de intensidad repartida lo más uniformemente posible sobre el área de interacción.

La cámara y los elementos iluminadores deben trabajar en el rango del infrarrojo cercano; debido a que el cambio de las imágenes generadas en el área de interacción por el sistema de visualización en el rango de la luz visible, no puede significar ruido en las imágenes capturadas por el sistema de detección.

Las zonas que están siendo tocadas de forma tangible por parte del usuario en el área de interacción, conocidas como “áreas de contacto”, deben ser estimadas por medio de un software de visión artificial cuya complejidad varía dependiendo de la técnica de iluminación usada en el sistema de detección.

Una vez extraídas digitalmente las áreas de contacto en las imágenes capturadas por la cámara; se utiliza el protocolo TUIO, explicado detalladamente en el capítulo 4, para comunicar las posiciones de las interacciones correspondientes en la interfaz gráfica mostrada por el sistema de visualización.

### **2.2.2. Comunicación Usuario-Computador**

Las imágenes obtenidas por la cámara son como la mostrada en la figura 2.3a, y deben ser filtradas para atenuar a una mínima proporción el ruido que se presenta en la señal de entrada, el cual en un dispositivo de este tipo provocaría una respuesta en la interfaz gráfica a “interacciones fantasma” que en realidad no están ocurriendo. El ruido que se presenta en la detección es dependiente de las variaciones de iluminación en el ambiente donde se sitúe el dispositivo.



(a)

(b)

Figura 2.3. Filtrado de señal a) Señal de entrada, b) Señal filtrada.

Después de la etapa de filtrado realizado por el software de visión artificial, se debe obtener una imagen como aquella que se observa en la figura 2.3b, para luego comunicar esta información por medio del protocolo TUIO de forma local utilizando el puerto 3333.

Existe una amplia gama de herramientas virtuales para desarrolladores encaminadas a mejorar la eficiencia, precisión y usabilidad del software de visión artificial encargado de la detección de los toques en el área de interacción.

Las más referenciadas en la literatura son el CCV (*Community Core Vision*) [21] y el *Touchlib* [22], que permiten filtrar las imágenes obtenidas por la cámara variando diferentes propiedades en una interfaz gráfica amigable, para aquellos que no acostumbran moverse en entornos de programación. Pero se presenta una desventaja al entrar grandes niveles de ruido cuando varía la iluminación de la habitación en donde esté funcionando el dispositivo.

### 2.2.3. Comunicación Computador-Usuario

Después de que la señal filtrada es comunicada al protocolo TUIO, este se encarga de enviar paquetes de mensajes UDP (protocolo para comunicaciones locales y en red), que contienen información de cada toque, como identificación, posición, velocidad, entre otras características. La aplicación desplegada en la interfaz gráfica sirve como cliente, que recibe estos mensajes para generar las interacciones que habitualmente corresponden a los gestos mostrados en la figura 2.4.

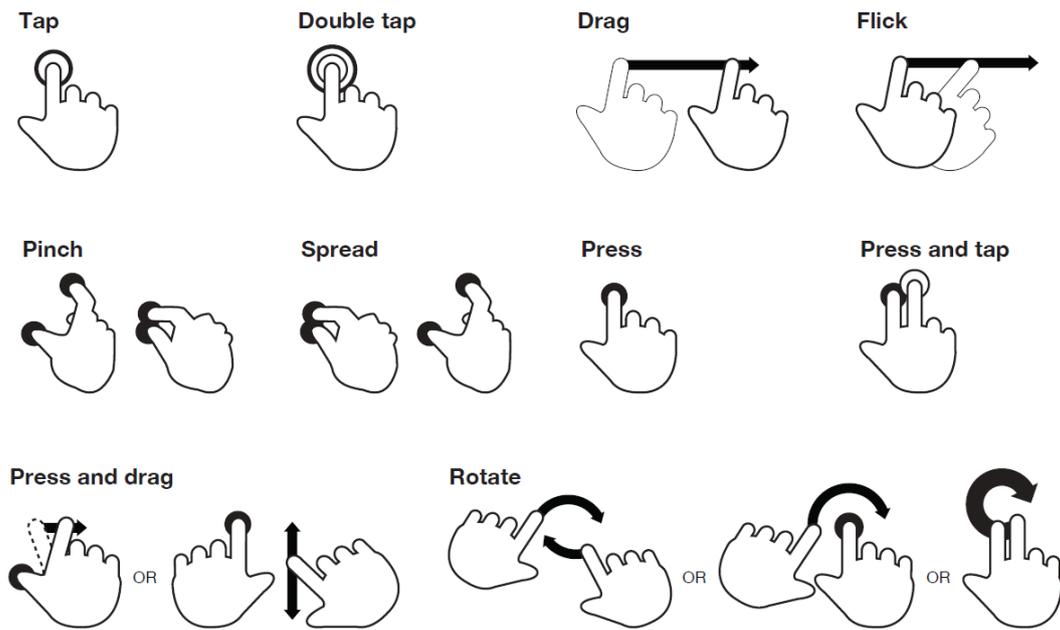


Figura 2.4. Gestos interactivos [23].

El *framework* seleccionado para evaluar el funcionamiento del dispositivo es *Processing 2.1* [24], que tiene un entorno de programación *open source* con una interfaz sencilla; que actualmente es utilizado por artistas, diseñadores, arquitectos y aficionados para la creación de interfaces gráficas básicas de proyectos multimedia e interactivos de diseño digital.

### 3. FUNDAMENTOS TEÓRICOS DEL HARDWARE

En general, los elementos básicos del hardware utilizado en el desarrollo del dispositivo fueron: una cámara infrarroja, iluminadores de lámpara LED infrarrojos, un proyector y una pantalla de proyección. El montaje de estos elementos está diseñado para cumplir con el esquema detallado de la figura 3.1.

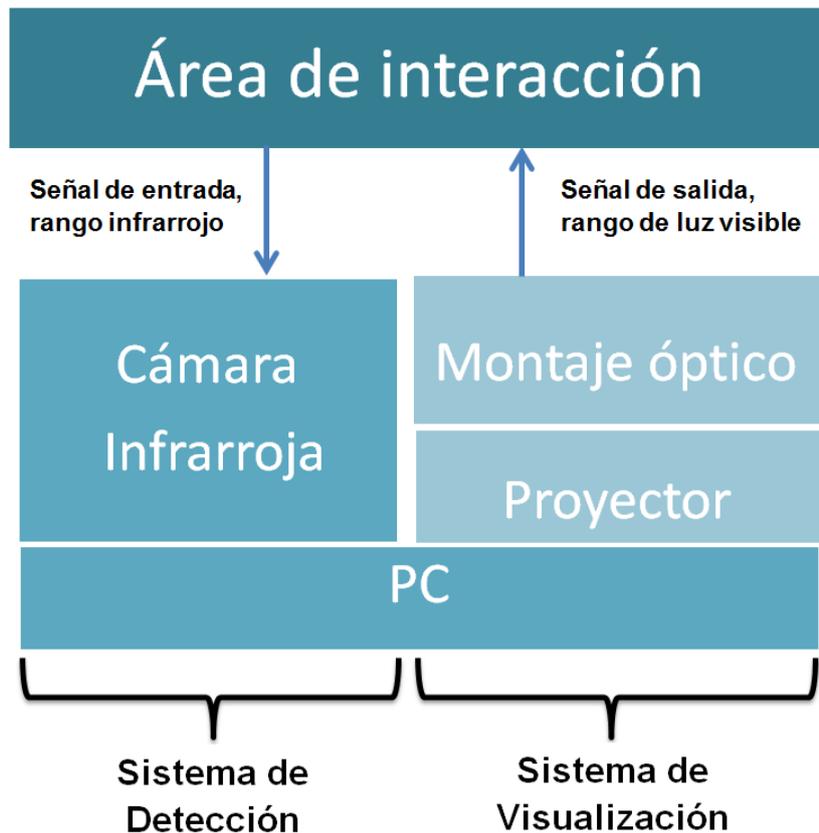


Figura 3.1. Esquema detallado del hardware.

Todo el hardware se divide en dos sistemas, uno de ellos es el de detección, funcional en el rango del infrarrojo cercano y sirve para obtener la señal de entrada desde el área de interacción. La configuración de este sistema depende de la que se implemente en el de visualización, que se encarga de entregar la señal de salida en el rango de la luz visible sobre la misma área de interacción, sin que interfiera con la otra señal y viceversa.

Por otro lado, el montaje óptico que hace parte del sistema de visualización, sirve para el buen desempeño de ambas etapas del hardware; ya que su configuración es utilizada para la obtención de la señal de entrada y para desplegar la señal de salida. Este montaje se encarga de aprovechar de la mejor manera posible las propiedades del proyector en el sistema de visualización y de los iluminadores en el sistema de detección.

### 3.1. SISTEMA DE VISUALIZACIÓN

La calidad y el impacto visual que se busque lograr por medio de la interfaz gráfica, dependen de las propiedades de los elementos del sistema de visualización, y de qué forma están siendo aprovechados en el diseño del dispositivo.

Como el dispositivo se diseñó para su uso como mesa interactiva, el área de interacción debe ubicarse de forma horizontal con el proyector apuntado hacia esta desde la parte inferior. Esto significa que la altura del dispositivo es dependiente de la magnitud que se busca para el área de interacción. Si se proyecta directamente sobre el área de interacción se obtendría una interfaz pequeña, haciendo que el dispositivo sea percibido como de uso personal.

Por esta razón, se decidió implementar un montaje de espejos, que ayude a optimizar el espacio ocupado por el camino óptico medio de los rayos emitidos desde el proyector.

#### 3.1.1. Proyector

Es el elemento representativo del sistema de visualización, ya que la cantidad de luz que emite medida en lumens, o luminiscencia; determina primordialmente las condiciones de iluminación máxima en el ambiente donde se sitúe el dispositivo, para que el usuario observe las imágenes de una forma óptima en cuanto a las condiciones de brillo. Por otro lado, el diseño estructural del sistema de visualización y por ende del dispositivo en general, debe desarrollarse teniendo en cuenta las características que se describen a continuación.

- **Throw ratio.** Esta propiedad relaciona el tamaño del área de interacción con respecto a su distancia con el lente del proyector [25]. Gracias al *throw ratio* se puede determinar la cantidad de espejos que deben montarse, dependiendo de la altura y el tamaño de pantalla que se busca obtener para el dispositivo.
- **Offset.** Se refiere a la calibración en cuanto a la forma de la imagen que permite cada proyector. Esta característica concede la posibilidad de utilizar el proyector a diferentes ángulos con respecto a la normal del área de interacción [26]. Dependiendo del *offset* del proyector utilizado en el sistema de visualización, se tiene un rango de ángulos posibles para disponer de este mismo en el diseño del dispositivo que se está desarrollando; si el proyector se acomoda fuera de este rango, la imagen se verá en forma de trapecio con los objetos pictóricos de la interfaz gráfica deformados.

En la figura 3.2 se observan los parámetros descriptores del proyector; donde  $D$  es el camino óptico medio de los rayos emitidos,  $L$  la altura de las imágenes desplegadas en el área de interacción, y  $\alpha$  es el ángulo con respecto a la normal de la misma donde el offset del proyector, permite obtener una imagen aproximadamente rectangular.

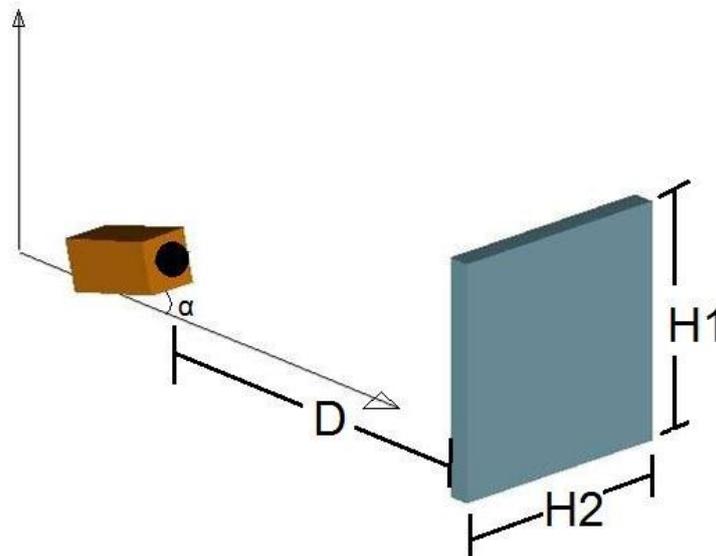


Figura 3.2. Parámetros del proyector.

### 3.1.2. Pantalla de proyección

El color de la pantalla determina la calidad de contraste que esta puede entregar dependiendo de la iluminación ambiental; es decir, la capacidad que esta tiene para poder formar una imagen donde el usuario logre diferenciar de manera óptima el brillo de los tonos blancos o de colores vivos, con respecto a las regiones de la imagen que se supone deben verse más oscuras [26].

En habitaciones iluminadas conviene utilizar una pantalla preferiblemente negra para que la luz exterior al dispositivo no aumente el brillo percibido por el usuario en las zonas oscuras de la imagen proyectada; mientras que para ambientes con poca iluminación se busca aprovechar la intensidad luminosa de las partes que más brillo presentan en la imagen por medio de una pantalla blanca.

### 3.2. SISTEMA DE DETECCIÓN

Los datos que componen la señal de entrada del sistema son imágenes capturadas por medio de la cámara infrarroja; cuyo sensor arroja una colección finita de medidas con valores asociados a un conjunto de números enteros conocidos como píxeles, pertenecientes a matrices bidimensionales conocidas

como planos. Esto es lo que se conoce como digitalización de la escena, que para el caso de este proyecto se centra en el área de interacción [27].

La fluidez de las interacciones del usuario con el dispositivo depende en gran medida de las propiedades con las que cuenta la cámara infrarroja, que es el elemento representativo en el sistema de detección. Por otro lado, la posibilidad de diferenciar por medio de visión artificial las interacciones en las imágenes capturadas, depende de la técnica de detección que se utilice y de la rigurosidad del desarrollador al implementarla; en el caso de este proyecto se utilizó iluminación difusa.

### **3.2.1. Cámara infrarroja**

La señal de entrada, es video que contiene una sucesión de imágenes de la escena, que son capturadas a cierta frecuencia de muestreo medida en cuadros (imágenes) por segundo, donde la utilidad de la información obtenida, depende de la cantidad de píxeles que están siendo usados para describir los niveles de intensidad infrarroja en el área de interacción; es decir, la resolución de esta última.

La cámara infrarroja debe acomodarse con el fin de que la resolución que describe al área de interacción sea la mayor posible, teniendo en cuenta que el tamaño de esta en la escena detectada, varía según la distancia de la cámara. Esta distancia es dependiente del ángulo de apertura ligado al lente que opera antes del sensor de esta última.

La resolución que se logre para el área de interacción influye en la precisión del reconocimiento de los toques de usuario; por otro lado, los cuadros por segundo que captura la cámara, determina la velocidad máxima en la que una interacción en movimiento va a ser detectada como una serie de toques individuales; es decir, la velocidad máxima en que se presenta identificación y seguimiento de los mismos.

### **3.2.2. Detección de toques por iluminación difusa**

En esta técnica se detectan los brillos de los toques que se presentan en el área de interacción, previamente iluminada desde la parte inferior de la misma. Los elementos iluminadores para esta técnica, son lámparas que deben ubicarse a cierta distancia del área de interacción; distancia que depende de la magnitud de intensidad luminosa que entregue la lámpara y de que tan uniformemente distribuida este la luz infrarroja, emitida por esta.

Según la configuración mostrada en la figura 3.3a, se busca que los haces luminosos incidan sobre el área de interacción desde varias direcciones; con el objetivo de favorecer en menor cantidad algunas zonas más que a otras sobre la misma [28], disminuyendo además la presencia de sombras. Al mencionar zonas se refiere a las resultantes cuando se separa el área de interacción con líneas divisoras como las de la figura 3.3b.

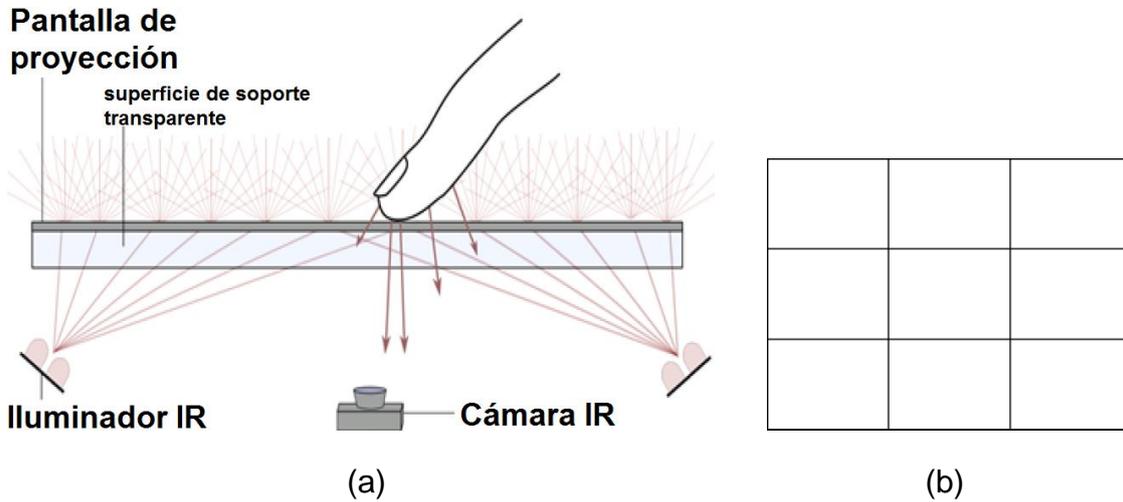


Figura 3.3. Detección de toques, a) Iluminación difusa [29], b) División por zonas del área de interacción, virtual vista superior.

Cuando el usuario toca la pantalla, la cámara captura un patrón de manchas brillantes, que sirven para identificar la pequeña superficie que está entrando en contacto con el área de interacción, y se vuelven borrosas conforme se alejan los dedos del usuario al otro lado de la pantalla; gracias a que esta última se compone por un material difusor semitransparente, donde los haces infrarrojos se encuentran con pequeñas rugosidades, que hacen que se transmitan en direcciones aleatorias [30].

Las áreas de contacto, deben ser estimadas por medio de la manipulación de las imágenes digitales capturadas por la cámara, dependiendo de la distancia de los objetos detectados con respecto a la pantalla.

#### 4. FUNDAMENTOS TEÓRICOS DEL SOFTWARE

El esquema detallado de los pasos que sigue el software desde que recibe la señal de video obtenida por la cámara, hasta la comunicación de los toques identificados como interacciones, se puede analizar en la figura 4.1.



#### □ Software de Visión Artificial

Figura 4.1. Esquema detallado del software.

En general, el software que necesita implementarse para el funcionamiento digital del dispositivo, se divide en dos tipos de procedimientos fundamentales. El primero se encarga de filtrar los niveles de ruido por medio del procesamiento de cada una de las imágenes de la señal de video, para evitar en mayor medida la ocurrencia de las interacciones fantasma detectadas durante la comunicación usuario-computador, dando como resultado un software de visión artificial como el desarrollado en este proyecto; mientras que el segundo procedimiento fundamental, se encarga de todos los pormenores a tener en cuenta antes de implementar la aplicación interactiva durante la comunicación computador-usuario, y se realiza a través de TUIO.

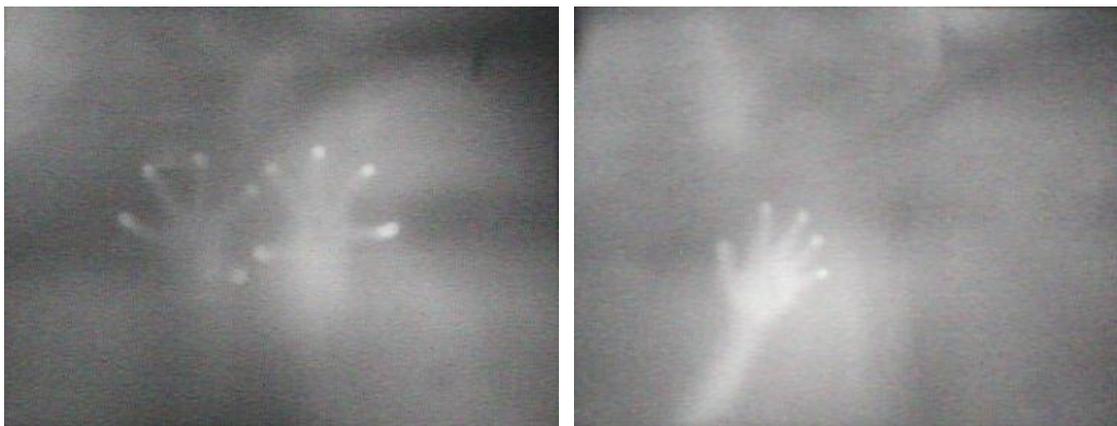
Dependiendo de los métodos utilizados durante el procesamiento digital de las imágenes, se logran alcanzar ciertos estándares para eliminar el ruido; ligado a las condiciones de iluminación exterior en el momento de realizar la captura de la imagen a modificar, y representado por pixeles que no describen valores asociados a las áreas de contacto. Además, con el procesamiento de las imágenes se compensa la falta de uniformidad en las condiciones de iluminación entre las diferentes zonas del área de interacción, que por cuestiones técnicas se presenta en el sistema de detección.

#### 4.1. SOFTWARE DE VISIÓN ARTIFICIAL

Las imágenes de salida después de todo el procesamiento digital son de naturaleza binaria; es decir, cuentan con la posibilidad de que se presenten solo dos colores, el blanco y el negro. El color blanco corresponde a las áreas de contacto, mientras que el color negro ha sido asignado al fondo, donde se ubican todos los píxeles del área de interacción de los que se puede afirmar, que la información aportada durante su análisis es considerada como ruido.

Existen dos fuentes de ruido en la imagen a analizar; esta aquella que proviene de las zonas que no presentan variaciones de intensidad bruscas, con respecto a los valores de iluminación inicial durante la detección, ruido considerado como de nivel bajo; y el ruido que proviene de los resplandores formados como consecuencia de la reflexión de los rayos infrarrojos que se presenta en los objetos al exterior del dispositivo. Este segundo tipo de ruido es el que más debe atenuarse durante el procesamiento digital de la imagen, ya que cuenta con una mayor susceptibilidad a las variaciones de iluminación exterior.

Al referirse a “ruido de resplandor” en este documento, se hace alusión a los niveles de intensidad detectados, que provienen específicamente de la reflexión de la luz infrarroja en los brazos del usuario; clasificándose como ruido de nivel medio y alto. El ruido de nivel medio es el que muestra valores de intensidad diferenciables con los mínimos presentados en las áreas de contacto, como se observa en el ejemplo de la figura 4.2a; mientras que el ruido de nivel alto surgen cuando el resplandor alcanza valores comparables con las áreas de contacto, y se presenta en imágenes como la de la figura 4.2b.



(a)

(b)

Figura 4.2. Ruidos de resplandor. a) Ruido de nivel medio, b) Ruido de nivel alto.

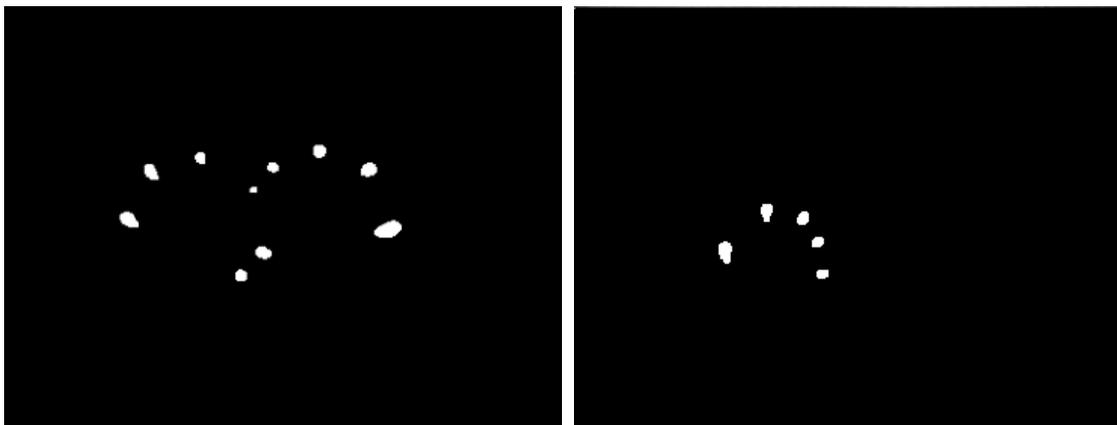
La primera etapa del software busca extraer las propiedades que sirven de forma más efectiva como descriptores en la imagen detectada. Posteriormente,

la segunda etapa atenúa todos los niveles de ruido, y resalta las áreas de contacto para que se diferencien efectivamente del fondo.

Posteriormente, se convierte la imagen a una de tipo binario; eliminando el ruido para lograr identificar las áreas de contacto y el fondo solo por medio de dos valores posibles, cero para el color negro que representa el fondo y uno para las áreas de contacto. A este paso se le conoce como umbralización de la imagen, en donde el número establecido como criterio para diferenciar a los pixeles entre sí, se le conoce como umbral.

En la última etapa, se eliminan las áreas pequeñas que han superado los filtros anteriores debido a los valores de intensidad que presentan desde la detección, pero que no pueden ser tomadas como información relevante por su tamaño; además de modificar la forma de las áreas de contacto, actuando como procedimiento de corrección final, que entrega información más precisa y consistente entre los cuadros de la señal de video, al protocolo TUIO.

Después del filtrado que debe realizar el software de visión artificial por medio de las etapas mencionadas; se deben obtener imágenes binarias como las que se muestra en la figura 4.3, que son resultados correspondientes a los tipos de ruido de resplandor utilizados en el ejemplo anterior.



(a)

(b)

Figura 4.3. Filtrado de imágenes. a) Ruido de nivel medio eliminado, b) Ruido de nivel alto eliminado.

## 4.2. PROCESAMIENTO DIGITAL DE IMÁGENES

Las imágenes están representadas numéricamente por medio de un conjunto de matrices de pixeles denominadas como planos, cuyos valores oscilan en una escala de 256 niveles de gris, relacionados a los niveles de intensidad luminosa en toda la escena; donde 0 es su valor más bajo correspondiente a una zona no iluminada o que absorbe todos los rayos de luz, y 255 es su valor

máximo representando una zona muy iluminada o que refleja todos los rayos que inciden sobre esta.

El procesamiento de la imagen se basa en una serie de transformaciones matemáticas de matrices bidimensionales que se denotan por letras mayúsculas, o también pueden ser representadas como funciones de la forma  $g(i,j)$ , refiriéndose al nivel de gris del pixel perteneciente a la coordenada espacial denotada por los valores  $(i,j)$  [31].

Las operaciones en los niveles de gris buscan como resultado una imagen modificada, donde el ruido que se plantea en la sección anterior está eliminado. El proceso de umbralización consiste en dividir la escala de niveles de gris en dos grupos; logrando así discriminar los valores de cada pixel según un número establecido, para luego asignarle uno de los dos valores posibles a su equivalente en las coordenadas de la imagen resultante.

Las operaciones morfológicas realizadas en las imágenes binarias que arroja el proceso de umbralización, se utilizan con el objetivo de lograr una mayor precisión en cuanto al número de toques detectados, además de lograr una efectiva congruencia en la ubicación espacial de los centros de masa de los toques, entre cuadro y cuadro de la señal de video enviada al protocolo TUIO.

#### **4.2.1. Espacios de color**

Son una representación matemática con la que se puede especificar, crear o visualizar cualquier color. Los espacios de color se representan como imágenes que combinan varios planos; que al ser analizados como diferentes canales de información en la señal de entrada, aportan cada uno descriptores diferentes para procesar la imagen.

Dependiendo de las limitaciones del tipo de sensor y de las aplicaciones en los que se usa la imagen que este entrega, se utilizan diferentes espacios de color [32]. En este proyecto se trabaja con los espacios de color RGB, HSV, Y'CbCr y XYZ.

#### **4.2.2. Operaciones sobre vecindades**

Son aquellas que se realizan para procesar bloques de la imagen en lugar de procesarla toda. Estos bloques son denominados kernel, matriz de convolución o máscara; y se componen de los pixeles correspondientes al entorno del pixel que se está analizando, generalmente situado en el centro del kernel y cuyo valor en la imagen resultante, varía dependiendo de los valores originales de

sus vecinos, según las operaciones matemáticas que se apliquen durante el proceso.

### 4.2.3. Operaciones en los niveles de gris

Son todas aquellas operaciones que se realizan para obtener una imagen que aún contenga los valores de sus píxeles repartidos en la escala de grises, pero con modificaciones planeadas para que la posterior etapa de umbralización arroje información efectiva. Las operaciones en niveles de gris utilizadas por el software desarrollado en este proyecto son las que se explican a continuación.

- **Manipulación de contraste.** Busca distanciar los valores de los píxeles en la escala de grises. Se utiliza una transformada que cambia el valor máximo y el mínimo de niveles de gris que tiene la imagen; reacomodando el resto de valores de forma lineal. Se tiene entonces como resultado una imagen con el contraste modificado cuyos píxeles están repartidos en un nuevo rango de niveles de gris [33].

Siendo  $m_1$  y  $m_2$  los valores mínimo y máximo respectivamente, se puede definir la imagen de salida  $G'$ , siguiendo la ecuación 1.

$$G' = \frac{\alpha}{m_2 - m_1} (G - m_1), \quad \text{ecuación 4.1.}$$

Donde  $G$  es la imagen de entrada y  $\alpha$  es el valor máximo que se desea en imagen de salida.

- **Filtro de media.** Uno de los filtros lineales más simples donde el valor de cada píxel es reemplazado por el valor medio del mismo y sus vecinos. Viene descrito por la ecuación 2.

$$G'(i, j) = \frac{1}{M} \sum_{k=i-m}^{k+m} \sum_{l=j-n}^{l+n} G(k, l), \quad \text{ecuación 4.2.}$$

Donde el kernel es de tamaño  $(2m+1) \times (2n+1)$ , y  $M$  es el número total de píxeles dentro de este.

- **Filtro espacial gaussiano.** En este tipo de filtro los valores del kernel varían intentando imitar la forma de una función gaussiana como la descrita en la ecuación 3.

$$G'(i, j) = e^{-\frac{(i+j)^2}{2\sigma^2}}, \quad \text{ecuación 4.3.}$$

Las variaciones de  $\sigma$  varían la amplitud resultante de los píxeles, y por ende el nivel de contraste en la imagen resultante.

- **Máscara de enfoque.** Es un filtro que utiliza una modificación de la imagen de entrada, como una especie de “plantilla” para atenuar ciertos rangos de niveles de gris y darle importancia a otros [34]. La ecuación 4 es la que se emplea para lograr este cometido.

$$G' = G + \beta * D, \quad \text{ecuación 4.4.}$$

Dependiendo del tratamiento previamente hecho a la imagen plantilla  $D$ , se logra una clase de efecto diferente en la imagen resultante  $G'$ ; mientras que la constante  $\beta$  establece la magnitud en que este mismo efecto se aplica.

#### 4.2.4. Operaciones morfológicas

Se basan en el cambio de los valores de los píxeles al otro valor posible en una imagen binaria, dependiendo del resultado que se busque obtener. Las operaciones morfológicas son de naturaleza lógica, ya que se compara el valor del píxel con los de sus vecinos por medio de algebra booleana para realizar el cambio anteriormente mencionado. Se realizan para los primeros cuatro u ocho vecinos, ubicando los píxeles objeto de análisis en el centro de alguno de los kernel mostrados en la figura 4.4.

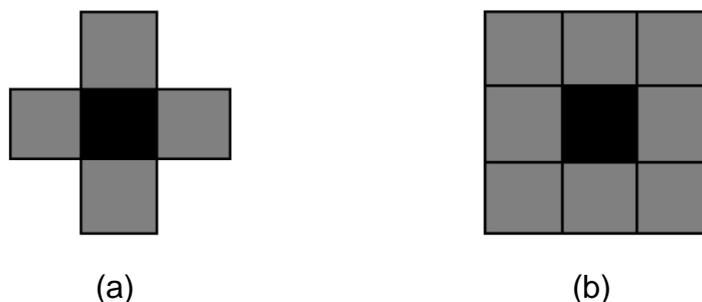


Figura 4.4. Análisis según los primeros vecinos [36]. a) Primeros cuatro vecinos, b) Primeros ocho vecinos.

Las operaciones morfológicas utilizadas en este proyecto son las que se reseñan a continuación:

- **Erosión.** Los píxeles a analizar son aquellos que tienen un valor igual a 1 en la imagen binaria, describiendo a las áreas de contacto. Se valora la existencia del número 0 entre los vecinos del píxel; de ser así, se asigna el otro valor posible en la misma coordenada de la imagen

resultante, de lo contrario se asigna el valor que tiene el pixel originalmente antes del análisis morfológico.

- **Dilatación.** Esta vez los pixeles a analizar son aquellos que tienen un valor igual a 0 en la imagen binaria, describiendo el fondo. el razonamiento para el análisis del pixel es equivalente al llevado a cabo durante la erosión, solo que en esta ocasión se valora la existencia del número 1 entre los vecinos del pixel.

### 4.3. PROTOCOLO TUIO

Es un protocolo diseñado para satisfacer los requerimientos de las plataformas que utilicen interfaces naturales de usuario [35], como en el dispositivo desarrollado en este proyecto.

TUIO fue creado por los programadores del proyecto reactIVision [36], y su implementación se hizo a partir de un protocolo más general denominado *Open Sound Control*, razón por la cual debe seguir su sintaxis. Este protocolo define propiedades comunes de las áreas de contacto, identificándolas como *fiducials* o gestos de los dedos y manos, para transmitir esta información desde una aplicación servidor hacia aplicaciones clientes.

La actual implementación del protocolo permite que las aplicaciones servidor o cliente se desarrollen independientemente en diferentes lenguajes de programación; utilizando transporte de datos UDP para hacer eficiente la transferencia [37].

#### 4.3.1. Transferencia de datos

El protocolo define dos clases de mensajes, *Set* y *Alive*. Los mensajes *Set* se utilizan para comunicar información sobre el estado del objeto, como su posición u orientación; mientras que los mensajes *Alive* indican el conjunto actual de objetos presentes tocando el área de interacción, usando una lista de indicadores numéricos [38]. Se empaquetan varios mensajes *Set* de forma redundante para completar la capacidad de transporte UDP, con el fin de evitar pérdidas de información, y en cada uno se incluye un mensaje *Alive*. El formato general para el envío y recepción es el siguiente para los mensajes *Set* y *Alive* respectivamente:

*/tuo/[Nombre del Perfil] set [número de identificador del objeto] [Parámetros].*

*/tuo/[Nombre del Perfil] alive [lista de identificadores de objetos activos] .*

Dependiendo del tipo de objetos que se estén identificando en el área de interacción, se tienen distintos Nombres de Perfil que se pueden utilizar en el protocolo; por ejemplo, un elemento detectado puede ser de dos tipos, *fiducial* (un área específica) o cursor (un solo punto), y estos a su vez pueden funcionar en dos o más “dimensiones”, dependiendo de los parámetros requeridos por la aplicación interactiva que sirve como cliente [37].

En este proyecto, se utiliza el envío de datos que identifican los gestos relacionados a elementos detectados de tipo cursor, correspondientes a los generados por los dedos del usuario que entran en contacto con el área de interacción.

## 5. DESARROLLO DEL DISPOSITIVO

El diseño del dispositivo está encaminado para que se obtenga una amplia área de interacción, siendo a la vez cómoda su utilización para el usuario y con la motivación de obtener un dispositivo como los mostrados en la figura 2.1. El funcionamiento del dispositivo se basa en sistemas ópticos por su coste de construcción, escalabilidad y robustez.

Es de notarse que se considera el área de interacción como el elemento principal de todo el hardware, ya que es el puente de comunicación con el usuario, por lo que se establece como el centro de atención del dispositivo cuando se encuentra en pleno funcionamiento.

Para la detección por iluminación difusa, el material necesario es accesible comparado con otras técnicas utilizadas comúnmente en este tipo de sistemas; por otro lado, es más difícil conseguir una iluminación uniforme, lo que trae como consecuencia la disminución de la precisión del software de visión artificial [39].

Por otro lado, en cuanto a la implementación del software que permite la respuesta a los toques del usuario que se presentan de forma tangible sobre el área de interacción, con interacciones digitales en la interfaz gráfica; se puede marcar diferencia con respecto a los proyectos realizados bajo la filosofía DIY, estableciendo una metodología de filtrado digital para las imágenes resultantes del sistema de detección, que sea exclusiva para el dispositivo desarrollado en este proyecto.

Tomando en cuenta que los lineamientos generales de seguimiento y comunicación local están establecidos por el protocolo TUIO, siendo un común denominador en la mayoría de los proyectos de este tipo; mientras que la generación de gestos interactivos es abordada en el *framework* utilizado para programar la aplicación interactiva (en el caso de este proyecto, el *framework* es Processing 2.1).

De los requerimientos iniciales como tamaño y ubicación del área de interacción dependen los parámetros iniciales que se deben tener en cuenta en el diseño general del dispositivo. Además, la eficiencia de la comunicación usuario-computador y por ende, la capacidad de lograr un vínculo cómodo y natural entre el humano y la máquina, depende primordialmente de los métodos utilizados para el filtrado digital de los toques durante el procesamiento de las imágenes detectadas en el rango del infrarrojo.

Al mencionar un vínculo cómodo y natural, se hace alusión a una detección de toques de manera no interrumpida a lo largo del tiempo en el que los dedos del

usuario entran en contacto con el área de interacción; y con una menor probabilidad de que se presenten interacciones fantasma.

## 5.1. DESARROLLO DEL HARDWARE

Los elementos en general para su implementación se muestran en la figura 5.1. La primera etapa para desarrollar todo el dispositivo, debe ser la implementación del sistema de visualización; ya que la variación de la posición de los elementos que lo componen, tiene consecuencias en las condiciones de iluminación infrarroja para la implementación del sistema de detección.

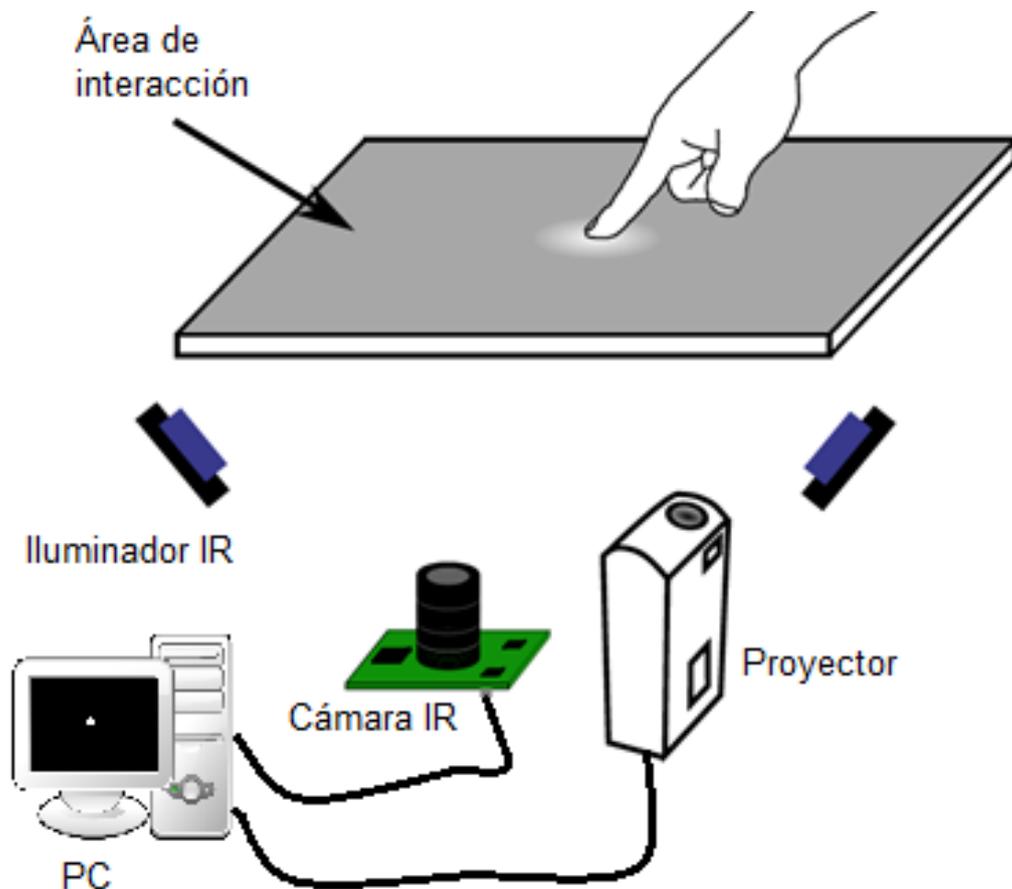


Figura 5.1. Componentes del hardware.

### 5.1.1. Elementos utilizados en este proyecto

El dispositivo utiliza un proyector de lámpara led de referencia *Favi Rio Led 3* mostrado en la figura 5.2a, que emite aproximadamente un promedio de 150 lumens sobre el área de interacción y cuenta con una resolución nativa de 800X600. Por otro lado, la pantalla de proyección utilizada en este proyecto se muestra en la figura 5.2b; y según el fabricante, está elaborada en su totalidad

con PVC, lo que le brinda a su textura propiedades elásticas que hay que tener en cuenta en el momento de realizar su montaje en el dispositivo.



(a)



(b)

Figura 5.2. Elementos del sistema de visualización, a) Proyector *Favi Rio Led 3*, b) Pantalla de proyección.

La pantalla es de tipo retroproyectivo, debido a que la fuente de las imágenes en el diseño a seguir, se encuentra en la parte posterior de esta misma con respecto al usuario [40]. Con este tipo de proyección visual la instalación del dispositivo luce limpia, ya que no se ve el proyector como una fuente distractora para el usuario; quien durante su interacción con el dispositivo, no puede generar sombras causadas por interferencias con los rayos que forman la interfaz gráfica, gracias a la implementación de un sistema basado en proyección trasera (retroproyección).

Esta clase de pantalla se distingue por estar elaborada de un material semitransparente que forma la imagen por medio de la difusión de la luz emitida por el proyector, aprovechando al máximo la luminiscencia del mismo [41].

En este tipo de material la difusión se presenta de tal forma que permite una distribución lo más homogénea posible de la luz sobre la superficie de la pantalla; disminuyendo la magnitud de los valores máximos de intensidad luminosa que se presentan en cierta zona de la misma, conocida como “hot spot” o punto caliente [42], además de aprovechar de manera eficiente los rangos presentados en el resto del área donde se forma la imagen; logrando que la interfaz gráfica no se muestre con las esquinas demasiado oscuras.

Con el objetivo de lograr los niveles de intensidad en el infrarrojo necesarios para el sistema de detección, se utilizan dos tipos de iluminadores; los primeros de referencia S8021-60-C-IR, mostrados en la figura 5.3a y que son denominados como tipo 1; reforzados por iluminadores desarrollados para este proyecto, denominados en adelante como tipo 2 y son mostrados en la figura 5.3b.



(a)



(b)

Figura 5.3. Iluminadores infrarrojos. a) Iluminador tipo 1, b) Iluminador tipo 2.

Para la construcción de los iluminadores tipo 2, cuya disposición electrónica se muestra en la figura 5.4, se hace uso de leds infrarrojos de 10 mm de diámetro, que emiten en una longitud de onda de 850nm aproximadamente.

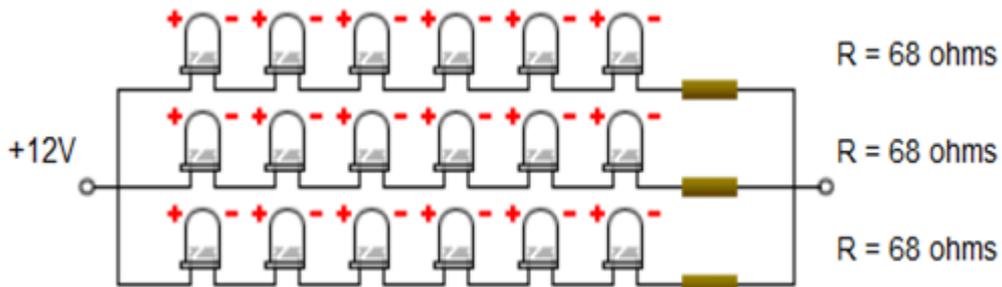


Figura 5.4. Disposición electrónica iluminador tipo 2.

La captura de las interacciones se realiza a través de una cámara de referencia *genius widecam 1050*, mostrada en la figura 5.5, adecuada para que solo detecte imágenes en el rango del infrarrojo. Esta cámara cuenta con un ángulo de apertura de 120° y es utilizada con una configuración de 640x480 de resolución a 30 cuadros por segundo.



Figura 5.5. *Genius widecam 1050*.

El ángulo de apertura de 120° permite la obtención de imágenes que cubren un área de interacción mayor, comparada con las webcam convencionales que por lo general tienen un ángulo de apertura de aproximadamente 60°.

Aunque esta cámara está diseñada para funcionar en el rango de la luz visible, su sensor alcanza a detectar ondas en el rango del infrarrojo cercano. Es por esta razón que el fabricante añade un filtro en el sistema óptico de la cámara, que consiste de una capa delgada de vidrio recubierta por un material usado como bloqueador de luz infrarroja, que además permite el paso de luz visible. Este filtro está ubicado entre el lente y el sensor; previniendo de esta forma los efectos no deseados de la luz infrarroja que dañan la imagen requerida por el consumidor durante la común utilización de la cámara.

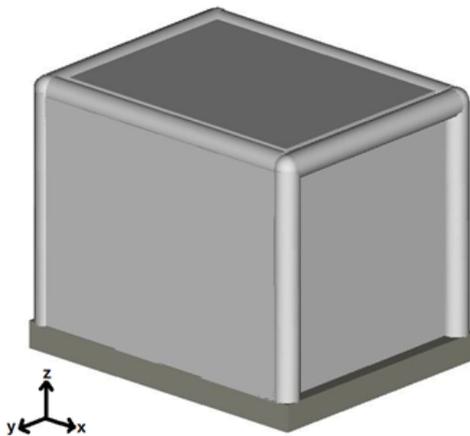
El computador base para el funcionamiento virtual del dispositivo; tiene un procesador *dual core* de 2,6 Ghz, 2 GB de memoria RAM de tipo ddr2, una tarjeta aceleradora ATI RAEDON de la serie 4300 y es utilizado con Windows 7 como sistema operativo.

### **5.1.2. Etapa 1, Sistema de visualización.**

El diseño del sistema de visualización se guía según los parámetros del proyector. Este diseño está desarrollado para obtener un área de interacción de aproximadamente 38" en la diagonal, con una altura de 70 cm, como se puede observar en la figura 5.6.

La magnitud del área de interacción, es directamente proporcional a la distancia medida desde el lente del proyector hasta el centro del área de interacción. Esta situación hace necesaria la optimización del espacio dentro del dispositivo, por medio del montaje de dos espejos para alcanzar las medidas establecidas en la figura 5.6.

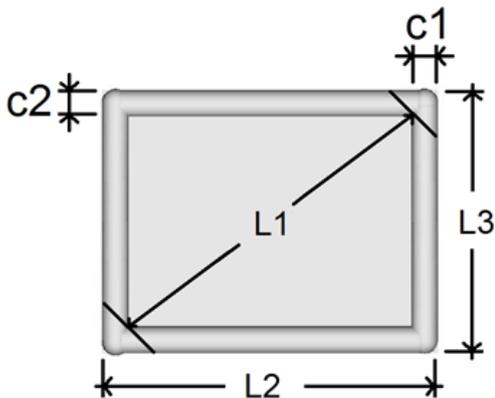
Debido a la contextura elástica de la pantalla de proyección, está se coloca en medio de dos vidrios transparentes con superficie esmerilada, como se muestra en la figura 5.7. Estos vidrios ayudan a disminuir los brillos añadidos por la iluminación del ambiente, que son percibidos por el usuario durante la proyección de la interfaz gráfica.



## DISEÑO ESTRUCTURAL

$L1 = 96.50$  (38"),  
 $L2 = 90.00$  cm,  
 $L3 = 70.00$  cm,  
 $L4 = 70.00$  cm,  
 $c1 = 7$  cm,  
 $c2 = 8$  cm,  
 $c3 = 5$  cm,  
 $c4 = 6$  cm.

Vista Superior



Vista Lateral

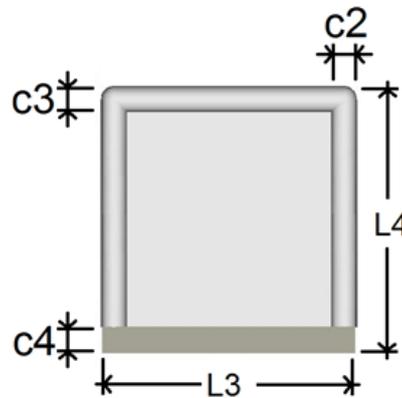


Figura 5.6. Diseño estructural del dispositivo.

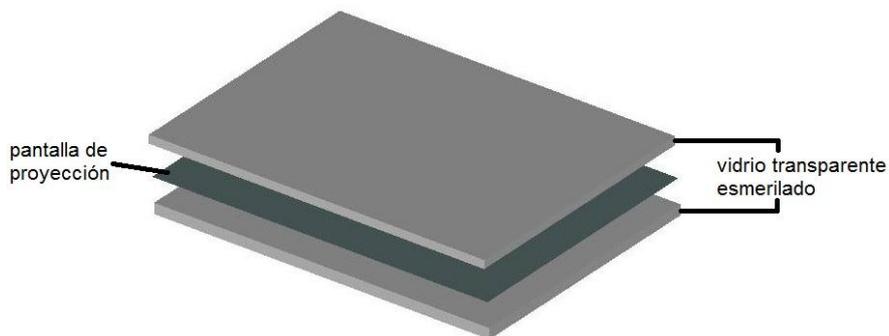


Figura 5.7. Componentes del área de interacción.

El material del elemento principal del área de interacción, una pantalla de proyección comercial; tiene cierto tipo de características que permiten brindar una experiencia satisfactoria durante la proyección de la interfaz gráfica; mientras que su color gris neutro, ayuda a obtener un buen nivel de contraste cuando se está proyectando la interfaz gráfica en habitaciones que presentan ciertas magnitudes medias de iluminación, permitiendo que el sistema de

visualización del dispositivo funcione de forma óptima en ambientes que no son completamente oscuros, como idealmente debería ser.

Esta pantalla de proyección, fue la seleccionada porque logra satisfacer necesidades de calidad visual en la interfaz gráfica; por esta razón, el área de interacción se toma como parte del sistema de visualización. Por otro lado, las propiedades de la pantalla de proyección también colaboran a que en el rango del infrarrojo cercano, se tenga una intensidad luminosa más uniformemente repartida.

Para el montaje de los espejos se buscó que las medidas arrojadas, logren que la interfaz gráfica no llegue deformada al área de interacción a la hora de construir el montaje real. Por esta razón, es necesario realizar mediciones que ayuden a tener un mejor conocimiento de las características del proyector (*Throw ratio* y *offset*), que se señalan en el capítulo 3; debido a que la posibilidad de aparición de sombras y reflejos en el área de interacción aumenta, conforme disminuye la precisión de las predicciones del diseño.

Dichas medidas observadas en la tabla 5.1, se realizan obteniendo valores para las variables señaladas en la figura 3.2, alejando gradualmente el proyector de una pared donde se generan las imágenes usando un offset fijo.

<b>D (m)</b> <b>±0.0005</b>	<b>H1 (m)</b> <b>±0.0005</b>	<b>H2 (m)</b> <b>±0.0005</b>	<b>α (grados)</b>
0,6500	0,3640	0,4320	-(4,8800 ± 0,0362)
0,7500	0,4080	0,4730	-(4,8800 ± 0,0362)
0,8500	0,4630	0,5710	-(4,8800 ± 0,0362)
0,9500	0,4890	0,6450	-(4,8800 ± 0,0362)
1,0500	0,5730	0,7020	-(4,8800 ± 0,0362)
1,1500	0,6310	0,7610	-(4,5500 ± 0,0361)
1,2500	0,6600	0,7930	-(4,5500 ± 0,0361)
1,3500	0,6910	0,8800	-(4,5500 ± 0,0361)
1,4500	0,7480	0,9570	-(4,8800 ± 0,0362)
1,5500	0,7860	0,9780	-(5,2000 ± 0,0362)
1,6500	0,8380	1,0700	-(5,8500 ± 0,0362)
1,7500	0,9020	1,1320	-(5,2000 ± 0,0362)
1,8500	0,9270	1,1680	-(5,5300 ± 0,0362)
1,9500	1,0120	1,2390	-(5,8500 ± 0,0362)
2,0500	1,0440	1,3150	-(7,1600 ± 0,0362)
2,1500	1,0970	1,3560	-(9,7700 ± 0,0363)
2,2500	1,1510	1,4320	-(13,0400 ± 0,0365)
2,3500	1,1920	1,4790	-(16,6700 ± 0,0367)
2,4500	1,2300	1,5720	-(21,9900 ± 0,0372)
2,6000	1,2980	1,6400	-(27,3900 ± 0,0378)

Tabla 5.1. Medidas del proyector con un offset fijo.

Con las medidas de la tabla 5.1, se busca establecer numéricamente los valores de tamaño del área de interacción, dependientes de su ubicación con

respecto al proyector. En ese sentido, las medidas ayudan a tener una noción más clara del *throw ratio*, y de los ángulos de los rayos que demarcan las fronteras de la interfaz gráfica en el área de interacción; por medio de aproximaciones logradas con las ecuaciones 5.1 y 5.2.

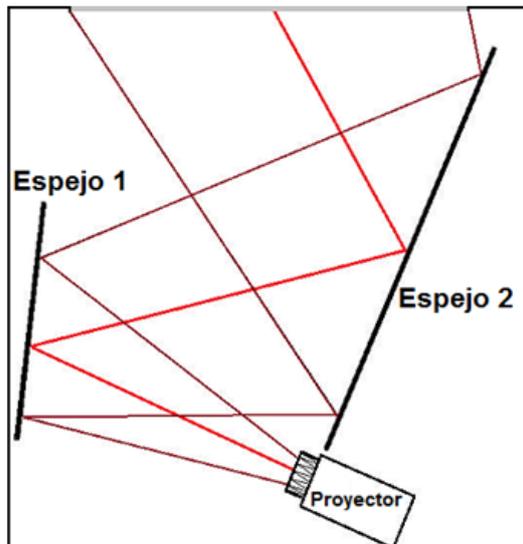
$$H1 = (0,4840 \cdot D) + 0,0490[\text{m}], \quad \text{ecuación 5.1.}$$

$$H2 = (0,6180 \cdot D) + 0,0390[\text{m}], \quad \text{ecuación 5.2.}$$

Las ecuaciones 5.1 y 5.2, además de usarse para establecer el tamaño del área de interacción dependiendo de la ubicación del proyector, ayudan a que se pueda determinar el tamaño de cada uno de los espejos; que también depende de los ángulos de los rayos de las fronteras de la interfaz gráfica.

Por otro lado, para obtener imágenes rectangulares para la interfaz del dispositivo, se utilizan como guía cuantitativa, los ángulos de incidencia en las fronteras que la demarcan; que son dependientes del ángulo en el que está ubicado el proyector dentro del sistema.

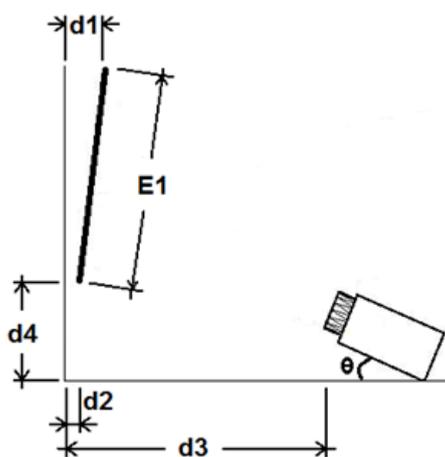
El diagrama de rayos del montaje de espejos diseñado, se observa en la figura 5.8. Las medidas de E1 y E2 arrojadas por el diagrama de rayos son de (30 cm) X (40 cm) para el denominado como espejo 1 que es representado por E1, y (70 cm) X (55 cm) para E2 correspondiente al denominado espejo 2. Estos valores son obtenidos por medio de las ecuaciones 5.1 y 5.2.



### DIAGRAMA DE RAYOS

$d1 = 5 \text{ cm}$ ,  
 $d2 = 2 \text{ cm}$ ,  
 $d3 = 40 \text{ cm}$ ,  
 $d4 = 11 \text{ cm}$ ,  
 $d5 = 7 \text{ cm}$ ,  
 $d6 = 12 \text{ cm}$ ,  
 $d7 = 35 \text{ cm}$ ,  
 $\theta = 20^\circ$ .

#### Vista Superior



#### Vista Lateral

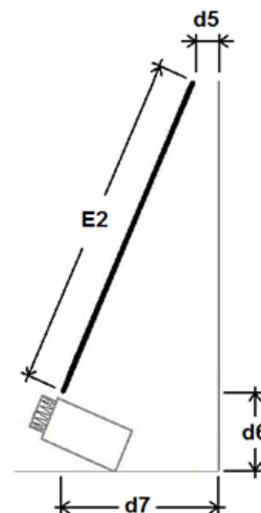


Figura 5.8. Diagrama de rayos para el montaje de espejos.

La estructura inicial de la que parte la construcción del dispositivo se puede observar en la figura 5.9a, mientras que su área de interacción es aquella que muestra la figura 5.9b. Para lograr una medida más precisa y reproducible, en cuanto al tamaño y la forma de la interfaz desplegada en el área de interacción, deben usarse monturas con tornillos de anclaje como se muestra en la figura 5.9c, que a su vez permiten la calibración de la imagen.

En la figura 5.9d, se puede ver como luce el sistema de visualización finalizado, determinando a su vez la apariencia general del dispositivo en cuanto a su estructura se refiere.



(a)



(b)



(c)



(d)

Figura 5.9. Implementación de la etapa 1. a) Estructura inicial, b) Área de interacción, c) Montaje de espejos, d) Sistema de visualización.

El revestimiento del sistema de visualización que constituye la parte exterior del dispositivo, es la que se puede mejorar si se busca lograr una apariencia más estética para el producto final, mostrado en la figura 5.10.



Figura 5.10. Revestimiento del sistema de visualización.

Con este revestimiento, se busca disminuir la llegada de luz del ambiente al área de interacción, para que en las condiciones de iluminación exterior disminuyan de alguna forma su potencial capacidad de reducir los niveles de contraste en la interfaz desplegada; al llegar solamente desde una dirección al área de interacción (desde arriba). Aunque cabe aclarar que las grandes diferencias en la calidad de brillo y contraste, se deben más al material difusor de la pantalla de proyección utilizada.

### **5.1.3. Etapa 2, sistema de detección.**

Primero debe acomodarse la cámara infrarroja, debido a que los reflejos y variaciones de brillo que esta detecte, dependen directamente de la posición en la que se ubique. La adecuación de la cámara se hace removiendo el filtro bloqueador de infrarrojos; para reemplazarlo por un negativo revelado de cámara fotográfica de 35mm, que sirve como bloqueador de luz visible y permite pasar los rayos infrarrojos, ubicados en el rango del espectro que alcanza a detectar el sensor de la cámara.

La cámara se ubica en la base del dispositivo, donde su ángulo de apertura le permite captar una escena con el área de interacción completa; además de que su montaje no introduce sombras en el sistema de visualización desde esta posición, y se logra detectar niveles de intensidad aproximadamente uniformes en todas las zonas del área de interacción, desde su punto de vista.

Por otro lado, la forma en que se logre la distribución de la intensidad de luz infrarroja proveniente de los iluminadores sobre el área de interacción; determina la calidad de detección en cuanto niveles de ruido se refiere, por lo que se debe internar que los toques del usuario sean identificables de forma equivalente en todas las zonas del área de interacción; además de que las pequeñas variaciones de luz infrarroja en el entorno del dispositivo, se noten en la menor cantidad posible en la señal de video obtenida por la cámara.

La iluminación infrarroja se acomodó de tal modo, que cada iluminador lograra cubrir determinadas zonas dentro del área de interacción, diferentes para cada uno; haciendo que los niveles de intensidad infrarroja valorados de forma cualitativa por los desarrolladores, no se diferencien entre sus valores máximos y mínimos. También se debe procurar que no se generen brillos en el lente de la cámara, que pueden entorpecer los análisis que hace el software de visión artificial explicado en la sección 5.2 de este capítulo.

Para lograr una distribución de intensidad de luz aproximadamente uniforme, a cada lado del proyector, se coloca una pareja de iluminadores compuesta por un ejemplar tipo 1 y uno tipo 2. Estos iluminadores son apuntados hacia las esquinas del espejo 1, procurando que sus puntos máximos de intensidad se

produzcan fuera del mismo. La disposición de iluminadores se puede analizar en la figura 5.11, que parte de la vista superior y frontal del diseño de la figura 5.6, sin tomar en cuenta el espejo 2 ni el área de interacción.

### DISPOSICIÓN DE ILUMINADORES INFRARROJOS

S1 = 45 cm, S2 = 65 cm.

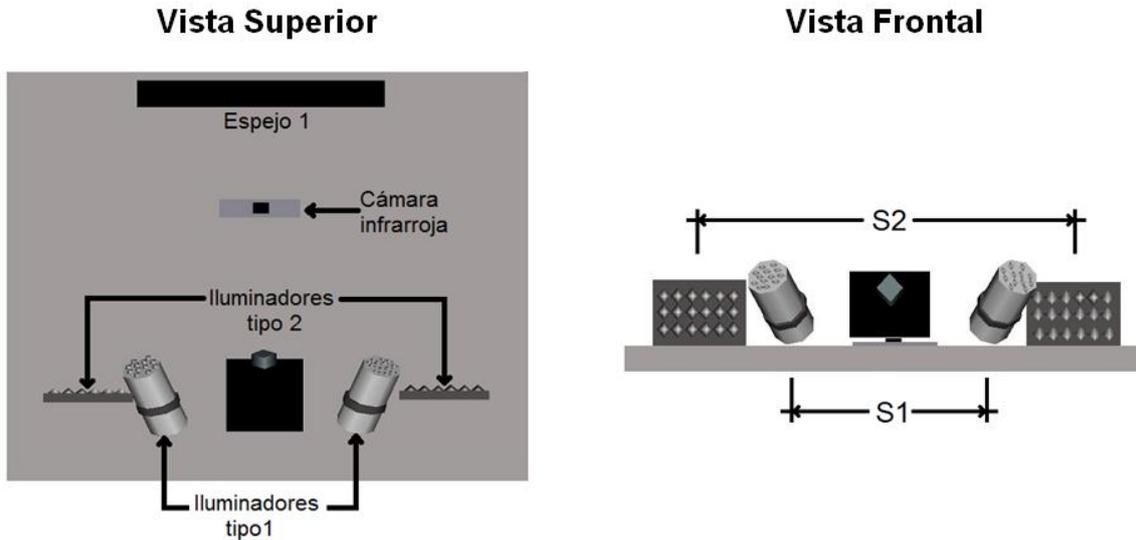


Figura 5.11. Disposición de iluminadores infrarrojos.

El tipo de imágenes que se obtienen son como la mostrada en la figura 5.12, y fue considerado por los desarrolladores de este proyecto como las adecuadas para iniciar la detección de interacciones del usuario con el dispositivo.



Figura 5.12. Área de interacción en el rango del infrarrojo.

## 5.2. DESARROLLO DEL SOFTWARE DE VISIÓN ARTIFICIAL

El software desarrollado en este proyecto tiene como objetivo obtener una detección libre de ruido para diferentes niveles de intensidad luminosa en el

entorno del dispositivo. Está implementado con bibliotecas de *OpenCV*, basándose en el análisis de contrastes para atenuar efectivamente el ruido del resplandor.

En ese sentido, se buscó resaltar las áreas de contacto en la imagen; disminuyendo el rango de la escala de grises en el que se presentan las mismas y variando los valores de los píxeles del fondo, a magnitudes donde se asume que es ruido a eliminar durante la umbralización; para luego entregar una adecuada imagen binaria, con uno de los dos valores posibles correspondiendo en su mayoría a las áreas de contacto y el otro al fondo.

La imagen a analizar debe mostrar principalmente los cambios de intensidad que se generan por los toques del usuario en el área de interacción; por esta razón, es necesaria la sustracción del *background* a las imágenes detectadas. Dicho *background* cuenta con los valores iniciales de niveles de gris correspondientes a la iluminación presentada antes de empezar la comunicación usuario-computador.

Por otro lado, antes de realizar cualquier procesamiento a las imágenes detectadas, es imperativo eliminar las zonas de la escena capturada que no corresponden al área de interacción; para evitar con esto, el desperdicio de recursos computacionales con del análisis de píxeles cuyo resultado, no aporta ninguna información útil. Esto se logra por medio de la extracción de una Región de Interés, o ROI por sus siglas en inglés.

### **5.2.1. Análisis general**

Como primer paso, son necesarias ciertas operaciones de adecuación de la imagen, cuyo resultado permite realizar un procesamiento de mayor calidad en las posteriores etapas de filtrado. Para diferenciar efectivamente las áreas de contacto del fondo, deben realizarse operaciones que modifiquen los valores de los píxeles en la escala de grises, atenuando el ruido y resaltando los toques; siendo esta etapa la de filtrado en los niveles de gris.

Específicamente, con el filtro de niveles de gris se busca que la diferencia entre el valor máximo y el mínimo que presentan los toques del usuario en las imágenes detectadas, sea la menor posible; es decir, se busca disminuir el rango dinámico en la escala de grises de los píxeles pertenecientes a las áreas de contacto.

Este filtro se utiliza para poder obtener una imagen donde se logre ubicar las áreas de contacto, en un rango mayor a cierto número en la escala de grises y cuya magnitud idealmente no debería ser superada por ninguno de los niveles de ruido; estableciéndose entonces dos grupos para todos los píxeles que conforman la imagen, uno para áreas de contacto y el otro para el fondo.

Al asignar los valores de 0 o 1 para cada uno de los grupos de píxeles mencionados se tiene como resultado una imagen binaria, siendo esta la umbralización. Si se ha procurado disminuir en la mayor cantidad posible el ruido durante la etapa de filtrado en los niveles de gris, se pueden obtener imágenes binarias más precisas al realizar la umbralización.

En este punto como se tiene solo dos valores posibles para cada píxel, la manipulación de la imagen solo puede significar el agregar o eliminar píxeles al área de contacto o al fondo; denominándose este tipo de filtro como morfológico o de forma.

El filtro morfológico es usado para eliminar píxeles solitarios, o pequeñas áreas que por su tamaño son consideradas como ruido que ha superado umbralización; también ayuda a corregir las áreas de contacto, que en algunos casos están deformadas por los resplandores cercanos a los toques provenientes desde afuera del dispositivo, como los observados en la figura 3.3, o por las pequeñas variaciones de iluminación en el entorno del mismo.

### 5.2.2. Pre-procesamiento

El ROI se extrae partiendo desde la imagen del *background* de la escena completa, que para el ejemplo práctico escogido en la explicación del procedimiento se muestra en la figura 5.13a. Cabe anotar que como herramienta de guía para este paso, se ponen sobre la pantalla cintas negras que demarcan las fronteras del área de interacción.

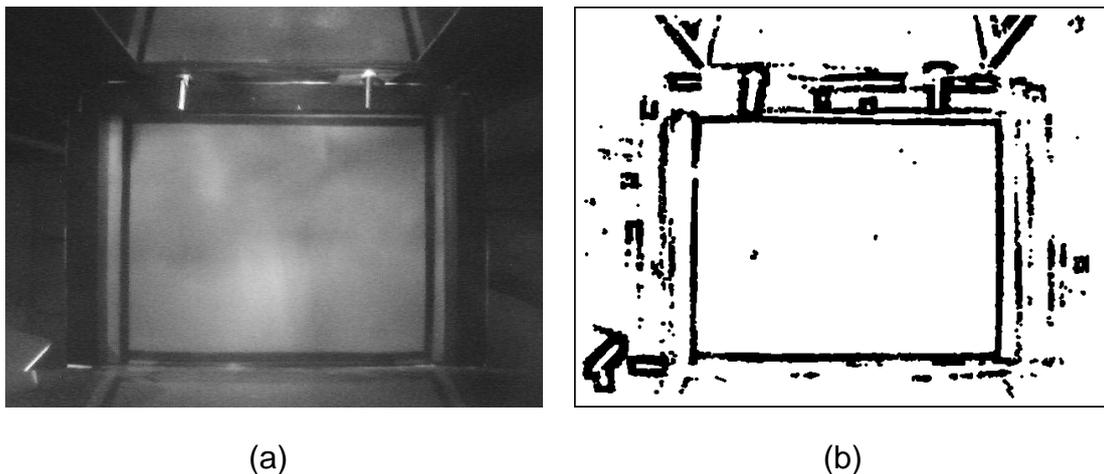


Figura 5.13. Escena a recortar, a) *Background* de la escena, b) Imagen binaria del *background*.

La imagen correspondiente a la figura 5.13a es una representación dada en un solo plano de niveles de gris; que parte de una imagen capturada en el espacio de color RGB y consta de la combinación de aproximadamente un 30% del

plano rojo, un 59% del verde y un 11% del azul. Esta imagen es convertida a una de tipo binario, como la que se muestra en la figura 5.13b; procurando que los píxeles del área de interacción correspondan mayoritariamente al color blanco, mientras que las cintas negras que la demarcan puedan distinguirse por el color negro.

Se deben seleccionar cuatro coordenadas correspondientes a aquellas donde se considera que se presentan las esquinas del área de interacción, para extraer esta misma como una imagen cuadrada más pequeña que la escena original; y cuyos píxeles describen los cambios de intensidad que son motivo de análisis durante el procesamiento.

Para determinar las coordenadas mencionadas, se utilizó el método de integrales proyectivas; usando dos vectores que tienen un número de posiciones correspondiente a las filas o columnas de la imagen (640 o 480). Las representaciones gráficas de los vectores de filas y columnas de la imagen escogida como ejemplo procedimental, se pueden analizar en la figura 5.14.

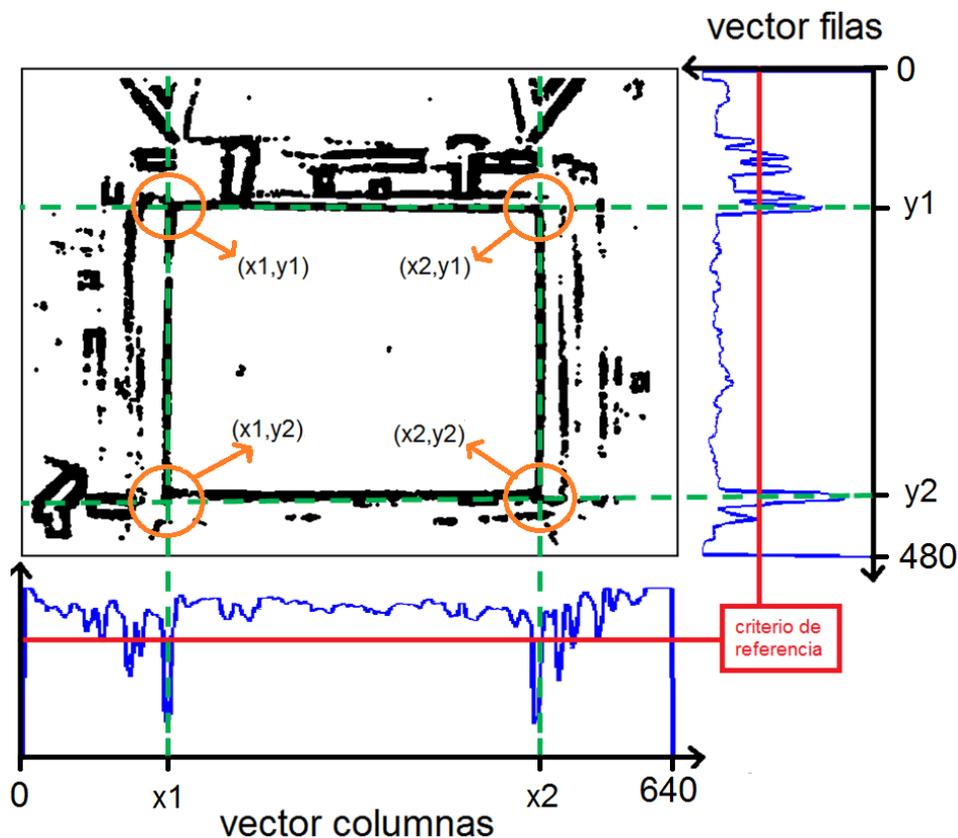


Figura 5.14. Método integrales proyectivas.

En cada uno de los dos vectores, se almacena la cantidad de veces en las que se repite el número 1 en las filas o columnas de la imagen binaria, dependiendo de la dirección en la que se está analizando la imagen; es decir, si se está valorando la existencia del número 1 para las filas o para las columnas. Así pues, en las posiciones del vector correspondiente a las filas, se colocan los

valores resultantes de la suma de píxeles en cada una de estas; y se sigue el mismo razonamiento para llenar el vector de columnas.

Luego se establece un valor equivalente a un porcentaje del máximo almacenado en cada uno de los dos vectores. Este valor es usado como criterio de referencia, para indicar los intervalos de las posiciones que lo superan, que luego son comparados entre sí escogiendo el más largo para cada vector.

Como se puede analizar en la figura 5.14, los dos intervalos seleccionados están acotados en los vectores, por los valores correspondientes de las filas y columnas que ayudan a establecer las coordenadas de las esquinas del área de interacción. Según estos valores, se realizan los cortes pertinentes para obtener una imagen, que comprende solamente la parte de la escena que debe ser procesada.

Los valores de las coordenadas usadas para realizar la extracción del área de interacción en la imagen del *background*, son utilizados para seguirla extrayendo en las posteriores imágenes pertenecientes a la señal de video proveniente del sistema de detección. De la imagen del *background* seguido como ejemplo se obtiene el resultado mostrado en la figura 5.15.

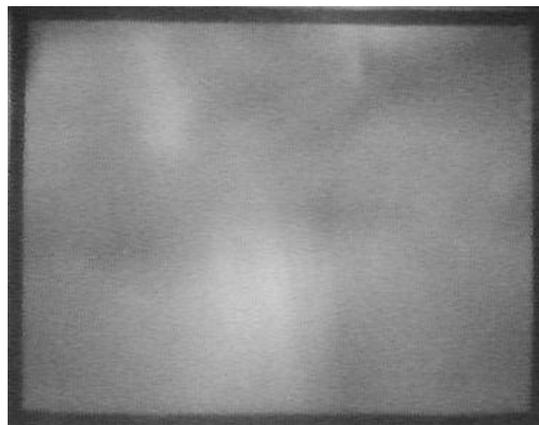


Figura 5.15. Región de interés (ROI).

Una vez establecida la región de interés, para cada imagen de la señal de video se procede a la sustracción del *background* en el espacio de color RGB, obteniendo una imagen resultante que se denota con la letra *I*.

Como ejemplo en esta ocasión se tiene la imagen mostrada en la figura 5.16a, que para este paso tiene como resultado la imagen de la figura 5.16b. Los resultados mostrados desde la figura 5.16 en adelante se centran en las zonas donde se presentan los toques; debido a que en todos los casos, las variaciones de niveles de gris en el resto del área de interacción no aportan información relevante para el análisis de ningún paso.



Figura 5.16. Imagen pre-procesada, a) Imagen detectada, b) Resta de *background*.

En la figura 5.16b, se puede observar que aunque los píxeles correspondientes a las áreas de contacto en la imagen  $I$ , tienen un buen nivel de contraste con los vecinos de su entorno, sus valores de intensidad en la escala de grises varían dependiendo de la zona del área de interacción que se esté analizando; trayendo como consecuencia que el rango dinámico sea muy grande.

Esto representa un inconveniente al momento de extraer un umbral que elimine el ruido de manera eficiente, debido a que los píxeles de las áreas de contacto en ciertas zonas son equiparables a los del ruido en otras; trayendo como consecuencia que el rango dinámico de los píxeles de las áreas de contacto, sea demasiado grande e incluya valores pertenecientes a los niveles de ruido detectados.

### 5.2.3. Etapa1, adecuación de la imagen

La imagen  $I$  fue convertida a los espacios de color Y'CbCr, HSV y XYZ; consiguiendo un grupo de doce planos en total. Luego de realizar un análisis cualitativo de los mismos, se determinó que los mejores planos descriptores que se pueden utilizar para la detección de toques, se obtienen según la combinación establecida en la ecuación 5.3.

$$I_1 = 0.4Y' + 0.25X + 0.33Y, \quad \text{ecuación 5.3.}$$

Donde  $Y'$  es el primer plano del espacio de color Y'CbCr,  $X$  y  $Y$  son el primer y segundo plano respectivamente del espacio de color XYZ. En la figura 5.17b se puede apreciar el tipo de imagen que se obtiene.

Posteriormente, con el objetivo de lograr unas condiciones óptimas en los niveles de gris de las áreas de contacto, se procede a aplicar un filtro espacial gaussiano sobre  $I_1$ ; logrando también la atenuación de una cantidad considerable de ruido de bajo nivel, si se analiza el ejemplo mostrado en la figura 5.17c.

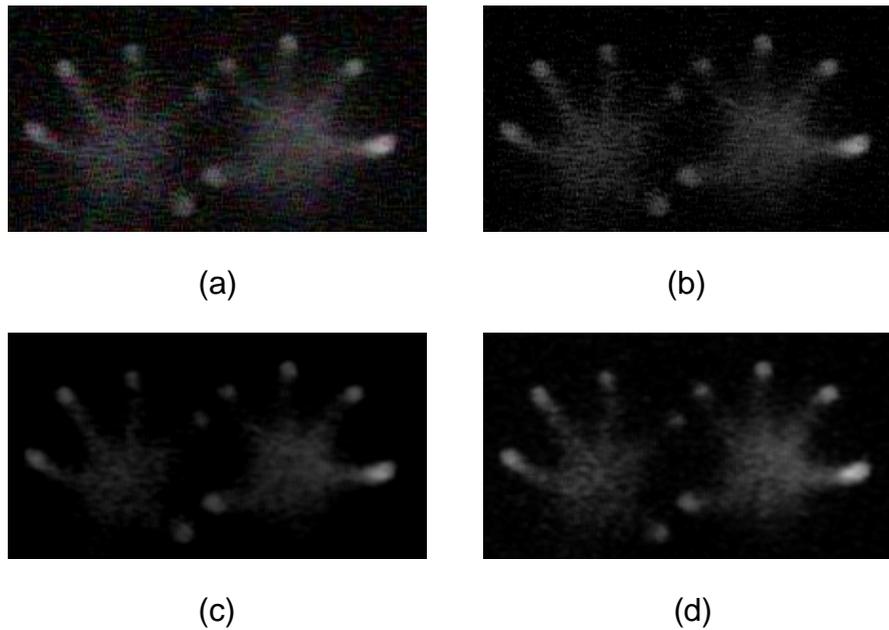


Figura 5.17. Pasos de la etapa 1. a) Imagen de entrada, b) Resultado de aplicar ecuación 5.3, c) Filtro de Gauss, d) Modificación de contraste.

Después, se lleva a cabo una modificación de contraste por medio de la ecuación 4.1, dando resultados como los de la figura 5.17d, donde se garantiza que el máximo de los píxeles cuando ocurren las interacciones, siempre esté ubicado en un valor de 255 en la escala de grises. De esta forma los siguientes pasos del procesamiento de la imagen, son menos susceptibles a pequeñas variaciones de iluminación en el entorno del dispositivo, logrando detectar los toques en condiciones más generalizadas.

#### 5.2.4. Etapa 2, filtro en los niveles de gris

Una vez logrado el tipo de imágenes de la figura 5.17d, se procede a la aplicación de un filtro de media con un kernel de 25x25, dando como resultado una imagen  $P_1$  como la observada en la figura 5.18a.

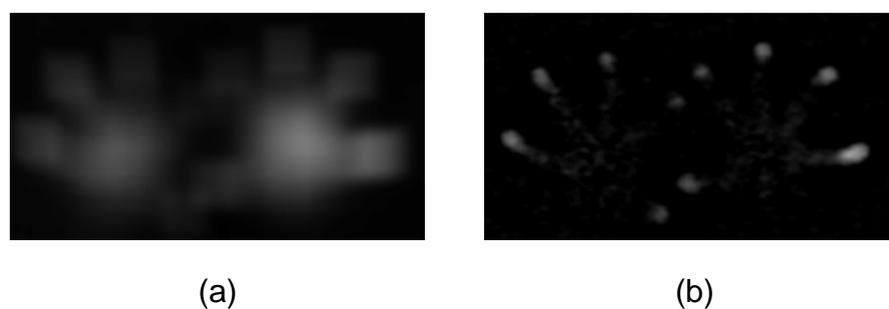


Figura 5.18. Máscara de enfoque, media. a) Filtro de media, b) Aplicación de la ecuación 4.4.

Al aplicar la ecuación 4.4 siendo  $I_1$  la imagen de entrada,  $P_1$  la imagen plantilla y  $\beta$  igual a -1; se obtiene una imagen de salida  $I_2$  que puede ser analizada en la figura 5.18b, donde los pixeles del ruido de nivel medio se han atenuado; y se despliegan las áreas de contacto con un aumento en sus valores de intensidad, es decir, que su rango dinámico se diferencia de los pixeles pertenecientes al ruido en el fondo.

En este punto, si se realiza una umbralización al tipo de imágenes como las mostradas en la figura 5.18b, se pueden obtener resultados aceptables en las imágenes binarias. Pero al detectarse ruido de nivel alto, se presentan interacciones fantasma debido a resplandores. Por otro lado, cuando se presentan leves aumentos de iluminación exterior, los valores del ruido de nivel medio aumentan de tal forma, que entran en el rango dinámico de los pixeles de las áreas de contacto en este paso; siendo necesario el seguir modificando la imagen en la escala de grises, hasta lograr atenuar efectivamente todos los niveles de ruido.

Debido a esto, se utiliza por segunda ocasión la ecuación 4.4, siendo nuevamente  $I_1$  la imagen de entrada; pero con el valor de  $\beta$  esta vez igual a -2, mientras que para la imagen plantilla se modifica  $I_2$  por medio de un filtro espacial gaussiano. La imagen que se obtiene en este paso será denominada como  $I_3$ .

Como ejemplo para la atenuación del ruido de nivel alto, se tiene la imagen mostrada en la figura 5.19a, cuyo resultado después de la etapa de adecuación corresponde a la figura 5.19b; para el cual la imagen  $I_2$  es como la observada en la figura 5.19c. El resultado  $I_3$  en este ejemplo es el de la figura 5.19d.

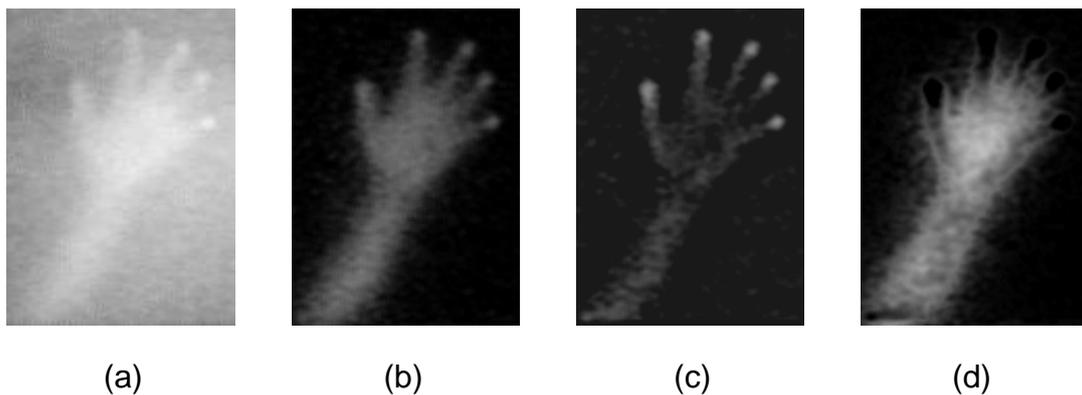


Figura 5.19. Ejemplo con ruido de nivel alto. a) Imagen detectada, b) Resultado de la etapa 1, c) Máscara de enfoque, media, d) Máscara de enfoque, Gauss.

Posteriormente, a  $I_2$  se le resta  $I_3$  para luego modificar el contraste de esta sustracción por medio de la ecuación 4.1, dando como resultado imágenes en donde se considera que todos los niveles de ruido han sido atenuados o eliminados; cumpliendo con el objetivo del filtro de niveles de gris y obteniendo el tipo de imagen requerido, para que la umbralización funcione efectivamente.

Para los ejemplos de ruido de nivel medio y alto analizados anteriormente, se obtuvieron imágenes de salida como las mostradas en las figuras 5.20a y 5.20b respectivamente.

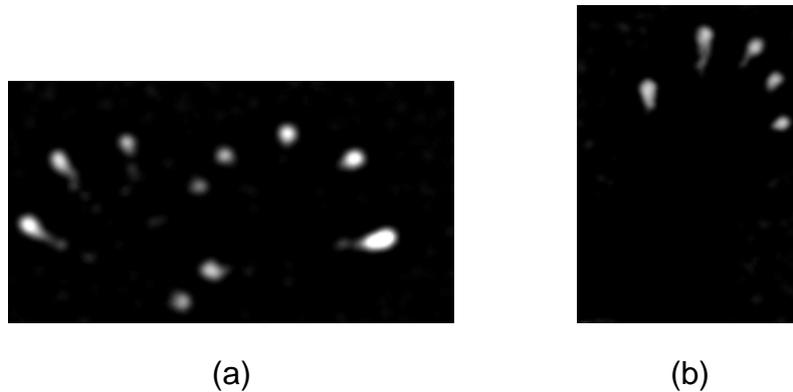


Figura 5.20. Atenuación de niveles de ruido, a) Atenuación de ruido de nivel medio, b) Atenuación de ruido de nivel alto.

### 5.2.5. Etapa 3, filtro morfológico

Después de la umbralización, se trabaja con imágenes binarias como las que se muestran en la figura 5.21.

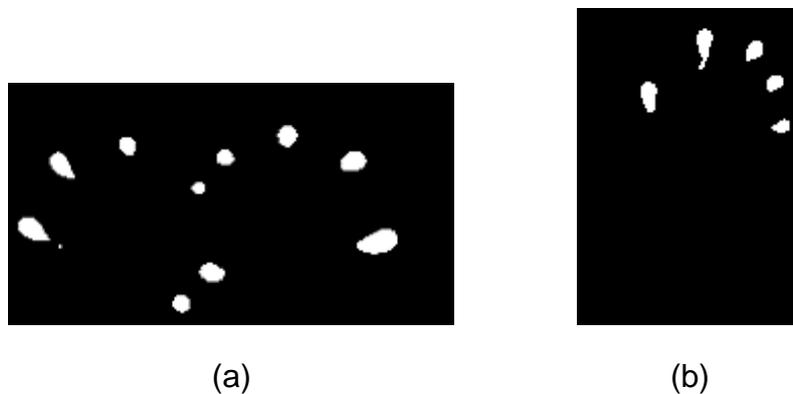


Figura 5.21. Umbralización. a) Ejemplo ruido de nivel medio, b) Ejemplo ruido de nivel alto.

En las imágenes binarias de la figura 5.21, se pueden observar dos efectos que causan error debido a los resplandores de las partes de los dedos que están muy cerca al área de interacción; que alcanzan valores de intensidad lo suficientemente altos para superar la umbralización.

El primero de estos errores se manifiesta como un pequeño grupo de píxeles representando un área de contacto inexistente y por ende, interacciones fantasma; error que se presenta cuando los brazos del usuario o cualquier otro objeto, se acercan mucho al área de interacción. Este tipo de error se observa

detalladamente en la figura 5.22a y corresponde a la imagen binaria mostrada en la figura 5.21a.

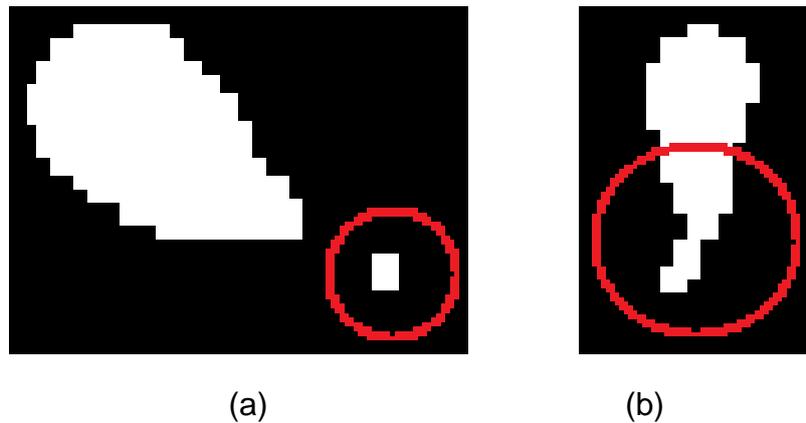


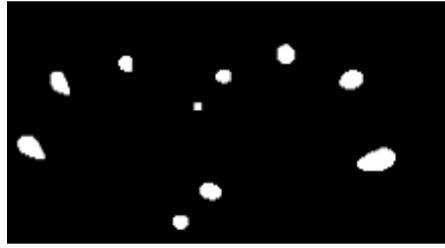
Figura 5.22. Errores en las imágenes binarias. a) Interacciones fantasma, b) Efecto cola.

El segundo error observado, proviene de las imágenes como las de la figura 5.21b, donde se presenta otro pequeño grupo de píxeles de error, esta vez como parte de las áreas de contacto formando una especie de “cola”, como la señalada en la figura 5.22b. Este error disminuye la precisión del número de píxeles y la forma en que se agrupan para cada área de contacto, que son los parámetros principales en los que se basa la extracción de la información, que se envía en los paquetes UDP del protocolo TUIO y corresponden a cada interacción identificada.

Esto puede dar como resultado, que los toques del usuario, sean respondidos con interacciones de forma errática dentro de un pequeño radio alrededor del área de contacto; debido a que los resplandores detectados como parte de esta, cambian constantemente en la señal de video.

Por estas razones, es necesario erosionar la imagen binaria analizándola según los primeros cuatro vecinos; lo que da como resultado la eliminación de buena parte de las colas de las áreas de contacto mencionadas como error, además de eliminar muchos de los píxeles causantes de interacciones fantasma.

Luego se busca aumentar la posibilidad de que las formas de las áreas de contacto, sean más parecidas entre cuadro y cuadro del video para la misma interacción, por medio de una dilatación según los primeros ocho vecinos. Las imágenes que se obtienen para los ejemplos escogidos se muestran en la figura 5.23.



(a)



(b)

Figura 5.23. Eliminación de errores debidos a resplandores. a) Eliminación de interacciones fantasma, b) Eliminación del efecto cola.

Observando este tipo de resultados, se considera que el filtro morfológico hace una buena labor para corregir la imagen binaria, dando por cumplido el objetivo del software de visión artificial de permitir una detección de interacciones tan limpia de ruido como sea posible.

## 6. RESULTADOS Y DISCUSIÓN

Al término de la construcción del hardware y el desarrollo del software de visión artificial para la detección de toques, se debe evaluar el funcionamiento de cada sistema que hace parte del dispositivo final; sistemas que se pueden clasificar como funcionales de modo visual (sistema de visualización) y de modo instrumental (sistema de detección y software de visión artificial).

Posteriormente, se debe lograr una sinergia entre la parte visual y la instrumental, para que de esta forma se cumpla con el objetivo de establecer una comunicación entre el computador y el usuario por medio de una interfaz natural.

### 6.1. EVALUACIÓN DE CADA SISTEMA

El funcionamiento propio de cada sistema desarrollado, visualización, detección y filtrado digital de la misma; se analiza siguiendo unos lineamientos propuestos en este proyecto, que ayudan a determinar cualitativamente la respuesta de todo el dispositivo, ante cambios bruscos de iluminación exterior.

Al aumentar la luz proveniente del exterior, se introducen variaciones notables visualmente para el usuario en el sistema de visualización; como también, su presencia al exterior del sistema, aporta pequeños niveles de intensidad infrarroja que se reparte por toda el área de interacción, variando los niveles de gris con respecto al *background* utilizado durante el procesamiento digital de las imágenes detectadas.

#### 6.1.1. Evaluación de la visualización

Su evaluación se hace para estimar el potencial impacto visual que puede lograrse al ser desplegada la interfaz gráfica. La fuente de luz exterior utilizada para los análisis, es una lámpara fluorescente compacta de 25W, ubicada según como se observa en la figura 6.1.

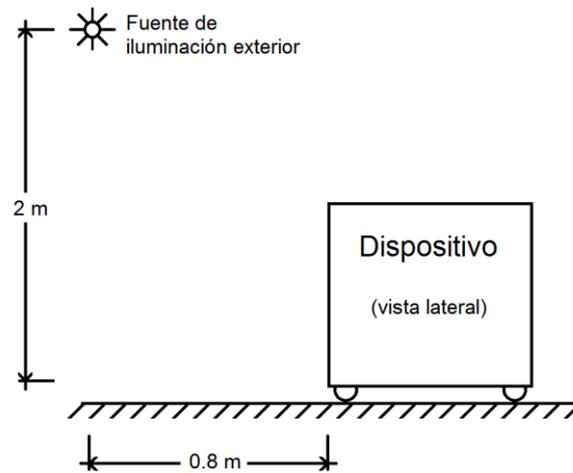


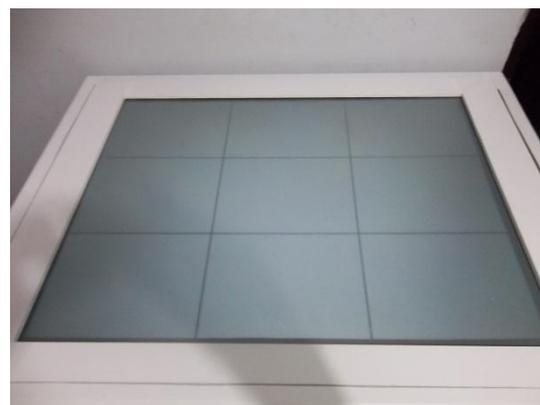
Figura 6.1. Ubicación de fuente de iluminación exterior.

El tamaño que se obtiene en la diagonal del área de interacción es de aproximadamente 34", siendo un 89.5% de la diagonal calculada en el diseño. Esto demuestra que el montaje de espejos está optimizando debidamente el espacio dentro de una mesa cuadrada de 70 cm de altura. Cada pixel de la interfaz gráfica está ocupando un estimado de  $0.35 \text{ mm}^2$  (0.70 mm de ancho y 0.50 mm de alto) en el área de interacción.

En cuanto a la valoración del área de interacción, se inicia analizando las características de la proyección de la interfaz, para estimar cualitativamente las condiciones de brillo y contraste de esta última. Las características difusoras de la pantalla se pueden entender de una forma clara, al analizar por zonas la calidad de la difusión de los rayos que forman la interfaz. Esto se logra proyectando una imagen como la de la figura 6.2, en los casos de oscuridad en el entorno y de aumento de iluminación exterior.



(a)



(b)

Figura 6.2. Análisis de luminiscencia por zonas, a) Oscuridad en el entorno, b) Aumento de iluminación exterior.

En la figura 6.2a, lo primero que se observa es que la interfaz desplegada no resulta incómoda, aunque el proyector está virtualmente apuntando hacia los ojos del usuario, siendo el área de interacción el único obstáculo para los rayos emitidos por este. Esto se debe a que la luminiscencia del proyector, está siendo aprovechada sin que se presenten incómodas concentraciones de brillo durante el despliegue de la interfaz gráfica, gracias a un punto caliente poco pronunciado, lo que significa una ventaja para el sistema de visualización.

Otra característica a detallar al analizar el área de interacción según las zonas demarcadas en la figura 6.2b, son los niveles de brillo que se observan en las esquinas. Estas últimas, son las zonas menos favorecidas en cuanto a niveles de luminiscencia del proyector, siendo las que más rápido tienden a desaparecer en este tipo de montajes de proyección visual ante aumentos de luz en el exterior. En el sistema desarrollado las zonas de las esquinas de la interfaz se opacan pero no desaparecen en el caso de iluminación exterior.

Como se proyecta una interfaz con una resolución equivalente a la resolución nativa del proyector (800X600), los elementos pictóricos y el texto desplegado se presentan en ciertos tamaños que requieren precisión en cuanto a la luminiscencia se refiere; para que sigan siendo entendibles en las imágenes proyectadas. Por esta razón, es necesario calibrar los valores controlables de brillo y contraste que influyen en la calidad de la imagen, para comparar los resultados en los dos casos de iluminación y oscuridad en el entorno del dispositivo.

Se utiliza un método de calibración visual sin utilizar instrumentos de medición. Las imágenes mostradas en las figuras 6.3a y 6.3b, son las plantillas utilizadas como pieza fundamental durante la calibración.

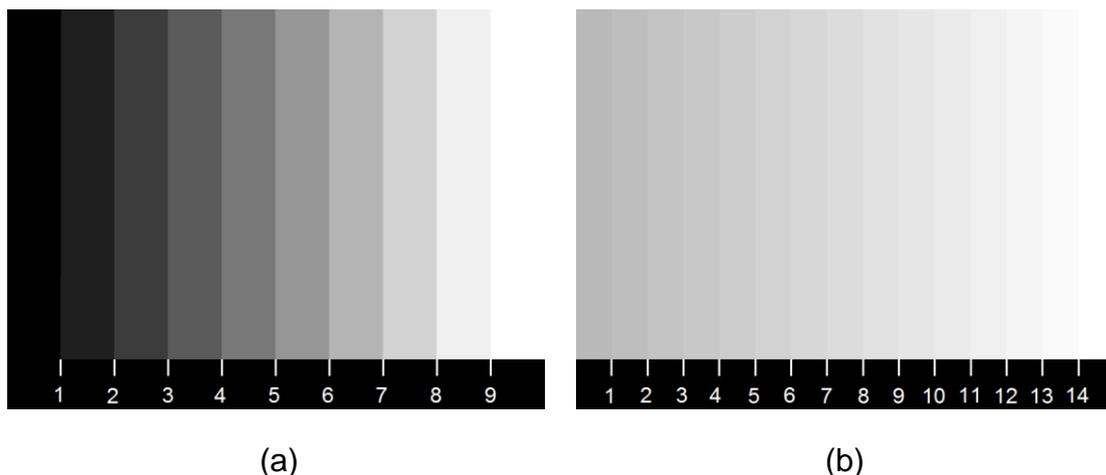


Figura 6.3. Plantillas de calibración. a) Plantilla de brillo, b) Plantilla de contraste.

La calibración consiste en comparar las barras con las que cuenta cada imagen plantilla al ser proyectada, viendo que tan definidos se presentan los rangos de

nivel de gris entre sus vecinos. Cuando en el área de interacción, se observan de formas claras y diferenciables entre sí las barras de la imagen utilizada como plantilla; significa que se tienen los valores óptimos de brillo o contraste.

En la prueba realizada, las relaciones de luminiscencia del proyector, se calibran siempre para el caso de la imagen desplegada en condiciones de oscuridad en el entorno; para luego someter al área de interacción a un aumento brusco de iluminación, por medio de la fuente mostrada en la figura 6.1. En este punto, se analizan los valores de brillo y contraste de las imágenes desplegadas en el área de interacción, sin variar la calibración inicial.

Para saber la calidad de la calibración, se suma la cantidad de barras identificables visualmente en cada plantilla proyectada; dando un valor ubicado en una escala de 25 números enteros, cuya magnitud máxima, 25; se establece como la correspondiente a una calibración ideal.

El máximo valor obtenido en la escala mencionada, se presenta cuando el número de barras identificables durante la proyección de las plantillas de brillo y contraste; es de 16 cuando el entorno del dispositivo está oscuro, y de 16 también para las condiciones de aumento de iluminación exterior.

Además, al observar entre que rangos de niveles de gris se visualiza cada barra identificada, se pueden establecer las pautas de funcionamiento a seguir, al momento que se quiera desarrollar interfaces gráficas, que aprovechen efectivamente las propiedades del sistema de visualización.

Los resultados obtenidos para los niveles de brillo y contraste después de la calibración, ayudan a inferir que el área de interacción puede soportar aumentos de intensidad luminosa en el exterior, sin que se deteriore la comunicación computador-usuario. Esto le brinda un amplio rango de funcionamiento a la parte visual del dispositivo, frente a las variaciones de iluminación en el entorno.

Seguido a esto, si se analizan imágenes desplegadas como las de la figura 6.4, se puede observar que la calibración de los valores de brillo y contraste, hace que se aproveche de una mejor forma las ventajas brinda la pantalla en el área de interacción; en cuanto a la distribución de luz se refiere.



(a)

(b)

Figura 6.4. Resultados de la calibración de luminiscencia. a) Antes de realizar calibración, b) Brillo y contraste calibrados.

### 6.1.2. Evaluación de la detección

Posteriormente, se valora el funcionamiento de la parte instrumental del dispositivo usando la misma fuente de luz exterior con la que se valoró el sistema de visualización. Al hablar de detección en esta sección, se hace alusión a los dos sistemas que operan para que los toques en el área de interacción, se presenten de forma identificable en imágenes binarias; es decir, se debe analizar el sistema de detección trabajando en conjunto con el software de visión artificial.

El diseño del montaje de espejos utilizado en el sistema de visualización, ayuda a que se tenga la posibilidad de poner los iluminadores en varias ubicaciones donde se puede hacer llegar una distribución de intensidad luminosa en el rango del infrarrojo; más o menos equivalente en todas las zonas del área de interacción.

Por otro lado, las altas concentraciones de luz proveniente de los iluminadores; son redirigidas en diferentes direcciones, debido a las rugosidades en la superficie del vidrio esmerilado, evitando así, la aparición de reflejos que dificultan la detección de los toques. Este razonamiento puede entenderse de una mejor forma al observar los resultados mostrados en la figura 6.5.



(a)

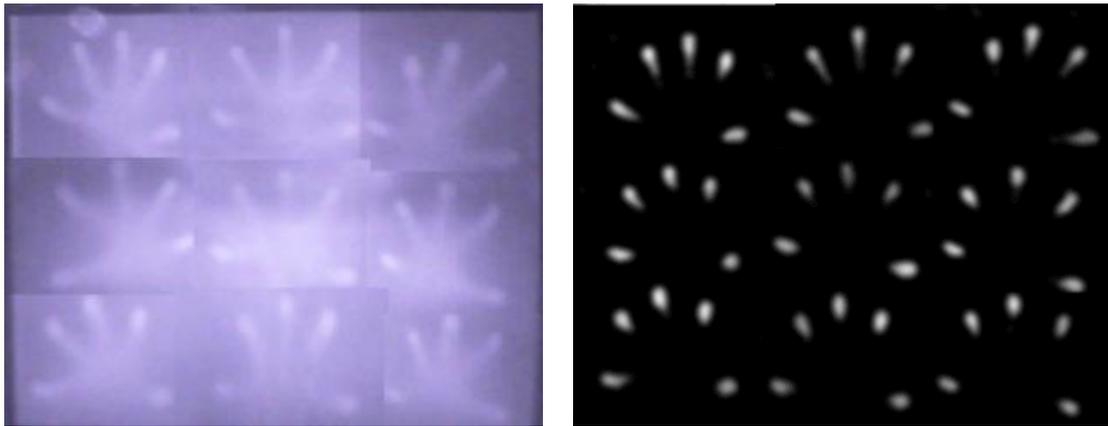
(b)

Figura 6.5. Disminución de reflejos de iluminadores infrarrojos, a) Superficie lisa, b) Superficie esmerilada.

También hay que tener en cuenta que el material difusor de la pantalla de proyección; colabora para que se presenten condiciones de iluminación infrarroja aceptables en cuanto a uniformidad se refiere. Además, la pantalla ayuda a que los rayos infrarrojos causantes del ruido de resplandores tengan patrones de reflejo que pueden ser diferenciables de los toques. De esta forma es posible realizar los análisis de contraste local con el software de visión artificial; es decir, que entre mejores propiedades difusoras presente la pantalla de proyección, se puede obtener una mayor precisión en el momento de detectar digitalmente las interacciones.

El tiempo de filtrado partiendo de la imagen de entrada hasta la obtención de imágenes binarias es de 15 milisegundos. En cuanto a la evaluación del filtro de niveles de gris del que depende mayormente la precisión del software, se inicia por estimar un promedio del rango dinámico de los toques en diferentes zonas del área de interacción; para compararlos con los obtenidos después de aplicar el filtro. Las estimaciones mencionadas, se pueden examinar para diferentes zonas en la figura 6.6, en los casos de la imagen detectada y después de pasar por el filtro de niveles de gris.

La figura 6.6a, corresponde a una recopilación de imágenes detectadas en el espacio de color RGB, presentando rangos dinámicos para los toques de 101 (123 a 224) en el plano rojo, 98 (115 a 213) en el plano verde y 98 (157 a 255) en el plano azul. Al analizar las magnitudes de los rangos dinámicos de los toques por cada zona; se puede concluir que los valores de nivel de gris de los pixeles identificables como toques, se presentan en rangos que tienden a incluir valores de ruido en todas las zonas del área de interacción.



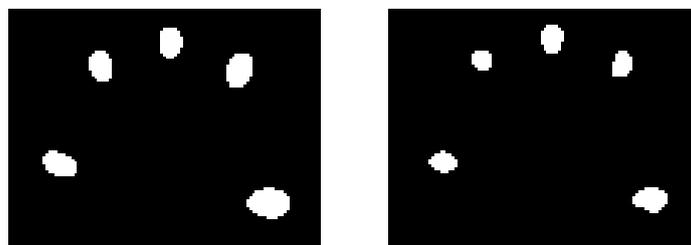
(a)

(b)

Figura 6.6. Evaluación de toques por zonas. a) Toques en la imagen capturada, b) Toques después del filtro de niveles de gris.

El rango dinámico de los toques en la imagen filtrada, se presenta entre 162 y 255 niveles de gris teniendo una magnitud de 93. Este valor se puede considerar como parecido a los rangos de la imagen detectada, pero al analizar la figura 6.7b se observa que el rango dinámico de los toques no es invadido por los niveles de gris del ruido. Por esta razón, el rango dinámico de los toques en la escala de grises después de la etapa de filtrado de niveles de gris; se puede diferenciar en todas las zonas con respecto a los niveles de ruido, debido a que durante el procesamiento digital de la imagen se elimina en su mayoría.

Como consecuencia, se tiene que las modificaciones del umbral determinado para obtener imágenes binarias, varían solamente la magnitud de las áreas de contacto de cada toque en el resultado; disminuyendo la probabilidad de que se presenten interacciones fantasma. Posteriormente, para evaluar la eficiencia del filtro morfológico, se debe comparar la forma de los toques en las imágenes binarias resultantes; en condiciones de oscuridad e iluminación exterior. Con este fin, se analiza en cada una de las zonas la morfología que presentan los toques, examinando imágenes como las del ejemplo de la figura 6.7.



(a)

(b)

Figura 6.7. Evaluación del filtro morfológico. a) Toques en condiciones de oscuridad en el entorno, b) Toques en condiciones de iluminación exterior.

Al analizar los toques en las figuras 6.7a y 6.7b, se puede observar que variando los patrones de resplandor causantes del ruido a eliminar; los toques se presentan de diferente tamaño pero con formas que se pueden considerar equivalentes. Esto tiene como consecuencia, que el centro geométrico de los toques después de aumentar la iluminación exterior, tengan coordenadas parecidas con los del caso de oscuridad en el entorno del dispositivo. Además, se resalta el hecho de que las interacciones se pueden seguir adquiriendo correctamente en un 85% de las ocasiones examinadas en toda el área de interacción.

El filtro morfológico también evita que se presenten interacciones con ligeros movimientos erráticos entre los cuadros de la señal de video; actuando ante la aparición de resplandores repentinos, que pueden causar un efecto de cola en las interacciones detectadas, y por ende, pequeñas variaciones en la posición de sus centros geométricos en cada imagen resultante.

Por otro lado, la respuesta del dispositivo para evitar que se presenten interacciones fantasma debido a los resplandores de los dedos del usuario cercanos al área de interacción; se examina según los eventos indicados en la figura 6.8, para cada una de las zonas.

En el evento mostrado en la figura 6.8a, se da inicio al movimiento del dedo índice hacia una posición de no interacción, que se puede observar en la figura 6.8b. Esta última figura, corresponde al primer cuadro en el que se deja de detectar el dedo índice. Se mide la distancia entre el dedo y la cara superior del área de interacción por medio de una escuadra, proceso que debe realizarse varias veces para cada zona.

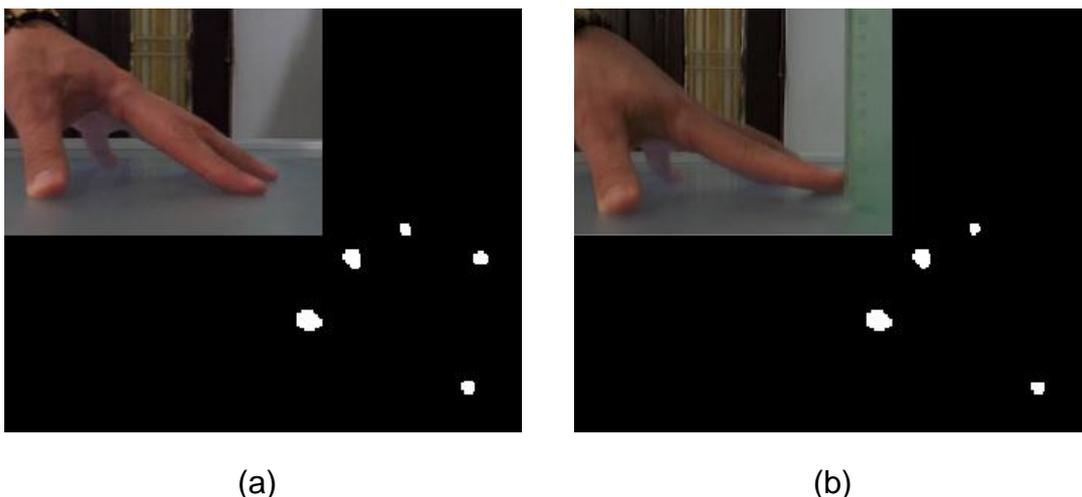


Figura 6.8. Evaluación de resplandores cercanos. a) Evento 1: interacción inicial, b) Evento 2: estado de no interacción.

En promedio, se tiene que el dispositivo deja de percibir los resplandores como interacciones fantasma, más allá de una distancia de 1.3 cm para los dos casos

de iluminación exterior analizados. Esta distancia se considera pequeña, si se compara con las dimensiones del área de interacción. De esta forma se demuestra que los elementos de esta última (vidrios y pantalla) cumplen una labor aceptable para diferenciar resplandores demasiado cercanos, lo que permite una mayor precisión en los resultados arrojados por los análisis de contraste del software de visión artificial, al momento de alejar los dedos del área de interacción.

Finalmente, los resultados del software desarrollado fueron comparados con las imágenes binarias que se obtienen con el CCV [21], para establecer las ventajas de la detección del dispositivo ante aumentos de iluminación exterior, en comparación con la que se obtiene en proyectos de este tipo que siguen la filosofía DIY. La detección lograda por medio del CCV, depende de las condiciones de iluminación infrarroja que se tengan para el sistema de detección, y utiliza ciertos valores iniciales como criterio para efectuar el filtrado; que deben ser establecidos por medio de consideraciones realizadas por los desarrolladores gracias a la interfaz que el software despliega con este fin. Como paso final, se umbraliza la imagen obtenida por medio de un valor establecido también “manualmente”.

La prueba se inicia estableciendo un *background* para la condición de oscuridad en el entorno del dispositivo. Luego se gradúan los valores iniciales para las propiedades en las que basa su análisis el CCV; hasta obtener un tipo de detección parecida a la que logra el software desarrollado. Ya con este paso terminado, se procede a encender la fuente de iluminación exterior; para observar la respuesta de la detección realizada por el CCV en este caso. En las figuras 6.9a y 6.9b, se puede examinar que tan efectiva es la detección del CCV ante variaciones de luz exterior. El ruido infrarrojo proviene de la misma fuente de iluminación exterior, que se ha venido utilizando durante la evaluación de los sistemas desarrollados en este proyecto.

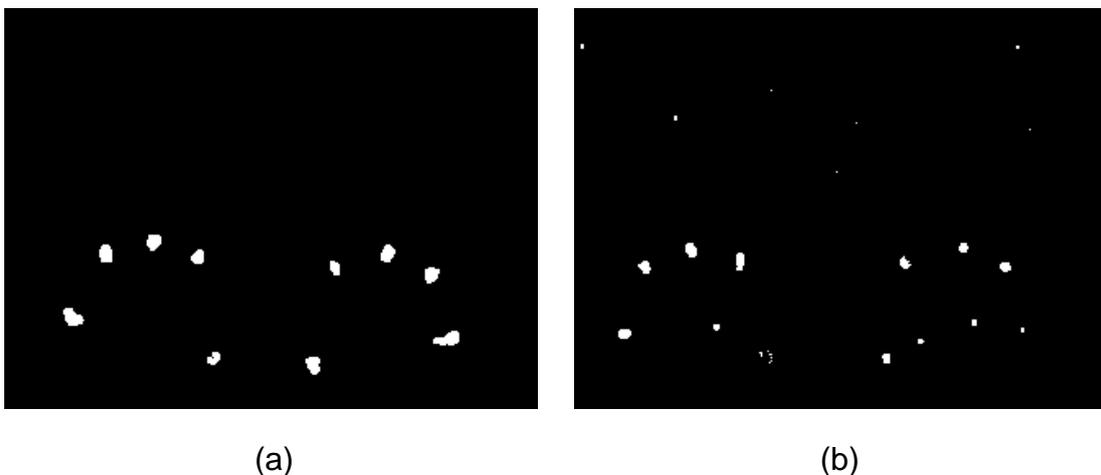


Figura 6.9. Detección del CCV, a) Condiciones de oscuridad, b) iluminación exterior.

Al analizar los resultados del CCV ante cambios en el *background* inicial, se observa que en las imágenes binarias obtenidas después de finalizar la etapa de filtrado digital; se presenta ruido en forma de pequeñas áreas que pueden ignorarse antes de la comunicación por TUIO, debido al bajo número de píxeles que las componen. Pero esto representa una complicación cuando las áreas de contacto detectadas han disminuido su magnitud por el aumento de la intensidad infrarroja en el área de interacción; ya que esto hace que en algunas ocasiones, ciertas áreas de contacto que son minúsculas comparadas con la resolución de la imagen, tiendan a desaparecer o a presentarse de manera intermitente, haciendo incómoda la futura comunicación usuario-computador.

Por otro lado, la imagen resultante del filtrado digital que realiza el CCV presenta interacciones fantasma debidas a ruido de resplandor, mucho más seguido que en el software de visión artificial desarrollado en este proyecto. Además en este último, las acciones de calibración inicial se reducen solamente a la modificación del umbral para obtener imágenes binarias.

Lo anterior puede ser consecuencia de que las propiedades que modifica el CCV, son susceptibles a las variaciones de los resplandores causantes del ruido. Esta situación también puede ocurrir durante los análisis de contraste que hace el software desarrollado, pero sus consecuencias se notan para niveles de iluminación exterior mucho más altos que el caso en estudio. Aunque cabe aclarar, que estas afirmaciones solo son válidas para trabajar con el hardware desarrollado en la parte instrumental de este proyecto.

Por estas razones, se concluye entonces que la mejor opción a utilizar para implementar la detección del dispositivo, es el software de visión artificial desarrollado en este proyecto. Esto permite cierta tolerancia en cuanto a variaciones experimentales en la ubicación de los iluminadores infrarrojos, hechas en pro de la búsqueda de mejores resultados en la detección de toques. En este punto, se puede proceder a la puesta en funcionamiento de todos los sistemas en conjunto.

## **6.2. FUNCIONAMIENTO COMO UN DISPOSITIVO INTEGRADO**

Después de realizada la validación del hardware, es de notarse que no se tiene aún un dispositivo completo, sino sistemas ópticos que funcionan cada uno por aparte con objetivos diferentes, debido a que aún falta la implementación del protocolo TUIO.

Para establecer el vínculo por medio de TUIO entre la parte visual y la parte instrumental del dispositivo; se propone seguir el orden de los pasos mostrados a continuación, con el fin de detectar errores de implementación lo más rápido posible. Las imágenes binarias que arroja el software de visión artificial, tienen

la resolución modificada a 800X600, que es la de la interfaz gráfica de las aplicaciones desarrolladas para evaluar el funcionamiento de los sistemas trabajando en conjunto.

### 6.2.1. Puesta en funcionamiento

Primero se debe constatar que se puede establecer una comunicación de forma local dentro del computador, por medio del envío de mensajes del protocolo TUIO desde un servidor simulado, que se comunica por el puerto 3333. Los mensajes son recibidos por una aplicación de prueba denominada como App1, que genera un pequeño círculo sobre una imagen por cada interacción detectada; siendo aprovechada entonces, para verificar que la información en los paquetes UDP recibidos por el cliente, corresponden a patrones de información que se enviarían desde un área de interacción real.

Este paso se realiza utilizando el *TUIOSimulator* [41], software que despliega la simulación de un área de interacción en forma de una imagen blanca, como se muestra en la figura 6.10a. Por cada acción generada sobre el simulador por medio de controles periféricos del computador, se envían mensajes con los identificadores de áreas de contacto y parámetros de las mismas hacia App1 que está sirviendo como cliente. La interfaz de App1, que se puede observar en la figura 6.10b; tiene la misma relación de aspecto que se despliega en el simulador, dando una mejor apreciación de las interacciones generadas.

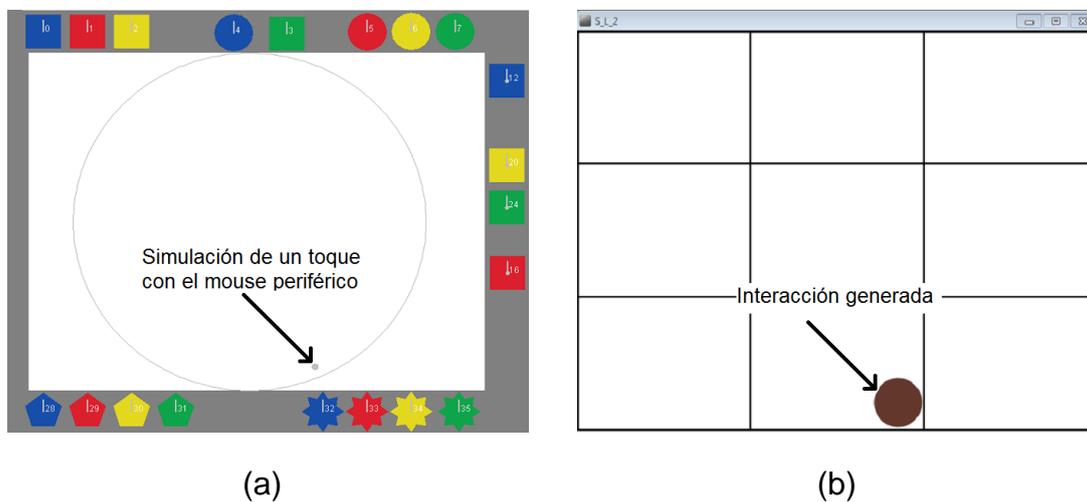


Figura 6.10. Estableciendo comunicación local, a) *TUIOSimulator*, b) App1.

Luego de verificar que la aplicación cliente, está en capacidad de recibir mensajes desde el área de interacción; se procede a constatar que también se presenta la transición de las imágenes resultantes del software de visión artificial a mensajes TUIO. Utilizando la aplicación interactiva denominada como App2, se comprueba que las imágenes binarias de la detección de toques sobre el área de interacción son bien correspondidos en la interfaz.

App2, escoge solo un toque entre todos los que se presentan en el área de interacción para generar una estela de círculos en la interfaz. Este efecto corresponde al trayecto del toque escogido en un intervalo de tiempo; mientras que para el resto de toques solo se presenta el despliegue de un círculo para cada imagen capturada en la señal de video. Así se comprueba que se presenta identificación de cada toque y seguimiento del mismo; ya que de no ser así, la interfaz de App2 siempre mostraría círculos sin estelas en cada cuadro desplegado, situación distinta a la exhibida en la figura 6.11.

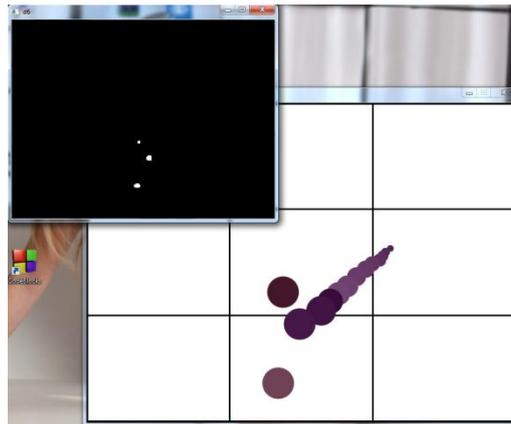
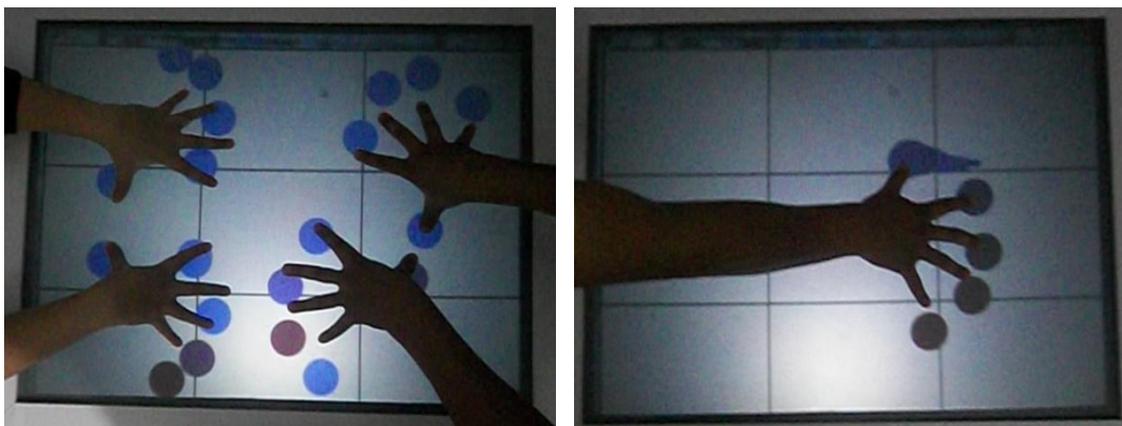


Figura 6.11. Prueba de identificación y seguimiento de toques.

Después de constatar la identificación y el seguimiento de los toques; se despliegan las interacciones en la interfaz gráfica, ya con el sistema de visualización en funcionamiento. En la figura 6.12a se observa la detección y respuesta a múltiples toques sobre el área de interacción al emplear App1, mientras que en la figura 6.12b se puede analizar el despliegue de App2 en el dispositivo integrado.



(a)

(b)

Figura 6.12. Dispositivo integrado, a) Respuesta a múltiples toques, b) Identificación y seguimiento de toques.

Se procede entonces a observar la continuidad con la que se presentan los toques en ciertos intervalos de tiempo haciendo uso de la aplicación que lleva como nombre App3. Al presentarse una interacción no interrumpida en una línea de tiempo, se genera un círculo que es más grande para cada cuadro en el que se identifica un área de contacto. El centro geométrico del área detectada corresponde al centro del círculo generado por la interfaz gráfica; círculo que se borra inmediatamente cuando se deja de presentar comunicación, percibiéndose un cambio brusco en los colores visualizados.

App3, está configurada para que el radio de los círculos desplegados en el área de interacción aumenten entre cuadro y cuadro. Bajo estas condiciones, se inicia entonces un análisis para verificar la identificación de toques a lo largo de ciertos periodos de tiempo. Los resultados son obtenidos bajo condiciones intermedias entre oscuridad y aumento de iluminación en el exterior; y son similares a los casos mostrados en la figura 6.13,

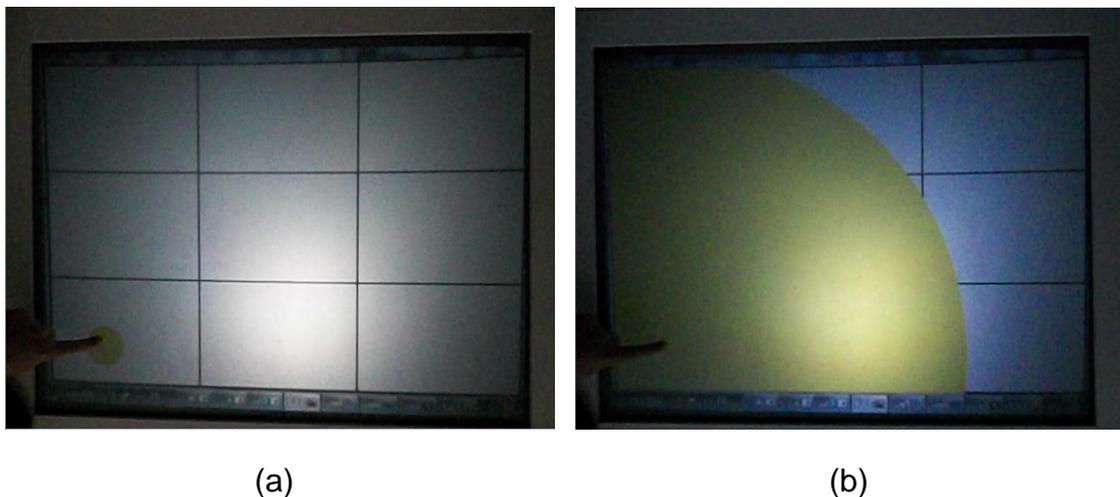


Figura 6.13. Continuidad de interacciones. a) Inicio de interacción, b) 30 segundos después.

Los círculos generados por la interfaz en la figura 6.13, responden a los toques del usuario hasta el punto en que sus diámetros superan la resolución de la interfaz gráfica; mostrando una pantalla de un solo color que cambia automáticamente cuando se deja de interactuar con el dispositivo.

### 6.2.2. Utilidad del dispositivo

Como paso final, el dispositivo debe ponerse a prueba para demostrar que cuenta con potenciales capacidades, útiles en instalaciones interactivas para obras de arte multimedia. En la figura 6.14, se muestra el dispositivo siendo utilizado por su primer usuario.



Figura 6.14. Capacidades inmersivas del dispositivo.

Es de notarse que la aplicación desplegada, capta la atención del usuario de prueba estableciendo un vínculo que se percibe como intuitivo, gracias a respuestas cromáticas a las interacciones que se presentan de forma caótica.

Para la parte visual, el óptimo aprovechamiento de las propiedades de la pantalla y el proyector, hace que la impresión lograda sobre el usuario sea equivalente en todas las zonas del área de interacción. Por otro lado, la parte instrumental permite que se reconozcan los toques del usuario de prueba, siendo tolerante a ciertas variaciones de iluminación exterior; haciendo que el vínculo logrado entre el humano y la máquina sea más difícil de romper ante la presencia de fuentes de luz en el entorno del dispositivo.

Además, la capacidad de reconocer múltiples toques, no solo permite aprovechar el despliegue de aplicaciones interactivas para formar un vínculo humano-máquina; ya que al contar con la posibilidad de interconectar las acciones de varios usuarios a la vez sobre la superficie de la mesa interactiva desarrollada, se logra que esta sirva como puente de comunicación entre varios participantes, sin la necesidad de que estos crucen palabra o tengan algún tipo de contacto físico. Esta situación se puede observar claramente en la figura 6.15.



Figura 6.15. Capacidades colaborativas del dispositivo.

## CONCLUSIONES

Se logra la obtención de una superficie *multitouch* que combina varios sistemas hardware y software, para que operando en conjunto puedan formar un vínculo satisfactorio cuando un usuario interactúa con el contenido digital de un computador.

El montaje de espejos logra un área de interacción a una altura apropiada para que el usuario pueda interactuar con la interfaz gráfica del dispositivo de forma cómoda y amigable.

Por otro lado, los materiales del área de interacción implementada con el fin de satisfacer prioritariamente, necesidades en cuanto a la calidad visual del dispositivo; ayudan a que la intensidad infrarroja para el sistema de detección, se presente de manera aceptable en un rango amplio de posibilidades de ubicación para los iluminadores infrarrojos; logrando que en las imágenes obtenidas, se puedan diferenciar digitalmente los niveles de ruido planteados en el capítulo 4, aún en casos que no siguen las indicaciones de diseño al píe de la letra, para posicionar los iluminadores.

El sistema de detección y el software de visión artificial permiten una detección limpia en cuanto a ruido infrarrojo se refiere, logrando interacciones que soportan pequeños cambios de iluminación exterior a diferencia de los dispositivos desarrollados con la filosofía DIY.

Además, se aportan conocimientos necesarios para el diseño, la construcción y la evaluación de superficies multitouch, que utilizan métodos de proyección en su parte visual y se basan en iluminación difusa para su parte instrumental.

## TRABAJOS FUTUROS

- Realizar una calibración de posicionamiento de toques.
- Elaborar nuevos diseños para el sistema de visualización, que varíen la estructura general del dispositivo, dependiendo del uso que se le vaya a dar.
- Estudiar nuevos métodos y materiales, que permitan disminuir la aparición de reflejos y resplandores para la parte instrumental, que causan el ruido en las imágenes detectadas.
- Implementar una interfaz gráfica para operar el software de visión artificial encargado de la detección, haciendo que su manipulación se pueda lograr de una forma más directa
- Desarrollar aplicaciones interactivas aprovechando las virtudes de integración social, que brinda el tamaño y los niveles de calidad visual logrados para el área de interacción.
- Realizar análisis sobre el comportamiento de grupos de varios usuarios, dependiendo de la aplicación interactiva desplegada.

## REFERENCIAS

- [1] Báscones Pere. "Arte e interactividad (I). Los logros de la reconciliación del arte, la ciencia y la tecnología". Universitat Oberta de Catalunya. 2003.
- [2] Gándara M. "¿Qué son los programas multimedios de aplicación educativa y cómo se usan?: Una introducción al modelo NOM". Uso de nuevas tecnologías y su aplicación en la educación a distancia, Módulos IV, V y VI. ULSA. México. 1999.
- [3] DeTar Charles. "Transgressive Interactive Art". Dartmouth College. 2007.
- [4] Pasman J., Pitrola C., Albuixech Y. "Arte y nuevas tecnologías ¿nuevos educadores?". Espacio Fundación Telefónica de Argentina. 2007.
- [5] Sardón M., Bender L. "Una aproximación a las obras interactivas como un sistema dinámico complejo". Espacio Fundación Telefónica de Argentina. 2006.
- [6] Alonso, Rodrigo. "Algunas propiedades de las instalaciones interactivas". Espacio Fundación Telefónica de Argentina. 2006.
- [7] Gándara M. La interfaz. Citado 16 de abril 2014. Disponible en línea en <http://roberto.dgme.sep.gob.mx/doc/multimedia/act5.pdf>.
- [8] Blake J. "Natural User Interfaces in .NET". Manning Publications. 2011. Page 1.
- [9] Marrero C. "Interfaz Gráfica de Usuario. Aproximación semiótica y cognitiva". Universidad de la Laguna. Tenerife. 2006.
- [10] EDUCAUSE authors. "7 things you should know about Multi-Touch Interfaces". EDUCAUSE. 2008.
- [11] Jakobsen M., Hornbaek K. "Up Close and Personal: Collaborative Work on a High-Resolution Multitouch Wall Display". University of Copenhagen. 2014.
- [12] Peltonen P., Kurvinen E., Salovaara A., Jacucci G., Ilmonen T., Evans J., Oulasvirta A., Saarikko P. "It's Mine, Don't Touch!: Interactions at a Large Multi-Touch Display in a City Centre". Helsinki Institute for Information Technology, Helsinki University of Technology, University of Helsinki. 2008.
- [13] Martorell E., de la Re A., Abad F. "Diseño y construcción de un banco de trabajo multitáctil de bajo coste". Universidad Politécnica de Valencia. 2010.
- [14] <http://www.circlsetwelve.com/products/diamondtouch.html>.

- [15] Hodges S., Izadi S., Bulter A., Rrustemi A., Buxton B. "ThinSight: Versatile Multi-touch Sensing for Thin Form-factor Displays". Microsoft Research Cambridge. 2007.
- [16] <http://www.microsoft.com/en-us/pixelsense/default.aspx>.
- [17] <http://www.reactable.com>.
- [18] [http://en.wikipedia.org/wiki/Microsoft\\_PixelSense](http://en.wikipedia.org/wiki/Microsoft_PixelSense).
- [19] Villanueva A. "Superficies multitáctiles interactivas: desarrollo de prototipo, software de control y paradigmas de aplicación a la visualización 3D". Universidade da Coruña. 2011.
- [20] <http://nuigroup.com/go/lite>.
- [21] <http://ccv.nuigroup.com>.
- [22] <http://nuigroup.com/touchlib/>.
- [23] Villamor C., Willis D., Wroblewski L. "Touch Gesture Reference Guide". Creative Commons. 2010.
- [24] <http://processing.org/>.
- [25] ©Da-Lite Screen Company. "The Secondary Reflection Issue". Angles of View. Volume VI, Issue 6. December, 2000.
- [26] CinemaSource Technical Bulletins. "Front Projection Screen Installation Methods". Copyright by CinemaSource, Inc. 2002.
- [27] ©Da-Lite Screen Company. "Front and Rear Projection?". Angles of View. Volume I, Issue 4. May, 1995.
- [28] ©Da-Lite Screen Company. "Diffusion Screens - A Closer Look". Angles of View. Volume III, Issue 10. October, 1997.
- [29] ©Da-Lite Screen Company. "Hot Spot or Not?". Angles of Reflection. Volume III, Issue 1. January, 2010.
- [30] ©Da-Lite Screen Company. "Vive la Contrast". Angles of Reflection. Volume III, Issue 3. May, 2010.
- [31] Olivares T., Cuenca P. "La Morfología Matemática en el Tratamiento Digital de Imágenes". Universidad de Castilla-La Mancha.
- [32] De la Escalera A. "Visión por Computador. Fundamentos y métodos". Pearson Educación S.A. Madrid. 2001. Página 15.

- [33] Singla R., Malhotra M., Agarwal D., Chopra D. "Multi Touch: An Optical Approach (Comparision of various techniques)". International Journal of Scientific and Research Publications. Volume III, Issue 4. April, 2013.
- [34] Rodriguez I., Lacasta A. "Apuntes de Física I de Tipografía. Óptica Geométrica". Universitat Politècnica de Catalunya. 1997. Pàgina 6.
- [35] Nui Group Authors. "Multitouch Technologies". Community Release, first edition. 2009. Page 13.
- [36] Carcedo y Franco A. "Programa de segmentación de regiones en imágenes médicas en MATLAB". Universidad de las Américas Puebla. 2004.
- [37] Guzmán C. "Detección y cuantificación de los cambios medioambientales mediante teledetección y procesamiento digital de imágenes a través de muestras satelitales obtenidas sobre el macizo colombiano". Universidad del Cauca. 2014.
- [38] Kumar A., Ramakrishnan M. "Personal Identification Using Foot Print Based on Sophisticated Contrast Enhancement". Journal of Computer Applications. Volume V, Issue 4. 2012.
- [39] Luft T., Colditz C., Deussen O. "Image Enhancement By Unsharp Masking The Depth Buffer". ACM Transactions on Graphics. Volume 25, Issue 3. July, 2006. Pages 1206-1213.
- [40] <http://www.tuio.org/?tuio20>.
- [41] <http://reactivision.sourceforge.net/>.
- [42] Ferrín C., Buchelly F. "Diseño y construcción de plataformas multitoque de tipo óptico, utilizando técnicas de visión artificial en tiempo real, para facilitar la interacción de los usuarios con el computador". Universidad del Cauca. 2010.
- [43] Kaltenbrunner M., Bovermann T., Bencina R., Costanza E. "TUIO: A Protocol for Table-Top Tangible User Interfaces". Universitat Pompeu Fabra, Bielefeld University, Medialab Europe. 2005.

## BIBLIOGRAFÍA

Grupo Proyecto Boipus. (2008). Software Open Source y el desarrollo de Nuevas Interfaces en el Arte Interactivo. Argentina: Tercer Simposio Prácticas de comunicación emergentes en la cultura digital.

Jensen M. (2011). Natural user interfaces from all angles. Denmark: Aalborg Universitet.

Blake J. (2011). Natural User Interfaces in .NET. ©Manning Publications Co.

L.Y.L. Muller. (2008). Multi-touch displays: design, applications and performance evaluation. Netherlands: Universiteit van Amsterdam.

Holzammer A. (2009). Combining Diffuse Illumination and Frustrated Total Internal Reflection for touch detection. Germany: Technische Universität Berlin.

Vargas A. (2010). Sistema multimedia interactivo de búsqueda de información por medio de una pantalla multitouch de bajo costo. Ecuador: Escuela Superior Politécnica del Litoral.

De la Re A. (2009). Interacción multitáctil en un sistema de generación procedural. España: Universidad Politécnica de Valencia.

NUI Group Authors. (2009). Multi-Touch Technologies. First edition. Community Release.

De la Escalera A. (2001). Visión por Computador. Fundamentos y métodos. España: Pearson Educación S.A.

Gonzalez R., Woods R. (2002). Digital Image Processing. Second Edition. United States of America: Prentice Hall.