

# INTERACCIÓN HUMANO-COMPUTADOR EN UN ESCRITORIO VIRTUAL

Andrés García – Miguel Carrasco

Escuela de Informática y Telecomunicaciones  
Universidad Diego Portales  
Vergara 432, 2do. piso, Santiago de Chile  
E-mail: {andres.garcia, miguel.carrasco}@mail.udp.cl

**Abstract:** Múltiples tecnologías de interacción humano-computador se han desarrollado en los últimos diez años, siendo una de las más extendidas las pantallas con sensibilidad táctil. Esta tecnología permite la interacción con dispositivos fijos o móviles empleando únicamente los dedos o algún marcador sobre un área delimitada. Precisamente, el área de trabajo en estos dispositivos constituye su principal limitación, ya que no pueden ser adaptados a otros tamaños o superficies. Esta investigación propone el diseño de un sistema de bajo costo que puede ser adaptado a una superficie de cualquier tamaño, emulando una interacción táctil. Nuestro prototipo utiliza un microproyector que proyecta un escritorio virtual, y una microcámara que captura las interacciones con el sistema. Dentro de las principales características de nuestro sistema, destacamos (1) la detección del movimiento de la mano en tiempo real, y (2) su uso intuitivo, el cual no requiere un entrenamiento previo del usuario. El sistema fue evaluado por diferentes usuarios evaluando aspectos como la usabilidad y la tasa de errores. De los resultados destacamos la facilidad de aprendizaje y el potencial de aplicabilidad en diferentes escenarios.

**KeyWords:** *interacción humano-computador, segmentación, escritorio virtual, seguimiento de la mano*

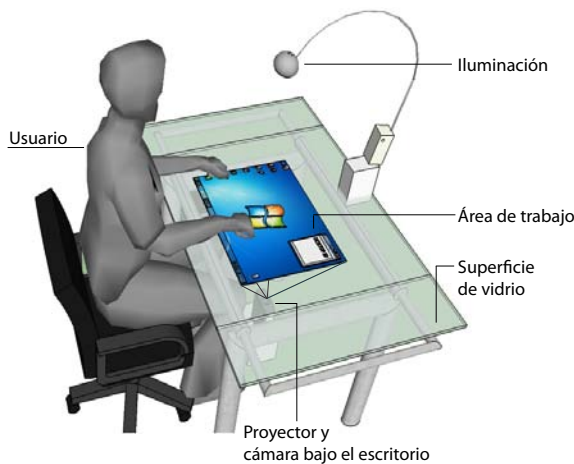
## 1. Introducción

La interacción humano-computador (HCI) ha sido un área de investigación muy relevante desde la creación de los primeros computadores. Uno de las propuestas más revolucionarias en este ámbito fue desarrollada en la década de los 60 con el *sketchpad* [1]. Este sistema permitía dibujar en la pantalla de un computador utilizando un lápiz de luz. Pese a las limitaciones de la tecnología de dicha época, este fue el primer sistema en proveer de una interacción de alta complejidad con el computador.

A comienzos de la década de los 90 aparecieron las primeras propuestas para generar interfaces más complejas y con una mayor interacción. Una de ellas fue el *Digital Desk* [2]. Este sistema permitía ver la imagen del computador proyectada sobre una mesa e interactuar con dichas proyecciones empleando únicamente los dedos del usuario. Lamentablemente debido al bajo poder de cómputo en esos años, este sistema fue desechado. Sin embargo, sentó las bases de un nuevo paradigma de comunicación, la *interacción con las manos*. Gracias a este primer avance, en 1995 fue desarrollado el sistema *finger paint* el cual, mediante el uso de un método de *finger tracking*, permitía dibujar con el dedo en el computador [3]. El mayor problema en estos primeros trabajos fue la baja capacidad de procesamiento del computador. En estos sistemas solo

se obtenía un seguimiento del dedo con un retardo visible. Actualmente, el incremento de la tecnología de los computadores ha permitido que este tipo de interacciones tengan el retardo menor a 50 ms [4], lo cual constituye un aplicación en tiempo real.

Uno de los actores que ha liderado un avance sustantivo en el área de HCI es la industria de la entretención. Empresas como Sony, Microsoft y Nintendo han desarrollado diversos tipos de tecnologías que permiten interactuar con videojuegos en forma natural empleando cámaras, sensores de aceleración y/o profundidad. De ellas destacamos (1) el PlayStation Move de Sony para la PlayStation 3, (2) Kinect de Microsoft para la X-Box360 y (3) el controlador de aceleración Wii para la consola de Nintendo. Aunque estos dispositivos trabajan de distinta forma, todos ellos permiten controlar los movimientos y acciones de los juegos en forma remota, empleando el movimiento del cuerpo como medio de interacción y comunicación. Lamentablemente este tipo de tecnologías se encuentran dirigidas principalmente al manejo interfaces en forma externa. De esta forma, no es posible interactuar con precisión sobre un área de trabajo. Por otra parte, la inserción de pantallas sensibles al tacto ha generado un inmenso desarrollo en aplicaciones y hardware para dispositivos móviles. Este tipo de interacción utiliza la detección del movimiento a



**Fig. 1** Escritorio virtual de interacción humano-computador propuesto para la interacción con el usuario.

través de la presión de uno o más dedos sobre una pantalla con sensibilidad táctil. Los ejemplos más directos se encuentran en las actuales *tablets* y *smartphones* los cuales tienen sistemas operativos optimizados para su uso. A pesar de estos avances, la mayor limitación en este tipo de tecnologías es que solo están diseñados para su uso en un área delimitada, la cual corresponde al tamaño del dispositivo.

Nuestra propuesta reúne ambas tecnologías al presentar un prototipo de interacción empleando gestos humanos. Así, nos basamos en parte en el paradigma de las interfaces de videojuegos, y al uso de superficies de interacción. En forma experimental, construimos un prototipo real con dos aplicaciones específicamente diseñados para incrementar la interacción humano-computador. La primera aplicación está diseñada para contar los dedos y ofrecer al usuario una interacción simple según la cantidad de dedos detectados sobre una superficie transparente. La segunda aplicación permite al usuario dibujar sobre una superficie empleando los dedos de una mano (Fig.1). Posterior a cada uso empleamos una encuesta con una escala psicométrica de Likert [5]. Los resultados indican que nuestra interfaz permite una interacción similar a la experiencia de una tablet. No obstante, el tipo de tecnologías empleadas es más cercano al paradigma de los videojuegos. En las siguientes secciones detallaremos los trabajos relacionados a nuestra propuesta, la metodología empleada, los resultados y conclusiones.

## 2. Trabajos relacionados

La invención de los computadores modernos requirió mecanismos de interacción que permitieran la entrada y salida de información. Como dispositivo de entrada fue adaptado el teclado del telégrafo en el primer

computador moderno en 1946 (ENIAC). Dos décadas más tarde, en el año 1964, los laboratorios MIT, Bell-labs y GE desarrollaron el monitor el cual permitió visualizar los comandos enviados por el teclado. En la misma década, diversos investigadores plantearon nuevas formas de interacción, algunos teóricos y otros más experimentales. Dentro de los más teóricos, las ideas propuestas por Licklider, Bush y Nelson apuntaron a que el computador debía adaptarse al trabajo de los usuarios [6]. En esta época los investigadores consideraban a los computadores como asistentes de documentos, donde también se podía almacenar información. Por otra parte, Licklider propuso que los usuarios y computadores se unieran para realizar investigaciones científicas empleando el computador para realizar los trabajos rutinarios, y a los usuarios para aportar nuevas ideas [6]. A pesar de la antigüedad, estos supuestos siguen siendo válidos aún en nuestros días.

En la línea de los trabajos experimentales, uno de los más sobresalientes fue el *Sketchpad* [1]. Este fue el primer dispositivo capaz de dibujar en el computador. Podía generar líneas, puntos, círculos dentro de la pantalla utilizando un lápiz de luz. Años más tarde, English, et al. crearon el *mouse*; un dispositivo que funciona para controlar puntero para la interfaz gráfica del sistema operativo [7]. La creación del mouse puso en incertidumbre el uso de otros dispositivos que cumplían la misma función. Estos fueron el Joystick isométrico, Step Keys y Text Keys. En busca del mejor dispositivo, Card, et al. [8] evaluaron la capacidad de los dispositivos para mover el puntero en ciertos puntos, haciendo líneas curvas y rectas. Los resultados determinaron que el mouse superó a los demás en posicionamiento, tasa de errores y velocidad de movimiento.

Aunque los computadores han aumentado su capacidad en todos los sentidos, la interfaz gráfica no ha tenido grandes cambios. Prácticamente se ha mantenido casi intacta a lo largo de los últimos 30 años, ya que sigue empleando el paradigma de interfaz WIMP (Windows, Icons, Menus, Pointing). Según Turró [6] y Beaudouin-Lafon [9] el paradigma WIMP ha llegado a su límite, provocando que no se puedan desarrollar nuevas HCI que logren reemplazar a los dispositivos como el mouse y el teclado. No obstante, en los últimos cinco años, y gracias a la expansión en la industria móvil y de la entretención se han desarrollado nuevos tipos de interacción humano-computador. A continuación analizaremos brevemente los distintos aspectos metodológicos y algorítmicos involucrados en nuestra investigación.

## 2.1. Interacción a través de medios virtuales

La interacción mediante el uso del seguimiento de manos y dedos con una cámara conectada al computador fue desarrollada a inicio de los años 90. Estas primeras aplicaciones analizaban el movimiento de la mano en áreas específicas dentro de una región. A continuación estudiaremos brevemente este tipo de tecnologías que ha sentado las bases de nuestra propuesta.

**Digital Desk:** Wellner [2] desarrolló en 1993 un sistema que une el escritorio virtual con uno físico. Para esto propuso tres sistemas, una *Calculadora*, *Paperpaint* y el doble *DigitalDesk*.

- *Calculadora:* es una aplicación con una interfaz simple donde el usuario introduce un número directamente sobre la interfaz o también puede introducir sobre un documento de papel que se encuentra sobre el *digitaldesk*. Esto último lo hace escaneando el área marcada y detectando el número.
- *PaperPaint:* es una aplicación que permite dibujar utilizando un lápiz en un papel que se encuentra en el escritorio. El sistema escanea la hoja dependiendo si el usuario retira el papel del área de trabajo, lo cual permite que el usuario cree un dibujo electrónico. Además permite copiar y pegar lo que se dibujó y así poder repetir la misma figura en otro documento. También permite copiar desde un documento y pegarlo en la hoja.
- *Doble DigitalDesk:* esta aplicación muestra la mano de un segundo usuario el cual interactúa con el primero. La idea es compartir el escritorio físico y virtual con un usuario que se encuentre en otro sitio; similar al framework propuesto por Microsoft en su proyecto *Illumishare* [4].

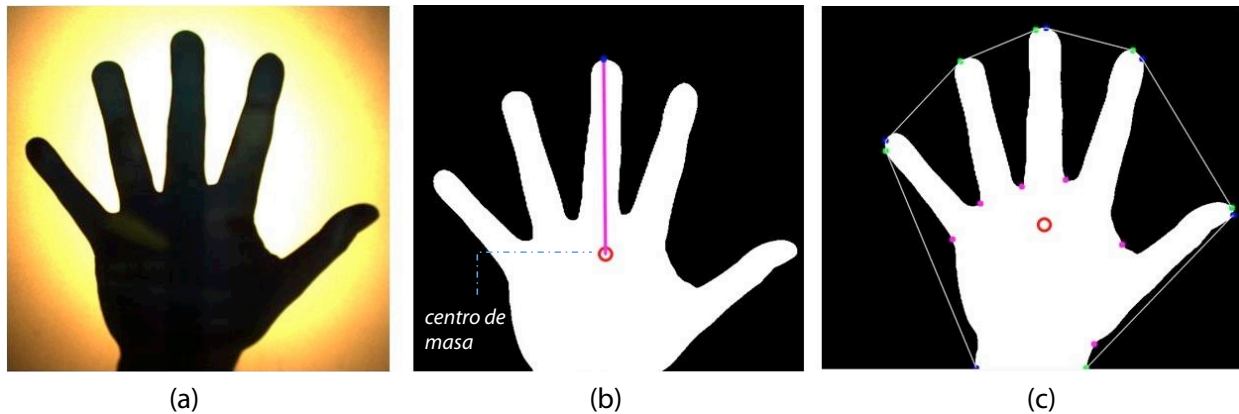
**Finger tracking as an input device:** Crowley, et al. [3] motivados por el *digitaldesk* de Wellner, determinaron que la operación básica consiste en el seguimiento del objeto a través de un puntero, como un dedo o lápiz. El seguimiento debe ser soportado por métodos que determinan si el sistema debe rastrear y detectar si hay un error. En forma simplificada los autores propusieron que el problema del seguimiento de un objeto se expresa como: dada la posición de un objeto en el frame  $k$ , se debe determinar la posición más probable en que se pueda encontrar el objeto en el frame  $k+1$ . Para rastrear el objeto se crea una máscara sobre el objeto a seguir y ésta se utiliza como plantilla.

**Finger Tracking for Digital Desk:** Brown y Thomas [10] diseñaron una actualización del sistema *Finger Tracking* basados en los trabajos previos de Wellner. Analizando los problemas que tuvieron los trabajos previos, implementaron un algoritmo el cual está estructurado en tres simples ideas.

- Encontrar la punta buscando la moda de las filas de la imagen, partiendo en la parte superior izquierda.
- La búsqueda se reduce limitando el uso de una región rectangular en la última posición conocida del dedo.
- La ubicación de la punta del dedo se puede acelerar mediante la exploración de la intersección de la muñeca con la parte inferior del escritorio siguiendo el contorno borde izquierdo de la mano hacia arriba y recogiendo el último punto encontrado en la punta del dedo.

Con estas tres ideas implementaron un algoritmo para la detección del dedo. El método depende de la inicialización del umbral y la detección de la punta del dedo por un barrido de trama previa [10]. Los resultados revelaron que su sistema es adecuado para el seguimiento de objetos o de la mano. Aunque en términos de precisión no es exacto, si permite detectar el dedo y proveer de interacción.

**Bare-Hand HCI:** von Hardenberg y Bérard [11] investigaron distintos métodos de HCI con la mano. La diferencia con otros trabajos era lograr que el sistema detectara el movimiento de la mano sin que ésta tuviera ningún dispositivo externo como marcador. Para lograr esto fue necesario encontrar ciertos requerimientos para que la interacción cumpla las necesidades del usuario. Los autores determinaron que en ciertos escenarios es mucho mejor el uso de detección y seguimiento de la mano comparando con un dispositivo de puntero para proveer de HCI. Un escenario planteado por los autores fue la incorporación del seguimiento de las manos con la imagen emitida por un proyector. Para encontrar la punta del dedo requirieron un algoritmo para extraer parámetros tales como el diámetro del dedo meñique y del dedo pulgar, tamaño de la sección rectangular que cubre el dedo entre otros. Luego efectuaron un análisis de la relación entre los dedos encontrados para determinar si pertenecen a la misma mano, u otra. Los resultados sobre la clasificación de los dedos resultó ser mejor efectiva, ya que detectaron las puntas de los dedos y su dirección, sin detectar el fondo donde se encuentra ésta.



**Fig. 2** (a) Imagen capturada por la cámara bajo el escritorio, (b) Imagen segmentada y punto de centro de masa, (c) Cálculo del Convex-Hull sobre la región y estimación de los dedos.

**Visual Tracking of Bare Fingers for Interactive Surfaces:** Letessier y Bérard [12] presentaron un sistema similar al anterior. Los autores propusieron un algoritmo de filtrado del tipo de segmentación de diferencias de imágenes. La diferencia respecto al sistema *Bare-Hand* radica en que la optimización del procesamiento no toma en cuenta la orientación del dedo. Este sistema es mucho más estable que los previos. Sin embargo, su uso es óptimo en un ambiente controlado. La contribución que hacen es el modelado de dos familias de algoritmos que son claves para el seguimiento de los dedos, además de un conjunto de opciones de diseño de sus instancias. La segmentación diferenciación de la imagen (IDS) es insensible a las sombras y el rápido rechazo del filtro (FRF) es invariante a la orientación de los dedos.

## 2.2. Detección del color piel

Diversas técnicas se han propuesto para detectar el color de la piel de la mano [13]. Las más extendidas son la detección por forma y la detección por color de la piel. En nuestra propuesta nos centramos en la segunda forma debido a que requiere un menor tiempo de cómputo. Para determinar el color de la piel, se buscan píxeles que cumplan características únicas y diferenciables de otros objetos. De esta forma, los modelos propuestos dependen del espacio de color empleado en la imagen (HSV, RGB, L.a.b) [14]. En cada espacio de color existen distintas técnicas, entre las más conocidas están las paramétricas, no paramétricas y región de piel explícitamente definida [15]. En particular, estudiamos el uso del modelo región de piel explícitamente definido dada su simplificación en la estimación de los valores de intensidad de la piel.

## 2.3. Seguimiento de objetos

Aunque en la literatura se han propuesto distintas técnicas de seguimiento o tracking de objetos, en este punto analizaremos la técnica seguimiento de punto, la cual fue aplicada en nuestra propuesta. Más detalles sobre otras técnicas se encuentran en Yilmaz, et al. [16]. El seguimiento de puntos consiste en detectar los objetos en imágenes consecutivas representados por puntos. La asociación de los puntos se basa en el estado anterior de los objetos que pueden incluir la posición del objeto y el movimiento. El seguimiento se puede formular como la correspondencia de los objetos detectados representados por puntos a través de marcos. La correspondencia de un punto tiene problemas complicados especialmente en la presencia de oclusiones en las entradas y salidas de los objetos.

En general, los métodos de punto de correspondencia se puede dividir en dos grandes categorías, los métodos deterministas y los métodos estadísticos. El método determinista provee de un costo de asociación dado por las restricciones del movimiento del objeto a observar que se encuentra en el frame anterior y el actual. En cambio, los métodos de correspondencia estadística utilizan el enfoque de espacio de estados para modelar las propiedades del objeto, tales como posición, velocidad y aceleración. Las mediciones suelen consistir en la posición del objeto en la imagen mediante un mecanismo de detección.

## 3. Metodología propuesta

Esta investigación pretende crear un escritorio virtual, el cual permita al usuario interactuar con el computador usando sus propias manos apoyadas sobre una superficie, sin tener que utilizar un dispositivo o marcador que controle el puntero. Los aspectos importantes de este proyecto son la interacción humano-computador y la detección de la mano. A continuación detallamos las principales fases del algoritmo propuesto.

### 3.1. Segmentación del color piel

Nuestra propuesta utiliza una microcámara dispuesta bajo el escritorio. Esto permite detectar en forma eficiente la sombra proyectada por la mano (Fig.2a). Recordemos que el sistema de iluminación se encuentra sobre la mesa, de esta forma, al encontrar en contacto la mano con la superficie de vidrio, la cámara captura una sombra de ésta (Fig.3). En la fase experimental, utilizamos el algoritmo de *región explícitamente definido* con la diferencia que algunos valores fueron adaptados experimentalmente respecto al algoritmo original. Por completitud, a continuación definimos conceptos básicos sobre una imagen digital. Sea  $\mathbf{I}$  una matriz de dimensión  $n \times m \times 3$  donde el valor de intensidad de cada píxel está el rango  $[0, \dots, 255]$ . La imagen  $\mathbf{I}$  es la composición de las matrices  $R$ ,  $G$  y  $B$  (Rojo, Verde y Azul). Para segmentar la piel, primero, evaluamos las siguientes reglas lógicas.

$$\begin{aligned} C_1 &= R < 40 \wedge G < 40 \wedge B < 40 \\ C_2 &= (\max(R, G, B) - \min(R, G, B)) > 15 \\ C_3 &= |R - G| > 15 \wedge R > G \wedge R > B \end{aligned} \quad (1)$$

donde  $C_1$ ,  $C_2$ , y  $C_3$  son tres condiciones requeridas por el algoritmo. Cada una de estas reglas puede tener un valor lógico {cero, uno}. Como resultado, la imagen binaria se construye a partir de la siguiente formulación

$$\mathbf{BN} = \begin{cases} 1 & \text{si solo si } \forall C_i = 1 \quad i = 1, 2, 3 \\ 0 & \text{caso contrario} \end{cases} \quad (2)$$

La imagen  $\mathbf{BN}$  se compone de valores ceros y unos, donde únicamente los valores verdaderos corresponderán a aquellos que definen la región de la mano, y separa la piel del fondo (Fig 2.b).

### 3.2. Punto medio y bordes de la mano

El objetivo del segundo paso es encontrar los bordes y el centro de masa o punto medio. Estas características nos permitirán encontrar los dedos de la mano. En una primera fase, empleamos el algoritmo de Canny [14] para buscar los contornos de la región. Posteriormente empleamos la imagen binaria  $\mathbf{BN}$  para buscar el centro de masa. Este proceso lo realizamos a través de los momentos estadísticos [14], definidos como

$$m_{rs} = \sum_{i,j \in \mathbf{R}} (i)^r (j)^s \quad (3)$$

donde  $(i, j)$  corresponde a una coordenada de la región binaria  $\mathbf{R}$ , cuyas posiciones son todas aquellas que tienen valor verdadero en la imagen binaria  $\mathbf{BN}$ .

Para obtener la posición del centro de masa, aplicamos los momentos estadísticos (3) según la siguiente formulación

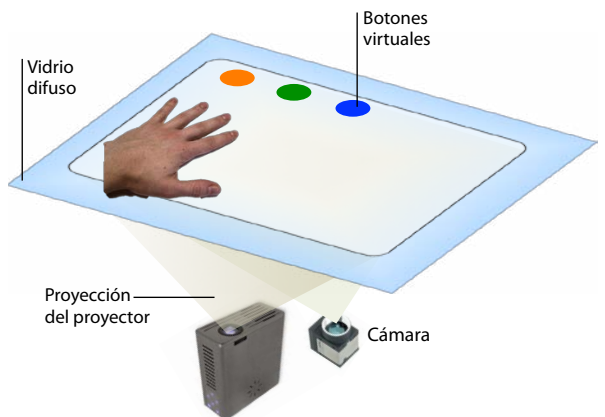
$$(\bar{i}, \bar{j}) = \left( \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (4)$$

donde  $(\bar{i}, \bar{j})$  corresponde al centro de masa de la región. La Fig.2b presenta un ejemplo del centro de masa de la región binaria destacada en un círculo.

### 3.3. Detección de dedos

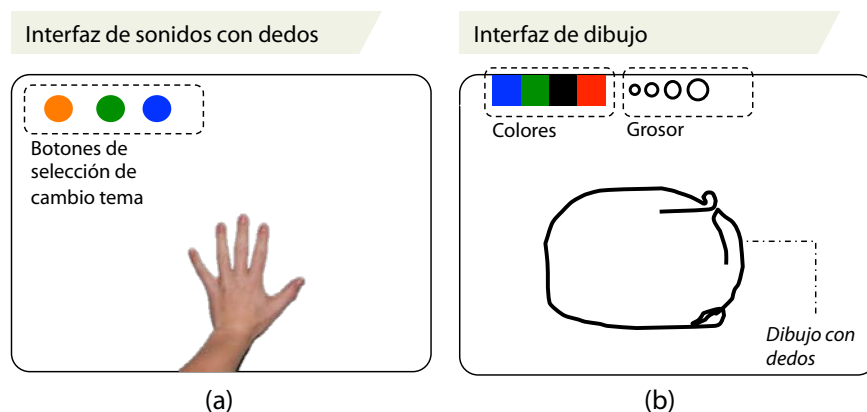
Para detectar los dedos formulamos una heurística que busca los cinco puntos más largos que existen entre el punto medio y los bordes. Este proceso lo realizamos a través del cálculo del *Convex Hull* el cual busca secciones convexas de un área geométrica binaria [15]. Esta técnica requiere determinar los bordes de la mano; obtenido en el paso anterior. Como resultado el algoritmo genera tres conjuntos de puntos: (1) *puntos de inicio*: los que indican donde empieza a formarse un espacio convexo, (2) *puntos finales*: los cuales marcan el fin del punto convexo, y (3) *puntos de profundidad*: los que junto a los dos pasos anteriores, determinan el espacio convexo en la región (Fig. 2c).

Nuestro algoritmo utiliza los puntos de inicio los cuales nos indican los dedos. Debido que los puntos no solo se encuentran en los dedos sino que se encuentran alrededor de toda la mano, aplicamos la siguiente condición de criterio. La coordenada  $(j)$  del punto de inicio tiene que ser menor a la coordenada  $(j)$  del punto medio de la mano. Esto se debe simplemente a que las posiciones de los píxeles van aumentando hacia abajo. En particular, la esquina superior izquierda tiene como coordenada la posición (1,1) y la esquina inferior derecha, la posición (800,600). Una vez aplicada dicha



**Fig. 3** Esquema de proyección y captura de información con el reflejo de la mano sobre el escritorio virtual.

**Fig. 4** Interfaces de interacción. (a) interfaz para modificar los sonidos con los dedos; (b) interfaz de dibujo sobre el escritorio



restricción, diferenciaremos los puntos que nos indican los dedos. Aún así nos planteamos indicar la punta del dedo más largo para efectos de que este punto sea utilizado como puntero para hacer el seguimiento en el escritorio virtual (Fig. 2c)

### 3.4. Seguimiento de la mano

Una vez detectada la mano con los puntos de los dedos y el punto medio, analizaremos su movimiento. En este proceso nos centramos en ubicar los puntos específicos de la mano. Por lo tanto, desarrollamos un seguimiento de punto correspondiente al dedo más largo de la mano. El seguimiento de punto analiza el o los puntos haciendo una comparación entre los estados del tiempo  $\{t-1\}$  y  $\{t\}$ . Para el algoritmo hemos considerado determinar ciertas restricciones al movimiento que el usuario debe generar para que así el sistema funcione en forma estable. Las restricciones son las siguientes:

1. La mano solo debe moverse dentro del área de trabajo.
2. La mano debe estar apoyada en el vidrio.
3. La mano debe estar en un rango de 120 a 60 grados respecto a al borde inferior del área de trabajo.

Dadas las restricciones del seguimiento, el sistema reconoce en un tiempo  $\{t\}$  la posición del dedo más largo y también la cantidad de dedos. Luego en el tiempo  $\{t+1\}$  hace la comparación tanto de la cantidad de dedos como de la posición del dedo más largo. De esta forma hacemos que el sistema pueda tener distintas acciones dependiendo de la cantidad de dedos que existan y la posición donde se encuentre.

### 3.5. Interacciones del sistema

Una vez determinados los dedos de la mano, y sus

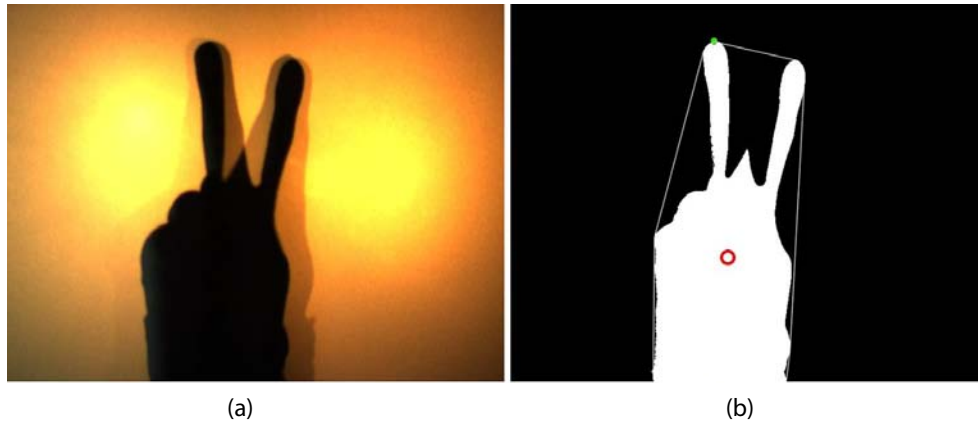
restricciones, diseñamos dos tipos de interacción, según el esquema propuesto en la Fig.3. A continuación describiremos cada una de éstas:

**Contar los dedos:** Esta aplicación permite en reproducir un sonido dependiendo de la cantidad de dedos detectados en el área de trabajo. La cantidad máxima de dedos que reproduce la aplicación es cinco. En la interfaz podemos ver una ventana de fondo blanco y en la parte superior se encuentran tres círculos (Fig. 4a). La funcionalidad de los círculos superiores permite que el usuario cambie los sonidos en forma virtual. La aplicación permite tener tres conjuntos de sonidos con un total de 15 archivos en formato ".wav". La principal utilidad de esta aplicación tiene por objeto que el usuario se adapte al manejo general del equipo. Para cambiar los sonidos, el dedo más largo debe posicionarse en cualquiera de los tres puntos sin importar la cantidad de dedos.

**Pintar con los dedos:** Esta aplicación permite que el usuario pueda dibujar utilizando un dedo. La interfaz tiene un fondo de color blanco y un menú en la parte superior donde se puede elegir el color, el grosor o borrar todo lo dibujado. Los colores para escoger son el blanco, negro, rojo, verde y azul. Los grosores de la línea son 3, 5 y 7 y 10, tal como ilustra la Fig. 4b. Al iniciar el programa el usuario comienza dibujando con el color negro, y el puntero del sistema operativo indica la posición del dedo. Para pintar se necesita colocar un solo dedo y el sistema seguirá su posición. Si el usuario coloca dos o más dedos solo se moverá el puntero sin pintar. Para cambiar el color, grosor o reiniciar la pantalla el usuario debe posicionar el dedo más largo en la opción que desea teniendo otro dedo detectado, y con dos dedos se selecciona una opción.



**Fig. 5** Defecto en el uso de dos sistemas de iluminación superior. (a) imagen original, (b) imagen binaria



#### 4. Pruebas y resultados experimentales

Esta sección presenta el análisis de los resultados. La primera evaluación mide la consistencia del programa en detalles específicos, observando concretamente el funcionamiento del algoritmo. La segunda evaluación presenta los experimentos a un grupo específico de personas. La finalidad de estas pruebas es comprobar la funcionalidad, rapidez, comodidad y adaptación de los usuarios con las aplicaciones, y también observar el modo de uso del escritorio virtual para futuras mejoras.

##### 4.1. Pruebas experimentales sobre la configuración

Estas pruebas se enfocan en la detección de la mano y los dedos. Como es común en este tipo de algoritmos, el factor más relevante en el algoritmo de segmentación tiene relación con el sistema de iluminación. En nuestro experimento empleamos dos focos de luz que proveen luz total en el área de trabajo. El primer foco corresponde a una ampollita alógena de 50w y el segundo foco a una luz incandescente de 60w. Lamentablemente la utilización de dos focos genera la proyección de dos sombras sobre el área de trabajo. Esto genera un solapamiento de dichas sombras creando una deformación en la imagen capturada por la microcámara (Fig. 5a). Como resultado, el algoritmo de binarización captura la deformación de la sombra (Fig.5b). A pesar de este inconveniente, esta deformidad no afecta totalmente la detección de la mano y sus dedos.

##### 4.2. Pruebas experimentales con usuarios

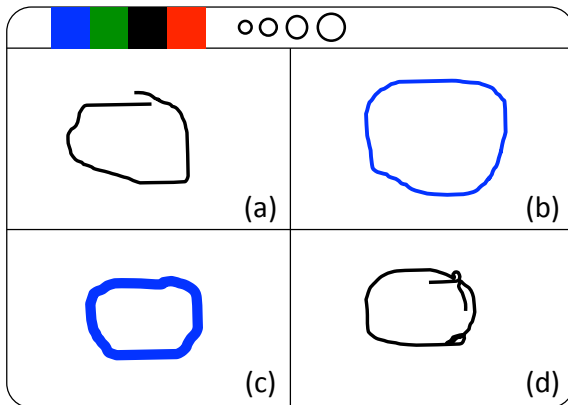
Las pruebas experimentales fueron aplicadas sobre ocho usuarios. La evaluación estuvo compuesta inicialmente por el uso y manejo de distintas interacciones con el fin de adaptarse al sistema, y posteriormente para determinar el grado de comodidad del usuario. Las aplicaciones diseñadas fueron (1) contar los dedos y reproducir un sonido y (2) pintar con los dedos.

**Contar los dedos:** El propósito de esta aplicación es entrenar al usuario para que aprenda a cambiar los sonidos utilizando los círculos de colores. Para ello, el usuario debe manipular el sistema según el dedo que tenga dispuesto sobre el escritorio virtual. Las evaluaciones en dicha tarea fueron las siguientes:

- Tiempo libre para experimentar con el sistema dando instrucciones generales.
- Reproducir los sonidos por defecto en orden ascendente y descendente según el número de dedos detectado por el escritorio.
- Los usuarios deben seleccionar el círculo que se encuentra al lado izquierdo y reproducir los sonidos colocando la cantidad de dedos en orden ascendente y descendente.
- Seleccionar los sonidos del círculo del medio y reproducir los sonidos colocando la cantidad de dedos en orden ascendente y descendente al número de dedos.
- Seleccionar los sonidos del botón del lado derecho y probar que al tener la misma posición de la mano y desplazarla por el área de trabajo, el programa no emite ningún sonido.

**Pintar con el dedo:** En esta aplicación el usuario debe pintar utilizando solo un dedo. Para seleccionar alguna opción el usuario debe posicionar su dedo más largo en el botón virtual correspondiente. Para evaluar esta prueba, dividimos el área de dibujo en cuatro cuadrantes, donde el usuario debe dibujar la misma figura en cada una de ellas cambiando el color y el grosor del pincel. Las pruebas consisten en:

- Tiempo libre entre uno y dos minutos para interactuar con el sistema.
- Dibujar cuadrados. En el primer cuadrante se dibuja un cuadrado de color verde con grosor 5, en el



**Fig. 6** Diagrama de tareas solicitadas al usuario en cuatro cuadrantes. En la primera tarea se solicita dibujar cuadrados en cuadro cuadrantes. La segunda tarea requiere dibujar círculos.

segundo cuadrante uno de color verde con grosor 3, en el tercer cuadrante un cuadrado de color rojo de grosor 7 y por último, en el cuarto cuadrante un cuadrado de color negro de grosor 3.

- Dibujar círculos: En el primer cuadrante tienen que dibujar un círculo de color negro y grosor 5, luego en el segundo cuadrante se hace un círculo de color azul y grosor 7, en el tercero se dibuja uno de color azul y grosor 3 y para finalizar un círculo en el último cuadrante de color negro y grosor 5 (ejemplo Fig.6).

La mayor problemática de los usuarios al dibujar consiste en mantenerse efectivamente en el cuadrante, sin dibujar o rayar otras secciones del área de dibujo. Otro punto analizado es la forma en cómo seleccionar las opciones y los dedos para interactuar con el sistema.

#### 4.3. Evaluación de la satisfacción

En el aspecto psicométrico realizamos una encuesta a los usuarios para tener respuestas concretas sobre la interacción con el escritorio digital una vez finalizada la prueba; según un modelo de escala de Likert no equilibrado [5]. Esta consulta se trató de una serie de preguntas con respuestas categorizadas en cuatro formas: (1) muy en desacuerdo, (2) en desacuerdo, (3) de acuerdo y (4) muy de acuerdo. Las respuestas permitieron tener una visión más clara de las falencias y virtudes que se encuentran en nuestro sistema. Las preguntas formuladas fueron las siguientes:

1. ¿Se sintió cómodo con la infraestructura del sistema?
2. ¿Es fácil hacer que el software haga exactamente lo que quiero?
3. ¿Se sintió cómodo con la iluminación del escritorio?

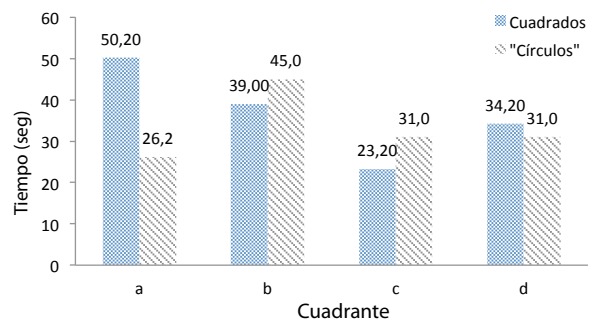
4. ¿Se acomodó rápido al uso de los programas?
5. ¿Es entendible el software utilizado?
6. ¿Considera simple la usabilidad para la interacción con el sistema?
7. ¿Le gustaría usar el escritorio diariamente?
8. ¿Considera adecuada la velocidad de respuesta del sistema?

En esta fase, las pruebas fueron realizadas sin informarles a los participantes la posición correcta que debía tener la mano en la aplicación. Por lo tanto, los usuarios debían hacer las tareas indicadas sin intervención de los diseñadores. Esto fue realizado teniendo para evaluar el uso intuitivo del sistema.

#### 4.4. Resultados de las pruebas experimentales

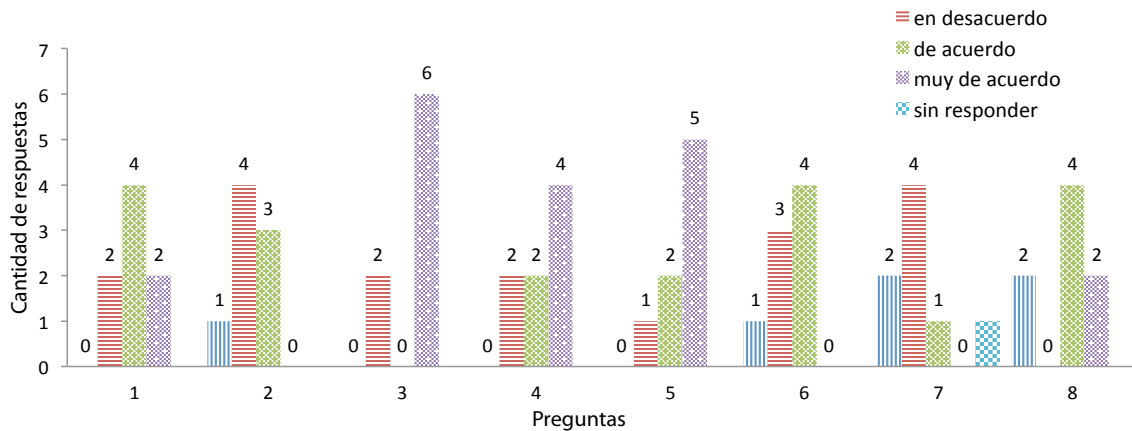
Según describimos anteriormente, la prueba de dibujo evaluó el tiempo requerido para lograr un dibujo en cada uno de los cuatro cuadrantes {a,b,c,d}; específicamente cuadrados y círculos (Fig.6). En el primer caso, observamos que el tiempo tiene una ligera disminución según el cuadrante seleccionado. El alza en el cuadrante {d} se puede deber a que el movimiento de la mano desde la esquina superior derecha hasta la esquina inferior izquierda tiene un trayecto más largo, y por lo tanto si los participantes se equivocaron tuvieron que realizar dos veces más el trayecto y también cambiar la opción correspondiente (Fig.7).

En la segunda tarea (dibujar círculos) se nota claramente una disminución del tiempo en el cuadrante {a}. Esto se debe a que en esta prueba el usuario solo debía cambiar el grosor del trazo a 7. Sin embargo, el círculo del cuadrante {b} requiere mayor tiempo ya que es necesario cambiar dos opciones. Los siguientes círculos que se dibujaron tienen un tiempo similar y se puede deber a un aprendizaje propio de la interacción con el sistema. Únicamente hay que notar un mayor tiempo en los cuadrantes {b} y {c} ya que en ellos hubo un cambio en el color y el grosor.



**Fig. 7** Tiempo promedio requerido para ejecutar un dibujo (cuadrado o círculo) en cada uno de los cuatro cuadrantes del escritorio.





**Fig. 9** Satisfacción de los usuarios al emplear el sistema en ocho preguntas (descritas en el punto 4.3)

En relación a los errores ejecutados por los usuarios, determinamos el número veces de error al ejecutar un dibujo variando la forma, color y grosor.

La mayor cantidad de errores al cambiar el grosor y color fueron cometidos al dibujar un cuadrado (Fig. 8). Esto se explica principalmente porque es la primera prueba. En cambio, la segunda presenta claramente una menor tasa de errores. Esto significa que los usuarios se van adaptando a la forma de utilizar el sistema. La mayor cantidad de errores se provocan por errores de uso de la interfaz gráfica, debido a que los colores son continuos (Fig. 6). De esta forma, al posicionar el dedo más largo en un color y luego, al seleccionar un grosor distinto provoca un cambio no intencionado del color.

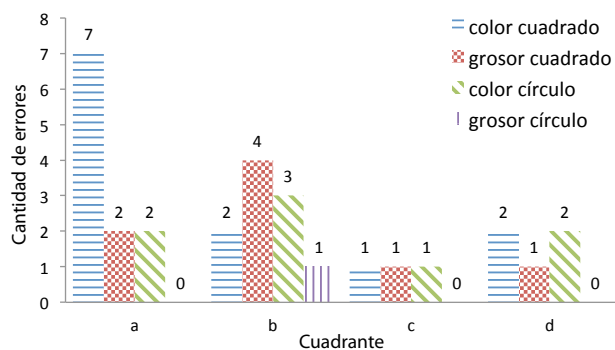
Existe una clara correlación entre el tiempo requerido para completar la tarea y la tasa de errores. Esto se evidencia en la alta tasa de errores en la primera tarea (dibujar cuadrados, Fig.8). Por el contrario, a medida que el usuario se adapta a la herramienta, la tasa de errores disminuye y requiere menos tiempo en efectuar la tarea solicitada.

#### 4.5. Resultados en la escala de satisfacción

Es importante destacar que a los usuarios no les fue enseñada una posición definida, ni tampoco los dedos a emplear en la interfaz (condición de inicio de la prueba). Por ello, cada usuario debió idear una manera para completar las pruebas. En general, la mayoría utilizó el dedo medio como el dedo principal y el dedo índice para señalar los dos dedos. Aunque esta elección no es incorrecta, no es la más eficiente porque al intentar pintar con el dedo medio es más difícil esconder los demás dedos y la postura de la mano es incómoda.

Otro detalle observado fue que los usuarios apoyaban solamente la punta de los dedos en el vidrio. Esto no es conveniente, ya que al estar la palma levantada y los dedos apoyados puede producir cansancio por uso a lo largo del tiempo. Por este motivo, después de interactuar con los participantes se decidió incluir a la prueba una pequeña sección explicativa sobre el uso correcto del dispositivo. Posterior a dicha explicación, se observó una mejoría notable en la utilización de la aplicación.

De acuerdo los resultados obtenidos en términos de satisfacción, a lo largo del experimento los usuarios se adecuaron al sistema, disminuyendo la tasa de errores y demorando menos tiempo. Según señala el gráfico de la Fig.9, los usuarios evaluaron en forma positiva las preguntas descritas en la Sec.4.3. En relación a los resultados negativos, éstas solo fueron obtenidas las preguntas #2 y #7. Para las demás preguntas las respuestas fueron valoradas positivamente, lo cual nos hace pensar que la usabilidad del sistema es viable.



**Fig. 8** Errores cometidos en promedio por los usuarios según color, grosor y forma (cuadrado y círculo).

## 5. Conclusiones

Esta investigación presenta el prototipo de un escritorio virtual manipulado a través de un sistema de interacción humano-computador. La mayor ventaja de nuestro sistema es la posibilidad de ser empleado en cualquier tipo de configuración, no únicamente un escritorio. Por otra parte nuestro sistema fue diseñado pensando un diseño de bajo costo ya que utiliza microproyectores y microcámaras, los cuales pueden ser fácilmente configurados y adaptados. Por otra parte, aunque el sistema de interacción tenga una similitud al presente en las pantallas táctiles, nuestro sistema tiene la ventaja de ser adaptable a una interfaz de cualquier tamaño. Además, nuestro algoritmo puede funcionar con múltiples dedos, o bien emplear el dedo más largo como puntero (en el caso de detectar múltiples dedos).

Las principales fases de nuestro algoritmo son: (1) detectar la mano, (2) seguir la mano y (3) interactuar con la aplicación. Primero, segmentamos la mano observando según el modelo de color definido por cada píxel de la imagen. Con esto obtenemos el punto medio de la mano; obtenemos las puntas de los dedos de la mano con la función *Convex Hull*, y finalmente la punta del dedo más largo. Segundo, para el seguimiento de la mano analizamos y comparamos la posición del dedo más largo en cada imagen del flujo de video. Tercero, la interacción se logra a través de una interfaz que funciona mediante la combinación entre la posición de la punta del dedo más largo y la cantidad de dedos detectada sobre el escritorio.

El sistema fue evaluado experimentalmente por ocho participantes. Las pruebas consistieron en la utilización de cada sistema y un conjunto de preguntas. A pesar de la baja cantidad de usuarios, las respuestas del sistema fueron positivas. Esta relación fue tan clara que nos permitió hacer un mejor análisis sobre la utilización del escritorio digital. En general, la mayoría de los usuarios no tuvo problemas al emplear el sistema, sin embargo, en una pregunta específica que indicaba si lo utilizarían diariamente, las respuestas fueron inclinadas en desacuerdo, por lo que concluimos que este sistema tiene un uso más experimental que cotidiano en la forma actual.

En conclusión, nuestro prototipo permite la integración con el usuario gracias a una rápida respuesta del sistema ya que funciona en tiempo real. Es muy importante destacar que nuestro sistema no requiere ningún tipo de entrenamiento previo para reconocer la mano, ya que puede funcionar para cualquier mano. Lamentablemente, si se detecta la presencia de un objeto ajeno en el área de trabajo que genere sombra sobre el escritorio, nuestro

algoritmo será ineficiente en el reconocimiento de la mano. Para que lo anterior no ocurra el usuario debe estar sentado en una posición correcta (Fig.1). Como trabajo futuro resta evaluar la aplicación en una mayor cantidad de usuarios, y aumentar las opciones de dibujo en el prototipo.

**Agradecimientos:** Este trabajo fue financiado por la Comisión Nacional de Ciencia y Tecnología (CONICYT) bajo el proyecto no.11100098, y por la Escuela de Informática y Telecomunicaciones de la Universidad Diego Portales.

## 6. Referencias

- 1 Sutherland, Ivan Edward. Sketchpad: A man-machine graphical communication system. In *Proceedings of the SHARE design automation workshop* ( 1964), ACM, 6329-6346.
- 2 Wellner, Pierre. Interacting with paper on digitalDesk. *Communications*, 36 (1993), 87-96.
- 3 Crowley, James L., Bérard, François, and Coutaz, Joëlle. Finger Tracking as an Input Device for Augmented Reality. In *IEEE International Conference on Automatic Face and Gesture Recognition - FGR* ( 1995), IEEE CS, 195-200.
- 4 Junuzovic, Sasa, Inkpen, Kori, Blank, Tom, and Gupta, Anoop. IllumiShare: sharing any surface. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* ( 2012), ACM, 1919-1928.
- 5 Likert, Rensis. A technique for the Measurement of Attitudes. *Archives of Psychology*, 140 (1932), 1-55.
- 6 Turró, Ribera. Evolución y tendencias en la interacción persona-ordenador. *El Profesional de la Información*, 14, 6 (2005).
- 7 English, William K, Engelbart, Douglas C, and Berman, Melvyn L. Display-Selection Techniques for Text Manipulation. *IEEE Transactions On Human Factors In Electronics*, HFE, 8 (1967), 5-15.
- 8 Card, Stuart K., English, William K., and Burr, Belt J. Evaluation of Mouse, Rare-Controlled Isometric Joystick, Step Keys, and Text Keys for Text Selection on a CRT. *Ergonomics* (1978), 601-613.
- 9 Beaudouin-Lafon, Michel. Designing interaction, not interfaces. In *Proceedings of the working conference on Advanced visual interfaces* (Italy 2004), ACM, 15-22.
- 10 Brown, Thomas and Thomas, Richard C. Finger Tracking for the Digital Desk. In *Proceedings of the First Australasian User Interface Conference* ( 2000), IEEE CS.
- 11 von Hardenberg, Christian and Bérard, François. Bare-hand human-computer interaction (. In *Proceedings of the 2001 workshop on Perceptive user interfaces* ( 2001), ACM, 1-8.
- 12 Letessier, Julien and Bérard, François. Visual tracking of bare fingers for interactive surfaces. In *Proceedings of the 17th annual ACM symposium on User interface software and technology* ( 2004), ACM, 119-122.
- 13 Vezhnevets, Vladimir, Sazonov, Vassili, and Andreeva, Alla. A Survey on Pixel-Based Skin Color Detection Techniques. In *Graphicon* ( 2003), ACM, 85-92.
- 14 Gonzalez, Rafael C. and Woods, R. E. *Digital Image Processing*. Prentice Hall, 2008.
- 15 Burger, Wilhelm and Burge, Mark J. *Digital Image Processing*. Springer, 2008.
- 16 Yilmaz, Alper, Javed, Omar, and Shah, Mubarak. Object Tracking: A Survey. *ACM Computing Surveys* , 38, 4 (2006).