

on del problema1.11Definición del problema1.1 on1.22Justificación1.2
rápida de objetos2.18Detección rápida de objetos usando características simples utilizando el
clasificador de cascada impulsada2.1

CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE EDUCACIÓN SUPERIOR DE ENSENADA, BAJA CALIFORNIA



PROGRAMA DE POSGRADO EN CIENCIAS EN CIENCIAS DE LA COMPUTACIÓN

Control de computadora basado en gestos con las manos en circunstancias de baja iluminación

Tesis

para cubrir parcialmente los requisitos necesarios para obtener el grado de
Maestro en Ciencias

Presenta:

América Ivone Mendoza Morales

Ensenada, Baja California, México
2015

Tesis defendida por

América Ivone Mendoza Morales

y aprobada por el siguiente comité

Dr. Vitaly Kober

Director del Comité

Dr. Hugo Hidalgo Silva

Miembro del Comité

Dr. Josué Álvarez Borrego

Miembro del Comité

Dra. Ana Isabel Martínez García

*Coordinador del Programa de
Posgrado en Ciencias de la Computación*

Dr. Jesús Favela Vara

Director de Estudios de Posgrado

Octubre, 2015

Resumen de la tesis que presenta América Ivone Mendoza Morales como requisito parcial para la obtención del grado de Maestro en Ciencias en Ciencias de la Computación.

Control de computadora basado en gestos con las manos en circunstancias de baja iluminación

Resumen elaborado por:

América Ivone Mendoza Morales

El reconocimiento de gestos con las manos ha sido un tema relevante en distintas áreas de las ciencias de la computación, por ejemplo en HCI es importante pues ayuda a crear una interacción natural entre la computadora y el usuario, por lo que se han desarrollado diversos métodos para encontrar el modelo que funcione en tiempo real y en diversas circunstancias. De manera que se pretende crear un modelo que fusione la información proporcionada por el dispositivo Kinect y haga el reconocimiento de gestos estáticos y dinámicos en tiempo real en circunstancias de baja iluminación y cuando existe oclusión. Dicho modelo será aplicado para crear un sistema que sirva como control de una computadora, es decir que los gestos puedan ser utilizados como el cursor de esta.

Palabras Clave: **Gestos con las manos, kinect, baja iluminación, oclusión.**

Abstract of the thesis presented by América Ivone Mendoza Morales as a partial requirement to obtain the Master of Science degree in Master in Sciences in Computer Science.

Computer control based in hand gestures in circumstances of low illumination

Abstract by:

América Ivone Mendoza Morales

The recognition of hand gestures has been prominent in different areas of computer science, eg. HCI is important because it helps create a natural interaction between the computer and the user, so have developed various methods to find the model that works in real time and in different circumstances. So it is to create a model that merges the information provided by the Kinect device, then the recognition of static and dynamic gestures in real time under conditions of low light and when there is occlusion. This model be applied to create a system that serves as a control computer, is that gestures can be used as the cursor.

Keywords: **Hand gestures, kinect, low illumination, occlusion.**

Dedicatoria

A ...

Agradecimientos

A ...

Al Centro de Investigación Científica y de Educación Superior de Ensenada.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por brindarme el apoyo económico para realizar mis estudios de maestría.

Tabla de contenido

Página

Resumen en español	iv
Resumen en inglés	v
Dedicatoria	vi
Agradecimientos	vii
Lista de figuras	x
Lista de tablas	xi
1. Introducción	1
1.1. Definición del problema	1
1.2. Justificación	2
1.3. Objetivo general	2
1.4. Objetivos específicos	2
1.5. Limitaciones y suposiciones	3
1.6. Reconocimiento de gestos con la manos	3
1.6.1. Etapas del reconocimiento de gestos	4
1.6.1.1. Detección	4
1.6.1.2. Seguimiento	5
1.6.1.3. Reconocimiento	6
1.7. Estado del arte	7
1.8. Organización de la tesis	7
2. Marco teórico	8
2.1. Detección rápida de objetos usando características simples utilizando el clasificador de cascada impulsada	8
2.1.0.4. Another title	8
3. Wow title	10
3.1. Stuff	10
3.2. Testing tables	11
3.3. Testing equations	11
3.4. Diseño de un método para identificar fatiga física	12
3.4.1. Extracción de parámetros	12
3.4.2. Cálculo de la demanda de oxígeno	13
4. Very título	14
5. Other much title	15
6. Título	16
7. Otro título	17
8. Conclusiones	18
Lista de referencias	19

Tabla de contenido (continuación)

A.	Apéndice
-----------	-----------------

20

Lista de figuras

Figura		Página
1.	Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam sit amet lobortis turpis. Praesent auctor mi metus.	9
2.	Wow sdf.	15

Lista de tablas

Tabla

Página

1. Mauris et imperdiet tortor. Maecenas consectetur lacus elit, dignissim eleifend dolor ornare ut. Aenean euismod porta nisi, et volutpat ex laoreet sit amet. Sed ac elit vestibulum neque ultrices feugiat 12

Capítulo 1. Introducción

La interacción entre humanos se lleva a cabo gracias a la comunicación que existe entre ellos, esta puede ser oral o escrita, generalmente, por no decir siempre, viene acompañada de gestos realizados con la cara, manos, cuerpo. Estos gestos sirven como complemento de la comunicación ya ayudan a que nuestra idea se percibida de manera correcta.

El creciente desarrollo de la tecnología, a llevado a crear y estudiar distintas áreas de las ciencias computacionales, particularmente HCI (por sus siglas en inglés), la área encarga del estudio, diseño e interacción del humano con la computadora. Uno de los objetivos principales es que la interacción sea de manera natural. Por lo que no es extraño que los investigadores de HCI se hayan interesado en los gestos corporales, en especial los gestos realizados con las manos, para crear un ambiente natural entre el usuario y la computadora. Para obtener una interacción natural, entre estos dos actuadores, se necesita hacer el reconocimiento de los gestos, esto ha sido cada vez más sencillo gracias al avance de la tecnología, en especial en los dispositivos de visión como distintos tipos de cámaras, y al crecimiento en la capacidad de procesamiento de las computadoras. Aunque existen diversos métodos y sistemas para lograr el reconocimiento, no existe ninguno que nos pueda dar un reconocimiento totalmente preciso en todas las situaciones que se presentan en el mundo real.

Es por eso que se propone crear un sistema que haga el reconocimiento de gestos con las manos, en situaciones donde existe baja iluminación y cuando tenemos oclusión causada por los dedos. El sistema se enfoca principalmente en atacar los problemas de gestos con las manos que no tienen movimiento, y después se abordarán los gestos con las manos que involucran movimiento. El sistema aplicará los gestos como control de la computadora, esto con ayuda del dispositivo Kinect como herramienta para capturar la información de entrada.

1.1. Definición del problema

nbkjghgjhyfkhfiy

1.2. Justificación

A finales de los años noventa se empezaron a desarrollar técnicas para reconocer gestos con las manos, las primeras fueron basadas en contacto y le siguieron las basadas en la visión, debido al avance de la tecnología estas fueron las más aceptadas.

Aunque existen diversos métodos para el reconocimiento de gestos con las manos, con buena precisión, sigue siendo un problema abierto ya que no es fácil tener un sistema que se adecue a todo tipo de situaciones como: amigable con el usuario, invariante a la iluminación, rotación, al fondo, que funcione en tiempo real o cuando exista oclusión.

1.3. Objetivo general

Desarrollar un sistema que permita controlar la computadora haciendo uso de gestos con las manos, estáticos y dinámicos. El sistema debe ser robusto, funcionar en circunstancias de baja iluminación, cuando exista oclusión en gestos dinámicos.

1.4. Objetivos específicos

- Identificar los métodos actuales de reconocimiento de gestos, estáticos y dinámicos cuando existe baja iluminación y en el caso de los gestos dinámicos cuando existe oclusión.
- Obtener conocimiento acerca del funcionamiento de sistema Microsoft Kinect.
- Desarrollar un sistema de reconocimiento de gestos estáticos y dinámicos, fusionando la información de los sensores de profundidad de dos dispositivos Kinect. El sistema desarrollado deberá funcionar en circunstancias de baja iluminación y también cuando existe oclusión, causada por los dedos.
- Analizar el sistema diseñado, en cuanto a su eficiencia presentada en base al reconocimiento de los gestos y tiempo de respuesta, en circunstancias de baja iluminación y oclusión. En el análisis del sistema se usará información real.
- Comparar los modelos propuestos con los existentes, en base al tiempo de respuesta y la eficiencia en cuanto al reconocimiento del gesto.

1.5. Limitaciones y suposiciones

Gran porcentaje de los trabajos previos en el área de reconocimiento de gestos con las manos basados en el modelo de la visión utilizan cámaras digitales o cámaras web. Esta investigación utiliza el dispositivo Kinect, para obtener la información de entrada del sistema.

De manera que las limitaciones del sistema propuesto están dadas por las características dicho dispositivo, tales como la distancia a la que se encuentra el dispositivo con el usuario (poner la distancia), la resolución de las imágenes a color (poner resolución) y la resolución del sensor infrarrojo (poner resolución).

También el sistema depende de dos sensores Kinect, que se utilizarán en el caso que exista oclusión.

Otra limitante es el número de gestos que podrá reconocer el sistema.

Se supone el área de trabajo como un cuarto estándar con buena iluminación (enfocado a pruebas con la cámara color del sistema Kinect)

1.6. Reconocimiento de gestos con las manos

La definición de gestos (?) son movimientos del cuerpo expresivos y significativos que involucren a los dedos, manos, brazos, cabeza, cara o cuerpo con la intención de transmitir información relevante o de interactuar con el ambiente. De acuerdo con la literatura (?) los gestos con las manos se clasifican en estáticos y dinámicos, los primeros están definidos como la posición y orientación de la mano en el espacio manteniendo esta pose durante cierto tiempo, por ejemplo para hacer una señal de aventón, a diferencia de los gestos dinámicos donde hay movimiento de la pose, un ejemplo es cuando mueves la mano en señal de adiós. De aquí en adelante entiéndase el término gestos con las manos, como gestos.

El reconocimiento de gestos se divide en tres fases(?), detección o segmentación; extracción de características seguimiento; dependiendo si los gestos son dinámicos, por último la etapa final el reconocimiento del gesto. Este se clasifican en dos modelos, basados en la visión y en contacto, esta clasificación depende de la manera en que son

capturados los datos, es decir la forma en que se obtiene el gesto, para posteriormente poderlo reconocer.

Los primeros acercamientos para llevar a cabo el reconocimiento de gestos fue usando modelos de contacto (?) y (?), como su nombre lo dice utilizan dispositivos que están en contacto físico con la mano del usuario, esto para capturar el gesto a reconocer, por ejemplo existen guantes de datos, marcadores de colores, acelerómetros y pantallas multi-touch, aunque estos no son tan aceptados pues entorpecen la naturalidad entre la interacción del humano y la computadora. Los modelos basados en la visión surgieron como respuesta a esta desventaja, estos utilizan cámaras para extraer la información necesaria para realizar el reconocimiento, los dispositivos van desde cámaras web hasta algunas más sofisticadas por ejemplo cámaras de profundidad.

En este trabajo, se toma el enfoque basado en la visión ya que se quiere obtener un sistema que para el usuario sea fácil de interactuar, y esta interacción sea natural y una manera de lograr esto es tomando este enfoque. Estos tienen mayor complejidad (acomoda este párrafo :P)

Los métodos basados en la visión se pueden representar por dos modelos (?), los basados en 3D, da una descripción espacial en 3D de la mano, y los basados en apariencia, como su nombre lo dice se basan en la apariencia de la mano. Los modelos basados en apariencia se dividen en dos categorías, los estáticos (modelo de silueta, de contorno deformables) y de movimiento (de color y movimiento).

1.6.1. Etapas del reconocimiento de gestos

Enseguida se describen las etapas del reconocimiento de gestos (detección, seguimiento y reconocimiento), con los métodos para llevar cada una de estas.

1.6.1.1. Detección

En esta etapa se detecta y segmenta la información relevante de la imagen (la mano), con la del fondo, existen distintos métodos para obtener dichas características como la de color de la piel, forma, movimiento, entre otras que generalmente son combinaciones

de alguna de estas, para obtener un mejor resultado. Enseguida se describe brevemente cada una de estas.

- **Color de la piel:** Se basa principalmente en escoger un espacio del color, es una organización de colores específica; como; RGB (rojo, verde, azul), RG (rojo, green), YCrCb (brillo, la diferencia entre el brillo y el rojo, la diferencia entre el brillo y el azul), etc. La desventaja es que si el color de la piel es similar al fondo, la segmentación no es buena, la forma de corregir esta segmentación es suponiendo que el fondo no se mueve con respecto a la cámara.
- **Forma:** Extrae el contorno de las imágenes, si se realiza correctamente se obtiene el contorno de la mano. Aunque si se toman las yemas de los dedos como características, estas pueden ser ocluidas por el resto de la mano, una posible solución es usar más de una cámara.
- **Valor de píxeles:** Usar imágenes en tonos de gris para detectar la mano en base a la apariencia y textura, esto se logra entrenando un clasificador con un conjunto de imágenes.
- **Modelo 3D:** Depende de cual modelo se utilice, son las características de la mano requeridas.
- **Movimiento:** Generalmente esta se usa con otras formas de detección ya que para utilizarse por sí sola hay que asumir que el único objeto con movimiento es la mano.

1.6.1.2. Seguimiento

Consiste en localizar la mano en cada cuadro (imagen). Se lleva a cabo usando los métodos de detección si estos son lo suficientemente rápidos para detectar la mano cuadro por cuadro. Se explica brevemente los métodos para llevar a cabo el seguimiento.

- **Basado en plantillas:** Este se divide en dos categorías (Características basadas en su correlación y basadas en contorno), que son similares a los métodos de detección, aunque supone que las imágenes son adquiridas con la frecuencia suficiente

para llevar acabo el seguimiento. Características basadas en su correlación, sigue las características a través de cada cuadro, se asume que las características aparecen en mismo vecindario. Basadas en contorno, se basa en contornos deformables, consiste en colocar el contorno cerca de la región de interés e ir deformando este hasta encontrar la mano.

- Estimación óptima: Consiste en usar filtros Kalman, un conjunto de ecuaciones matemáticas que proporciona una forma computacionalmente eficiente y recursiva de estimar el estado de un proceso, de una manera que minimiza la media de un error cuadrático, el filtro soporta estimaciones del pasado, presente y futuros estados, y puede hacerlo incluso cuando la naturaleza precisa del modelo del sistema es desconocida; para hacer la detección de características en la trayectoria.
- Filtrado de partículas: Un método de estimación del estado de un sistema que cambia a lo largo del tiempo, este se compone de un conjunto de partículas (muestras) con pesos asignados, las partículas son estados posibles del proceso. Es utilizado cuando no se distingue bien la mano en la imagen. Por medio de partículas localiza la mano la desventaja es que se requieren demasiadas partículas, y el seguimiento se vuelve imposible.
- Camshift: Busca el objetivo, en este caso la mano, encuentra el patrón de distribución mas similar en una secuencia de imágenes, la distribución puede basada en el color.

1.6.1.3. Reconocimiento

Es la clasificación del gesto, la etapa final del reconocimiento, la clasificación se puede hacer dependiendo del gesto. Para gestos estáticos basta con usar algún clasificador o empatar el gesto con una plantilla. En los dinámicos se requiere otro tipo de algoritmos de aprendizaje de máquina. A continuación se encuentran los principales métodos para llevar acabo el reconocimiento del gestos.

- K-medias: Consiste en determinar los k puntos llamados centros para minimizar el error de agrupamiento, que es la suma de las distancias de todo los puntos al

centro de cada grupo. El algoritmo empieza localizando aleatoriamente k grupos en el espacio espectral. Cada píxel en la imagen de entrada es entonces asignada al centro del grupo mas cercano

- K-vecinos cercanos (KNN, por sus siglas en inglés): Este es un método para clasificar objetos basado en las muestras de entrenamiento en el espacio de características.
- Desplazamiento de medias: Es un método iterativo que encuentra el máximo en una función de densidad dada una muestra estadística de los datos.
- Máquinas de soporte vectorial (SVM, por sus siglas en inglés). Consiste en un mapeo no lineal de los datos de entrada a un espacio de dimensión más grande, donde los datos pueden ser separados de forma lineal.
- Modelo oculto de Markov (HMM, por sus siglas en inglés) es definido como un conjunto de estados donde un estado es el estado inicial, un conjunto de símbolos de salida y un conjunto de estados de transición. En el reconocimiento de gestos se puede caracterizar a los estados como un conjunto de las posiciones de la mano; las transiciones de los estados como la probabilidad de transición de cierta posición de la mano a otra; el símbolo de salida como una postura especifica y la secuencia de los símbolos de salida como el gesto de la mano.
- Redes neuronales con retraso: Son una clase de redes neuronales artificiales que se enfocan en datos continuos, haciendo que el sistema sea adaptable para redes en linea y les da ventajas sobre aplicaciones en tiempo real.

1.7. Estado del arte

1.8. Organización de la tesis

Capítulo 2. Marco teórico

2.1. Detección rápida de objetos usando características simples utilizando el clasificador de cascada impulsada

El método desarrollado por (ref Viola Jones) consiste en detectar un objeto

2.1.0.4. Another title

El significado etimológico de fatiga proviene del latín *fatigare*; *fatim* que significa "con exceso", y *agere* que significa "hacer". Es típicamente definida como la reducción en la capacidad fisiológica de un tejido u órgano con manifestación física y/o psíquica generada por la demanda prolongada de **actividad física y/o mental**; respectivamente.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam sit amet lobortis turpis. Praesent auctor mi metus, sed bibendum ligula efficitur eu. Suspendisse ut ante id erat interdum accumsan. Pellentesque eget hendrerit eros, et ullamcorper elit. Proin a lacus et sem hendrerit efficitur. Praesent eget eros sed tellus dapibus bibendum sit amet vel justo. Maecenas finibus porttitor dictum. Fusce lacinia dictum interdum ??



(a) Mi metus, sed bibendum ligula efficitur eu.



(b) Mi metus, sed bibendum ligula efficitur eu.



(c) Mi metus, sed bibendum ligula efficitur eu.

Figura 1: Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam sit amet lobortis turpis. Praesent auctor mi metus.

Capítulo 3. Wow title

3.1. Stuff

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam sit amet lobortis turpis. Praesent auctor mi metus, sed bibendum ligula efficitur eu. Suspendisse ut ante id erat interdum accumsan. Pellentesque eget hendrerit eros, et ullamcorper elit. Proin a lacus et sem hendrerit efficitur. Praesent eget eros sed tellus dapibus bibendum sit amet vel justo. Maecenas finibus porttitor dictum. Fusce lacinia dictum interdum.

cell1 cell2 cell3
[a] cell4 cell5 cell6
cell7 cell8 cell9

[b]

cell1	cell2	cell3
cell4	cell5	cell6
cell7	cell8	cell9

[c]

Col1	Col2	Col2	Col3
1	6	87837	787
2	7	78	5415
3	545	778	7507
4	545	18744	7560
5	88	788	6344

[d]

The aligning options are m for middle, p for top and b for bottom.	cell2	cell3
cell1 dummy text dummy text dummy text	cell5	cell6
cell7	cell8	cell9

Country List			
Country Name or Area Name	ISO ALPHA 2 Code	ISO ALPHA 3 Code	ISO numeric Code
Afghanistan	AF	AFG	004
Aland Islands	AX	ALA	248
[e] Albania	AL	ALB	008
Algeria	DZ	DZA	012
American Samoa	AS	ASM	016
Andorra	AD	AND	020
Angola	AO	AGO	024

[f]	col1	col2	col3
	Multiple row	cell2	cell3
		cell5	cell6
		cell8	cell9

3.2. Testing tables

Nunc hendrerit justo vitae leo imperdiet, eu egestas nunc tristique. Etiam eget risus purus. Suspendisse sagittis tellus eu ipsum ultrices porttitor. Aliquam iaculis, metus sed ullamcorper blandit, justo nibh vehicula ipsum, vitae finibus diam orci vitae magna. Donec sit amet orci a dui laoreet euismod. Sed sed justo eget metus fermentum lacinia quis eget tellus. Pellentesque nibh metus, auctor id felis sed, lobortis condimentum urna. Nullam vel pharetra nisi. Sed volutpat nisi at efficitur blandit. Nulla interdum dictum dui, nec laoreet diam vulputate non. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Suspendisse non lobortis elit, vel bibendum tellus. Praesent gravida feugiat metus, non ultricies nunc mattis ut 1

3.3. Testing equations

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam sit amet lobortis turpis. Praesent auctor mi metus, sed bibendum ligula efficitur eu. Suspendisse ut ante id erat

Tabla 1: Mauris et imperdiet tortor. Maecenas consectetur lacus elit, dignissim eleifend dolor ornare ut. Aenean euismod porta nisi, et volutpat ex laoreet sit amet. Sed ac elit vestibulum neque ultrices feugiat

	FFS	SOFA	FQ	CIS20R	FACIT
1	TAF	TAF	PF	PF	PF
2	TAF	CM	CS	EE	PF
3	PF	TAF	CS	CM	EE

interdum accumsan. Pellentesque eget hendrerit eros, et ullamcorper elit. Proin a lacus et sem hendrerit efficitur. Praesent eget eros sed tellus dapibus bibendum sit amet vel justo. Maecenas finibus porttitor dictum. Fusce lacinia dictum interdum (Ecuación 1).

$$K = R + H + V$$

$$R = \text{consumo de oxígeno} \times kg^{-1} \times min^{-1} \quad (1)$$

$$H = \text{constante horizontal} \times \text{velocidad de desplazamiento}$$

$$V = \text{constante vertical} \times \text{velocidad de desplazamiento}$$

3.4. Diseño de un método para identificar fatiga física

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam sit amet lobortis turpis. Praesent auctor mi metus, sed bibendum ligula efficitur eu. Suspendisse ut ante id erat interdum accumsan. Pellentesque eget hendrerit eros, et ullamcorper elit. Proin a lacus et sem hendrerit efficitur. Praesent eget eros sed tellus dapibus bibendum sit amet vel justo. Maecenas finibus porttitor dictum. Fusce lacinia dictum interdum.

3.4.1. Extracción de parámetros

Ecuación 2.

$$A_d = -g - \frac{\sum F}{mass} \quad (2)$$

Donde A_d representan la aceleración que se aplica a un dispositivo, g la constante de gravedad de 9.81 m/s^2 , y $\sum F$ las fuerzas que se aplican al propio sensor.

3.4.2. Cálculo de la demanda de oxígeno

Text 3 and more O_2 text:

$$\begin{aligned}
 K &= R + H + V \\
 R &= 3.5 - (0.0367 \times BMI) - (0.0038 \times age) + (0.1790 \times gender) \\
 H &= 0.1 \times \text{velocidad de desplazamiento} \\
 V &= 1.8 \times \text{velocidad de desplazamiento}
 \end{aligned} \tag{3}$$

Donde 1 + 2 representan el consumo de O_2 en reposo personalizado al usuario ($ml \times kg^{-1} \times min^{-1}$) (?), H el componente horizontal relativo a la velocidad de desplazamiento (m/min), V el componente vertical relativo a la velocidad (m/min) y pendiente de desplazamiento (%).

Velocidad:

Para obtener la velocidad de desplazamiento se utiliza el número de pasos realizados por el usuario como se muestra a continuación (Ecuación 4).

$$S_k = D_k / W \quad D_k = ST_k \times SL \quad SL = D_{total} / ST_{total} \tag{4}$$

Capítulo 4. Very título

Sobre referencias. CICESE pide este formato (Adleman, 1998)

Capítulo 5. Other much title



Figura 2: Wow sdf.

Capítulo 6. Título

blablabla

Capítulo 7. Otro título

blah blah blah (?)

Capítulo 8. Conclusiones

Amazing conclusions!!

Lista de referencias

Adleman, L. M. (1998). Computing with DNA. *Scientific american*, **279**(8): 34–41.

Lakin, M. R., Youssef, S., Cardelli, L., y Phillips, A. (2012). Abstractions for DNA circuit design. *Journal of The Royal Society Interface*, **9**(68): 470–486.

Apéndice A. Apéndice

El apéndice...