

ESTIMACIÓN DE POSE DE LA CABEZA HUMANA

RAJIV GONZÁLEZ

Resumen

Just a test.¹

ÍNDICE

1	Resumen	1
2	Objetivos	1
2.1	objetivo general	1
2.2	objetivos específicos	1
3	Introducción	2
4	Marco teórico	2
5	Mapeo de coordenadas a ángulos	4
6	Ecuaciones	5
6.1	Desviación de ϕ_y y ϕ_x	6
7	Análisis del comportamiento de los ángulos ϕ_x y ϕ_y	7
8	Simulación processing	10
9	Dispositivo para adquirir datos	11
9.1	IMU	11
9.2	Microcontrolador	11
9.3	Xbee	12
9.4	Diseño de la tarjeta	12
9.5	tarjeta electrónica	13
10	Rotaciones en los pares de posiciones	13
10.1	Cálculo de todas las rotaciones posibles	15
	Referencias	15

RESUMEN

OBJETIVOS

objetivo general

Detectar el rostro de las personas y estimar su pose relativa a un plano virtual situado enfrente de ellas, asociando regiones en dicho plano con información de textura de los rostros detectados y su posición. Lo anterior debe ser capaz de funcionar en tiempo real.

objetivos específicos

- Generar una base de datos de rostros humanos con la información de pose asociada.
- Desarrollar el sistema para que funcione en un rango de distancia amplio entre la persona y la cámara.

¹ This is a footnote.

- Implementación de un algoritmo clasificador para asociar la región que se observa con poses determinadas.

INTRODUCCIÓN

Conocer la posición, orientación y movimiento de la cabeza son cuestiones de gran importancia ya que pueden ayudar a las personas a interactuar con las computadoras de una forma más natural a como se realiza actualmente, dando como resultado en útiles en aplicaciones como: video conferencias, controles o interfaces hombre-máquina especiales, aplicaciones de realidad virtual. Además, si se toma en cuenta hacia dónde está mirando la persona y el tiempo que lleva en cierta posición puede proporcionar información adicional que ayudaría a inferir si está muy atenta a lo que está observando o si la persona está durmiendo.

Por medio del análisis de imágenes capturadas a personas y en combinación con algoritmos de aprendizaje automático es posible conocer la posición y orientación de sus cabezas con respecto a un marco de referencia, la información obtenida puede ayudar a conocer qué es lo que están observando las personas.

MARCO TEÓRICO

En el área de visión computacional la estimación de la pose de una cabeza es el proceso de inferir la orientación y la posición de una cabeza humana a partir de las imágenes [Murphy et al, 2009]. El estimador debe ser robusto a distorsiones causadas por la cámara, a las expresiones faciales, al cabello, al vello facial y a objetos que ocluyan la cabeza como por ejemplo unos lentes o sombreros.

Existen algunos aspectos que se deben tomar en cuenta al realizar investigación en de este tema: se necesita previamente localizar la cabeza de las personas; la cabeza humana puede ser modelada como un objeto rígido sin tomar en cuenta el resto del cuerpo humano; la estimación de la pose de la cabeza se hace con respecto a un marco de referencia centrado en la cámara y de manera más general a un sistema global de coordenadas; para estimar la mirada de las personas con precisión en cualquier configuración un sistema seguidor de ojos debe ser complementado con el sistema de estimación de la pose de la cabeza. [Wang et al, 2009].

Actualmente se han realizado una gran diversidad de métodos que se pueden emplear para solucionar el problema de la estimación de la pose de la cabeza, razón por la cual motivó a Murphy-Chutorian y Manubhai a desarrollar una taxonomía sobre los enfoques más importantes, en ella se presenta la clasificación de los métodos tomando como principal parámetro de clasificación el enfoque fundamental sobre el cual subyace la implementación del método. Las clasificaciones son:

- *Métodos de apariencias de plantillas.*- Este método utiliza métricas de comparación basadas en imágenes para hacer correspondencias de la pose de una cabeza con un conjunto de ejemplos (plantillas) con las poses etiquetadas, hace la correspondencia con las más similares de las plantillas. El método no requiere entrenamiento con ejemplos negativos, solo requiere los ejemplos etiquetados.

Desventajas:

- (a) Solo estiman posiciones discretas
- (b) Se requiere saber dónde se localiza el rostro
- (c) Ineficiente con muchos ejemplos
- (d) Ineficiente con imágenes de la misma persona del entrenamiento y durante la prueba

- *Métodos detectores en forma de arreglo.*- Se entrenan múltiples detectores de rostro con diferentes poses discretas, la diferencia con el método anterior es que la imagen de entrada es evaluada con un detector entrenado con bastantes imágenes y un algoritmo supervisado.

Ventajas:

- (a) La detección y localización no son etapas diferentes
- (b) Son buenos con imágenes con alta y baja resolución

Desventajas:

- (a) Se requieren bastantes imágenes de entrenamiento
- (b) Es difícil implementar el arreglo con un número de detectores extenso en un sistema en tiempo real
- (c) Podrían darse ambigüedades en la clasificación (múltiples clasificaciones positivas).

- *Métodos de regresión no lineal.*- Estiman la pose mediante el aprendizaje de una función no lineal de mapeo desde un espacio de imágenes a una o más direcciones de pose. Con un conjunto de entrenamiento se puede construir un modelo que estime de forma discreta o continua una pose.

Desventajas:

- (a) No es claro cómo se utilizará la herramienta específica de la regresión
- (b) La estimación realizada es tosca (coarse) en las ubicaciones discretas
- (c) Son propensas al error

Las ventajas de utilizar este método con redes neuronales son:

- (a) Bastante rápidos
- (b) Solo requieren imágenes etiquetadas para entrenamiento
- (c) Arrojan los resultados más exactos en la práctica.

- *Métodos de variedad-embedida o manifold.*- Este método busca manifolds de baja dimensión que modelen la variación continua en la pose de la cabeza. Nuevas imágenes pueden ser incrustadas en estos manifolds y después utilizarlos para la estimación con técnicas como la regresión en el espacio embebido. Cualquier algoritmo de reducción de dimensionalidad puede ser considerado como un intento de un manifold embebido, pero la dificultad yace en crear un algoritmo que exitosamente recupere la pose de la cabeza mientras ignora otras fuentes de variación en la imagen.

Desventajas:

- (a) Al igual que con los detectores en forma de arreglo la estimación de la pose es tosca ya que la estimación se deriva de un conjunto discreto de mediciones.
- (b) La heterogeneidad en los datos de entrenamiento que es común en muchos escenarios de entrenamiento en el mundo real, representan un problema. Para contrarrestar lo anterior es necesario entrenar un manifold con múltiples personas, pero a menudo es imposible obtener un muestreo regular de poses de cada individuo.
- (c)

Ventajas:

- (a)

(b)

(c)

Así como existen muchas diferencias entre los métodos también hay unas características que prevalecen:

- Como métrica para conocer el desempeño de los métodos se utiliza el error de la media absoluta angular para pitch, roll y yaw.
- Mientras más poses se deseen estimar la clasificación se vuelve más difícil y propensa a cometer más errores.
- Un gran número de los enfoques en algún punto del proceso de estimación utilizan un algoritmo de aprendizaje automático, los más comunes son las redes neuronales [Brown et al, 2002], support vector machine [Huang et al, 1998], y los métodos de boosting como el floatboost [Zhang et al, 2006].
- Se logra mucha precisión en la estimación mediante la utilización de cámaras de visión estéreo.
- La variación de la pose está restringida a un rango que contiene todas las características visibles del rostro (que se utilizan para la clasificación) desde una vista frontal.

Como ya se ha mencionado anteriormente para estimar la mirada con precisión se necesita un sistema de detección y seguimiento de ojos además de múltiples cámaras y el sistema de estimación de pose de la cabeza, habiendo señalado lo anterior cabe destacar que el proyecto que se desarrollará más que estimar la mirada con precisión (además de la pose de la cabeza) representándola como un vector normal al iris, se pretende estimar una región en un plano virtual enfrente de ellas, la región indicaría lo que están observando las personas de la escena que tienen enfrente.

MAPEO DE COORDENADAS A ÁNGULOS

Uno de los problemas que hay que tomar en cuenta y definir en el presente trabajo de tesis antes de realizar los experimentos, es la relación que existe entre la ubicación en metros de lo que está mirando el sujeto de estudio en la pantalla y la pose de la cabeza de esta persona medida en ángulos, tomando en cuenta diversos aspectos como: la estatura de la persona, la distancia de la persona a la pantalla, cuál es el marco de referencia, etc. Esta sección consiste en describir cómo se realiza el mapeo (o la ecuación) de dicho objeto desplegado en pantalla y la orientación de la cabeza, primero se describirá la geometría del escenario.



Figura 5.1:

La pantalla que observa el sujeto del experimento se coloca en una pared a una distancia de L_y metros del piso, la pantalla tiene unas dimensiones de

W m. de ancho y L m. de alto. El sistema de coordenadas tiene su origen en la cámara y se ubica encima de la pantalla y a la mitad del ancho W , la recta que se encuentra a lo largo de W es paralela al eje X y la que está a lo largo de L es paralela a Y . Las personas se encontrarán paradas en un mismo plano (piso) y enfrente de la pantalla a una distancia D_z . Hay un punto en específico del rostro de las personas que es de interés y es el que se utiliza para realizar el análisis, este punto se define como P y se encuentra a la mitad de la línea que une los ojos. El vector \vec{v} sale del punto P y es ortogonal al plano de la cara, el punto P se encuentra a una distancia H_y con respecto al piso y es dependiente de la estatura de la persona. Otro parámetro que se debe mencionar es el que hace referencia a la distancia que hay entre la persona y el origen sobre el eje X , es decir, la componente X del punto P de la persona y lo denotaremos como a .

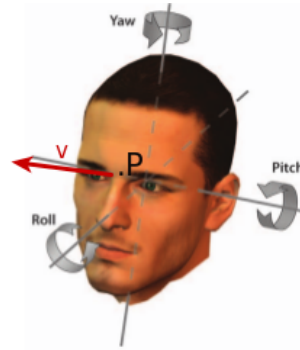


Figura 5.2:

Como se puede observar en la figura 1.2 las personas tenemos tres tipos de movimiento de la cabeza: yaw, pitch y roll; sin embargo para el análisis de este proyecto solo se toman en cuenta el movimiento de yaw y pitch, ya que estos son los movimientos que se dan cuando una persona mueve su cabeza para observar el objeto en dos dimensiones de la pantalla.

El objeto de la pantalla que la persona observa es representado como S y el vector \vec{v} representa la mirada de la persona o lo que es lo mismo \vec{PS} , cuando la persona se encuentre mirando S el vector \vec{v} debe apuntar a S , esto quiere decir que si se proyecta el vector sobre la misma dirección hacia la pantalla debe intersectar S .

El objeto S se desplazará a través de toda la pantalla (X, Y) metros, tomando como marco de referencia la esquina superior izquierda de la pantalla, a esta esquina le corresponde la coordenada $(0, 0)$, (X, Y) es la esquina inferior derecha y $(\frac{X}{2}, \frac{Y}{2})$ el centro de la pantalla. La pose de la cabeza se mide mediante dos ángulos: ϕ_x y ϕ_y .

- ϕ_y .- Observando desde el plano $Y - Z$ representa que tanto se desvía el vector \vec{v}
- ϕ_x .- Observando desde el plano $X - Z$ (desde arriba de la persona y la pantalla) representa que tanto se desvía el vector \vec{v}

De igual manera el ángulo ϕ_x indica el movimiento yaw que la persona realiza al mirar como se traslada el objeto S y el ángulo ϕ_y el movimiento pitch.

ECUACIONES

Una vez definido el escenario y los aspectos a tomar en cuenta, ahora falta demostrar cual es la relación que se puede hallar entre la posición del objeto en la pantalla y la pose de la cabeza mediante los ángulos ϕ_x y ϕ_y .

Para facilitar el análisis en esta sección se ilustra el escenario con todos los elementos en la figura 2.1

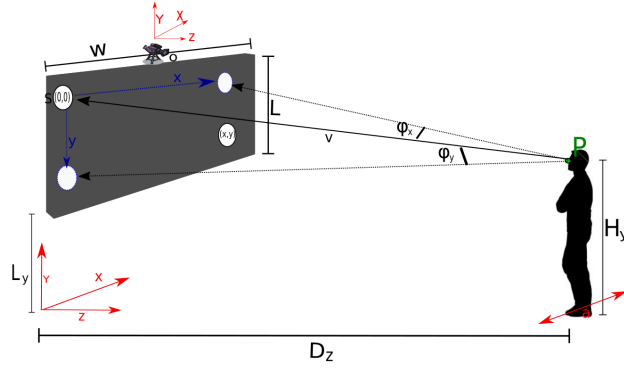


Figura 6.1: Escenario y aspectos a tomar en cuenta

El análisis comienza observando la escena desde el plano $Y - Z$, primero se definen dónde se encuentran con respecto al sistema de coordenadas los puntos más importantes, los cuales son: S y P . Supongamos que el objeto a observar está en la coordenada (x, y) de la pantalla, como se había mencionado previamente el sistema de coordenadas se encuentra centrado en la cámara, a la mitad de la pantalla con respecto al eje X y a una pequeña altura c_y m. con respecto al eje y , el movimiento de S yace sobre el plano $X - Y$ por lo que su componente en z es igual cero, entonces:

$$S = [S_x, S_y, S_z]^T = [x - W/2, -(c_y + y), 0]^T \quad (6.1)$$

Para hallar las componentes de P se toma la distancia D_z a la que se encuentra del origen sobre el eje Z , como se había mencionado la persona se desplaza a m sobre el eje X y se encuentra a una distancia sobre el eje y de $L_y + L - H_y$:

$$P = [P_x, P_y, P_z]^T = [a, -(L_y + L - H_y), D_z]^T \quad (6.2)$$

Ahora que ya se definieron donde se encuentran los puntos que se utilizarán con respecto al marco de referencia centrado en la cámara, se necesitan los ángulos entre la posición del vector $\vec{v} = \vec{PS}$ y cada uno de los ejes que son de interés, estos ángulos se conocen como cosenos directores del vector \vec{v} y únicamente son necesarios los que se forman con el eje Y y el X . Las fórmulas de los cosenos directores son:

$$\cos(\phi_y) = \frac{\vec{v}_y}{|\vec{v}|} \quad (6.3)$$

$$\cos(\phi_x) = \frac{\vec{v}_x}{|\vec{v}|} \quad (6.4)$$

donde $|\vec{v}|$ se halla mediante:

$$\sqrt{\vec{v}_x^2 + \vec{v}_y^2 + \vec{v}_z^2} = \sqrt{(x - \frac{W}{2} - a)^2 + (-c_y - y + (L_y + L - H_y))^2 + (-D_z)^2} \quad (6.5)$$

Por lo tanto los ángulos de desviación son:

$$\phi_y = \cos^{-1}\left(\frac{-c_y - y + (L_y + L - H_y)}{\sqrt{(x - \frac{W}{2} - a)^2 + (-c_y - y + (L_y + L - H_y))^2 + (-D_z)^2}}\right) \quad (6.6)$$

$$\phi_x = \cos^{-1}\left(\frac{x - \frac{W}{2} - a}{\sqrt{(x - \frac{W}{2} - a)^2 + (-c_y - y + (L_y + L - H_y))^2 + (-D_z)^2}}\right) \quad (6.7)$$

Durante la etapa de entrenamiento se requiere capturar bastantes imágenes de personas de frente a la pantalla, con el rostro en diferentes orientaciones debido a que está mirando el objeto en pantalla en diferentes posiciones, esto quiere que durante la etapa de captura de datos se debe variar la posición del objeto en pantalla que la persona del experimento está mirando y también se debe variar la posición de la persona en el lugar del experimento. Sea una instancia de captura de datos la captura de la imagen de una persona anotando su ubicación en el lugar del experimento, la posición del objeto y la orientación de la cabeza (ángulos ϕ_x y ϕ_y).

Se llega a dar una extensa cantidad de instancias de captura por cada persona variando los parámetros mencionados, lo que llevaría a bastante tiempo de captura de instancias por cada persona y sería muy inconveniente para ellas, por lo tanto se realizó un análisis antes de realizar los experimentos del comportamiento de los ángulos en diferentes circunstancias, el análisis fue realizado mediante el software matemático Octave y a continuación se discuten los resultados. Suponiendo que es una pantalla de 42 pulgadas y el marco de referencia se encuentra en la cámara encima de la pantalla.

- Ancho de la pantalla: $W = 0,9282m$
- Largo de la pantalla: $L = 0,5523m$
- Distancia de la pantalla al piso: $L_y = 1,5m$
- Distancia de la cámara a la pantalla: $C_y = 0,05m$
- Estatura del sujeto del experimento: $H_y = 1,6m$
- Movimiento de la persona a través del eje Z: de $0,1m$ a $3m$ con pasos de $0,51786m$
- Movimiento de la persona a través del eje X: de $-3m$ a $3m$ con pasos de $0,109m$
- Movimiento de la figura en pantalla a través del eje X: de $-\frac{W}{2}m$ a $\frac{W}{2}m$ con pasos de $0,01658m$
- Movimiento de la figura en pantalla a través del eje Y: de $-C_y$ a $-(C_y + L)m$ con pasos de $0,051786m$

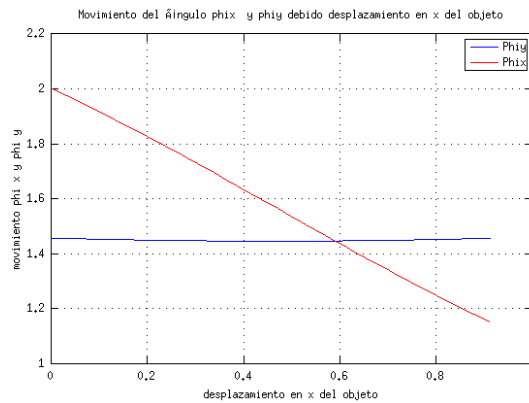


Figura 7.1:

En la gráfica de la figura 3.1 se puede observar la variación de los ángulos ϕ_x y ϕ_y con la persona en una única posición: $P = [0, -(L_y + L - H_y), 1]^T$, la variación de la figura en pantalla (en el eje x y y) de lo que observan es graficada con el marco de referencia centrado en la esquina superior izquierda de la pantalla, sin embargo los cálculos se hacen tomando la

posición de la cámara como marco de referencia. El mayor cambio que ocurre en los ángulos es con respecto a ϕ_x de 2 a 1.1 radianes que son 0.9 radianes o 51.5662 grados, lo cual parece tener mucho sentido ya que el desplazamiento se hace en el eje de ϕ_x .

Lo que resulta peculiar es que a pesar de que no hay desplazamiento en el eje y de la figura si hay una ligera variación en el ángulo ϕ_y de la mirada de la persona, esto se debe al punto de fuga. El punto de fuga es el lugar geométrico en el cual las proyecciones de las rectas paralelas a una dirección dada en el espacio, no paralelas al plano de proyección, convergen, lo anterior se puede ver ilustrado en la figura 3.2, el punto de fuga es el lugar donde convergen todas líneas "paralelas" de color verde, y la línea del horizonte es la recta horizontal de color azul. Cuando la figura inicialmente se encuentra en

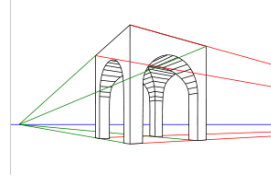


Figura 7.2:

una posición superior a la de los ojos y se mueve hacia uno de los extremos de la pantalla, ésta pareciera moverse hacia abajo (hacia el horizonte) y cuando inicialmente la figura se encuentra en una posición inferior a la de los ojos de la persona y se mueve hacia un extremo, ésta pareciera moverse hacia arriba por lo tanto la mirada de las personas en el eje y tiende a moverse.

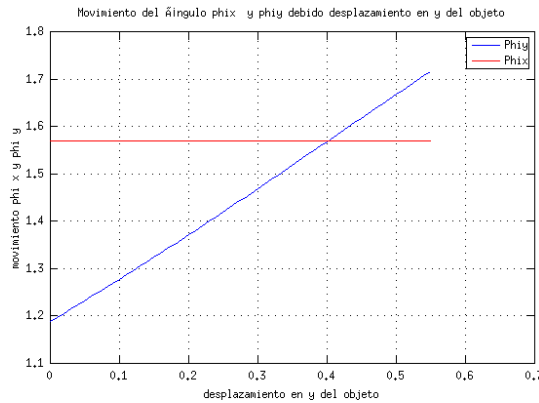


Figura 7.3:

En la gráfica de la figura 3.3 se puede apreciar que el objeto que observan las personas en pantalla únicamente se desplaza en el eje y , en este caso el único desplazamiento de la mirada es con respecto al ángulo ϕ_y , al desplazarse el objeto de $0.5523m$ la mirada varía de 1.2 a 1.7 radianes, es decir, 0.5 radianes o lo que es lo mismo 28.6479 grados.

La gráfica 3.4 describe el caso en el únicamente la persona se mueve sobre el eje X de -3 a $3m$, el mayor desplazamiento lo tiene el ángulo ϕ_x , de esta gráfica se puede concluir que la mayor variación en el ángulo se da entre -1 y $1m$ la cual la mirada varía de 0.8 a 2.36 radianes, si la persona se mueve más allá de esta distancia la mirada tiende a estabilizarse y no valdría la pena hacer experimentos en estas zonas. Lo anterior se ve con mayor claridad en la figura 3.5, aquí se amplió el rango en el que se mueve la persona sobre el eje X y en consecuencia ϕ_x tiende a estabilizarse.

El último análisis corresponde al caso cuando la figura en pantalla no se mueve y la persona únicamente se mueve sobre el eje Z . La figura que observa se encuentra en el centro de la pantalla, es decir en: $[S_x, S_y, S_z]^T = [\frac{W}{2}, -(c_y + \frac{L}{2}), 0]^T$ y la persona se encuentra en $[P_x, P_y, P_z]^T = [0, -(L_y + L - H_y), vecZ]^T$ donde $vecZ$ es un vector que va desde 0.1 a $3m$.

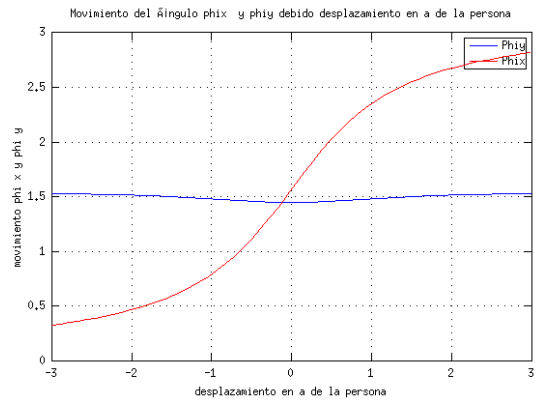


Figura 7.4:

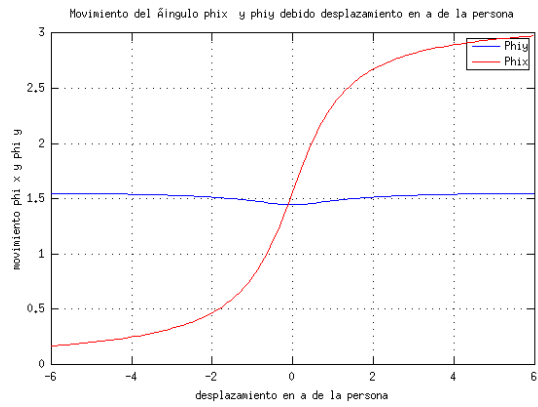


Figura 7.5:

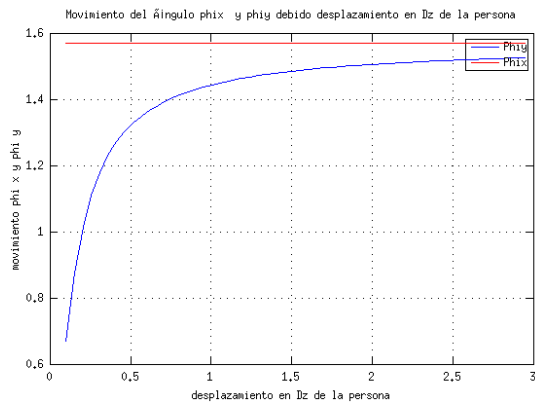


Figura 7.6:

En la gráfica 3.6 se puede apreciar que mientras se va alejando la persona el único ángulo de la mirada que varía es ϕ_y , va de 0,6 a 1,55 radianes, esto quiere decir que hay una variación en el ángulo de 0.95 radianes o 51.56 grados hacia abajo cuando la persona se aleja, además se puede observar que a partir de 1.5m la variación que hay es muy pequeña, de 0 a 0.8m es cuando se da la mayor variación. Lo anterior se puede explicar con la línea de fuga, mientras la persona se aleja el objeto pareciera moverse hacia abajo y establecer con la mirada de la persona en 1.6 radianes o 91 grados, en el horizonte.

SIMULACIÓN PROCESSING

Se realizó una simulación mediante el software processing del análisis anterior, esto se hizo para poder observar el comportamiento de los ángulos en diferentes posiciones de las personas al mismo tiempo y variar en tiempo real y de manera interactiva la posición en el eje X y Y de la figura en pantalla. Los parámetros del escenario son los mismos a los utilizados en la análisis anterior: W , L , LY , Cy y Hy .

En la figura 4.1 se puede apreciar una figura de la simulación, en ella se despliegan 36 posiciones de personas separadas por 1m en el eje X y 0.5m en el eje Y. La variación en el eje Y se da por medio de la barra deslizante que se encuentra en la esquina inferior derecha, va de 0 a 0.5523m y el desplazamiento en el Y se logra haciendo click en la figura (óvalo blanco) en pantalla que se encuentra en la parte superior de la ventana de la simulación y sin soltar el mouse se arrastra la figura de izquierda a derecha. Los ángulos ϕ_x y ϕ_y además de ser mostrados en cada una de las posiciones son representados mediante el radio del círculo que representa cada persona y la línea que sale de dicho ángulo, ϕ_x es la línea y ϕ_y es el radio.

Mediante la simulación de igual manera se demostró como al arrastrar

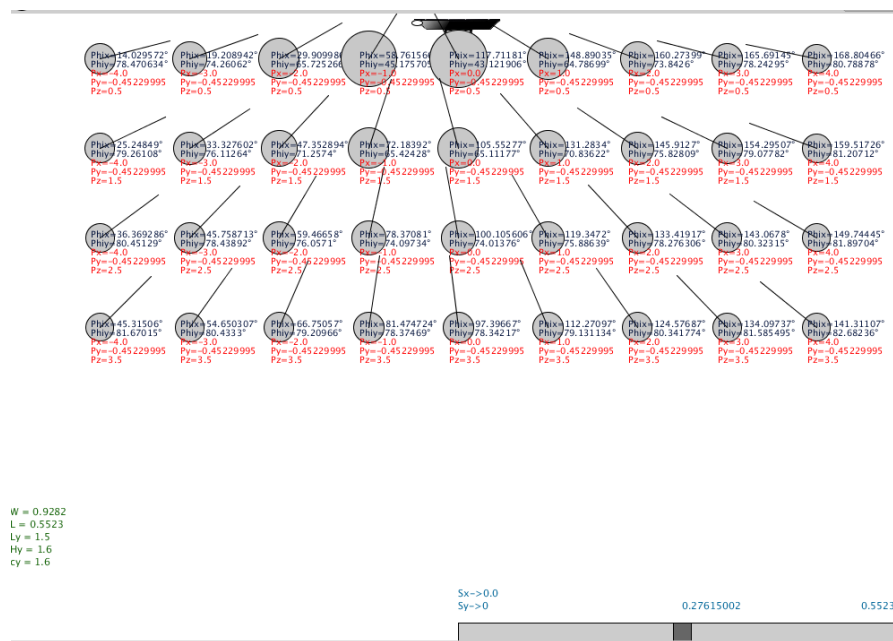


Figura 8.1:

la figura únicamente en X no solo varía ϕ_x sino también ϕ_y de todas las instancias, esto se puede corroborar en los valores de los ángulos y en el área del círculo.

Durante la etapa de adquisición de datos para el algoritmo de entrenamiento, además de conocer la ubicación de la persona y la de la figura que observan en pantalla es necesario capturar una imagen del rostro de la persona y conocer (mediante un aparato de medición) la orientación de la cabeza, lo anterior es necesario para evaluar y comparar el sistema de estimación de pose.

Como señalan en [Murphy-Chutorian, 2009] el método que se utilice para obtener las medidas de la orientación de la cabeza en el conjunto de entrenamiento debe ser preciso y entre los métodos más eficientes proponen los sistemas captura ópticos de movimiento y los sensores inerciales, sin embargo, los primeros son muy costosos, por lo que en el desarrollo del presente trabajo de tesis se optó por utilizar un sensor inercial, el cual consiste de acelerómetros y giroscopios a menudo acoplados con algún tipo de filtro para reducir ruido.

En este apartado se describirá el proceso de diseño y fabricación del dispositivo que realiza la medición de la orientación.

De modo que la tarjeta electrónica debe ir en la cabeza de las personas, durante el diseño se tomó mucho en cuenta el tamaño y el peso que tendría, ya que de ser muy grande y pesado podría ser muy incomodo para las personas durante los experimentos.

IMU

El imu seleccionado para este proyecto es el BNO055 de Bosch (figura 5.1). El sensor utiliza algoritmos que mezclan los datos del acelerómetro, magnetómetro y giroscopio en una salida estable de la orientación de los tres ejes, el sensor es capaz de arrojar los datos en cuaterniones, vectores y ángulos de Euler. El BNO055 se alimenta con voltaje de 3.3 a 5v, mide 2.67x2.032cm y para obtener los datos utiliza comunicación I2C.

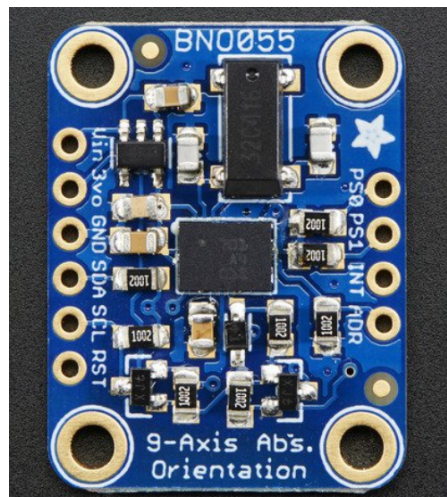


Figura 9.1:

Microcontrolador

Debido a las restricciones de peso y tamaño se utilizó un arduino micro, el cual es la versión más pequeña del arduino con 4.8x1.8cm y esto se debe a que es una versión más limitada en cuanto a periféricos y módulos, sin embargo, es adecuada para el uso que se le da en el proyecto, únicamente se utilizará para obtener los datos del imu mediante comunicación I2C y enviarlos a la computadora por comunicación serial. El arduino se alimenta

mínimo con 7v para que su regulador interno regule a 5v el microcontrolador que usa.



Figura 9.2:

Xbee

Durante los experimentos las personas se estarán moviendo en diferentes posiciones de la escena y mirando diferentes lugares en la pantalla, como se había mencionado la tarjeta electronica se colocará en su cabeza, por lo que utilizar un cable (de comunicación serial) entre la computadora y la tarjeta para obtener los datos de la pose sería muy inadecuado ya que éste debe ser bastante largo, le pesaría a la tarjeta y afectaría su posición en la cabeza, y finalmente podría afectar la experiencia de la persona el tener un cable muy extenso cerca de ella. Tomando en cuenta las consideraciones anterior se decidió trabajar con módulos xbee.

Los xbee son módulos inalámbricos creados para la comunicación inalámbrica entre ellos, su finalidad es la eliminación de cables en la comunicación serial. Para este proyecto se utilizaron los xbee s1 (figura 5.3) los cuales son los más pequeños, de bajo consumo y simples de utilizar debido a que su configuración punto a punto es bastante sencilla y únicamente se requiere declarar un xbee como emisor y el otro como receptor. Los xbee s1 se alimentan con 3.3v.



Figura 9.3:

Diseño de la tarjeta

Para alimentar todo el circuito es necesario al menos 7v, ya que como se había mencionado eso necesita el Arduino micro y es el dispositivo que requiere mayor voltaje, el imu funciona con 5v por lo que se le puede conectar el pin de 5v que tiene el Arduino, sin embargo aún se tiene el inconveniente de que el xbee y su comunicación funcionan con 3.3v, y las baterías recargables de iones de litio como la que se pretende utilizar por su reducido tamaño y peso, arrojan solo 3.7v.

Para solucionar el problema de los 7v se utilizó la bomba de carga TPS61093, figura 5.4, en una configuración para elevar el voltaje a 7v. La comunicación entre el arduino y el xbee (diferentes niveles de voltaje) se logró mediante

el cambiador de nivel TXS0102, y finalmente para obtener 3.3v de la batería se utilizó el regulador TPS73633. Se pudo alimentar el xbee a través del pin de 3.3v del arduino, sin embargo, el xbee consume hasta 60mA lo cual es demasiado para el regulador de 3.3v del Arduino y podría afectar su funcionamiento.



Figura 9.4:

En la imagen 5.5 se puede ver el esquemático de la tarjeta y en la 5.6 el pcb. Adicionalmente se creo una pequeña tarjeta para la recepción de

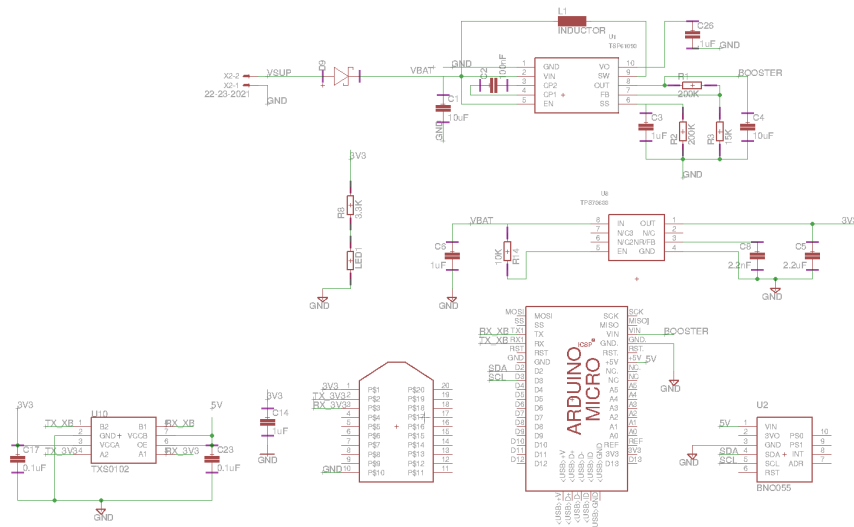


Figura 9.5:

datos a través del otro xbee y envío a la computadora mediante un cable con convertidor de rs232 a usb. Más adelante se detallará el programa para la recepción de datos.

tarjeta electrónica

En las siguientes imágenes se presentan: la tarjeta electrónica con el imu, arduino y xbee ya fabricada; la batería recargable de 3.7v y la tarjeta que conecta el xbee receptor de datos con la computadora para poder enviar los datos del imu hacia la computadora.

ROTACIONES EN LOS PARES DE POSICIONES

Se eligió como forma de representación de la pose de la cabeza los cuaterniones, los cuales son obtenidos mediante el IMU, como se había mencionado el dispositivo se colocará en la cabeza de las personas durante el entrenamiento para la generación del "ground truth", sin embargo, aún falta considerar los siguientes detalles:

- ¿Cuáles posiciones en el piso con respecto al marco de referencia centrado en la cámara se utilizaron para capturar la información(imagen rostro y pose en cuaterniones)?.
- ¿A qué distancia estarán las posiciones entre si y hasta que distancia máxima se tomarán datos, es decir, cual es el mínimo en X y Z que se debe seleccionar? ya que como se observó en la sección de *Análisis del*

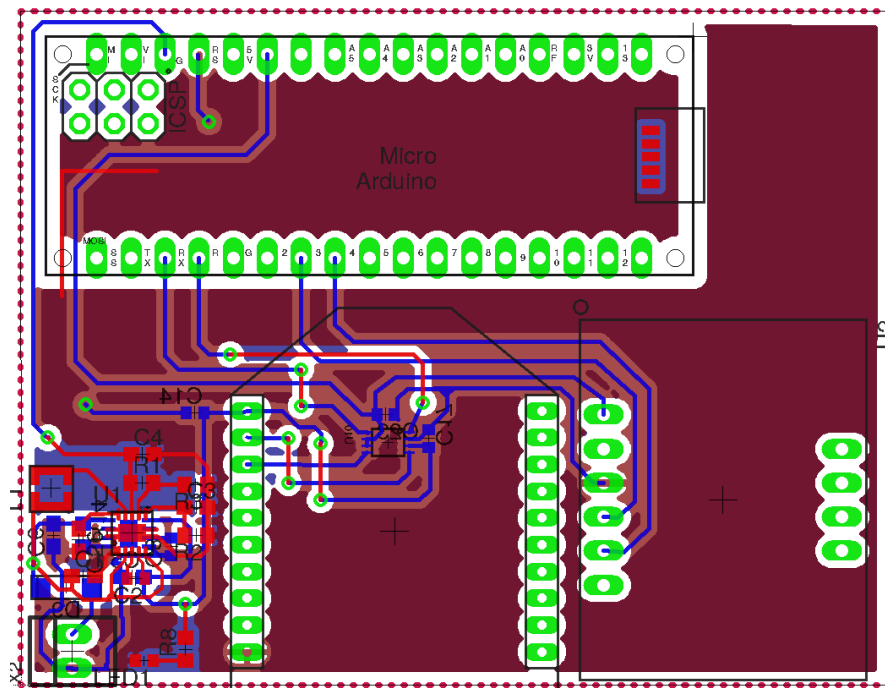


Figura 9.6:

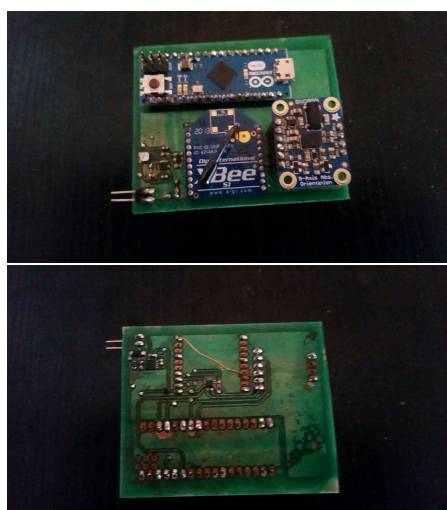


Figura 9.7: tarjeta con el IMU

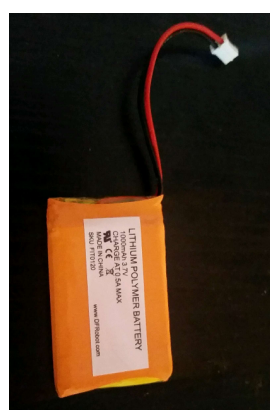


Figura 9.8: batería recargable

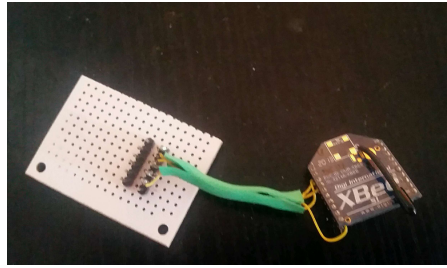


Figura 9.9: tarjeta receptora de datos

comportamiento de los ángulos ϕ_x y ϕ_y dependiendo de los parámetros del experimento hay valores en los cuales ya no vale la pena hacer análisis ya que las variaciones en los ángulos es muy pequeño o inexistente.

- ¿Cuáles posiciones para la figura en pantalla se deben utilizar?.
- ¿Cuántas posiciones de figura en pantalla y personas en el piso se considerarán?
- ¿Cómo lidiar con las pequeñas variaciones en las mediciones del IMU que se den en diferentes personas en los mismas instancias de experimentación?

Cálculo de todas las rotaciones posibles