

# DETECCIÓN DEL ROSTRO HUMANO, ESTIMACIÓN DE SU POSE Y DE SU MIRADA

## Avances de Tesis

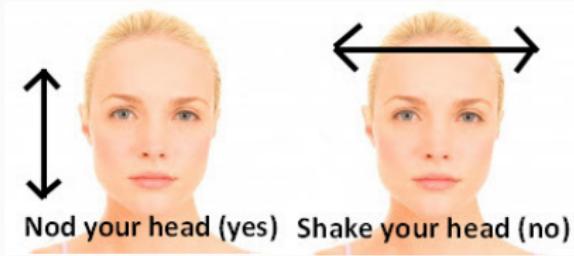
---

Rajiv González

31 de mayo de 2017

Maestría en Ciencias de la Computación, UADY

- ¿Por qué una persona mueve su cabeza?
- Importancia de conocer la posición y orientación
- Mediante imágenes capturadas de personas es posible conocer la posición y orientación de su cabeza.



Detectar el rostro de las personas y estimar su pose mediante una cámara monocular con el objetivo de inferir la región que estén mirando en un plano virtual enfrente de ellas.

# ¿Qué es la estimación de la pose de una cabeza?

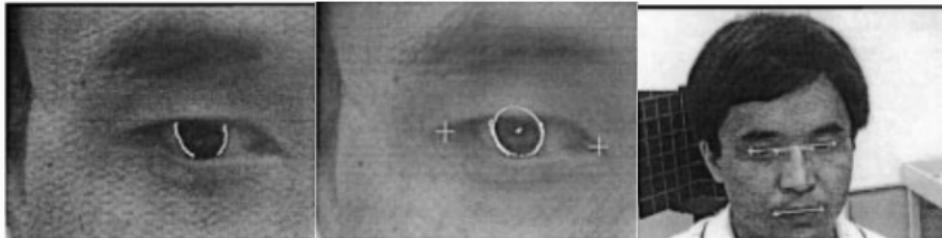
## Visión computacional

Es el proceso de inferir la orientación y la posición de una cabeza humana a partir de imágenes [Murphy et al, 2009].



## Aspectos a tomar en cuenta

- Previamente se debió localizar el rostro.
- La cabeza humana puede ser modelada como un objeto rígido.
- Marco de referencia centrado en la cámara.
- Para estimar la mirada con precisión se debe complementar con un sistema seguidor de ojos, [Wang et al, 2009].
- Un gran número de enfoques en algún punto del proceso utilizan aprendizaje automático.

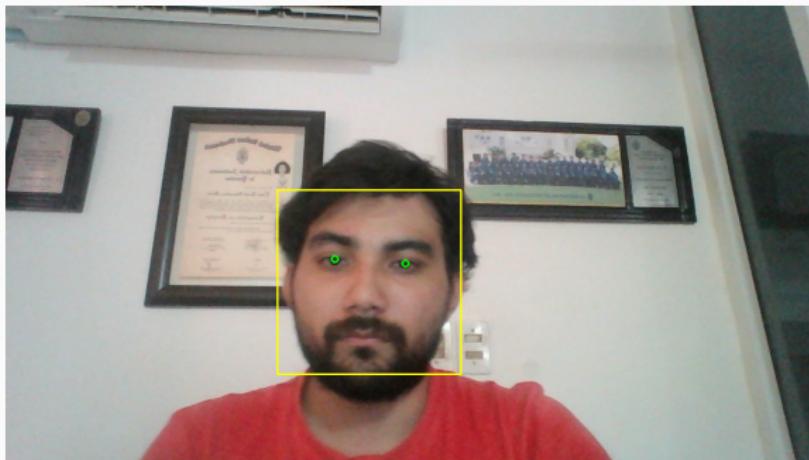


# Etapas del desarrollo del proyecto de tesis

- Localizar el rostro en la imagen
- Geometría del lugar de experimentación
  - Relación entre la ubicación de la cabeza de la persona y la región que está observando
  - Graficación de posibles casos
- Diseño y fabricación de tarjeta de adquisición de datos
- Algoritmo generador de mejores posiciones (puntos 3d) en la escena
- Proyección de puntos en el plano del piso
  - Descomposición de la matriz  $H$  para hallar la  $R$  y  $T$ , y el plano
  - Optimización con Levenberg-Marquardt
  - Reproyección de puntos e intersección con el plano
- Captura de datos
- Red Neuronal
- Experimentos y resultados

# Localizar rostros en la imagen

OpenBR

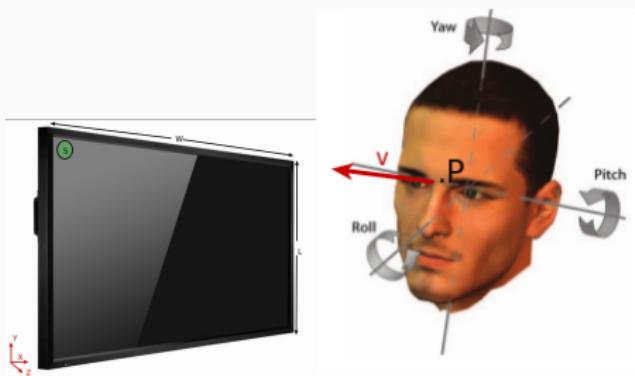


# Geometría del lugar de experimentación

Relación entre la ubicación de la cabeza de la persona y la región que está observando

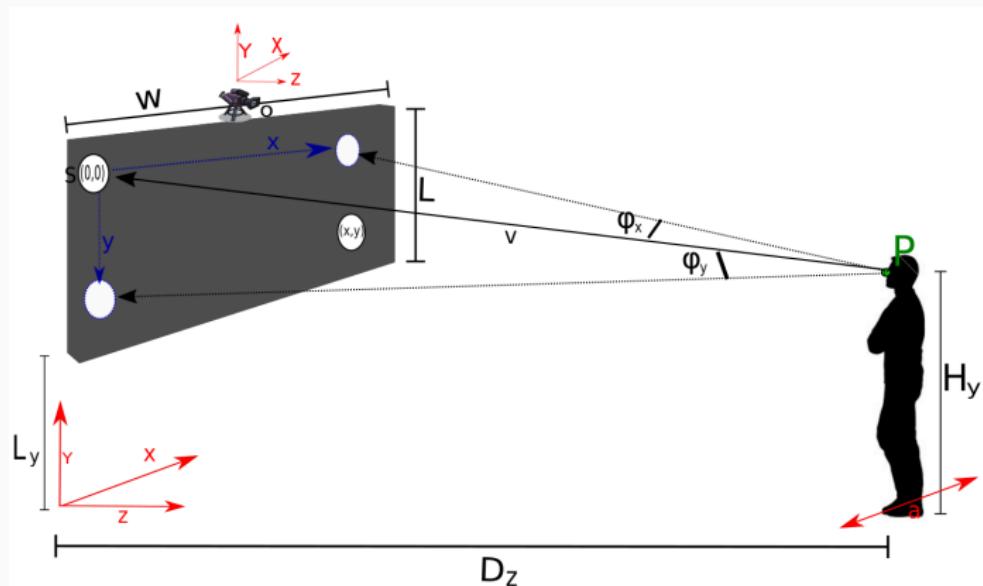
**Encontrar la posición de la persona y de lo que mira en pantalla con respecto a la cámara. Parámetros a considerar:**

- Estatura de la persona
- Dimensiones de la pantalla donde se desplegará la figura
- Distancia de la cámara a la pantalla (eje Y)
- Distancia de la pantalla(eje Y) al plano del piso



# Geometría del lugar de experimentación

Relación entre la ubicación de la cabeza de la persona y la región que está observando

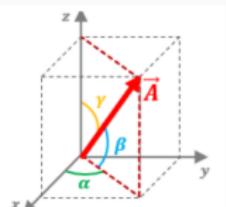


$$S = [S_x, S_y, S_z]^T = [x - W/2, -(c_y + y), 0]^T \quad (1)$$

$$P = [P_x, P_y, P_z]^T = [a, -(L_y + L - H_y), D_z]^T \quad (2)$$

# Geometría del lugar de experimentación

- Ecuaciones que describen el desplazamiento de la figura en pantalla y la mirada de la persona:  $\phi_x$  y  $\phi_y$
- Cosenos directores



$$\cos(\phi_y) = \frac{\vec{v}_y}{|\vec{v}|} \quad (3)$$

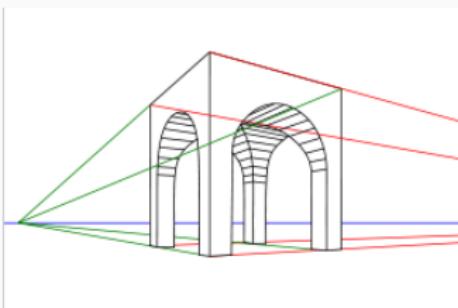
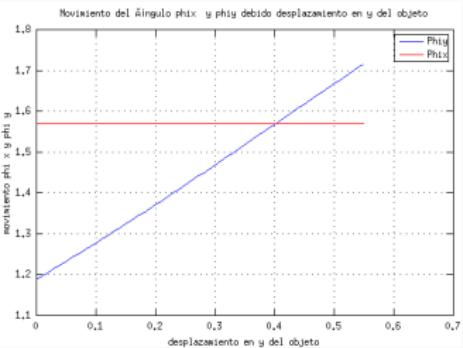
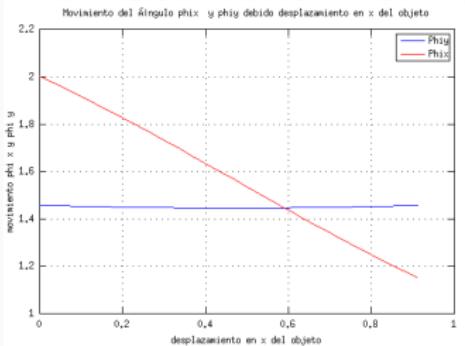
$$\cos(\phi_x) = \frac{\vec{v}_x}{|\vec{v}|} \quad (4)$$

$$\phi_y = \cos^{-1}\left(\frac{-c_y - y + (L_y + L - H_y)}{\sqrt{(x - \frac{w}{2} - a)^2 + (-c_y - y + (L_y + L - H_y))^2 + (-D_z)^2}}\right) \quad (5)$$

$$\phi_x = \cos^{-1}\left(\frac{x - \frac{w}{2} - a}{\sqrt{(x - \frac{w}{2} - a)^2 + (-c_y - y + (L_y + L - H_y))^2 + (-D_z)^2}}\right) \quad (6)$$

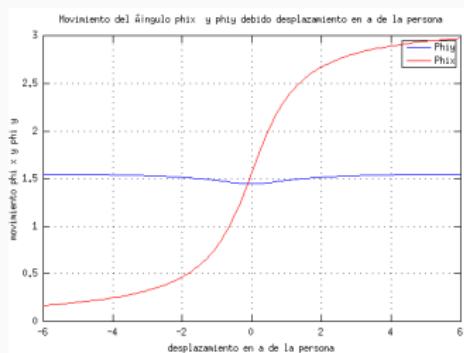
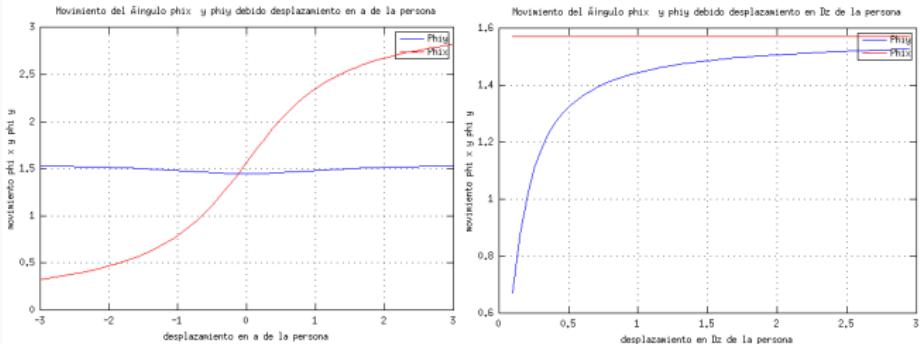
# Geometría del lugar de experimentación

## Graficación de posibles casos

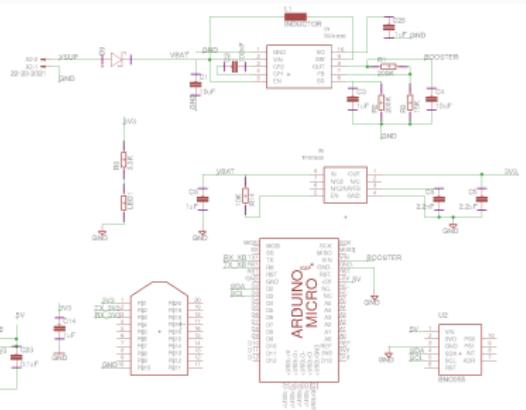
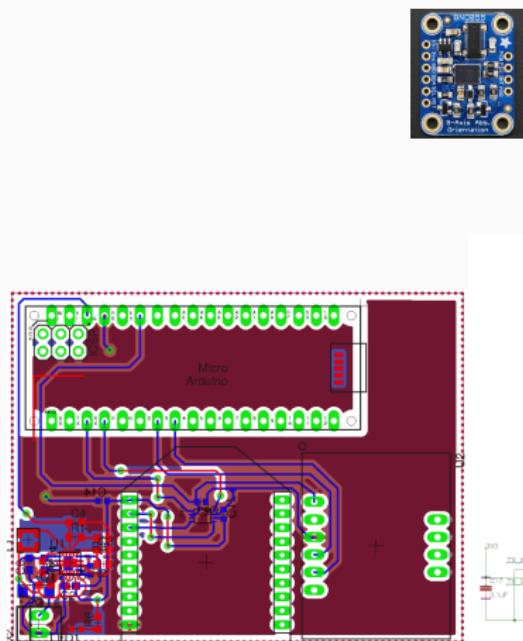


# Geometría del lugar de experimentación

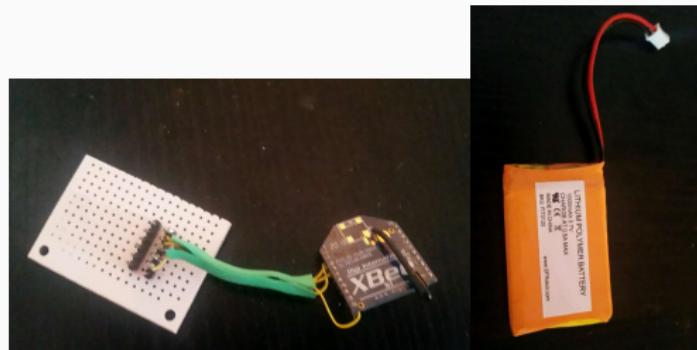
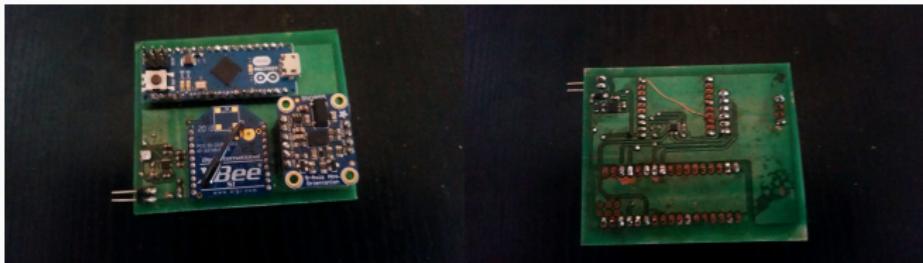
## Graficación de posibles casos



# Tarjeta de adquisición de datos

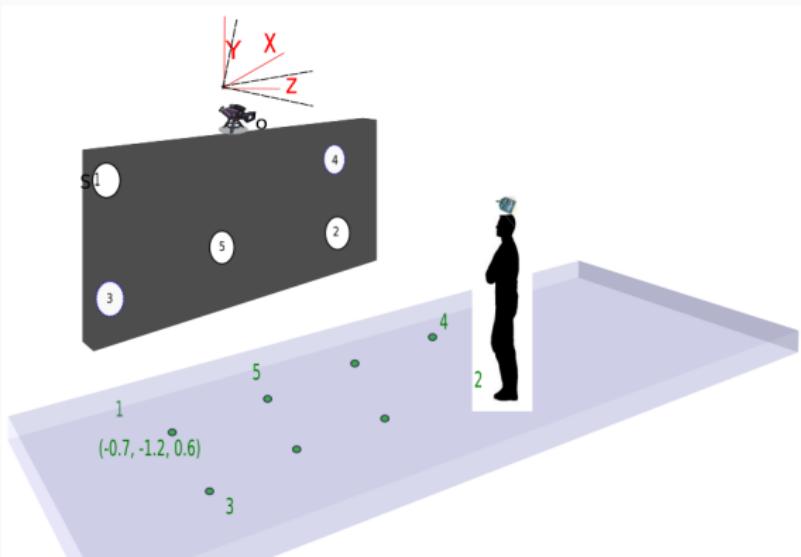


# Tarjeta de adquisición de datos



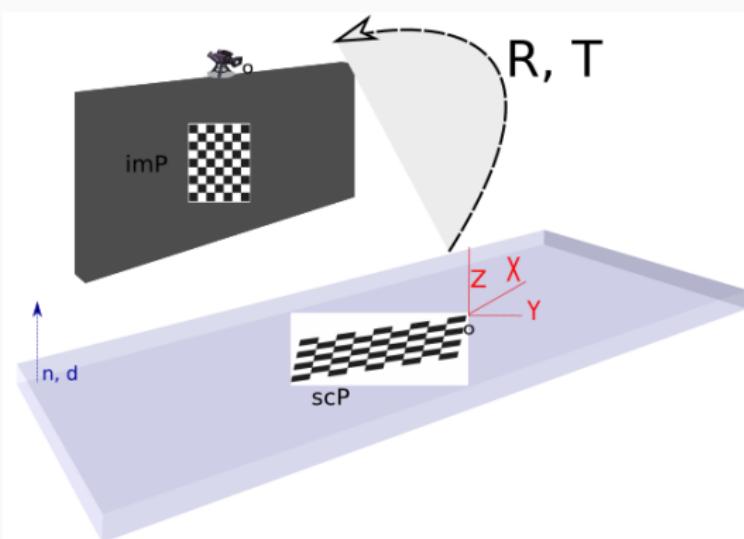
# Algoritmo generador de mejores posiciones

- Se toman en cuenta todos los parámetros: W, P, S, L, H...
- A cada posición de la figura en pantalla le corresponden dos posiciones en el plano del piso(dos poses de la cabeza)
- El algoritmo genera las posiciones en pantalla y en el plano del piso más significativos
- El marco de referencia está rotado!



# Proyección de puntos en el plano del piso

El primer paso es conocer dónde se encuentra la cámara con respecto al plano del piso.



$$scP = \begin{bmatrix} 0 & 0 & 0.12... \\ 0 & 0.12 & 0.12... \\ 0 & 0 & 0... \\ 1 & 1 & 1... \end{bmatrix}$$

### Homografía

Es toda transformación proyectiva que determina una correspondencia entre dos figuras geométricas planas

A partir de la matriz  $H$  y los dos conjuntos de puntos de los planos es posible hallar la matriz de rotación.

$$scP_{3 \times 48} = H_{3 \times 3} * imP_{3 \times 48} \quad (7)$$

# Proyección de puntos en el plano del piso

Descomposición de la homografía

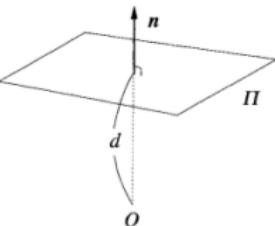


¿ $R$  y  $T$  son correctas?

# Proyección de puntos en el plano del piso

Plano del piso

Conociendo la  $R$  y  $T$  se calcula la ecuación del plano.



$$n = R * \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

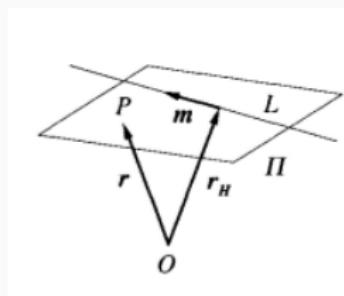
$$scPRot_1 = R * scP_1 + T \quad (8)$$

$$d = \langle N, scPRot_1 \rangle \quad (9)$$

# Proyección de puntos en el plano del piso

Intersección de rectas en el plano

Verificación de  $R$  y  $T$  mediante la intersección las rectas de los puntos de la imagen proyectados con el plano del piso

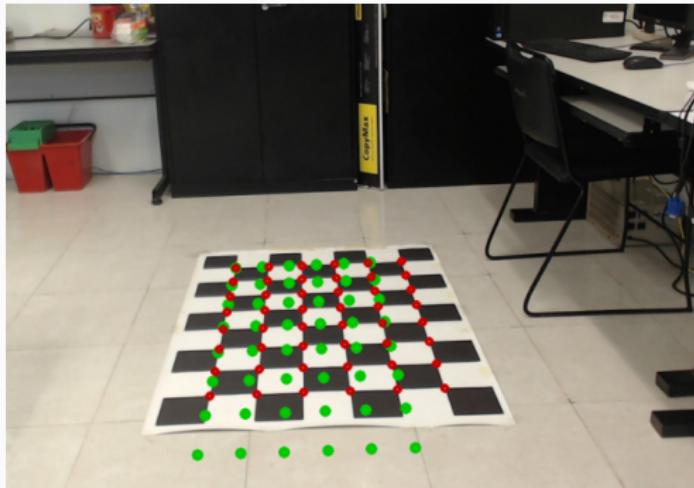


$$r = r_H + \frac{d - \langle n_{\Pi}, r_H \rangle}{\langle n_{\Pi}, m \rangle} m \quad (10)$$

# Proyección de puntos en el plano del piso

Intersección de rectas en el plano

La rotación y traslación calculadas presentan bastante error



Error de 26cm en X, 2.68 en Y y 12.8 en Z

- Algoritmo Levenberg–Marquardt
- Minimización de:

$$F(p) = \frac{1}{2} \sum_1^m (f_i(p))^2 \quad (11)$$

donde  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i = 1, \dots, m$  son las funciones dadas y  $m \geq n$ .

- Las funciones a optimizar deben tener los parámetros de  $R$  y  $T$

$$f_{xyz}(p) = \|X_{rt} - X_{intersec}\| \quad (12)$$

$$f_{xyz}(p) = \{R * X_{scn} + T\} - \{\frac{d}{\langle n_{\Pi}, X_{img} \rangle} X_{img}\} \quad (13)$$

En problemas de optimización numérica la redundancia de las matrices de rotación es incoveniente y a menudo es preferible una representación mínima.

## Fórmula de Rodrigues

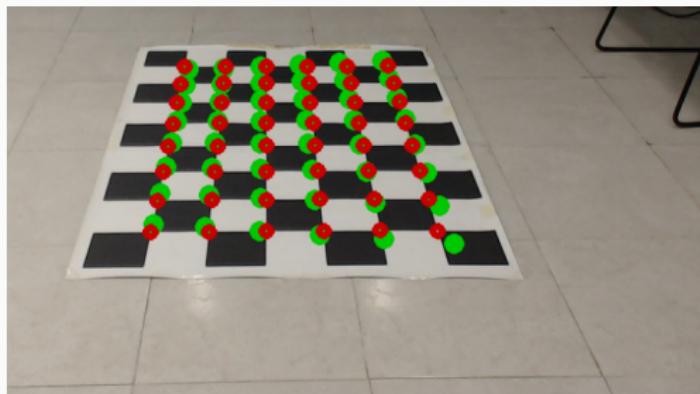
Cualquier matriz de rotación se puede realizar mediante la rotación alrededor de un eje fijo por un cierto ángulo  $\|\omega\|$ :

$$R = I + \frac{\hat{\omega}}{\|\omega\|} \sin(\|\omega\|) + \frac{\hat{\omega}^2}{\|\omega\|^2} (1 - \cos(\|\omega\|)) \quad (14)$$

$$p = [\omega_x, \omega_y, \omega_z, \|\omega\|, T_x, T_y, T_z] \quad (15)$$

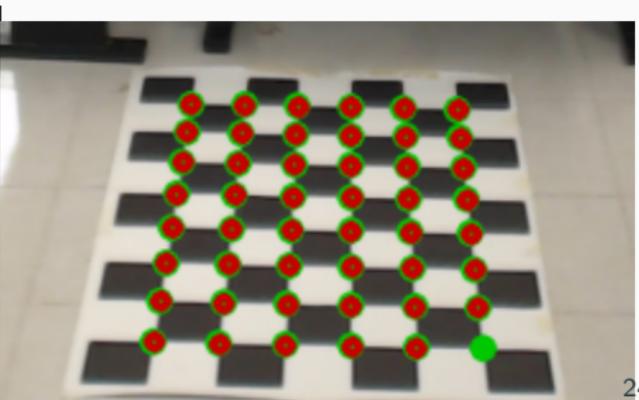
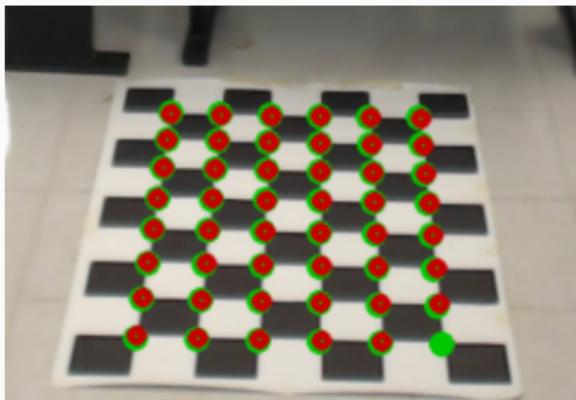
# Optimización numérica de R y T

## Resultados

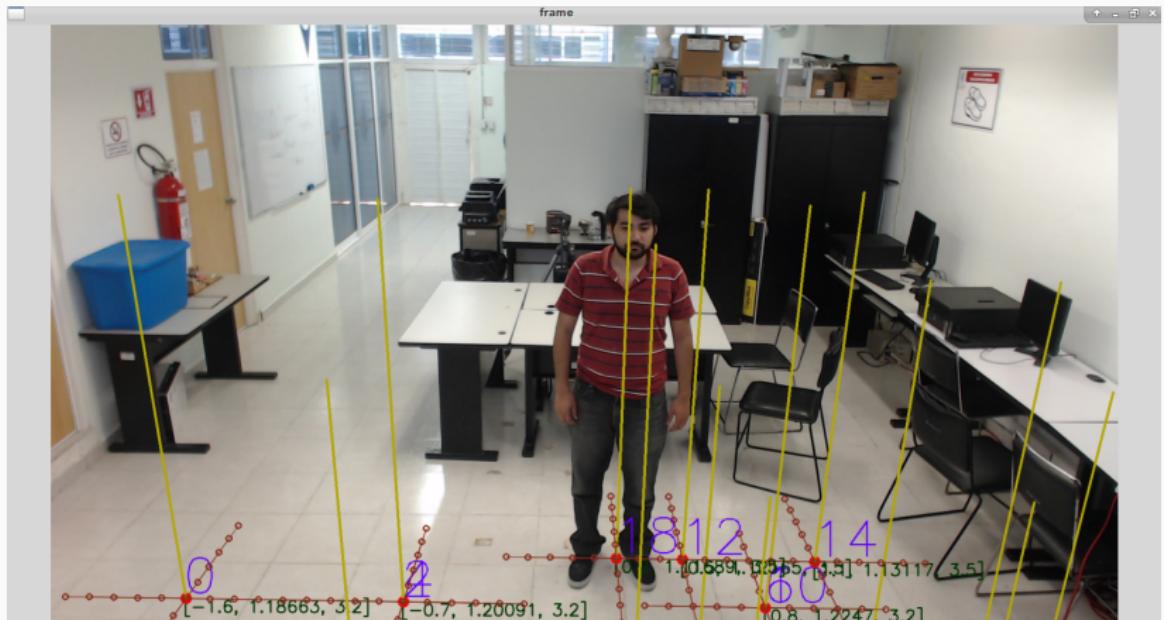


# Optimización numérica de R y T

Recalcando matriz de calibración y corrigiendo la imagen



# Proyección de marcas en el plano del piso



- Localizar el rostro en la imagen
- Geometría del lugar de experimentación
- Diseño y fabricación de tarjeta de adquisición de datos
- Algoritmo generador de mejores posiciones en la escena
- Proyección de puntos en el plano
  - Optimización numérica
- **Captura de datos**
- Red neuronal
  - Implementación, entrenamiento y resultados

-  Murphy-Chutorian, E., & Trivedi, M. M.  
**Head pose estimation in computer vision: A survey.**  
*Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(4), 607-626, 2009
-  Viola, P., and Jones, M.  
**Robust Real-Time Face Detection.**  
*International Journal of Computer Vision*, 57, 137-154, 2004.
-  Wang, J. G., & Sung, E.  
**Study on eye gaze estimation.**  
*Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 32(3), 332-350, 2002.
-  Papageorgiou, C. P., Oren, M., & Poggio, T.  
**A general framework for object detection.**  
*In Computer vision, 1998. sixth international conference on*, (pp. 555-562). IEEE., 1998
-  Kalal, Z., Mikolajczyk, K., & Matas, J.

*¿Preguntas?*