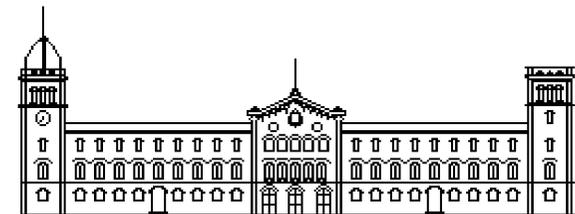


Proyecto Final de Carrera:

RGBD Análisis para el Reconocimiento Automático de acciones con Kinect.

Autor: Gabriel Domínguez Cisternes.
Directores: Sergio Escalera y Miguel Reyes.





Índice

- Contexto y Motivación.
- Antecedentes.
- Propuesta.
 - Auto-calibración.
 - Esqueleto.
 - DTW.
- Validación.
- Conclusiones.
- Trabajo futuro.



Contexto y Motivación

- ¿Qué queremos hacer?
 - Automatizar la tarea del reconocimiento de acciones.
- ¿Por qué?
 - La tendencia a automatizar tareas esta en alza.
 - Facilitará el trabajo en algunos campos.
- ¿Cómo?
 - Con un dispositivo capaz de extraer información del movimiento de una persona.
 - Y con una aplicación capaz de analizar la información extraída con el dispositivo.

Contexto y Motivación

- Kinect: Es un dispositivo capaz de medir la distancia que hay entre la cámara y el entorno, generando lo que se conoce como mapas de profundidad.



Dispositivo Kinect.



Ejemplo de mapa de profundidad.

Antecedentes

¿Qué hay hecho?
(sin sensor de profundidad).

■ Hasta el momento existen dos formas:

Sensores.

Obtienen la información del sensor colocado en el cuerpo del modelo.

Visión artificial.

Obtienen la información del cuerpo del modelo utilizando técnicas de procesamiento de imágenes.



Posteriormente se aplican algoritmos para el reconocimiento (Dynamic Time Warping, Hidden Markov Models, etc.).

Antecedentes

¿Qué hay hecho?
(con sensor de profundidad).

- Juegos de XBOX 360.
- Aplicaciones de traslación de movimiento persona/avatar persona/robot.



- Otras...

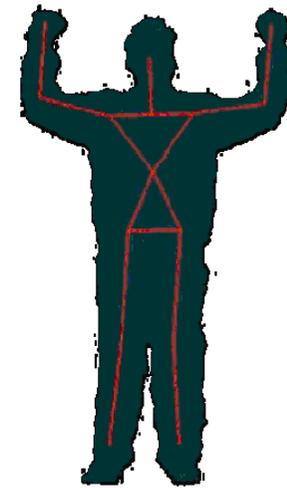
Propuesta

auto-calibración.

Automatización del proceso de calibración clásica.

■ Calibración clásica.

- Detección del cuerpo.
- Reconocimiento de la postura Psi (ψ).
- Seguimiento de la persona.



Denominaremos auto-calibración si evitamos el segundo punto de la calibración clásica.

Propuesta

auto-calibración.

- Se consigue utilizando los métodos de la clase XnSkeletonCaptability:

```
SaveCalibrationDataToFile(userID, nombreFichero);  
LoadCalibrationDataToFile(userID, nombreFichero);
```

- **Salvar:**

- Calibración clásica
- Salvar en fichero.

- **Cargar:**

- Detección de la persona
- Cargar un fichero.
- Seguimiento de la persona.

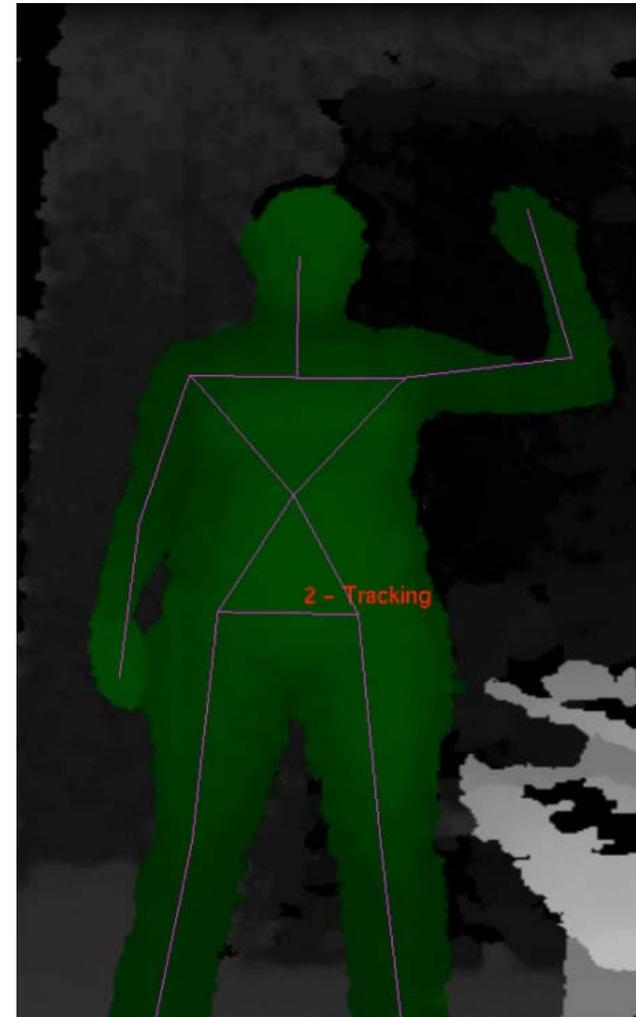
Permitirá hacer el seguimiento de una persona en videos donde no exista la calibración clásica.

Propuesta

- Una vez aplicada una calibración obtenemos un esqueleto compuesto por 15 partes:

- Cabeza
- Cuello
- Hombro Izq/Der
- Codo Izq/Der
- Mano Izq/Der
- Torso (CM)
- Cadera Izq/Der
- Rodilla Izq/Der
- Pie Izq/Der.

esqueleto.





Propuesta

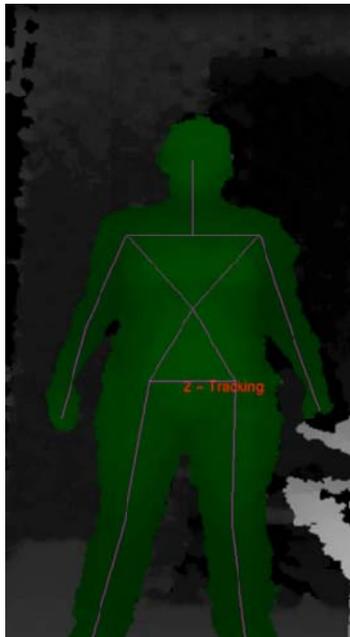
esqueleto.

- Para poder comparar poses en diferentes lugares y con modelos diferentes, necesitamos normalizar los puntos obtenidos.
 - Referenciar todos los puntos respecto a uno (cuello).
 - Escalar a una medida determinada para que todos tengan el mismo tamaño.

Propuesta

esqueleto.

- Una acción es un conjunto de poses ordenadas procesadas como el apartado anterior:



Pose 1

...



Pose N/2

...



Pose N

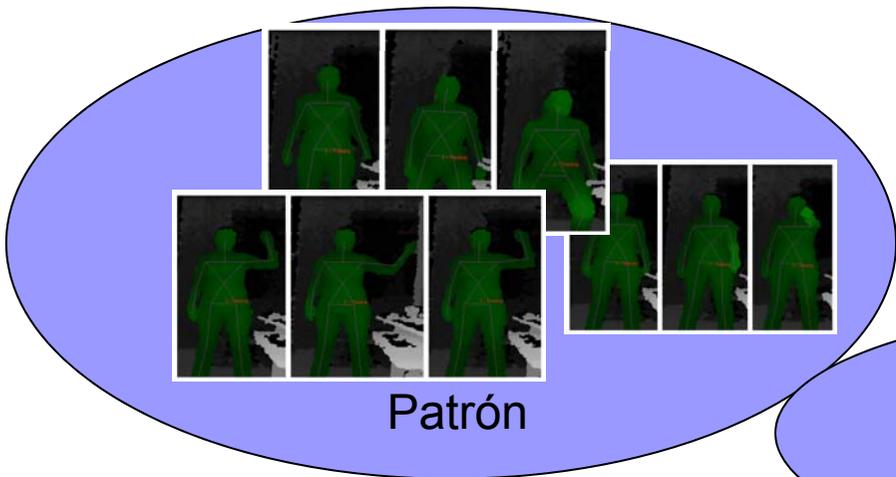
Propuesta

DTW.

■ Dynamic Time Warping:

□ Algoritmo de detección de patrones, con tres parámetros:

- Patrón. (¿Qué buscamos?)
- Flujo de datos de entrada. (¿Dónde buscamos?)
- Valor de Corte. (¿Cuándo se acepta como válido?)



Propuesta

DTW.

		DatoEntrada1	DatoEntrada2	DatoEntrada3	DatoEntrada4
Dato1Patron	0	0	0	0	0
Dato2Patron	∞	a_{11}	a_{12}	a_{13}	a_{14}
...	∞
DatoNPatron	∞	a_{n1}	a_{n2}	a_{n3}	a_{n4}

```
inc DTWDistance(char s[1..n], char v[1..m]) {
  declare inc DTW[0..n, 0..m]
  declare inc i, j, cost

  for i := 1 to m
    DTW[0, i] := infinity
  for i := 1 to n
    DTW[i, 0] := infinity
  DTW[0, 0] := 0

  for i := 1 to n
    for j := 1 to m
      cost := d(s[i], v[j])
      DTW[i, j] := cost + minimum(DTW[i-1, j], // insertion
                                  DTW[i, j-1], // deletion
                                  DTW[i-1, j-1]) // match

  return DTW[n, m]
}
```

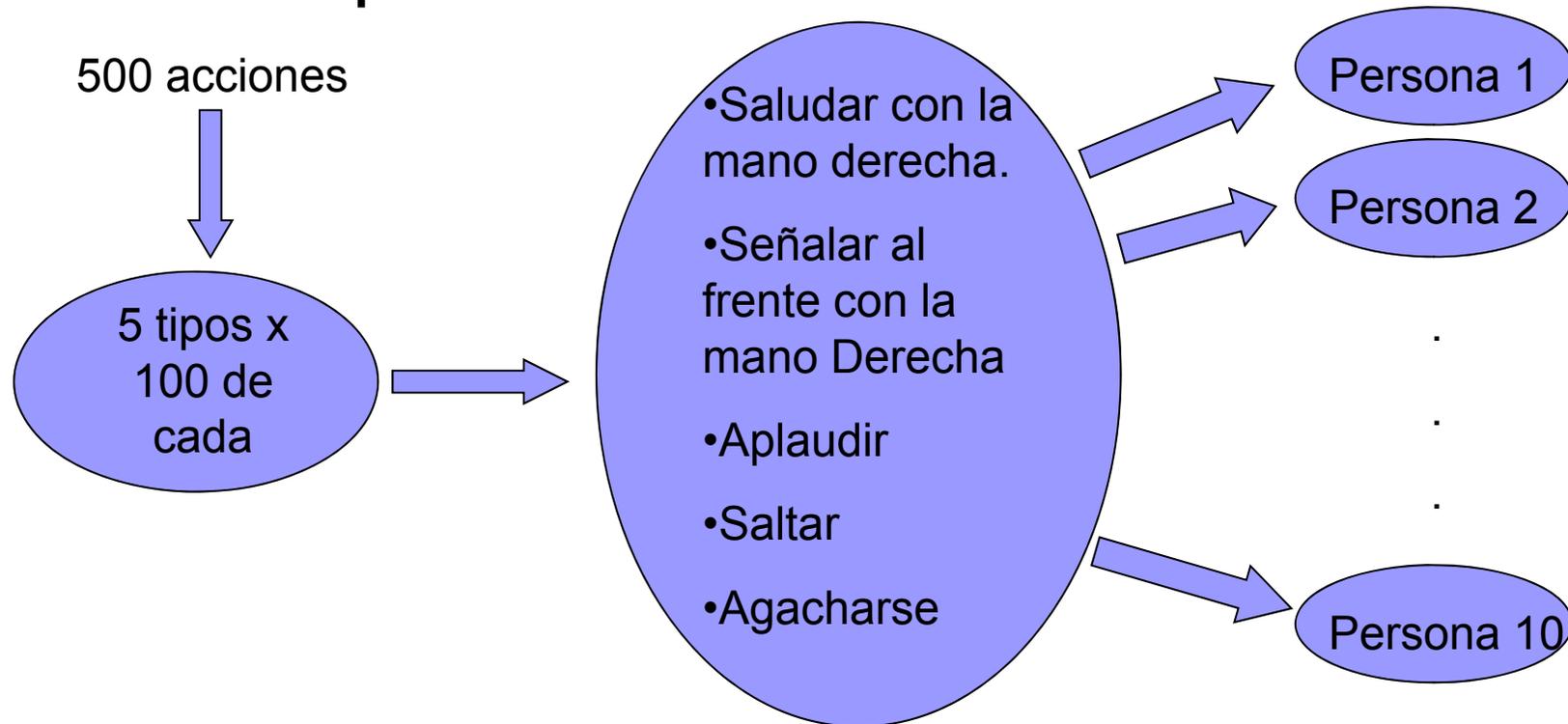
$$d_{22} = \sqrt{\sum (Dato2Patron_i - DatoEntrada2_i)^2}$$

$$a_{22} = d_{22} + \text{Min}(a_{11}, a_{12}, a_{21})$$

Validación

Creación de una BD.

- Para la validación se diseña una base de datos que contiene:





Validación

Creación de una BD.

- Se capturan las acciones:
 - Partimos de la aplicación de muestra NiUserTracker.
 - Añadimos el código necesario para extraer la información.
 - Guardamos la información en ficheros de texto para después poder procesarla.



Validación

Parámetro de corte.

- No se puede definir de entrada.
- ¿Cómo se calcula? Aplicando DTW.
 - Se define el patrón (acción).
 - Fijamos el valor de corte a 1.
 - Como datos de entrada, tenemos las 500 acciones de la base de datos.
 - Contamos en que casos la acción es aceptada.
 - Calculamos la efectividad del parámetro como:

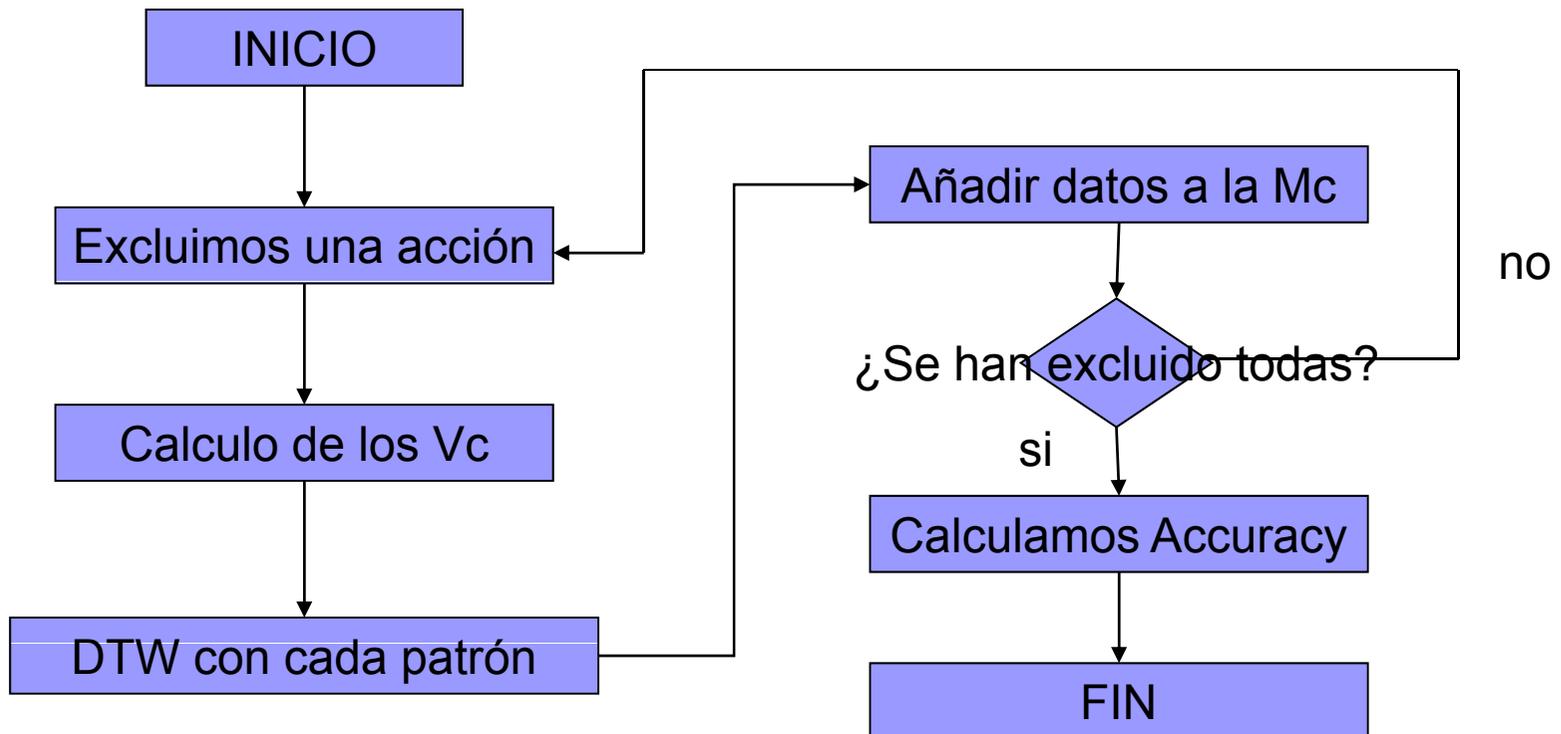
$$efectividad = \frac{accionesBienClasificadas}{accionesAccion + accionesMalClasificadas}$$

- Incrementamos el valor de corte y repetimos el proceso.
- Finalmente nos quedamos con el valor que nos de una efectividad más elevada.

Validación

Resultados.

- Generamos una matriz de confusión utilizando el método “leave-one-out”. Que consiste en dejar una de las acciones fuera del cálculo y luego aplicar los resultados a la acción excluida.



Validación

Resultados.

- Resultada al aplicar el procedimiento anterior a nuestra base de datos.

$$\text{MatrizConfusion} = \begin{pmatrix} 99 & 0 & 0 & 0 & 0 \\ 0 & 86 & 0 & 15 & 0 \\ 0 & 5 & 95 & 6 & 0 \\ 0 & 0 & 0 & 13 & 0 \\ 0 & 0 & 0 & 7 & 57 \end{pmatrix}$$

$$\text{efectividadClasificacion} = \frac{99+86+95+13+57}{99+86+95+13+57+15+5+6+7} = 0,91$$

$$\text{efectividadReal} = \frac{99+86+95+13+57}{500+15+5+6+7} = 0,66$$

Validación

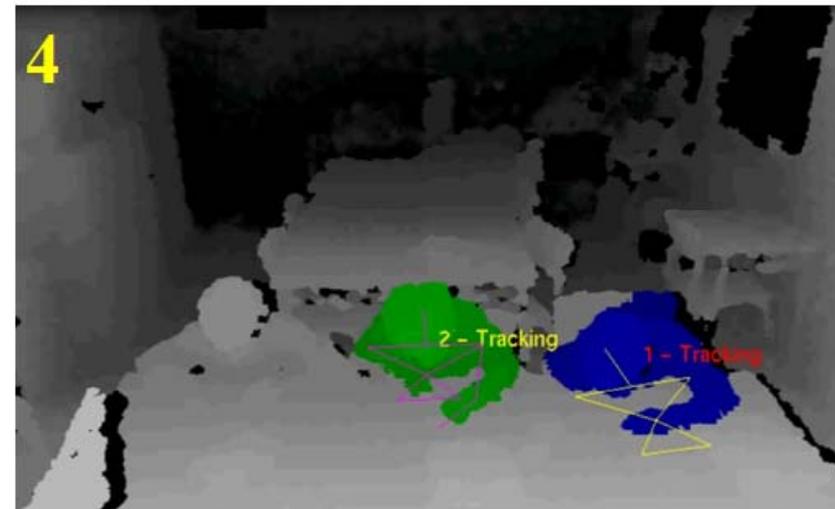
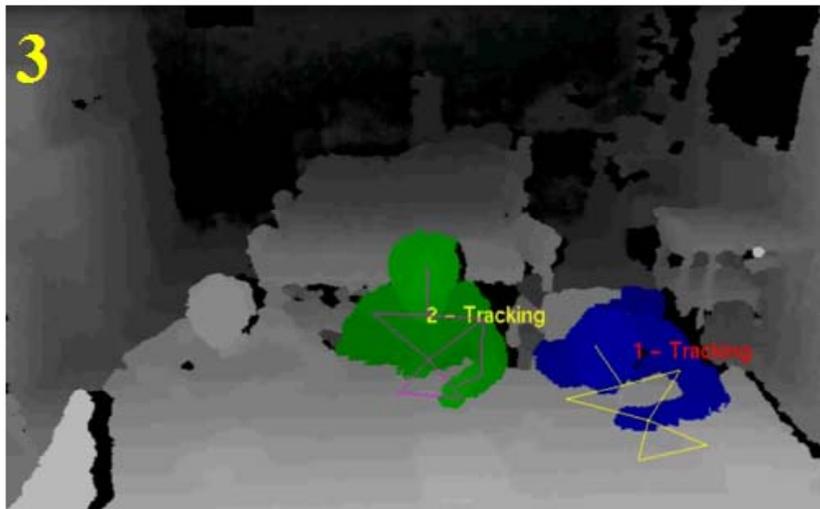
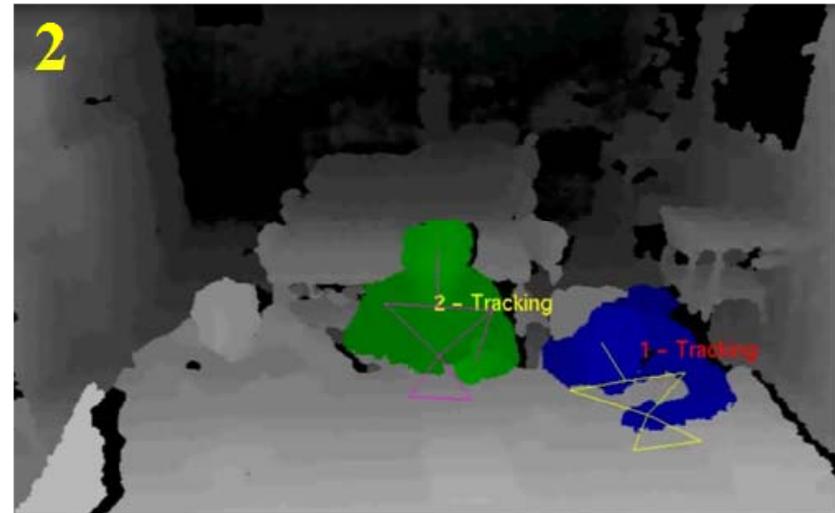
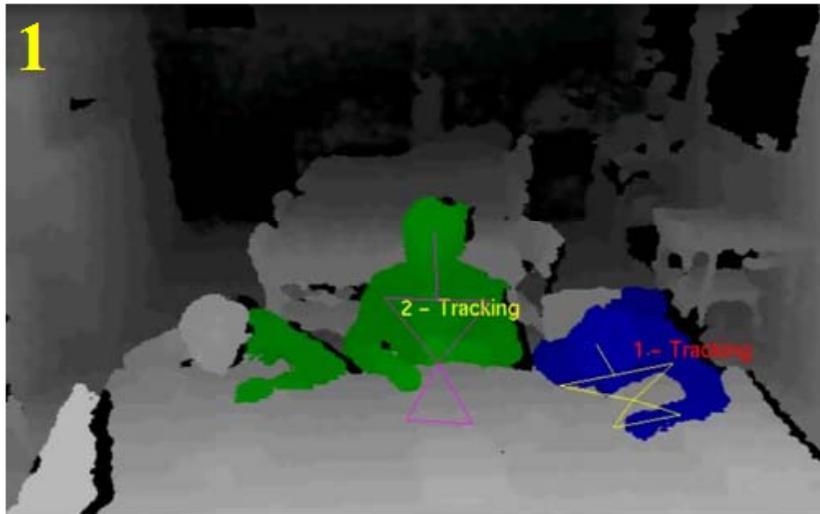
Resultados.

- Aplicación en un caso real (niños con trastorno de déficit de atención por hiperactividad TDAH).
- En este caso definimos un nuevo gesto (agacharse sobre la mesa) y lo buscamos en los videos.



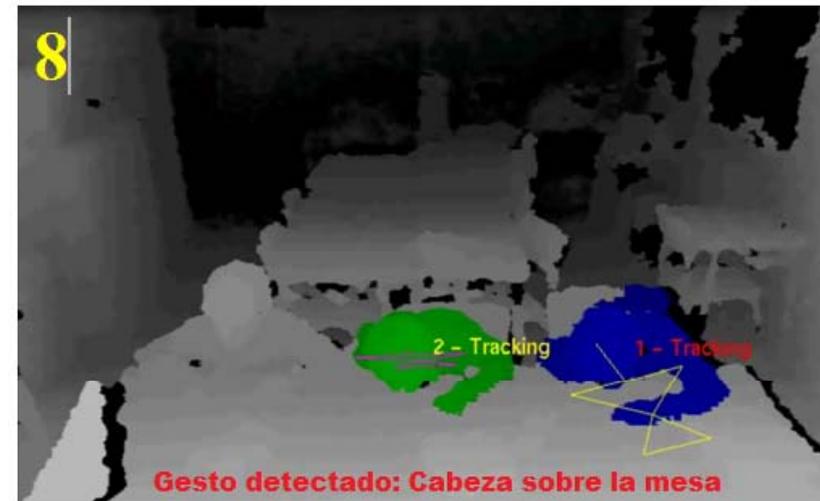
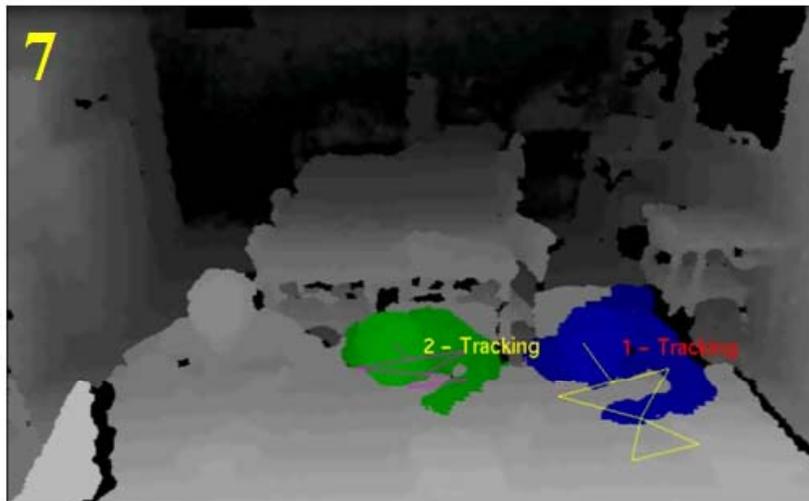
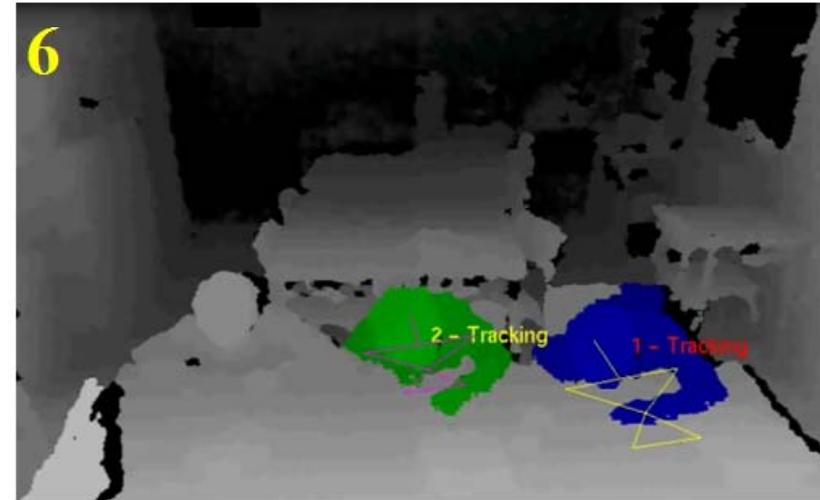
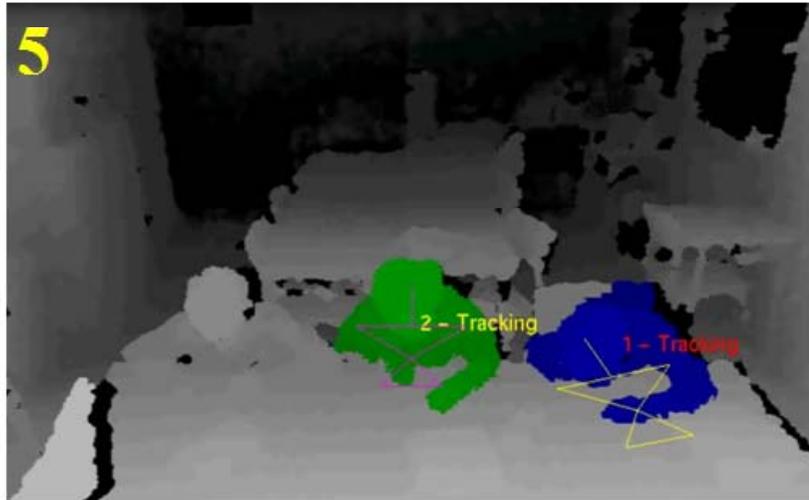
Validación

Resultados.



Validación

Resultados.



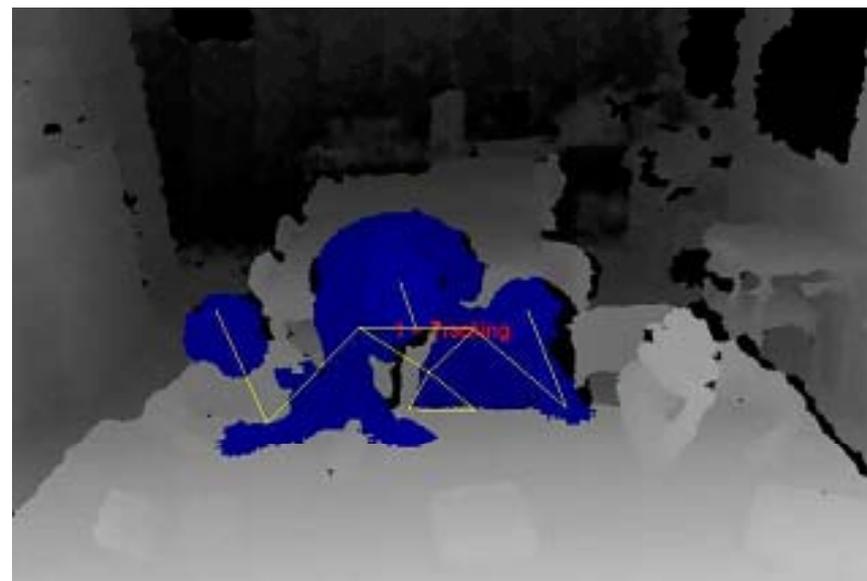
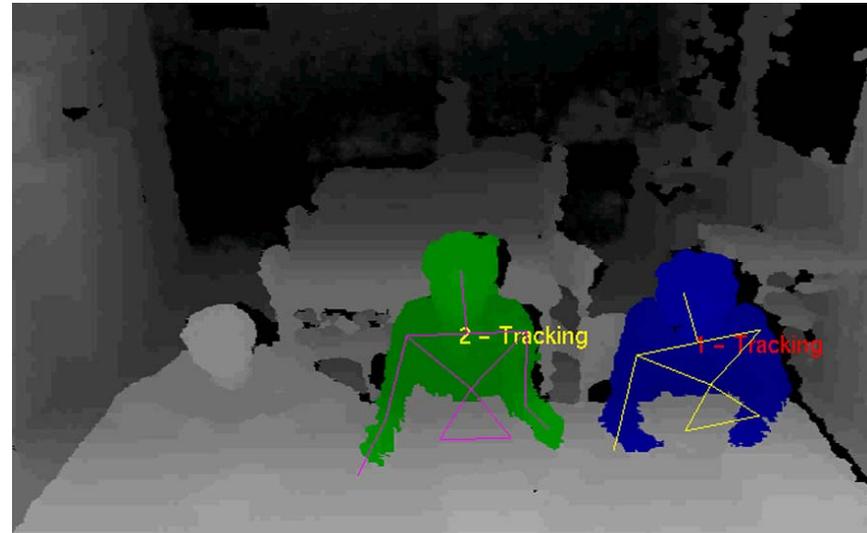
Conclusiones

- Auto-calibración:
 - ↑ Permite extraer información de videos donde no existe la calibración clásica.
 - ↓ No siempre funciona, o no es correcta.

Si es posible siempre es mejor utilizar el método de calibración clásica.

Auto-calibración.

Extracción de características de esqueleto.





Conclusiones

Testeo.

■ Dynamic Time Warping

- ↑ Funcionamiento correcto siempre que la secuencia que se busca sea muy similar al patrón.
- ↓ Clasificación incorrecta debido a que el parámetro de corte es muy elevado.
- ↓ No clasificación de la acción por diferencias en la posición de articulaciones no relevantes al gesto.



Trabajo futuro

Mejora en el análisis de DTW.

- Se propone una mejora del proyecto que consiste en:
 - Dar una importancia (peso) diferente para cada articulación. Para corregir la influencia de articulaciones no relevantes en la acción.
- Se han estudiado e implementado varias formas de calcular los pesos y se está preparando un publicación para la ICCV.

