



Universitat
de Barcelona

Multi-clasificación Discriminativa de Partes Corporales basada en Códigos Correctores de Errores

Realizado por: J. Tomás Pérez Yarza

Dirigido por: Sergio Escalera Guerrero

M. Ángel Bautista Martín

Barcelona, 26 de Septiembre de 2013



Introducción

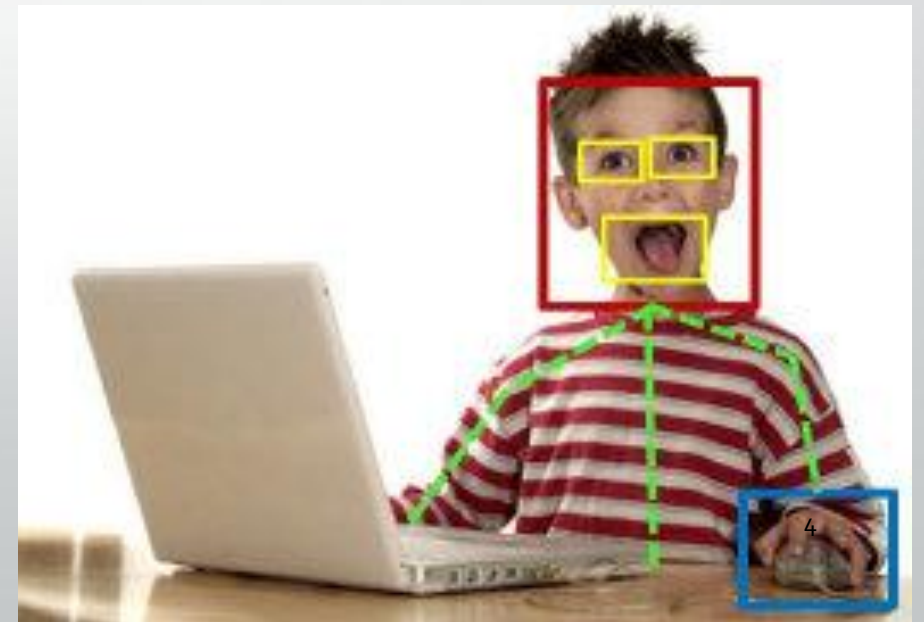
Introducción

- La visión por computador es un área desafiante y apasionante de investigación.
- Dificultades: Cambio de iluminación, oclusiones parciales, cambios en el punto de vista de referencia, deformaciones ...
- Gran abanico de aplicaciones.



Introducción

- Sistema de Reconocimiento de personas y segmentación de extremidades.
- Aplicación de distintas técnicas de Visión Artificial.
- Aprendizaje por parte de los Clasificadores.
- La clave: Amplio conjunto de entrenamiento.



Contenidos

- Base de datos HuPBA.
- Metodología.
- Resultados.
- Conclusiones.



Base de Datos HuPBA

Base de Datos HuPBA

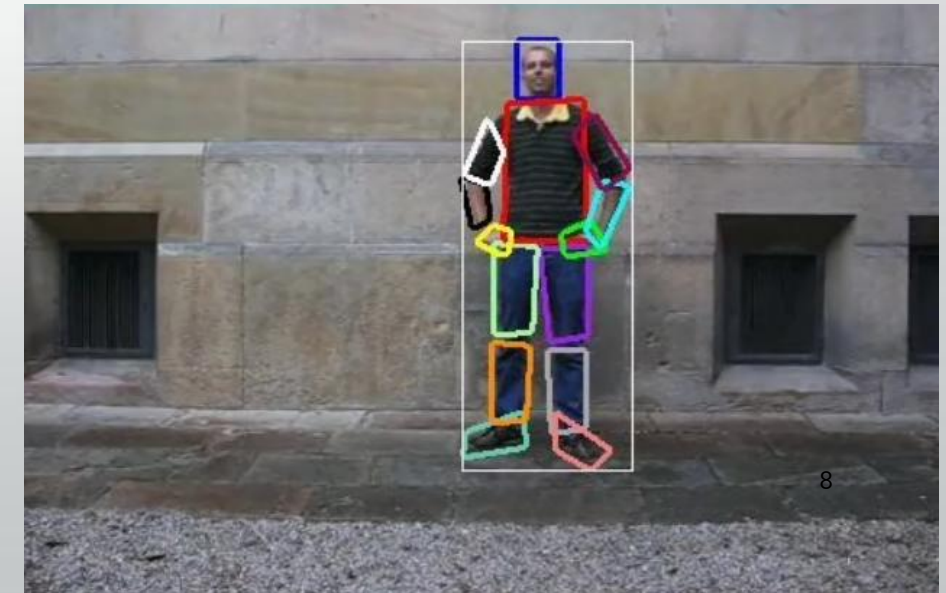
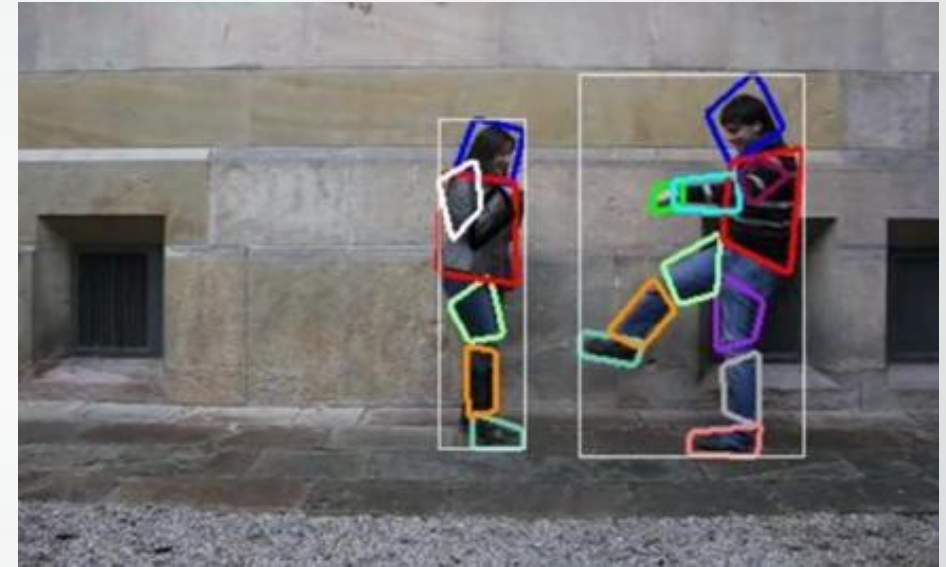


- **HuPBA: Human Pose Recovery and Behavior Analysis (Análisis de Pose y Comportamiento Humano).**
- **¿Por qué?**
Escasez de grandes Bases de Datos.
Contribuir en esta área de investigación.



Base de Datos HuPBA

- 9 vídeos de entorno controlado.
- Los actores realizan acciones: saludar, saltar, correr
- Descomposición de los vídeos en imágenes RGB (Total 8234).
- Etiquetaje de hasta 14 partes corporales.



- | | | |
|-----------------------|--------------------|-------------------|
| • Cabeza | • Brazo izquierdo | • Muslo izquierdo |
| • Torso | • Brazo derecho | • Muslo derecho |
| • Mano izquierda | • Pie izquierdo | |
| • Mano derecha | • Pie derecho | |
| • Antebrazo izquierdo | • Pierna izquierda | |
| • Antebrazo derecho | • Pierna derecha | |

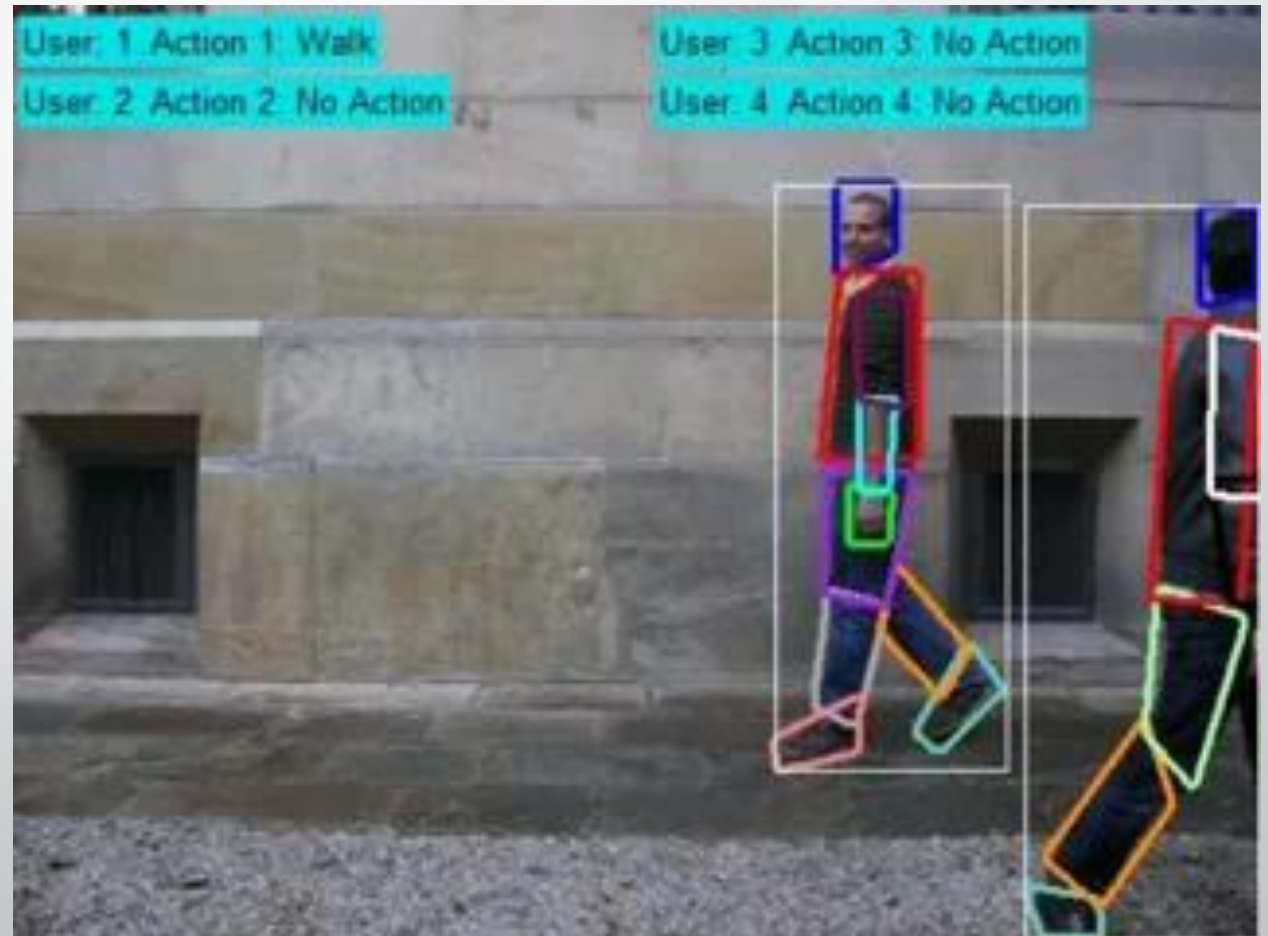
Base de Datos HuPBA



Base de Datos HuPBA

- Se han generado más de 120.000 máscaras.
- Fichero CSV con los gestos.
- Fichero CSV con las Bounding Boxes.
- Vídeos de Validación.

ID	Gesto
1	Saludar
2	Señalar
3	Aplaudir
4	Agacharse
5	Saltar
6	Caminar
7	Correr
8	Dar la mano
9	Abrazar
10	Dar dos besos
11	Pelear





Metodología

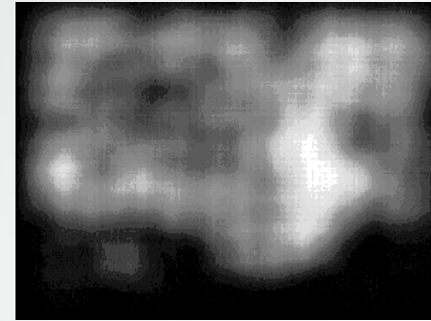
Metodología



Detección de Personas

Cascadas Adaboost
+ Haar-like Features
+ ECOC

Fase 1

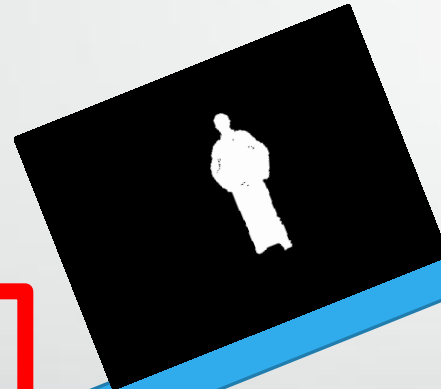


Mapa de probabilidad

Segmentación Binaria

Graph cuts: grabcut

Fase 2



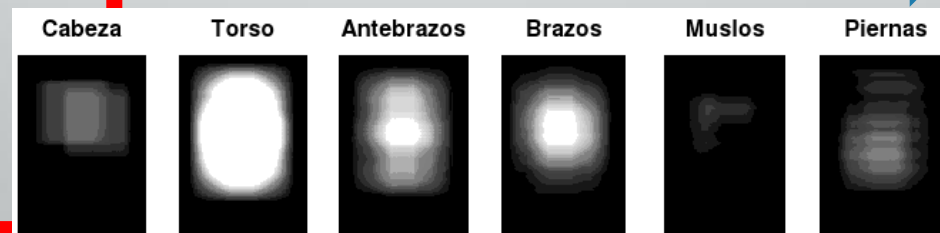
Máscara binaria

Detección de Extremidades

Clasificadores SVM
+ Descriptor HOG +
ECOC

Fase 3

6 imágenes de probabilidad



Multi-segmentación

Alpha-beta-swap
graph cuts

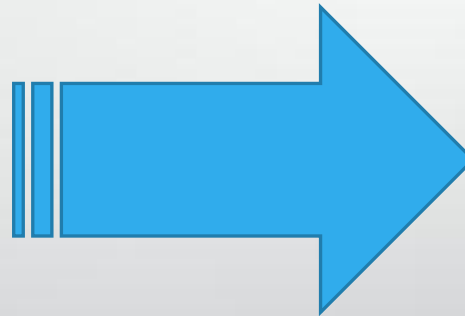
Fase 4



Metodología: Pasos previos

- Se ha reducido el número de extremidades.
 - Similares por simetría (brazos, antebrazos, muslos y piernas).
 - Tamaño muy pequeño, problemas redimensión (pies y manos).

ID	Extremidad
1	Cabeza
2	Torso
3	Mano izquierda
4	Mano derecha
5	Antebrazo izquierdo
6	Antebrazo derecho
7	Brazo izquierdo
8	Brazo derecho
9	Pie izquierdo
10	Pie derecho
11	Pierna izquierda
12	Pierna derecha
13	Muslo izquierdo
14	Muslo derecho



ID	Extremidad
1	Torso
2	Cabeza
3	Antebrazos
4	Brazos
5	Muslos
6	Piernas

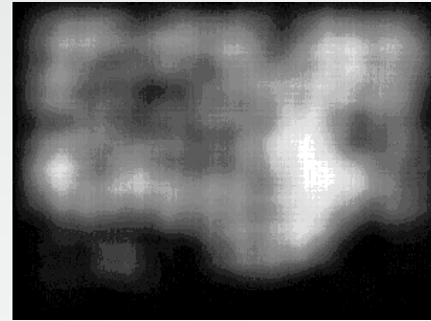
Metodología: Fase 1



Detección de Personas

Cascadas Adaboost
+ Haar-like Features
+ ECOC

Fase 1

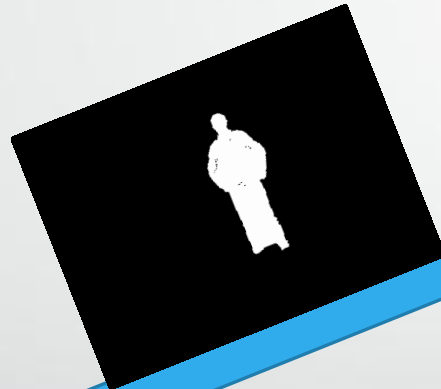


Mapa de probabilidad

Segmentación Binaria

Graph cuts: grabcut

Fase 2



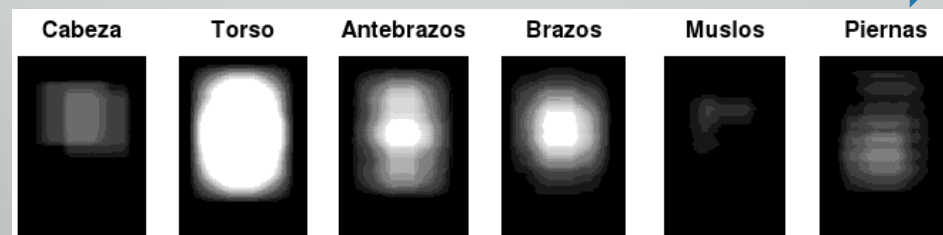
Máscara binaria

Detección de Extremidades

Clasificadores SVM
+ Descriptor HOG +
ECOC

Fase 3

6 imágenes de probabilidad



Multi-segmentación

Alpha-beta-swap
graph cuts

Fase 4

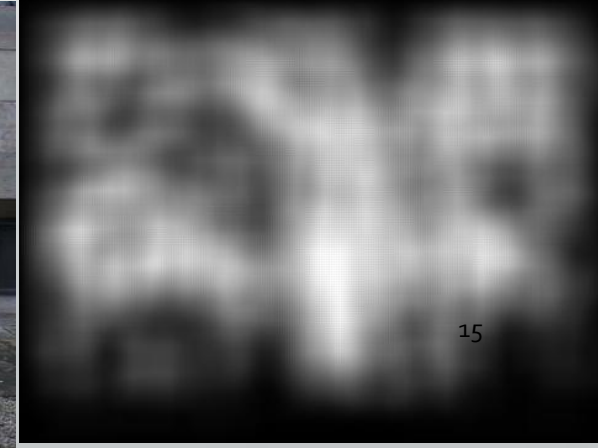
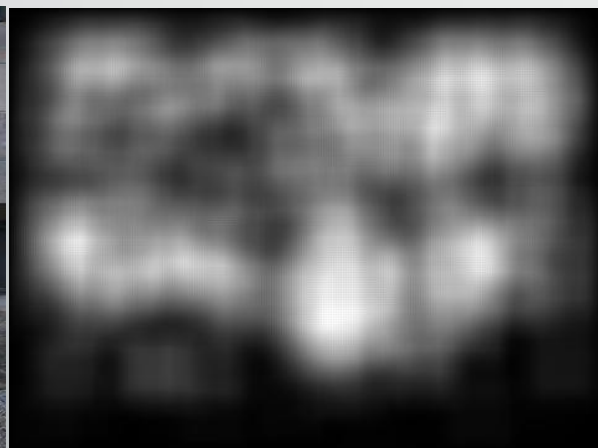
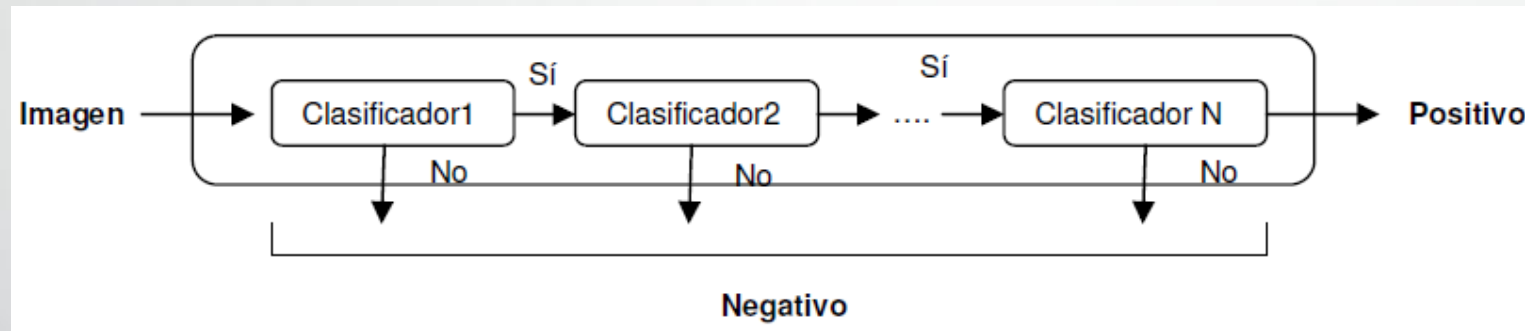


Metodología: Fase 1 - Detección de Personas

- Descripción de la imagen: características Haar-like.
- Cascada de clasificadores *binarios*: Adaboost.
- Conjunto de clasificadores “débiles” en distintas etapas.
- La gran mayoría de ejemplos negativos son eliminados al principio.

Cascadas Adaboost
+ Haar-like Features
+ECOC

Fase 1



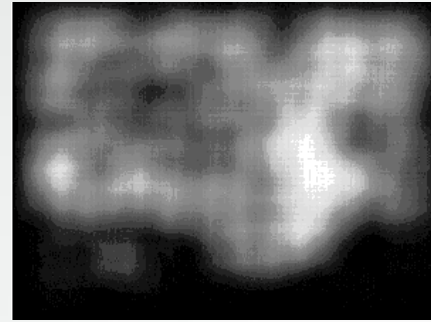
Metodología: Fase 2



Detección de Personas

Cascadas Adaboost
+ Haar-like Features
+ ECOC

Fase 1

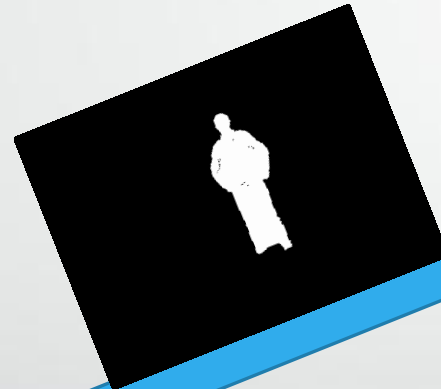


Mapa de probabilidad

Segmentación Binaria

Graph cuts: grabcut

Fase 2



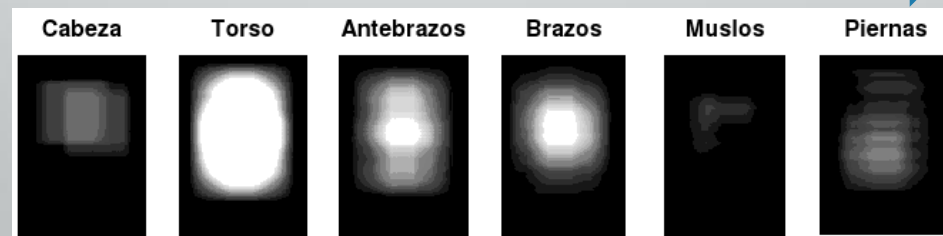
Máscara binaria

Detección de Extremidades

Clasificadores SVM
+ Descriptor HOG +
ECOC

Fase 3

6 imágenes de probabilidad



Multi-segmentación

Alpha-beta-swap
graph cuts

Fase 4



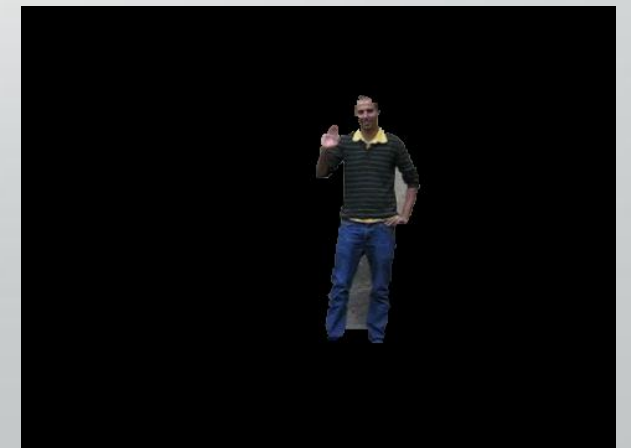
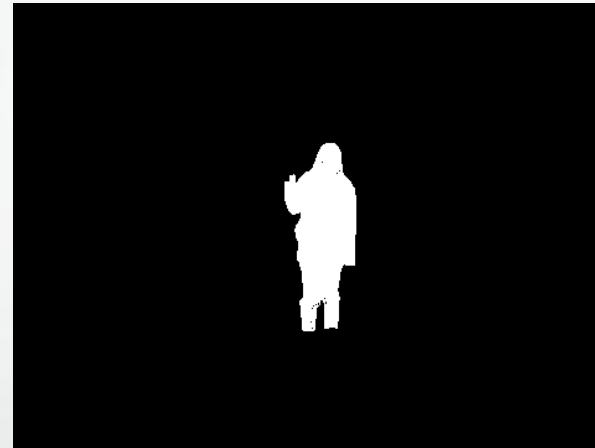
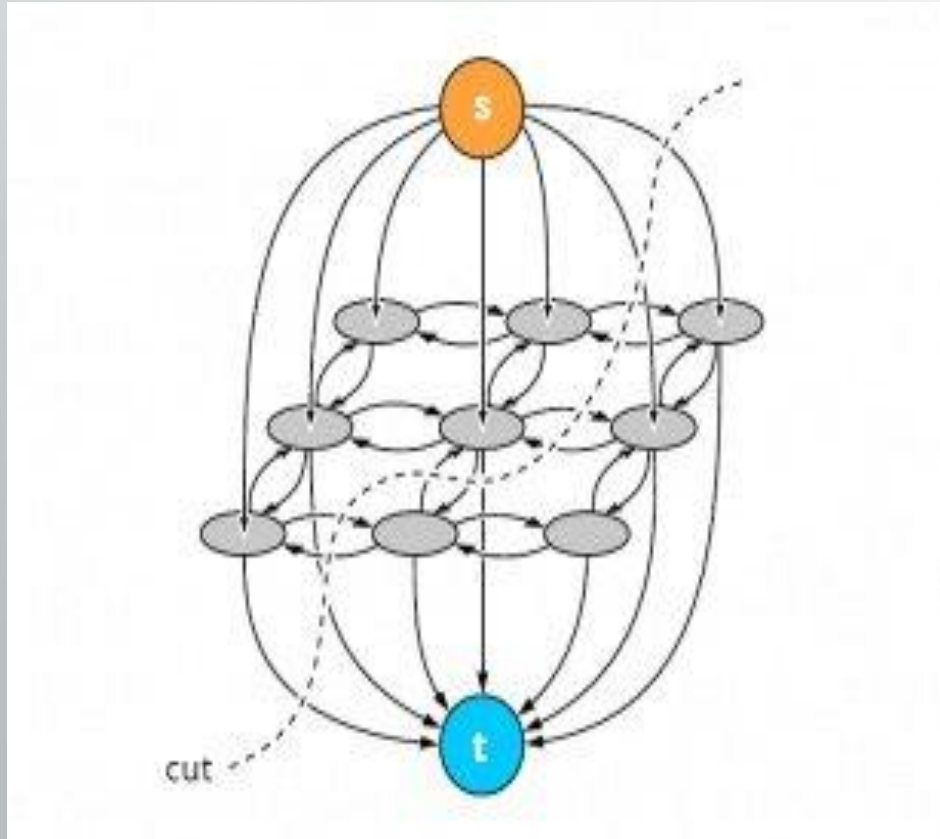
Metodología: Fase 2 – Segmentación Binaria

- Aplicación del algoritmo Ford-Fulkerson hasta su saturación
- Las fronteras saturadas corresponden al contorno de segmentación resultante

Segmentación Binaria

Graph cuts: grabcut

Fase 2



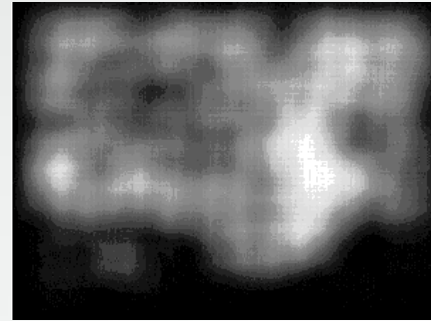
Metodología: Fase 3



Detección de Personas

Cascadas Adaboost
+ Haar-like Features
+ECOC

Fase 1

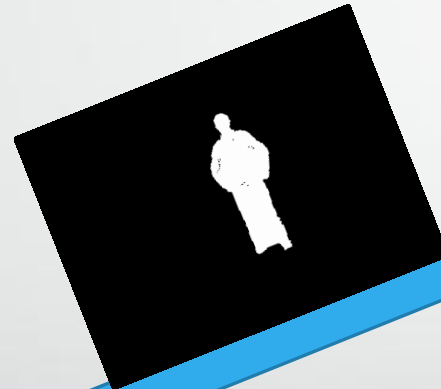


Mapa de probabilidad

Segmentación Binaria

Graph cuts: grabcut

Fase 2



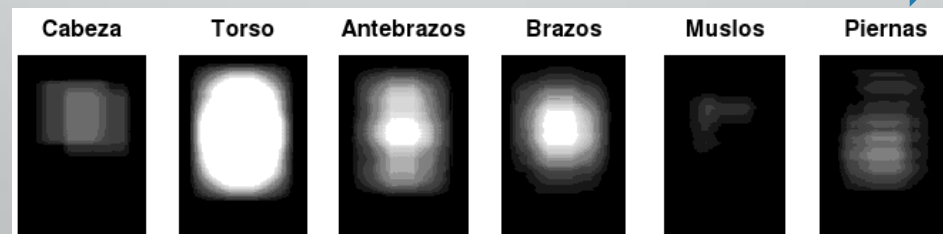
Máscara binaria

Detección de Extremidades

Clasificadores SVM
+ Descriptor HOG +
ECOC

Fase 3

6 imágenes de probabilidad



Multi-segmentación

Alpha-beta-swap
graph cuts

Fase 4



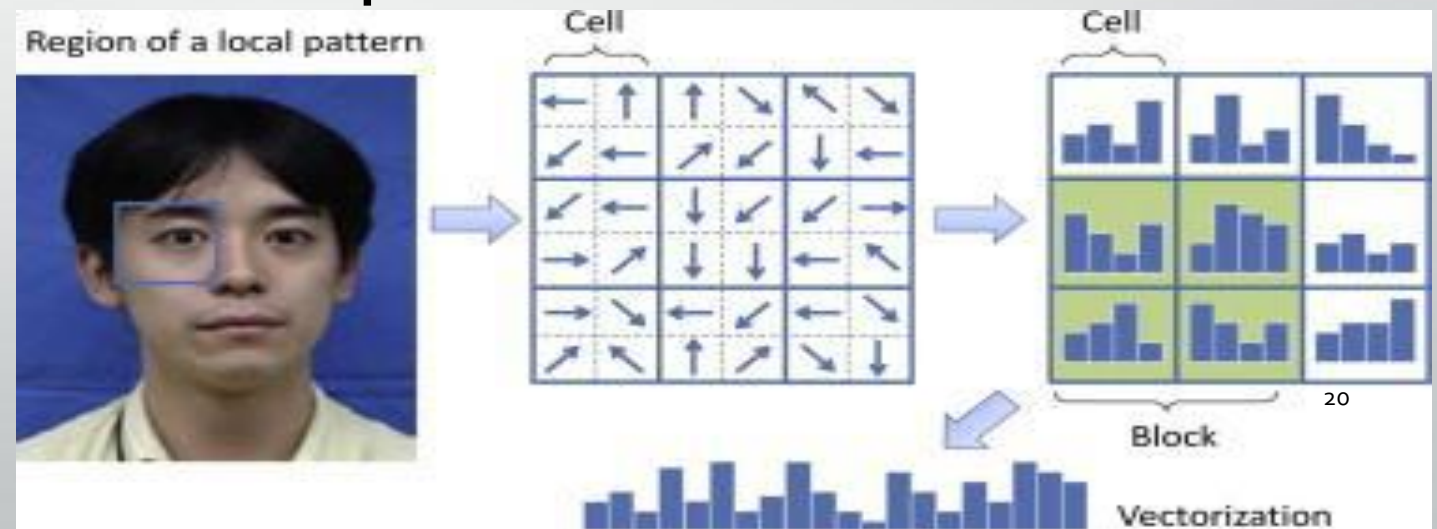
Metodología: Fase 3 – Detección de Extremidades

Detección de Extremidades

- **Descriptor HOG: Histogramas de Orientación de Gradientes**
(más discriminativo que Haar-like).
- Una imagen se describe como una distribución de las direcciones de cambio de intensidad.
- Divide la imagen en celdas.
- Se calcula un histograma para cada celda con todas las orientaciones detectadas.
- La combinación de histogramas forma el descriptor.

Clasificadores SVM
+ Descriptor HOG +
ECOC

Fase 3



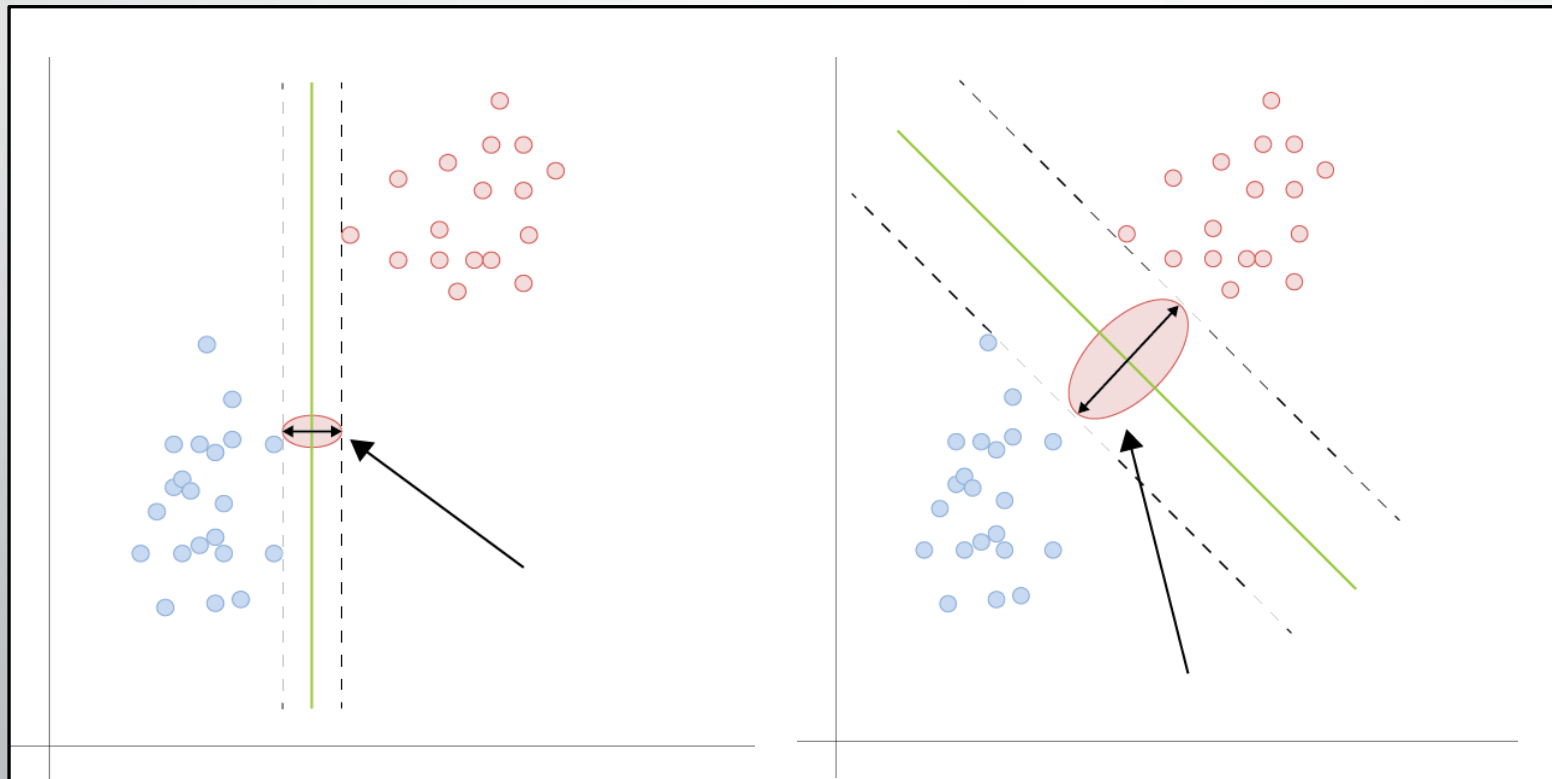
Metodología: Fase 3 – Detección de Extremidades

- Clasificador SVM Support Vector Machine (Máquinas de Soporte Vectorial).
- Un hiperplano separa los datos de cada clase.
- La solución óptima será la que tenga un mayor margen entre los puntos más próximos de cada clase.

Detección de Extremidades

Clasificadores SVM
+ Descriptor HOG +
ECOC

Fase 3

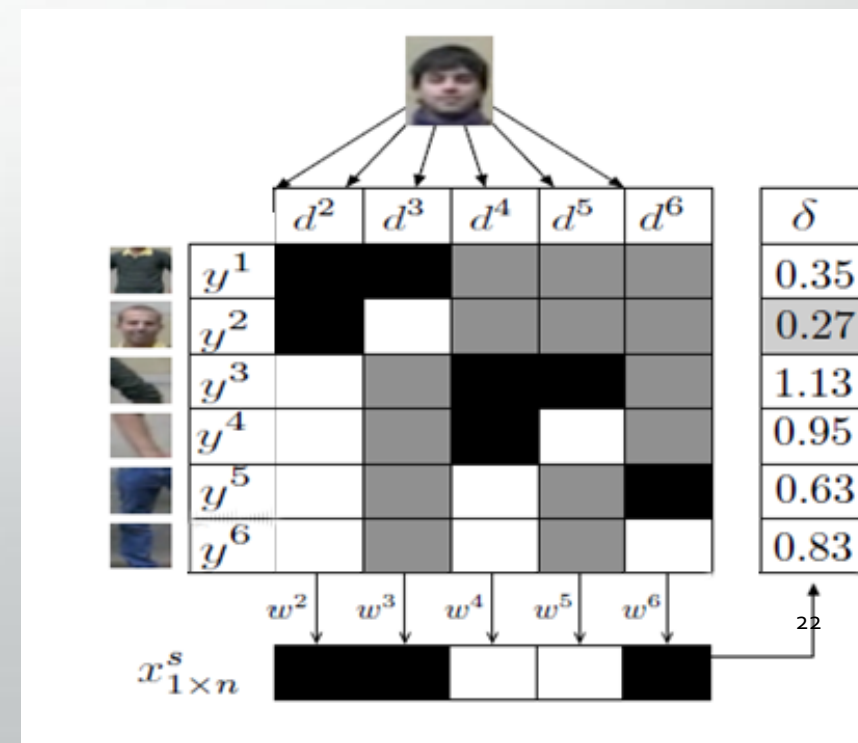
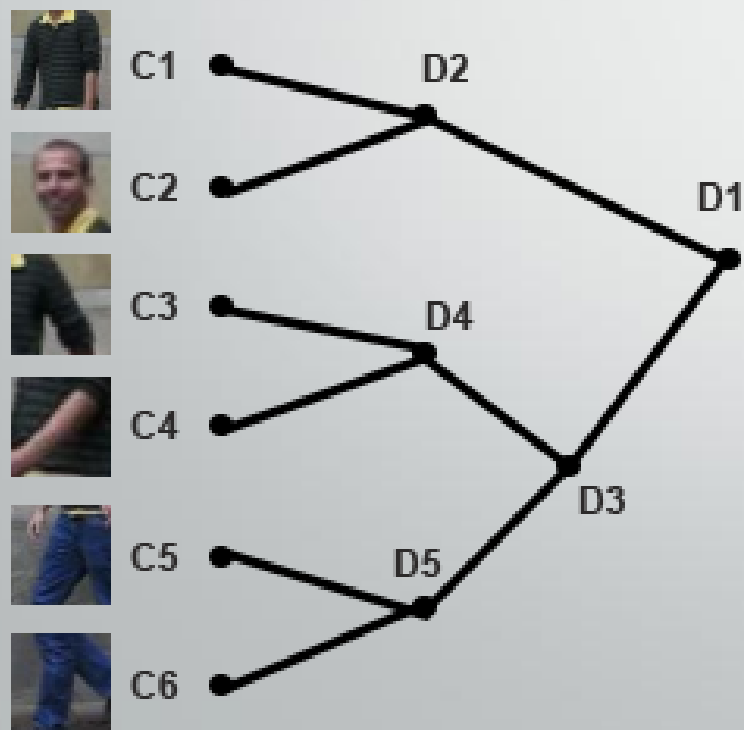


Metodología: Fase 3 - ECOC

- Error Correcting Output Codes (Códigos de Corrección de Errores).
- La combinación de clases codifican la matriz ECOC. (Sin fondo)
- Se descodifica mediante la Expresión: $d(x, y) = \frac{\sum_{i=1}^5 -x^i \cdot y^i \cdot W^i}{\sum |Y|}$

Clasificadores SVM
+ Descriptor HOG +
ECOC

Fase 3



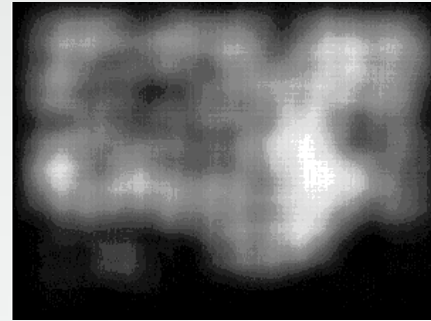
Metodología: Fase 4



Detección de Personas

Cascadas Adaboost
+ Haar-like Features
+ECOC

Fase 1

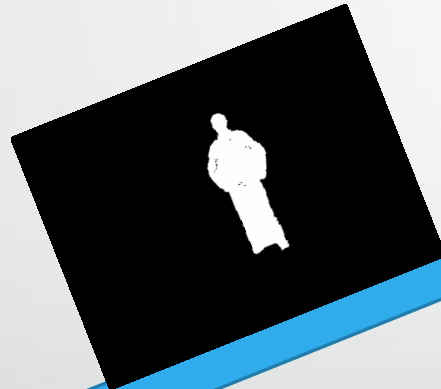


Mapa de probabilidad

Segmentación Binaria

Graph cuts: grabcut

Fase 2



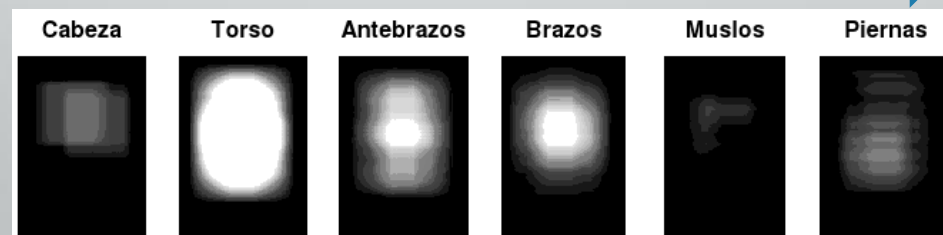
Máscara binaria

Detección de Extremidades

Clasificadores SVM
+ Descriptor HOG +
ECOC

Fase 3

6 imágenes de probabilidad



Multi-segmentación

Alpha-beta-swap
graph cuts

Fase 4



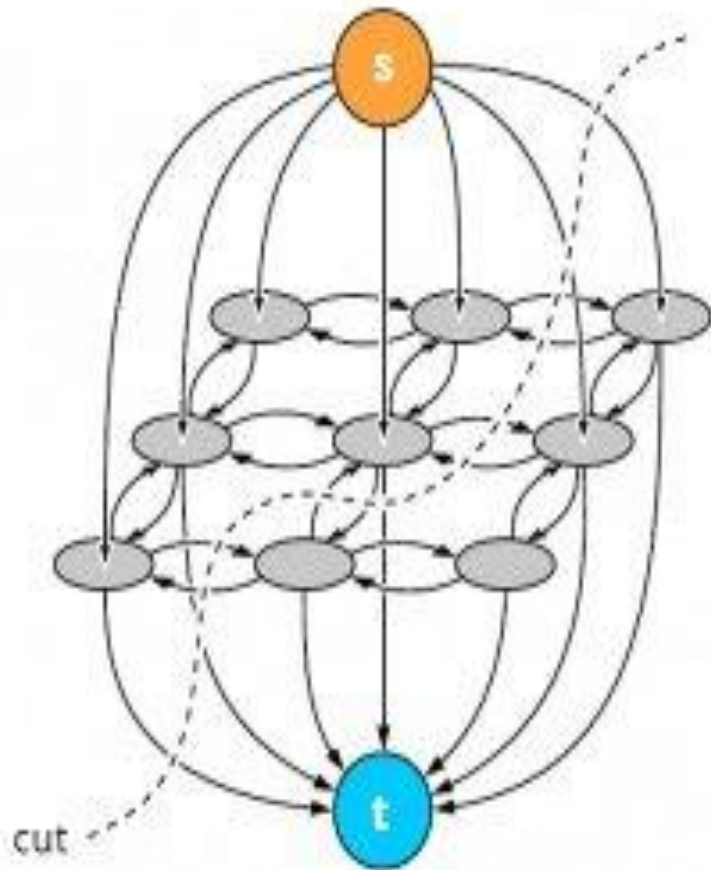
Metodología: Fase 4 – Segmentación Binaria

- Lanzamiento iterativo de graph cuts.

Multi-segmentación

Alpha-beta-swap
graph cuts

Fase 4





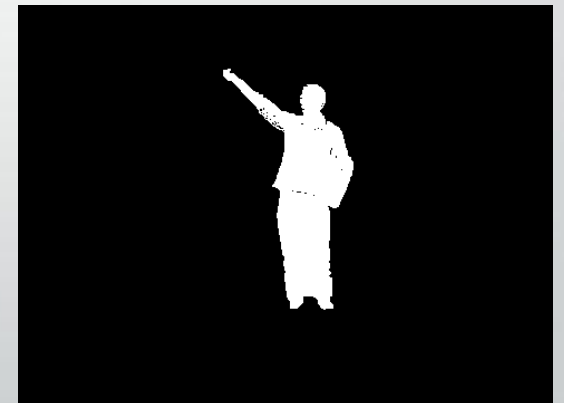
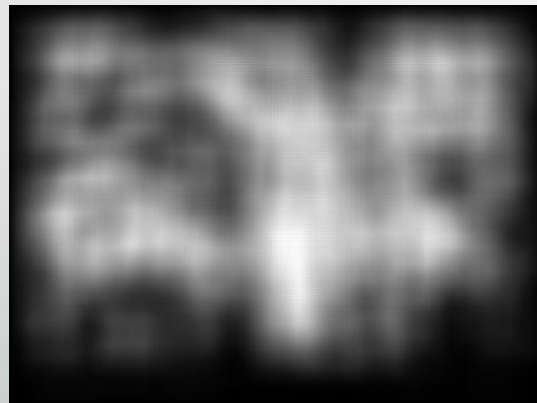
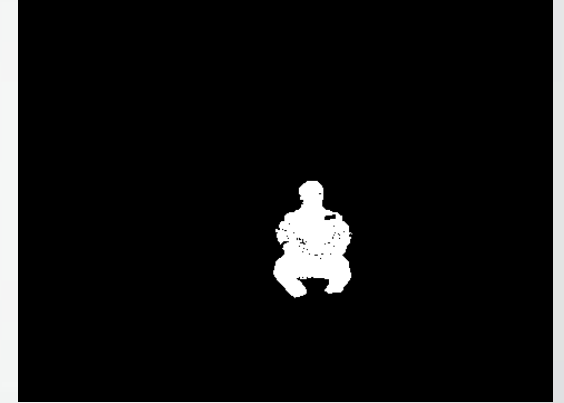
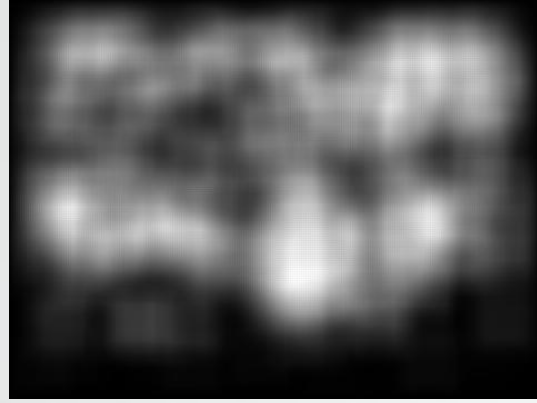
Resultados

Resultados: Parámetros y Configuración

- **Overlap:** Evaluación de la segmentación. **Overlap = 1** óptimo.
- **Accuracy:** Evaluación de la clasificación. **Accuracy = 100%** óptimo.
- 6 clasificadores en 8 Etapas (Cascada Adaboost).
- Tamaño de descriptor de 324 elementos (HOG).
- *Cross Validation* optimización parámetros *C* y *Gamma* (SVM).
- Experimentación *Leave One Out* (8 + 1).
- Reducción del conjunto de entrenamiento hasta el 10% del total (tiempo).



Resultados: Fase 1 + Fase 2



Fase 1: Detección de Personas

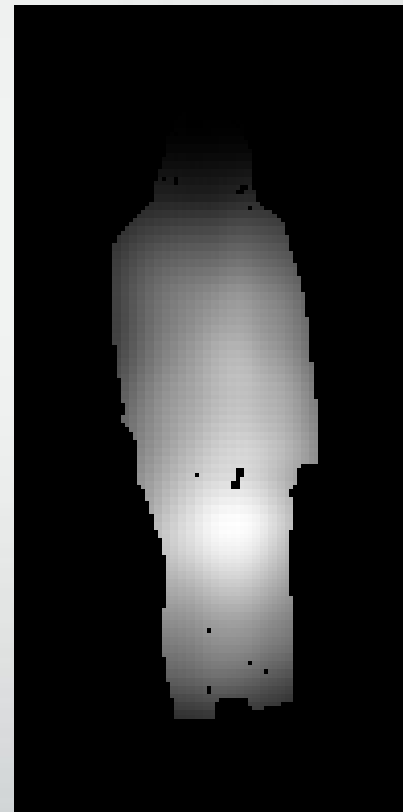
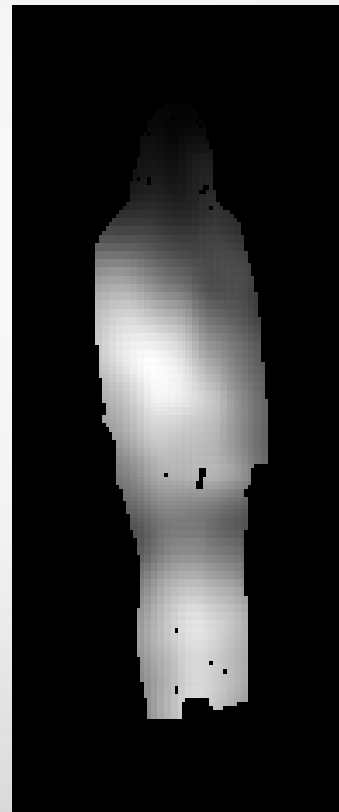
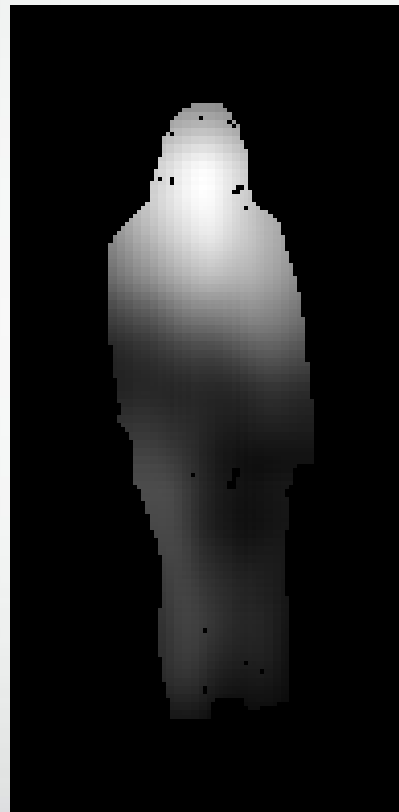
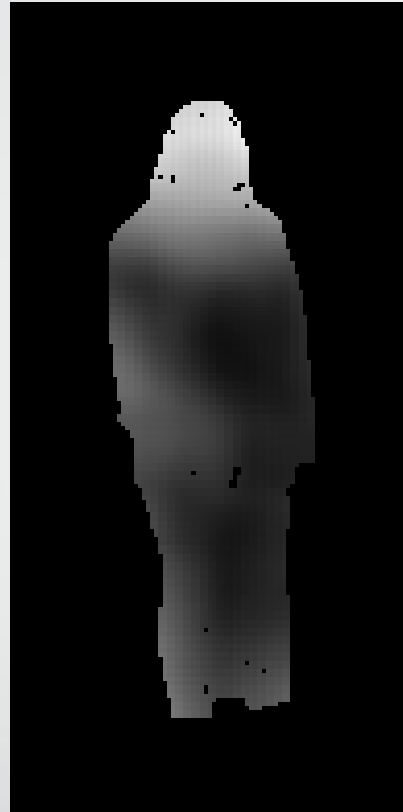
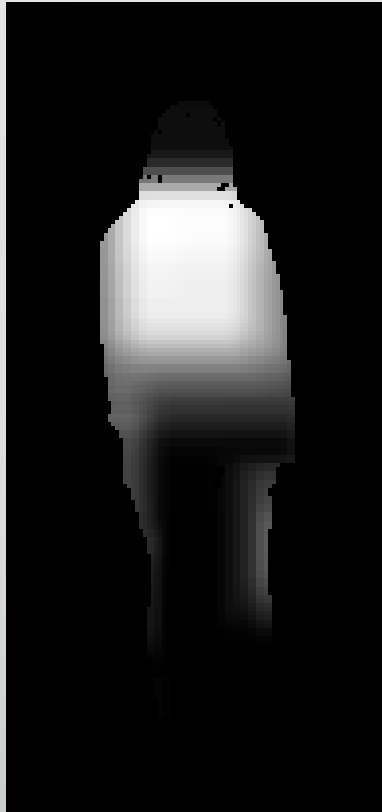
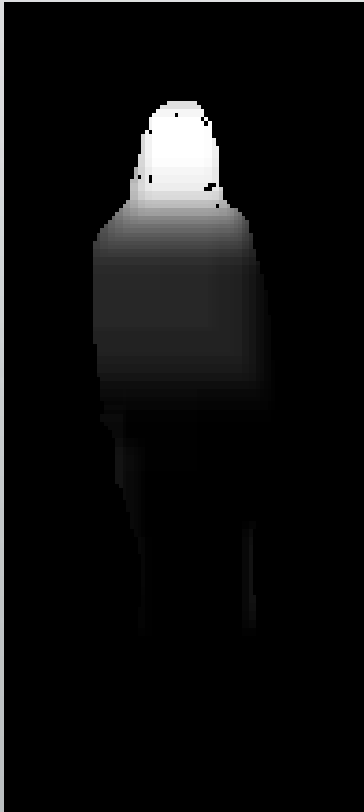
Fase 2: Segmentación binaria²⁷

Resultados: Fase 3 – Multi-clasificación de Extremidades

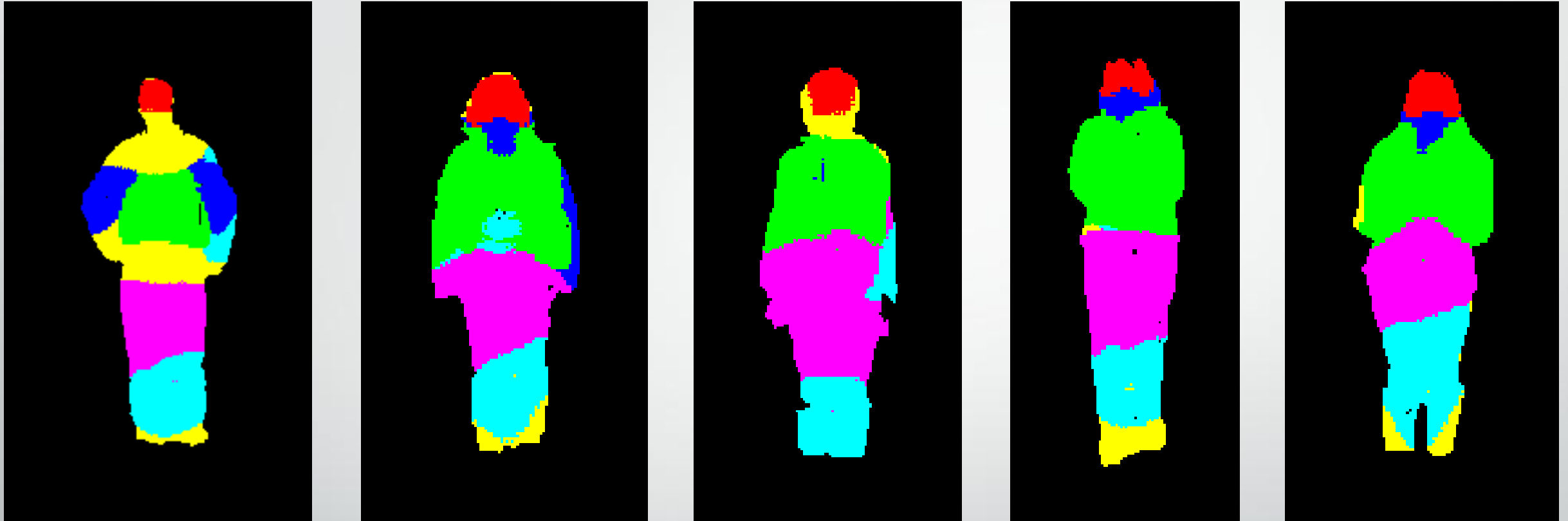
- Se han lanzado un total de 9 experimentos.
- Matriz de confusión promedia de los 9 experimentos:

Clase	C1	C2	C3	C4	C5	C6
C1 (torso)	5%					
C2 (cabeza)		80%				
C3 (brazos)			22%			
C4 (antebrazos)				44%		
C5 (muslos)					90%	
C6 (piernas)						58%

Resultados: Fase 3 – Multi-clasificación de Extremidades



Resultados: Fase 4 – Multi-segmentación



Cabeza	Torso	Brazos	Antebrazos	Muslos	Piernas
0,2737	01667	0,0921	0,3990	0,3021	0,3884



Conclusiones

Conclusiones: Resumen

- Se ha diseñado una interfaz de etiquetaje y se han etiquetado múltiples regiones y gestos en secuencias de datos RGB de múltiples usuarios (HuPBA).
- Se han entrenado y utilizado varios clasificadores: Clasificación binaria / Multi-clasificación.
- Se han aplicado técnicas de segmentación: Segmentación binaria / Multi-segmentación.
- Se han obtenido unos resultados prometedores.
- Contribución con la publicación de la Base de datos HuPBA.


Conclusiones: Handicap

- Un gran Handicap: los tiempos.
 - Reducción del total de ejemplos de entrenamiento.
 - Acotar el tamaño de ventana de la Sliding Window.
 - Aumentar el desplazamiento entre ventanas de la Sliding Window.
 - Experimentos: 3 Aulas
 - 31 PC x 2 núcleos = 62 Núcleos
 - 24 PC x 8 núcleos = 192 Núcleos
 - 25 PC x 2 núcleos = 50 Núcleos
- TOTAL: 304 Núcleos ~ 1 semana**

Conclusiones: Trabajo futuro

- Afrontar el Handicap.
- Combinar otras metodologías del Estado del Arte.





Multi-clasificación Discriminativa de Partes Corporales basada en Códigos Correctores de Errores

Gracias

Realizado por: J. Tomás Pérez Yarza

Dirigido por: Sergio Escalera Guerrero
M. Ángel Bautista Martín