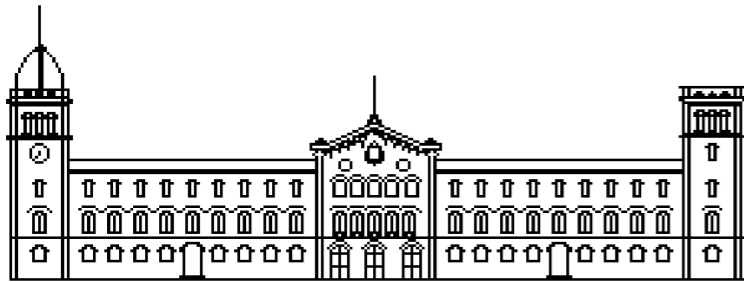


**Programació Dinàmica per al
Reconeixement de Gestos:
Aplicació a una nova base de dades de
Subjectes, Extremitats, Accions i
Interaccions**



Facultat de Matemàtiques
Universitat de Barcelona

Alumne: Jordi Suñé Fontanals
Directors: Dr. Sergio Escalera Guerrero
Miguel Ángel Bautista Martín
Realitzat a: Departament de Matemàtica
Aplicada i Anàlisi. UB
Barcelona, 13 de febrer de 2012

Continguts:

1 Introducció

2 Objectius

3 Base de dades

4 Anàlisi

5 Disseny i implementació

6 Resultats

7 Conclusions

1 Introducció

2 Objectius

3 Base de dades

4 Anàlisi

5 Disseny i implementació

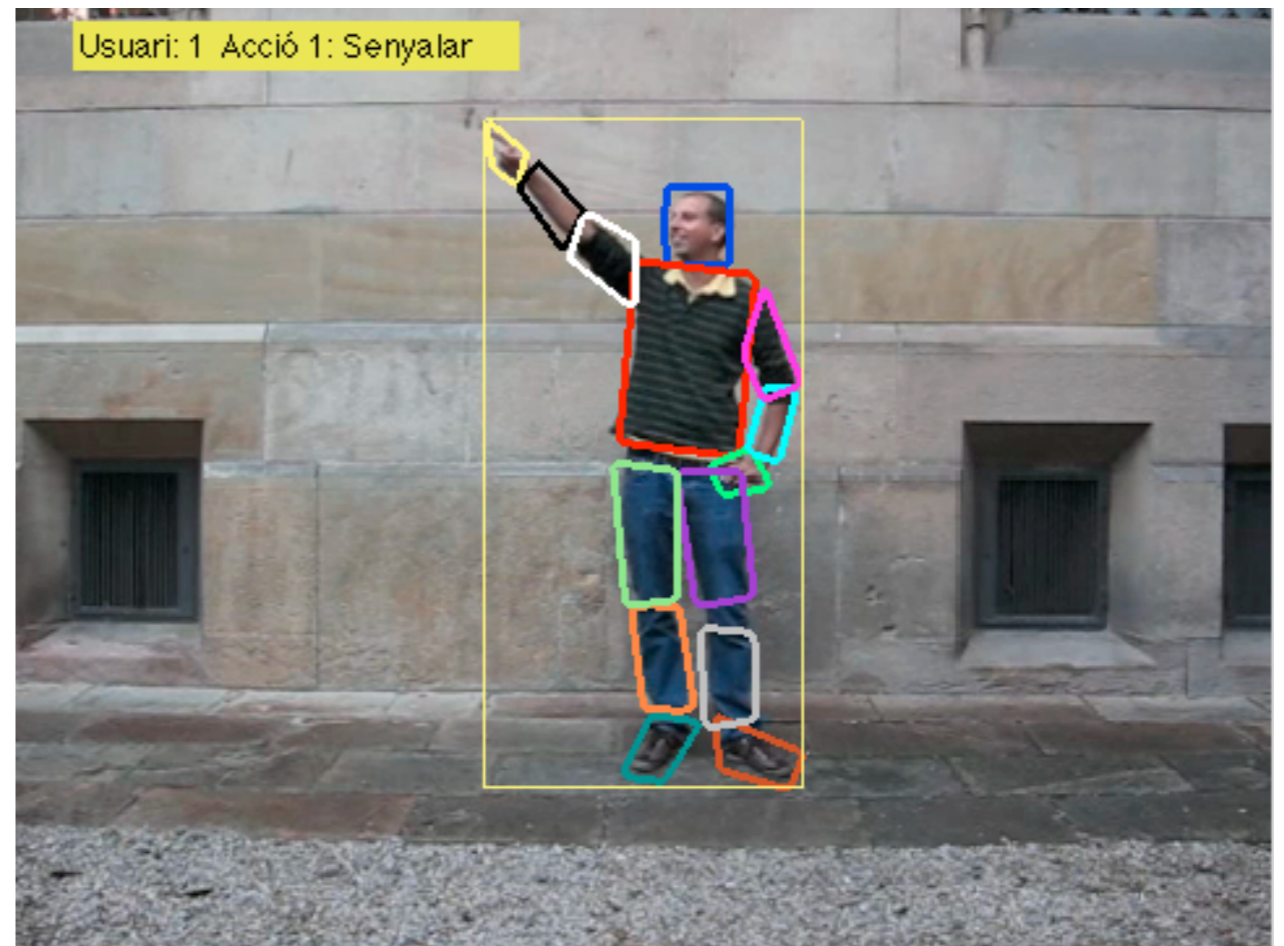
6 Resultats

7 Conclusions

Introducció:

Conceptes principals

- ▶ Què és el reconeixement d'accions?
 - ▶ Estat actual de la tecnologia
- ▶ Què és una acció?
- ▶ Com obtenim les mostres de la Base de dades?



Introducció:

Resum proposta

- ▶ Realització d'una base de dades amb suficients mostres per dur a terme els experiments
- ▶ Definir i testejar el millor mètode de càlcul de descriptors dels 4 plantejats
- ▶ Aplicar una tècnica de reconeixement de gestos

1 Introducció

2 Objectius

3 Base de dades

4 Anàlisi

5 Disseny i implementació

6 Resultats

7 Conclusions

Objectius:

Objectius

- ▶ Dividir vídeos en fotogrames
- ▶ Etiquetar mostres i creació de la Base de Dades
- ▶ Obtenir les matrius de descriptors
- ▶ Implementar i aplicar Dynamic Time Warping sobre les matrius de descriptors obtingudes
- ▶ Calcular el percentatge d'encert dels mètodes emprats i veure quin és el millor

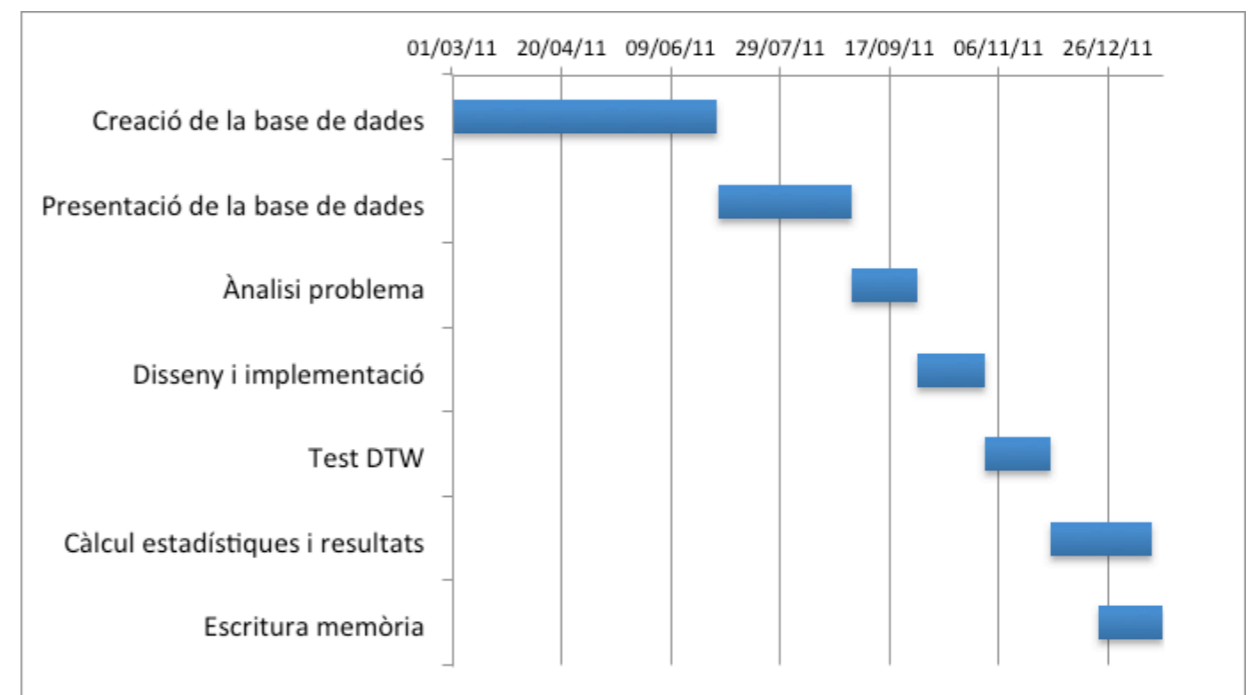
Objectius:

Planificació:

Taula de planificació:

Activitats	Data inici	Data fi	Dies
Creació de la BD	01/03/11	30/06/11	121
Presentació de la BD	01/07/11	31/08/11	61
Anàlisi del problema	31/08/11	30/09/11	30
Disseny i implementació	30/09/11	31/10/11	31
Test DTW	31/10/11	30/11/11	30
Càlcul d'estadístiques i resultats	30/11/11	15/01/12	46
Escritura memòria	22/12/11	20/01/12	29

Diagrama de Gantt:



1 Introducció

2 Objectius

3 Base de dades

4 Anàlisi

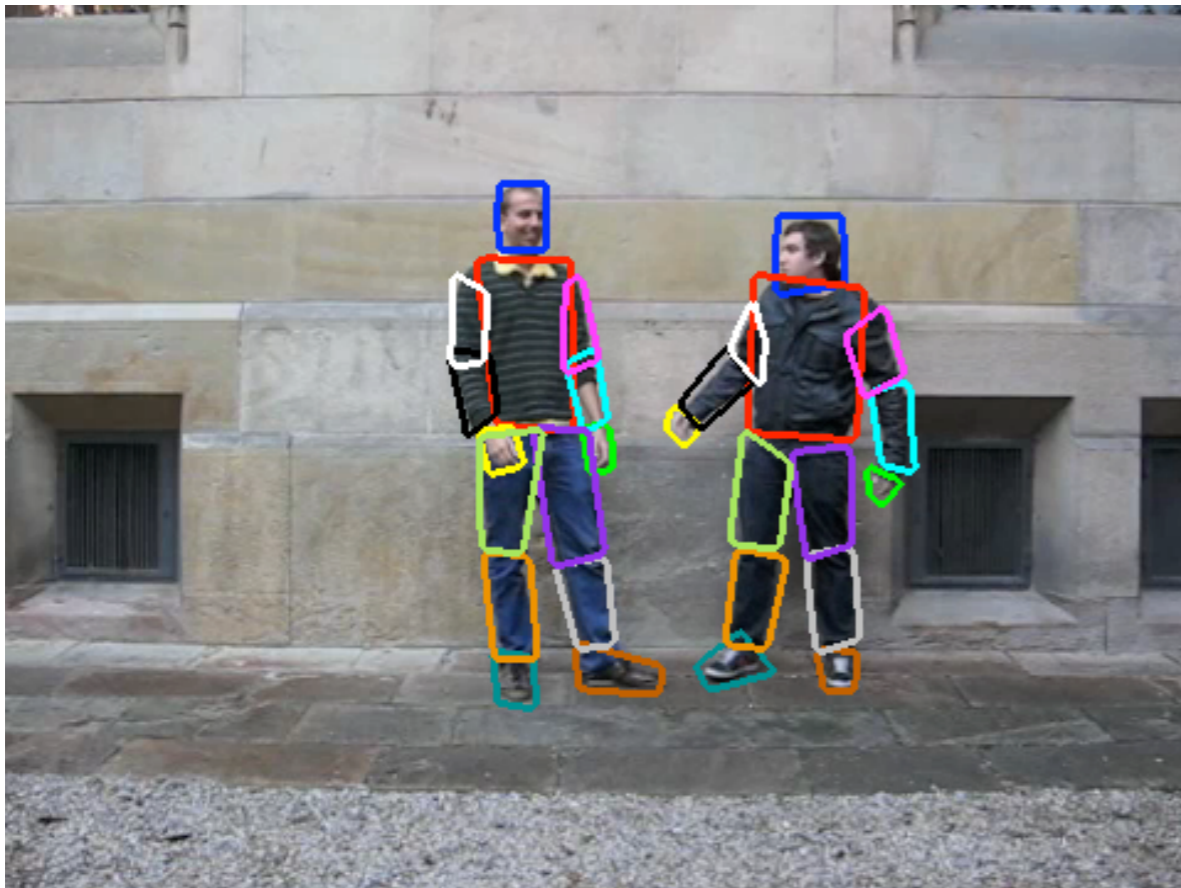
5 Disseny i implementació

6 Resultats

7 Conclusions

Base de dades:

Extracció de màscares:



Parts etiquetades

- ▶ Cap
- ▶ Tronc
- ▶ Mà esquerra
- ▶ Mà dreta
- ▶ Avantbraç dret
- ▶ Avantbraç esquerra
- ▶ Braç dret
- ▶ Braç esquerra
- ▶ Peu dret
- ▶ Peu esquerra
- ▶ Cama dreta
- ▶ Cama esquerra
- ▶ Cuixa dreta
- ▶ Cuixa esquerra

Base de dades:

Condicions etiquetatge

- Etiquetatge de parts del cos
 - Regla del 50%
- Etiquetatge d'accions
 - Regla d'inici de moviment

Casos sorgits durant l'etiquetatge:

Cas 1: No està visible el 50% de la part del cos

X Mal etiquetatge



Cas 2: Està visible el 50% de la part del cos

✓ Bon etiquetatge



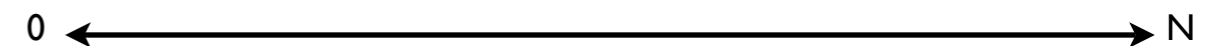
Base de dades:

Vectors d'accions:

Obtenció dels vectors d'accions

- ▶ 1 Vector per acció (11 vectors per individu)
- ▶ Obtenció manual dels vectors d'accions
- ▶ Extracció dels vectors d'accions a través de les matrius de descriptors

Composició del vector:



N = núm. de fotogrames del vídeo

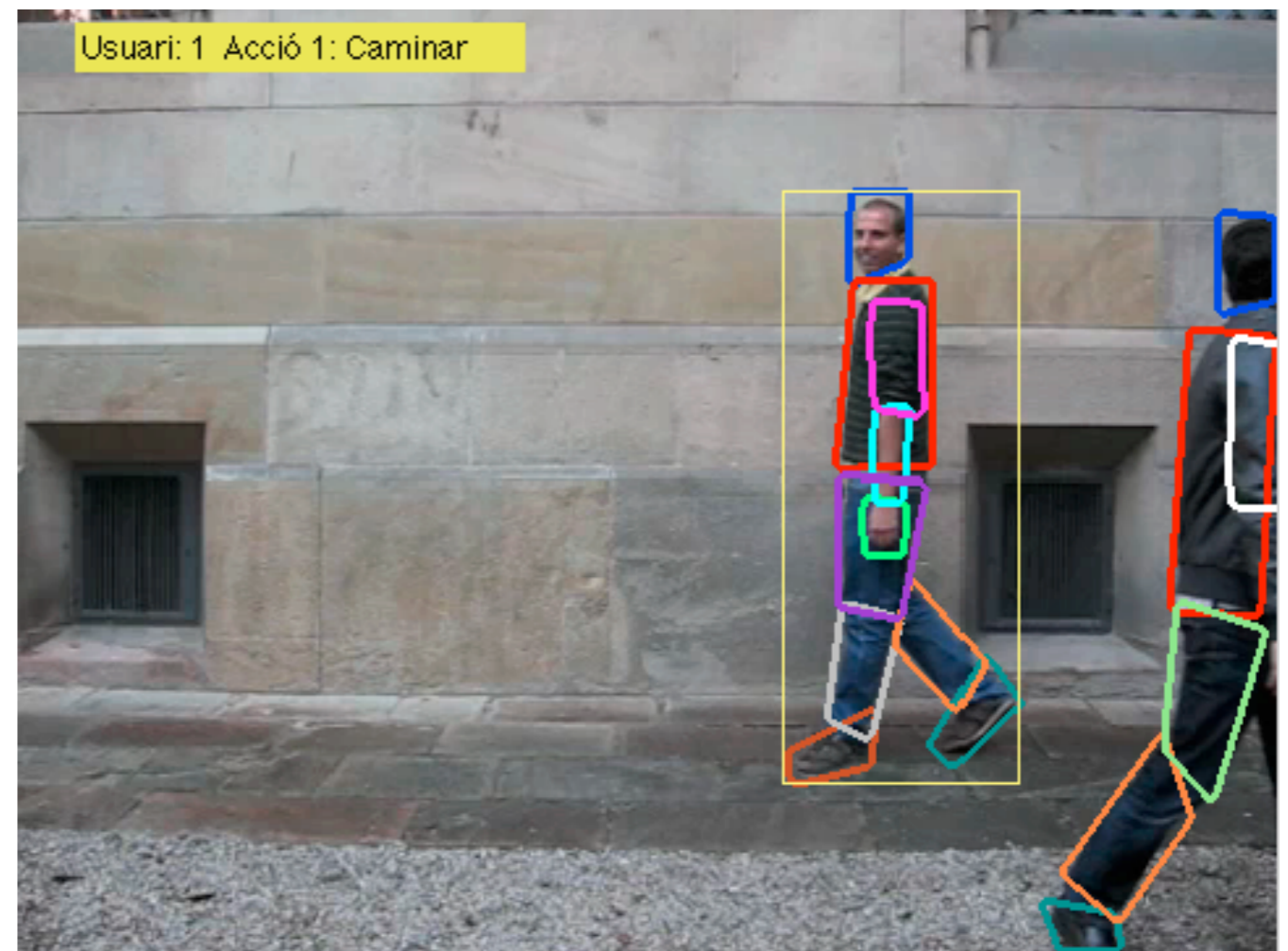
Valors de cada posició $\left\{ \begin{array}{l} 1 \text{ si apareix l'acció} \\ 0 \text{ si no apareix l'acció} \end{array} \right.$

Base de dades:

Dades rellevants de la BD

- ▶ Núm. de vídeos etiquetats: 10
- ▶ Màscares etiquetades manualment: 14410
- ▶ Màscares totals: 288827
- ▶ Format escollit per guardar les màscares: BMP
 - ▶ Avantatges
 - ▶ Inconvenients

Vídeo de mostra:



1 Introducció

2 Objectius

3 Base de dades

4 Anàlisi

5 Disseny i implementació

6 Resultats

7 Conclusions

Anàlisi:

Patrons d'accions que intentarem reconèixer en mostres desconegudes:

Accions

Accions individuals:

- Saludar
- Senyalar
- Aplaudir
- Saltar
- Ajupir
- Caminar
- Córrer

Accions col·lectives:

- Donar la mà
- Abraçar
- Donar dos petons
- Barallar

Anàlisi:

Algorisme utilitzat pel reconeixement de patrons d'accions:

Dynamic Time Warping

- ▶ Definició
- ▶ Requeriments:
 - ▶ Matriu de la seqüència completa
 - ▶ Matriu amb patró a comparar
 - ▶ Valor de tall

Pseudocodi:

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

  for i := 1 to n
    for j := 1 to m
      cost := d(s[i], t[j])
      DTW[i, j] := cost + minimum(DTW[i-1, j ],
                                  DTW[i , j-1],
                                  DTW[i-1, j-1])

  return DTW[n, m]
}
```

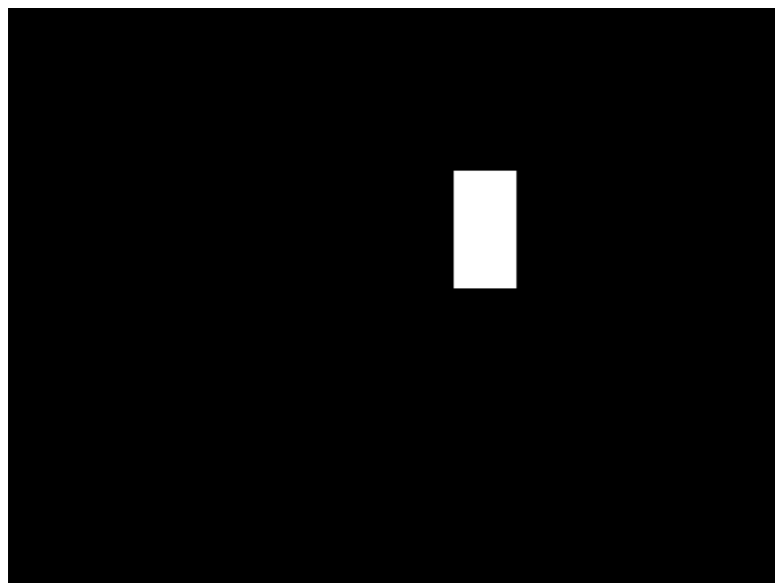

Anàlisi: Preprocessat de les dades

Dades extretes de les màscares

Per cada màscara fem:

- Àrea
- Centroide

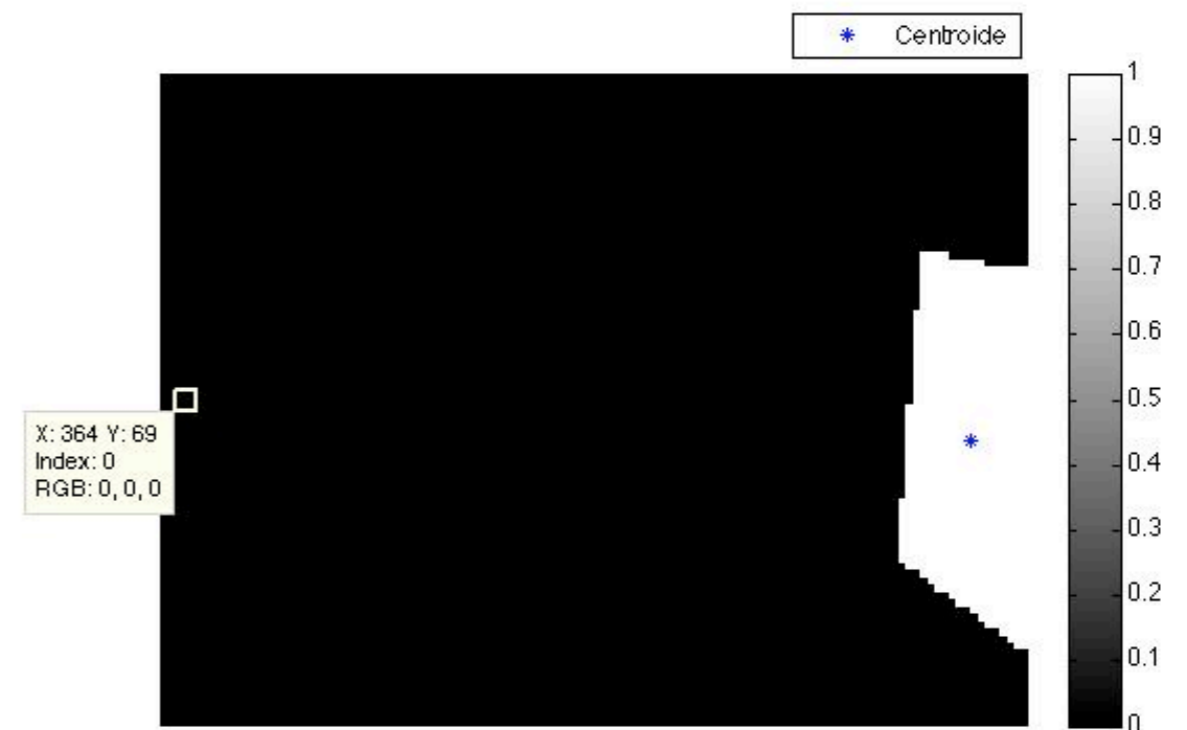
Àrea:



Àrea màscara:

$\text{Sum}(\text{Sum}(\text{màscara}))$

Centroide:



El centroide de la màscara és el punt de referència que fem per mesurar les distàncies

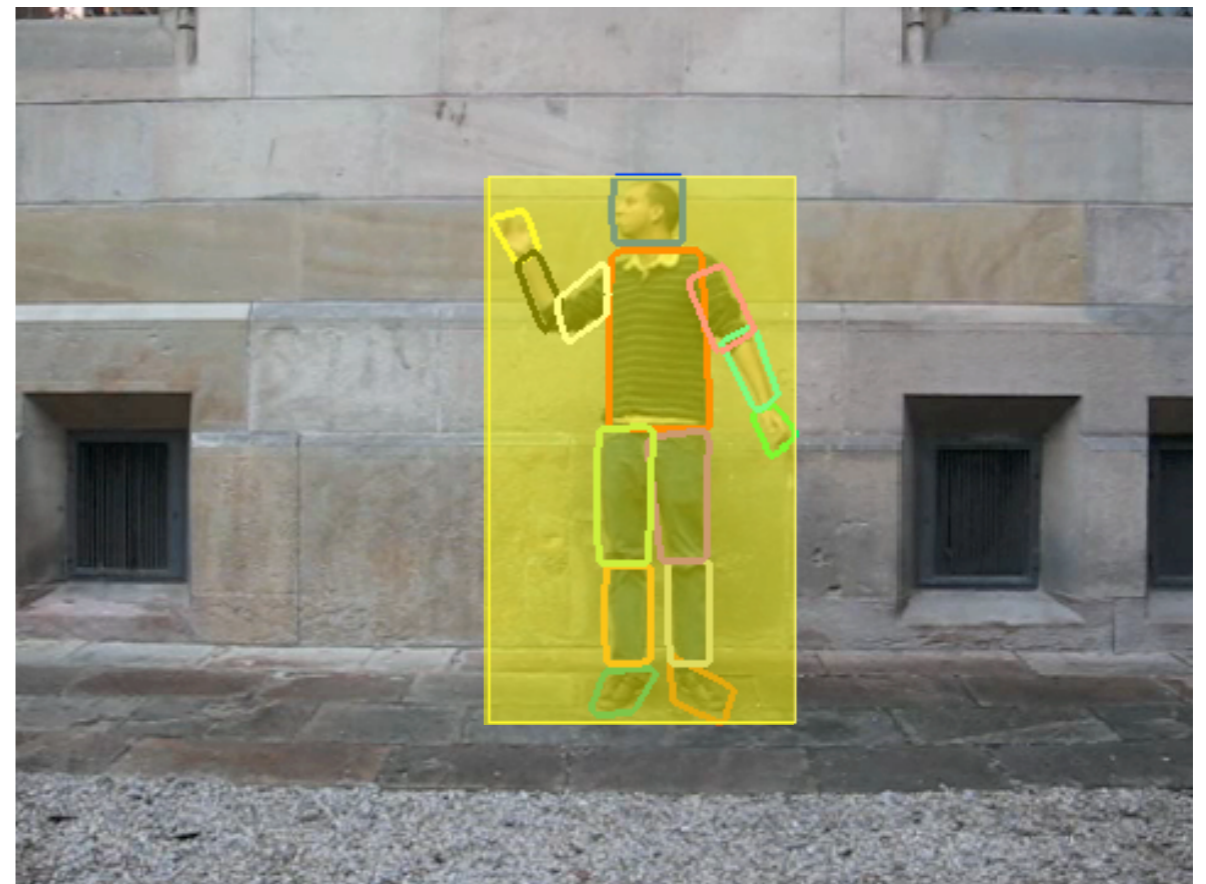
Anàlisi: Preprocessat de les dades

Bounding Box

Mirem per cada fotograma:

- ▶ Extrems de cada part etiquetada:
 - ▶ Valor en l'eix X més petit (X_{min})
 - ▶ Valor en l'eix X més gran (X_{max})
 - ▶ Valor en l'eix Y més petit (Y_{min})
 - ▶ Valor en l'eix Y més gran (Y_{max})
- ▶ Tracem línies entre els punts i guardem la seva àrea en una estructura

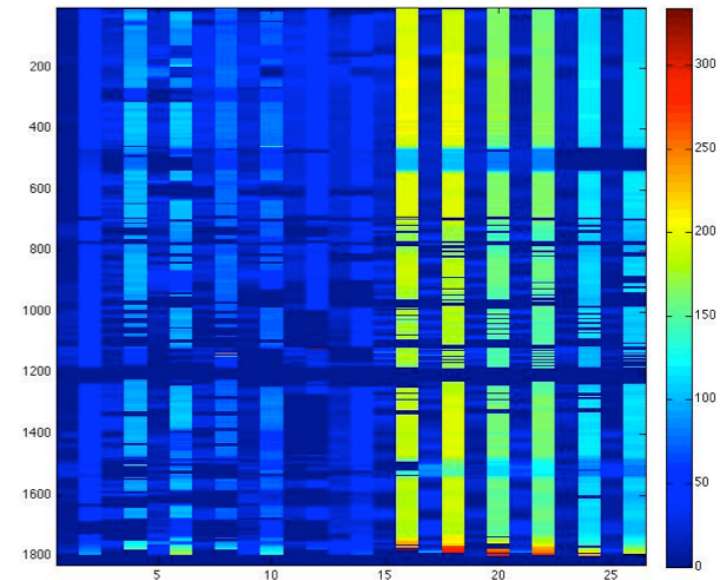
$$\text{Àrea bounding box} = (X_{max} - X_{min}) \times (Y_{max} - Y_{min})$$



Anàlisi:

Descriptors

- Què són?
- Fòrmules emprades



Tipus descriptor	Fòrmula utilitzada
1	$ \partial(C(\text{cap}), C(\text{extremitat})) $
2	$\frac{ \partial(C(\text{cap}), C(\text{extremitat})) }{A(\text{bounding box})}$
3	$\frac{ \partial(C(\text{cap}), C(\text{extremitat})) }{A(\text{cap})}$
4	$\frac{ \partial(C(\text{cap}), C(\text{extremitat})) }{\sum(\text{\`arees extremitats})}$

∂ : Distància euclidiana

$$\sqrt{(x - y)^2} = |x - y|$$

C: Centroide

```
s = regionprops(BW, 'centroid');
centroids = cat(1, s.Centroid);
```

A: Àrea

```
Sum(Sum(màscara))
```

1 Introducció

2 Objectius

3 Base de dades

4 Anàlisi

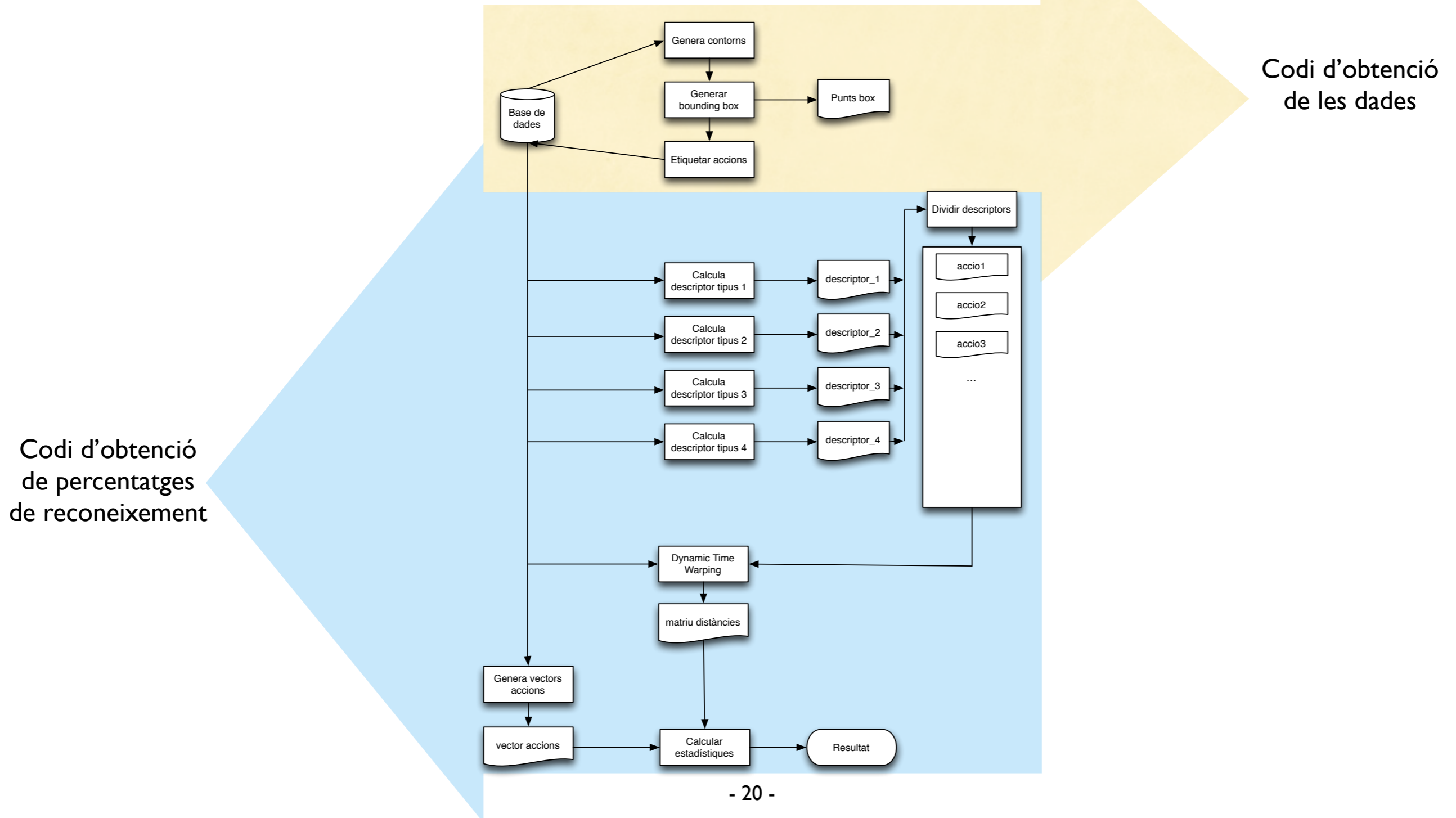
5 Disseny i implementació

6 Resultats

7 Conclusions

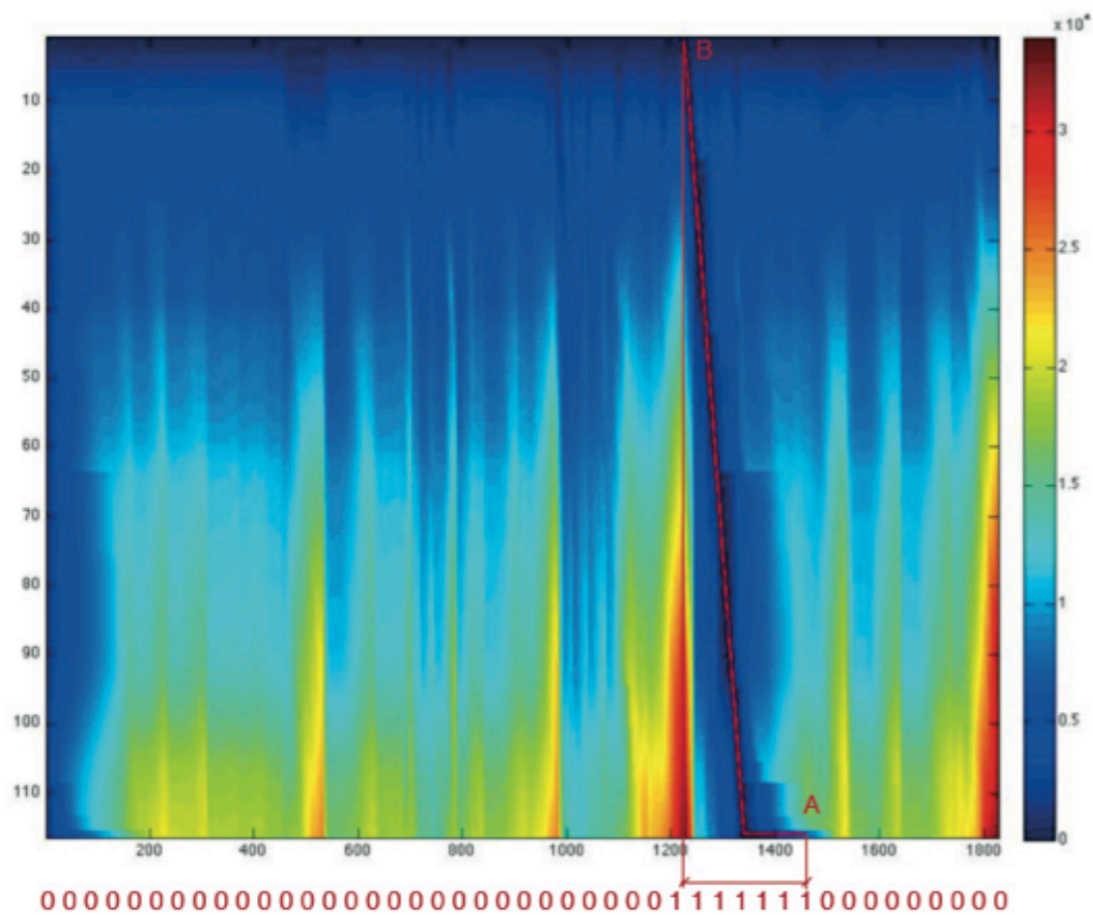
Disseny i implementació:

Diagrama de flux dels diversos mòduls del projecte



Disseny i implementació:

Matriu de distàncies amb camí identificat:



Vector obtingut amb el recorregut del camí

Taula amb el camí:

Valor de tall: 9,5

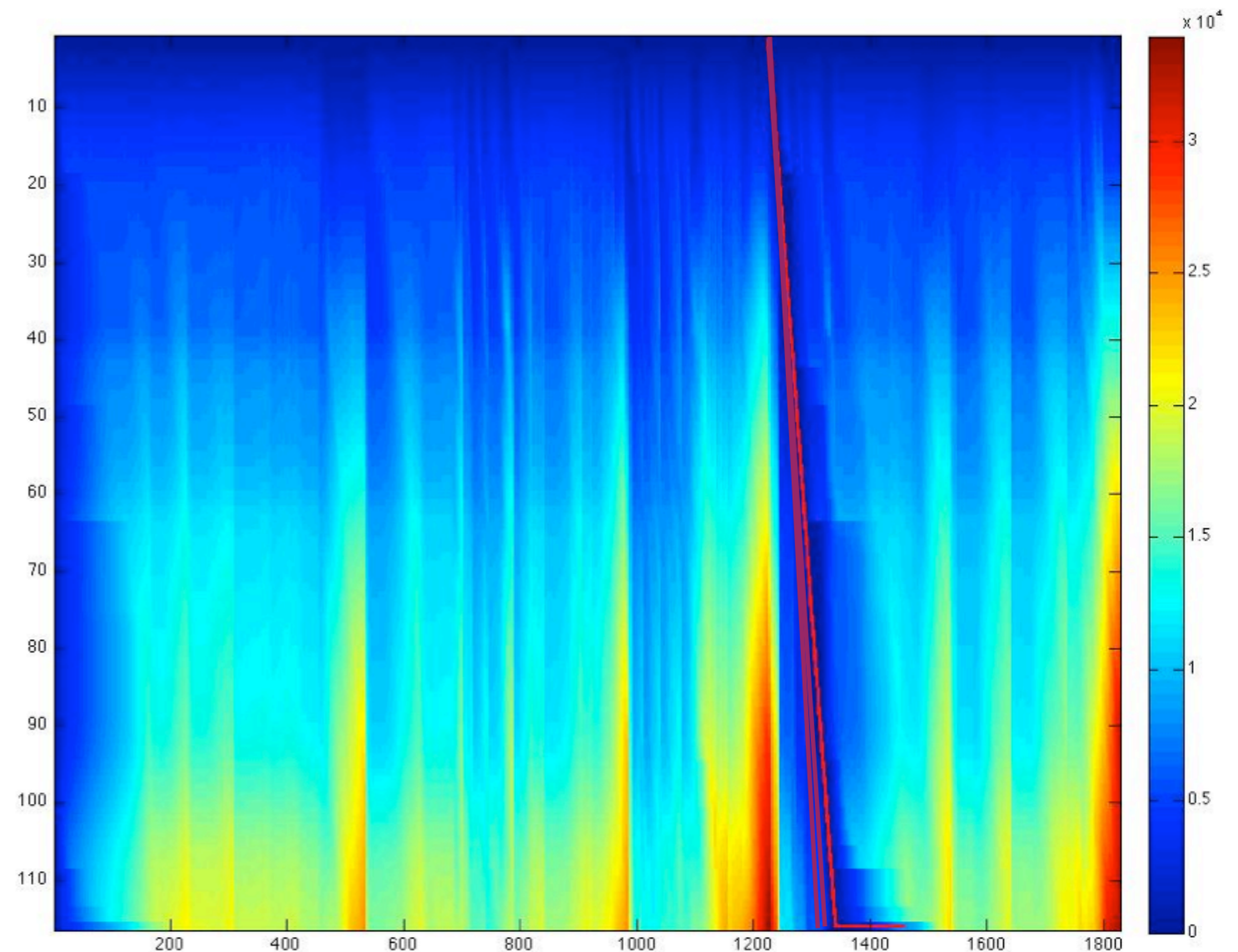
	Dada 1	Dada 2	Dada 3	Dada 4	Dada 5	Dada 6
	0	0	0	0	0	0
Dada Acció1	1,1023	2,977	3,0002	6,3425	9,5436	13,225
Dada Acció2	4,7645	4,2345	3,5267	5,1231	11,123	14,133
Dada Acció3	3,4345	8,3275	4,7563	4,076	10,554	5,2345
Dada Acció4	7,2346	9,4742	6,1298	9,3424	9,562	12,456
Dada Acció5	13,765	14,956	10,035	12,348	9,562	9,384

Disseny i implementació:

Obtenció del valor de tall òptim:

Què necessitem?

- ▶ Recorre la última fila matriu de distàncies fins a trobar el valor de tall especificat
- ▶ Realitzar el camí de tornada per cada punt que s'ha trobat
- ▶ Per cada valor es genera un vector d'accions que és comparat amb l'obtingut manualment
- ▶ Es fa el càlcul estadístic per aquest valor



Disseny i implementació:

Càlcul d'estadístiques:

Mètodes

- ▶ Cas 1: una mostra
 - ▶ Resultat únic
- ▶ Cas 2: múltiples mostres
 - ▶ Mitja de varis resultats

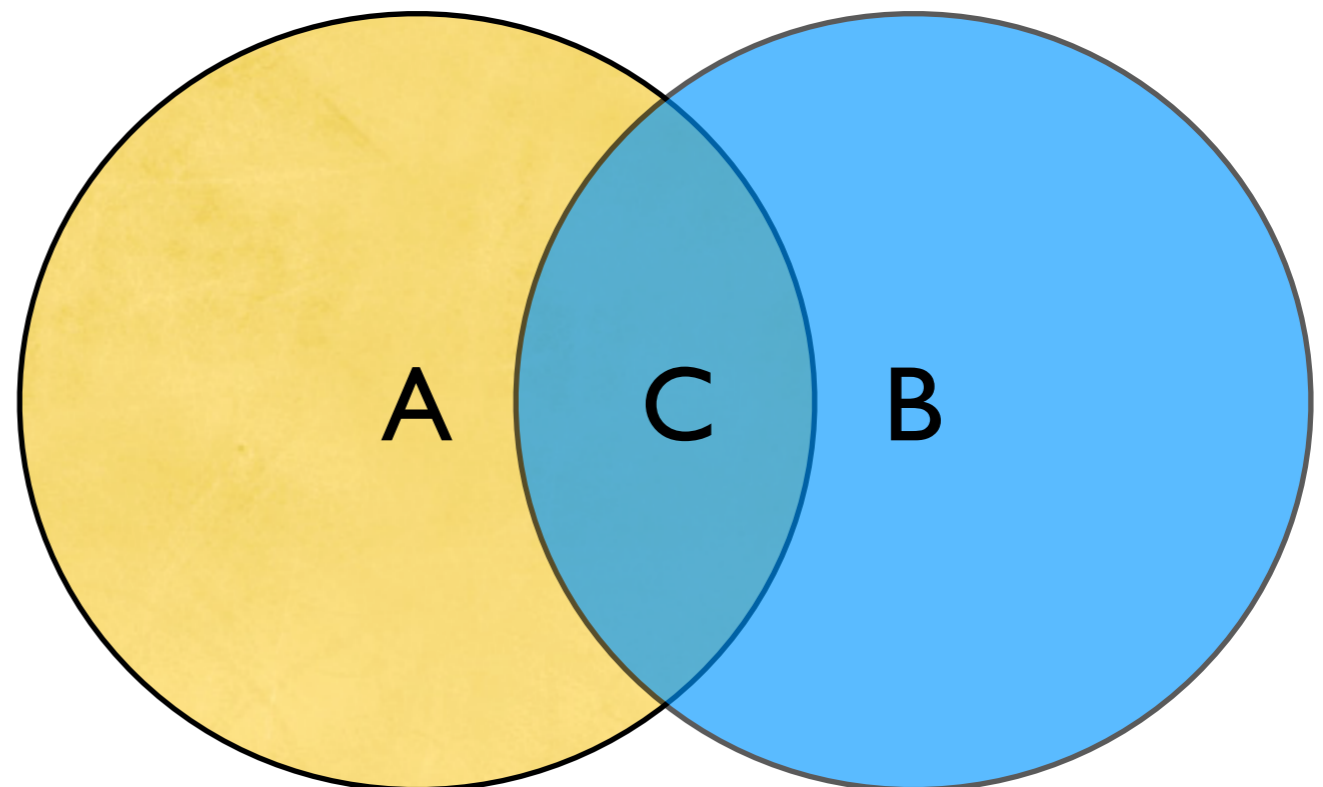
$A =$ Valor del vector d'accions extret manualment

$B =$ Valor del vector d'accions obtingut

$C = A \cap B$

Fòrmula:

$$\frac{C}{A \cup B}$$



1 Introducció

2 Objectius

3 Base de dades

4 Anàlisi

5 Disseny i implementació

6 Resultats

7 Conclusions

Resultats:

Llenguatge emprat per l'elaboració dels scripts: MATLAB

Dades emprades en la validació

- ▶ Actors participants: 6
- ▶ Accions a reconèixer: 11
- ▶ Màscares totals: 51171
- ▶ Seqüències completes: 2
- ▶ Matrius de descriptors a reconèixer: 99
- ▶ Vectors d'accions: 66

Scripts realitzats

- ▶ Base de dades:
 - ▶ Generar contorns
 - ▶ Generar Bounding Box
 - ▶ Etiqueta accions
- ▶ Càlcul de descriptors i estadístiques:
 - ▶ Calcular descriptors
 - ▶ Dynamic Time Warping
 - ▶ Generar vectors d'accions
 - ▶ Calcular estadístiques

Resultats:

Percentatges de reconeixement emprant el tipus 1:

Mitjana d'encert = 24,936%

Individu / acció	Saludar	Senyalar	Aplaudir	Saltar	Ajupir	Caminar	Córrer	Abraçar	Donar mà	Donar petons	Barallar
Vídeo 1 actor 1	9,0909	14,7914	16,4728	3,3273	64,8855	38,1473	15,9788	44,2722	30,3222	16,7621	23,4571
Vídeo 1 actor 2	-	-	-	-	-	40,2014	-	26,3002	17,7384	24,4689	21,4356
Vídeo 1 actor 3	-	-	-	-	-	63,1234	-	-	-	-	-
Vídeo 1 actor 4	-	-	-	-	-	53,1840	-	-	-	-	-
Vídeo 2 actor 1	22,9730	39,8844	14,4958	13,7143	17,9775	29,3889	26,5246	29,1266	10,1455	8,3278	14,3187
Vídeo 2 actor 2	-	-	-	-	-	24,7517	-	21,3463	22,4168	16,3341	12,1431

Percentatges de reconeixement emprant el tipus 2:

Mitjana d'encert = 10,607%

Individu / acció	Saludar	Senyalar	Aplaudir	Saltar	Ajupir	Caminar	Córrer	Abraçar	Donar mà	Donar petons	Barallar
Vídeo 1 actor 1	10,4265	3,2491	3,2343	3,6738	12,5690	20,1753	2,9174	14,2602	9,1457	8,2403	8,1457
Vídeo 1 actor 2	-	-	-	-	-	17,6281	-	8,4588	13,5139	9,5741	3,4571
Vídeo 1 actor 3	-	-	-	-	-	15,4526	-	-	-	-	-
Vídeo 1 actor 4	-	-	-	-	-	18,2214	-	-	-	-	-
Vídeo 2 actor 1	12,3064	2,7675	4,1250	9,2231	9,7780	12,8733	13,6645	19,1376	8,2346	8,4302	8,0230
Vídeo 2 actor 2	-	-	-	-	-	16,1292	-	12,580	18,2300	7,1470	15,6512

Resultats:

Percentatges de reconeixement emprant el tipus 3:

Mitjana d'encert = 27,694%

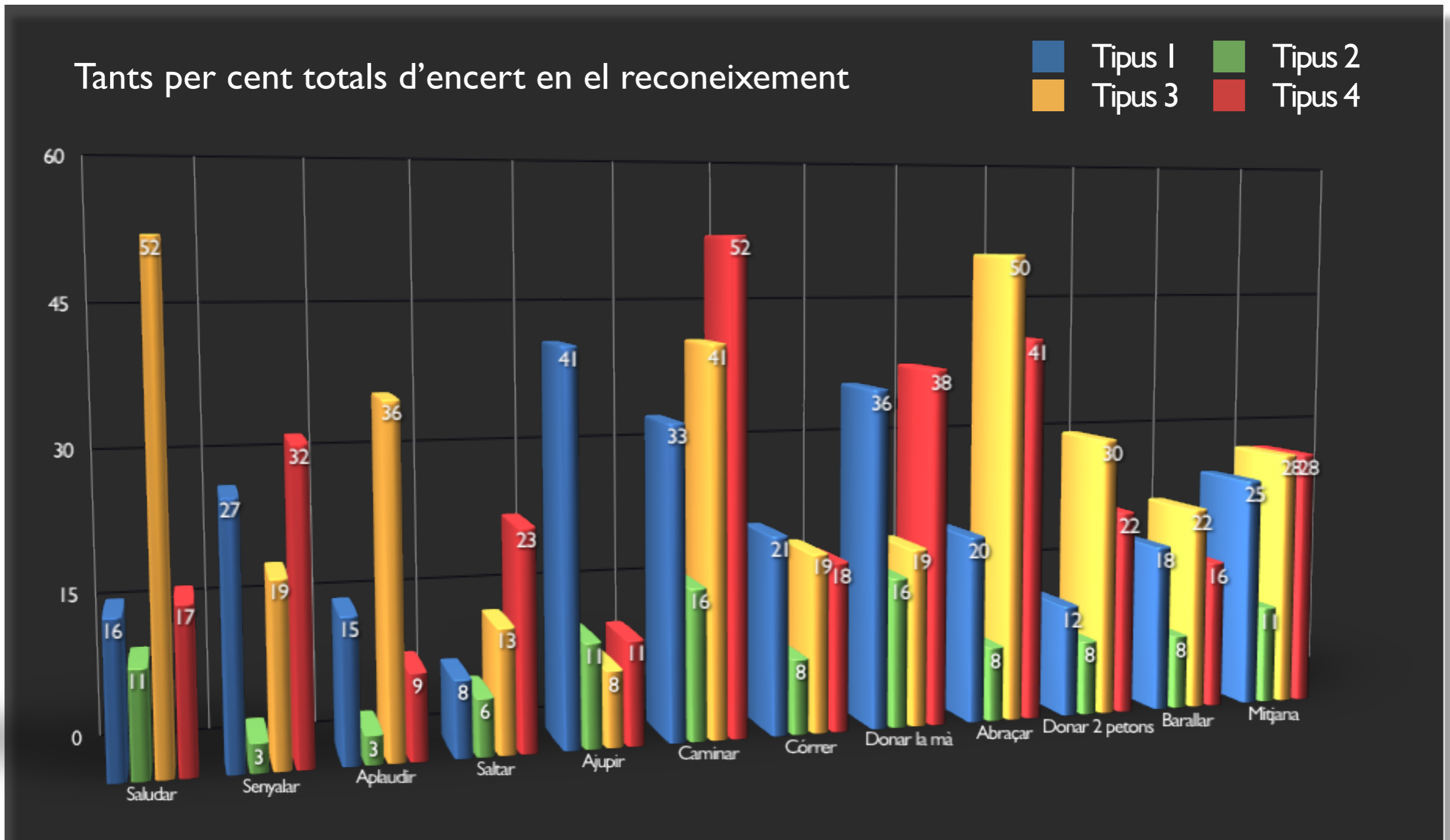
Individu / acció	Saludar	Senyalar	Aplaudir	Saltar	Ajupir	Caminar	Córrer	Abraçar	Donar mà	Donar petons	Barallar
Vídeo 1 actor 1	54,4304	19,2488	43,1373	11	6,6779	49,1629	20,1839	30,1376	48,4677	36,7332	30,4218
Vídeo 1 actor 2	-	-	-	-	-	60,1234	-	12,3342	22,5625	26,3451	12,6398
Vídeo 1 actor 3	-	-	-	-	-	40,1299	-	-	-	-	-
Vídeo 1 actor 4	-	-	-	-	-	27,1538	-	-	-	-	-
Vídeo 2 actor 1	50,6024	20,4348	30,4348	15,7377	9,7654	33,1457	17,9141	9,1014	53,3112	24,7913	13,7456
Vídeo 2 actor 2	-	-	-	-	-	26,7685	-	16,7991	40,6731	8,3426	30,1225

Percentatges de reconeixement emprant el tipus 4:

Mitjana d'encert = 27,708%

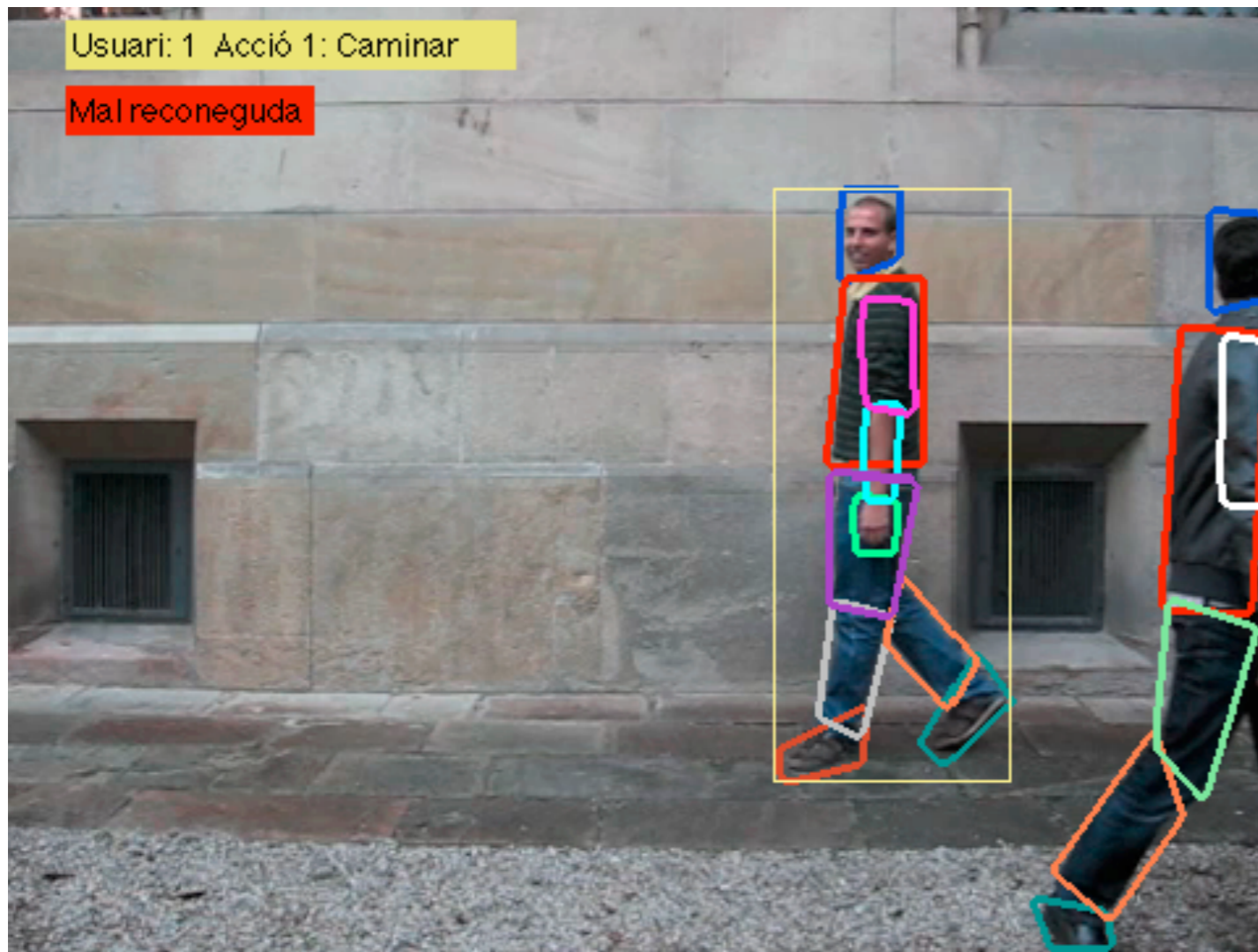
Individu / acció	Saludar	Senyalar	Aplaudir	Saltar	Ajupir	Caminar	Córrer	Abraçar	Donar mà	Donar petons	Barallar
Vídeo 1 actor 1	17,9688	23,2955	5,7579	6,4252	17,7734	67,4582	21,1644	57,2281	60,3001	24,7126	18,4500
Vídeo 1 actor 2	-	-	-	-	-	45,2374	-	49,3780	12,2892	32,4753	20,1412
Vídeo 1 actor 3	-	-	-	-	-	42,1346	-	-	-	-	-
Vídeo 1 actor 4	-	-	-	-	-	50,6788	-	-	-	-	-
Vídeo 2 actor 1	16,4062	40,7407	12,3693	39,7306	6,1657	38,4137	16,5517	19,0912	22	19,4208	14,7183
Vídeo 2 actor 2	-	-	-	-	-	23,1435	-	55,1347	30,4584	25,3870	11,4715

Resultats:



Resultats:

Vídeo amb els resultats obtinguts:



	Acció ben reconeguda
	Acció mal reconeguda

1 Introducció

2 Objectius

3 Base de dades

4 Anàlisi

5 Disseny i implementació

6 Resultats

7 Conclusions

Conclusions:

Objectius complerts

- ✓ ▶ Realització d'una base de dades amb suficients mostres per dur a terme els experiments
- ✓ ▶ Definir 4 mètodes d'obtenció de descriptors
- ✓ ▶ Testejar els descriptors obtinguts amb l'algorisme DTW i comprovar quin ens dóna millors percentatges d'encert
- ✓ ▶ Codi modular i BD reutilitzables per a futurs projectes
- ✓ ▶ Millors resultats gràcies a la normalització del descriptor per l'àrea d'algunes de les seves parts

Conclusions:

Treball futur

- ▶ Afegir característiques més discriminants alhora d'obtenir els descriptors, per tal d'obtenir millors resultats
 - ▶ Millorar l'algorisme del DTW per que les parts que tinguin més participació en una acció tinguin més pes
 - ▶ Emprar altres tipus de descriptors com HOG, SIFT, etc.
- ▶ Comparar amb altres mètodes de reconeixement, com per exemple el Hidden Markov Model (HMM), que es basa en probabilitats en comptes de distàncies

Preguntes