

Aplicación de la lógica difusa a la identificación forense

MÁSTER UNIVERSITARIO EN INGENIERÍA INFORMÁTICA

Carlos Morales Aguilera

Carlos Santiago Sánchez Muñoz

4 de Enero de 2020

Inteligencia Computacional



**UNIVERSIDAD
DE GRANADA**

E.T.S. de Ingenierías Informática y de Telecomunicación



Introducción

Estado del arte

Introducción

Uso de técnicas

Métodos y herramientas

Modelo

Superposición cráneo-cara

Decisión

Resultados

Etapa de superposición

Etapa de toma de decisiones

Etapa de toma de decisiones

Experimentos realizados

Conclusiones

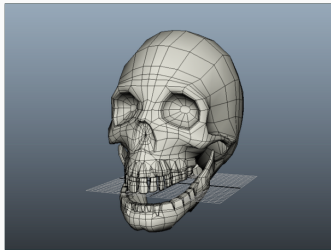
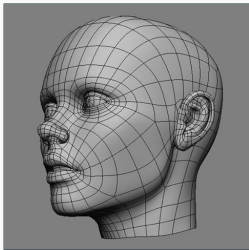
Introducción

Objetivos:

- Analizar el problema y posibles mejoras.
- Entender por qué aplicar lógica difusa.
- Comprender la mejora que se obtiene.
- Comentar nuevas vías de desarrollo.

Escenario

- Técnicas principales no son suficiente.
- No se poseen medidas exactas.



Estado del arte

- Proyecto MEPROCS: tenía como objetivo proponer una Marco de la UE para permitir la amplia aplicación de la técnica CFS en escenarios prácticos de identificación forense.
- Cuestionario enviado a 600 científicos durante 3 meses. Se reciben 97 respuestas de las cuales 45 afirman haber usado CFS.
- Encuesta estuvo compuesta por 16 preguntas relacionadas con el número de casos emprendidos en los últimos 30 años:
 - Escenarios de identificación.
 - número de horas empleadas en la realización de CFS.
 - resultados de identificación.
 - Materiales, herramientas y técnicas utilizadas.
 - Principales problemas enfrentados en la aplicación del CFS.

- De la encuesta MEPROCS, se reportaron 3854 casos de CFS: 2744 (71,2 %) fueron presentados en un tribunal de justicia y 1246 (32,3 %) de ellos resultaron en una identificación positiva.
- La mayoría de los casos (2489) no fueron clasificados en ninguna de las categorías sugeridas en la encuesta.
- Más del 50 % de los investigadores han implementado CFS en solo un pequeño número de casos (1 a 10 casos durante 10 años) mientras que los restantes informaron un mayor número de casos.

⇒ Se van a analizar en profundidad las respuestas de los participantes.

Distribución geográfica: India es el país en el que se han notificado la mayoría de los casos (1471), seguido de Perú (980). En España solo 31.

Uso de técnicas: Escenarios de identificación

En función del número de candidatos y coincidencias:

PIOI: Presunta identidad de un solo candidato.

PICL: Identidad de un individuo dentro de una lista cerrada de candidatos.

UICL: Identidad desconocida de un individuo dentro de una lista cerrada.

UIOL: Identidad desconocida de un individuo dentro de una lista abierta.

CNC: Casos no clasificados por los respondedores.

Escen.	Id. positiva	Exclusión	Indeter.	No clasif.	Total
PIOI	451 (11.7 %)	33 (0.9 %)	61 (1.6 %)	–	545 (14.1 %)
PICL	309 (8.0 %)	4 (0.1 %)	19 (0.5 %)	–	332 (8.6 %)
UICL	336 (8.7)	–	2 (0.1 %)	–	338 (8.8 %)
UIOL	150 (3.9 %)	–	–	–	150 (3.9 %)
CNC	–	–	–	2489 (64.6 %)	2489 (64.6 %)
Total	1246 (32.3 %)	37 (1.0 %)	82 (2.1 %)	2489(64.6 %)	3854(100 %)

Cuadro 1: Uso de CFS reportado por 45 investigadores

Uso de técnicas: Escenarios de identificación

En función de la naturaleza de la investigación:

Desastres masivos: eventos naturales o provocados por el hombre, que causan la muerte o lesiones a un gran número de personas.

Terrorismo: actividad humana diseñada para causar daño, utilizando violencia premeditada con fines políticos o religiosos.

Personas desaparecidas: según el RCIC son aquellas personas de las que sus familiares no tienen noticias y/o que han sido reportadas como desaparecidas.

Fosas comunes: lugares donde se entierra un número indeterminado de personas como consecuencia de la guerra o desastres masivos.

Nature of the investigation	Cases
Terrorist attack	15 (0.4 %)
Missing persons	458 (11.9 %)
Mass casualty	130 (3.4 %)
Mass grave	87 (2.2 %)
Other	3164 (82.1 %)
Total	3854 (100.0 %)

Cuadro 2: Clasificación según la naturaleza de la investigación

Se pidió a los investigadores que informaran sobre el método (manual, semiautomático o completamente automático) y el tipo de técnica de CFS utilizada, así como la fuente del material (fotografía, imagen de vídeo o digital) para la superposición.

Técnicas:

	Superposición fotográfica	Superposición de vídeo	Superposición asistida por computadora	Respuestas totales
Manual	20	3	8	31
Semiautomatic	10	3	5	18
Automatic	11	1	2	14
Total	41	7	15	63

Cuadro 3: Distribución de técnicas reportadas

Herramientas necesarias durante la implementación de la técnica:

Respuesta	Porcentaje	Cantidad
Rayos X	12	4
CTs	12	4
Escáner 3D	12	4
Sistema de vídeo	15	6
Banco de calaveras	15	5
Photoshop	61	21
Otros	39	13
	Total de repuestas	34

Cuadro 4: Herramientas implementadas en técnica

Landmark: punto craneofacial en donde se va a realizar la medición.

Response	Frequency	Count
Glabella	38 %	11
Nasion	38 %	11
Rhinion	17 %	5
Nasospinale-Subnasale	28 %	8
Alare-Alare	31 %	9
Subspinale-Superior labial sulcus	24 %	7
Prosthion-Upper lip border (Labiale superius)	34 %	10
Infradentale superius-Lower lip border	17 %	5
Supramentale-Labiamentale	14 %	4
Gonion	38 %	11
Zygion	31 %	9
Orbitale-Infraorbital	17 %	5
Porion	17 %	5
Dacryon-Endocanthion	34 %	10
Gnathion-Menton	21 %	6
Gnathion	38 %	11
Pogonion-Mental	14 %	4
Pogonion	14 %	4
Prosthion	24 %	7
Ectoconchion	38 %	11
Frontozygomatic suture-Ectocanthion	21 %	6
Other	62 %	18

Literatura

- Las estadísticas obtenidas en la encuesta fueron similares a los datos encontrados en la literatura. Clasificación de acuerdo a la naturaleza de la investigación, escenarios de identificación y resultados.
- 26 artículos habían abordado la naturaleza de la investigación y, posteriormente, habían sido sometidos a escrutinio judicial.
- Una gran cantidad de los casos incluyeron identificación por superposición, corroborada por otras técnicas de identificación (ADN).
- Cuando no se pudo obtener corroboración, se utilizó como método de identificación la superposición, junto con el perfil antropológico.
- Identificación de personajes históricos como Mozart.

Conclusión

- CFS se considera una técnica de identificación confiable y eficiente, aunque se implementa junto con otro método en muchos casos.

Modelo

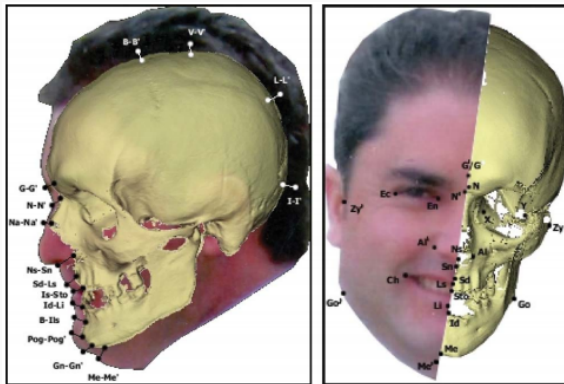
Modelos previos

- Primeros modelos superposición calavera sobre fotografía.
- Minimización distancia puntos de referencia craneofaciales.
- Modelo anatómicamente incorrecto.
- Se necesita un modelo que considere la piel.

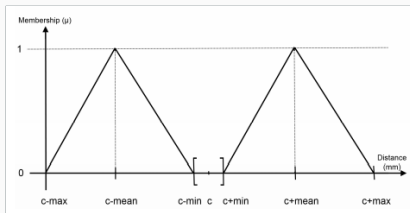
Superposición cráneo-cara

Primera propuesta

- Se ubica una esfera en el punto de referencia craneal.
- Se calculan *max*, *min* y *mean* en 3D.



Representación una dimensión en lógica difusa



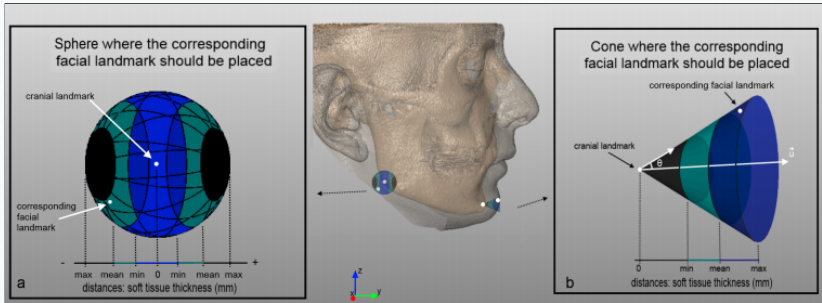
Fórmula

$$\tilde{B}_p = \begin{cases} 1 - \frac{|p - c + mean|}{max - mean}, & \text{if } c - max \leq p < c - mean \\ 1 - \frac{|p - c + mean|}{mean - min}, & \text{if } c - mean \leq p \leq c - min \\ 1 - \frac{|p - c - mean|}{mean - min}, & \text{if } c + min \leq p \leq c + mean \\ 1 - \frac{|p - c - mean|}{max - mean}, & \text{if } c + mean < p \leq c + max \\ 0, & \text{otherwise} \end{cases}$$

Superposición cráneo-cara

Enfoque no válido

- Anatómicamente incorrecto.
- Inconsistencia entre puntos.



Se define el **vector normal** como $\vec{v} = (x_n, y_n, z_n)$.

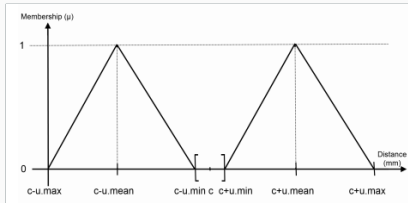
Por otro lado se define la **magnitud** de \vec{v} como

$$||\vec{v}|| = \sqrt{x_n^2 + y_n^2 + z_n^2}.$$

El **vector unitario** de \vec{v} se define entonces de la manera:

$$\vec{u} = \left(\frac{x_n}{||\vec{v}||}, \frac{y_n}{||\vec{v}||}, \frac{z_n}{||\vec{v}||} \right) = (u_x, u_y, u_z)$$

Representación una dimensión en lógica difusa



Fórmula

$$\tilde{B}_p = \begin{cases} 1 - \frac{|p - c + u \cdot \text{mean}|}{u(\text{max} - \text{mean})}, & \text{if } c - u \cdot \text{max} \leq p \leq c - u \cdot \text{mean} \\ 1 - \frac{|p - c + u \cdot \text{mean}|}{u(\text{mean} - \text{min})}, & \text{if } c - u \cdot \text{mean} \leq p \leq c - u \cdot \text{min} \\ 1 - \frac{|p - c - u \cdot \text{mean}|}{u(\text{mean} - \text{min})}, & \text{if } c + u \cdot \text{min} \leq p \leq c + u \cdot \text{mean} \\ 1 - \frac{|p - c - u \cdot \text{mean}|}{u(\text{max} - \text{mean})}, & \text{if } c + u \cdot \text{mean} < p \leq c + u \cdot \text{max} \\ 0, & \text{otherwise} \end{cases}$$

Métrica de distancia

- Distancia entre conjuntos difusos:

$$d'(x, B) = \frac{\sum_y d(x, y) \cdot B(y)}{\sum_y B(y)}$$

- Distancia entre punto y conj. difuso:

$$d''(A, B) = \frac{\sum_x \sum_y d(x, y) t[A(x), B(y)]}{\sum_x \sum_y t[A(x), B(y)]}$$

- Función fitness *Fuzzy Mean Error*:

$$FME = \frac{\sum_{i=1}^{N_{crisp}} (d'(x_i, f(C^i))) + \sum_{j=1}^{N_{fuzzy}} (d''(F^j, f(C^j)))}{N}$$

Métrica de distancia sin puntos exactos

$$D_{FMD} = \frac{\bigcup_{\alpha \in (0,1]} \sum_{i=1}^N D(F^i, f(C^i))_{\alpha}}{N}$$

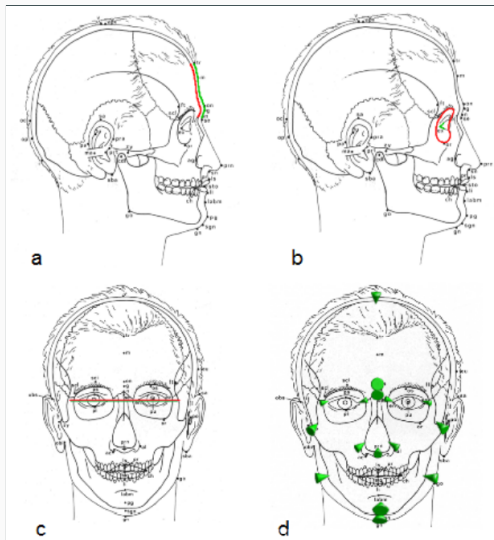
Por lo tanto, el número obtenido se aplica a la función de fitness de la forma:

$$FME = \frac{\sum_{\alpha_{min}}^{\alpha_{max}} \alpha \cdot D_{FMD}(\alpha)}{\sum_{\alpha_{min}}^{\alpha_{max}} D_{FMD}(\alpha)}$$

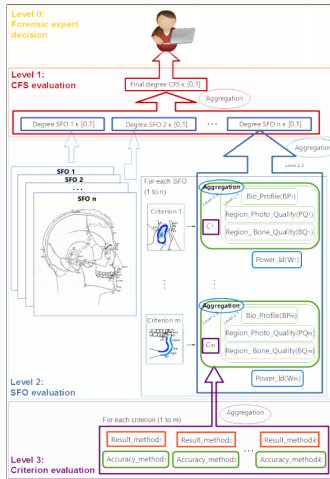
Criterios de satisfacción

- Análisis consistencia anatómica de las curvas morfológicas del cráneo y cara.
- Evaluación consistencia anatómica por correspondencia posicional.
- Localización linear y comparación para análisis de la consistencia anatómica.
- Evaluación de la consistencia del espesor del tejido suavizado entre los puntos de referencia craneofaciales.

Criterios de satisfacción

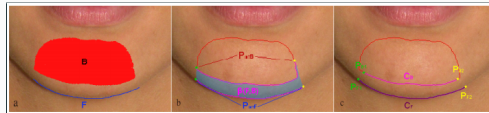


Sistema de soporte de decisiones



Nivel 3: Evaluación de criterios

- Modelado de la consistencia anatómica de los contornos craneofaciales.
- Medición de la posición relativa de dos regiones para estudiar consistencia anatómica.
- Modelado de la consistencia anatómica por localización linear y comparación.



- Modelado de la consistencia anatómica de una posición de dos regiones óseas.

Nivel 2: Evaluación SFO

Nivel 2.1: Función de agregación para combinar evaluaciones de calidad de material y perfil biológico.

$$O_{Level2,1}(PQ_m, BQ_m, BP_m) = \{min, prod, mean\}$$

Nivel 2: Evaluación SFO

Nivel 2.2: Función de agregación para combinar el grado de correspondencia y el valor de incertidumbre del nivel 2.1

$$O_{Level2,2}(C_m, Output_{Level2,1}) = \{wam, wgm\}$$

Nivel 2: Evaluación SFO

Nivel 2.3: Función de agregación para combinar el poder de identificación el grado del nivel 2.2

$$g_{\lambda}(A \cup B) = g_{\lambda}(A) + g_{\lambda}(B) + \lambda g_{\lambda}(A) \mu(B)$$

- *Integral discreta de Choquet:*

$$C_g(x) = \sum_{i=1}^n [x_{(i)} - x_{(i-1)}] g(H_{(i)})$$

- *Integral de Sugeno:* $S_g(x) = \max_{i=1, \dots, n} \min(x_{(i)}, g(H_{(i)}))$

Nivel 2: Evaluación SFO

Nivel 2.3: Función de agregación para combinar el poder de identificación el grado del nivel 2.2

$$O_{Level2,3}(Output_{Level2,2_i}, PowerID(W_i)) = \{wam, wgm, choq, sug\}$$

Nivel 3: Evaluación CFS

Agregación de los diferentes *SFOs* obtenidos en un único resultado:

$$Agg_Nreg = \frac{\sum_{n=1}^N (D_SFO_n \cdot Nreg_n)}{\sum_{n=1}^N Nreg_n}$$

Resultados

Etapas de superposición: modelado del estudio de distancias difusas e incertidumbre de coincidencia de puntos de referencia

- El procedimiento automático, basado en CV y SC, realiza la SFO de la forma más natural replicando el escenario original en el que se tomó la fotografía.
- El procedimiento técnico automatiza este enfoque propenso a errores.
- La orientación del cráneo y su ajuste de tamaño se puede realizar de forma automática simplificando el trabajo de los expertos.
- Reduce el tiempo para obtener una superposición válida y lo automatiza.
- La evaluación ha usado 18 estudios de caso. Se logra un emparejamiento competitivo entre pares de puntos de referencia craneales y faciales.
- RCGA-c-45 supera significativamente a los demás métodos SFO.
- RCGA-c-30 también logra una buena clasificación.

Etapas de superposición: modelado del estudio de distancias difusas e incertidumbre de coincidencia de puntos de referencia

- Por lo tanto, las nuevas propuestas obtienen mejores resultados que los enfoques que no consideran el grosor de los tejidos blandos faciales.
- No obstante, aunque se obtienen resultados buenos aún se identifica un error residual al evaluar un resultado de SFO contra su superposición.
- La elección de una buena métrica de distancia es crucial para nuestro método, ya que puede afectar significativamente a la calidad de las soluciones.
- El enfoque SFO presenta el mejor rendimiento utilizando la distancia media ponderada en la mayoría de los casos.
- Resultados más precisos y robustos que las otras implementaciones.
- Esta distancia considera información espacial relacionada con la distancia, pero no la posición relativa direccional entre ellos. La relación entre los puntos de referencia craneales y faciales debe satisfacer un criterio de posición específico en función de la pose \Rightarrow Trabajo futuro.

Etapas de toma de decisiones: desarrollo jerárquico del sistema de apoyo a las decisiones

Correspondencia anatómica cráneo-cara en tres niveles: evaluación de criterios, evaluación de SFO y evaluación de CFS.

Clasificamos las fuentes de incertidumbre y los grados de confianza involucrados en este proceso en relación con el hueso, la imagen, las superposiciones del cráneo y la cara, los aspectos morfológicos y los métodos utilizados.

Etapas de evaluación del criterio

- 1) La coherencia entre los contornos craneales y faciales.
- 2) Consistencia anatómica por relación posicional.
- 3) Ubicación y comparación de la línea.
- 4) Consistencia del grosor del tejido blando entre los puntos de referencia.

Primera familia Mentón y contorno craneal.

Segunda familia Órbita craneal y posición central del globo ocular, borde superior y posición de la oreja...

Tercera familia Paralelismo de líneas de ectocantión, frontales y centrales.

Etapas de toma de decisiones: desarrollo jerárquico del sistema de apoyo a las decisiones

La evaluación de correspondencia craneofacial comprende todo el rostro. Se han considerado de una a ocho regiones y de cero a 18 líneas interdentes. Utilizando técnicas de CV, se han estudiado diferentes métodos para medir la consistencia entre cada región ósea y facial.

Etapas de evaluación de la SFO

Estudio experimental para elegir función de agregación.

Primer subnivel: agrega las fuentes de incertidumbre del hueso y la imagen, y la variabilidad del perfil biológico. Se propone utilizar *min*, *prod* y *media*.

Segundo subnivel: integrar esta agregación con el grado de coincidencia del cráneo y la cara. Uso de las medias aritmética y geométrica: *wam* y *wgm*.

Tercer subnivel: grado de correspondencia craneofacial SFO agregando los diferentes valores para todas las regiones.

A la vista de los resultados obtenidos en los experimentos, concluimos que el mejor desempeño actual en el nivel 2 se obtiene utilizando la combinación de funciones de agregación *mean-wgm-wgm*.

Etapa de toma de decisiones: desarrollo jerárquico del sistema de apoyo a las decisiones

El mejor comportamiento se logra cuando se agregan los diferentes grados de SFO de un mismo caso ponderando por el número de regiones consideradas.

DSS clasificó correctamente un número relativamente alto de casos positivos (13 de 16, es decir, 81,25 %) y negativos (186 de 224, es decir, 83,03 %).

Se puede concluir que el sistema actual es apto para:

- Casos de filtrado (clasificación-listado): en el 62,5 % de las ocasiones el caso positivo está en primera posición. Además el caso positivo se incluyó en el 60 % de los candidatos mejor clasificados con un 100 % de probabilidad y dentro del 27 % con un 92 % de probabilidad.
- Establecimiento de la exclusión: la capacidad para determinar una identidad negativa se realizó con un 97,4 % de precisión (de la tasa de TN sobre el conjunto de prueba y 83,03 % sobre el de entrenamiento).

El DSS diseñado puede considerarse la **primera herramienta automática para clasificar parejas de rostros y cráneos con una precisión similar a la de los expertos forenses** con mejor desempeño.

Exp. 1: Correspondencia de puntos de referencia con incertidumbre

Calcular la distancia entre un punto de referencia craneal y su facial.

30 ejecuciones independientes para comparar la robustez.

Case, pose	pl	il	CCGA	RCGA	RCGA-s	RCGA-c-0	RCGA-c-10	RCGA-c-30	RCGA-c-45
1,f	7	7	4.565	2.750	2.727	4.116	4.330	3.078	3.164
1,l	5	4	16.588	7.406	6.938	7.400	6.693	5.788	5.828
2,f	8	5	4.906	3.690	3.486	3.823	4.212	3.337	3.331
2,l	3	2	8.299	8.605	8.871	8.871	8.871	5.286	3.688
3,f	8	7	3.815	3.629	3.664	6.594	6.594	3.007	2.941
3,l	4	4	9.367	10.643	10.815	11.073	11.073	7.604	7.001
4,f	7	6	4.635	3.647	3.505	3.356	3.356	3.223	3.045
4,l	4	3	14.027	14.327	13.494	13.847	13.847	12.446	12.171
5,f	10	6	2.996	2.436	2.453	2.204	2.406	2.641	2.600
5,l	5	4	8.545	6.865	6.097	6.846	6.846	3.715	2.825
6,f	8	7	4.654	3.784	3.644	2.618	3.076	2.878	2.921
6,l	3	4	16.486	12.959	11.665	2.204	11.126	11.009	10.626
7,f	10	5	4.253	3.639	3.617	4.695	4.695	4.100	3.714
7,l	5	4	10.212	12.663	10.909	9.649	9.649	9.729	10.345
8,f	9	6	5.269	4.409	3.579	3.492	3.492	2.856	2.882
8,l	4	4	6.840	8.555	7.812	19.935	19.935	6.262	5.412
9,f	10	4	6.174	5.541	5.757	6.660	6.660	5.071	4.887
9,l	3	5	10.210	11.701	11.700	19.852	19.852	10.391	9.509
Average (mm)			7.778	6.976	6.548	7.759	8.416	5.585	5.281

Imagen 1: Resultados frente al estado del arte (30 ejecuciones y 9 casos)

Modelos que utilizan el modelo del cono (c – *angulo*) son los que obtienen los mejores resultados y los del modelo esférico (s) también son prometedores.

Exp. 1: Correspondencia de puntos de referencia con incertidumbre

Se han utilizado cuatro variantes del enfoque cónico, en los que se consideran cuatro amplitudes diferentes del ángulo del cono, considerando que el que mejor resultados ha obtenido es el modelo con $\pm 45^\circ$.

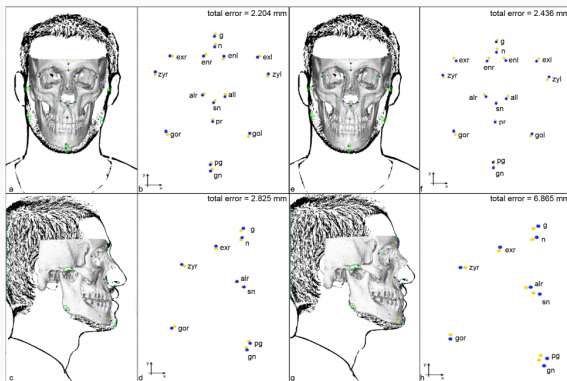


Imagen 2: Las imágenes a y c corresponden al enfoque *RCGA-c-45* mientras que las imágenes e y g corresponden al enfoque *RCGA* clásico

Exp. 2: Distancias difusas

La elección de **Weighted Mean distance** era arbitraria.

Mismos parámetros y ejemplos pero con modelo *RCGA-c-45*.

Error (en mm) y desviación para cada caso de estudio.

Case, post	pl	il	Weighted Mean	Nearest Point	Nearest Point E1	Nearest Point E2	Mean α -cuts (Nearest point)	Mean α -cuts (Sym. mean)	Mean α -cuts (Hausdorff)	Rosenfeld
1,f	7	7	3.17±0.31	3.28±0.65	2.97±0.81	2.07±0.87	2.51±1.00	1.98±0.89	3.03±0.53	3.04±0.91
1,l	5	4	5.83±2.59	6.68±2.84	8.13±3.25	7.63±3.24	7.73±3.17	7.78±3.14	7.42±3.85	9.39±1.47
2,f	8	5	3.33±0.45	3.18±1.42	3.14±1.88	2.15±0.42	2.45±1.12	3.27±1.78	4.24±1.18	3.90±0.36
2,l	3	2	3.69±0.23	9.25±2.14	8.24±1.90	8.19±2.76	7.64±2.22	6.52±2.71	6.48±2.62	8.57±2.78
3,f	8	7	2.94±0.23	4.81±0.69	4.55±0.56	4.46±0.56	4.35±0.65	4.17±0.67	3.66±0.65	4.64±0.77
3,l	4	4	7.00±0.09	7.20±0.56	7.35±0.81	6.94±0.39	7.05±0.51	6.98±0.37	7.54±0.59	9.59±1.50
4,f	7	6	3.05±0.06	4.24±0.18	3.90±0.27	3.90±0.34	4.17±0.40	3.67±0.30	4.10±0.13	4.46±0.67
4,l	4	3	12.17±2.83	11.64±2.71	11.95±2.90	11.93±2.94	10.83±2.51	11.38±2.66	12.01±2.08	12.65±2.92
5,f	10	6	2.60±0.13	5.21±0.65	5.14±0.84	5.34±0.71	5.26±0.66	4.12±0.33	4.42±0.46	7.77±1.66
5,l	5	4	2.83±0.15	5.15±2.33	5.05±2.07	5.69±2.14	6.14±2.15	4.49±1.71	6.85±2.60	8.57±2.28
6,f	8	7	2.92±0.12	4.98±0.37	3.78±0.74	3.98±0.88	3.70±0.54	3.23±0.53	3.97±0.57	2.93±0.88
6,l	3	4	10.63±3.02	12.64±3.69	11.19±4.10	11.88±3.24	9.89±2.74	10.41±3.35	12.05±2.46	11.34±2.38
7,f	10	5	3.71±0.17	3.93±1.64	4.42±2.12	4.12±2.06	3.50±1.77	3.19±0.81	5.05±0.64	5.33±1.09
7,l	5	4	10.35±1.97	10.96±1.21	10.48±1.15	10.36±0.88	9.91±0.93	10.20±0.89	10.95±0.90	10.85±1.33
8,f	9	6	2.88±0.24	4.35±0.49	4.21±0.46	4.39±0.40	4.36±0.21	4.43±0.33	3.56±0.16	5.85±1.35
8,l	4	4	5.41±0.15	8.64±1.96	8.19±1.76	9.15±1.55	8.10±1.77	6.87±1.70	7.40±1.85	10.60±1.93
9,f	10	4	4.89±0.15	5.72±0.44	5.79±0.33	5.74±0.36	5.61±0.40	5.19±0.34	4.54±0.25	5.25±0.72
9,l	3	5	9.51±0.69	11.68±2.84	11.02±2.19	11.22±1.98	11.41±2.06	10.22±1.15	10.53±1.23	13.26±2.31
Average			5.38	6.86	6.64	6.62	6.37	6.01	6.54	7.67

Imagen 3: Resultados de las diferentes distancias (30 ejec. y 9 casos)

Resultados similares y según el caso es mejor una distancia u otra.

Fila final \Rightarrow *Weighted Mean Distance* es la mejor distancia.

Conclusiones

Posibles vías de desarrollo

- Personas con alguna deformidad craneal.
- Personas con problemas de peso.
- Personas con lesiones.

Gracias por su atención