



Trabajo de Fin de Grado

GRADO DE INGENIERIA INFORMÁTICA

**Facultad de Matemáticas
Universidad de Barcelona**

DETECCIÓN DE PERSONAS CON MODELOS ARTICULADOS

Irene Morales Osorio

Tutor: Luis Garrido y Oriol Pujol
Barcelona, 30 de junio de 2016

Índice

| | |
|---|----|
| 1. Introducción y motivación..... | 3 |
| 1.1 Motivación | 3 |
| 1.2 Objetivos | 5 |
| 1.3 Estructura de la memoria | 5 |
| 2. Herramientas | 6 |
| 2.1 Coordenadas Afines | 6 |
| 3. Desarrollo | 8 |
| 3.1 Análisis | 8 |
| 3.2 Diseño | 11 |
| 3.2.1 Energía intra-región | 13 |
| 3.2.2 Energía inter-región | 17 |
| 3.2.3 Energía total | 20 |
| 3.2.4 Inicialización de muelles | 21 |
| 4. Pruebas i Resultados | 23 |
| 4.1 Análisis | 23 |
| 4.2 Pruebas de verificación | 23 |
| 4.2.1 Energía intra-región | 24 |
| 4.2.1.1 Imagen Sintética | 24 |
| 4.2.1.2 Imagen de la base de datos | 26 |
| 4.2.1.3 Inicialización de las regiones | 30 |
| 4.2.2 Inicialización de las distancias | 33 |
| 4.2.3 Energía inter-región | 34 |
| 4.2.4 Energía total | 39 |
| 4.2.5 Energía total con imágenes sintéticas | 45 |
| 5. Conclusiones y Líneas futuras | 48 |
| 5.1 Conclusiones | 48 |
| 5.2 Líneas Futuras: Texturización | 49 |
| 6. Bibliografía | 52 |
| Apéndice: Estructura código | 53 |

1. Introducción y Motivación

1.1 Motivación

Uno de los problemas más complejos de la visión por computador es el de “scene understanding”, hacer que un ordenador sea capaz de entender una escena. Este problema implica reconocer los componentes de la escena, y su relación con el entorno. Uno de los componentes más complejos de reconocer son los seres humanos, debido al hecho de ser un organismo articulado. Por estos motivos en este trabajo se ha querido realizar un modelo articulado (Figura 1) capaz de adaptarse a cualquier postura que pueda realizar una persona a partir de un algoritmo matemático. Para ello cuenta con la ayuda de un conjunto de imágenes que nos dan información sobre la situación de las partes del cuerpo de la persona.



Figura 1: Modelo Articulado Humano

Éste modelo articulado puede utilizarse en proyectos futuros, con la finalidad de lograr reconocimiento de gestos, caras , etc. En el caso de disponer de un conjunto de imágenes que forman un vídeo, se podría conseguir, con dicho modelo, un seguimiento de una o varias personas. Aunque en todo el proyecto se hable sobre modelo articulado de personas, se podría utilizar el mismo algoritmo para animales (Figura 2) o hasta para objetos inanimados.

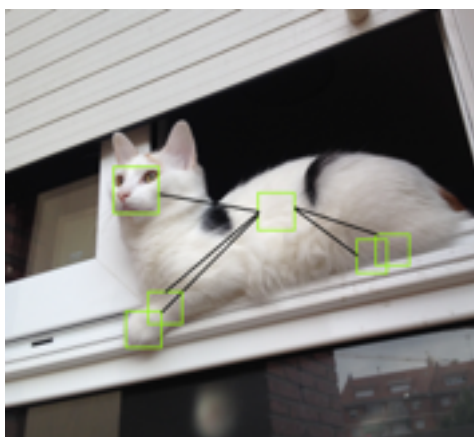


Figura 2: Modelo Articulado Animal

En la literatura de visión por computador uno de los modelos clásicos es el modelo por componentes o “pictorial model”. [*Pictorial Structures for Object Recognition*] Un *pictorial model* es un modelo articulado donde cada parte corresponde a un componente del objeto a reconocer y se establece una conexión entre las partes que puede modelarse como un modelo en muelle, es decir, un modelo articulado conectado por muelles entre algunas de sus regiones, los cuales no permiten que las regiones se separen ni se junten demasiado, por tanto, mantienen mínimamente la forma. La realización algorítmica de este tipo de modelos se puede formular como un modelo deformable [*Image Segmentación with Cage Active Contours*]. Un modelo deformable es un modelo en el cual un contorno o polígono se va adaptando dependiendo de su interior, para ello utiliza la interpolación de coordenadas afines.

Este proyecto se p matemático que siga los principios de los modelos deformables para realizar un modelo pictórico. Para ello se cuenta como entrada de un conjunto de imágenes previamente obtenidas a partir de un algoritmo de aprendizaje automático y que corresponden a la probabilidad de encontrar una parte del modelo en una imagen, véase Figura 3. A partir de estas partes se realizará un modelo articulado deformable basado en muelles que se adapte a las conjunto de imágenes obtenidas.

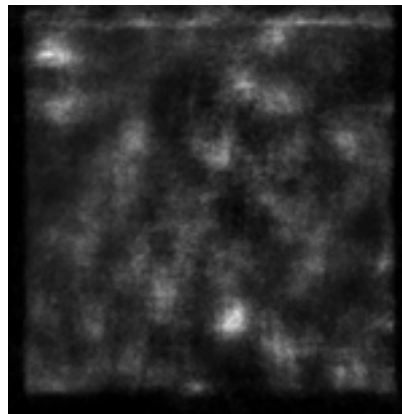


Figura 3: Imagen de probabilidad obtenida por aprendizaje automático

1.2 Objetivos

El objetivo principal de este trabajo es conseguir realizar un algoritmo matemático que sea capaz de detectar personas a partir de un conjunto de imágenes de probabilidad. Este algoritmo tiene que adaptarse perfectamente a la silueta de la persona sea la que sea esta. Dentro del objetivo principal habrían diversos subobjetivos los cuales se pueden dividir en tres partes:

- Obtener una primer aproximación de las articulaciones del cuerpo a partir de imágenes de probabilidad.
- Mejorar la localización de las articulaciones del cuerpo mediante las relaciones que existen entre ellas, también llamadas muelles.
- Aplicación de textura.

Todos estos objetivos se han cumplido a excepción de la aplicación de la textura que solo se ha realizado de forma teórica por falta de tiempo.

1.3 Estructura de la memoria

La memoria se divide en 5 apartados que consisten en Introducción y motivación, Herramientas, Desarrollo, Pruebas y Conclusiones.

En el capítulo 2 se realiza una explicación de todas las herramientas matemáticas externas que son necesarias para realizar cálculos que forman parte del algoritmo.

En el capítulo 3 se realiza una detallada explicación sobre todo el desarrollo del algoritmo matemático utilizado. Mientras que en el capítulo 4, hay una extensa demostración del algoritmo, donde se muestran un conjunto de resultados variados a partir de diferentes situaciones.

2. Herramientas

En este apartado se explican las herramientas matemáticas externas necesarias (que se han necesitado utilizar) para poder realizar el algoritmo que se explicará en el apartado 3.

2.1 Coordenadas Afines

En un polígono cerrado, compuesto por un número finito de vértices, n , en el cual se desea interpolar puntos las coordenadas afines son necesarias. Las coordenadas afines nos permite localizar unívocamente un punto del interior del polígono en función de las coordenadas de los vértices.

Cada uno de los puntos que se encuentran dentro de un contorno cerrado se puede definir tal i como se muestra en la Ecuación 1.

$$p = \sum_{i=1}^n \alpha_i(p) V_i$$

Ecuación 1: Definición Punto

Donde α corresponde al conjunto de coordenadas afines y V_i al conjunto n de vértices.

Las coordenadas afines de un determinado punto se pueden calcular a partir de la Ecuación 2, mostrada a continuación:

$$\alpha_i(p) = \frac{w_i}{\sum_j^n w_i}$$

Ecuación 2: Definición coordenadas afines de un determinado punto P

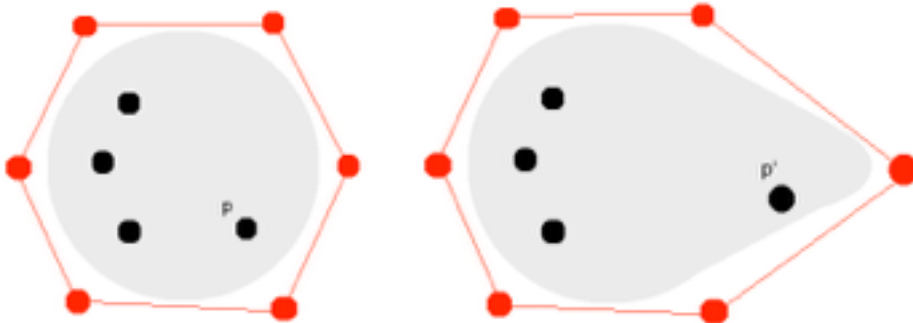


Figura 4: Punto inicial formado por seis coordenadas afines, una por vértice.

Figura 4b: Nuevo punto a partir de las coordenadas afines al mover un vértice

De la anterior fórmula se tiene que calcular previamente el valor de W_i , a continuación podemos ver la Ecuación 3 que permite el cálculo:

$$w_i = \frac{\tan(\gamma_{i-1}/2) + \tan(\gamma_i/2)}{\|V_i - p\|}$$

Ecuación 3: Definición del componente W_i de las coordenadas afines

Donde $\|V_i - p\|$ es la distancia que se encuentra entre el vértice V_i y las coordenadas del punto p . Finalmente, γ corresponde al ángulo que forman $[V_i, p, V_{i+1}]$.

Las coordenadas afines van entre 0 y 1. Además la suma de todas las coordenadas afines de un punto, para cada uno de los vértices, equivale a un total de 1. El cálculo de la coordenada afín de un determinado punto, tendrá un valor más cercano a 1 cuanto más cerca se encuentre un punto del vértice, mientras que, por lo contrario, cuanto más alejado se encuentre un punto del vértice tendrá un valor más cercano a 0.

3. Desarrollo

3.1 Análisis

El trabajo consiste en adaptar un modelo articulado, tal y como se muestra en la Figura 5, a partir de diferentes partes del cuerpo de una persona. Al observar la Figura 6, se muestra de forma más clara como el modelo articulado se adapta a las diferentes articulaciones de la persona que se encuentra en la imagen.

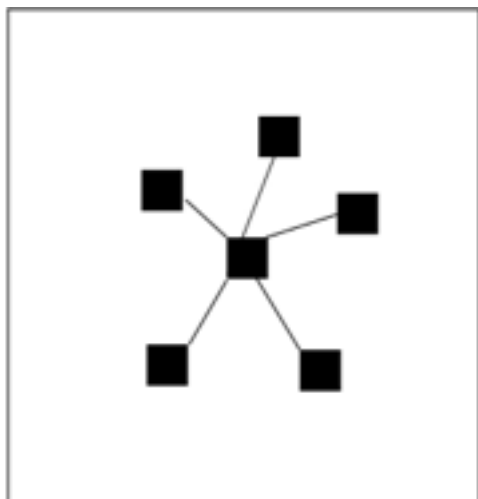


Figura 5: Modelo Articulado



Figura 6: Modelo articulado en imagen original

Suponiendo que se parte de un modelo inicial como el de la Figura 7. Éste tendrá que adaptarse hasta encontrar la posición de una persona en la imagen de forma automática, tal y como se observa en la Figura 8. Este modelo tiene que ser adaptable a cualquier posición que una persona pueda realizar.

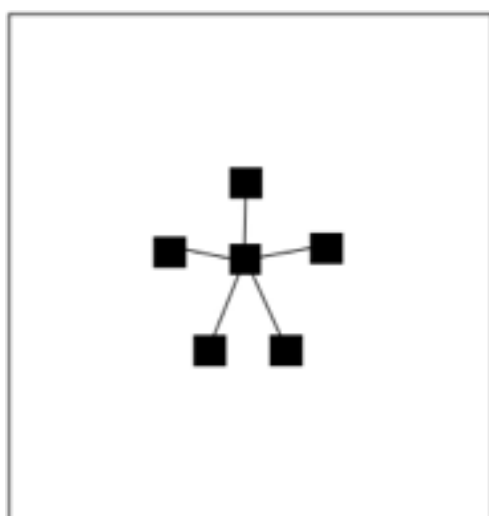


Figura 7: Modelo articulado

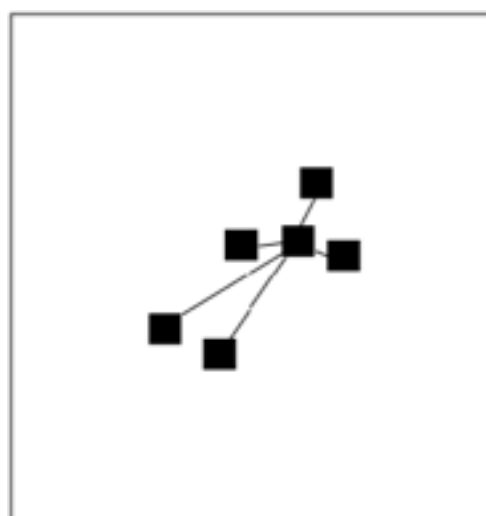


Figura 8: Modelo articulado 2

Para ello, partimos de una base de datos compuesta por varias imágenes. Las cuales nos dan información sobre la probabilidad de que haya una determinada parte del cuerpo. Como se observa en la Figura 9 a mayor nivel de gris existe una mayor probabilidad de que esa zona de la imagen corresponda a una parte del cuerpo.

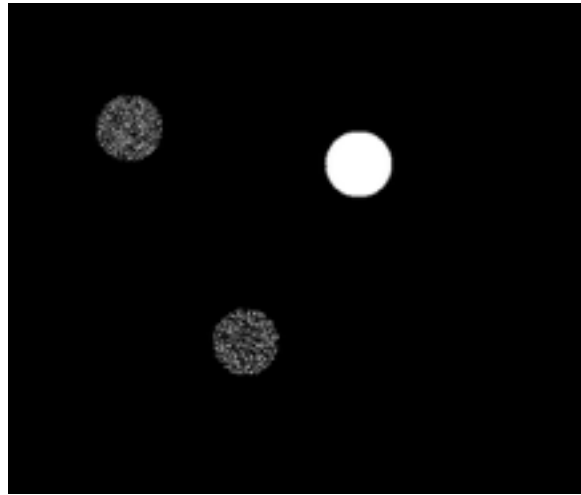


Figura 9: Imagen de probabilidad sintética

La Figura 9 es una imagen de probabilidad creada de forma artificial. Suponiendo que esta imagen corresponde a una imagen de probabilidad de la cabeza, se destacan diversos núcleos que podrían ser la cabeza real. Pero hay una zona de las tres que contienen un nivel de gris más elevado que destaca del resto, tal y como se señala en la Figura 10, por lo tanto, esa zona tiene una mayor probabilidad de que sea la cabeza, y las otras dos pueden ser falsos positivos.

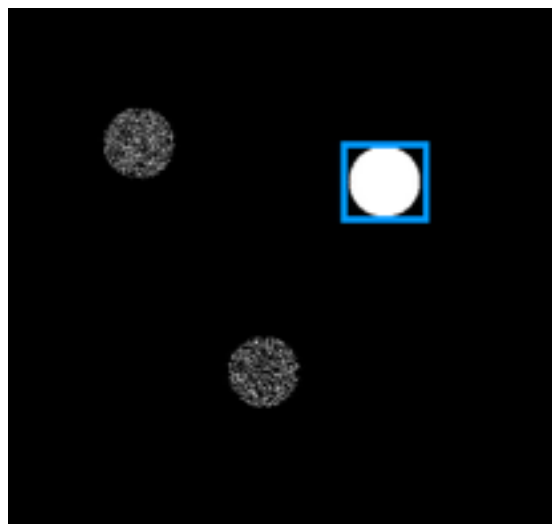


Figura 10: Imagen de probabilidad sintética destacada

Cada una de las imágenes de la base de datos se compone de 15 canales. Los cuales corresponden a cada una de las partes del cuerpo en el siguiente orden: cabeza, cuello, hombro izquierdo, codo izquierdo, muñeca izquierda, hombro derecho, codo derecho, muñeca derecha, cadera izquierda, rodilla izquierda, tobillo izquierdo, cadera derecha, rodilla derecha, tobillo derecho y fondo. Cada uno de estos canales corresponde a una imagen de probabilidad en escala de grises.

A partir de la imagen original, Figura 11, tenemos en la base de datos sus correspondientes imágenes de probabilidad, tres ejemplos de ellos se encuentran en las Figuras 12, 13 y 14.



Figura 11: Imagen original

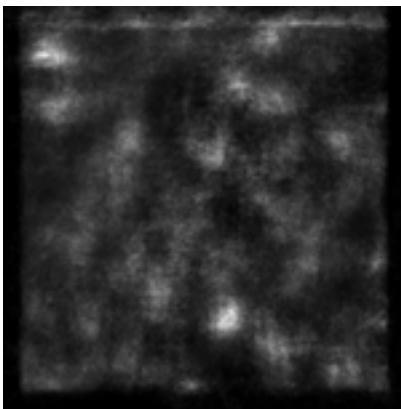


Figura 12: Imagen de probabilidad de la cabeza

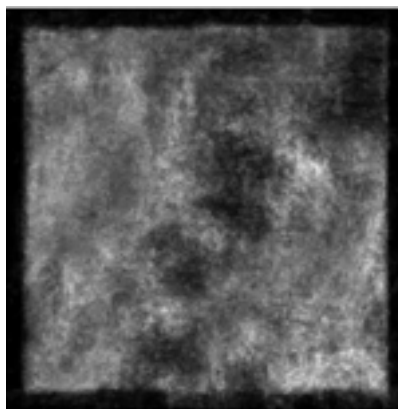


Figura 13: Imagen de probabilidad de la muñeca izquierda

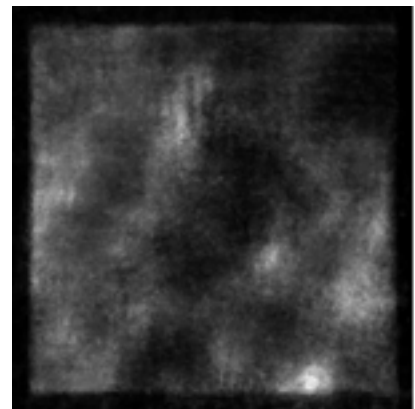


Figura 14: Imagen de probabilidad del tobillo izquierdo

Las imágenes de probabilidad reales (Figura 8, 9 y 10) no son tan fáciles de tratar como puede ser la de la Figura 6. Todo lo contrario, resultan imágenes muy complejas y no se puede observar de forma sencilla y clara la situación de una parte del cuerpo determinada. Por estos motivos, es necesario realizar el diseño de un algoritmo que ayude a resolver este problema. De manera que facilite la detección de las diferentes partes del cuerpo de la persona.

3.2 Diseño

A las diferentes partes del cuerpo o articulaciones, que se utilizan para obtener el modelo articulado, las llamamos regiones.

Cada región está constituida por un conjunto de vértices V_i , compuestos de las coordenadas X e Y, y un número finito de puntos p . El algoritmo es válido para que una región tenga cualquier forma pero en el trabajo se centra en regiones con forma cuadrada. En particular las regiones utilizadas, al ser cuadradas como en la Figura 15, están constituidas por 4 vértices (V_1 , V_2 , V_3 y V_4) y un número finito de puntos delimitados por los vértices.

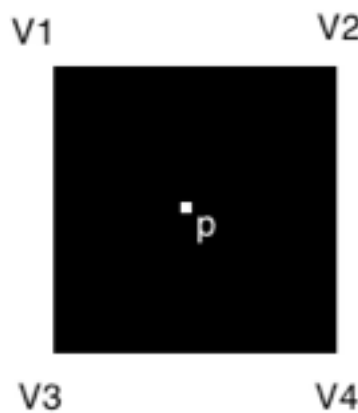


Figura 15: Región cuadrada

Los puntos de una región los podemos definir como el sumatorio de sus coordenadas afines, explicadas en el apartado 2.2, respecto a los vértices, Ecuación 4.

$$p = \sum_{i=1}^n \alpha_i(p) V_i$$

Ecuación 4: Definición Punto

Donde α corresponde al cálculo de las coordenadas afines de un determinado punto (p).

En la Figura 12 tenemos un ejemplo de un modelo articulado con 6 regiones, las cuales tienen 4 vértices. Los puntos de cada una son todos los píxeles que se encuentran dentro del cuadrado que forma la región. Los modelos creados en el trabajo son de tipo estrella como en la siguiente figura, Figura 16. Contienen un nodo central el cual se encuentra conectada al resto de nodos. Se utiliza el tipo estrella con la finalidad de facilitar la formulación del modelo matemático, así como, encontrar una solución a éste.

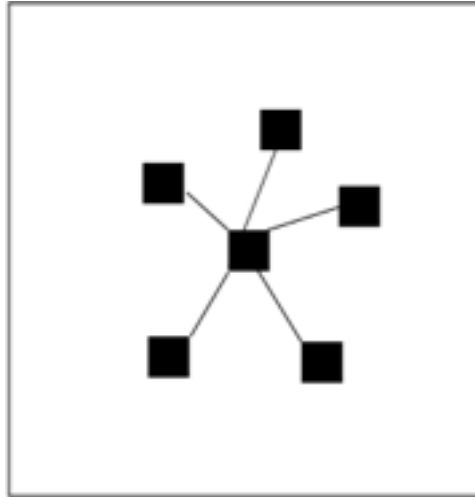


Figura 16: Modelo articulado tipo estrella

A continuación se presenta la formulación matemática de la energía que se utiliza para adaptar el modelo articulado a la persona. Se basa en minimizar una energía formada por la combinación lineal de 2 términos: una energía intra-región y otra inter-región.

3.2.1 Energía Intra-región

La energía intra-región es la energía que tiene una región concreta ella sola, sin tener en cuenta si está conectado a otra u otras regiones diferentes.

Esta energía consiste en la minimización del sumatorio de todos los puntos de una región concreta (Ecuación 5).

$$E = - \sum_{p \in \mathbb{R}} I(p)$$

Ecuación 5: Energía Intra-Región

Nos referimos con energía (E) a la cantidad de probabilidad de que en una región de la imagen (I) se encuentre una parte del cuerpo determinada. El signo negativo es debido a que al trabajar con imágenes de probabilidad, se quiere maximizar dicha probabilidad. Con lo que se coge el signo negativo para convertirlo en un problema de minimización.

Para poder obtener el mínimo en cada dirección (x,y), y así encontrar la variaciones de energía, tendremos que calcular la derivada de la energía (Ecuación 6) respecto a la dirección suponiendo que la perturbación(dx,dy) tiende a 0.

$$\frac{\partial E}{\partial d} = \left(\frac{\partial E}{\partial d_x}, \frac{\partial E}{\partial d_y} \right)$$

Ecuación 6: Energía Intra-Región

Para obtener la derivada de la energía en cada una de las direcciones se utiliza la regla de la cadena. Para realizar la regla de la cadena primero se definen los vértices que tendrá la región. Así como, las nuevas posiciones que tendrán. La región se compone de cuatro vértices (V1, V2, V3 y V4), tal y como se muestra en la Figura 17:

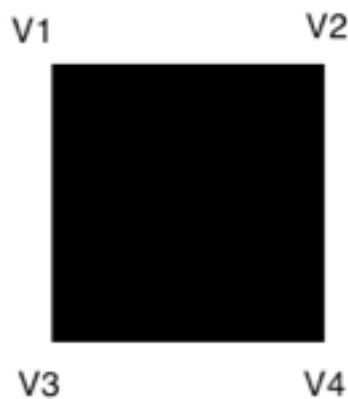


Figura 17: Vértices de una región

Los vértices se definen por sus propias coordenadas más la perturbación que sufren en cada una de las direcciones (Ecuación 7).

$$Vi' = Vi + (dx, dy)$$

Ecuación 7: Vértices con perturbación en dirección X y/o Y

Por tanto, la formula explicada anteriormente, al principio de este apartado, es la energía intra-región delimitada por los vértices cuando la perturbación en las dos direcciones es 0 (Ecuación 8).

$$E(dx = 0, dy = 0) = - \sum_{p \in \mathbb{R}} I(p)$$

Ecuación 8: Energía Intra-Región sin perturbar

Mientras que, si la perturbación en alguna de las dos direcciones tiene cualquier otro valor, la formula de la energía queda definida como en la Ecuación 9.

$$E'(dx \neq 0, dy \neq 0) = - \sum_{p \in \mathbb{R}} I(p')$$

Ecuación 9: Energía Intra-Región perturbando mínimo en una dirección

La p' corresponde a cada uno de los nuevos puntos que contiene la región. Por lo tanto, intervienen en el cálculo de la nueva energía. Estos puntos se definen a partir del sumatorio de sus coordenadas afines respecto a los nuevos puntos que estarán perturbados en, al menos, una de las dos direcciones como se muestra en la Ecuación 10.

$$p' = \sum_{i=1}^n \alpha(p)(Vi + (dx, dy))$$

Ecuación 10: Nuevo punto que contiene la región

Al realizar el cálculo de la cadena quedan las dos siguientes formulas, una para cada dirección donde se realiza la perturbación (Ecuación 11 y 12).

$$\frac{\partial E}{\partial dx} = - \sum \frac{\partial I}{\partial x} \frac{\partial p_x}{\partial dx}$$

Ecuación 11: Derivada energía Intra-Región X

$$\frac{\partial E}{\partial dy} = - \sum \frac{\partial I}{\partial y} \frac{\partial p_y}{\partial dy}$$

Ecuación 12: Derivada energía Intra-Región Y

Se observa que la derivada de la imagen, respecto la coordenada X e Y, corresponde a los cambios de gradientes. Por lo tanto, previamente se tienen que calcular estos gradientes de la imagen. El cálculo de gradientes se realiza a partir de un filtro Prewitt, donde la mascara utilizada en la dirección X corresponde a la Ecuación 13 y en la dirección Y a la Ecuación 14.

$$\text{Dirección } X = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \quad \text{Dirección } Y = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Ecuación 13: Mascara dirección X

Ecuación 14: Mascara dirección Y

Todos los vértices pueden moverse todos hacia la dirección predominante entre todos ellos al realizar la suma de todas las coordenadas afines o que cada uno se mueva libremente según su coordenada afín.

La derivada de la energía donde todos los puntos se mueven hacia la misma dirección es la siguiente:

$$\frac{\partial E}{\partial d_x} = - \sum_{p \in \mathbb{R}} \left(\frac{\partial I}{\partial x} * \left(\sum_{i=1}^n \alpha_i(p) \right) \right)$$

Ecuación 15: Derivada energía intra-región moviendo los puntos en una única dirección perturbando en X

$$\frac{\partial E}{\partial d_y} = - \sum_{p \in \mathbb{R}} \left(\frac{\partial I}{\partial y} * \left(\sum_{i=1}^n \alpha_i(p) \right) \right)$$

Ecuación 16: Derivada energía intra-región moviendo los puntos en una única dirección perturbando en Y

Tal y como se muestra en la Figura 18 todos los vértices se moverán en la misma dirección.

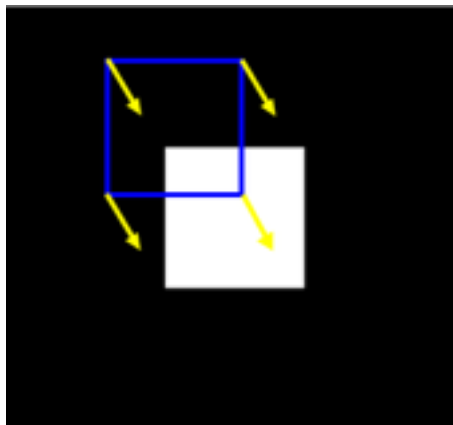


Figura 18: Movimiento de los vértices hacia una misma dirección

En el caso de que los vértices se muevan libremente, el cálculo de la derivada de la energía es ligeramente diferente. La componente n de la fórmula corresponde a cada uno de los vértices que tiene la región, en este caso concreto iría del 1 al 4:

$$\frac{\partial E_n}{\partial d_x} = - \sum_{p \in \mathbb{R}} \left(\frac{\partial I}{\partial x} * \alpha_n(p) \right)$$

Ecuación 17: Derivada energía intra-región moviendo los puntos libremente perturbando en X

$$\frac{\partial E_n}{\partial d_y} = - \sum_{p \in \mathbb{R}} \left(\frac{\partial I}{\partial y} * \alpha_n(p) \right)$$

Ecuación 18: Derivada energía intra-región moviendo los puntos libremente perturbando en Y

Tal y como se observa en la Figura 19, cada uno de los vértices se moverá en una dirección según la coordenada afín calculada previamente.

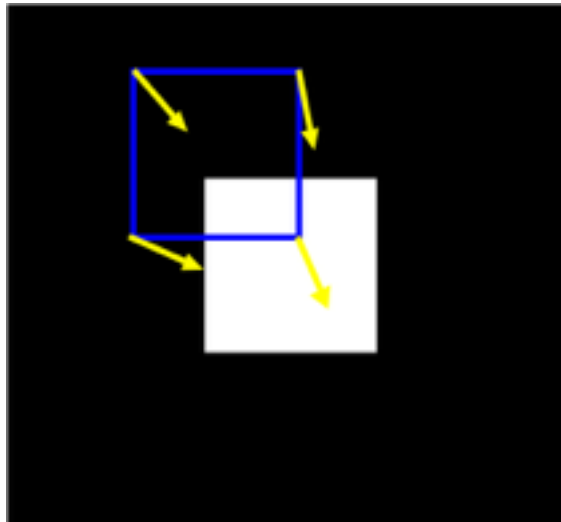


Figura 19: Movimiento de los vértices en diversas direcciones

3.2.2 Energía Inter-región

La energía inter-región es la que existe entre diferentes regiones o partes del cuerpo, las cuales consideramos que están conectadas. A estas conexiones entre regiones les llamamos muelles.

Para calcular esta energía es necesario un nodo central que será el torso de la persona. Desde éste saldrán diversos muelles hacia las articulaciones. De esta forma se obtiene un modelo de tipo estrella.

En la Figura 20, se observa un ejemplo de modelo articulado, en el cual se resaltan con un punto blanco el centro de cada una de las regiones. Las cuales corresponden a las diferentes partes del cuerpo. El C0 corresponde al nodo central, torso, mientras que el resto de nodos, C1 a C5, corresponden a la cabeza, muñeca izquierda, muñeca derecha, tobillo izquierdo y tobillo derecho respectivamente.

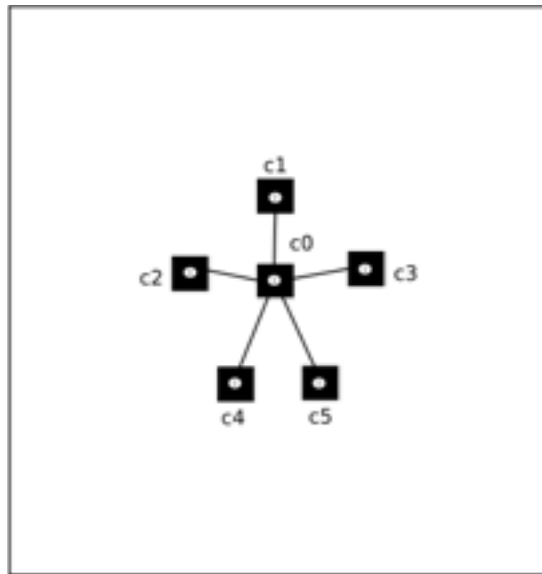


Figura 20: Centros en el modelo articulado

Para cada región es necesario definir una función que sea capaz de dar información sobre si el muelle, o los muelles, que la conecta con otras regiones, tienen una distancia razonable, o por lo contrario, una distancia muy lejana o cercana. Para ello se tiene que calcular la función para cada uno de los muelles. Por lo tanto, se calcula la Ecuación 19 para cada una de las parejas de centros existentes.

$$f(c_0, c_i) = (||c_0 + d_1 - c_i||^2 - d_{0,i}^2)^2$$

Ecuación 19: Función entre los centros de dos regiones

Para que la función se comprenda de una forma más sencilla, se escribe a continuación la misma ecuación pero quitando la notación vectorial:

$$f(c_0, c_i) = (((c_{0,x} + d_{0,x}) - c_{i,x})^2 + ((c_{0,y} + d_{0,y}) - c_{i,y})^2 - d_{0,i}^2)^2$$

Ecuación 20: Función entre los centros de dos regiones sin notación vectorial

Seguidamente se calcula la derivada de la función en el caso de perturbar cualquier región que no sea el torso. En el caso de perturbar en la dirección X se utiliza la Ecuación 21, mientras que si se quiere perturbar en dirección Y se utiliza la Ecuación 22.

$$\frac{\partial f(c_i, c_0)}{\partial d_{i,x}} = 2 * (((c_{i,x} + d_{i,x}) - C_{0,x})^2 + ((c_{i,y} + d_{i,y}) - c_{0,y})^2 - d_{0,i}^2) * 2 * ((c_{i,x} + d_{i,x}) - c_{0,x}) * (-1)$$

Ecuación 21: Derivada de la función entre los centros de dos regiones en dirección X

$$\frac{\partial f(c_i, c_0)}{\partial d_{i,y}} = 2 * (((c_{i,x} + d_{i,x}) - C_{0,x})^2 + ((c_{i,y} + d_{i,y}) - c_{0,y})^2 - d_{0,i}^2) * 2 * ((c_{i,y} + d_{i,y}) - c_{0,y}) * (-1)$$

Ecuación 22: Derivada de la función entre los centros de dos regiones en dirección Y

Si se perturba el torso, el cálculo de la derivada cambia ligeramente como queda demostrado en las Ecuaciones 23 y 24.

$$\frac{\partial f(c_0, c_i)}{\partial d_{0,x}} = 2 * (((c_{0,x} + d_{0,x}) - c_{i,x})^2 + ((c_{0,y} + d_{0,y}) - c_{i,y})^2 - d_{0,i}^2) * 2 * ((c_{0,x} + d_{0,x}) - c_{0,x})$$

Ecuación 23: Derivada de la función entre los centros de dos regiones en dirección X en el caso del torso

$$\frac{\partial f(c_0, c_i)}{\partial d_{0,y}} = 2 * (((c_{0,x} + d_{0,x}) - c_{i,x})^2 + ((c_{0,y} + d_{0,y}) - c_{i,y})^2 - d_{0,i}^2) * 2 * ((c_{0,y} + d_{0,y}) - c_{0,y})$$

Ecuación 24: Derivada de la función entre los centros de dos regiones en dirección Y en el caso del torso

Hasta ahora el modelo articulado que se ha tenido en cuenta es el conformado de 6 regiones con un único nodo central. Pero un modelo articulado puede tener muchas más regiones que 6, así como diversos nodos centrales. No obstante, todas estas características comportan que los cálculos se compliquen considerablemente. Por el contrario, aumentar el número de regiones y nodos centrales que el modelo sea más preciso y la silueta de la persona quede mejor definida tal y como se muestra en la Figura 21.

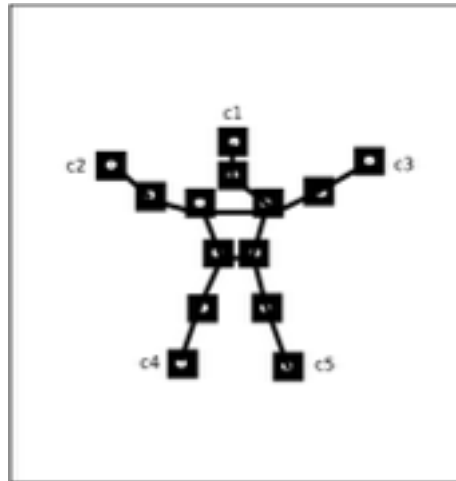


Figura 21: Modelo articulado complejo

3.2.3 Energía Total

La energía total para el conjunto de las regiones se compone por el sumario de las energías inter-región e intra-región (Ecuación 25).

$$E = \sum_i (\sum_{p \in \mathbb{R}} I(p) + \beta f(C_0, C_i))$$

Ecuación 25: Energía total de todas las regiones

Por lo tanto, para cada una de las regiones su energía total se calcula según la Ecuación 26.

$$E = \sum_{p \in \mathbb{R}} I(p) + \beta f(C_0, C_i)$$

Ecuación 26: Energía total para cada región individualmente

En esta fórmula hay un componente β para dar mayor importancia o menor a la energía inter-región respecto la energía intra-región, ya que los resultados obtenidos en la energía inter-región son números demasiado elevados comparándolos con la energía intra-región. De esta forma, si β tiene un valor de 1, prácticamente, solo se tendrá en cuenta la energía inter-región en los resultados, mientras que si el valor es demasiado pequeño pasará exactamente lo contrario. Es necesario encontrar un valor idóneo que produzca un equilibrio entre las dos energías.

A continuación se normaliza la derivada de la energía total según la Ecuación 27. Una vez normalizada se conoce la dirección por pixel hacia la que se moverá un cierto vértice.

$$\frac{\partial E}{\partial d} = (x, y) = \left(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}} \right)$$

Ecuación 27: Normalización de la derivada de la energía total

Una vez calculada la energía total normalizada, el siguiente paso consiste en calcular el descenso de gradiente como se muestra en la Ecuación 28. Y de esta manera obtener la nueva posición de cada uno de los vértices.

$$\vec{V}_i^n = \vec{V}_i^{n-1} - \alpha * \frac{\partial E}{\partial d}$$

Ecuación 28: Descenso de gradiente

Podemos observar en la Ecuación 28, que para obtener la nueva posición de cada uno de los vértices, se le aplica el descenso de gradiente a la posición actual del vértice, pero este descenso de gradiente es multiplicado por el símbolo α . Este símbolo α afecta a la cantidad de pixeles que se desplazará el vértice en la dirección obtenida por el descenso de gradiente. Se puede controlar la cantidad de pixeles que se mueve cada vértice gracias a la normalización explicada previamente.

3.2.4 Inicialización de Muelles

En la energía inter-región aparece el problema de indicar que distancia es la ideal entre regiones, para solucionarlo previamente se ha tenido que inicializar estas distancias.

La inicialización se realiza a partir de la imagen del hombre de Da Vinci que corresponde a la Figura 22. Se utiliza este hombre porque, se dice que, tiene las proporciones y medidas perfectas entre las distintas partes del cuerpo.

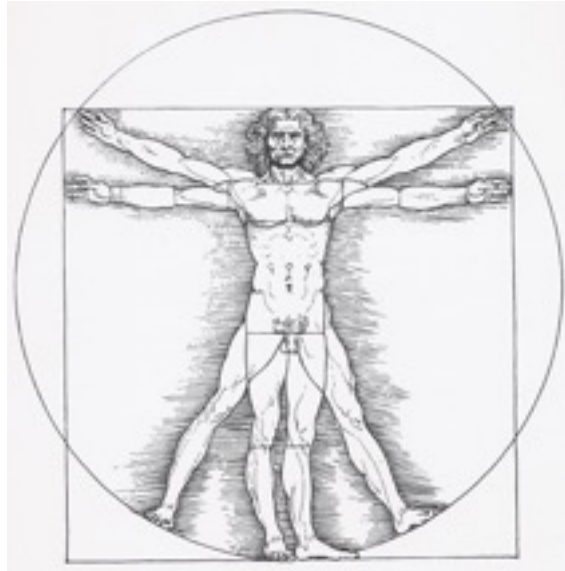


Figura 22: Hombre de Da Vinci (Vitruvio)

Para encontrar las distancias se necesitan conocer previamente la localización de los centros de las regiones que se tendrán en cuenta. Tal y como se muestra en la Figura 23, esos serán nuestros 6 centros.

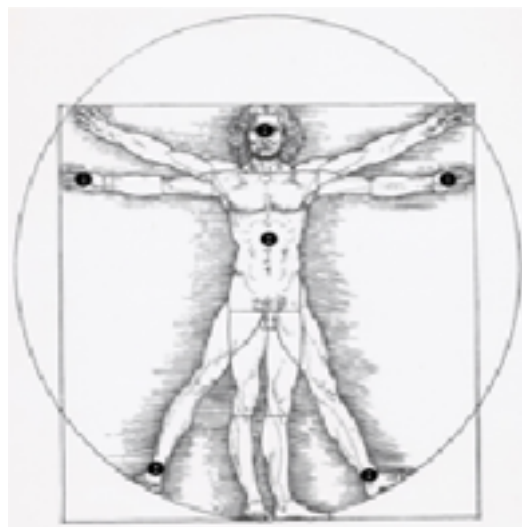


Figura 23: Centros de las regiones en hombre Vitruvio

Posteriormente, se calculan las distancias (Ecuación 29) entre todas las regiones que se encuentren conectadas, estas conexiones o muelles están resaltadas en la Figura 24.

$$d_{0,i}^2 = (c0_x - ci_x)^2 + (c0_y - ci_y)^2$$

Ecuación 29: Distancia entre centros

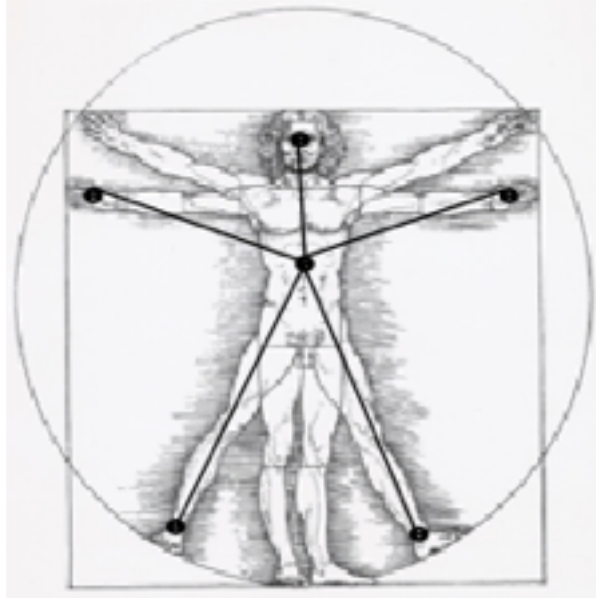


Figura 24: Muelles del hombre Vitruvio

Observase que la gran mayoría de personas que podemos encontrar en las imágenes no tendrán la misma altura que el hombre de Da Vinci. Una fácil solución es guardar las distancias en función de la altura, para ello se dividen las distancias entre la altura del hombre Vitruvio (Ecuación 30). A continuación, se tendrá que multiplicar los valores por la altura aproximada que tendrá la persona de las imágenes de la base de datos (Ecuación 31).

$$d_{0,i} = d_{0,i}/h$$

Ecuación 30: Distancias normalizadas entre 0 y 1

$$d_{0,i} = d_{0,i} * x$$

Ecuación 31: Distancias en función de la altura

Donde h es la altura del hombre vitruvio y x la altura de la persona que estamos buscando.

4. Pruebas y Resultados

4.1 Análisis

Para realizar las pruebas y no tener tanta cantidad de articulaciones, la cual cosa hace que los cálculos sean más complicados, solo se tendrán en cuenta las 6 principales regiones que son: cabeza, muñeca izquierda, muñeca derecha, torso, tobillo izquierdo y tobillo derecho. Estas seis regiones serán las que formarán parte del modelo articulado como se observa en la Figura 25.

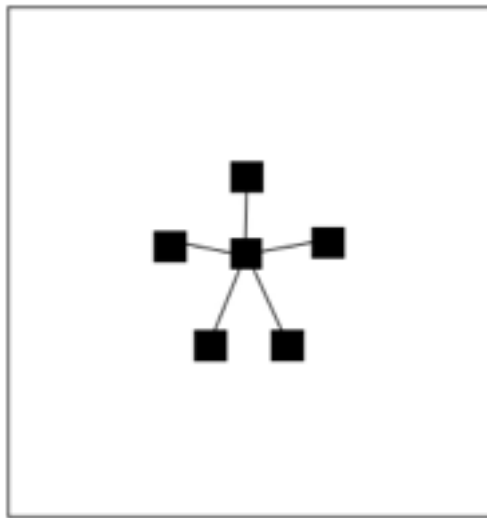


Figura 25: Regiones del modelo articulado

4.2 Pruebas de verificación

En los siguientes sub-apartados se realizarán diversas pruebas para demostrar el funcionamiento y comportamiento de las diversas secciones explicadas en el apartado 3 de Desarrollo.

4.2.1 Energía Intra-región

Comenzaremos haciendo pruebas con la energía intra-región, es decir, pruebas en las que únicamente se tendrá en cuenta la energía de las propias regiones a partir de las imágenes de probabilidad.

4.2.1.1 Imagen sintética

Las primeras pruebas de la energía intra-región se han realizado con una imagen sintética. Esta imagen es completamente blanca a excepción de un cuadrado negro, Figura 26.

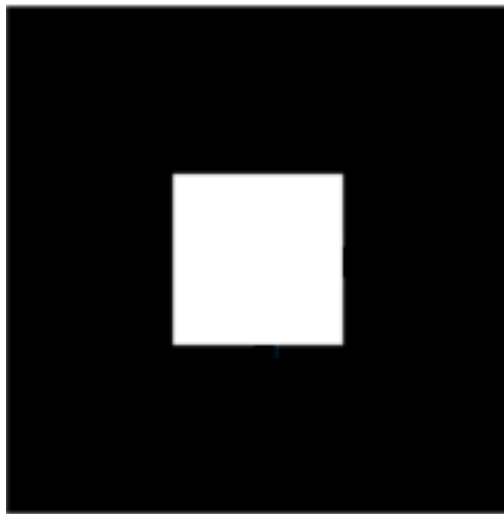


Figura 26: Imagen sintética

Para realizar la prueba, primero, se inicializa un región con forma cuadrada (formada por 4 vértices y todos los puntos que quedan delimitados por estos) relativamente cerca de la zona donde hay mayor probabilidad como se muestra en la Figura 27.

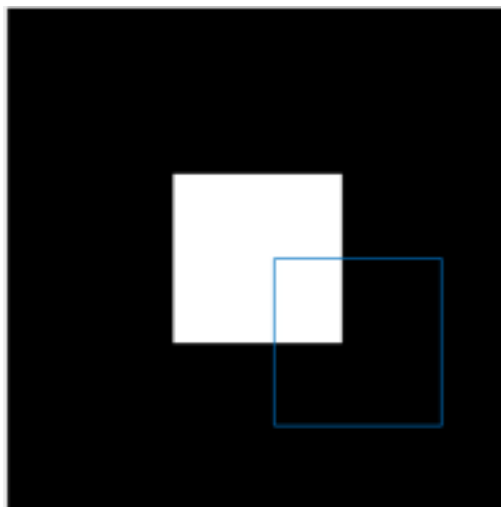


Figura 27: Inicialización región

El objetivo de esta prueba es que la región azul vaya moviéndose de forma progresiva hasta colocarse en la zona de mayor probabilidad de la imagen, zona de color blanco en este caso.

Como se observa en la Figura 28, a partir de las iteraciones, la región azul se va moviendo acercándose a la zona de mayor probabilidad. Finalmente, en la Figura 29 queda el recuadro azul justo encima de la zona de probabilidad, aunque se intente calcular más veces, la energía intra-región ya no se mueve de esa zona porque todo su alrededor tiene menor probabilidad. Por consecuencia el objetivo está cumplido ya que el comportamiento de la región es el esperado y deseado.

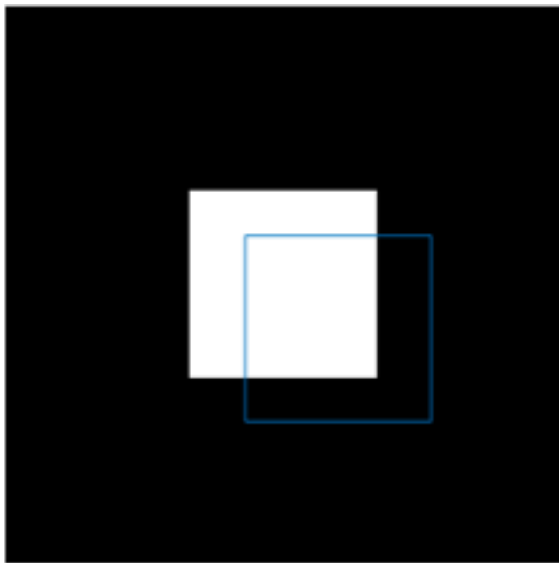


Figura 28: Movimiento de la región

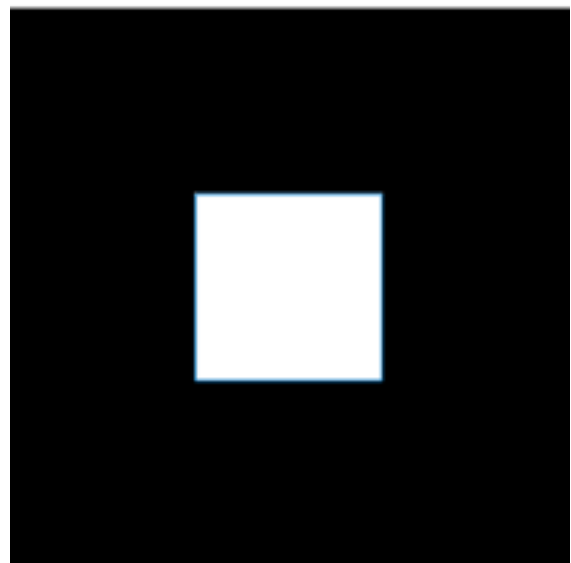


Figura 29: Prueba intra-región

4.2.1.2 Imagen de la base de datos

A continuación, la siguiente prueba es similar. Pero esta vez, no se utilizara una imagen sintética, sino que se cogerá una de las imágenes de probabilidad que conforman la base de datos, concretamente la imagen de probabilidad del cuello, Figura 30, de una de las imágenes. La elección del cuello es completamente arbitraria, es una de las partes del cuerpo que se disponen de la persona, aunque para el modelo articulado realizado no se tiene en cuenta.

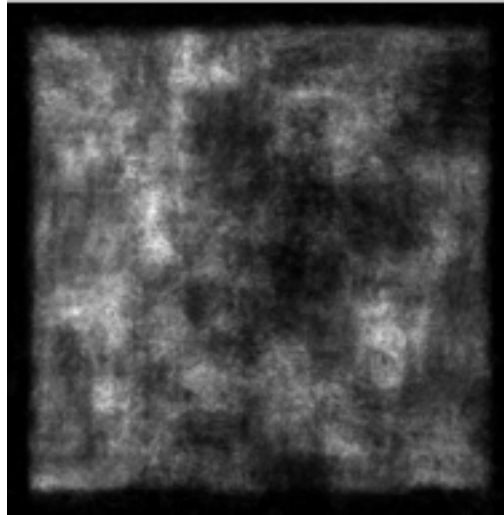


Figura 30: Imagen de probabilidad del cuello

A simple vista, en la Figura 30, se observan varias zonas donde hay una gran cantidad de probabilidad, por lo tanto, no es tan fácil saber hacia qué dirección se dirigirá la región inicializada en la Figura 31.

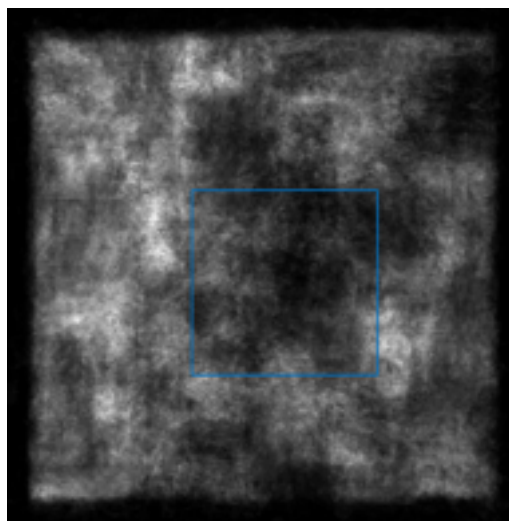


Figura 31: Inicialización región cuello en imagen de probabilidad

El comportamiento esperado es que la región se mueva hacia el oeste o sudeste que son las regiones destacadas de energía. Como se observa en la Figura 32, efectivamente, la región se dirige a la zona sudeste donde hay una gran cantidad de energía. Se dirige hacia está porque esa zona está más cerca a la inicialización que la de la izquierda.

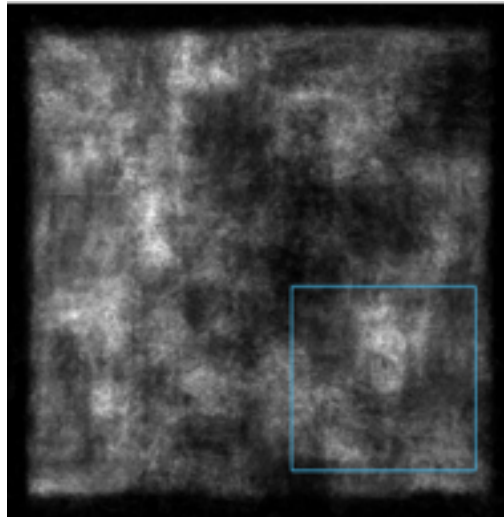


Figura 32: Resultado prueba 1

Aunque la prueba ha realizado el comportamiento que se deseaba, realmente lo esperado es que la región se dirija al lugar donde se encuentra el cuello en la imagen original. Pero si se compara la Figura 32 con la Figura 33, se observa como la región que señala la Figura 35 no corresponde con el recuadro marcado en la Figura 33.



Figura 33: Imagen original resaltando el cuello

El mal comportamiento que se ha encontrado al comparar los resultados con la imagen original, hace que se vuelva a realizar una nueva prueba, pero con una región de un tamaño aproximado a la medida del cuello. La inicialización de esta segunda prueba se puede observar en la Figura 34. Lo que se espera es que se dirija hacia la localización del cuello. El comportamiento realizado en la prueba se observa en la Figura 35. Donde si que se dirige hacia la zona más cercana donde encuentra mayor probabilidad pero no es la localización del cuello.

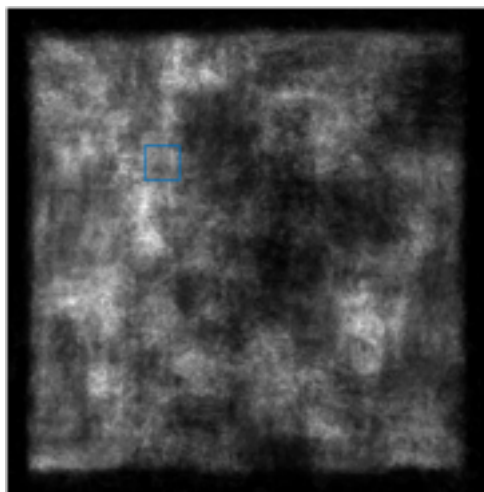


Figura 34: Inicialización del cuello 2

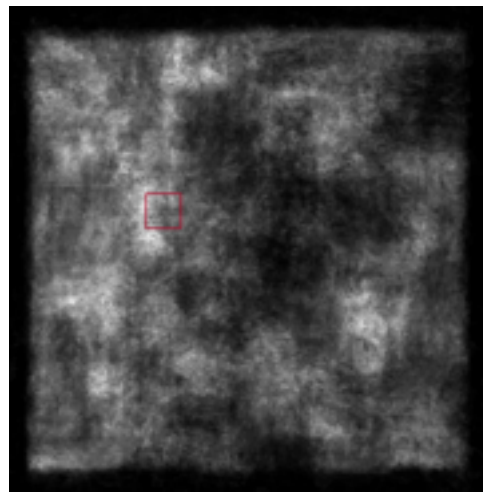


Figura 35: Resultado prueba 2

A partir de la prueba anterior, queda comprobado que la inicialización de una región es muy importante, ya que si se encuentra muy lejano nunca conseguirá llegar al sitio que deseamos. Por este motivo en la nueva prueba se inicializa relativamente cerca la región, Figura 36. El objetivo vuelve a ser que la región acabe situándose donde realmente se encuentra el cuello. No obstante, como se observa en la Figura 37, se sitúa la región en la zona de mayor energía cercana pero no exactamente donde se encuentra el cuello, Figura 38. Esto sucede porque si se observa bien la imagen de probabilidad la zona concreta del cuello no es la que tiene mayor probabilidad de toda la zona de alrededor.

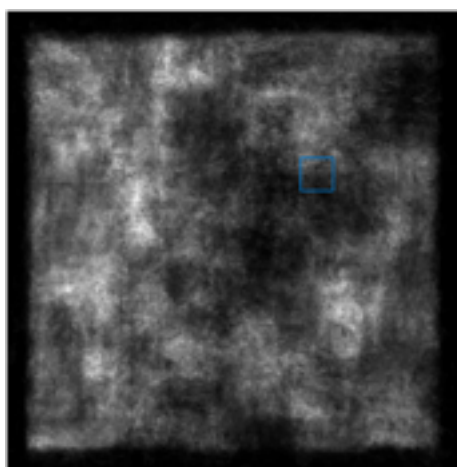


Figura 36: Inicialización del cuello 3

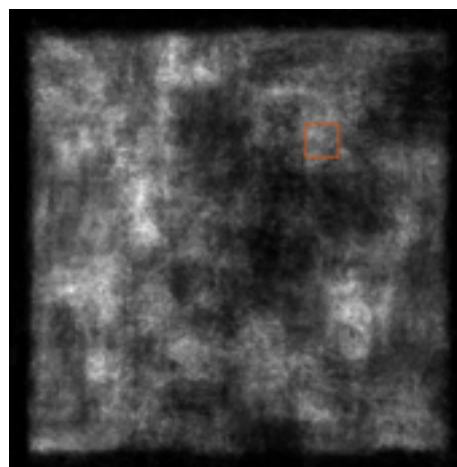


Figura 37: Resultado prueba 3



Figura 38: Imagen original resaltando el cuello

Con todas estas pruebas realizadas se puede afirmar que es importante colocar una región de una medida similar a la de la región real. Pero, también, inicializar la región en una zona cercana para ayudar al algoritmo a encontrar la zona del cuerpo que buscamos. Con estos resultados se llega a la conclusión de que, no solo es importante una medida parecida al tamaño real para encontrar la zona correcta, sino que una lejana inicialización hace que el resultado sea erróneo. Esto sucede porque en las imágenes de la base de datos hay demasiadas zonas que dan la probabilidad de que se encuentre una determinada parte del cuerpo y no solo la región correcta.

Todas las pruebas han sido realizadas moviendo los 4 vértices de la región en la misma dirección, pero tal como se explica en el apartado 3.2.1 los vértices se pueden mover libremente según el cálculo de las coordenadas afines.

Para adaptar una región a la zona de energía, se mueven libremente los vértices, de esta forma se evita que la región tenga zonas donde no haya ninguna probabilidad.

Si nos fijamos en la inicialización, Figura 39, se observa que si se moviesen los vértices libremente, la región tendría que estar compuesta por gran parte de blanco y no tanto gris como en la inicialización. La Figura 40 cumple con el comportamiento deseado, ya que la región se adapta a la zona de mayor probabilidad sin necesidad de preservar los ángulos de la región inicial.

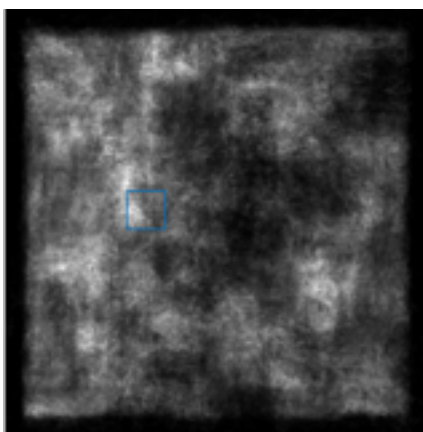


Figura 39: Inicialización región

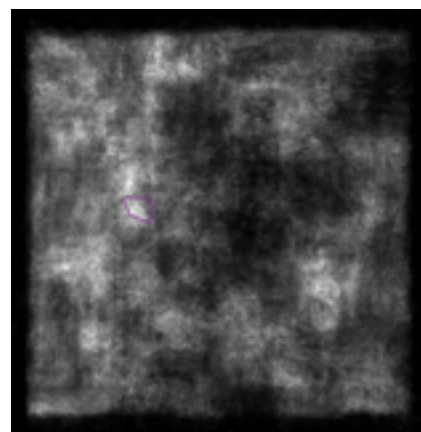


Figura 40: Resultado vértices libres

4.2.1.3 Inicialización de las regiones

Para que el cálculo de la energía intra-región sea más efectivo se ha comprobado en las anteriores pruebas que es muy importante una inicialización para que el algoritmo funcione y encuentre de una forma correcta las articulaciones de la persona.

Proponemos un algoritmo de dos pasos. Donde, en un primer paso, encontramos una posición aproximada de las regiones y en un segundo, refinamos la solución encontrada en el primero aplicando la minimización.

Para solucionar el primer paso del algoritmo se cogen las imágenes de probabilidad correspondientes a cada una de las partes del cuerpo de una imagen. En este caso corresponderán a las imágenes de probabilidad que conforman la Figura 41.



Figura 41: Inicialización región

Para cada imagen de probabilidad, se inicializa una región cuadrada colocada en las coordenadas (0,0) de la imagen y con una determinada anchura, tal y como, se muestra en la Figura 42.

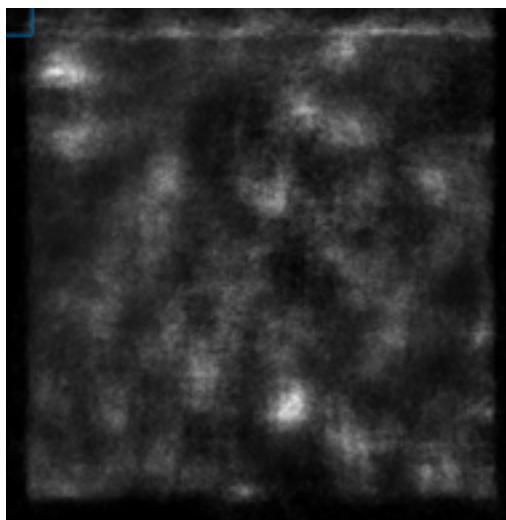


Figura 42: Inicialización región

Esta región se irá desplazando por toda la imagen. Este proceso se denomina ventana deslizante, en cada movimiento de esta ventana se calcula la energía intra-región. Una vez se ha desplazado toda la ventana deslizante por la imagen nos quedamos con la región que contenga la mayor probabilidad. En el caso de la cabeza la región con mayor probabilidad sería la resaltada en la Figura 43. Pero como se puede observar la región utilizada es demasiado grande y el algoritmo no consigue que la región envuelva la zona de mayor probabilidad

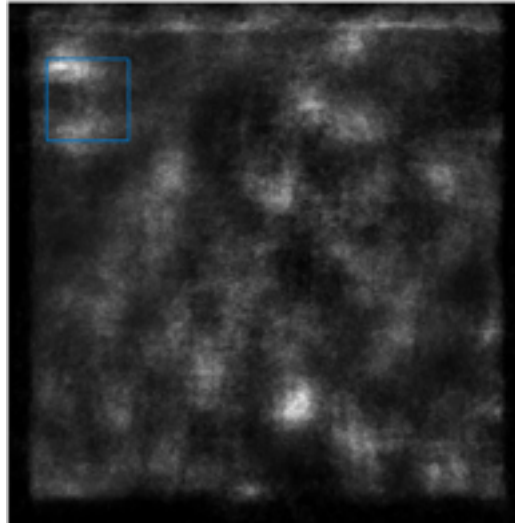


Figura 43: Inicialización cabeza

Si se utiliza una región más pequeña, se consigue un mejor resultado en la búsqueda de la zona donde hay más energía intra-región. Ya que la medida es más similar a la articulación real, Figura 44.

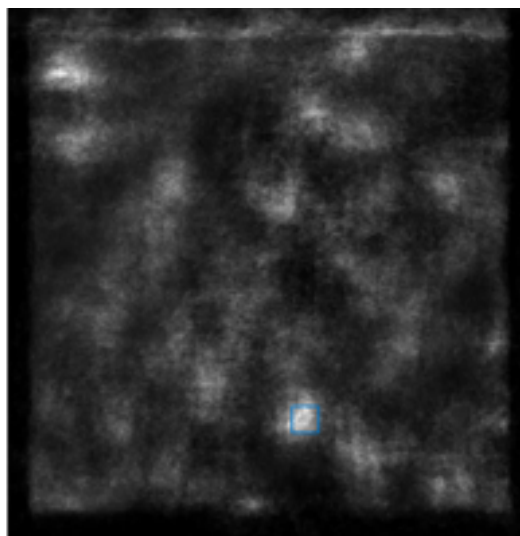


Figura 44: Inicialización cabeza

Una vez obtenida la región con mayor energía intra-región, Figura 44, se compara con la situación de la cabeza real, Figura 45, y se observa que la localización es completamente incorrecta, las dos regiones se encuentran muy distantes. Por tanto, queda demostrado que las imágenes de probabilidades no son perfectas y cuesta mucho encontrar las distintas partes del cuerpo.



Figura 45: Imagen original resaltando la cabeza

Al no se conseguir obtener el resultado querido, se propone encontrar las tres mejores opciones de una determinada imagen de probabilidad. Las tres localizaciones encontradas corresponden a las Figuras 46, 47 y 48 pero todas ellas quedan muy lejos de donde esta situada la cabeza.

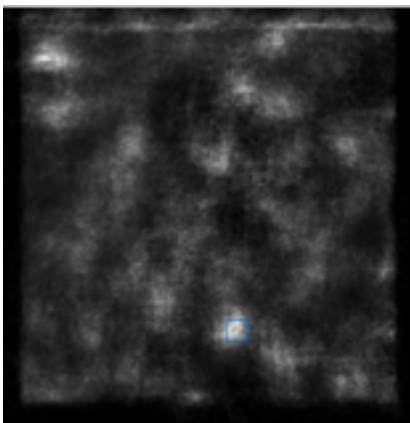


Figura 46: Inicialización cabeza 1

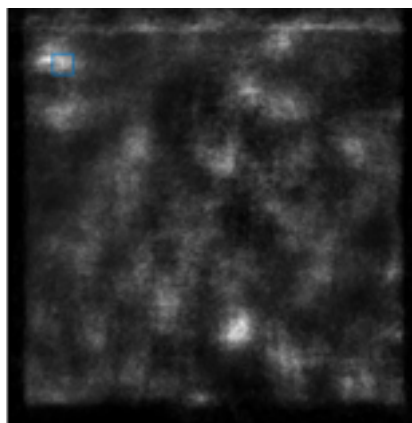


Figura 47: Inicialización cabeza 2

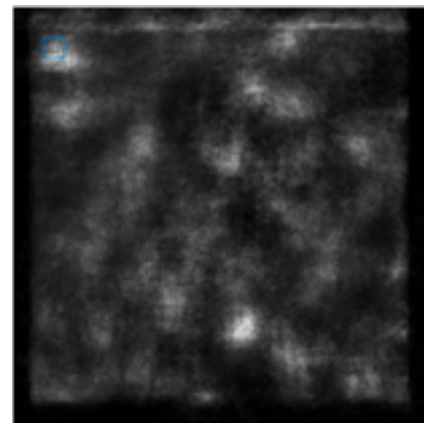


Figura 48: Inicialización cabeza 3

4.2.2 Inicialización de distancias

Para encontrar una primera relación entre las diferentes articulaciones que se van a tratar, utiliza como modelo el hombre de Vitruvio, ya que se dice que tiene las medidas perfectas.

De forma manual, se encuentran los centros de las articulaciones, que se encontrarán en el punto central de los cuadrados negros dibujados en la Figura 49:

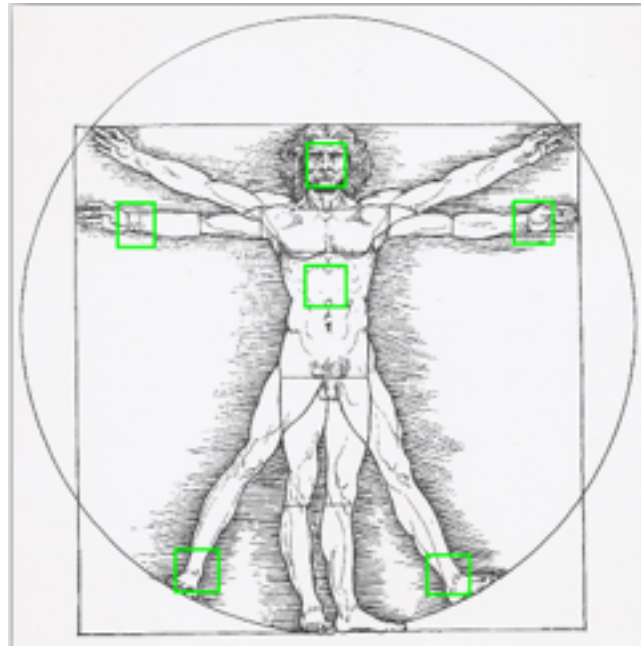


Figura 49: Articulaciones que se tendrán en cuenta del hombre Vitruvio

Obteniendo la siguiente relación entre el torso, nodo central, y el resto de articulaciones:

| | Cabeza | Muñeca Izquierda | Muñeca Derecha | Pierna Izquierda | Pierna Derecha |
|-------|--------|------------------|----------------|------------------|----------------|
| Torso | 0.79 | 1.02 | 0.84 | 0.98 | 0.96 |

Los resultados de la tabla vienen dados al dividir todas las distancias entre la altura del hombre (380 píxeles en el caso de la Figura 52). De esta forma, obtenemos las distancias aproximadas que debería tener si midiese un píxel de altura. Por lo tanto, tenemos las distancias calculadas de forma relativa. Por lo que si una persona tiene una altura de 80 px solo hace falta multiplicar todas las medidas por 80.

4.2.3 Energía Inter-Región

En las pruebas en las que se utiliza la energía inter-región no se tiene en cuenta la energía intra-región. Por lo tanto, no se tienen en cuenta las imágenes de probabilidad.

Para visualizar de forma más clara y simple las regiones utilizadas en las pruebas de la energía inter-región, se pintan las regiones en escala de grises sobre un fondo blanco, como se muestra en la Figura 50.



Figura 50: Modelo articulado

Para saber a qué corresponde cada región, a continuación, se ha creado la siguiente leyenda de colores:

| | |
|---|-------------------|
|  | Cabeza |
|  | Muñeca Izquierda |
|  | Muñeca Derecha |
|  | Torso |
|  | Tobillo Izquierdo |
|  | Tobillo Derecho |

Para calcular la energía inter-región es necesario conocer, aproximadamente, las distancias teóricas que deberían tener cada uno de los muelles, para poder realizar los cálculos. Estas distancias se calculan a partir del hombre de Da Vinci, como se explica en el apartado 3.2.2. Como ya se ha calculado un algoritmo que proporciona las distancias en forma relativa, en estas pruebas solo es necesario multiplicar estas distancias relativas por la altura, distancia entre cabeza y tobillos, de la persona para obtener las distancias teóricas.

Para la primera prueba indicamos que se calculen las distancias teóricas con una altura de la persona de 2 px. Esta prueba se realiza de forma exagerada para observar el comportamiento del algoritmo creado. El objetivo que se desea para este estudio es que todas las regiones se junten, aunque no queden completamente sobrepuestas, ya que la distancia es muy pequeña. A partir de la inicialización de la Figura 51, se obtiene el resultado de la Figura 52. En ésta se obtiene el resultado es el esperado al juntarse todas las regiones y tener una distancia entre los centros muy pequeña.



Figura 51: Imagen inicial 1

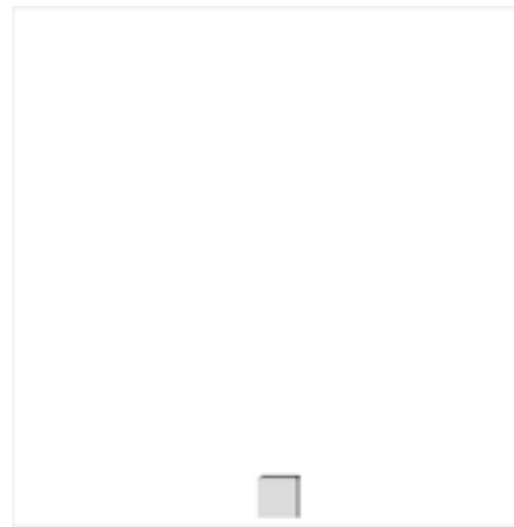


Figura 52: Imagen final 1

Si se realiza una prueba entre las distancias ideales entre centros y la que existe se puede ver que la diferencia es realmente pequeña:

| | Cabeza | Muñeca Izquierda | Muñeca Derecha | Pierna Izquierda | Pierna Derecha |
|------------------------------------|--------|------------------|----------------|------------------|----------------|
| Torso (Distancias Teóricas) | 0.79 | 1.02 | 0.84 | 0.98 | 0.96 |
| Torso (Distancias Reales) | 1.41 | 1.41 | 2.24 | 1.0 | 1.41 |

Para la siguiente prueba se aumenta la altura de la persona a 80px para obtener las distancias teóricas. El comportamiento esperado es que las regiones se junten hasta cumplir las distancias teóricas pero sin estar tan cerca como en el ejemplo anterior. A partir de la inicialización de la Figura 53 obtenemos el resultado de la Figura 54, que es bastante cercano a lo que se esperaba.

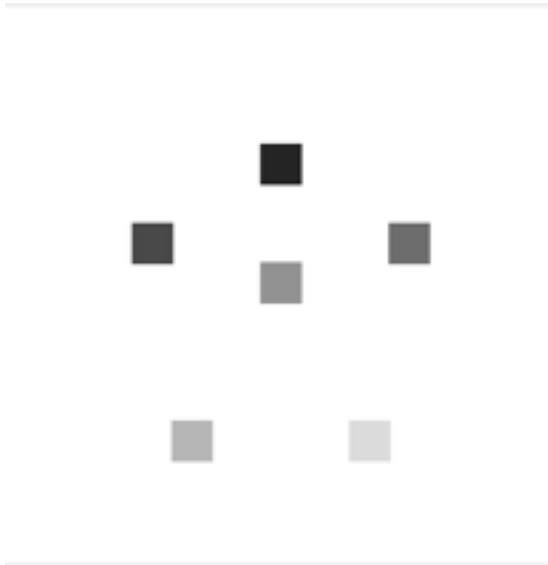


Figura 53: Imagen inicial 2



Figura 54: Imagen final 2

Al comparar las distancias teóricas con las reales, se puede observar que la diferencia entre unas y otras es muy pequeña. Por lo tanto, el algoritmo se está ejecutando de forma correcta.

| | Cabeza | Muñeca Izquierda | Muñeca Derecha | Pierna Izquierda | Pierna Derecha |
|------------------------------------|--------|------------------|----------------|------------------|----------------|
| Torso (Distancias Teóricas) | 31.53 | 40.62 | 33.7 | 39.22 | 38.58 |
| Torso (Distancias Reales) | 27.65 | 38.05 | 30.59 | 38.20 | 36.5 |

Para una tercera prueba, se calculan las distancias teóricas con una altura de la persona de 180px. Ahora, se espera que las regiones queden aún más separadas que en la prueba anterior. La inicialización se puede observar en la Figura 55, a partir de ésta se llega al resultado de la Figura 56. Donde, efectivamente, al poner distancias teóricas más elevadas, las regiones se van separando respecto al torso porque es el nodo central.



Figura 55: Imagen inicial 3

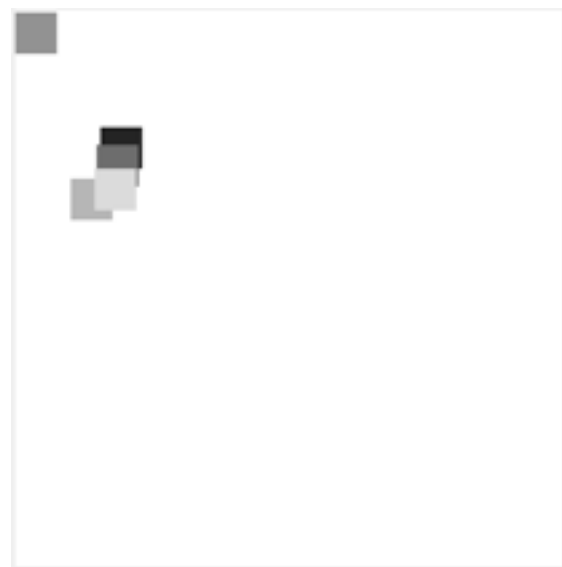


Figura 56: Imagen final 3

A simple vista, parece que el resultado es completamente correcto, como se ha podido observar. Pero si se comprueban las distancias teóricas con las reales se observa que una de las distancias está a 0.0. El motivo matemático de este comportamiento es desconocido por el momento, se sabe que tiene relación al coincidir los centros en la misma posición.

| | Cabeza | Muñeca Izquierda | Muñeca Derecha | Pierna Izquierda | Pierna Derecha |
|------------------------------------|--------|------------------|----------------|------------------|----------------|
| Torso (Distancias Teóricas) | 71.0 | 91.4 | 75.82 | 88.25 | 86.81 |
| Torso (Distancias Reales) | 72.2 | 0.0 | 78.54 | 88.54 | 88.54 |

Por último, se realiza una prueba con una altura de persona de 280px para corroborar el comportamiento que tienen los muelles. Se espera que queden las regiones aún más separadas que en el caso anterior. En la Figura 57 está la inicialización y en la 58 el resultado final. Éste continua siendo el esperado, a excepción de una región que vuelve a suceder lo mismo que en el caso anterior. Al ir desplazando las zonas en algún momento se ponen una encima de la otra y el resultado de la función es 0.



Figura 57: Imagen inicial 4

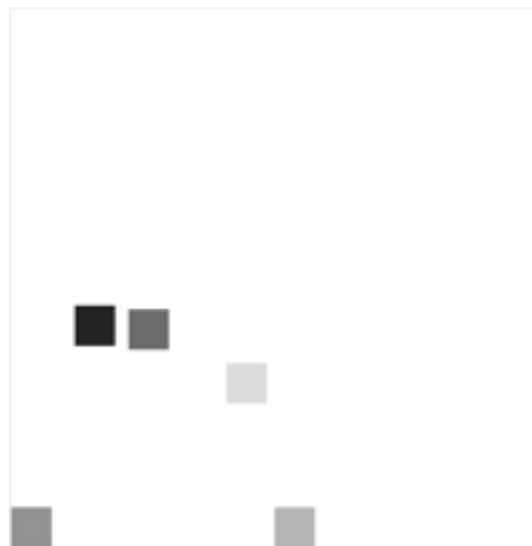


Figura 58: Imagen final 4

| | Cabeza | Muñeca Izquierda | Muñeca Derecha | Pierna Izquierda | Pierna Derecha |
|------------------------------------|--------|------------------|----------------|------------------|----------------|
| Torso (Distancias Teóricas) | 110.36 | 142.16 | 117.95 | 137.28 | 135.04 |
| Torso (Distancias Reales) | 110.06 | 0.0 | 119.7 | 137.0 | 134.79 |

Como conclusión, destacar que las distancias teóricas y reales solo varían en un margen de error de 3 píxeles, lo cual hace bastante fiable y preciso el algoritmo. Como mejora, en un futuro, se podría arreglar el error de la superposición de dos regiones para conseguir la perfección en el algoritmo.

4.2.4 Energía Total

La energía total tiene en cuenta tanto la energía intra-región como inter-región.

Antes de realizar las pruebas de la energía total se prepara una inicialización tanto de la energía intra-región como de la inter-región.

En el caso de la energía intra-región se calculan la localización de cada región, gracias a una ventana deslizante, la cual nos da la posición con máxima energía como se explica en el apartado 4.2.1.3. Para que la inicialización sea más exacta se calculan las tres mejores posiciones de cada región y se busca la mejor combinación entre todas las regiones. Es decir, se calculará la energía inter-región para cada combinación. Se determinará la combinación que dé una energía total mayor. Esta será la inicialización que utilizaremos para realizar las pruebas de la energía total, que en este caso corresponde a la Figura 59.



Figura 59: Inicialización intra-región

| | Cabeza | Muñeca Izquierda | Muñeca Derecha | Pierna Izquierda | Pierna Derecha |
|--|--------|------------------|----------------|------------------|----------------|
| Torso (Distancias Teóricas) | 118.24 | 152.31 | 126.37 | 147.09 | 144.68 |

Recordatorio de la leyenda de colores ya explicada en anteriores apartados, para saber a que corresponde cada una de las regiones:



En la inicialización mostrada en la Figura 59 se puede observar que únicamente hay 5 regiones pero realmente están las 6. Esto sucede a causa de que dos regiones se encuentran una encima de la otra de forma casual al calcular la inicialización. Como se muestra en las Figuras 60 y 61, las regiones se encuentran en el mismo lugar. Al realizar las pruebas siguen quedando una encima de la otra. Esto es debido a que la distancia entre cada una de ellas al torso es la misma, al ser una de estas regiones el tobillo izquierdo y la otra el tobillo derecho. Añadir que, al estar en la misma zona, tienen también la misma energía intra-región. Por lo tanto, al tener la misma intra-región e inter-región, tienen la misma energía total.

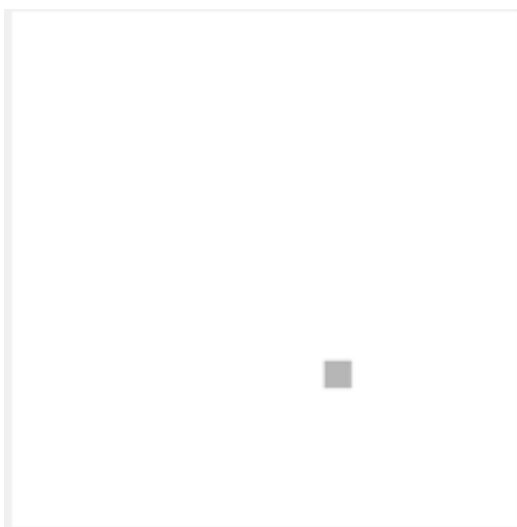


Figura 60: Inicialización tobillo izquierdo

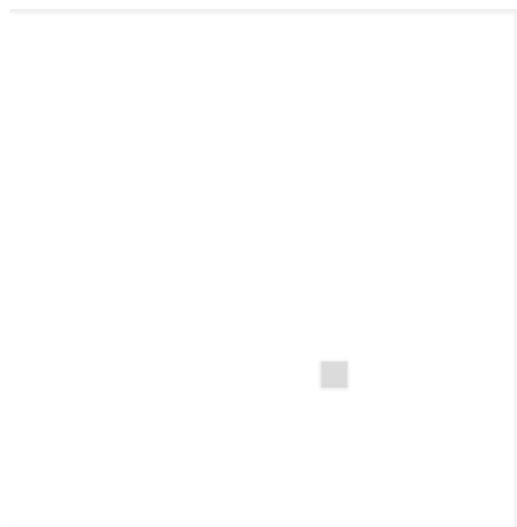


Figura 61: Inicialización tobillo derecho

A continuación, hay que realizar una inicialización de las distancias teóricas del modelo articulado, que nos servirán para los cálculos de la energía inter-región. La explicación de cómo se encuentran estas distancias está en el apartado 3.2.2. La inicialización de las distancias será con una altura de la persona de 300 px.

Una vez inicializado tanto las regiones como los muelles, ya está todo listo para poder realizar los cálculos de la energía total.

Para la energía total (Ecuación 32) de cada una de las regiones, hay que tener en cuenta que en la fórmula se puede observar un componente β . Este valor es el encargado de equilibrar el valor de la energía inter-región, con el fin de que se tengan en cuenta las dos energías.

$$E = \sum_{p \in \mathbb{R}} I(p) + \beta f(C_0, C_i)$$

Ecuación 32: Energía Total de cada región

En la primera prueba se considera un valor de $\beta = -1$. Para ver el comportamiento que tiene el algoritmo y saber si una energía está teniendo más importancia que la otra. Lo que se quiere conseguir es tener en cuenta por igual las dos energías. A partir de la inicialización de la Figura 62, se obtiene el resultado de la Figura 63. El cual, a priori, no tiene forma de persona,. Pero para saber si ha tenido en cuenta una energía más que otra nos fijamos en la Figura 64 que es el resultado de coger la misma inicialización pero teniendo solo en cuenta la energía inter-región. Por lo tanto, se confirma que solo se está teniendo en cuenta la energía inter-región en la energía total. Éste no es el comportamiento que se desea obtener.



Figura 62: Inicialización inicial

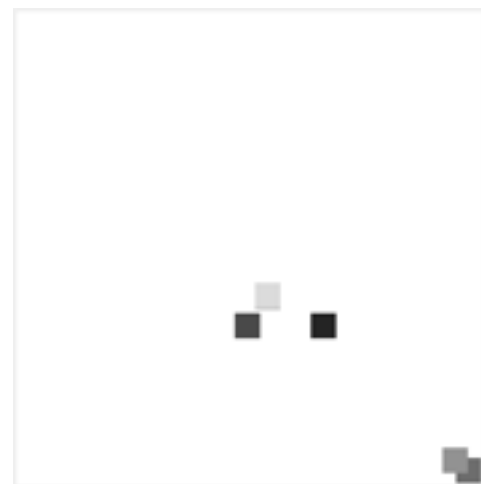


Figura 63: Resultado Energía Total 1



Figura 64: Resultado Energía Inter-Región

Con las observaciones de la prueba anterior se llega a la conclusión de que los valores que obtenemos de la energía inter-región son demasiado elevados comparados con los obtenidos de la intra-región.

Si ponemos un valor de β demasiado pequeño como por ejemplo -0.0000015 el comportamiento esperado sería justamente el contrario al de la prueba anterior. Se conseguiría que prácticamente la energía inter-región no tenga importancia y en cambio tenga mucha la intra-región. A partir de la inicialización de la Figura 65 se obtiene el resultado de la Figura 66, donde se puede observar que no tiene en cuenta prácticamente los muelles.



Figura 65: Inicialización inicial

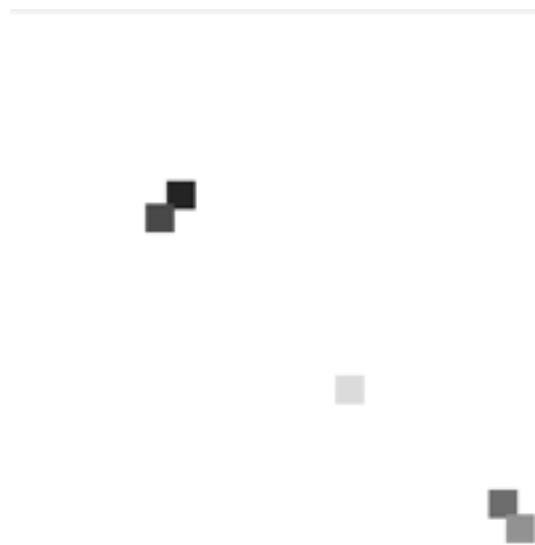


Figura 66: Resultado Energía Total 2

Para demostrar que la energía intra-región tiene mucho peso en esta prueba, nos fijamos en la Figura 67. Se ha dibujado la región final de la cabeza sobre su respectiva imagen de probabilidad. Se puede observar como se ha quedado la región en una zona de gran probabilidad. También se realiza la misma prueba en la Figura 68 pero teniendo en cuenta el cuello y su imagen de probabilidad.

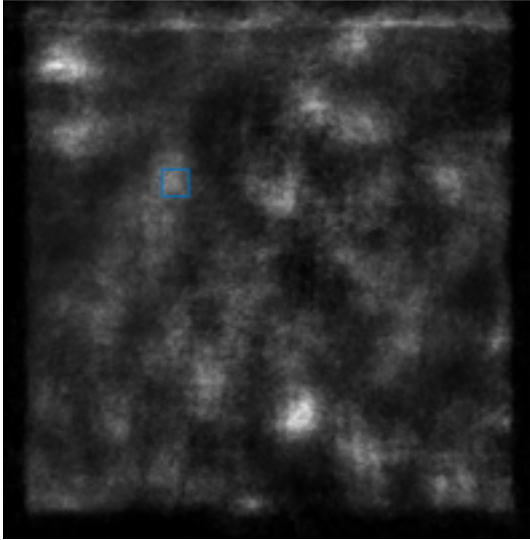


Figura 67: Región final de la cabeza

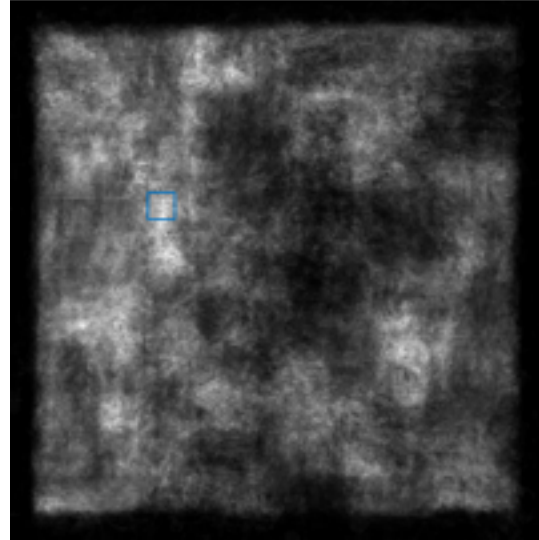


Figura 68: Región final del cuello

Para intentar equilibrar las dos energías, se varía el valor de β , utilizando un valor intermedio al de las dos pruebas realizadas anteriormente. En esta prueba le daremos un valor de -0.00015 , con este valor se espera un buen equilibrio de energías. A partir de la inicialización de la Figura 69 se obtiene el resultado de la Figura 70. Se observa que aun que tiene en cuenta la energía de los muelles también interfiere la energía intra-región. Se puede comprobar, a partir de los muelles mostrados en la Figura 67, que hay diferencias en las imágenes. Éstas son provocadas por la energía intra-región.



Figura 69: Inicialización inicial

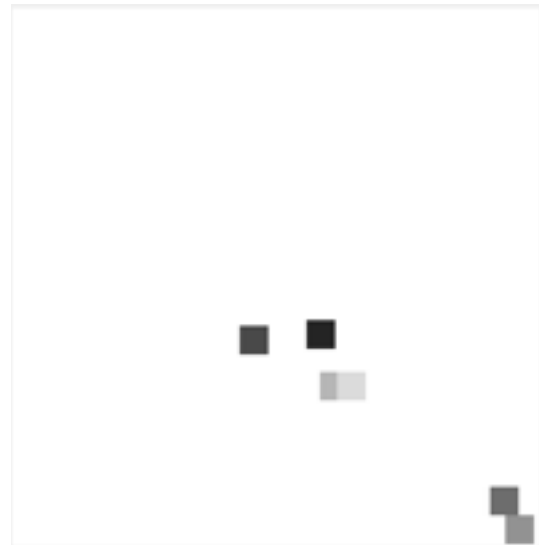


Figura 70: Resultado Energía Total 3

Se puede afirmar que se ha conseguido el comportamiento esperado: encontrar un valor de β que consiga un equilibrio entre las dos energías. Aunque el objetivo principal del trabajo es obtener la silueta de la persona, la cual a simple vista se puede afirmar que no cumple. Si se observa el modelo realizado de forma sintética de la Figura 71 y se compara con el resultado final de la Figura 72, no hay parecido alguno.



Figura 71: Modelo articulado sintético



Figura 72: Modelo articulado de la prueba

Se intenta una solución, para facilitar al algoritmo encontrar un resultado correcto. Consiste en realizar la inicialización de las regiones a un lugar cercano a donde están situadas realmente. Esta inicialización se realiza de forma manual para colocar las regiones tal i como se muestra en la Figura 73. Pero, aun así, el resultado continúa siendo incorrecto como se muestra en la Figura 74.



Figura 73: Modelo articulado sintético



Figura 74: Modelo articulado de la prueba

El algoritmo realiza los pasos deseados pero las imágenes de probabilidad no ayudan a que funcione correctamente, por ello es muy complicado que sea posible obtener un buen resultado.

4.2.5 Energía Total con Imágenes Sintéticas

Tal y como se ha explicado en el apartado anterior, las imágenes de probabilidad son muy complicadas de tratar. Esto provoca que adaptar un modelo articulado sea prácticamente imposible.

Para comprobar que el algoritmo funciona, pese a las imágenes de probabilidad, se crean unas imágenes de forma sintética a partir del proceso de dilatación, Figura 75. Estas imágenes son guardadas en `cmap="gray"` para poder tratarlas como imágenes de probabilidad, Figura 76.

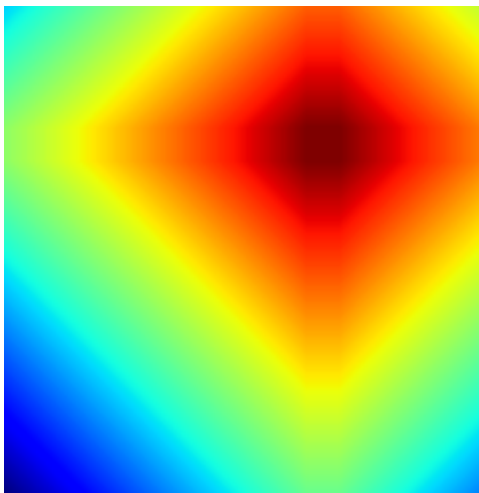


Figura 75: Imagen de probabilidad sintética

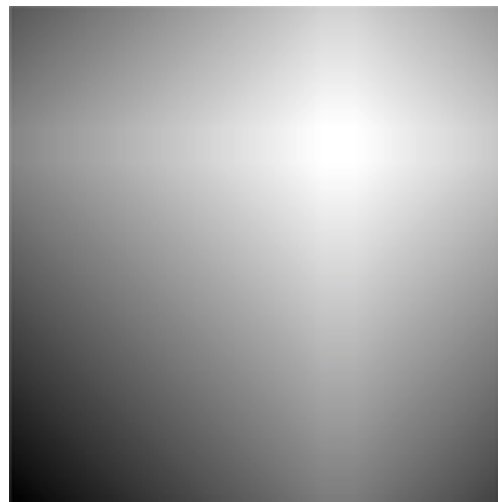


Figura 76: Imagen de probabilidad sintética con `cmap="gray"`

A partir de estas imágenes de probabilidad se realiza el algoritmo matemático explicado en el capítulo 3.

Para una primera prueba se utiliza la inicialización mostrada en la Figura 77. El comportamiento esperado al utilizar las imágenes de probabilidad sintéticas es que la prueba sea un éxito, a partir de esta inicialización debería de adaptarse las regiones al lugar correcto, sin gran variación, ya que al comparar la Figura 77 con la imagen original 78, se comprueba que las regiones se encuentran cercanas a la realidad.



Figura 77: Modelo articulado sintético



Figura 78: Modelo articulado encima de la imagen original.

A continuación, se aplica el algoritmo y como resultado se obtiene la Figura 79. En dicha figura se observa que las regiones quedan adaptadas, pero hay un problema con las muñecas y torso, esto es debido a que las distancias teóricas entre el torso y las muñecas es mayor a la que existe en la imagen que se ha seleccionado. En el caso de tratar un número mayor de regiones se tendría en consideración el codo y se podría obtener el modelo correcto.



Figura 79: Resultado Energía Total

En el caso de realizar una mala inicialización de las distancias, colocar una altura de la persona en 50px, se puede observar que sigue funcionando correctamente el algoritmo al juntarse todas las regiones a causa de los muelles, Figura 80. Aunque la inicialización de las regiones sea correcta, Figura 77.



Figura 79: Resultado Energía Total 2

Cuando se intentaba calcular la energía total en el capítulo 4.2.1.3 se inicializaban las regiones con la ayuda de la energía intra-región. Esta inicialización que se obtenía no era daba un resultado correcto, pero si se intenta realizar la misma prueba pero con las imágenes sintéticas se obtiene una inicialización de regiones como la de la Figura 80. Se puede observar que ahora si podemos afirmar que funciona ya que sitúa de forma bastante exacta donde se encuentran las regiones. El tobillo izquierdo y el derecho se encuentran solapados, podría haber sucedido porque se encuentran uno muy cerca uno del otro.

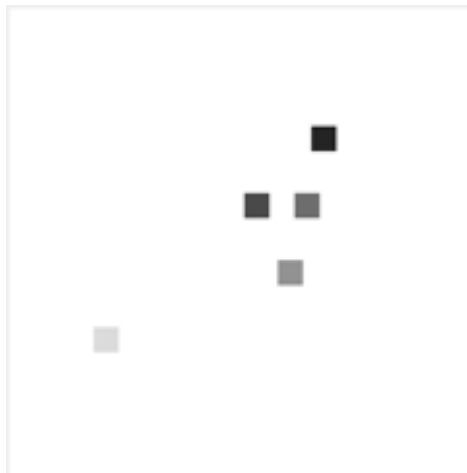


Figura 80: Inicialización de las regiones a partir de las imágenes sintéticas

Se puede llegar a la conclusión que efectivamente influye, de forma notable, las imágenes de probabilidad con las que se parte inicialmente.

5. Conclusiones y Líneas Futuras

5.1 Conclusiones

Se ha desarrollado un algoritmo que permite adaptar modelos articulados a partir de un conjunto de imágenes de probabilidad, las cuales dan información sobre las distintas partes del cuerpo. Los resultados de este algoritmo han realizado el comportamiento esperado, pero no han sido satisfactorios. Esta insatisfacción viene producida por la complicidad de trabajar con las imágenes de probabilidad, de las cuales se partía en el trabajo.

Pese que no encuentra de forma satisfactoria a la persona de la imagen, hay que destacar que el comportamiento que realiza el algoritmo sobre las imágenes es correcto y todas las pruebas realizadas hacen lo que se supone que debería ocurrir si únicamente se supiese los puntos máximos de probabilidad.

Al realizar el algoritmo se han tenido en cuenta dos tipos de energías, en las cuales sus fórmulas correspondientes dan valores muy distintos. Mientras, el resultado de la intra-región da valores pequeños, el de la inter-región da resultados demasiado elevados. Esto implica que es muy importante encontrar un equilibrio entre estas dos. Para ello depende en gran parte de las imágenes de probabilidad, ya que las imágenes de la base de datos y las creadas sintéticamente provocan una variación en la energía intra-región. Por este motivo hay que modificar la cantidad de importancia que se le da a una energía y a otra dependiendo de la situación.

Es realmente necesario disponer de estas dos energías porque se ha podido comprobar que la energía de la propia región por si sola no es suficiente. Tienen que existir unas restricciones que ayuden a no perder la forma la persona.

De haber dispuesto de mayor tiempo, habría sido muy interesante poder aplicarle al modelo articulado una textura. Objetivo que no se ha podido implementar por falta de tiempo. Pero podría ser una línea futura que mejoraría el proyecto. No obstante, la texturización se ha analizado de forma teórica en el capítulo 5.2.

También, en un futuro se podría mejorar el algoritmo de aprendizaje que ha generado las imágenes de probabilidad, para evitar que existan tantos falsos positivos. De esta forma los resultados serían mucho más óptimos, como se ha demostrado en las últimas pruebas del capítulo 4.2.5 al tener unas imágenes con una mejor exactitud.

5.2 Líneas Futuras: Texturización

Las líneas futuras se centran en realizar una texturación del modelo articulado obtenido.

El objetivo de la texturación consiste en coger una textura de otra persona y aplicar está al modelo articulado obtenido. De tal forma que, cualquier persona que se encuentre en una cierta postura se deforma hasta obtener la postura del modelo articulado al que aplicamos la texturización.

Una vez se obtiene el modelo articulado de la persona, como por ejemplo el de la Figura 81, ya se puede realizar una texturización de éste.

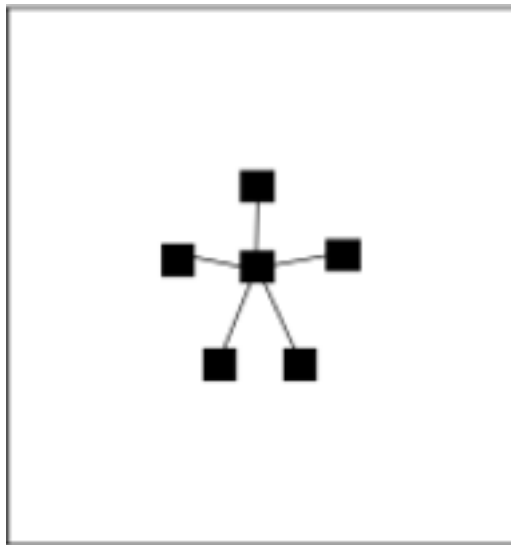


Figura 81: Modelo articulado

El primer paso consiste en crear una malla que envuelva todas las regiones, y obtener a partir de esta malla un polígono con la forma de la silueta como el de la Figura 82.

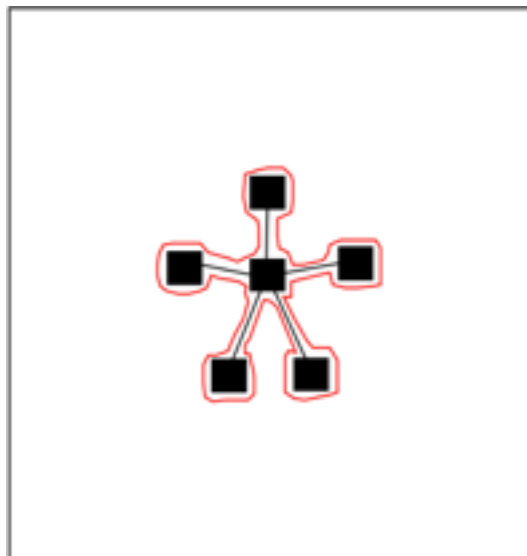


Figura 82: Malla del modelo articulado

En la Figura 82, se puede observar como la malla aún teniendo que tiene forma de persona puede ser demasiado simple para realizar una texturización. Quedaría mucho mejor con una malla mejor definida, esto se puede conseguir si el número de regiones es mayor como en la Figura 83. Una mejora aún superior sería una adaptación de la región más concreta si los vértices se mueven libremente.

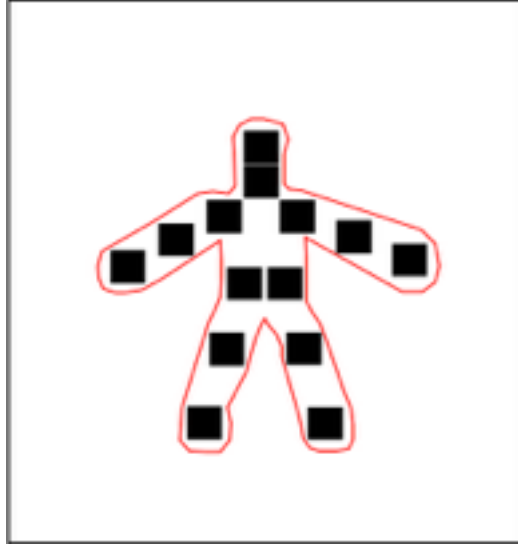


Figura 83: Malla del modelo articulado complejo

Para cada uno de los puntos que forman parte del polígono, que envuelve el modelo articulado, malla roja de la Figura 83, se tiene que encontrar a qué punto corresponde en la textura. En la imagen de la textura, también, es necesario obtener el polígono que envuelva la persona con la misma cantidad de vértices para relacionarlo.

Para encontrar la relación entre un punto del modelo articulado y el punto al que corresponde en la textura, se tiene que calcular las coordenadas afines de la imagen inicial y multiplicarlas por los vértices del polígono de la imagen de la textura. De esta forma, se encuentra una relación entre las dos mallas. Para ello se utiliza la derivada de los puntos (Ecuación 33), donde α corresponde a las coordenadas afines de la posición obtenida por todo el algoritmo, mientras que V_i son cada uno de los vértices del polígono que forma la silueta en la textura.

$$\frac{\partial p}{\partial d} = \sum_{i=1}^n \alpha_i(p) * V_i$$

Ecuación 33: Malla del modelo articulado complejo

De forma sintética, se puede observar que si tenemos una imagen inicial, como la de la Figura 84. Seguidamente se calcula el modelo articulado, Figura 85, es posible aplicar una textura de ejemplo como la de la Figura 81. El resultado obtenido será la Figura 87, la silueta de la Figura 84 pero con la textura de la Figura 86.



Figura 84: Imagen inicial sintética

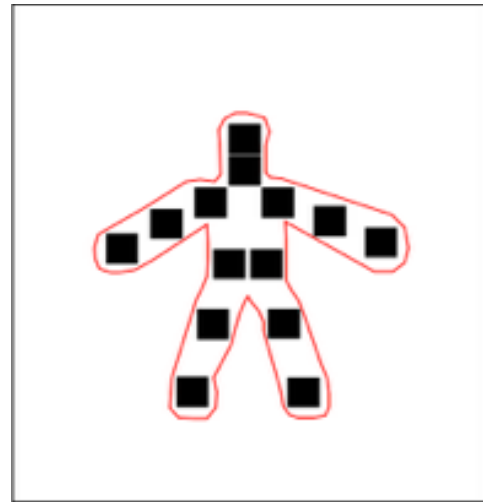


Figura 85: Modelo articulado



Figura 86: Textura sintética



Figura 87: Aplicación de la textura sintética en el modelo articulado

La parte de texturación sería una línea futura del proyecto, por lo tanto, no se han realizado pruebas reales sobre este apartado.

6. Bibliografía

[1] Pedro F. Felzenszwalb y Daniel P. Huttenlocher, “Pictorial Structures for Object Recognition”, Septiembre 2004.

[2] Lluís Garrido, Marité Guerrieri y Laura Igual, “Image Segmentation with Cage Active Contours”, Diciembre 2012.

Apéndice: Estructura código

El código esta escrito en Matlab y esta estructurado en diversas clases que son las explicadas a continuación:

- main: es la clase principal del código. En esta función hay un valor beta que se modifica para dar mayor o menor importancia a la energía inter-región, también hay una variable beta que es posible de modificar para controlar que cantidad de pixeles se mueve cada uno de los vértices.
- mean_value: calcula las coordenadas afines de un punto P respecto un conjunto de vértices C.
- e_cuadrado: es la función encargada de calcular la energía intra-región de una región concreta.
- máximo: Función que con la ayuda de una ventana deslizante es capaz de encontrar las tres mejores localizaciones, es decir, donde se encuentra una mayor energía intra-región. Para ello calcula en cada localización posible con la función e_cuadrado la energía intra-región.
- inicializar: desde esta función se realizan tantas llamadas como regiones a la función máximo(). Para cada una de las llamadas a la función se pasa por parámetros la imagen de probabilidad correspondiente. Una vez obtenido los vértices y centros de todas las regiones se guardan en una estructura de matriz, una matriz para vértices y otra para centros.
- distancias: Inicialización de los muelles a partir de una imagen correspondiente al hombre de Da Vinci, también calcula todos los centros de cada una de las regiones obtenidas. Las distancias obtenidas son por pixel de altura de la persona, por tanto, están en función de la altura.
- combination: Recibe las tres mejores localizaciones para cada una de las regiones, a continuación calculando la energía total para cada una de las combinaciones se queda con la de mayor energía total.
- interRegion: calcula la derivada de la energía inter-región tanto en desviación en x, como en y.
- main2: es una clase principal secundaria, con ella se realizan todas las pruebas que utilizan imágenes sintéticas.
- Imágenes sintéticas de probabilidad.ipynb: Notebook escrito en python que se encarga de la creación de las imágenes sintéticas creadas para utilizarlas como imágenes de probabilidad.

Para ejecutar el código solo hay que darle al play desde Matlab a la clase principal main. Mientras que si se quiere observar pruebas con las imágenes creadas de forma sintética, hay que darle al play desde la clase main2.