

SEGMENTACIÓN DE VIDEO

PARTE 1

Francisco J. Hernández López

fcoj23@cimat.mx



SEGMENTACIÓN DE VIDEO

- Parte importante en muchas de las aplicaciones de visión computacional tales como: video conferencias, video vigilancia, edición de video, etc.
- Consiste en partir el video en regiones espaciales, temporales o espacio-temporales que son homogéneas en alguna característica (intensidad, color, textura, movimiento, etc.)
- Requiere de una apropiada medida de distancia para comparar las características seleccionadas. Diferentes características y criterios generalmente dan diferentes resultados en la segmentación

FACTORES A CONSIDERAR

- **Interacción con el usuario**
Para introducir conocimiento adicional en la segmentación, esto típicamente da mejores resultados
- **Rendimiento en tiempo-real**
El tiempo disponible para lograr la segmentación restringe las técnicas que se pueden utilizar
- **Complejidad del contenido**
Podemos tener diferentes soluciones y análisis dependiendo de la complejidad tanto espacial como temporal de un video. Escenas complejas requieren sofisticados algoritmos de segmentación
- **Calidad en la segmentación**
Para algunas aplicaciones es necesario dar una segmentación muy precisa (contornos bien definidos, etc.) y para otras basta con una caja envolvente que cubra el objetivo por ejemplo

Paulo L Correia and Fernando Pereira. Classification of video segmentation application scenarios. Circuits and Systems for Video Technology, IEEE Transactions on, vol. 14, no. 5, pages 735-741, 2004.

Segmentación de video. Francisco J. Hernández-López

Agosto-Diciembre 2016

CONTORNOS ACTIVOS

- Curvas que tratan de ajustarse al contorno de los objetos, modelan las fronteras entre los objetos y el fondo de la imagen
- Se pueden clasificar en:
 - Snakes → (Kass et. al. 1988) Minimización de una energía con una curva spline 2D, la cual evoluciona (se mueve) a través de las características de la imagen (bordes fuertes)
 - Intelligent scissors → (Mortensen and Barrett 1995) Permite al usuario dibujar en tiempo real una curva que se adhiera al contorno del objeto
 - Level set → (Osher and Sethian 1988) La idea es representar la evolución del contorno usando una función característica, permite cambiar la topología e incorporar estadística basada en regiones

Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 1(4), 321–331.

Mortensen, E. N. and Barrett, W. A. (1995). Intelligent scissors for image composition. In *ACM SIGGRAPH 1995 Conference Proceedings*, pages 191–198.

Osher, Stanley, and James A. Sethian. "Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations." *Journal of computational physics* 79.1 (1988): 12-49.

Segmentación de video. Francisco J. Hernández-López

Agosto-Diciembre 2016

SNAKE

- Involucra la minimización de una energía interna

$$\mathcal{E}_{int} = \int \alpha(s) \left\| \frac{\partial f(s)}{\partial s} \right\|^2 + \beta(s) \left\| \frac{\partial^2 f(s)}{\partial s \partial s} \right\|^2 ds$$

donde s es la longitud de arco a lo largo de la curva

$$f(s) = (x(s), y(s))$$

$\alpha(s)$ y $\beta(s)$ son funciones peso de la primera y segunda derivada

- Podemos discretizar esta energía de la siguiente manera

$$E_{int} = \sum_i \alpha(i) \frac{\|f(i+1) - f(i)\|^2}{h^2} + \beta(i) \frac{\|f(i+1) - 2f(i) + f(i-1)\|^2}{h^4}$$

donde h es el tamaño de paso

SNAKE (C1)

- Además involucra una energía externa

$$\mathcal{E}_{image} = w_{line}\mathcal{E}_{line} + w_{edge}\mathcal{E}_{edge} + w_{term}\mathcal{E}_{term}$$

donde:

- *line* → atrae al snake a crestas oscuras
 - *edge* → atrae al snake a gradientes fuertes (bordes)
 - *term* → atrae al snake a terminaciones de línea
- En la practica se suele utilizar solo el termino *edge*, el cual puede ser:
 - $E_{edge} = \sum_i -\|\nabla I(f(i))\|^2$
 - $E_{edge} = \sum_i -\|(G_\sigma * \nabla^2 I)(f(i))\|^2$

SNAKE (C2)



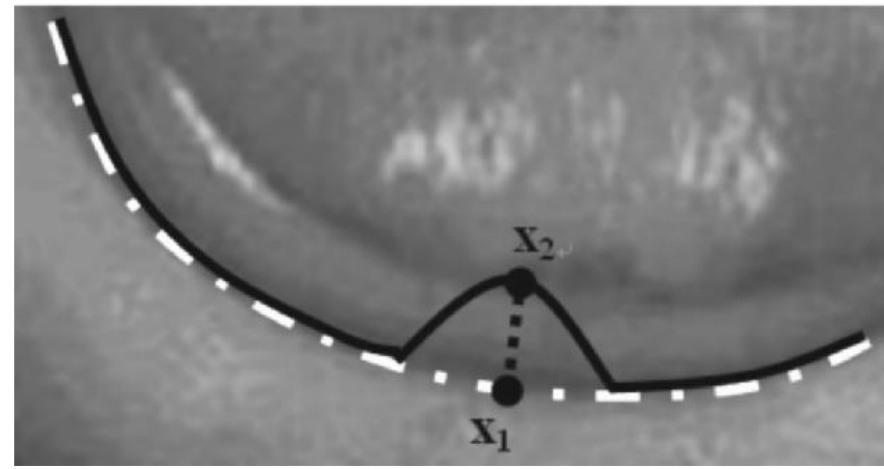
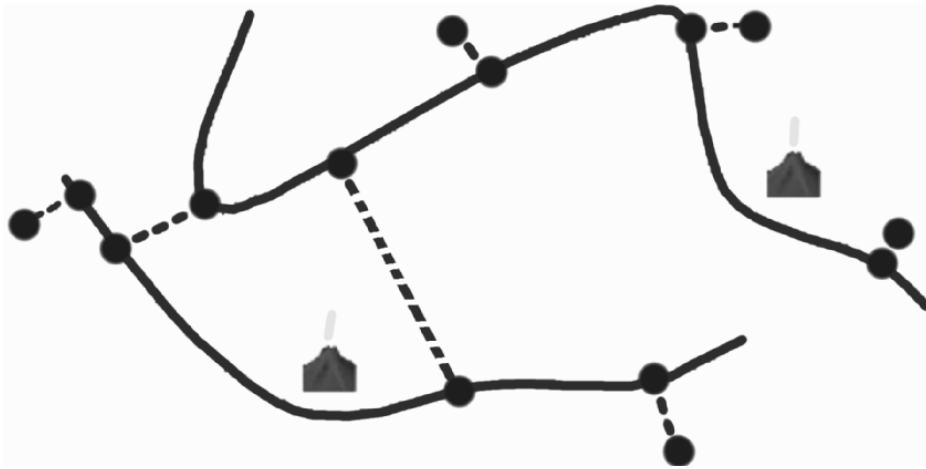
Lip tracking (Kass et al. 1988)

- Aplicación interactiva: “snake pit”

- Resortes en puntos de anclaje (Fuerzas de atracción)

$$E_{spring} = k_i \|f(i) - d(i)\|^2$$

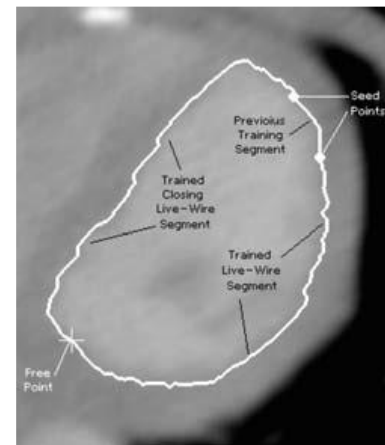
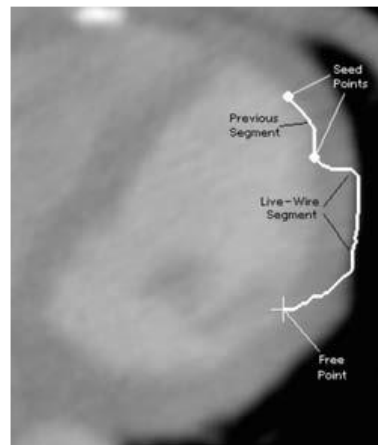
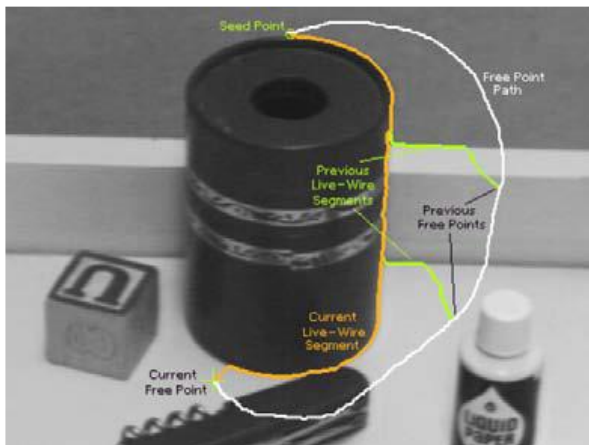
- Volcanes en puntos de anclaje (Fuerzas de repulsión) $1/r$



Farag, A. Deformable Models: Theory and Biomaterial Applications, Topics in Biomedical Engineering. International Book Series, Springer, 2007

INTELLIGENT SCISSORS

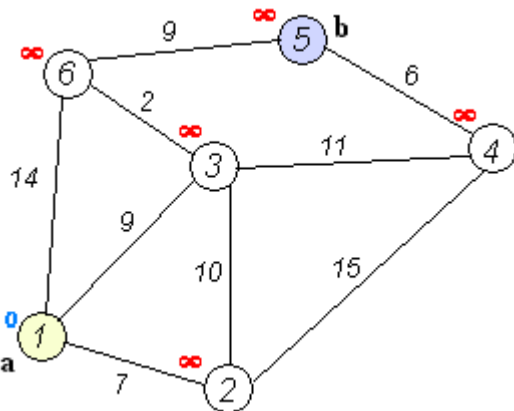
- La idea es ir optimizando el contorno del objeto mientras el usuario dibuja la curva
- El sistema trata de adherir la curva dibujada por el usuario a los contrastes altos (bordes)



Mortensen and Barrett 1995

INTELLIGENT SCISSORS (C1)

- Calcula el camino de menor costo entre un punto semilla inicial y la posición actual del apuntador del mouse
- Pre-procesamiento para asociar costos bajos a los bordes que son altamente probables de estar en el contorno
- Calcular el costo más bajo para ir de un punto a otro usando Dijkstra (Algoritmo de programación dinámica breadth-first)



https://es.wikipedia.org/wiki/Algoritmo_de_Dijkstra

<http://staff.ustc.edu.cn/~xjchen99/teaching/ImageUnderstanding/ImageUnderstanding.html>

Richard Szeliski. 2010. Computer Vision: Algorithms and Applications (1st ed.). Springer-Verlag New York, Inc., New York, NY, USA.

Segmentación de video. Francisco J. Hernández-López

Agosto-Diciembre 2016

INTELLIGENT SCISSORS USANDO GIMP



1. Damos los puntos para dibujar la curva
2. Cerramos la curva dibujando el último punto sobre el primero
3. Damos click en la parte interna del contorno calculado
4. Obtenemos el recorte



LEVEL SET

- Una limitante de los métodos de contornos activos basados en parametrizar la curva de la forma $f(s)$ es que llega a ser todo un reto *modificar la topología* de la curva a medida que evoluciona y además puede ser necesario una *re-parametrización*
- Una alternativa para contornos cerrados es usar level set
 - La curva es definida por los cruces por cero de una función característica $\phi(x, y)$
 - Evolucionan para ajustarse a los objetos de interés modificando $\phi(x, y)$ en lugar de la curva $f(s)$

CONTORNOS ACTIVOS GEODÉSICOS

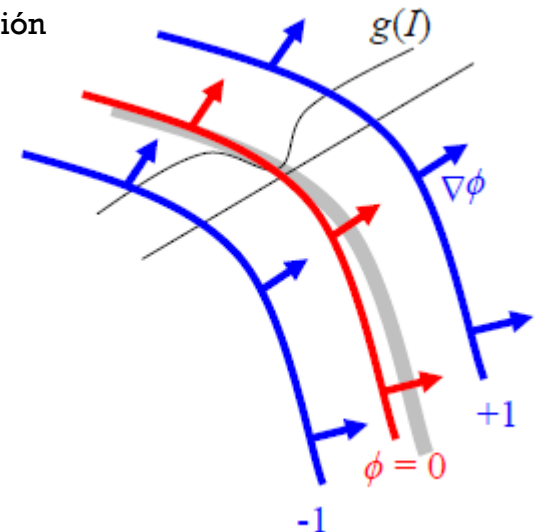
- Minimizan la longitud de una curva en \mathbb{R}^2 con la siguiente ecuación diferencial parcial

$$\frac{d\phi}{dt} = |\nabla\phi| \operatorname{div} \left(g(I) \frac{\nabla\phi}{|\nabla\phi|} \right) = \underbrace{g(I) |\nabla\phi| \operatorname{div} \left(\frac{\nabla\phi}{|\nabla\phi|} \right)}_{\text{Mueve la curva en dirección de su curvatura}} + \underbrace{\nabla g(I) \cdot \nabla\phi}_{\text{Mueve la curva en dirección del gradiente}}$$

con

$g(I) \rightarrow$ versión generalizada del potencial de bordes

$$E_{edge} = \sum_i -\|\nabla I(f(i))\|^2$$

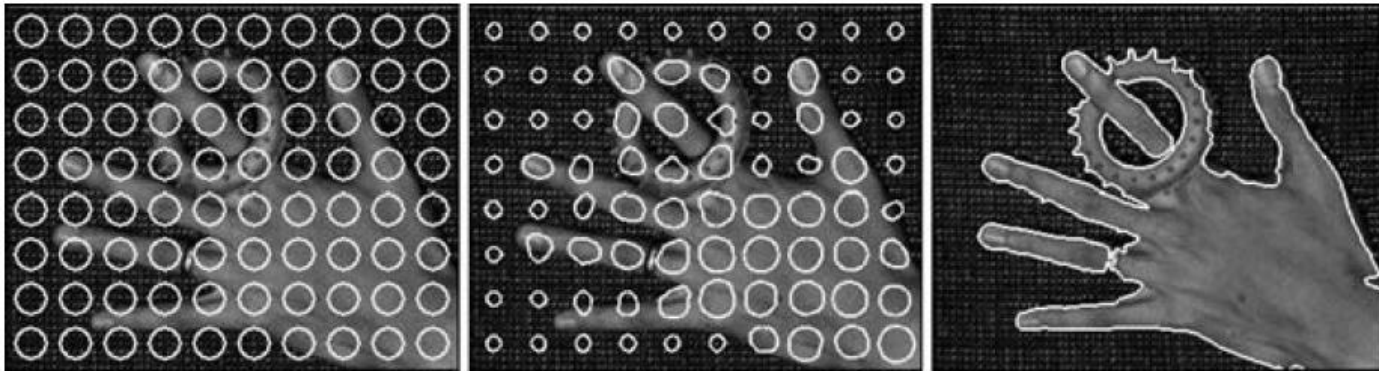


Caregnato, Alejandro Frangi, and Vicent Caselles Costa. "La segmentación de imágenes: el método de los contornos activos geométricos." *Encuentros multidisciplinares* 8.23 (2006): 14-26.

Segmentación de video. Francisco J. Hernández-López

Agosto-Diciembre 2016

APLICACIONES



Segmentación con Level Set, Cremers et. al. 2007

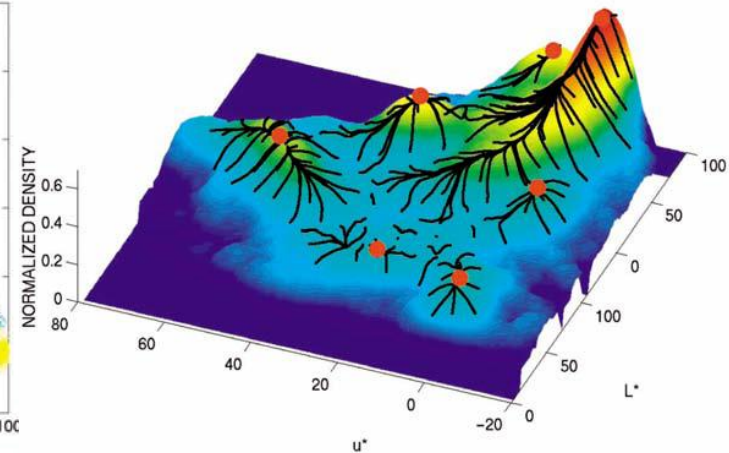
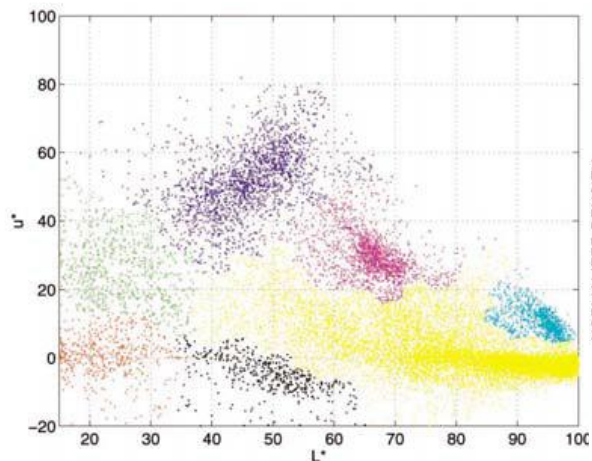


Rotoscopia, Agarwala et. al. 2004



MEAN SHIFT

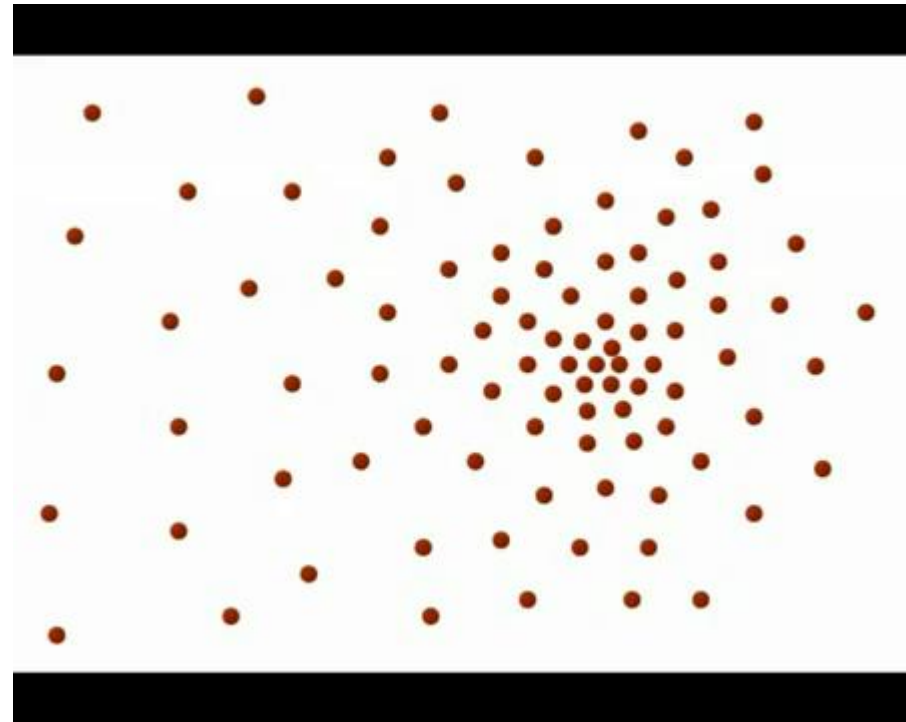
- Introducido originalmente por Fukunaga y Hostetler, 1975
- Modela el vector de características (asociado a un pixel de la imagen) como muestras a partir de una función de densidad de probabilidad desconocida, con la finalidad de encontrar agrupamientos (clusters) en dicha distribución



Segmentación de imágenes usando mean shift, Comaniciu and Meer 2002

MEAN SHIFT (C1)

- Evita la estimación de la función de densidad de probabilidad y consiste en dos pasos:
 1. Filtro preservando discontinuidad
 2. Agrupamiento
- Idea:
 - Se posiciona (de forma aleatoria) una región de interés S con centro \vec{c}
 - Se calcula el centro de masa \vec{c}_m de S
 - Movemos \vec{c} hacia \vec{c}_m



<https://www.youtube.com/watch?v=kmaQAsotT9s>

MEAN SHIFT (C2)

- Parámetros: Tamaño y forma de la región de interés S , o definición de un kernel simétrico

$K(\vec{x}) = \alpha k(\|\vec{x}\|^2)$, con α una constante positiva que hace que la integral de $K(\vec{x})$ sea 1:

- $K_N(\vec{x}) = \alpha \exp\left(-\frac{1}{2}\|\vec{x}\|^2\right)$, **Kernel Gaussiano**
- $K_E(\vec{x}) = \begin{cases} \alpha(1 - \|\vec{x}\|^2) & \text{si } \|\vec{x}\| \leq 1, \\ 0 & \text{otro caso,} \end{cases}$ **Kernel Epanechnikov**
- Dado n puntos \vec{x}_i de dimensión d , el estimador $\tilde{f}_{h,K}$ en el punto \vec{x} es

$$\tilde{f}_{h,K}(\vec{x}) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{\vec{x} - \vec{x}_i}{h}\right),$$

donde h representa el tamaño del kernel

MEAN SHIFT (C3)

- Queremos identificar el punto \vec{x} para el cual

$$\nabla \tilde{f}_{h,K}(\vec{x}) = 0$$

- Considerando el kernel Gaussiano

$$\tilde{f}_{h,K}(\vec{x}) = \frac{1}{nh^d} \sum_{i=1}^n \alpha \exp\left(-\frac{1}{2} \frac{\|\vec{x} - \vec{x}_i\|^2}{h^2}\right),$$

- La derivada es

$$\nabla \tilde{f}_{h,K}(\vec{x}) = \frac{\alpha}{nh^{d+2}} \sum_{i=1}^n -(\vec{x} - \vec{x}_i) \exp\left(-\frac{1}{2} \frac{\|\vec{x} - \vec{x}_i\|^2}{h^2}\right) = 0$$

- Entonces iteramos la siguiente **relación de recurrencia**:

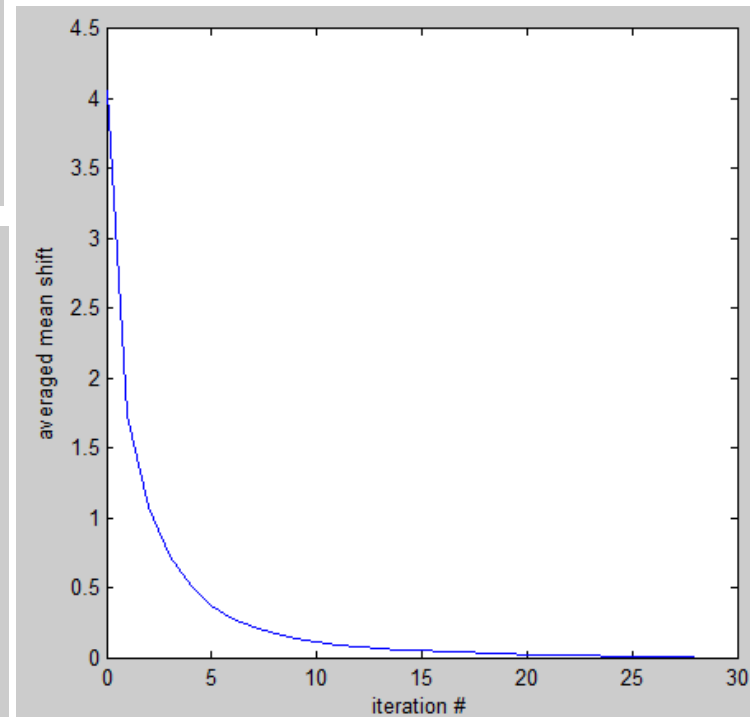
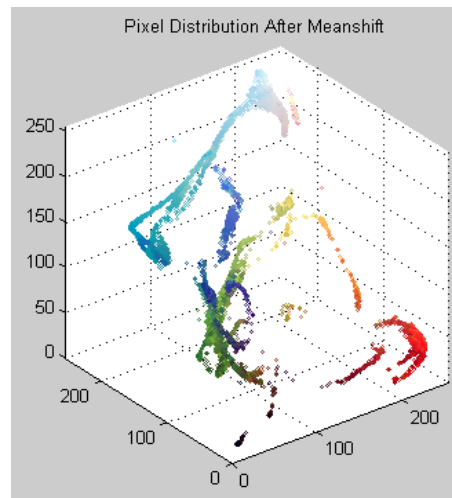
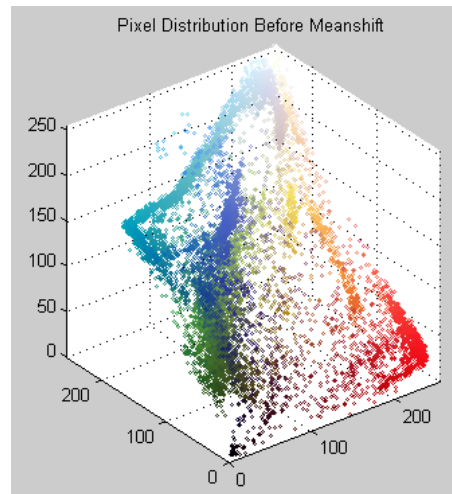
$$\vec{x}_{t+1} = \frac{\sum_{i=1}^n \vec{x}_i \exp\left(-\frac{1}{2} \frac{\|\vec{x}_t - \vec{x}_i\|^2}{h^2}\right)}{\sum_{i=1}^n \exp\left(-\frac{1}{2} \frac{\|\vec{x}_t - \vec{x}_i\|^2}{h^2}\right)}$$

ALGORITMO MEAN SHIFT PARA SEGMENTACIÓN

1. Para cada pixel \vec{x}_i (con \vec{x}_i^S la posición espacial) en la imagen I , inicializar $t = 1$ y $\vec{x}_{i,0} = \vec{x}_i$
2. Calcular $\vec{x}_{i,t+1}$ con la **relación de recurrencia** definida anteriormente hasta convergencia
3. Calcular la imagen filtrada: $F(\vec{x}_i^S) = I(\vec{x}_{i,t+1}^S)$
4. Determinar los clusters $\{C_p\}_{p=1,\dots,m}$ agrupando los valores de F , es decir, fusionando las regiones concentradas en los puntos de convergencia
5. Asignar etiquetas: $L(\vec{x}_i^S) = \{p | F(\vec{x}_i^S) \in C_p\}$
6. Pos-procesamiento: Eliminar las regiones más pequeñas a P píxeles (Fusionarla con la región adyacente)

MEAN SHIFT PIXEL CLUSTER

Yue Wu, 2013 (Actualización: 02/Dic/2013)



<http://www.mathworks.com/matlabcentral/fileexchange/40990-mean-shift-pixel-cluster/content/meanShiftPixCluster.m>

Segmentación de video. Francisco J. Hernández-López

Agosto-Diciembre 2016