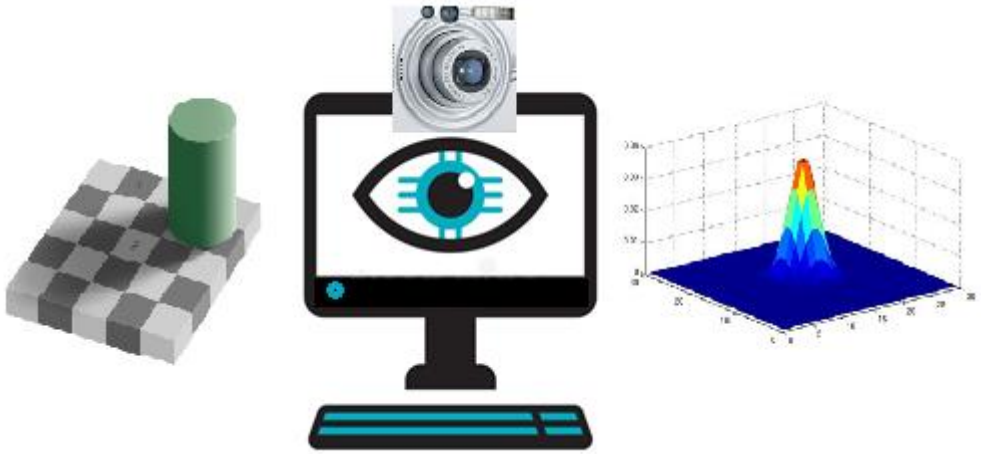


Computer Vision
Lecture 11
Segmentation & Grouping

Dr. Ali Mahmoud Mayya

Computer Science Dept.

AL Manara University, Syria

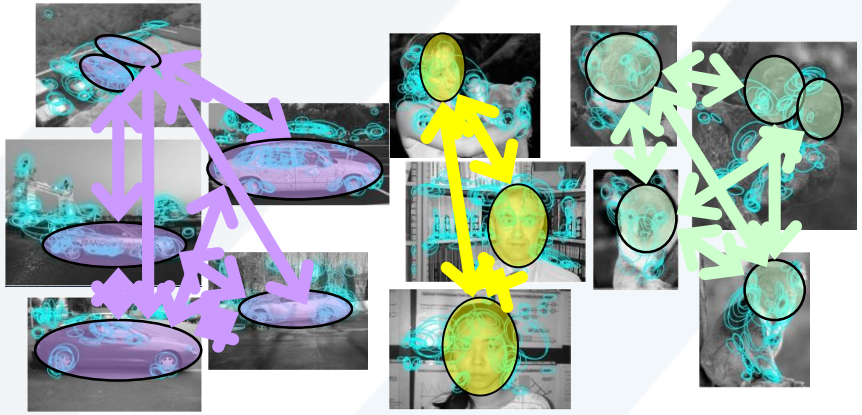
2024

What is Grouping in Computer Vision?

- Gather features that belong together
- تجميع السمات التي تنتمي إلى نفس المجموعة معاً ضمن مجموعة واحدة.
- Obtain an intermediate representation that compactly describes key image or video parts.
- الحصول على تمثيل وسيط يصف تماماً الأجزاء الرئيسية في الصورة أو الفيديو.

Determine image regions

Group video frames into shots



Object-level grouping

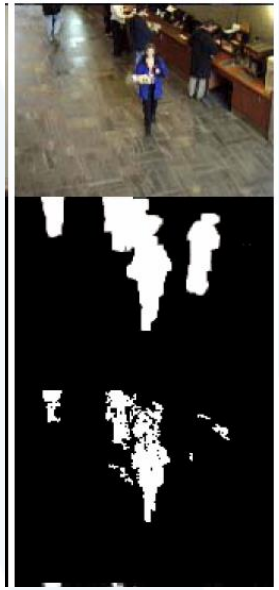
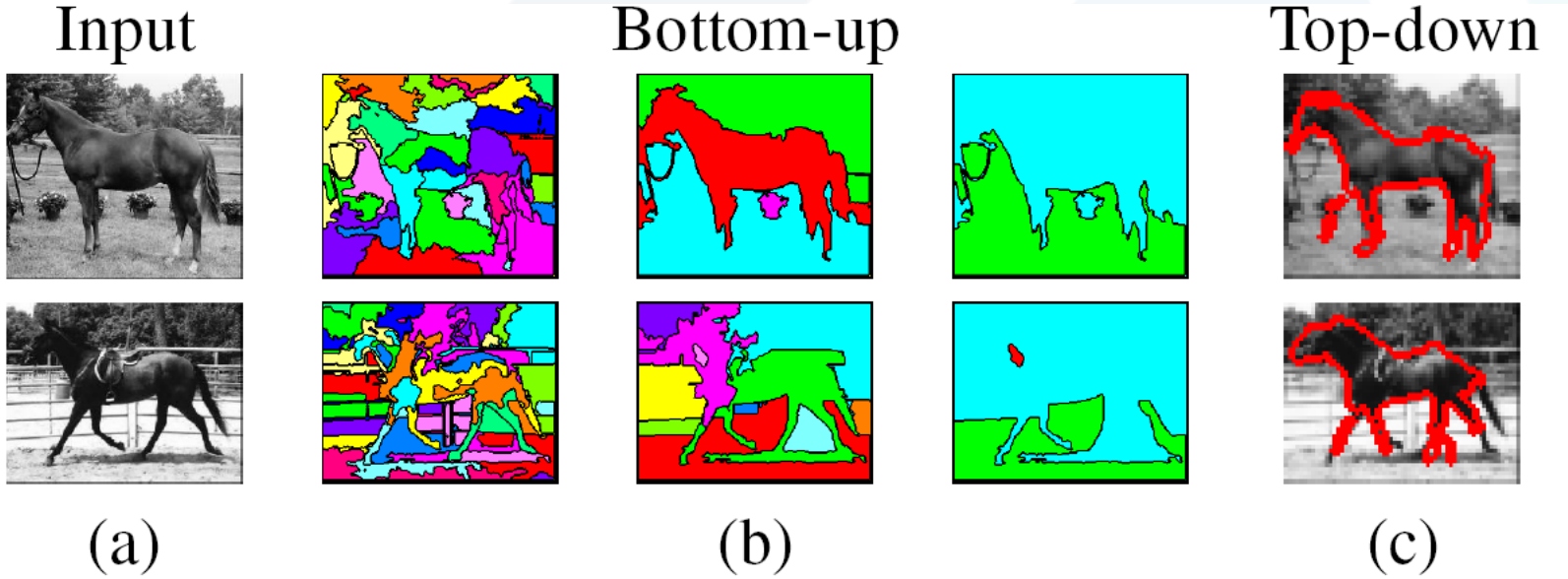


Figure-background

Grouping & Segmentation

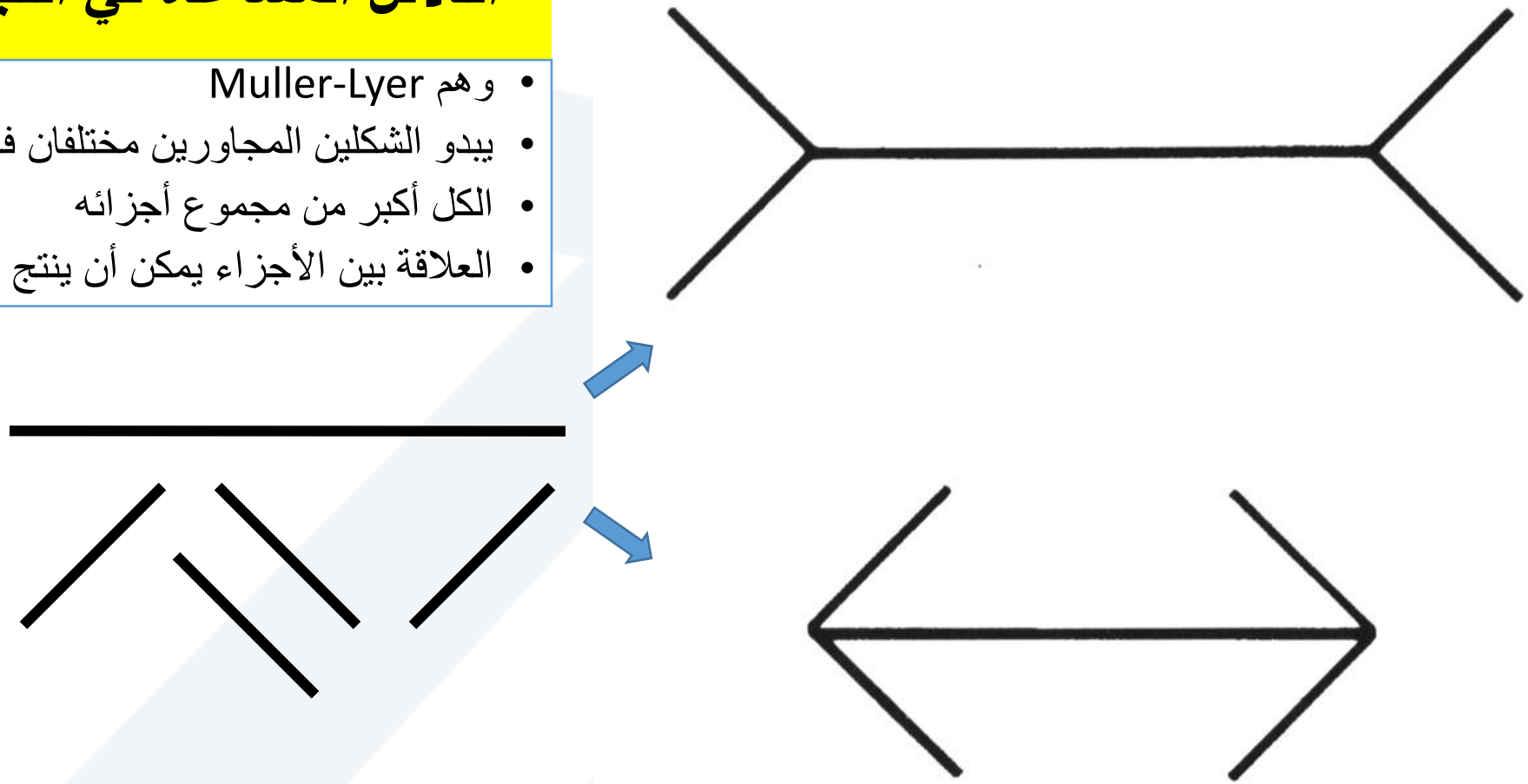
- **Top down vs. bottom up segmentation التجزيء أعلى-أسفل مقابل التجزيء أسفل-أعلى**
  - Top down: pixels belong together because they are from the same object
    - من الأعلى للأسفل: تنتمي البكسلات لنفس المجموعة لأنها من نفس المكون.
  - Bottom up: pixels belong together because they look similar.
    - من الأسفل للأعلى: تنتمي البكسلات لنفس المجموعة لأنها تبدو متشابهة.
- **Hard to measure success كيفية قياس نجاح التجزيء**
  - What is interesting depends on the app. يعتمد على التطبيق



Grouping & Segmentation

- **Meta Clues الدلائل المساعدة في التجزيء**

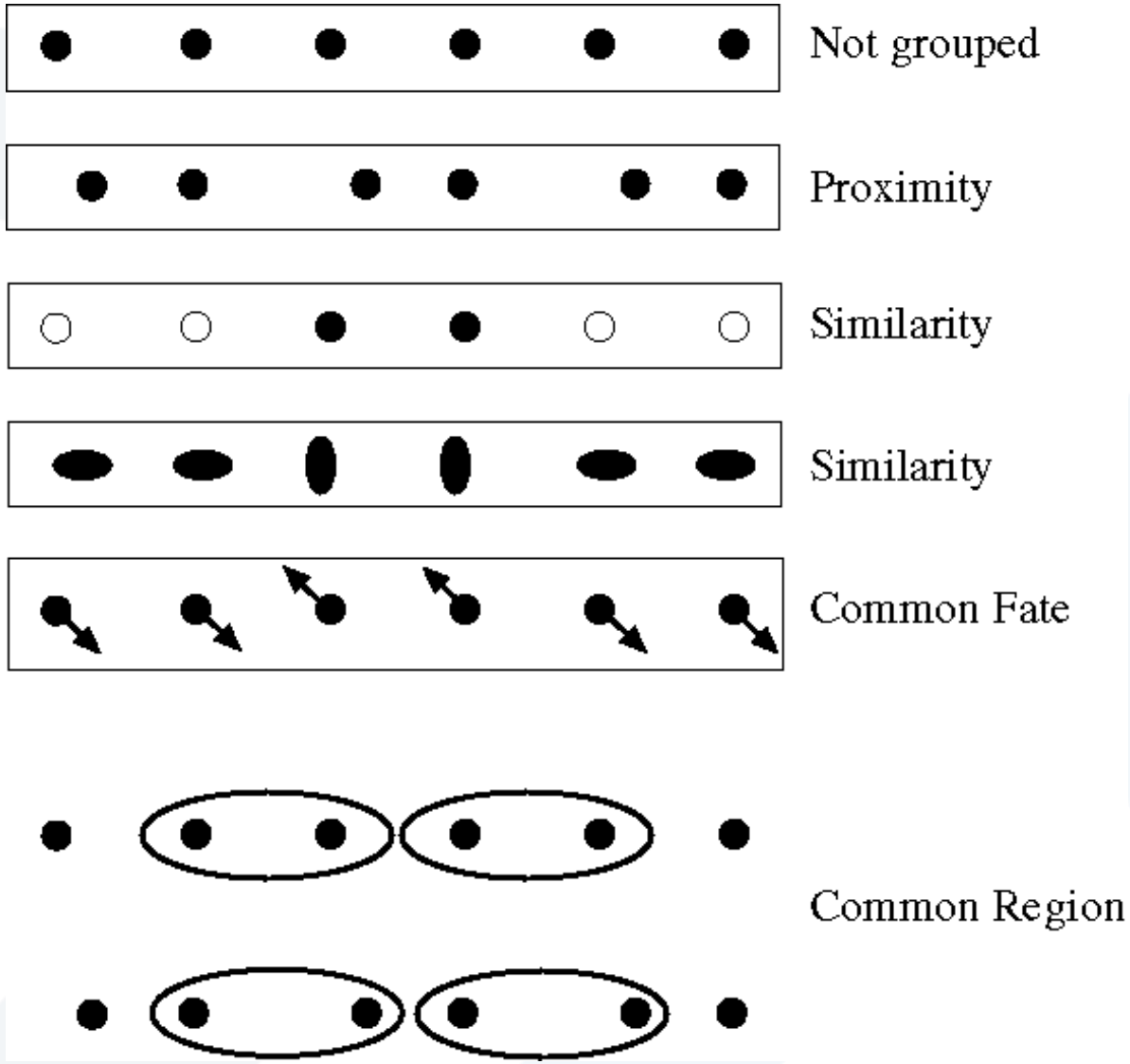
- وهم Muller-Lyer
- يبدو الشكلين المجاورين مختلفان في الطول لكن طولهما واحد.
- الكل أكبر من مجموع أجزائه
- العلاقة بين الأجزاء يمكن أن ينتج عنها سمات جديدة.



The Muller-Lyer illusion

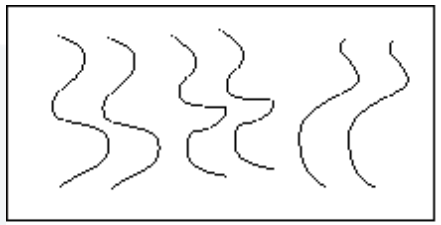
Principles of perceptual organization
مبادئ التنظيم الإدراكي

Grouping & Segmentation

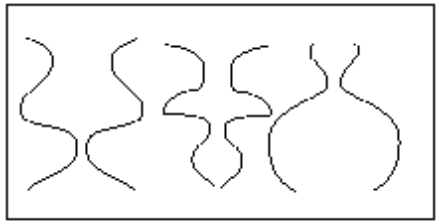


From Steve Lehar: The Constructive Aspect of Visual Perception

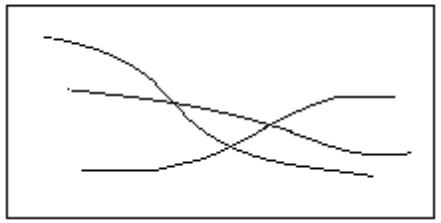
Principles of perceptual organization
مبادئ التنظيم الإدراكي



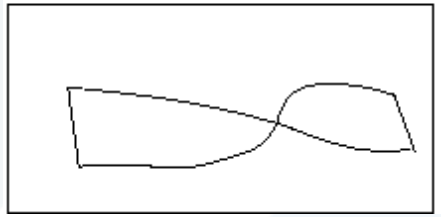
Parallelism



Symmetry



Continuity

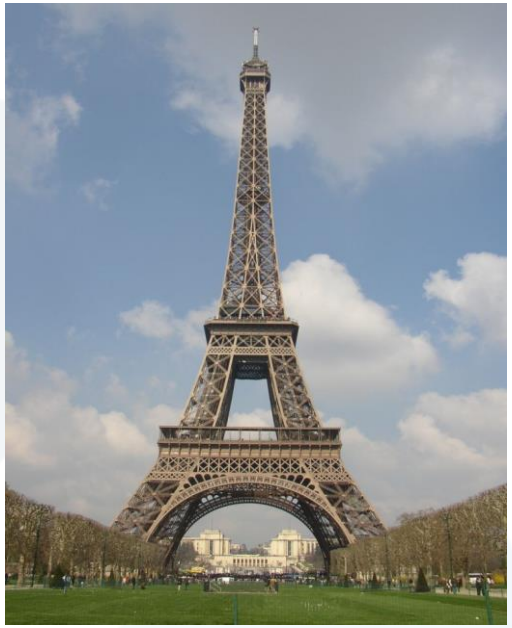


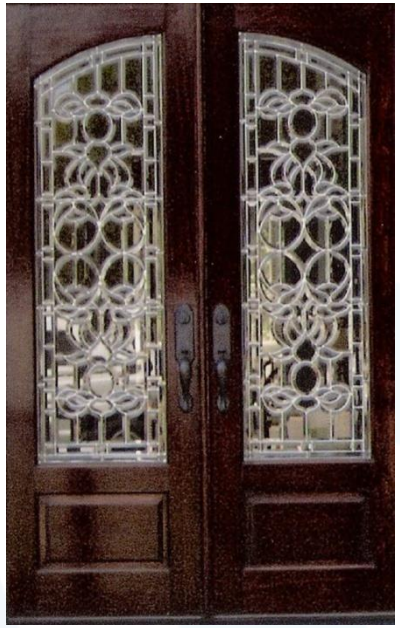
Closure

Grouping & Segmentation

- **Similarity التشابه**

- **Symmetry التناظر**

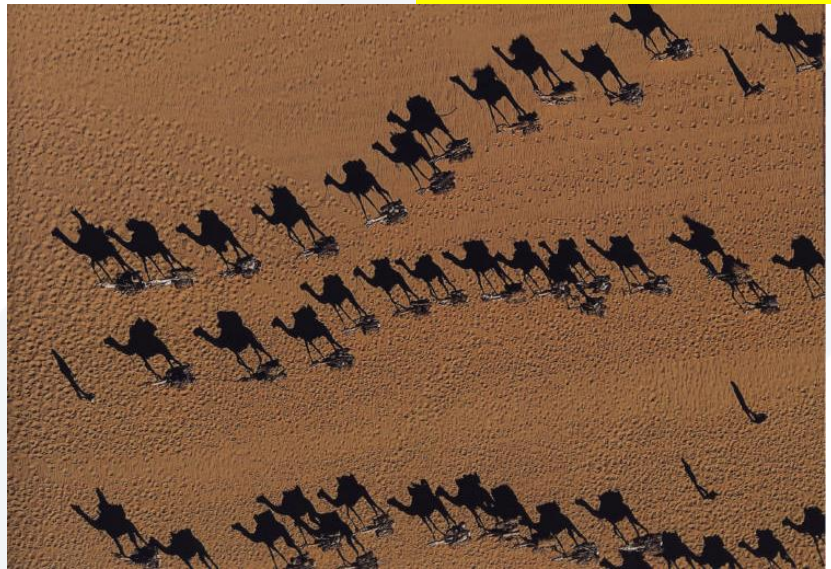




Grouping & Segmentation

- **Proximity التقارب**

- **Common Fate المصير المشترك**



Bottom-up Segmentation التجزيء من الأسفل للأعلى

- Bottom-up segmentation via clustering
  - Algorithms: كخوارزميات
    - Mode finding and mean shift: k-means, EM, mean-shift
    - Graph-based: normalized cuts
  - Features: color, texture, … كسمات
    - Quantization for texture summaries

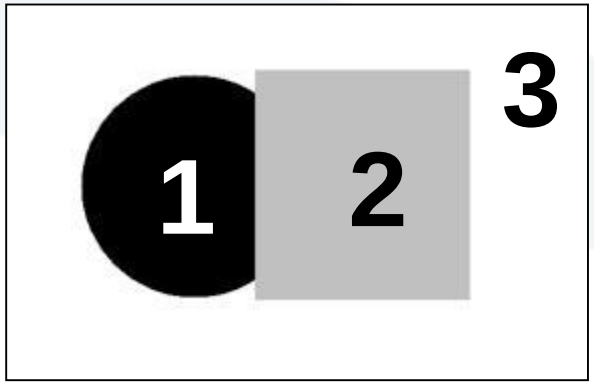
Aim of Segmentation:

Separate image into coherent “objects”

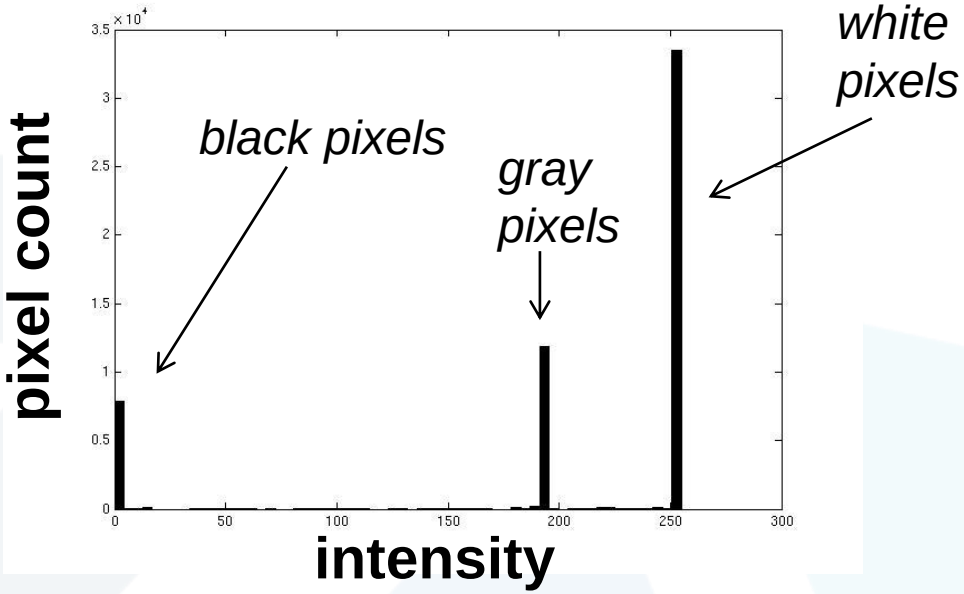
تهدف عملية التجزيء إلى تقسيم الصورة إلى أجزاء مترابطة

image human segmentation

Bottom-up Segmentation التجزيء من الأسفل للأعلى

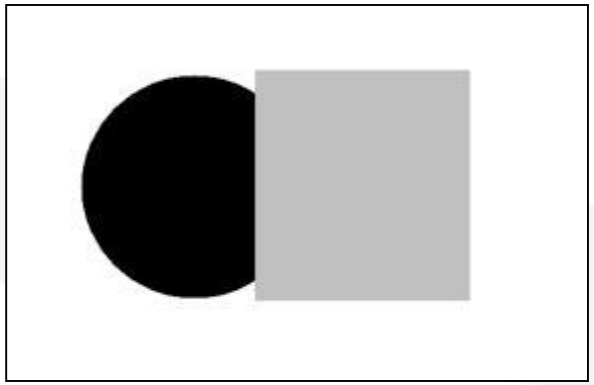


input image

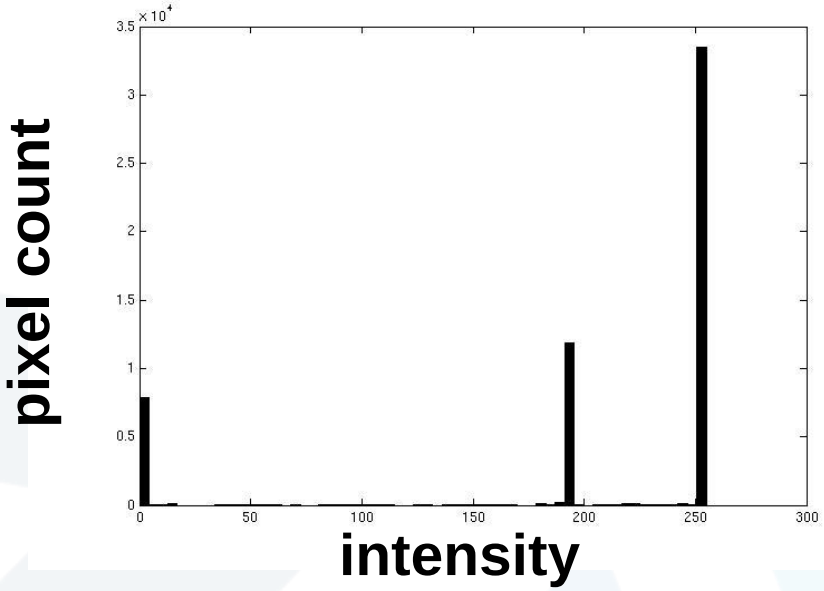


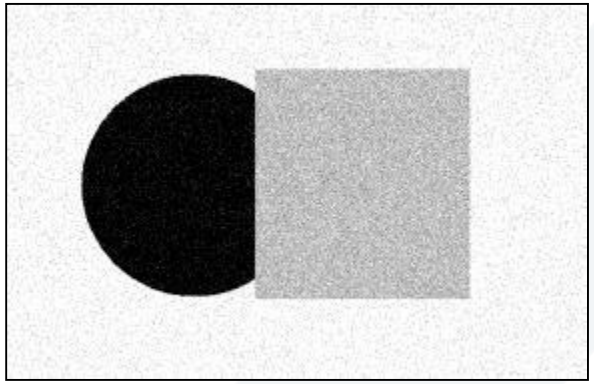
- These intensities define the three groups.
- We could label every pixel in the image according to which of these primary intensities it is.
  - i.e., *segment* the image based on the intensity feature.
    - تجزيء الصورة بالاعتماد على سمة السوية اللونية
- What if the image isn't quite so simple?

Slide credit: Kristen Grauman

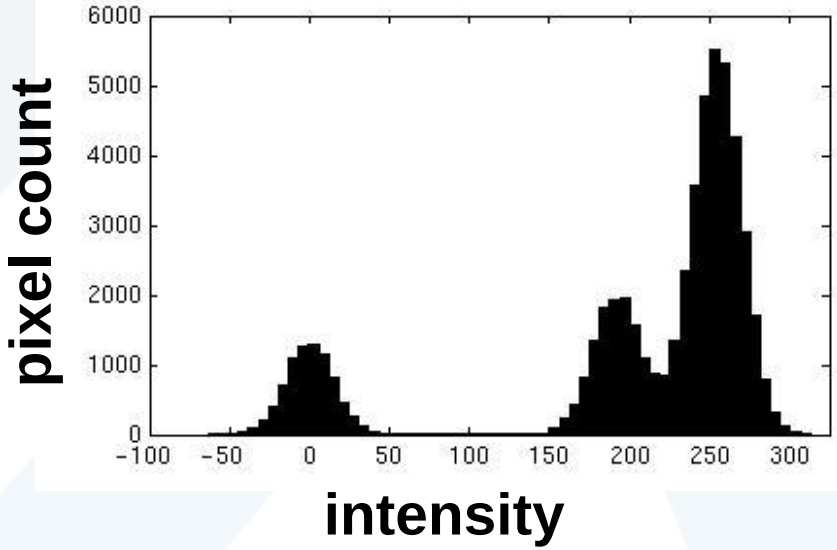


input image

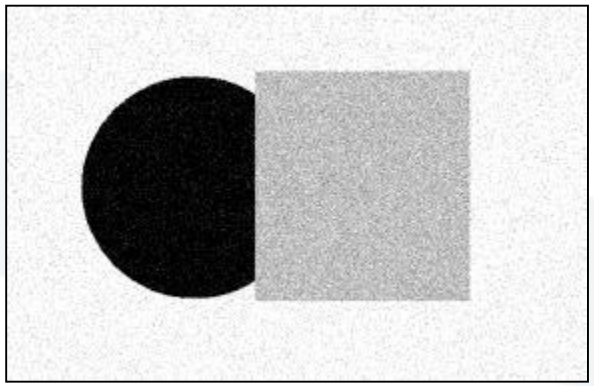




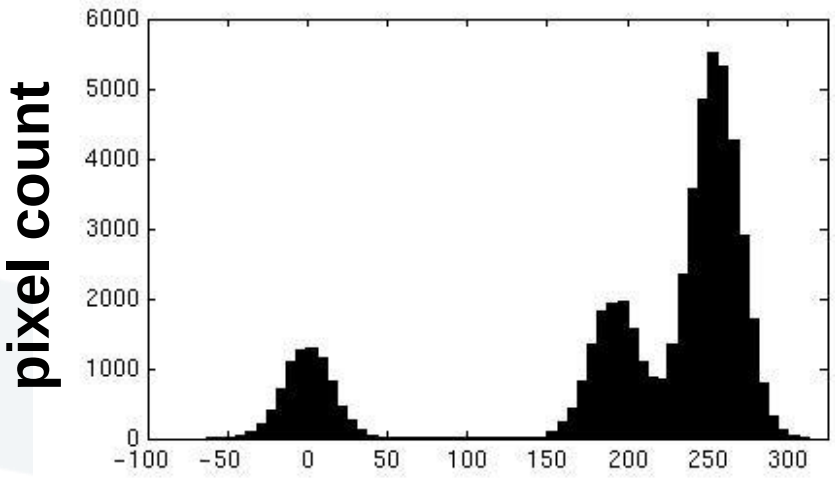
input image



Slide credit: Kristen Grauman



input image

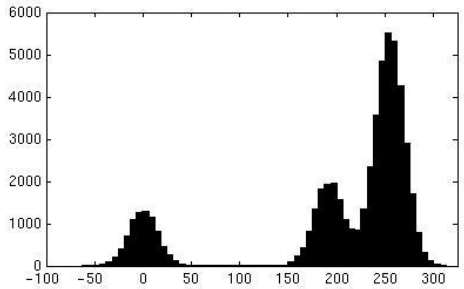


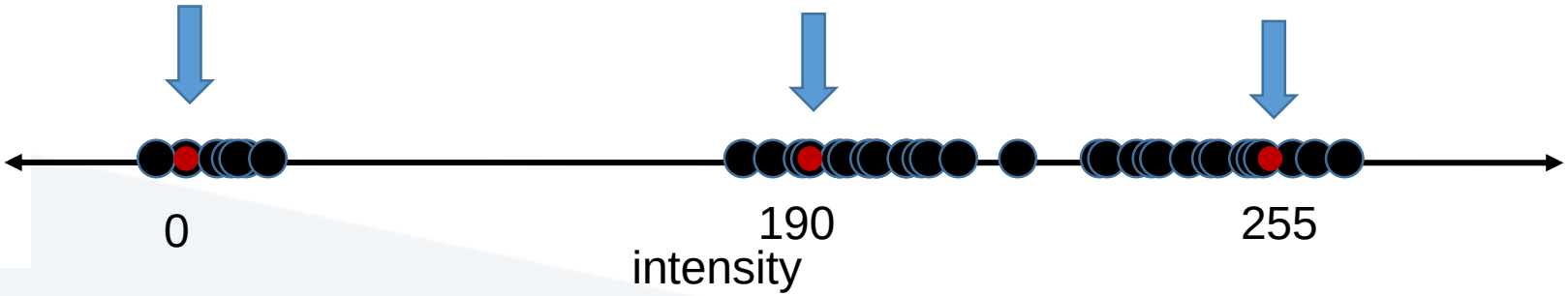
intensity

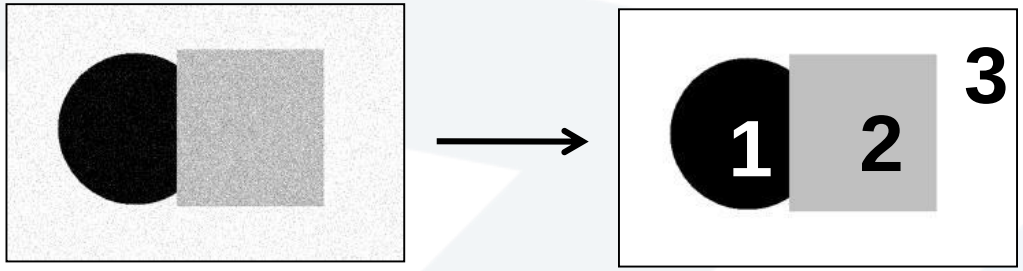
- Now how to determine the three main intensities that define our groups?
- We need to ***cluster.***

- كيف نحدد السويات اللونية الثلاثة الأساسية التي تحدد المجموعات الثلاثة ؟ نحتاج إلى عملية **العنقدة**

Slide credit: Kristen Grauman







- Goal: choose three “centers” as the representative intensities, and label every pixel according to which of these centers it is nearest to.
- الهدف اختيار **مراكز** السويات الممثلة لكل منطقة ثم **عنونة** كل بكسل برقم المركز **الأقرب إليه**.
- Best cluster centers are those that minimize SSD between all points and their nearest cluster center c_i:
- أفضل المراكز التي يتم اختيارها تلك التي تقلل خطأ SSD بين كل نقاط المجموعة والمركز.

$$\sum_{\text{clusters } i} \sum_{\text{points p in cluster } i} ||p - c_i||^2$$

Slide credit: Kristen Grauman

Bottom-up Segmentation التجزيء من الأسفل للأعلى

Clustering العنقدة

- Clustering algorithms:
  - Unsupervised learning تعلم بدون إشراف
  - Detect patterns in unlabeled data تقوم بكشف النماذج في البيانات غير المعنونة
    - E.g. group emails or search results مثل تجميع رسائل البريد الالكتروني أو تجميع نتائج البحث
    - E.g. find categories of customers مثال آخر: تصنيف الزبائن إلى مجموعات
    - **E.g. group pixels into regions التطبيق الأهم: تجميع عناصر الصورة إلى مناطق**

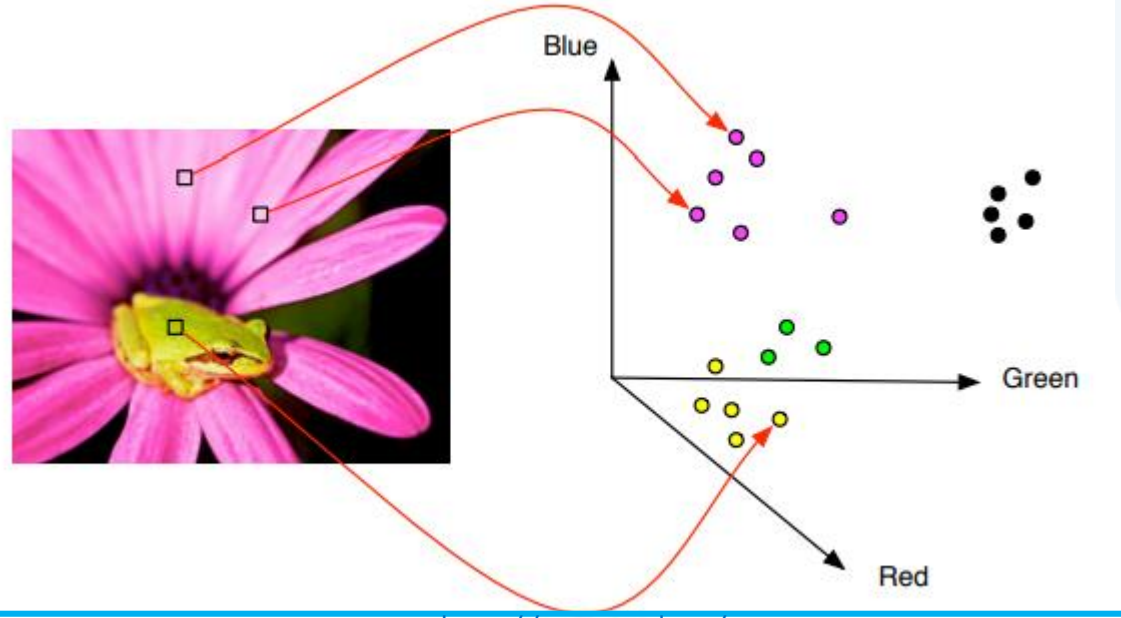
  - Useful when don't know what you're looking for Requires data, but no labels
- دائماً تنفع طريقة التجزيء هذه (العنقدة) في حال نبحث عن بيانات ما ضمن الصورة لكن ليس لدينا أصناف محددة لهذه البيانات.

Clustering العنقدة

Bottom-up Segmentation

K-means

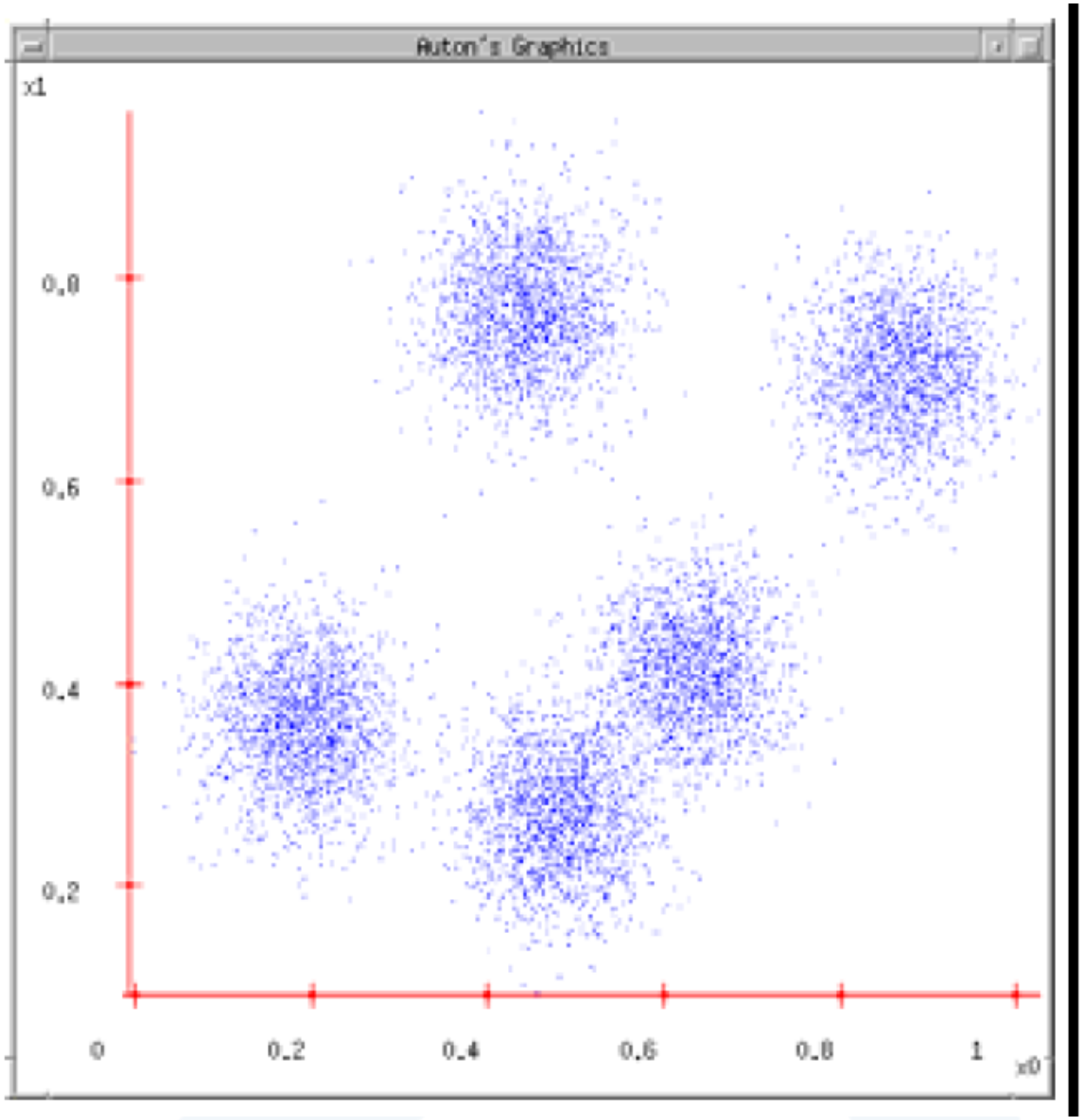
- Randomly initialize the *k* cluster centers, and iterate between the two steps we just saw.
  1. Randomly initialize the cluster centers, $c_1, \ldots, c_K$. اختيار المراكز الابتدائية عشوائياً
  2. Given cluster centers, determine points in each cluster من أجل كل مركز، حدد نقاط العنقود الأقرب لكل مركز
     - For each point p, find the closest c_i. Put p into cluster i
  3. Given points in each cluster, solve for c_i تحديث مراكز العناقيد لتكون متوسط قيم عناصر العنقود
     - Set c_i to be the mean of points in cluster i
  4. If c_i have changed, repeat Step 2 في حال تغير المراكز أعد تكرار الخطوات وإلا توقف



Source: Steve Seitz

K-means

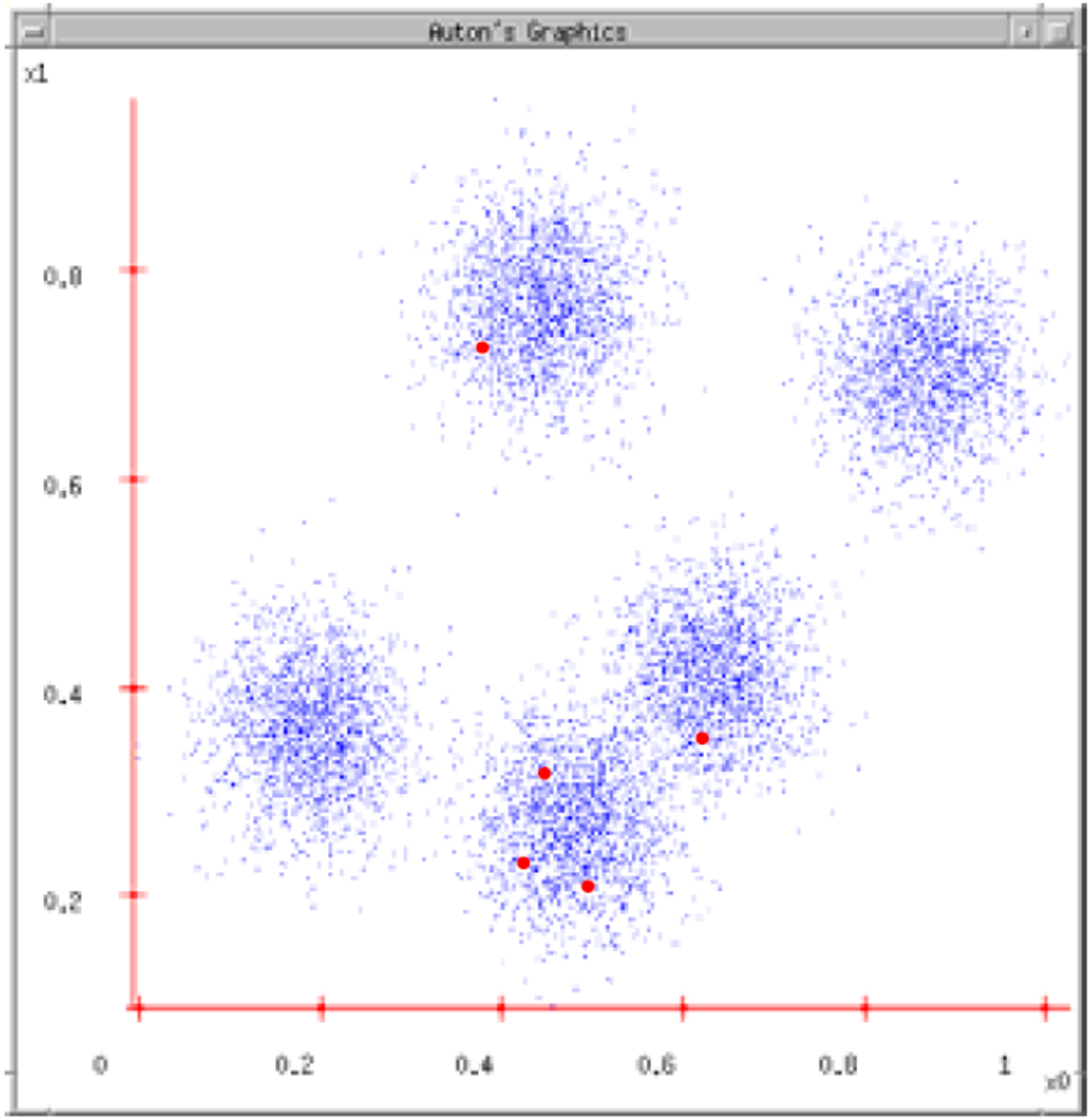
1. Ask user how many clusters they'd like. *(e.g. k=5)*



Slide credit Andrew Moore

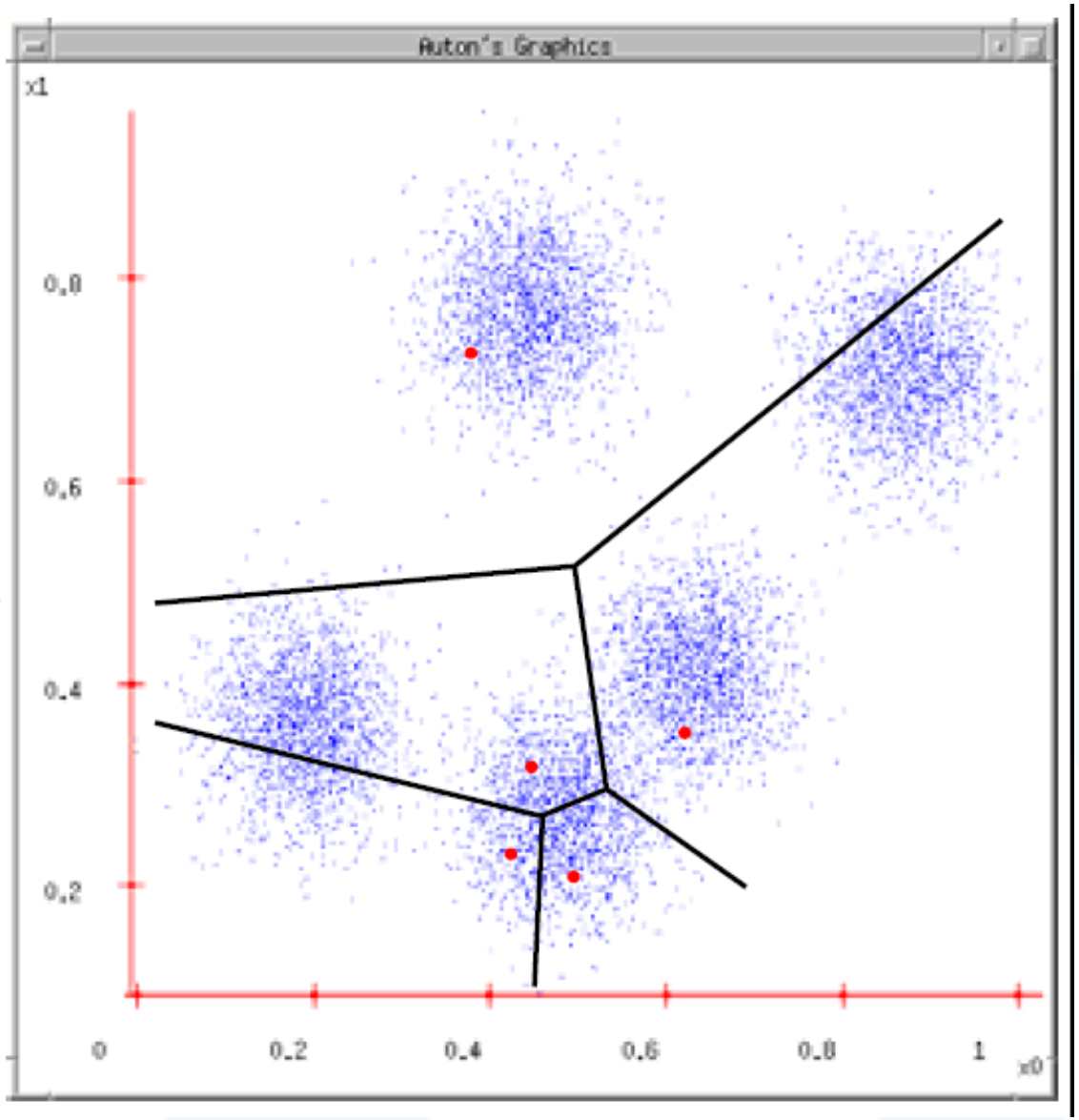
K-means

1. Ask user how many clusters they'd like. *(e.g. k=5)*
2. Randomly guess k cluster Center locations



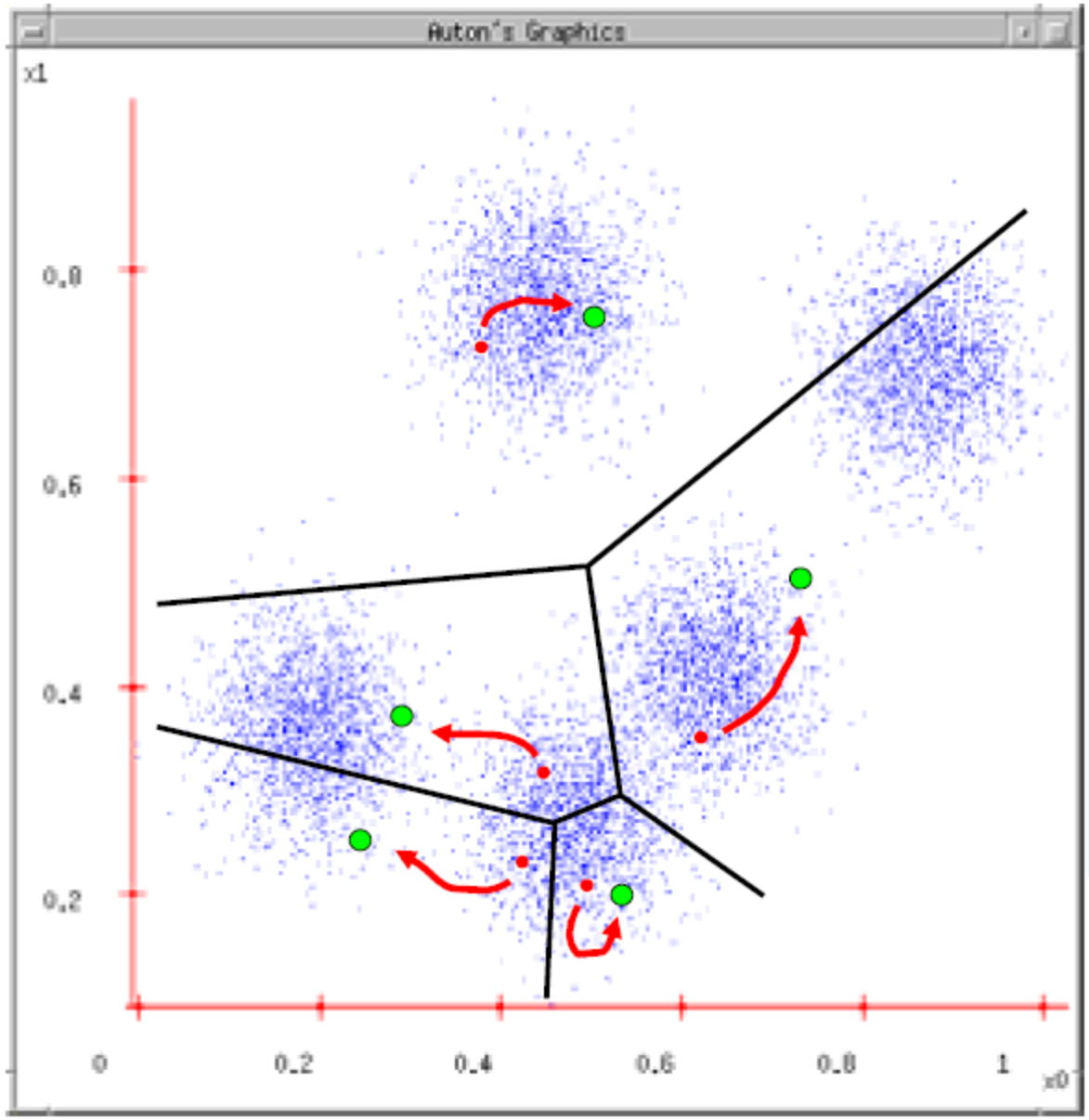
K-means

1. Ask user how many clusters they'd like. *(e.g. k=5)*
2. Randomly guess k cluster Center locations
3. Each datapoint finds out which Center it's closest to. (Thus each Center "owns" a set of datapoints)



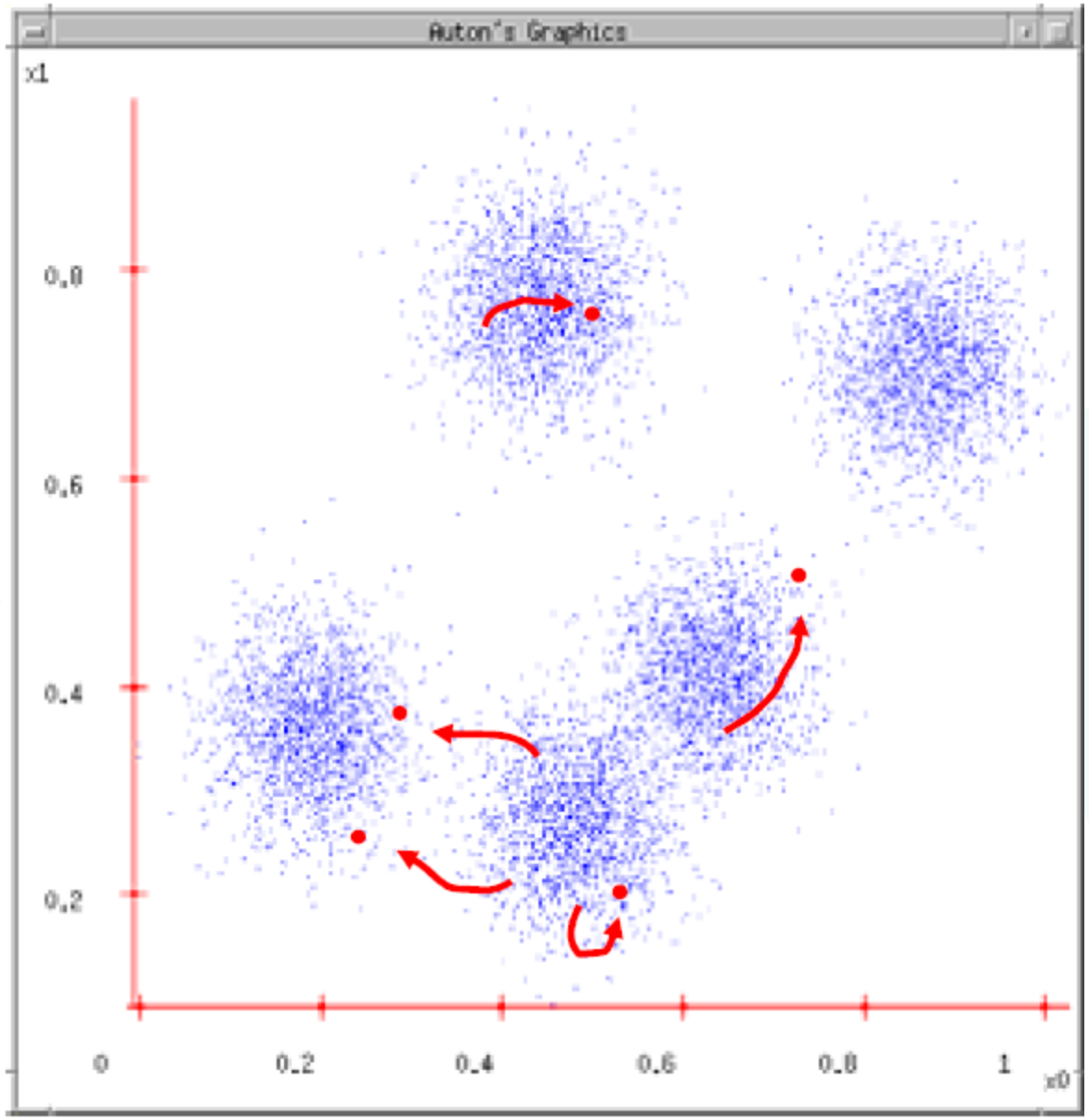
K-means

1. Ask user how many clusters they'd like. *(e.g. k=5)*
2. Randomly guess k cluster Center locations
3. Each datapoint finds out which Center it's closest to.
4. Each Center finds the centroid of the points it owns



K-means

1. Ask user how many clusters they'd like. *(e.g. k=5)*
2. Randomly guess k cluster Center locations
3. Each datapoint finds out which Center it's closest to.
4. Each Center finds the centroid of the points it owns…
5. …and jumps there
6. …Repeat until terminated!



Bottom-up Segmentation

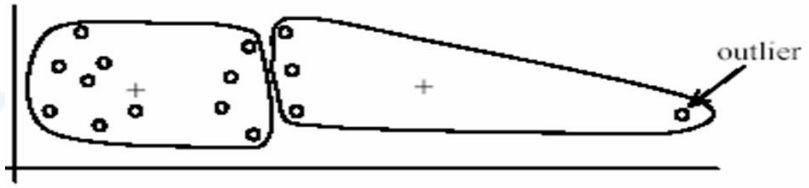
K-means: pros and cons

Pros المحاسن

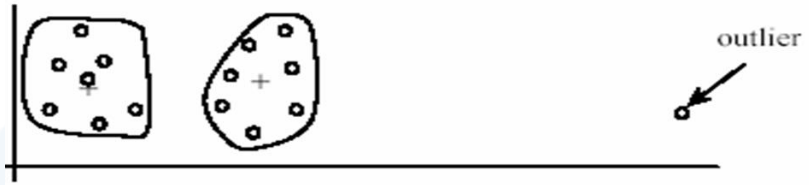
- Simple, fast to compute سهلة وسريعة

Cons/issues المساوئ

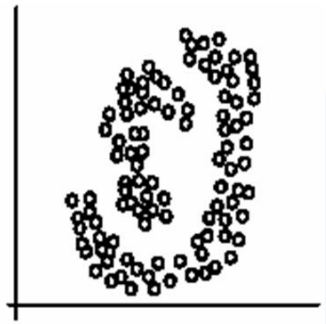
- Setting k? معرفة قيمة عدد العناقيد الصحيح
- Sensitive to initial centers حساسة لاختيار المراكز الابتدائية
- Sensitive to outliers حساسة للضجيج
- Detects spherical clusters كشف العناقيد الكروية صعب



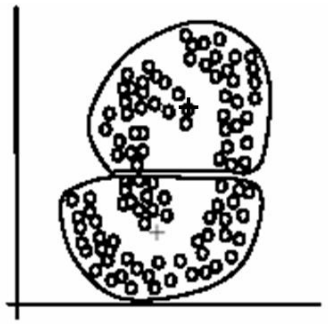
(A): Undesirable clusters



(B): Ideal clusters



(A): Two natural clusters



(B): *k*-means clusters

Clustering العنقدة

Bottom-up Segmentation

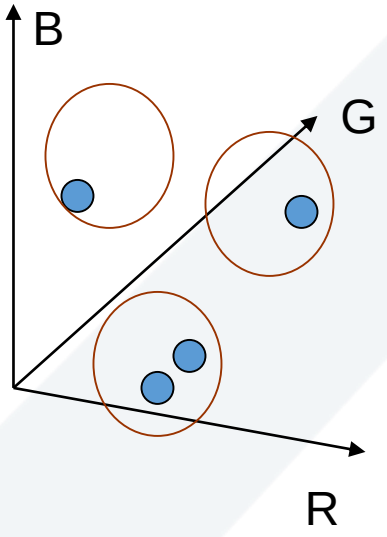
K-means

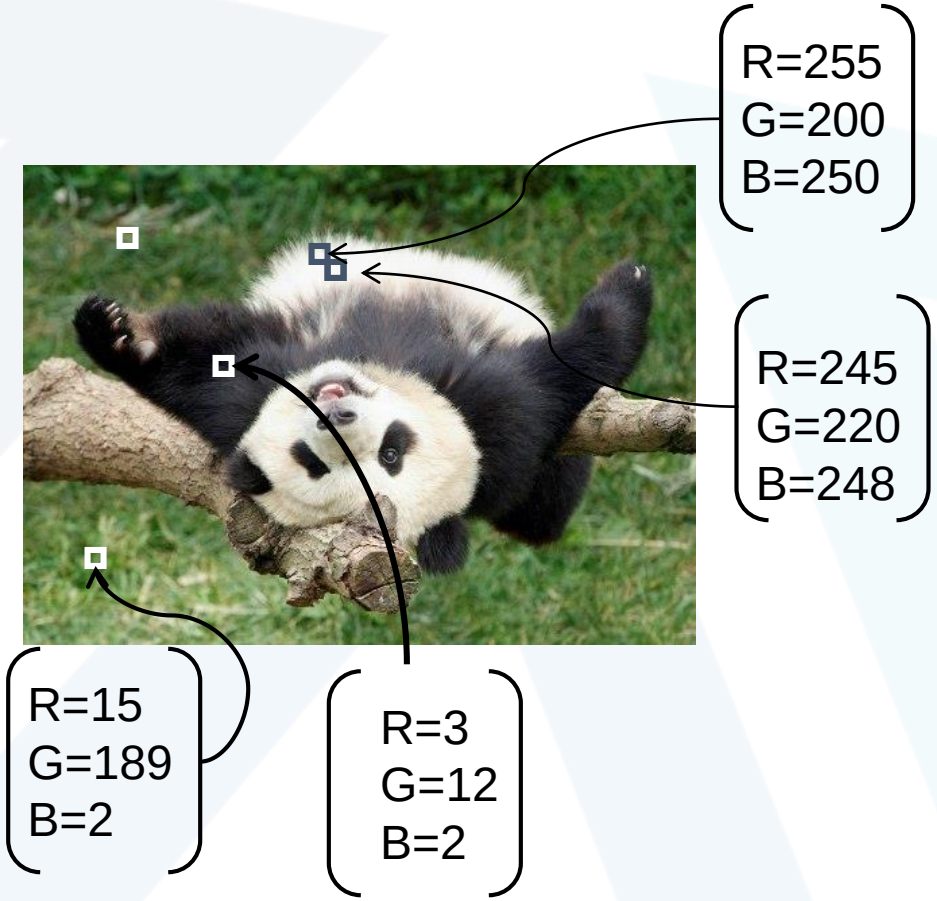
Depending on what we choose as the *feature space*, we can group pixels in different ways.

تعتمد على ما نختاره كفضاء سمات، يمكن تجميع البكسلات بعدة طرق (السوية اللونية، اللون، النسيج، الموقع،..)

Grouping pixels based on **color** similarity

تجميع البكسلات بالاعتماد على التشابه اللوني





Feature space: color value (3-d)

فضاء السمات: قيمة اللون (ثلاث قيم أحمر أخضر، أزرق)

Slide credit: Kristen Grauman

Clustering العنقدة

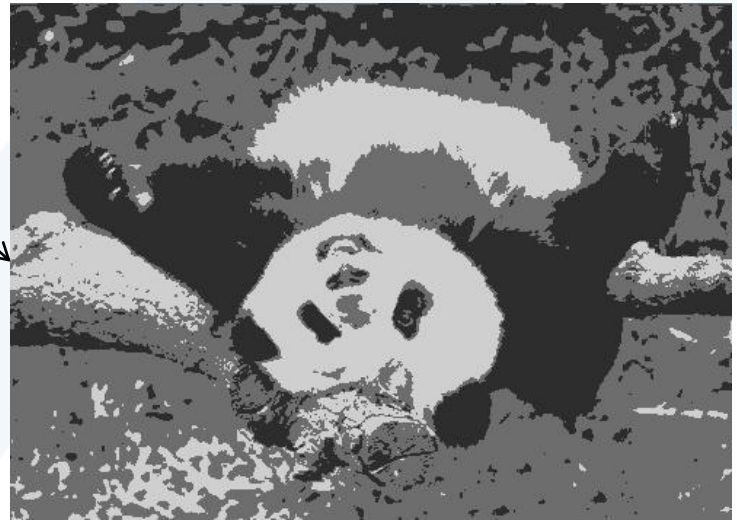
Bottom-up Segmentation

K-means

K=2

2 clusters
2 colors
(white, black)

K=3



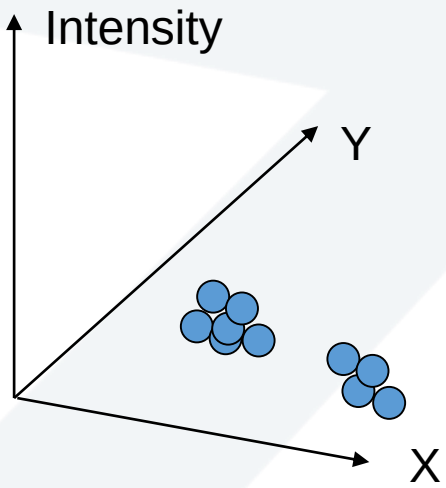
3 clusters
3 colors
(white, gray, black)

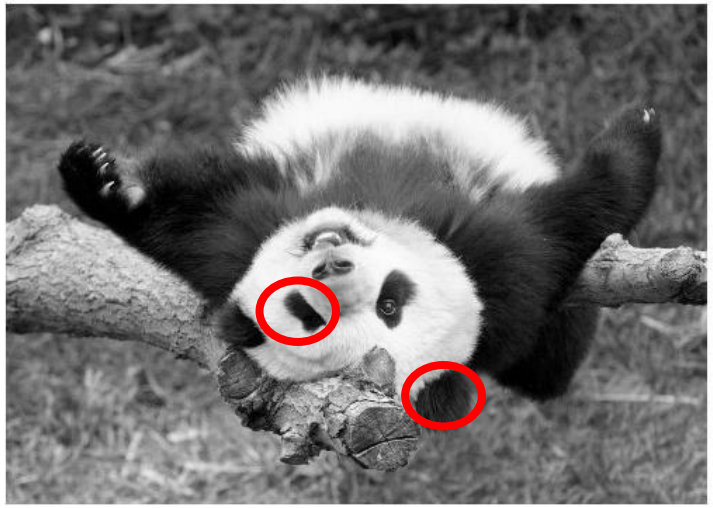
العنقدة هنا تتم وفق السويات الرمادية
intensity

Clustering العنقدة

Bottom-up Segmentation

K-means





العنقدة هنا تتم وفق السويات الرمادية + الموقع

intensity + position

Both regions are black, but if we also include **position (x,y),** then we could group the two into distinct segments; way to encode both similarity & proximity.

Clustering العنقدة

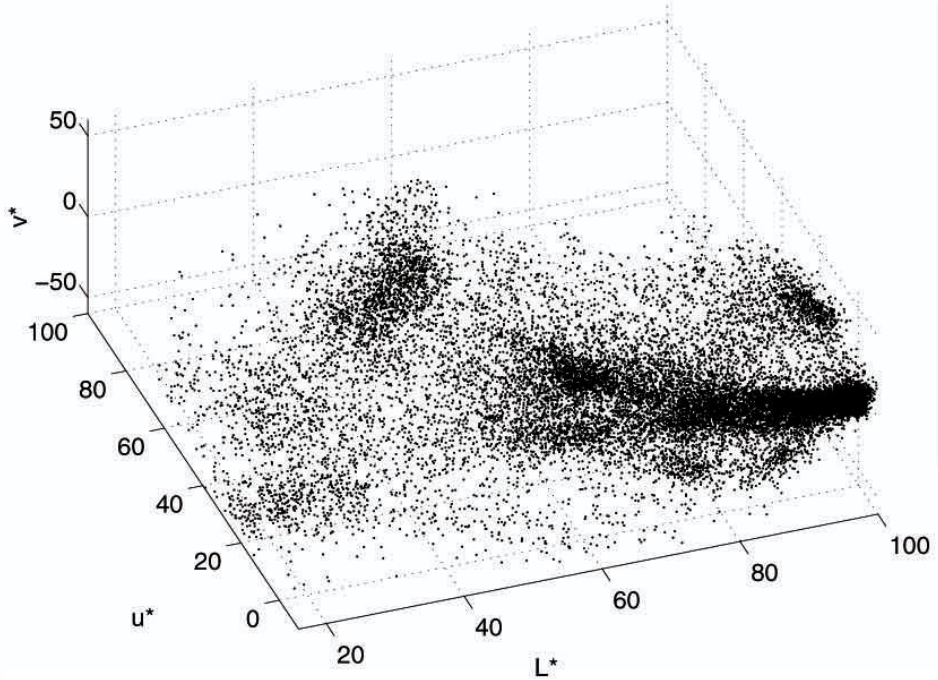
Bottom-up Segmentation

Mean shift algorithm

- The mean shift algorithm seeks *modes* or local maxima of density in the feature space
- تبحث خوارزمية إزاحة المتوسط على نقاط modes أو النقاط العظمى المحلية للكثافة في فضاء السمات.

image

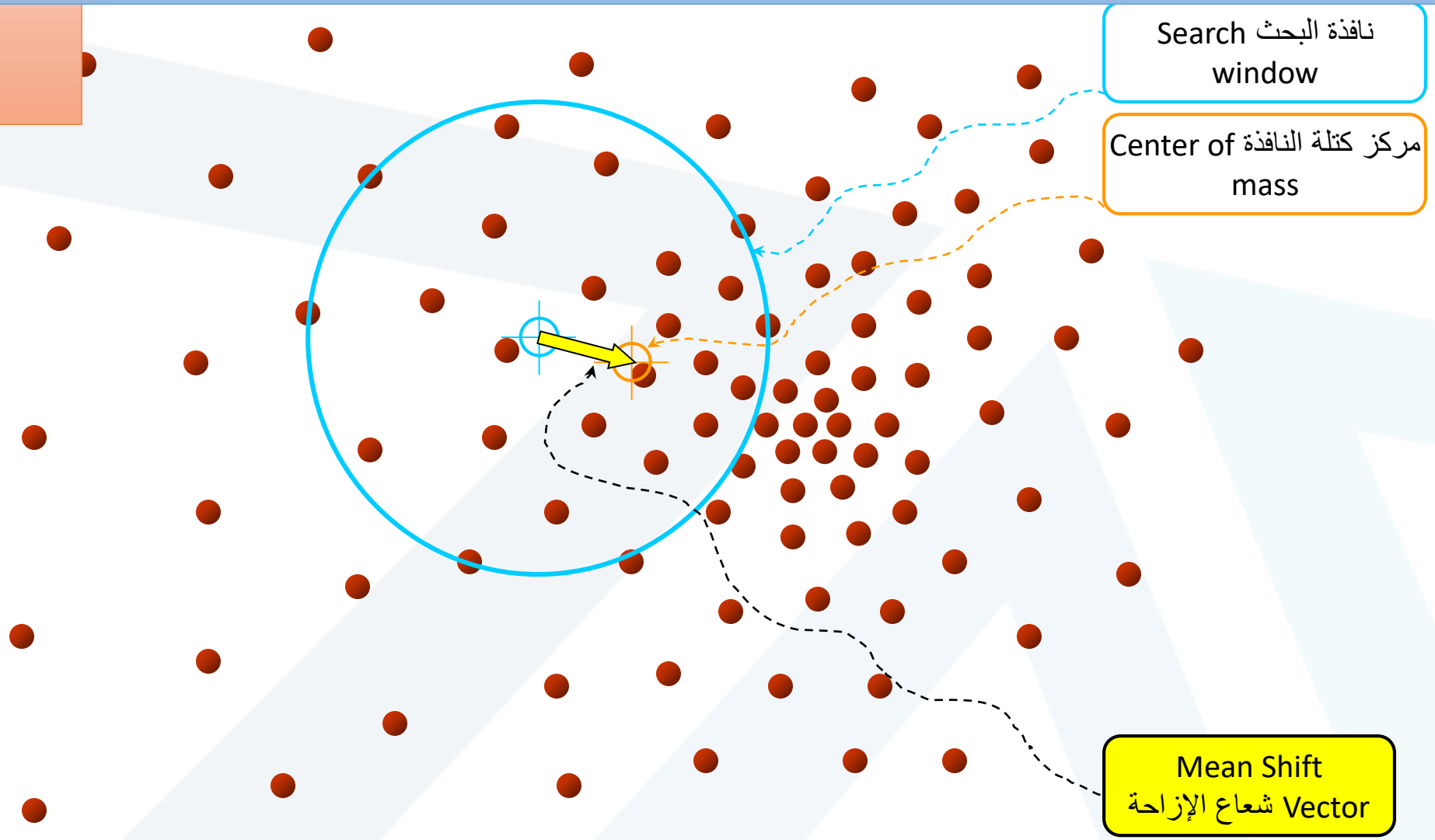
فضاء السمات Feature space (L*u*v* color values)



Clustering العنقدة

Bottom-up Segmentation

Mean shift algorithm

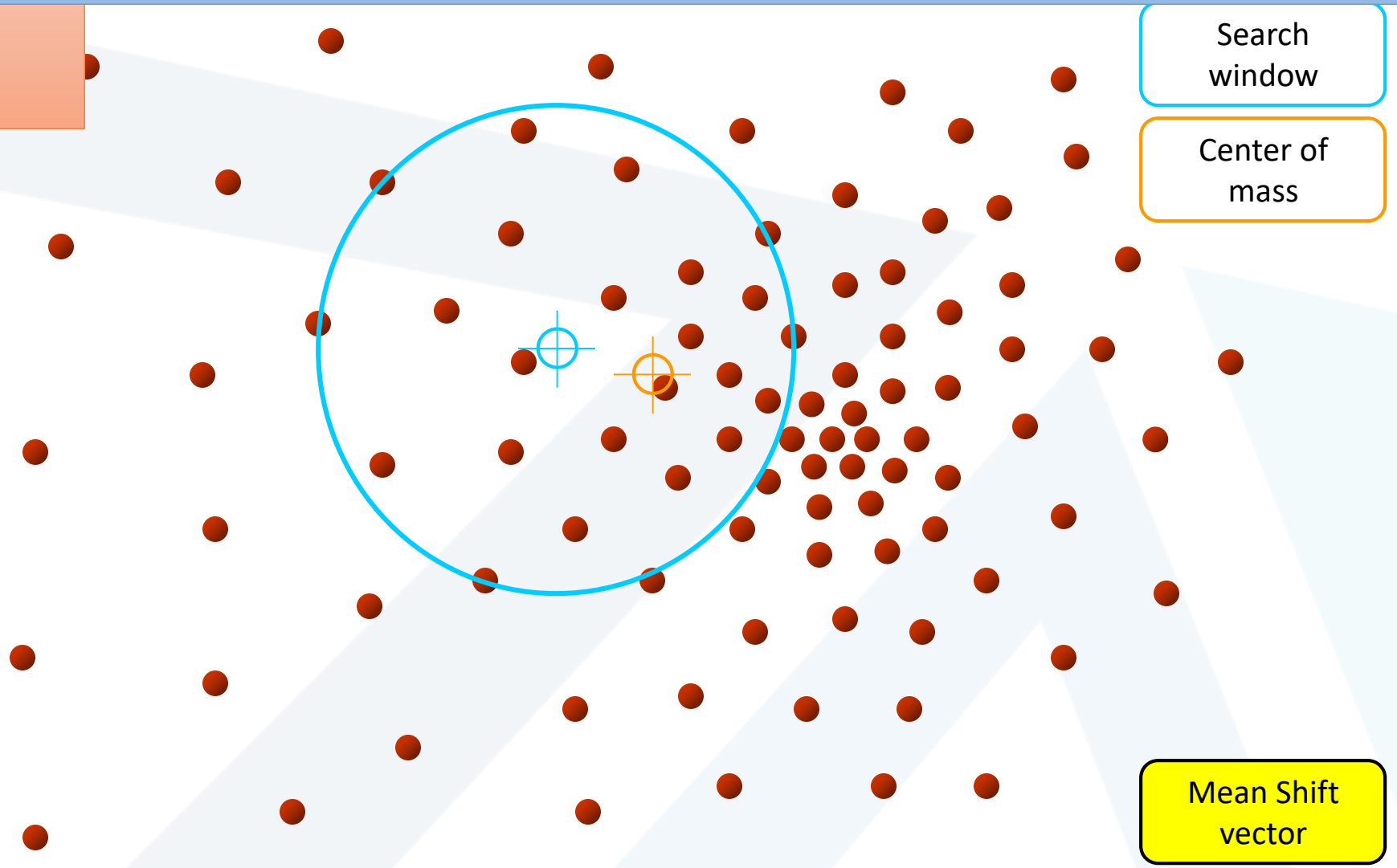


Slide by Y. Ukrainitz & B. Sarel

Clustering العنقدة

Bottom-up Segmentation

Mean shift algorithm

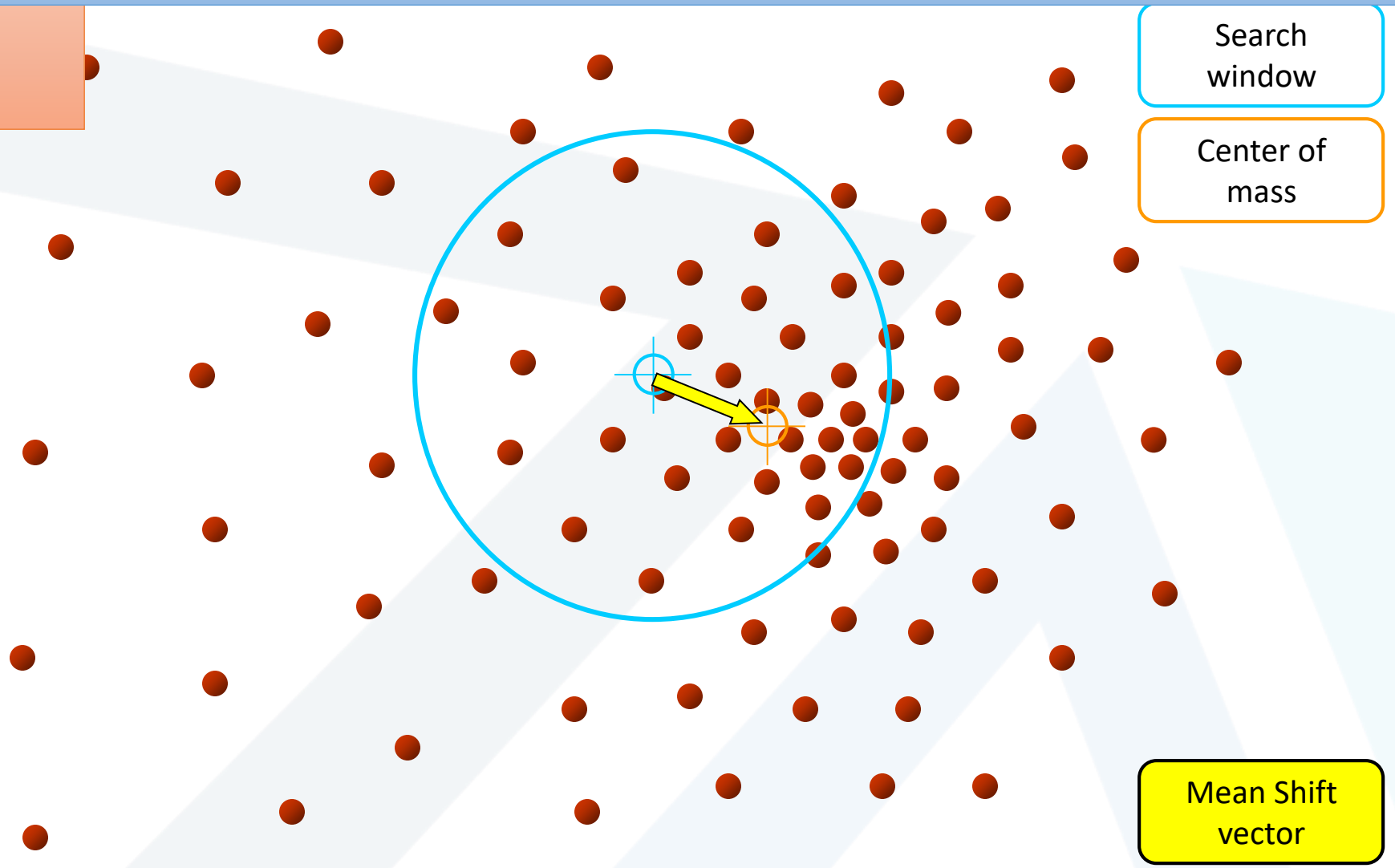


Slide by Y. Ukrainitz & B. Sarel

Clustering العنقدة

Bottom-up Segmentation

Mean shift algorithm

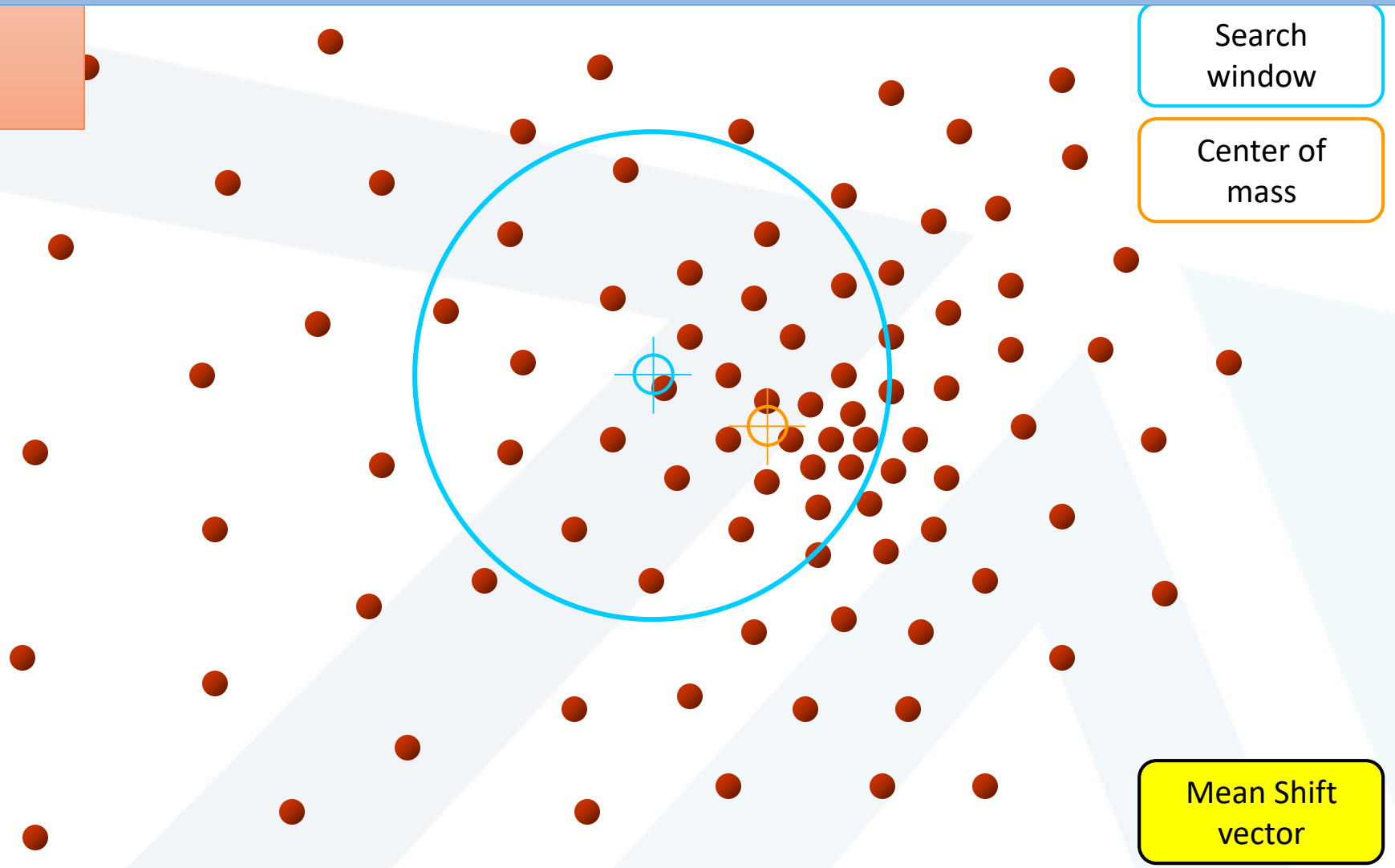


Slide by Y. Ukrainitz & B. Sarel

Clustering العنقدة

Bottom-up Segmentation

Mean shift algorithm

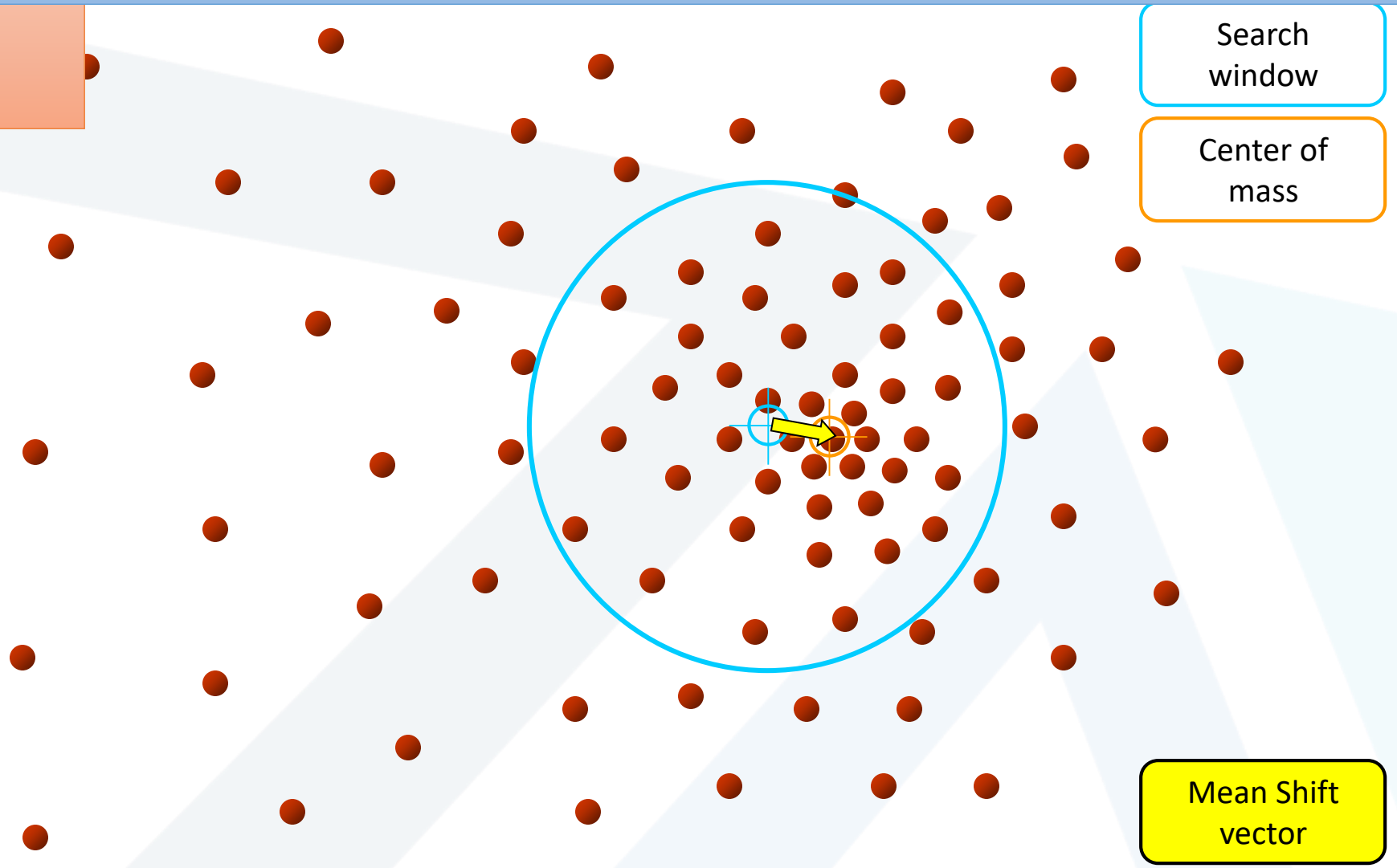


Slide by Y. Ukrainitz & B. Sarel

Clustering العنقدة

Bottom-up Segmentation

Mean shift algorithm

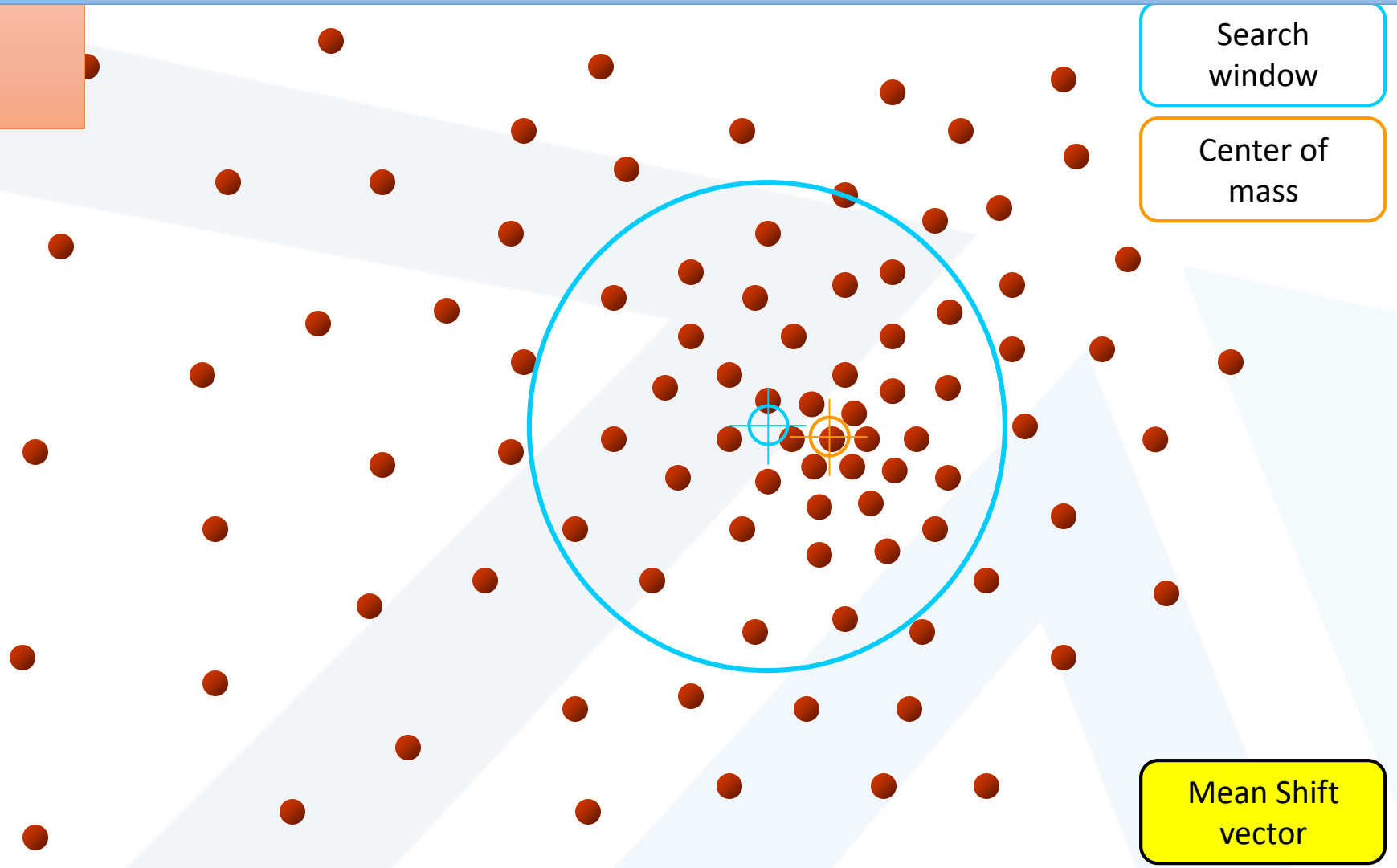


Slide by Y. Ukrainitz & B. Sarel

Clustering العنقدة

Bottom-up Segmentation

Mean shift algorithm

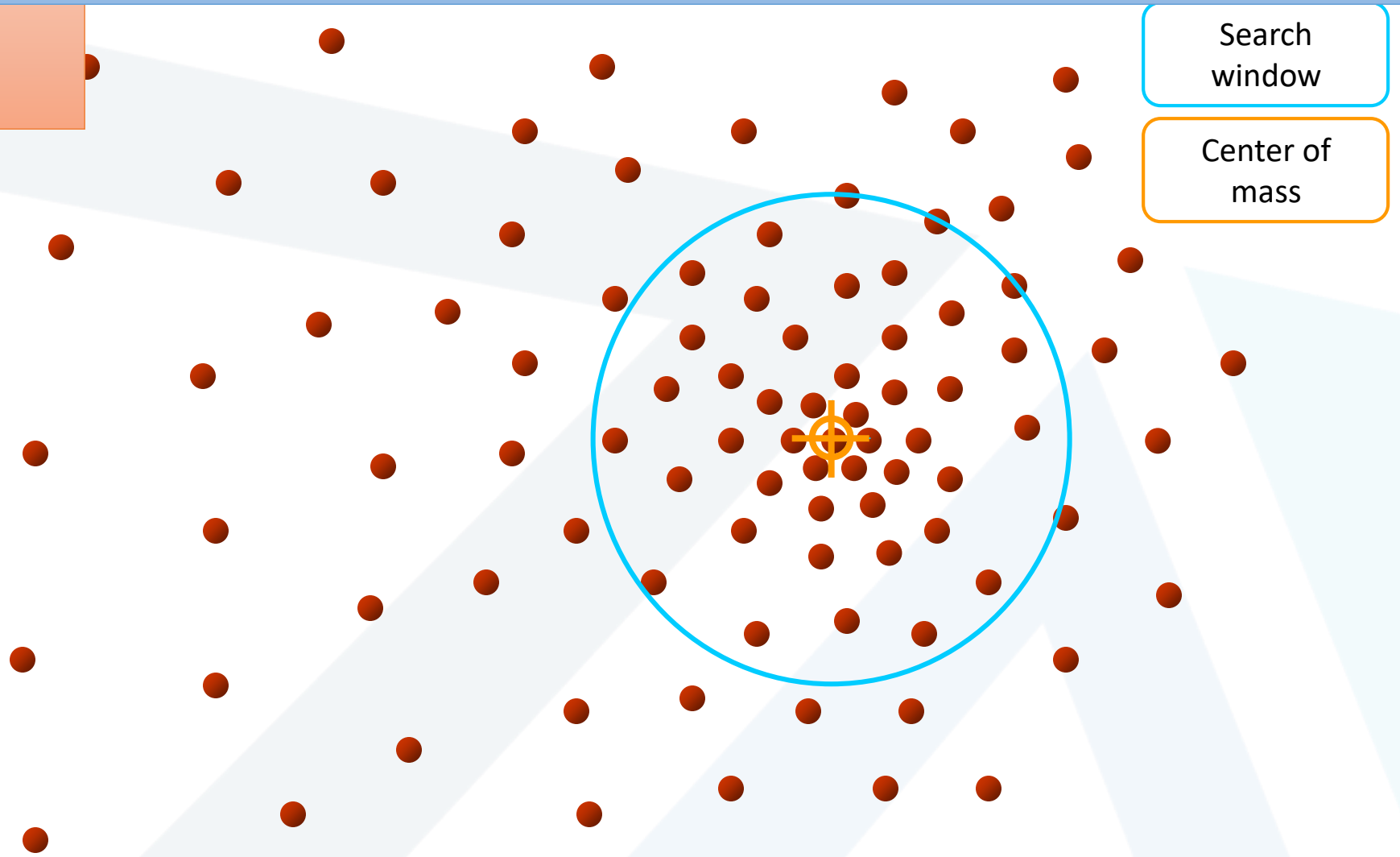


Slide by Y. Ukrainitz & B. Sarel

Clustering العنقدة

Bottom-up Segmentation

Mean shift algorithm



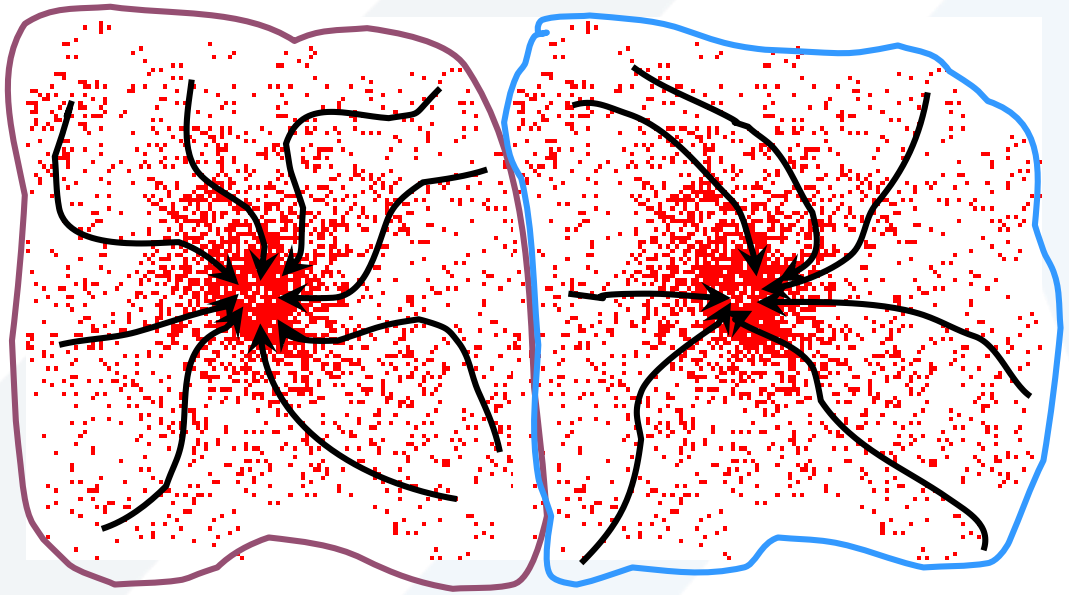
Slide by Y. Ukrainitz & B. Sarel

Clustering العنقدة

Bottom-up Segmentation

Mean shift algorithm

- Cluster: all data points in the attraction basin of a mode
- **العنقود:** كل نقاط البيانات تتوضع في المنطقة المحيطة بنقطة mode.
- Attraction basin: the region for which all trajectories lead to the same mode
- **منطقة الجذب (التأثير):** المنطقة التي تقود كل المسارات فيها إلى نقطة mode.



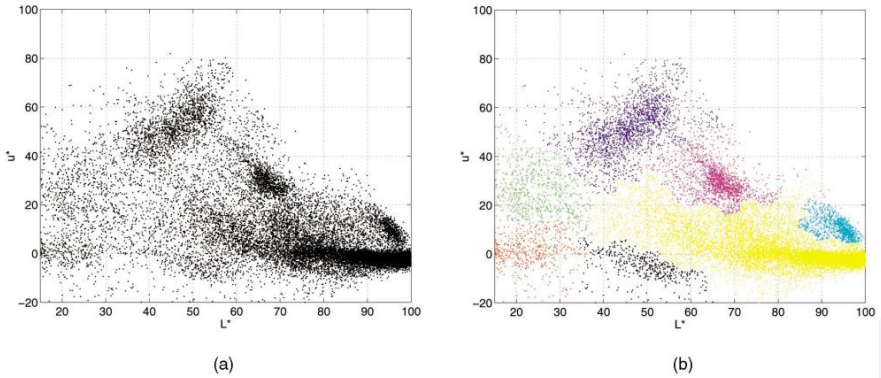
Slide by Y. Ukrainitz & B. Sarel

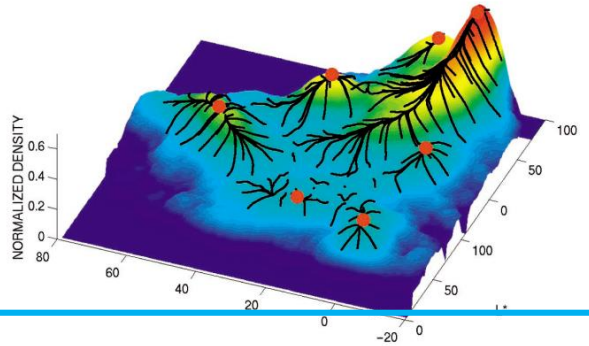
Clustering العنقدة

Bottom-up Segmentation

Mean shift algorithm

- Find features (color, gradients, texture, etc)
- إيجاد السمات (الألوان، التدرجات، النسيج، الخ).
- Initialize windows at individual feature points
- تهيئة النافذة بقيم نقاط سمات ابتدائية.
- Perform mean shift for each window until convergence
- تطبيق mean shift لكل نافذة حتى يتقارب المنحني.
- Merge windows that end up near the same “peak” or mode
- دمج النوافذ التي تتقارب عند نفس نقطة Mode (التي لها قمم متقاربة).





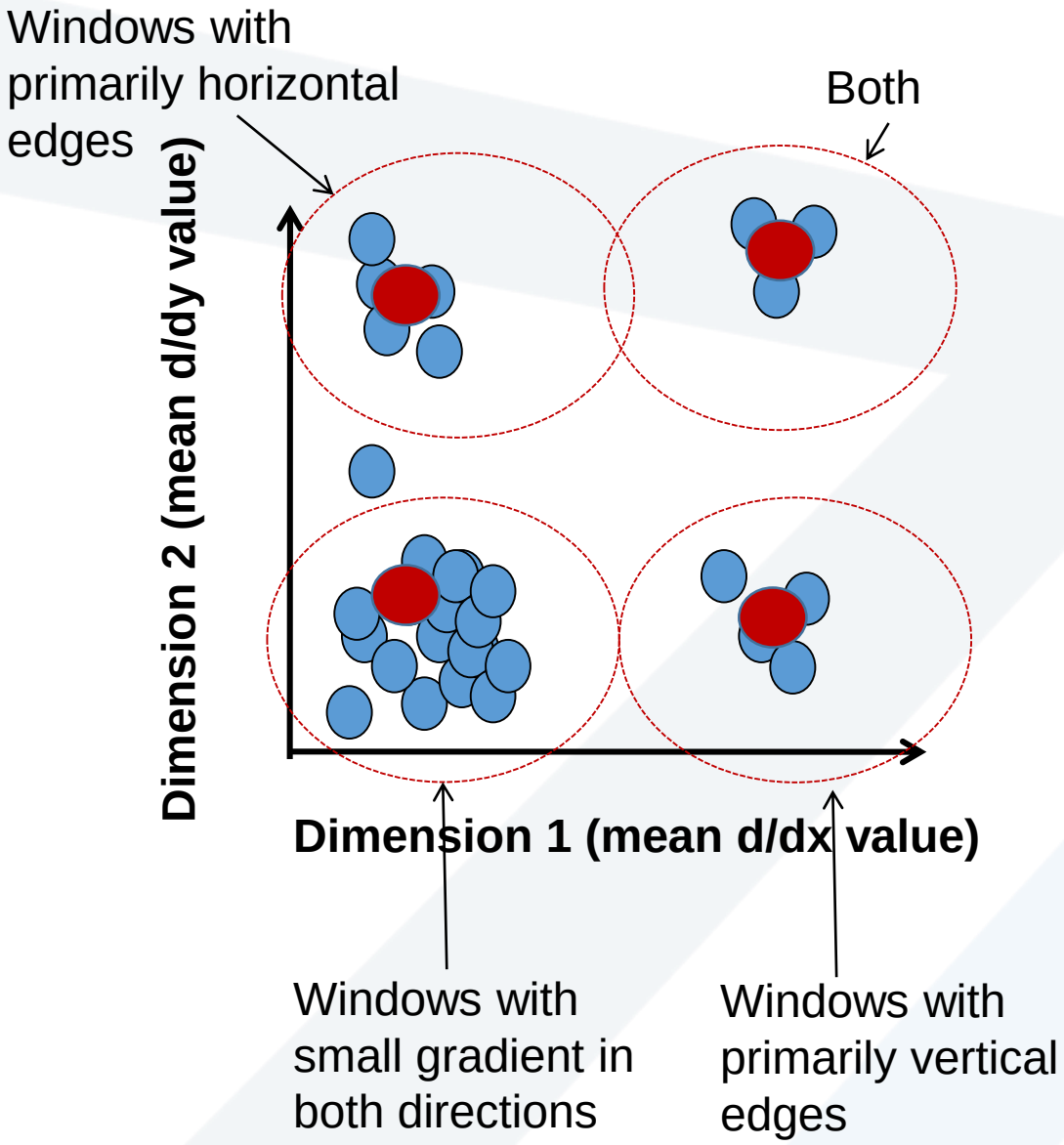
Bottom-up Segmentation

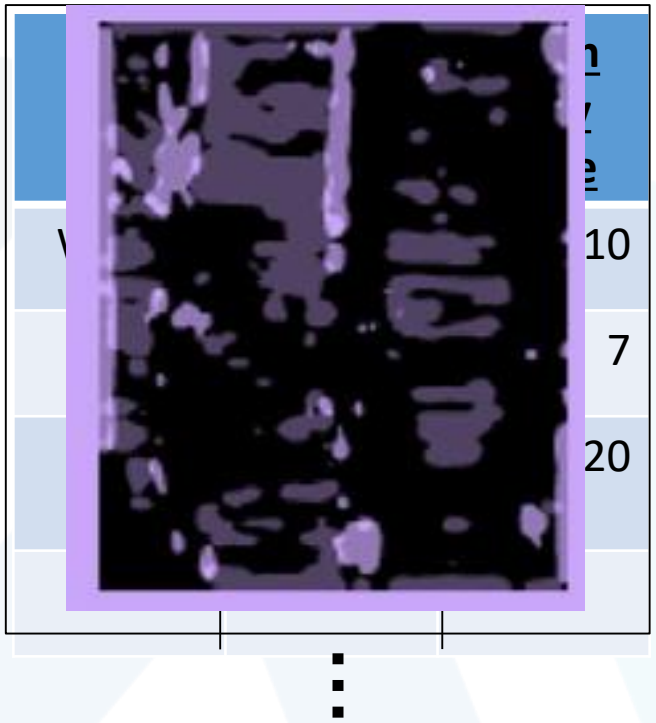
Mean shift algorithm

- Pros:
  - Does not assume shape on clusters لا تأخذ بالحسبان الشكل أثناء حساب العناقيد
  - One parameter choice (window size, aka “bandwidth”) نحتاج لتحديد بارامتر واحد فقط (حجم النافذة)
  - Generic technique طريقة عامة يمكن استخدامها في كل التطبيقات
  - Find multiple modes
- Cons:
  - Selection of window size يؤثر حجم النافذة المختارة على نتيجة العنقدة
  - Does not scale well with dimension of feature space تتأثر بمشكلة تغير الحجم

التجزيء بالاعتماد على النسيج

Bottom-up Segmentation





statistics to summarize patterns in small windows

Slide credit: Kristen Grauman

التجزيء بالاعتماد على النسيج

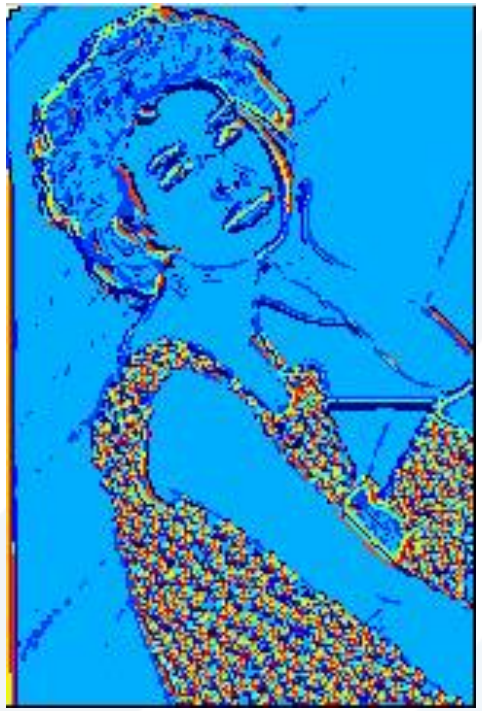
Bottom-up Segmentation

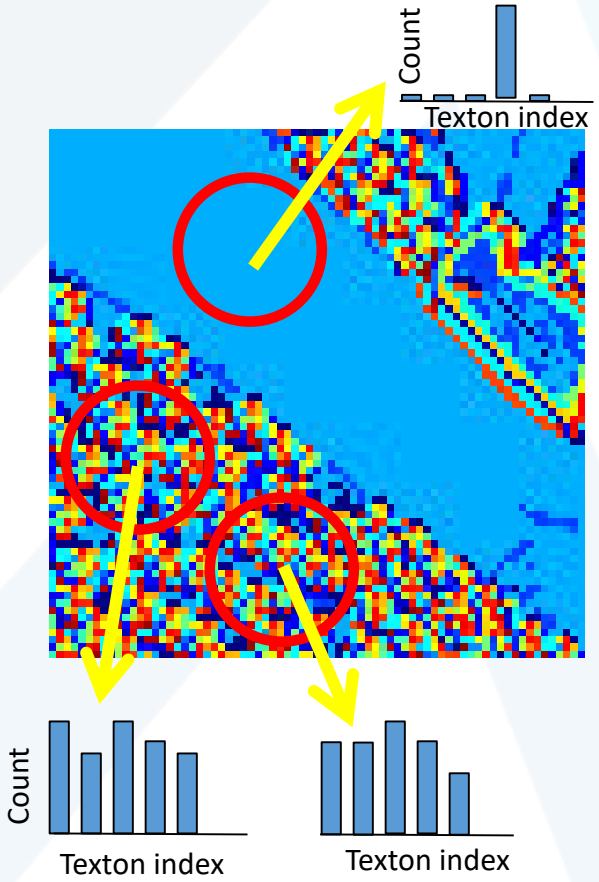
طريقة إيجاد السمات النسيجية (النسجات) Textons باستخدام بنك المرشحات (سؤال مهم في الامتحان).

- 1- أخذ نوافذ ضمن الصورة بحجم محدد.
- 2- حساب نتيجة الترشيح ضمن هذه النافذة باستخدام مرشح أو عدة مرشحات من بنك المرشحات Filter Bank.
- 3- وصف السمات النسيجية ضمن نافذة محددة بالاعتماد على هيستوغرام النسجة.

Image

Texton map



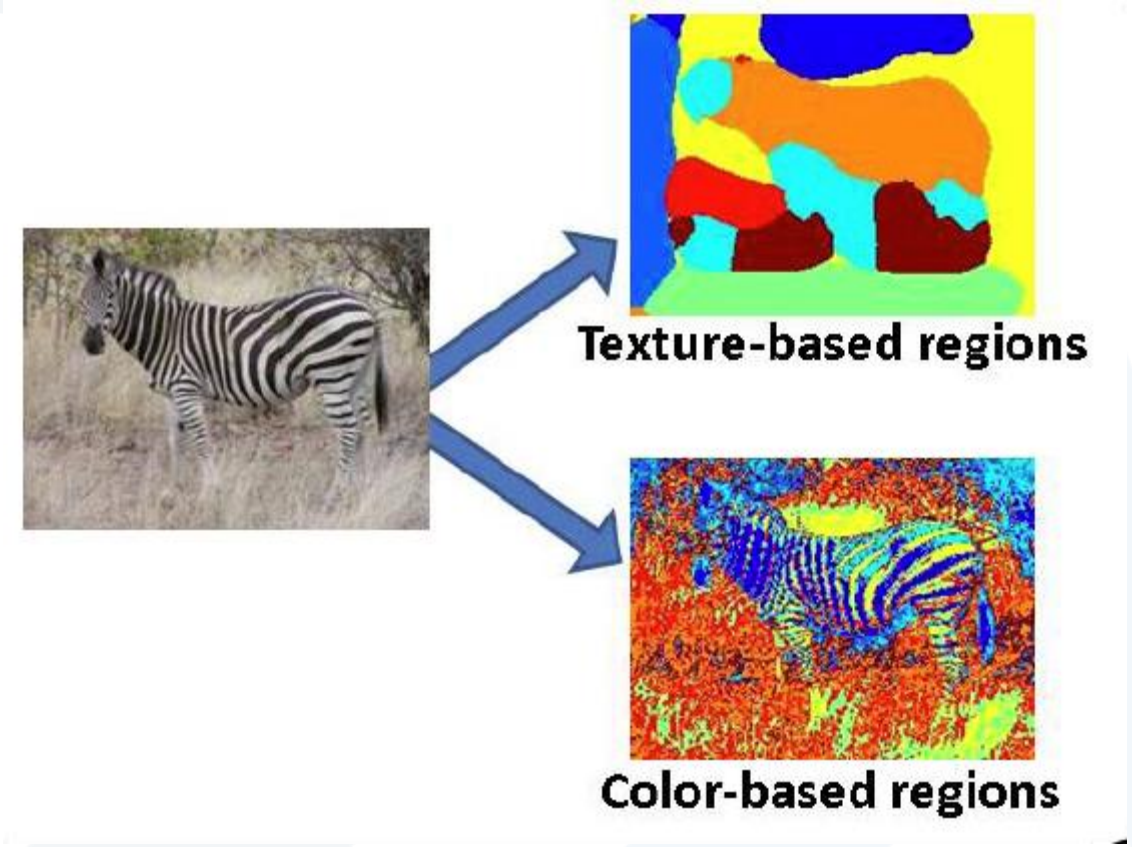


Malik, Belongie, Leung and Shi. IJCV 2001.

Reference: Lana Lazebnik

التجزيء بالاعتماد على النسيج

Bottom-up Segmentation



Slide credit: Kristen Grauman

التجزيء بالاعتماد على النسيج

Bottom-up Segmentation

تبدو السمات اللونية لهذه الصور كلها متشابهة

لكن السمات النسيجية لها مختلفة تماماً

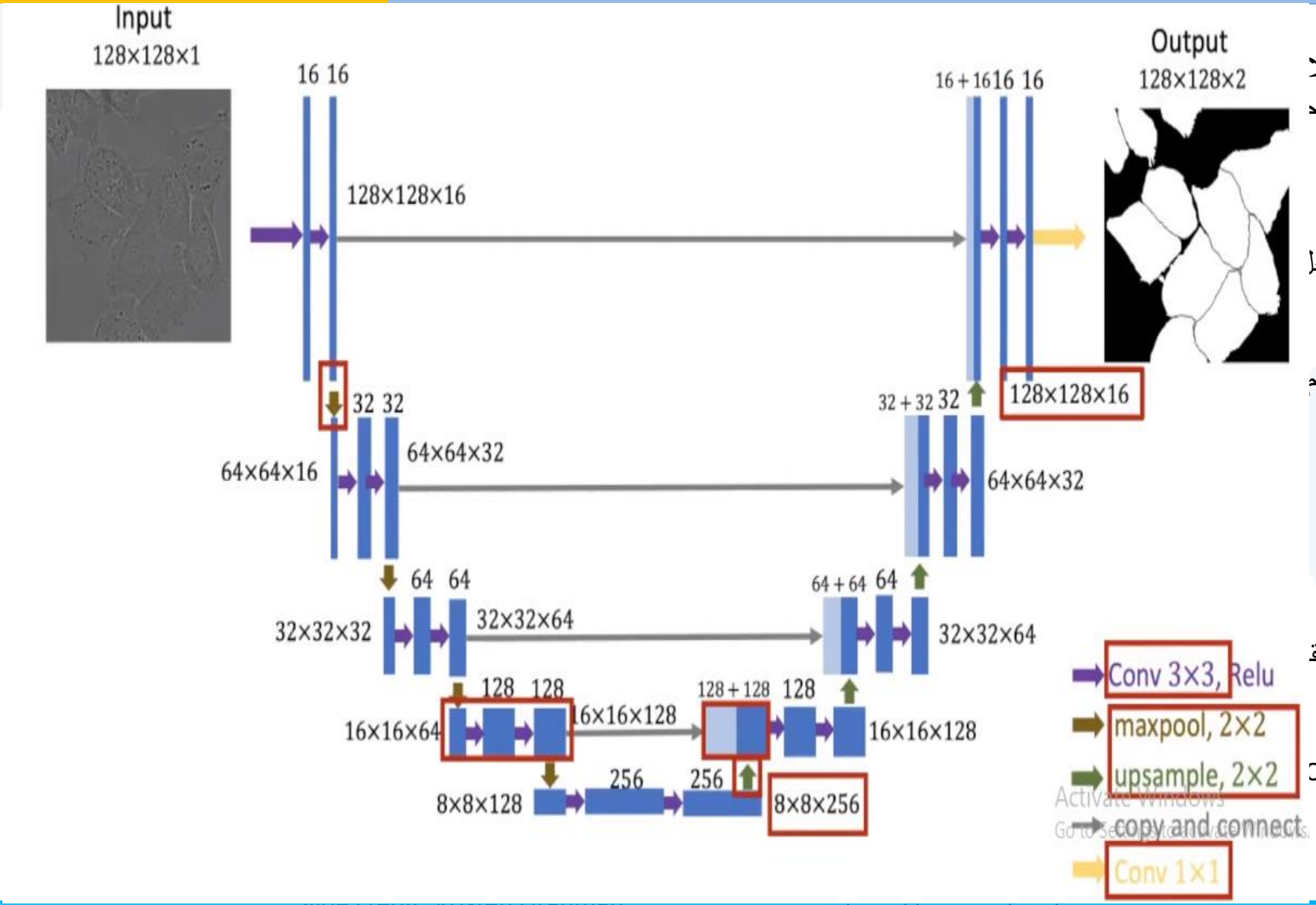
لاستخدام هذه السمات في استرداد الصور Image Retrieval لا يمكن الاعتماد على واحدة فقط!

نحتاج كلا النوعين من السمات للاستعلام عن صورة بشكل صحيح

Slide credit: Kristen Grauman

Unet model

Deep Learning-based semantic Segmentation



المقصود بمصطلح Semantic Segmentation

التقسيم الدلالي للصور حيث يتم دراسة كل بكسل من بكسلا الصورة على حدة ثم يتم إسناد Label لهذا البكسل يمثل أد الأصناف Segments التي يمكن تقسيم الصورة لها.

أشهر نماذج التعلم العميق المستخدمة في هذا النوع من التجزيء هو **شبكة Unet**

وهي شبكة تعلم عميق دخلها صورة وخرجها صورة (الدخل هو الصورة الأصلية أما الخرج فهو صورة Ground Truth أو الصورة التي تتضمن عنونة لكل بكسل إلى أحد الأصناف التي تمثل Segments المراد تجزيء أو تقسيم الصورة وفقاً لها)

تتعلم الشبكة Unet بحيث تربط الدخل (الصورة الأصلية) مع الخرج (الصورة المعنونة) بأقل خطأ ممكن.

سميت بشبكة Unet لأن لها شكل حرف U حيث تتضمن قسمين يسمى الأول **Encoder** في حين يسمى الثاني **Decoder**

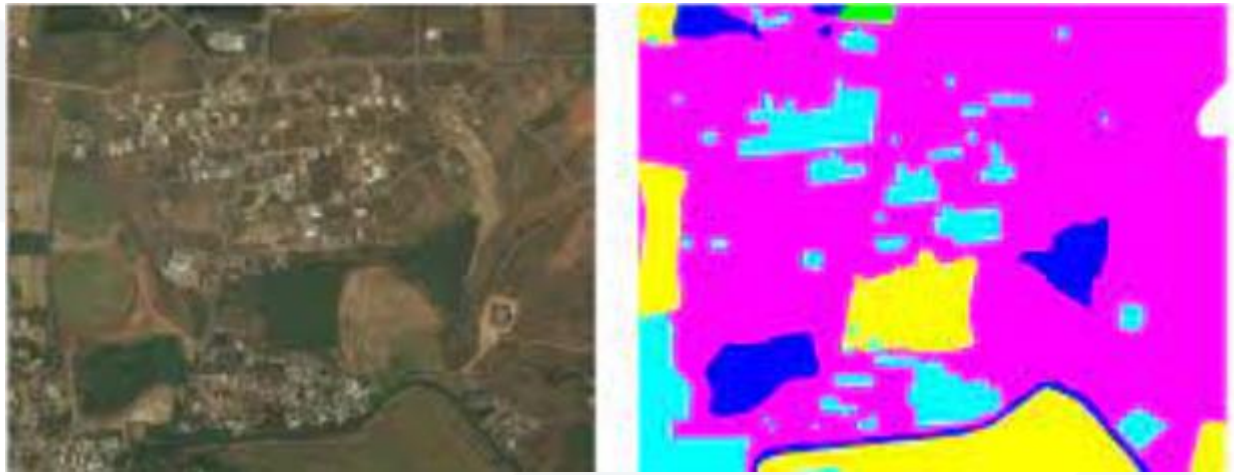
تتضمن طبقات التفافية Convolutional Layers وطبق تخفيض Down Sampling في الجزء الأيسر Encoder

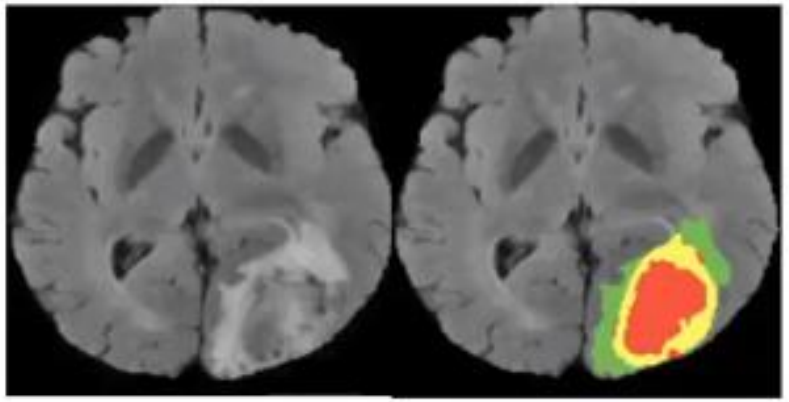
في حين تتضمن طبقات التفافية Convolutional Layers وطبقات إعادة بناء Up Sampling في الجزء الأيمن Decoder

Unet applications

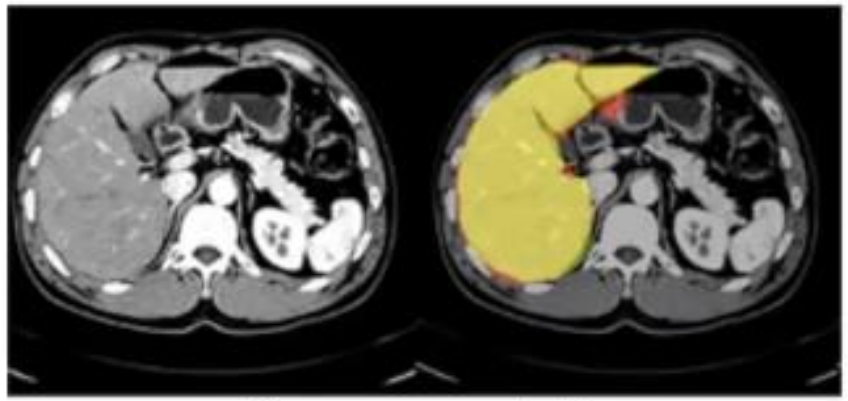
Deep Learning-based semantic Segmentation

Aerial Image Segmentation





Brain tumor segmentation



Liver segmentation