UNIVERSITÀ
DEGLI STUDI
FIRENZE

Mid-term programming assignments

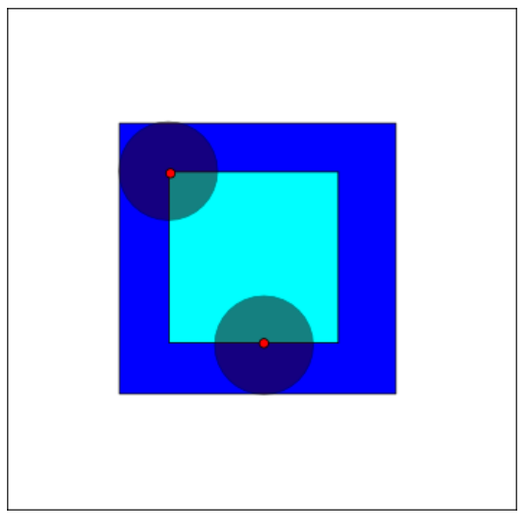
Parallel Programming

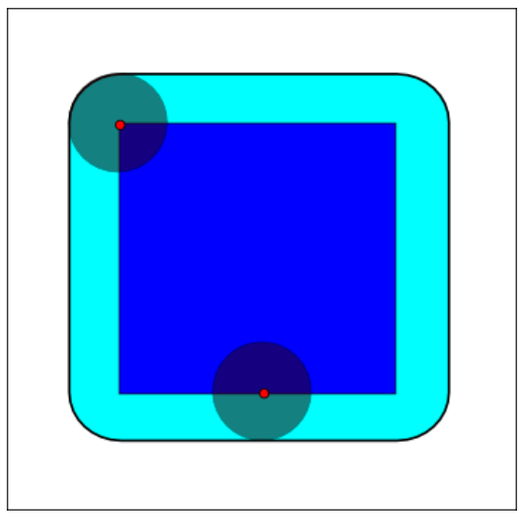
What to do

- It’s possible to implement the assignments in pairs (only 9 credit course)
- Implement a programming assignment using one of the approaches presented (parallelization in OpenMP + vectorization). Implement a sequential and a parallel version.
- Write a tech report with performance analysis (speedup) and prepare a presentation
- Keep everything on a public Github/Bitbucket/Gitlab

Morphological image processing

- Image processing technique
- Evaluate on different image sizes.
- See https://en.wikipedia.org/wiki/Mathematical_morphology





Bigrams / trigrams

- Compute histograms of bigrams/trigrams on texts (eg. Wikipedia / Gutenberg project documents)
  - Consider bigrams/trigrams of characters and words
  - How to deal with tokens ? And splitting tokens ?
- If needed re-use more and more times the collected data to create a large enough corpus

ANN retrieval

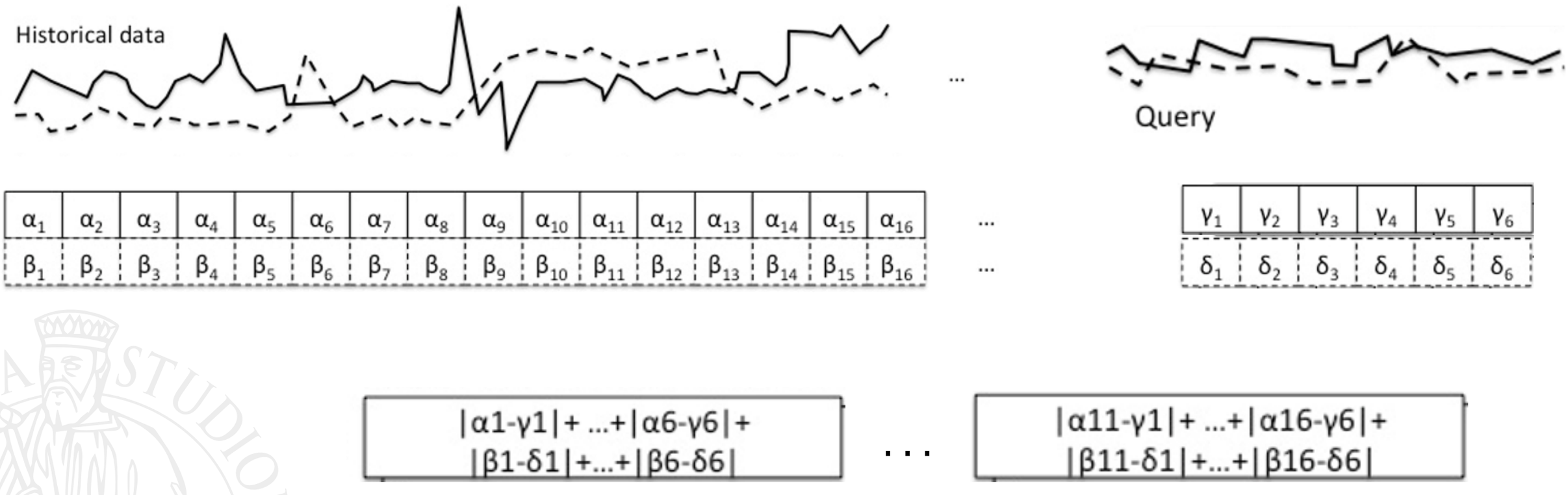
- Consider high dimensionality vectors (eg. 128)
- Given a specified query find the *k nearest neighbors* in a database of vectors
- Can use LSH for approximate indices, eg. using mlpack o LSHkit
- Can get sample datasets from http://corpus-texmex.irisa.fr

Password decryption

- Decrypt passwords encrypted using crypt (DES)
  - Consider 8 characters lengths in the set [a-zA-Z0-9./]
  - Library funds may be deprecated or not available. E.g. CUDA does not provide the crypt C-library function. Must use a special implementation (check Moodle).
  - It’s a (slowed down) search problem.
  - Can choose to attack a specific password in a list (in different positions), or decrypt a full list of passwords (in this case reduce the search space)
  - Can consider only dates or common 8-chars passwords

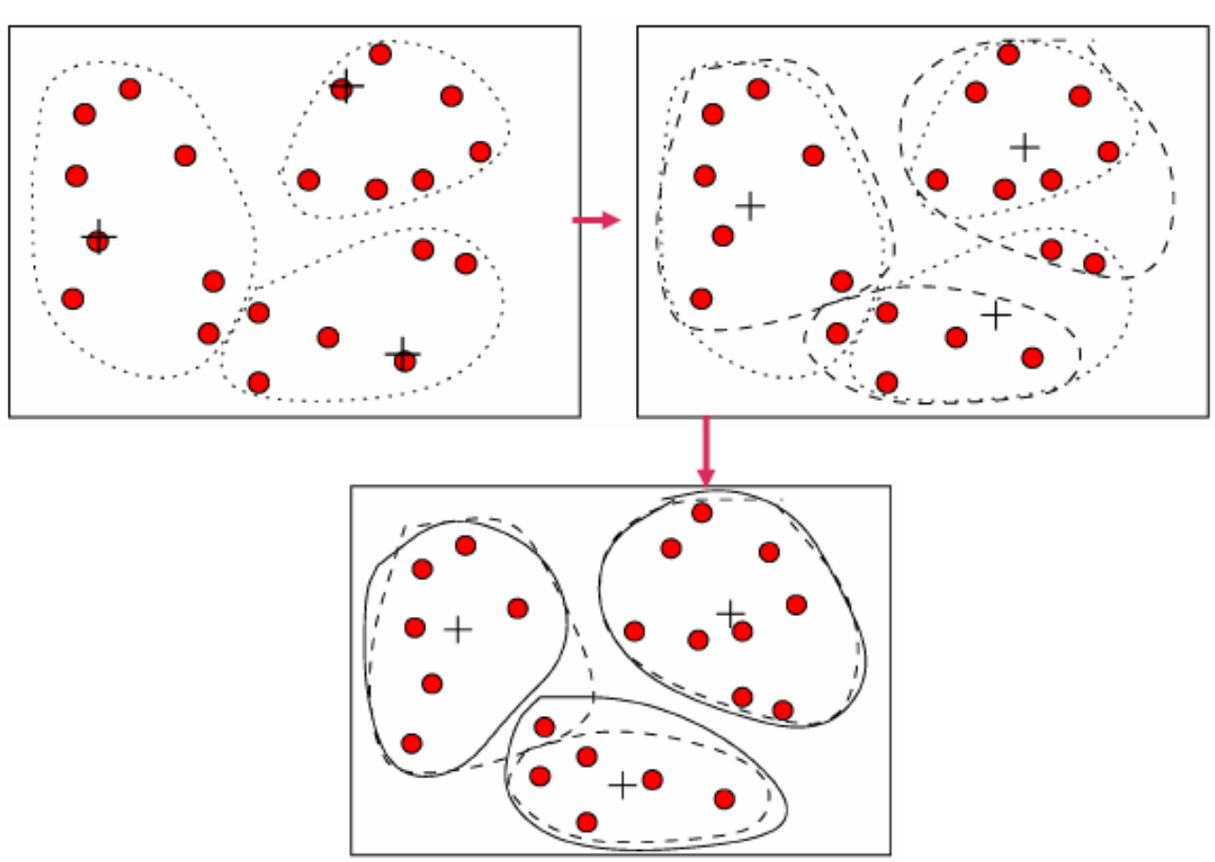
Pattern recognition

- Search a given time series within a larger and longer set of time series
  - Get time series from existing machine learning datasets or generate using: https://github.com/cyrilou242/mockseries
- Use SAD as the metric to evaluate the match



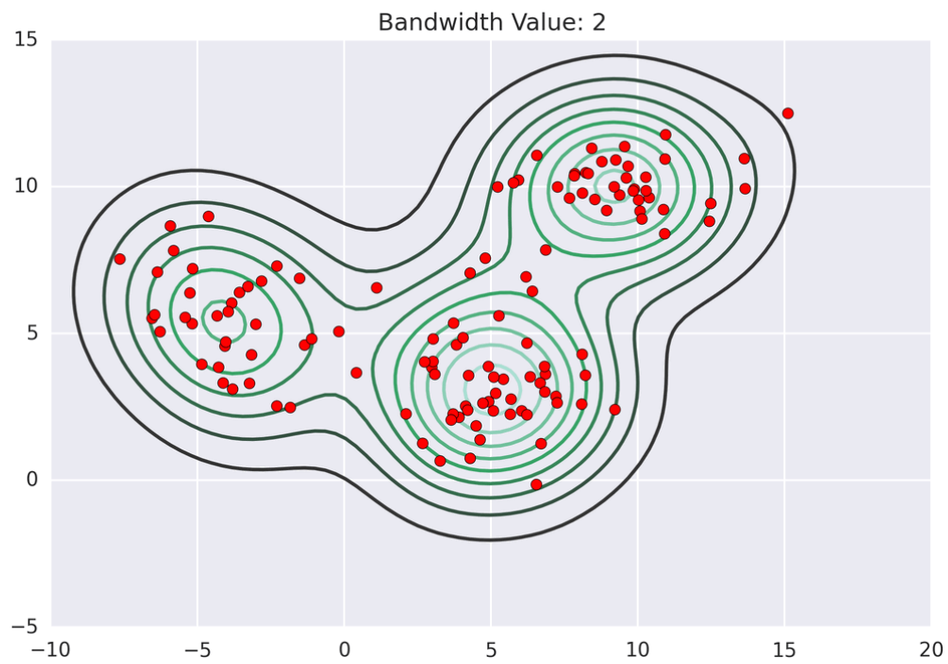
k-means

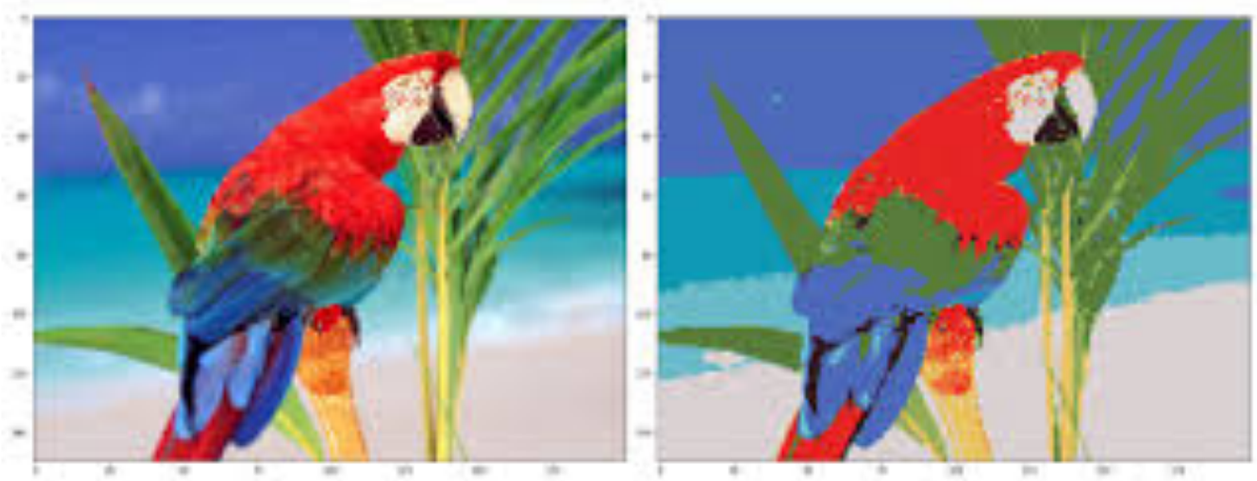
- Implement a parallel version of k-means clustering
  - There’s no need to implement K-means++ centroid assignments
    - Careful with experiment reproducibility: use fixed iterations or the same starting points
  - Feel free to chose to operate on 2D or larger dimensionality vectors



Mean shift clustering

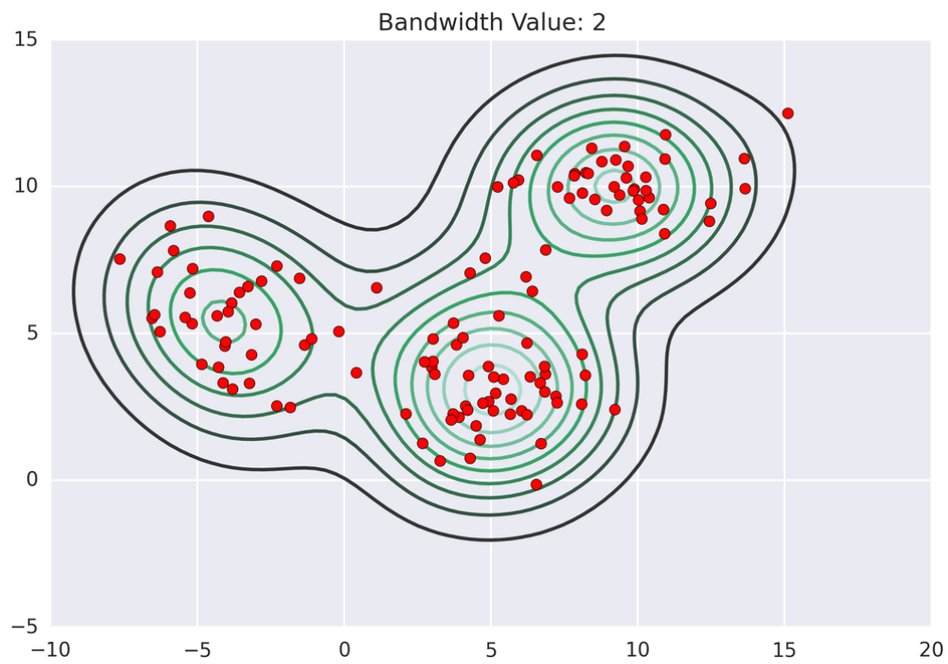
- Implement a parallel version of mean-shift clustering. This clustering doesn't require to specify the number of desired clusters (it uses KDE, so needs a bandwidth parameter) but its complexity is $O(N^2)$
- Can be used to segment images

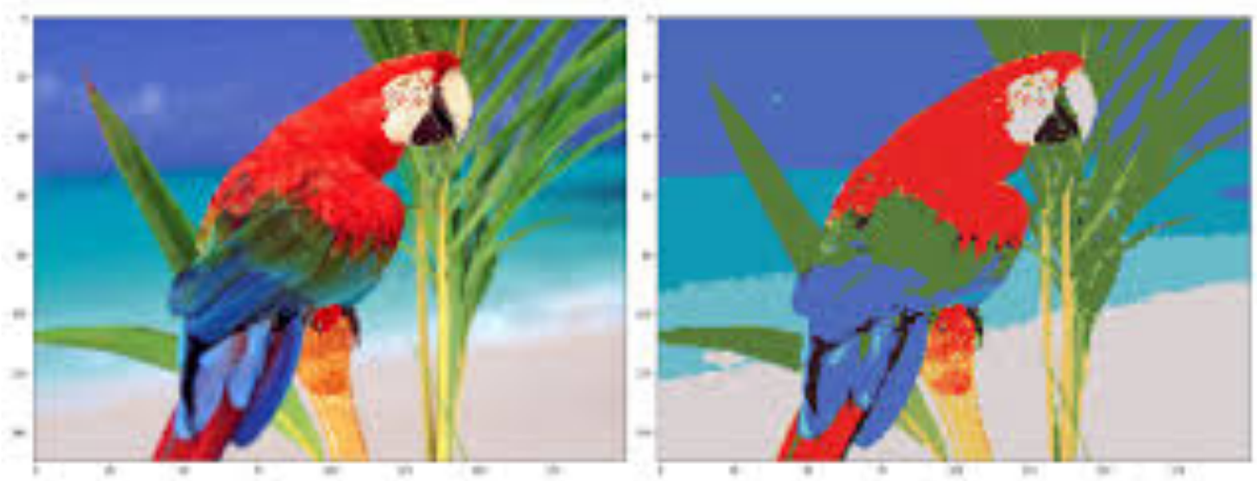




Mean shift clustering

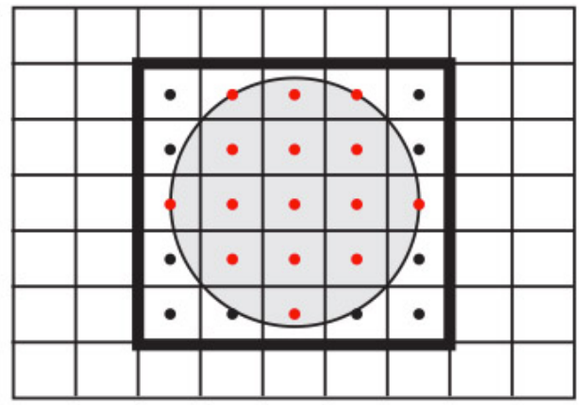
- Implement a parallel version of mean-shift clustering. This clustering doesn't require to specify the number of desired clusters (it uses KDE, so needs a bandwidth parameter) but its complexity is $O(N^2)$
- Can be used to segment images



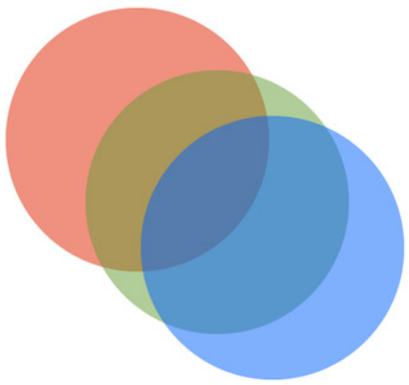


Renderer

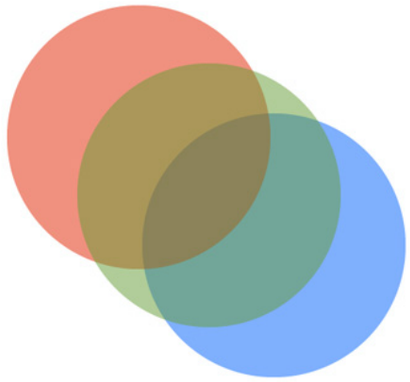
- Implement a renderer that draws (semi) transparent circles (or other shapes). Circles have 3D coordinates and the order along Z axis matters for the correct rendering



- A pixel belongs to a circle if its center is within the circle.



Correct order:
blue over green over red



Incorrect order

Random 10K

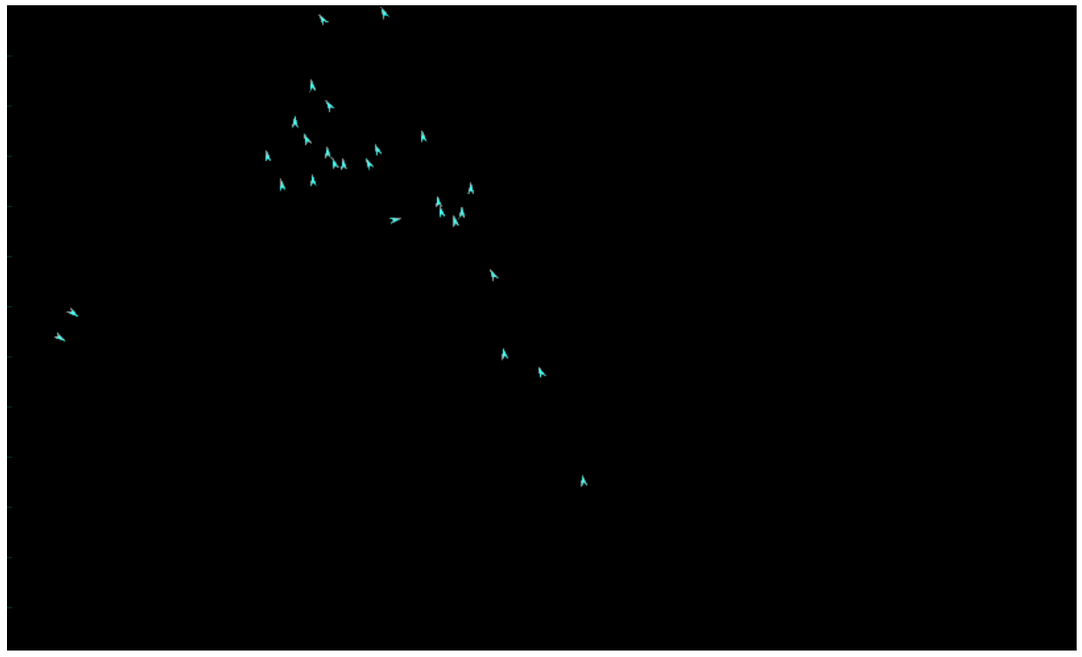
Random 100K

Image composition

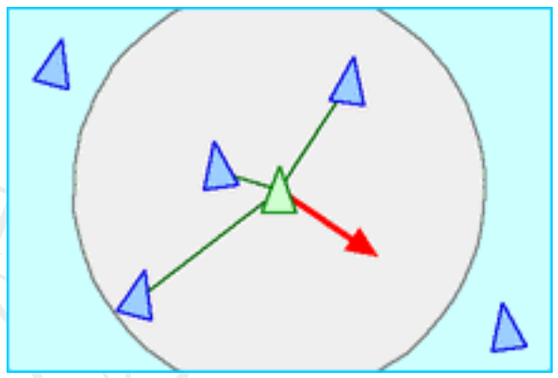
- It’s a variation of the Render assignment.
- Implement a parallel alpha composition of images
- Useful to create image augmentation to train Convolutional Neural Networks, e.g. for detection (see Blend augmentation in ImgAug library)
- Can parallelize applying the same object to different images, or in different positions in the same image

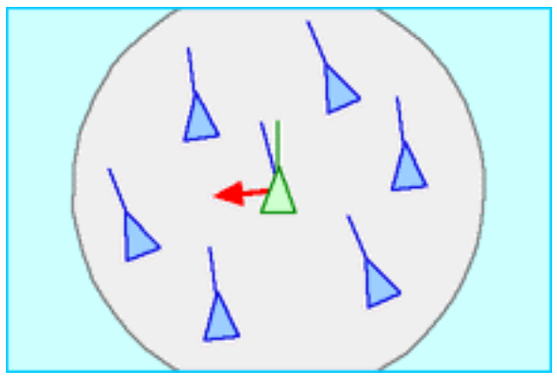
Boids

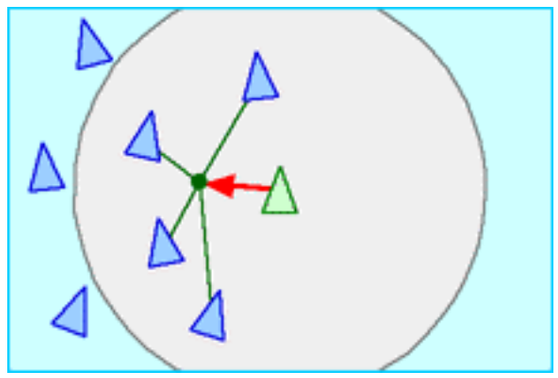
- Simulate the behavior of a flock / crowd.
  - Boids is an artificial life program which simulates the flocking behaviour of birds, and related group motion.



- Check https://vanhunteradams.com/Pico/Animal_Movement/Boids-algorithm.html

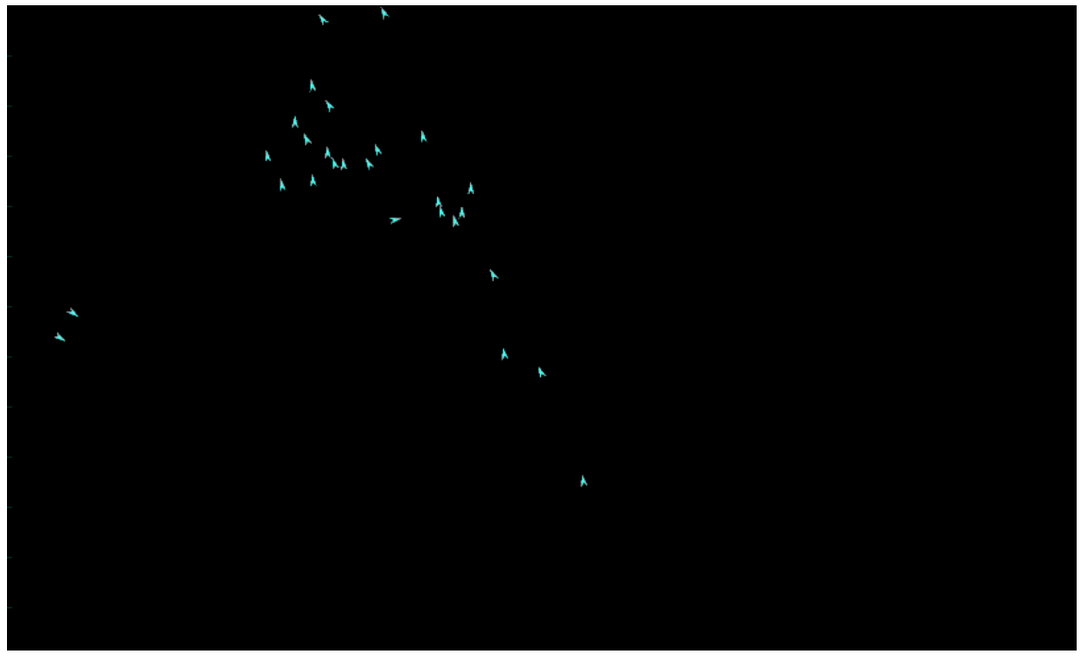






Boids

- Simulate the behavior of a flock / crowd.
  - Boids is an artificial life program which simulates the flocking behaviour of birds, and related group motion.



- Check https://vanhunteradams.com/Pico/Animal_Movement/Boids-algorithm.html