A novel trajectory segmentation technique based on the K-means algorithm.

Features:
- Accurate: it guarantees the convergence to a minimum-distortion segmentation;
- Robust: each run for a given K always yields in the same result;
- Fast: only two clusters are inspected in each classification step;
- Near-linear complexity: $\Theta(kd)$ instead of $\Theta(nkd)$;
- Not extra input parameters: just the same as in K-means;
- Real-time, large DBs: the number of data points does influence when updating the centroids;
- On-line learning: clusters can be updated while new samples arrive without affecting the previous data structure.

Why?
Leverage K-means strengths (simplicity, fast convergence) minimizing drawbacks (initialization, instance order).

How?
Initialize data with Trace Segmentation [2] and impose temporal constraints to the sequential version of K-means [1].

Trace Segmentation

Input: Trajectory $x_1, \ldots, x_N$; Traces $M \geq 2$
Output: Normalized trajectory $y_1, \ldots, y_M$

- $L_1 = 0$
- for $n = 2$ to $N$ do
  - $L_n = \lambda_{n-1} + \|x_n - x_{n-1}\|$
  - $\lambda = \frac{L_n}{\lambda_{n-1}}$
  - $y_1 = x_1$
  - $n = 2$
- for $m = 2$ to $M-1$ do
  - while not $L_{n-1} \leq (m-1)\lambda \leq L_n$ do
  - $n++$
  - $y_m = x_{n-1} + (x_n - x_{n-1})\frac{(m-1)\lambda - L_{n-1}}{L_n - L_{n-1}}$
  - $y_M = x_N$
  - // last point

Sequential Clustering

Input: Trajectory $X$; No. clusters $C$
Output: Partition $\{\mu_1, \ldots, \mu_C\}$; Distortion $J$

- $J^0 = \{X_1, \ldots, X_C\}$ // use trace segmentation
- for $c = 1$ to $C$ do
  - $\mu_c = \frac{1}{N_c}\sum_{x \in X_c} x$
  - $J = J + \sum_{x \in X_c} ||x - \mu_c||^2$
  - while transfers do
    - transfers = false
    - for all $i : x \in X_i$ do
      - if $n_i > 1$ then
        - $y^* = \arg \min_{y \in X_i} \frac{1}{n_i} \sum_{x \in X_i} ||x - \mu_y||^2$
        - $\Delta J = \sum_{x \in X_i} ||x - \mu_y||^2 - \frac{1}{n_i} \sum_{x \in X_i} ||x - \mu_x||^2$
      - if $\Delta J < 0$ then
        - $\mu_x = \mu_x - \frac{\Delta J}{||x - \mu_x||^2}$
        - $\mu_y = \mu_y + \frac{\Delta J}{||x - \mu_y||^2}$
        - $X_i = X_i - \{x\}$
        - $X_y = X_y + \{x\}$
        - $J = J + \Delta J$

Steps Explained

References

Acknowledgements
Work supported by the Spanish research programme Consolider Ingenio 2010: MIPRCV (CSD2007-00018).