Random Forest Clustering and Application to Video Segmentation

Frank Perbet
http://frank.perbet.org
Bjorn Stenger
http://mi.eng.cam.ac.uk/~bdrs2/
Atsuto Maki
atsuto.maki@crl.toshiba.co.uk

Toshiba Research Europe
Cambridge Research Laboratory
http://www.toshiba-europe.com/research/crl/cvg/

Clustering is a partitioning of a data set into subsets so that the data points in each subset are homogeneous and more similar to each other than those from different subsets. This paper considers the problem of automatically clustering very large data sets in a high-dimensional space. We introduce the notion of a Graph of Superimposed Partitions (GSP) that represents a remapped space of the input data where regions of high density are mapped to a larger number of nodes (See Figure 1 (a) and (b)).

Using a random decision forest [2], we first generate multiple partitions of the same input space, one per tree. A node is then assigned to each intersection of the partitions which are obtained by merging the outputs of the trees, and subsequently the graph is constructed by linking the nodes which are spatially adjacent. Generating such a graph essentially turns a space clustering problem into a graph clustering one which we solve by employing the Markov cluster algorithm (MCL) [6].

The proposed algorithm is able to capture non-convex structure while being computationally efficient, capable of dealing with millions of data points. We show the performance on synthetic data and apply the method to the task of automatic video segmentation.

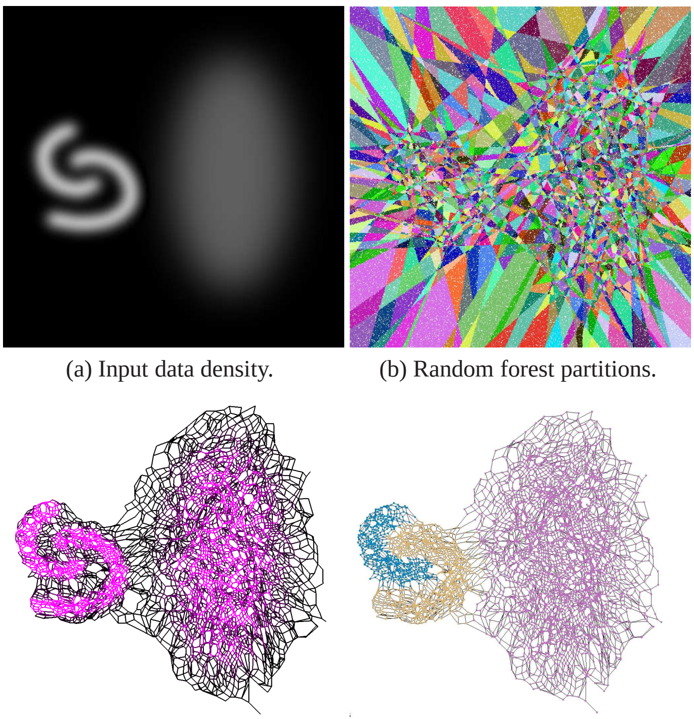
In the first step a random decision forest [2] is used to generate multiple partitions of the same input data space, one partition per tree. This process spreads the density of data samples evenly, similar to histogram equalization but applied to a high dimensional space. In the second stage, we superimpose these partitions and identify their *intersections* which constitute our compact *partitions*. We assign a node to each intersecting small volume and subsequently construct the graph by linking the nodes that are spatially adjacent. We define this as the Graph of Superimposed Partitions (GSP).

Since each node of the GSP stands for all the data within each intersection, generating the GSP essentially turns the problem of space clustering into that of a graph clustering problem for which there exist a range of techniques, see [6, 7] for examples. In this paper the Markov cluster algorithm (MCL) [6] is employed owing to its scalability.

As the decision forest plays an important role in our algorithm, it is worth mentioning *clustering trees* [1, 3] that are designed such that the leaves define a spatial partitioning of the high dimensional space. While a single tree using linear split functions is not capable of finding non-convex volumes by its nature, the motivation for using a decision forest is that by combining the partitions of multiple trees we can potentially deal with non-convex clusters. It should also be noted that decision forests have been applied for the task of creating visual codebooks for bag-of-feature image classification [4, 5]. However, in this case the leaf indices from multiple trees are simply stacked into a vector, and not integrated. The resulting partitions are useful for the task of codebook generation, but do not necessarily capture the underlying structure of the data.

The contributions of the proposed algorithm are: (i) its ability to cluster data on non-convex manifolds, (ii) it does not require the number of clusters, nor kernel radius as parameters, and (iii) its scalabililty allows it to deal with millions of data points.

As an example application, we apply the GSP algorithm to the video segmentation problem. The data space is six dimensional and includes the pixel coordinates, time and the colour values: $[x,y,t,r,g,b]$, each coordinate scaled to the range $[0,1]$. The clustering algorithm is run on the complete space-time volume and therefore insures that segmented regions are consistent over time.



(a) Input data density. (b) Random forest partitions.

(c) Graph structure. (d) Clustering result.

Figure 1: **Overview of the proposed clustering algorithm.** *(a) PDF of a 2D synthetic data set where brighter regions correspond to higher density. (b) Compact partitions (intersections of multiple partitions from decision forest). (c) The edges of the Graph of Superimposed Partitions (GSP), brighter colour corresponds to larger edge weight. (d) The nodes of the GSP coloured by the final cluster ID.*

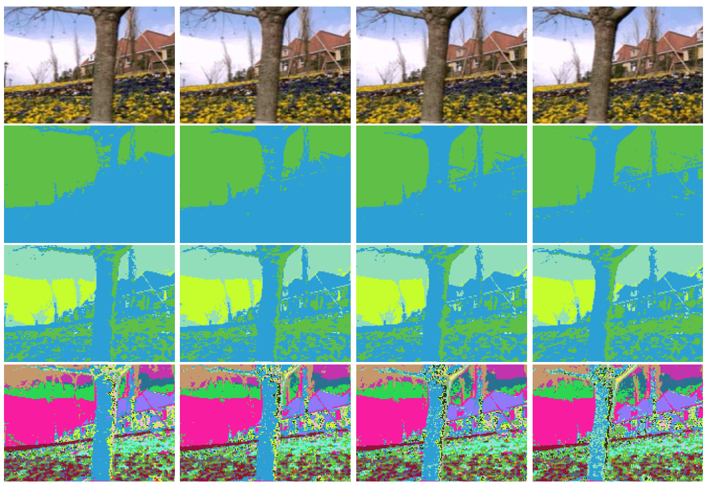


Figure 2: **Results on Flower Garden Sequence.** *Top: input frames. Rows below: segmentation with increasing cluster resolution parameter.*

[1] H. Blockeel, L. D. Raedt, and J. Ramong. Top-down induction of clustering trees. In *ICML*, pages 55–63, 1998.

[2] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[3] B. Liu, Y. Xia, and P. S. Yu. Clustering through decision tree construction. In *CIKM*, pages 20–29, 2000.

[4] F. Moosmann, B. Triggs, and F. Jurie. Fast discriminative visual codebooks using randomized clustering forests. In *NIPS*, pages 985–992, 2006.

[5] J. Shotton, M. Johnson, and R. Cipolla. Semantic texton forests for image categorization and segmentation. In *CVPR*, 2008.

[6] S. van Dongen. *Graph Clustering by Flow Simulation*. PhD thesis, University of Utrecht, 2000.

[7] Z. Wu and R. Leahy. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE-PAMI*, 15(11):1101–1113, 1993.