ImageQuest: From Multiple Inaccurate Trajectories to One Accurate Trajectory

Aparna Das, Stephen Kobourov, Sankar Veeramoni and Yunhao Xu
University of Arizona – Department of Computer Science

Introduction

Scientists can capture images of objects from the microscopic scale to the planetary scale but manual time-consuming work is often required to process these images. Meanwhile, the public's interaction with image processing has exploded, e.g. Facebook users annotate pictures an incredible 80 million times each day. People also spend millions of hours playing games like Angry Birds and Farmvile.

Goal. Develop platform to allow citizen scientists to trace

- static objects such as trees in images
- trajectories of moving objects such as ants and bees in videos

Traces of static objects can provide quantitative measurements such as size, shape and appearance. Traces of trajectories in video data are useful to biologists studying behavioral patterns in insect colonies.

Data Sets

Tree Annotation.

Video of Ant Colony.

Problem Statement

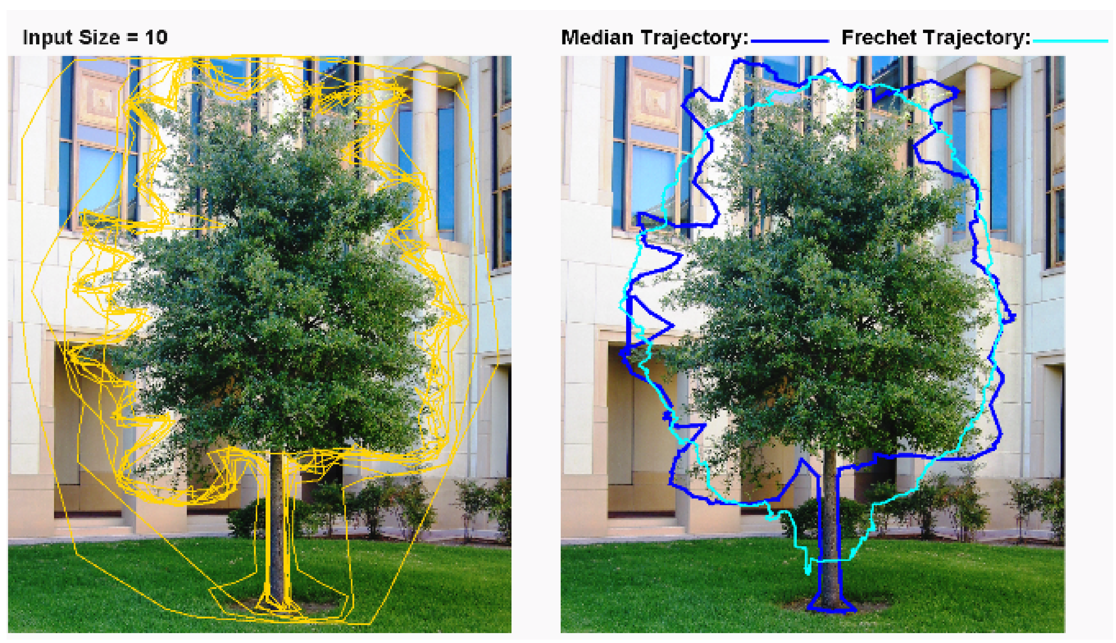
For each object we will collect traces of the object from multiple citizen scientists and compute a *consensus* of these traces. We study the problem of how to extract

- an accurate outline of a static object from multiple (inaccurate) traces
- an average trajectory for a moving object from multiple (inaccurate) traces.

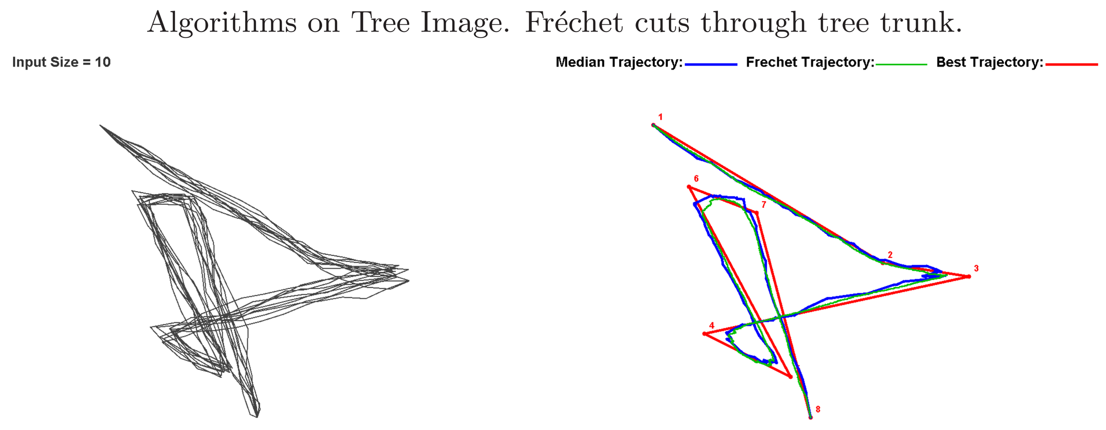
Comparison of Consensus Algorithms

We test two **consensus** algorithms for the above problems: the **Sample Fréchet** and **Median Trajectory** algorithms. We use two static data sets, the tree image data set and a generated data set of trajectories.

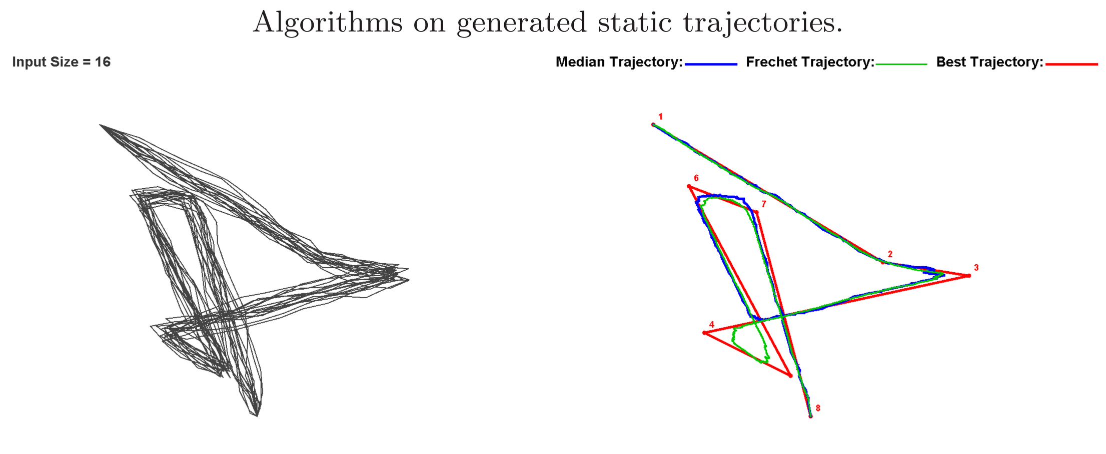
The Median trajectory algorithm works well when the arrangement of curves have faces of similar sizes, as in the tree data. If the arrangement has faces of many different sizes, Median trajectory can miss small faces. When the input curves have large Fréchet distance in certain regions and small Fréchet distance in other regions the final output of the Fréchet Sample algorithm can look very different from the inputs in the regions with small Fréchet distance.



Algorithms on Tree Image. Fréchet cuts through tree trunk.



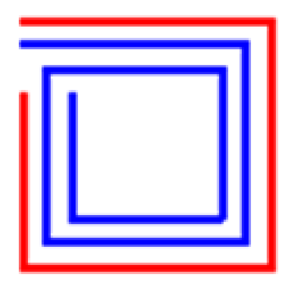
Algorithms on generated static trajectories.

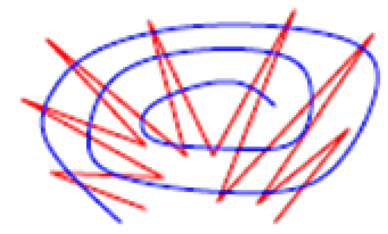


Median Trajectory misses small face.

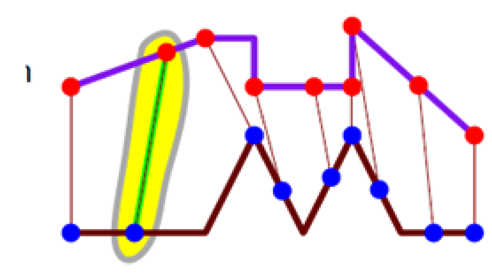
Sampling the Fréchet Alignment

Based on the approach of sampling and averaging, first we align the traces from multiple users so the most similar parts are aligned together (e.g. similar to alignment of multiple DNA sequences). We use Fréchet distance to measure similarity between traces which considers the overall structure of the trace better than nearest neighbor based similarity measures such as Hausdorff distance. Informally the Fréchet distance of two traces is the *dog-leash* distance, where a man walks along one trace and dogs along the other traces. The Fréchet distance is the minimum length leash necessary for the man to walk the dog remaining connected at all times by the leash. The computation of Fréchet produces an alignment of the traces: at each step the position of the man is mapped to the position of the dog [2].





Two curves with small Hausdorff and large Fréchet distance.

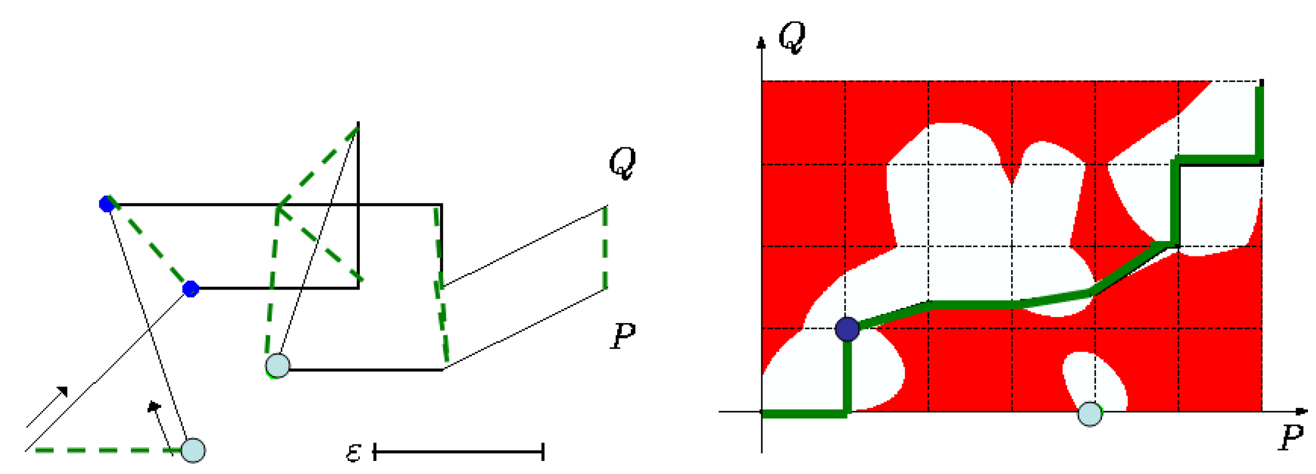


Fréchet distance minimizes leash length.

Given the Fréchet alignment of two traces we compute their consensus by taking the midpoint of the leash over time, as the man and dog complete their walk. To find the consensus of a set of traces T, we repeatedly take two traces from T compute their consensus and replace the two traces with their consensus, thus reducing the size of T by one. We repeat the process until T contains one trace.

Fréchet Sample Algorithm
Input:Set of traces T
Output: The consensus of T

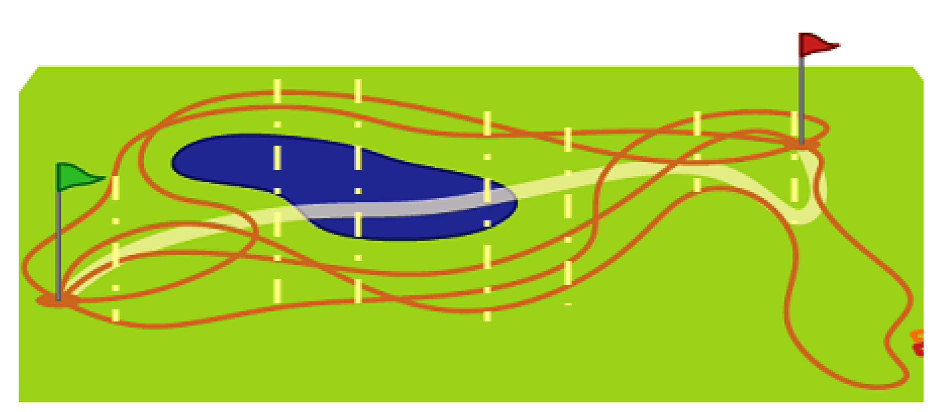
while T has more than one element **do**
  Let P and Q be two elements in T and A be their Fréchet alignment.
  for each edge of alignment A **do**
    Sample a point from P and a point from Q and find their midpoint
  end for
  Define the consensus C of P, Q by connecting midpoints in the order as they appear in the alignment.
  Replace P, Q with C reducing the size of T by one.
end while



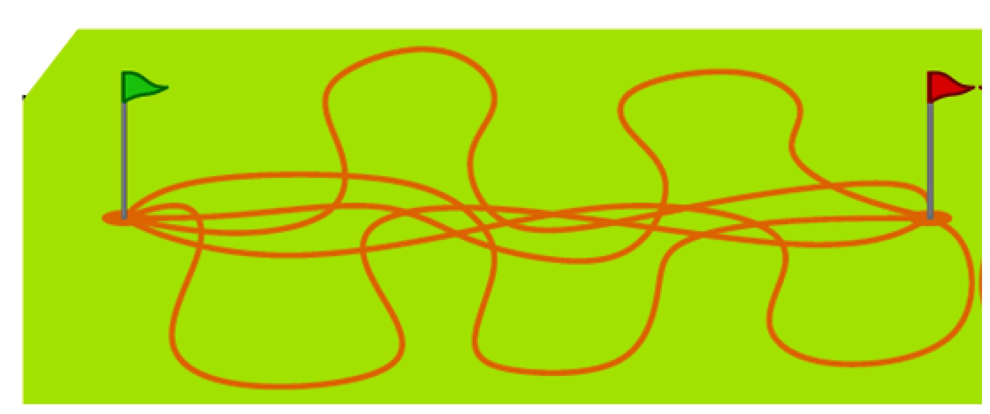
The green curve is the Fréchet alignment P, Q.

Median Trajectory

Averaging locations, one from each trace, parallels the way we measure the average of a data set of numbers. However averaging locations of traces may cause the consensus to interfere with environmental obstacles such as lakes. Thus in the second approach we extend the notion of a computing a median of a set of numbers by constraining ourselves to picking a consensus which always stays on one of the input traces. However for some inputs there may be no one input trace which is a good representative of the set of inputs.



Averaging locations leads to a consensus trajectory that goes through a lake.



No input trace is a good representative of the inputs.

The Median Trajectory Algorithm proposed by Buchin et. al. [3] constructs a consensus trace using actual pieces of different input traces. Assuming all traces have a common start and end point, consider the **arrangement** of traces. The consensus starts by taking the piece of the trace that is currently in the middle, beginning with the common start point. At the next intersection point in the arrangement (possibly) another trace is in the middle and the consensus takes a piece from that trace. The process continues at each intersection point in the arrangement until the common end point is reached.

Median Trajectory Algorithm [3]
Input: set of traces T with a common start and common end point.
Output: consensus of T

1. Consider the arrangement of all the traces in T.
2. Start at the common start and travel along the **middle** trajectory choosing the middle trajectory at each intersection point until common end point is reached.

Consensus using different pieces of inputs.

Future Work

- **Tracking Moving Objects.** For traces of moving objects, each point of the trace is also associated with a time stamp. The algorithms we have considered do not consider the time stamp. We plan to define the consensus by aligning input traces by time stamps, finding the centers per timestamp, and connecting centers.
- **Building Paths Together.** The above algorithms build the path of each ant one at a time. We also plan to design algorithms which build a graph of the paths of **all** ants in the video together.

References

[1] C. Alpert, C. Grimm, S. Kobourov, J. O'Neil-Dunne, R. Pless, R. West. ImageQuest: Citizens Advancing Biology with Calibrated Imaging and Validated Analysis. NSF Proposal 10-538, 2010.
[2] H. Alt and M. Godau. Computing the Frchet distance between two polygonal curves. In *Internat. J. Comput. Geom. Appl 1995.*
[3] K. Buchin, M. Buchin, M. Van Kreveld, M. L´offler, R.I. Silveira, C. Wenk, L. Wiratima. Median Trajectories. In *ESA' 10.*