# Distributed Scene Understanding

We have reached a technology inflection point, in which the availability of miniature, network-enabled devices (*motes*), along with the availability of small battery-operated cameras, provides a tremendous opportunity for using collections of camera sensor motes to observe a scene and carry out automatic visual analysis in a fully decentralized manner.

Our work envisions a new paradigm of “Distributed Computer Vision” that uses the distributed albeit low processing power of battery operated, wireless camera sensors to perform high-level complex vision tasks. In this new paradigm, cameras no longer transmit image data or low-level feature information to a single server that performs the analysis in a centralized fashion. Instead, each mote analyzes data locally and the motes collaborate to reach a global analysis of the scene by sharing their respective analyses with their neighbors in the *vision graph* (i.e., who can see what).

Given this new paradigm, it is necessary to develop the theory behind distributed algorithms for solving several computer vision problems. In particular, we are developing robust consensus algorithms on manifolds for solving computer vision problems such as camera localization, 3D reconstruction, object tracking and object recognition in a *fully decentralized* fashion.

**Distributed pose estimation.** In [4] we addressed the problem of estimating the pose of an object from multiple views acquired by a calibrated camera sensor network. Each camera obtains a local estimate of the object’s pose using classical techniques from multiple view geometry. A global estimate of the object’s pose is obtained by finding the average pose (Karcher mean in $SE(3)$) in a distributed fashion. This is done using extensions of Euclidean consensus to non-Euclidean spaces.

**Distributed localization.** In [2] we proposed a distributed algorithm for calibrating a camera sensor network from local point correspondences. Each camera uses classical techniques to estimate its pose relative to its neighbors. The estimation of the absolute camera poses is then posed as a distributed consensus problem on $SE(3)$.

**Distributed face recognition.** In [1] we considered the problem of distributed face recognition under varying viewpoint. Each camera can access only a subset of the global training set and uses classical techniques such as eigenfaces and tensorfaces to classify a face. However, such estimates may be wrong due to the limited training set, low-resolution of the images or incorrect view point. Our method proceeds by averaging local estimates of the face’s pose using the method in [4]. This global pose estimate significantly improves the performance of the local classifiers.

**Distributed structure from motion.** In [3] we showed that the affine scene reconstruction problem can be posed as a distributed optimization problem involving distributed least squares and distributed PCA. We showed how such problems can be solved using consensus.

## References

[1] R. Tron and R. Vidal. Distributed face recognition via consensus on SE(3). In *Workshop on Omnidirectional Vision*, 2008.

[2] R. Tron and R. Vidal. Distributed 3-d localization in camera networks. In *Conference on Decision and Control*, (submitted) 2009.

[3] R. Tron and R. Vidal. Distributed affine structure from motion. In *IEEE International Conference on Computer Vision*, (submitted) 2009.

[4] R. Tron, R. Vidal, and A. Terzis. Distributed pose averaging in camera networks via consensus on SE(3). In *International Conference on Distributed Smart Cameras*, 2008.