CLAIMS

1. A method of controlling a vehicle in an autonomous driving mode, the method comprising:

receiving, by one or more processors, sensor data identifying a plurality of objects;

identifying pairs of objects of the plurality of objects;

determining, by the one or more processors, for each identified pair of objects of the plurality of objects, a similarity value which indicates whether the objects of that identified pair of objects can be responded to by the vehicle as a group;

clustering, by the one or more processors, the objects of one of the identified pairs of objects based on the similarity value; and

controlling, by the one or more processors, the vehicle in the autonomous driving mode by responding to each object in the cluster in a same way.

2. The method of claim 1, wherein the clustering is further based on a distance between the objects of the one of the identified pairs of objects.

3. The method of claim 1, wherein the clustering is further based on a similarity between object types of the objects of the one of the identified pairs of objects.

4. The method of claim 1, wherein the clustering is further based on a similarity between past and current motion of the objects of the one of the identified pairs of objects.

5. The method of claim 1, wherein the similarity values are determined further based on a similarity between predicted future motion of the objects of the one of the identified pairs of objects.

6. The method of claim 1, wherein the clustering is further based on a relative location of the objects of the one of the identified pairs of objects to a feature in an environment.

7. The method of claim 6, wherein the feature is a crosswalk.

8. The method of claim 6, wherein the feature is a bicycle lane.

9. The method of claim 1, wherein the clustering is further based on whether one object of each identified pair of objects appears to be following another object of that identified pair of objects.

10. The method of claim 1, wherein the clustering is further based on whether the objects of each identified pair of objects are identified as belonging to a predetermined semantic group.

11. A system for controlling a vehicle in an autonomous driving mode, the system comprising one or more processors configured to:

receive sensor data identifying a plurality of objects;

identify pairs of objects of the plurality of objects;

determine, for each identified pair of objects of the plurality of objects, a similarity value which indicates whether the objects of that identified pair of objects can be responded to by the vehicle as a group;

cluster the objects of one of the identified pairs of objects based on the similarity value; and

control the vehicle in the autonomous driving mode by responding to each object in the cluster in a same way.

12. The system of claim 11, wherein the one or more processors are further configured to cluster the objects of the one of the identified pairs of objects further based on a distance between the objects of the one of the identified pairs of objects.

13. The system of claim 11, wherein the one or more processors are further configured to cluster the objects of the one of the identified pairs of objects based on a similarity between object types of the objects of the one of the identified pairs of objects.

14. The system of claim 11, wherein the one or more processors are further configured to cluster the objects of the one of the identified pairs of objects further based on a similarity between past and current motion of the objects of each identified pair of objects.

15. The system of claim 11, wherein the one or more processors are further configured to cluster the objects of the one of the identified pairs of objects further based on a similarity between predicted future motion of the objects of the one of the identified pairs of objects.

16. The system of claim 11, wherein the one or more processors are further configured to cluster the objects of the one of the identified pairs of objects further based on a relative location of the objects of the one of the identified pairs of objects to a feature in an environment.

17. The system of claim 16, wherein the feature is a crosswalk.

18. The system of claim 11, wherein the one or more processors are further configured to cluster the objects of the one of the identified pairs of objects further based on whether one object of the objects of the one of the identified pairs of objects appears to be following another object of the objects of the one of the identified pairs of objects.

19. The system of claim 11, wherein the one or more processors are further configured to cluster the objects of the one of the identified pairs of objects further based on whether the objects of the one of the identified pair of objects are identified as belonging to a predetermined semantic group.

20. The system of claim 11, further comprising the vehicle.