
Shinko Y. Cheng  Mohan M. Trivedi

Abstract: We examine the feasibility of a multi-camera voxel based occupant posture estimation system. Several new considerations are made to allow this tested human body modeling system to work reliably in the passenger seat of a vehicle, including camera position, segmentation, and body modeling with voxel reconstructions, all from a constrained 4 camera setup. To describe occupant posture, a partial human body model consisting of a head and torso is proposed. The accuracy of the estimation of this model is compared against ground-truth.

Camera Placement
Camera placement is limited due to the confines of the front seat. Four viewpoints have been used here and judging by the results, appear to be adequate for the purpose of extracting the head and torso locations and orientations. Suitable places for these cameras are in the middle of the front dash, the passenger side-view mirror, the driver rear-view mirror, and the top of the passenger window.

Partial Body Modeling and Model Estimation
It has been shown in an indoor environment that a complete articulation of the human head, torso, upper arms, lower arms, upper legs and lower legs from human voxel data is possible. For the car environment, the parts of the human body that can be found most consistently are the head and torso.

Estimating location and orientation consists of

- Template matching a spherical shell kernel against the surface voxels of the voxel reconstruction of the occupant
- Kalman Filter tracks the location of the detected head
- Torso growing procedure, fixing a cylinder with radius and length of 20cm and 30cm anchored on one side at the base of the neck.

An issue with body modeling from voxel data is the melding of limbs into the body. From voxel data alone, at times there is not enough information for even the human observer to decide a particular portion of a volume to be a limb or torso. To partially address the issue, the arms are detected when they are in view, and then tracked back into the volume blob. It can be shown to work well in momentary situations. This is currently active research.

Silhouette Generation
Silhouette generation by image segmentation is critical component in shape-from-silhouette techniques. Lighting in a car is variable and highly volatile, posing a difficult problem for several color-based segmentation algorithms.

To partially address the issues of volatile changes in the color, the system uses a statistical background subtraction technique that assumes that the background pixel values are gaussian distributed and only varies along the "chromaticity line" when a shadow is cast over or lifted from it.

Conclusion: Given the voxel reconstruction of the occupant alone, it is shown to provide fairly reliable location estimates of the head and torso. Experiments demonstrate an estimation accuracy of average 7.55cm from ground-truth for head positions.

Special considerations arise when placing this tested system into the interior of the car with regards to camera placement, camera calibration, voxel reconstruction and body modeling, and most notably image segmentation for silhouette generation. The last issue remains only partially addressed.

For more detailed posture information, cues such as appearance of arms and legs, and depth measurements within the silhouettes or a more elaborate model of the human subject may be required. Despite this, partial body modeling using voxel reconstructions show promising results.

In future work, we intend to incorporate knowledge of the boundaries of occupant in-position, out-of-position and critically-out-of-position regions from the new airbag firing requirements set forth by the NHTSA in the FMVSS 208 towards the design of a classification system.