A Novel Automated Method for 3D Reconstruction using   
RGB/IR Image Sequences from Multi-modal Cameras

In recent years, the use of multi-modal camera rigs consisting of an RGB sensor and an infrared (IR) sensor have become increasingly popular for use in surveillance and robotics applications. The advantages of using multi-modal camera rigs include improved foreground/background segmentation, wider range of lighting conditions under which the system works, and richer information (e.g. visible light and heat signature) for target identification. However, the traditional computer vision method of mapping pairs of images using pixel intensities or image features is often not possible with an RGB/IR image pair.

We introduce a novel automated method for finding correspondences in RGB/IR image pairs from multi-modal cameras, using the optical flow fields. In applications where there is motion between the camera rig and the scene (scanning security camera, camera mounted on a vehicle, cameras on a moving robot, etc.) and where the scene exhibits visual texture, our method finds correspondences between multi-modal image sequences without using intra-camera pixel intensities or features.

From these optical flow based correspondences we align the image pairs and use the resulting alignment to estimate dense depth maps with accuracies similar to, and in some cases, substantially better than, methods that align pairs of RGB images based on image features or pixel intensities.

We test our method on synthetic optical flow fields and on real image sequences taken with a traditional multi-modal binocular stereo RGB/IR camera rig. We demonstrate our method's accuracy by comparing against a precision ground truth.