Air and space-borne imaging is critical for surveillance, reconnaissance, and interpretation efforts. Most imaging is performed using single 2D modalities, such as visible electro-optical, infrared, and synthetic aperture radar (SAR), that tend to provide incomplete information of complex scenes. By simultaneously collecting information with multiple modalities, most individual sensing limitations can be overcome. However, because the selected sensing modalities measure very disparate attributes of the same scene, interpretation of the combined imagery becomes more challenging, as a feature in one modality may map to many or none in another modality.

With funding through the Information, Computation, and Exploitation Line Portfolio, Laboratory staff have developed an approach to address these shortcomings in multimodal 2D imagery interpretation by translating 2D imagery to a combined 3D image that more accurately captures the complex 3D, multi-surface geometry of an underlying scene. The project, Multimodal Vision for 3D Scene Interpretation (MMV3D), expands on prior 2D domain transfer techniques to learn a generalizable 2D to 3D translation by training a deep neural network on co-collected 2D and 3D lidar imagery. The method can then be employed on 2D-only imagery to generate high-quality 3D imagery, with resulting 3D products shown to be more readily interpretable by trained imagery analysts.

The approach has been demonstrated on operational scenes in South America to translate 2D SAR imagery to 3D and more readily identify hidden man-made encampments under dense foliage. The technique was also successfully applied to urban imagery and for recovering 3D geometry of military vehicular targets. The team is adapting this current approach to additional modalities, which may enable significant improvements in multimodal imagery interpretation for a variety of mission areas.