Dr. Thomas M. Strat
Information Exploitation Office (IXO)
Precision Target Identification

The goal of air campaigns has always been to destroy the right target—and only that target. And that has become increasingly important in today's precision strike operations.

Avoidance of unintended effects—so-called collateral damage—is a major driver for the Rules of Engagement that have been adopted in recent campaigns. Fully-effective air power that avoids the risk of collateral damage requires precise identification (ID) of targets, particularly mobile targets.

The need to identify large numbers of targets in very short periods of time places unachievable demands on today's largely manual interpretation of reconnaissance imagery. Instead, automated target identification is sorely needed, and DARPA's Information Exploitation Office (IXO) is working toward that objective.

This leads to two themes in IXO:

- The development and exploitation of data from close-in sensors —sensors that have the resolution and discrimination to enable precision ID.

- And second, the identification of targets based largely on 3-dimensional geometry. Historically, the target recognition community has attempted to recognize targets in visible, infrared, and radar images, which are two-dimensional projections of the three-dimensional world.

For example, the template-based approach to Automatic Target Recognition uses large collections of 2D projections that are matched against 2D images. The model-based approach to ATR was devised to deal more effectively with the problematic combinatorics of pose, depression angle, and articulations.

For example, the DARPA MSTAR program used detailed 3D geometric models to predict the 2D projections, which were then matched against the 2D image. This was very effective in addressing the combinatorics, but insufficient as a means to cope with the wide variability within each target class. In the future it will be desirable to match 3D geometric models against 3D volumetric data—rather than 2D projections—to achieve unambiguous target identification even under the most difficult circumstances.

Several programs are underway or planned in IXO to realize this goal.

Jigsaw, managed by Robert Hauge, is an ongoing project to build a laser radar (or LADAR) sensor to produce high-resolution 3D data. E3D, for exploitation of 3D Data, is a new program managed by Bob Hummel. The goal of E3D is to recognize targets in 3D data such as that produced by Jigsaw. And VIVID (for Video Verification and Identification) is a new concept that we're exploring for target identification in visible and infrared motion imagery.

The objective of Jigsaw is to design and build an imaging laser radar device to produce very high-resolution, 3D, volumetric data of targets. Such a sensor would enable the detection and identification of targets day or night as well as through foliage, camouflage, and even windows. Cued to the existence of a possible target by a foliage-penetrating radar, the Jigsaw sensor will be carried to the vicinity by a small UAV to obtain detailed 3D data in which the target can be identified.

As an illustration of the potential of LADAR for precision ID through foliage, the Jigsaw program recently collected a series of range images from a variety of angles of a scene that included vehicles parked behind a dense tree line at Fort AP Hill. Shown here is a mosaic of color images obtained along with the range data. It is nearly impossible to discern the vehicles by viewing these images.

Here we show the individual range images, color-coded to show the height above the ground plane. Each individual LADAR range image pokes through gaps in the foliage but samples less than 10% of the target so the vehicles cannot be seen in this data either. If we combine the range data from all range images by registering the 3D information to a common coordinate system, the 3D structure of the vehicles becomes readily apparent—as seen here. This same 3D data can be computationally rotated to view the scene from other perspectives. The vehicles then become plainly visible.

The Jigsaw Program has been underway since July of last year. The first phase recently concluded, and two contractor teams have been selected to build prototype LADAR sensors with the resolution and frame-rate to meet design goals for use as sensors for precision identification. Once we have 3D data such as that produced by the Jigsaw sensor, how can we use it? The E3D program, whose goal is to recognize targets automatically within 3D volumetric data, provides one answer.

This new program is just getting underway and represents the first major effort to use 3D data for precision target recognition. It has its foundations in prior work that analyzed range data from LADAR or stereo processes—work that had been focused on the identification of manufactured parts on an assembly line or on the navigation of robotic vehicles.

To enable precise identification of targets, E3D matches 3D geometric models to the 3D data. It employs a "Recognition by Components" approach as a means to deal with the expected variability of observed targets in the real world. By specifically looking for key discriminators, such as a radar dish on an air defense system or the main gun tube on an armored vehicle, we can attempt to classify the vehicle even in the presence of other variations, such as gear stowed on the roof or a fender that is missing. In fact, the recognition of those variations themselves can serve as key identifiers to help with "fingerprinting" individual vehicles.

The direct exploitation of 3D data is also useful for recognizing targets that are partially obscured. The M109 self-propelled howitzer shown here cannot be recognized by conventional ATR techniques that attempt to match a description of the entire vehicle. But 3D information permits detection of the occlusion, and the remainder of the vehicle can be matched to the corresponding parts of the 3D model for successful recognition.

E3D is a brand new program with a Broad Agency Announcement-based competition completed last spring and contracts getting underway this summer.

Now let me turn to an idea that we hope to pursue in the near future. As I mentioned earlier, the avoidance of collateral damage has become a major driver during precision strike planning and execution. Let me show you one example that has now become infamous. The target of the strike in this clip is a railroad bridge over one of the major rivers in Kosovo. The train you see had the misfortune of crossing the bridge at exactly the wrong time. Our precision munition in this case hit exactly where it was intended, but the result was catastrophic unintended collateral damage.

To avoid incidents such as this one, we propose to use video sensors on loitering platforms, such as unmanned aerial vehicles. We call this concept VIVID for Video Verification and Identification. GPS-guided weapons are extremely reliable at hitting the grid coordinate that they have been assigned, but they are blind. VIVID provides the eyes for these weapons.

The VIVID concept employs two complementary strategies: First, it will identify and track the target. As shown here, VIVID matches a three-dimensional model to many video frames. If the proper model is used, as shown on the left, a good match is obtained. If an incorrect model is used, the match is poorer and the targets can be distinguished.

VIVID will employ this model-matching procedure throughout the tracking process in order to keep continuous track on vehicles even in dense urban traffic and frequent occlusions. Once we have identified a target, we should never lose it!

Second, VIVID will use its video sensor to observe the surrounding area for potential collateral damage situations. It will search for noncombatant vehicles or people that are within the splash radius of the weapon and will also enforce protection of structures contained in a no-strike database. VIVID is the only program proposed to look not only at the target but also at the surrounding area.

VIVID is a concept in the formulation stage—we are interested in your ideas for technological solutions to this set of problems. How can we confirm the target identity with sufficiently high reliability? How can we maintain continuous track on moving targets even in dense traffic and frequent occlusions? How can we detect the presence of non-combatants within a potential impact area?

If successful, VIVID will support the engagement of larger numbers of difficult targets while reducing the frequency of so-called "CNN events" caused by the unintended effects of precision strikes.

Let me conclude by contrasting the present practice of airborne reconnaissance with a vision that is not too far off. Today, our warfighters work to determine how to get the image they need. Relying predominantly on scarce sensing resources from standoff platforms, the resulting analysis is image-centric, and the ability to identify a target is inconsistent.

In the future, precision ID will be target-centric. Given the need to identify a potential target, the warfighter will have access to an array of available imagery, including 3D volumetric and video data from relatively low-altitude UAVs and micro air vehicles. Target identification will be unambiguous.

I have described several initiatives within IXO that are contributing to this vision. Jigsaw: a 3D imaging laser-radar to be used for close-in identification of targets. E3D: target identification algorithms for exploiting 3D volumetric data. and VIVID: the identification and tracking of targets and non-targets in visible and infrared motion imagery.

Together with our partners in government and industry we are working to make unambiguous precision identification and collateral damage avoidance a reality. Your ideas on how to do this more accurately and more expediently will help our military remain the dominant fighting force in the world.

Thank you.