As the unmanned reconnaissance aircraft cruising through the skies of Afghanistan have demonstrated, advanced imaging is becoming ever more important to U.S. military operations. Thanks to improvements in both sensor technology and computational abilities, researchers have developed military sensors that capture and process increasing amounts of information. Passive sensors that simultaneously collect images from multiple spectral bands allow processing that exploits differences in the characteristics of those bands. The U.S. Army Research Laboratory (Adelphi, MD), along with several industry and academic partners, is developing dual-band infrared (IR) sensors and processing techniques to use this capability to detect military vehicles. The aim of this technology demonstration system is to build a single unit that combines sensor components with advanced signal processing and target recognition tools.

Dual-band imaging has the advantage that it can better detect targets in clutter. It can distinguish between targets and decoys and has the ability to defeat IR countermeasures such as smoke, camouflage, and flares. In most previous multiband efforts, the pixels of the individual colors either were not spatially registered or the outputs were read out sequentially rather than simultaneously.

The Army Research Laboratory program has produced moderate-format (256 × 256 pixels) staring, dual-band focal plane...
target detection setups
Typically, data fusion involves combining the data from different sources, which can include data from multiple spectral channels of a single sensor, images taken at different times by the same sensor, and data collected by different sensors at the same time. We can perform data fusion at pixel level (measurements or signals), feature level (attributes), decision level (rules), or a mixture of these three levels. In our experiments, we combined the information from the two bands at either the pixel or the feature level and studied the effects of each approach on detection accuracy and computational cost.

We designed a set of similar eigen-neural-based target detection architectures, which differ in the type of their input data. Each of the four target detectors consists of an eigenspace transformation followed by a multilayer perceptron, a commonly used neural network (see figure). Their inputs are image chips (target-sized image portions) extracted either from an individual band or from both the midwave and longwave bands simultaneously. The principal component analysis eigenspace transformation is used to extract features and reduce dimensions on the image chips. The perceptron then processes the transformed data and produces a confidence number that measures the likelihood that the image chip contains a target.

For fusion, we use both dual-band pixel-level architecture (DPA) and dual-band feature-level architecture (DFA). For pixel-level fusion, we combine the midwave and longwave chips into one vector before the principal component analysis' eigenspace transformation, so that all of the pixel values are processed as one entity. After the transformation, the perceptron processes the transformed data.

For feature-level fusion, the midwave and longwave chips are transformed separately, and then the perceptron processes transformed data from both sources. We use both architectures because the relative performance of these architectures is data dependent. In particular, performance depends on the amount of training samples available in relation to the variability and information content of the underlying distributions. We expect that as the amount of training data increases, the performance of the dual-band pixel-level architecture will increase relative to the dual-band feature-level architecture. For the purposes of this study, we were more concerned whether either of these architectures improves performance than with which of these architectures is superior on this data set.

detection performance
Using a set of dual-band images, we examined the target detection capability of the proposed architectures. There were 461
pairs of longwave-midwave matching frames, with 572 legitimate targets at ranges between 1 and 4 km. We chose 231 frames as the training set and the remaining 230 frames as the testing set. From the training frames, we first extracted a number of training image chips that contain the signature of either a legitimate target or a competitive clutter. These training chips were used to train all four architectures to differentiate targets from clutter.

Each trained architecture performed the target detection task by taking data chips that were extracted across the whole input frame and producing a corresponding output response map that was an image of confidence numbers. The global maximum and a predefined number of local maxima of the confidence image were picked as potential detections. High scores within a small neighborhood were combined and represented by the highest-scoring pixel among them.

To score the detections, we defined an adaptive acceptance window as a rectangle centered at the ground-truth location of a given target. The size of the acceptance window was inversely proportional to the ground-truth range of the target, which measured around 40 × 26 pixels at the 2 km range. Any detection that fell within the acceptance window was declared a hit; otherwise we classed it as a false alarm. We adjusted the number of hits and false alarms per frame by setting different cutoff thresholds on the confidence levels of the detections, below which detections were simply discarded. Using these hit/false-alarm pairs, we plotted the corresponding receiver operating characteristic curve for each target detector (see figure 3).

The vertical axis indicates the hit rate, which is the percentage of targets found by the detector. The horizontal axis is the total number of false alarms divided by the number of frames in the data set. For these curves, we focus at the low false-alarm portion of the curve, which is the most critical for many practical applications. The single-band longwave architecture detector clearly outperformed the single-band midwave architecture detector, especially at the region with very low false-alarm rates. The dual-band detectors show better performance than the single-band detectors for the more difficult targets (those near or to the right of the knee of the curve). The dual-band feature-level architecture outperforms the dual-band pixel-level architecture, which suggests that more training data is needed to increase the accuracy of the pixel-level architecture.

Figure 1 A simultaneous dual-band image of an M60 tank acquired with the QWIP FPA shows difference between midwave (left) and longwave (right) infrared detection.

Figure 2 Four target detection architectures tested include single-band midwave architecture (SMA), single-band longwave architecture (SLA), dual-band pixel-level architecture (DPA), and dual-band-feature-level architecture (DBA), which transform input data and process them through a multilayer perceptron (MLP) neural network.

Figure 3 The various architectures show different rates of hits and false alarms in both the training frames (top) and testing set (bottom).

References