Automatic Building Identification using GPS and Machine Learning

Robert Woodley – 21st Century Systems, Inc., robert.woodley@21csi.com
Warren Noll – 21st Century Systems, Inc., warren.noll@21csi.com
Joseph Barker – 21st Century Systems, Inc., joseph.barker@21csi.com
Donald C. Wunsch II – Missouri University of Science and Technology, dwunsch@mst.edu

ABSTRACT

Video sensor capabilities and sophistication has improved to the point that they are being utilized in vast and diverse applications. Many such applications are now on the verge of providing too much video information reducing the ability to review, categorize, and process the immense amounts of video. Advancement in other technology areas such as Global Positioning System (GPS) processors and single board computers have paved the way for a new development of smart video sensors. A need exists to be able to identify stationary objects, such as buildings, and register their location back to the GIS database. Furthermore, transmitting large image streams from remote locations would quickly use available band width (BW) precipitating the need for processing to occur at the sensor location. This paper addresses the problem of automatic target recognition. Utilizing an Adaptive Resonance Theory approach to cluster templates of target buildings processing and memory requirements can be significantly reduced allowing for processing at the sensor. The results show that the network successfully classifies targets and their location in a virtual test bed environment eventually leading to autonomous and passive information processing.

1. Introduction

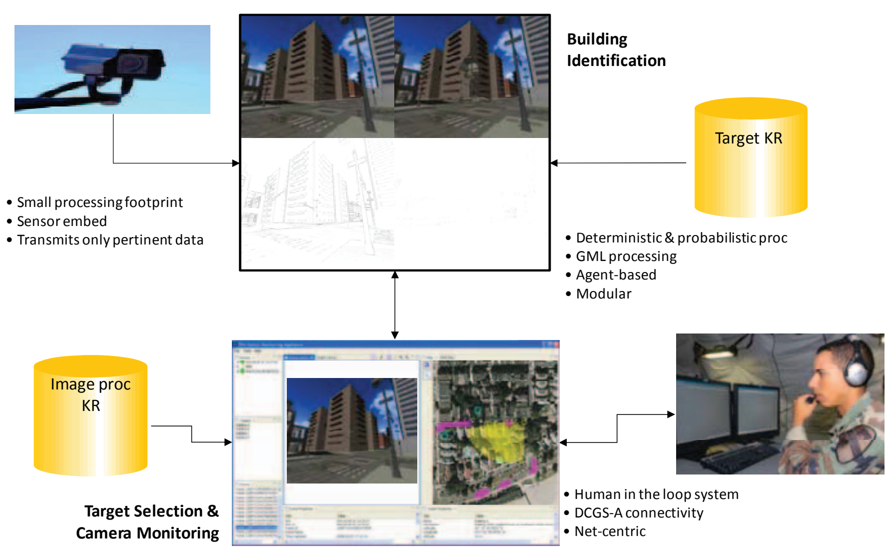

Figure 1: Target Recognition Assistant (TRA).

To address the problem of automatic video processing for stationary object classification, 21^{st} Century Systems, Inc.® (21CSI®) has developed a concept known as the Target Recognition Assistant (TRA). As depicted in Figure 1, TRA has the capabilities of processing video in real-time to collect intelligence on stationary targets and communicate with the operator for proper accounting of battle damage assessment. From the outset we made the decision to use commercial off the shelf (COTS) hardware for this project. We primarily made this decision to keep costs down and secondly to not reinvent wheels that are already available. There are three main components of the device: GPS unit, digital compass, and processing unit. In addition to these, we made the device connect to video systems using standard Ethernet connectivity. We also designed the system such that the device may be monitored from any standard mobile device with wi-fi connectivity.

What makes TRA different than any other geo-registration system [1,2] is the use of a machine learning capability designed to not only identify objects, but to identify the objects under adverse conditions. In particular, we are looking at bomb damage where portions of the structure may be destroyed, yet we still wish to identify the target. TRA takes a series of images and compares these against known targets for a given location. The images are classified by an ART network clustering algorithm.

2. TRA Image Comparison Using Machine Intelligence

A computational intelligence architecture was designed to distinguish targets from non-targets. The project's ultimate goal is to distinguish damaged from non-damaged structures, and to classify the extent of damage to structures that have been damaged. The architecture chosen is based on the Adaptive Resonance Theory (ART) of neural network learning. This approach was chosen due to its scalability to large input patterns, fast run time (often O(n)), existence of learning stability theorems, large body of related applications literature, applicability to both supervised and unsupervised learning approaches. We chose Adaptive Resonance Theory [3,4]. This approach was chosen due to:

- Scalability to large input patterns [5,6]
- Fast run time, often O(n) [6,7]
- Existence of learning stability theorems [4]
- Large body of related applications literature [5, 7,8]
- Applicability to both supervised and unsupervised learning approaches. [9]

We begin by computing possible targets given the GPS location and orientation of the camera. Figure 2 data can be trained heteroassociatively with Figure 3 data to increase confidence in building identification. By performing feature extraction from both Figures 2 and 3 to isolate the subregion(s) of interest, we develop the training information. Enhanced training techniques round out the training [9,10].

Figure 2: Undamaged buildings data, including range and orientation data.

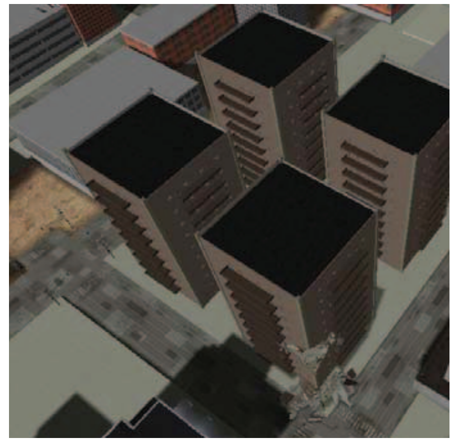

Figure 3: Data analogous to the undamaged building except after damage.

3. Results

For use in the final system, the ARTMAP is trained from imagery generated by a 3D simulation of the target and its surroundings. Images are selected from 360 degrees around the target (at multiple distances) in 1 degree increments. Any images where the target is occluded are discarded. The training is done offline and the resulting template set uploaded to the embedded system on demand.

Figure 4: Sample input image.

Figure 5: Template building match (note the superimposed damage information)..

When integrated into the embedded system, the ARTMAP classification sub-system aids in the selection of quality intelligence imagery for storage and upload. Due to the highly integrated nature of the system, it is difficult to evaluate the ARTMAP performance in isolation, but in testing the combined system achieved over 90% accuracy in selecting good imagery with the primary failure mode being occluded targets. Examples of the processing steps are shown in Figure 4 and 5. Performance is another key factor that was required since the embedded system has a PDA-level processor. Even processing over 50 templates per frame, the ARTMAP is able to achieve more than 1 frame per second.

4. Conclusion

The algorithm development was focused on vision analytic approaches to identifying targets. The building identification algorithm uses ART to create template examples of the targets. If the video image matches the template the algorithm returns a match indication. The ART algorithm is very light weight and can be run on a Gumstix single board computer (SBC). The ART algorithm also provides primary occlusion control by recognizing that the object in the image is not the target. The ART network will return a no-match if the structure is occluded.

The experiments described herein demonstrate feasibility of Adaptive Resonance Theory for target/non-target classification from urban battleground imagery. It is expected that a similar procedure will work for more difficult classification tasks such as assessing the degree of battle damage, need for more data, fusion of multiple data sources, and other tasks. These opportunities deserve further investigation. The generalization ability of this approach also needs to be tested when larger and more diverse datasets become available.

5. References

[1] Lisa Gottesfeld Brown, "A Survey of Image Registration Techniques," ACM Computing Surveys (CSUR), Vol. 24, Issue 4, pp. 325 – 376, Dec. 1992.

[2] Barbara Zitova and Jan Flusser, "Image Registration Methods: A Survey," Image and Vision Computing, Vol. 21, Issue 11, pp. 977 – 1000, Oct. 2003.

[3] Gail Carpenter and Stephen Grossberg, "Adaptive Resonance Theory," in The Handbook of Brain Theory and Neural Networks, 2nd Edition, Michael Arbib, (Ed.), MIT Press, pp 87-90. Available in preprint form online: http://cns.bu.edu/Profiles/Grossberg/CarGro2003HBTNN2.pdf, 2003.

[4] Carpenter, G.A. and Grossberg, S., "A massively parallel architecture for a self-organizing neural pattern recognition machine," Computer Vision, Graphics, and Image Processing, 37:54-115, 1987.

[5] Caudell, T.P., Smith, S.D.G., Escobedo, R., and Anderson, M., "NIRS: Large scale ART–1neural architectures for engineering design retrieval," Neural Networks, 7:1339-1350, 1994.

[6] Thomas Caudell, Scott Smith, Craig Johnson, and Richard Escobedo, "A data compressed ART1 neural network algorithm," Proc. SPIE Conference on Aerospace Sensing, April 1990.

[7] Sam Mulder and Donald C. Wunsch II, "Million City Traveling Salesman Problem Solution by Divide and Conquer Clustering With Adaptive Resonance Neural Networks," Neural Networks, July 2003.

[8] Rui Xu and Donald C. Wunsch II, "Clustering," IEEE Press / Wiley, Oct. 2008.

[9] Carpenter, G.A., "Distributed learning, recognition, and prediction by ART and ARTMAP neural networks," Neural Networks, 10:1473-1494, 1997.

[10] Carpenter, G.A., Grossberg, S., and Rosen, D.B., "Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system," Neural Networks, 4:759-771, 1991.

[11] G.A. Carpenter, S. Grossberg, N. Markuzon, J.H. Reynolds, and D.B. Rosen, "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps," IEEE Trans. Neural Networks, vol. 3, pp. 698-713, Sep. 1992.