A Hybrid Vision System for Object Recognition and Application in a Systems Integration Sandpit

Can Huynh, Quoc Do and Lakhmi Jain

Presenter: Dr Quoc Do
Presentation Overview

- Introduction to the Systems Integration Sandpit
- System evolution and systems upgrade
  - Hybrid Vision System
    - Memory Feedback Modulation
    - SUF Feature Description
    - Integrated/hybrid system
- Results
- Future work and summary
Introduction to the Systems Integration Sandpit

- Originally established and named Microcosm in 2006.
- Renamed as the Systems Integration Sandpit in 2009.
Introduction to the Systems Integration Sandpit

The aims are to provide:

- DASI, DSTO and DMO with an environment where SE&SI teaching and research can be carried out.
- A focus for training staff in SE&SI practices.
- A “Sandpit” which can be used to develop and evaluate system configurations and operation, stage demonstrations, and conduct experiments.
- An environment where tangible systems can demonstrate the challenges of SE&SI to Australian defence organisations.
- An environment that can develop student appreciation of system interfaces, in particular those used in real-time networked software-intensive-systems.
Systems Integration Sandpit (SIS)
Education and Training

Defence and Systems Institute
Systems Integration Sandpit (SIS)
Operational Scenario - OV1
Systems Integration Sandpit (SIS)
Stage One and Stage Two

- **Stage One** demonstration was conducted at UniSA, with the presence of DSTO CDS and CSIO in December/January 2009.
- **Stage Two** demonstration was conducted at Defence and Industry Conference at Adelaide Convention Centre, June 2009.
- **Stage Three (part one)** will be demonstrated in June 2010.
Systems Integration Sandpit

- The SIS adopts the Spiral development model with five phases:
  - Requirements phase
  - Design phase
  - Build phase
  - Test and Evaluation phase
  - Technology Insertion phase
Vision capability for object recognition was identified as a potential technology insertion.

This led to the development of the Hybrid Vision System for object detection and recognition.

The aims were:

- to demonstrate the integration of “best-of-breed” pre-existing vision components to create the new system, rather than developing the system from scratch.
- to illustrate of how detailed design is driven by top-down systems engineering approach.
SIS Capability Gap

- The physical decomposition of the mobile platform for the SIS-project is shown below.
- Top-down systems engineering has mapped the object detection and recognition functionality onto the physical component “Vision System”.

The hardware architecture of the Robot One sub-system
Hybrid Vision System

- The hybrid vision system integrates concepts from was derived from:
  - Human vision system
  - Memory feedback modulation
  - Neural networks

Human Vision System

- Object recognition in the human vision system is believed to be performed in two distinctive stages:
  - Pre-attentive stage is a selective processing mechanism that selectively focuses on prominent/significant areas or regions of interest.
  - Attentive stage is a much slower and computationally intensive process that analyses and recognises familiar objects.
Hybrid Vision System
Memory Feedback Modulation

- Biological Neural Networks
  - Adaptive Resonance Theory (ART)
  - Selective Attention Adaptive Resonance Theory (SAART) [1].
    - Able to recognise an object in a cluttered scene.
    - This neural network is a dynamic system.
    - Computationally intensive
    - Not suitable for real-time applications.

(Lozo, 1995)
Hybrid Vision System
Memory Feedback Modulation [1]

(Do, 2010)
Memory feedback enhancement of early stages

One possible implementation of the enhancement is achieved by applying the below formula and then apply lateral competition to achieve object-background separation.

\[ E(x(i, j)) = S(i, j) - G(S(i, j) * F(i, j)) \]
**SIFT (Scale-Invariant Feature Transform) descriptor:**

- Proposed by David Lowe in 1999. This algorithm detects local image feature and extracts a set of descriptors suitable for image matching and object recognition.

- The SIFT descriptors are invariant to scale, orientation, affine distortion and partially invariant to illumination changes (Lowe 1999).

- Mikolajczyk conducted a comparison study in 2005, proven that SIFT is the most successful algorithm for extracting local descriptors.
Inspiring from this approach many other SIFT-like algorithms have been developed. For instance:

- **SURF** (Speeded Up Robust Features) by Bay et al. in 2006,
- **PCA-SIFT** (Principal Components Analysis-SIFT) by Ke, and Sukthanka in 2004,
- **GLOH** (Gradient Location-Orientation Histogram) by Mikolajczyk and Schmid in 2005,
- And the most recent **ASIFT** (Affine-SIFT) by Morel and Yu in 2009.
The hybrid vision system combines:

- Memory feedback modulation concept.
- Current state-of-the-art SIFT and SURF feature descriptors and their descendents to achieve object recognition at multiple scales and viewpoints.
Preliminary Results

- Recognition in different backgrounds

Recognition of object with varying in scale and rotation
Preliminary Results

- Recognition in cluttered backgrounds

Object recognition under varying levels of occlusion
Summary

- This work based on the concept of “systems engineering in the small as a precursor to systems engineering in the large”

- The project demonstrated:
  - the integration of “best of breed” pre-existing vision components to create the new system, rather than developing the system from scratch.
  - to illustrate of how detailed design is driven by top-down systems engineering approach.

- Future work:
  - Extend the integrated vision system to achieve real-time object detection and recognition.
  - Integrate the vision system into the SIS robotic platform.
References