

TECHNICAL REPORT

**AMOR: An Adaptive, Multimodal
Architecture for Visual Object
Recognition
(01-01-10-SF0176)**

Dr James Mountstephens

25th March 2014

PERPUSTAKAAN
UNIVERSITI MALAYSIA SABAH



UMS
UNIVERSITI MALAYSIA SABAH

SYNOPSIS

The general objective of this research was to develop a novel architecture for the difficult, but crucial, problem of recognising objects in visual scenes. An Adaptive, Multimodal architecture for Object Recognition (AMOR) was developed that extends existing unimodal (visual-only) systems in order to increase their accuracy. The specific objectives of this research were to design, implement and evaluate the proposed architecture. The architecture design was formalised mathematically and algorithmically in the form of "reclassification", implemented in custom Matlab code and evaluated by comparison to existing visual-only methods for object recognition. A scene dataset consisting of 8,750 visual scenes containing 575 object classes was developed and used for testing. Experiments demonstrated that when using a combination of SIFT visual features and an SVM classifier, AMOR could achieve an average of 30.5% accuracy in object recognition, which was a 5.5% improvement over a standard visual-only approach. Also developed was a novel measure of 'compatibility' between visual confusion and object co-occurrence that attempts to quantify the extent to which context can compensate for visual confusion. A reasonable level of correlation was found between compatibility and adaptive multimodal improvement in performance.

Objektif umum kajian ini adalah untuk membangunkan seni bina baru untuk masalah yang sukar, tetapi penting, mengiktiraf objek dalam adegan visual. Seni bina Adaptive, Multimodal untuk Objek Pengiktirafan (AMOR) telah dibangunkan yang merangkumi unimodal (visual sahaja) sistem yang sedia ada untuk meningkatkan ketepatan mereka. Objektif khusus kajian ini adalah untuk mereka bentuk, melaksana dan menilai seni bina yang dicadangkan. Reka bentuk seni bina telah dirasmikan secara matematik dan algorithmically dalam bentuk "Pengelasan", dilaksanakan dalam kod Matlab adat dan dinilai oleh perbandingan dengan kaedah visual sahaja yang sedia ada bagi pengiktirafan objek. A dataset tempat kejadian yang terdiri daripada 8,750 adegan visual yang mengandungi 575 kelas objek telah dibangunkan dan digunakan untuk ujian. Eksperimen menunjukkan bahawa apabila menggunakan gabungan menapis ciri-ciri visual dan pengelas SVM, AMOR boleh mencapai purata ketepatan 30.5% pada pengiktirafan objek, yang merupakan peningkatan 5.5 % berbanding pendekatan visual sahaja standard. Juga dibangunkan adalah langkah novel daripada 'keserasian' antara kekeliruan visual dan objek bersama-kejadian yang cuba untuk menyatakan kuantiti sejauh mana konteks boleh mengimbangi kekeliruan visual. A yang berpatutan korelasi didapati antara keserasian dan peningkatan pelbagai mod penyesuaian dalam prestasi.

