Self Adaptive Approach to Face Recognition

Paul Robertson & Robert Laddaga
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: Our problem is to build a face recognition system which is robust against unforeseen changes in pose and lighting. There are now many face recognition systems that work extremely well in constrained situations, but in natural environments, where lighting and pose can vary widely, they perform poorly.

Motivation: Image understanding programs in general have tended to be very brittle and perform poorly in situations where the environment cannot be carefully constrained. On the other hand, natural vision systems in humans and other animals are remarkably robust, and robust vision is essential for many applications. Natural environments contain rich variety. It is very hard to build a single algorithm that can deal with the natural range of possibilities, but two aspects of the world provide a means of meeting the challenge: 1) in any particular situation one only has to deal with a subset of that variety; and 2) the subsets are not random but tend to occur in clusters or contexts.

Approach: Our approach to robust vision applications is to use a self adaptive software architecture [7, 2], combined with a large and diverse collection of vision operators. This project aims to produce a robust face identification program based on the self adaptive (GRAVA) architecture [5]. GRAVA has already been successfully applied to satellite image interpretation [3, 4] and is now being used to identify faces in video images with unconstrained pose and lighting [6]. The face identification system has many useful applications that include intelligent spaces and security.

The GRAVA architecture is aware of the context in which it is operating, as determined by the images being processed. GRAVA supports diversity by constructing a network of agents that can interpret the current context, and by rebuilding the network of agents as the context changes. In this way, the running system of agents is always well matched with the image being interpreted.

In addition to knowledge of faces that the agents themselves use to recognize faces, GRAVA brings to bear two other kinds of knowledge. One of these is knowledge of the contexts that can be presented by the environment (contextual awareness), and the other is self-awareness, or meta-knowledge about the state of the program and the agents. Self awareness provides the knowledge necessary for the program to make changes to itself. Contextual awareness provides the knowledge necessary to know when it is necessary to change and what it is necessary to change to.

Most face identification and recognition systems work by measuring a small number of facial features given a canonical pose and matching them against a database of known faces. By building a face recognizer that can seamlessly switch between different contexts such as pose and lighting we can construct a recognizer that is robust to normal changes in the natural environment. This permits a much wider application of face recognition technology.

Our application involves recognizing people as they move about an intelligent space [1] in an unconstrained way. To better understand contexts consider the face “pose” contexts:

Figure 1 shows four pose contexts: “profile”, “oblique”, “off-center”, and “frontal”. The profile view is supported by agents that measure points along the profile of the face, the corner of the eye, and the lips. The oblique view with ear supports measurements of the ear and measurements of the position of the ear, eye, and nose. The triangle formed by the eye, ear, and nose help to determine the angle of the face to the camera which allows measurements to be normalized before recognition. similarly, the off-center and frontal views each have their own measurement contexts.

The recognizer supports a collection of face candidate finders, face models, feature finders, and normalization...
algorithms implemented as agents. The face recognition process is shown schematically in Figure 2. Face candidate finder agents look for face like shapes in the image and generate evidence that supports the selection of a set of contexts based on the shape and shading of the face candidate. Agents appropriate to the context are selected to make a special purpose face recognizer. If the recognizer doesn’t succeed in finding appropriate features where they are expected to be the system self-adapts by using available evidence to select a more appropriate context, constructing a new recognizer, and trying again. Convergence on the right set of contexts is rapid because evidence in support of a context is collected each time an agent runs.

What is unique about the recognizer outlined above is that it has multiple ways of recognizing faces, it divides up a complex space of lighting, age, race, sex, and pose into contexts that can be composed in a huge number of ways and self-adapts the recognizer at runtime.

**Impact:** The applications for robust vision are myriad. Robust vision is essential for many applications such as mobile robots, where the environment changes continually as the robots move, and robustness is essential for safe and reliable operation of the robot.

**Future Work:** Little is known about the stability of systems built this way and there are as yet no known guidelines for building systems that are intended to be stable. More work remains to be done in this area.

**Research Support:** This research is supported by the DARPA Mobies program, under contract F33615-00-C-1702, and by MIT Project Oxygen.

**References:**


