Particle Swarm Optimization & Facial Emotion

H. Desa, R. Nagarajan and B. M. Ghandi

School of Mechatronic Engineering, Universiti Malaysia Perlis UniMAP

E-mail : hazry@unimap.edu.my

INTRODUCTION

There are several applications that can be derived from efforts in improving human-computer interaction so that computers can have the intelligence to perceive the emotional state of a human user and react accordingly. Such as an intelligent welfare robot to provide support and comfort to bed-ridden / highly disabled people. This is important given the modern life style of declining children population and busier middle-age which leaves the senior citizens and the disabled to fend for themselves.

Particle Swarm Optimization (PSO) is an algorithm that has been found to be very efficient and effective in solving a variety of optimization or searching problems. PSO is a population-based search algorithm that was first developed to simulate the social behavior of birds as they fly in a group searching for food. It has since been used for the classical travelling salesman problem, electrical power systems, neural networks training, image clustering, data clustering, gene clustering, underwater acoustics, task assignment and combinational logic circuits design. However a modified PSO, renamed as Guided Particle Swarm Optimization (GPSO), is here successfully applied in detecting facial emotions.

EMOTION

Six basic emotions are identified as universal and independent of cultural background, both in terms of how they are expressed and how they are perceived. These are happiness, anger, sadness, surprise, disgust and fear. One approach to facial expressions classification is to recognize the underlying facial muscle activities and interpret these in terms of categories such as emotions, attitudes or moods. Facial Action Coding System (FACS), the most commonly used system developed for human observers to describe facial activity, allots Action Units (AU) to visually observable facial muscle actions. FACS allows unique decomposition of facial expression to 44 AUs.

The methodology used is based on studying the underlying AUs involved in expressing the different types of emotions. Then the specific AUs movements can be observed using luminous markers, placed on the subject's face. A video record of the subject is taken as the different types of emotions are expressed. Fig. 1 shows some digital stills extracted from such a video. The aim here is to identify the expressed emotion for each frame in the video by simply observing the changes in the positions of the AUs. Upon obtaining the video clip, the first step is to digitize the clip and obtain the AU positions in 2D over time.

The second step is a training session for a particular subject, where the program is manually taught the approximate positions of the AUs for each of the emotions we wish to detect. Finally, the program, a direct implementation of GPSO, is executed for the full length video clip where a visual display of the emotion expressed at each frame is obtained continuously.

PSO and GPSO

A. Particle Swarm Optimization (PSO)

A PSO algorithm maintains a swarm of particles, where each particle represents a potential solution. Particles are "flown" through a multi-dimensional search space, where the position of a particle is adjusted according to two factors:

- Its own successful experience
- The successful experiences of its neighbours.

The velocity vector drives the optimization process, and reflects both the experiences of the particle and that of its neighbours. The experiential knowledge of the Particle is referred to as the cognitive component, and is proportional to the distance of the particle from its own best position. The socially exchanged information is referred to as the social component of the velocity equation and is represented by two PSO algorithms, gbest and lbest,. The social networking employed by gbest PSO reflects a star topology, where the social component of the velocity reflects the information obtained from the entire swarm. The lbest, is similar to gbest, except that it uses a ring social network topology, where smaller neighborhoods are defined for each particle. The social component reflects information exchanged within the neighborhood of the particle. The two versions of PSO algorithms are similar because the social component of the velocity updates causes both to move towards the global best, with two main differences:

- the larger particle interconnectivity in gbest converges faster but at the cost of less diversity.
- larger diversity in lbest, results in more coverage of the search space, and less prone to being trapped in local minima.

B. *Guided Particle Swarm Optimization (GPSO)*
Emotion detection is a search problem, to identify which of the possible emotion does the current facial expression represent. In order to apply PSO for emotion detection, definable parameters of the algorithm are:

- search space and its dimension
- representing a particle in the emotion-detection setting
- representing position and velocity of a particle
- objective function to be minimized by the PSO

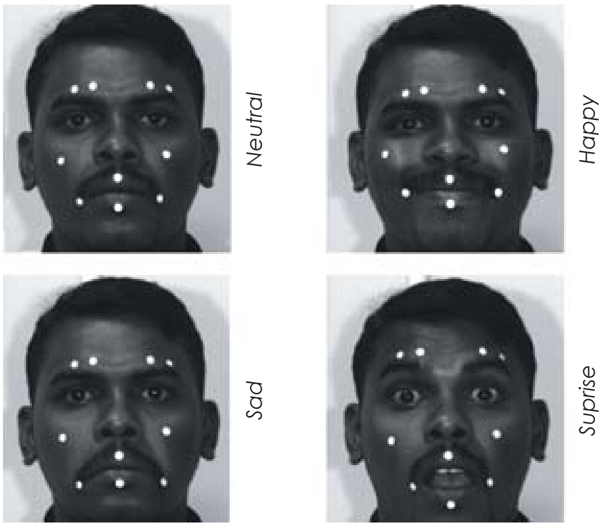

Fig. 1 : Positions of AUs in different emotions

The approach to emotion detection, is to monitor the changes in the positions of the action units, placed on the face of a subject over a period of time, from which we can then determine the emotion expressed at each point in time.

***Definition:* Search space and its dimension:**
If Action Units (AUs) are denoted by, q1, q2, …, qn then D1, D2, …, Dn represents the domains of the AUs, where Dj represents a 2-dimensional rectangular neighborhood window consisting of the possible points that qj can be assigned to. Then the search space is a n-tuple, Rn where the dimension of the search space is n, which represents the number of action units being observed.

***Definition:* Particle position and velocity:**
A particle is an abstract object in the Rn search space with a position and a velocity and represents a possible solution. In emotion detection there are a number of possible emotions which can be encountered at any time. In order to solve a multi-target problem, multiple swarms, one for each possible emotion is used. Since each swarm has a different target to reach, the objective function of each swarm, the Euclidean distance between its current position and its target, is defined differently.

***Definition:* Objective function for a swarm:**
In each iteration of the PSO algorithm, each swarm will update the positions of its particles as usual. These positions are then compared to find the swarm that is closest to its target, which is considered to have found a solution.

This gives data about the positions of the action units and if the particles can take advantage of this knowledge, then they are likely to reach their target sooner than if they rely solely on their cognitive and experiential knowledge. Changing the algorithm such that the position of the AUs should always be represented as one of the particles in each swarm. Incorporating these changes, the particles are effectively guided to converge towards the path of the action units. hence the term Guided Particle Swarm Optimization (GPSO).

EXPERIMENTAL RESULTS

The GPSO algorithm, implemented in C# language under the .NET framework, has two modes: learning and detection. In the learning mode, a video clip is run to capture the approximate positions of the AUs corresponding to each of the basic emotions. Once a particular emotion is observed, the video is paused and the identified positions of the AUs are saved into a file as the coordinate values for the particular emotion. The learning session is over as soon as the data for each of the relevant emotions is obtained. In detection mode, the system will take as input a video clip, the digitized data for the video clip and the positions of the AUs corresponding to the various emotions as captured in the training session. The system initializes a swarm by creating random particles within the domain of each of the AUs. The GPSO algorithm is then executed to detect the emotions expressed in each frame of the video clip and the detected emotion is visually displayed on the screen.

The algorithm modifications introduced, enabled particles to converge very quickly towards the AUs and identify the emotion being expressed. Three of the six universal basic emotions, namely happy, surprise and sad are considered. These three, plus the neutral state gives four possible states that the GPSO system can detect presently. The system is tested with 6 different video clips of different subjects with each video clip about 30 seconds long or of 200 frames, displayed at about 7 frames per second. In order to test the performance of the detection algorithm, the system is made to pause at each frame, where a manual identification is also made and for each frame, the emotion detected by the system as well as manually is recorded in a file. The average success rates - where the auto-detection and the manual detection coincide - recorded after taking the data 10 times for each video clip, is quite good, ranging from 85% to 95%.

Conclusions

A modification of the Particle Swam Optimization (PSO) algorithm for the purpose of emotion detection has been shown, appropriately named , Guided Particle Swarm Optimization (GPSO). It has been implemented and tested, with promising success rates. The algorithm is very efficient in terms of the speed with which particles converge to identify the emotion being expressed in each video frame. This is in part due to the concurrent nature of the PSO algorithm where multiple particles are involved in searching different portions of the search space in parallel, thus increasing the chances of finding a solution sooner. Another equally important factor contributing to the efficiency of the GPSO algorithm is the fact that it made particles to be guided by the actual positions of the AUs as the video clip is played.

This article is a modified version of the original which will be presented in 2009 IEEE Symposium on Industrial Electronics & Applications (ISIEA 2009)", Kuala Lumpur, 4th-6th October 2009