Jay Hatcher
Article 2

Dynamically Adaptive Tracking of Gestures and Facial Expressions

Summary:

The goal of this particular DDDAS application is the interpretation of behavioral indicators in human beings using video analysis. To achieve this goal a variety of techniques are used together to analyze the video in real time. Both 2D and 3D techniques are used to track motion of the face, hands, and shoulders. Two cameras are used simultaneously, one for the upper body, and one for the face.

The motions of the head and face are tracked and analyzed using a combination of color matching, Kalman tracking, a KLT Tracker, Active Shape Models using Principle Component Analysis to detect shape and local profile variations, and a 3D model parameterized to adjust to features detected by the other methods. These methods reinforce and correct each other so that tracking is maintained across a variety of motions of varying speed. Color matching locates the face, and Kalman tracking helps to track the head as it moves while a KLT Tracker combined with the ASM models identify facial features as they move, with the 3D model assisting as the face changes orientation.

The same color matching and Kalman tracking methods are employed to find and track the motion of the hands. Hands are distinguished from arms using edge detection with the Canny method, as the edge density of the hands is much greater than that of the arms. Shoulders are also located using edge detection by detecting all the edges to the left and right of the head and filtering out undesired edges. The position of the hands helps to provide a bound on where the shoulders could be. Using a block-based approach, edges thought to be near the shoulders are filtered by estimating the texture at all blocks centered on the pixels normal to the edges. This information is used to produce a texture function along each edge. These functions are then analyzed to determine if the edges correspond to a shoulder.

Images demonstrating the system in action show that the head, shoulders, and hands are successfully located and that the face is correctly identified, including moving facial features such as the eyes and mouth. Various facial expressions and gestures are demonstrated, all of which are correctly detected according to the overlaid images.

Application of DDDAS:

The methods used to do tracking of various features are where DDDAS begins to show up in this system. Different methods are used at different times depending on the current need. Initial locations are determined with color matching and Kalman filtering on color gradients to better isolate the face and hands. Once these are located, other techniques are employed to keep the analysis focused on the correct regions as the subject moves around and changes expression. The way these techniques are applied together dynamically is based on the current information known, so as each technique adds to the data available, other techniques can take advantage of this information and provide more useful information. The symbiotic relationship between techniques employed and the image data informing the model of the persons facial and body expressions fits the kind of symbiosis that occurs in DDDAS. The other dynamic relationship is the interaction of the two cameras with the model, as each camera's image data informs the model and helps it to make a better analysis of the other camera's image. For example, head and face analysis (camera 1) help to locate both the face and hands (camera 2) by color matching, and the location of the hands (camera 2) help to locate the shoulders (camera 1) by providing probable places to look for edges between the head and the hands.

Ways to Improve DDDAS Application:

Some of the techniques employed could possibly be refined if the system analyzes its past subjects in addition to using its supplied data for the color matching algorithm and 3D model. Data from an ever growing sample set as more subjects are monitored could be used to dynamically refine the color matching algorithm and 3D model representation. However, since the current data seems to work well in practice this may be unnecessary. Another area in which DDDAS might be useful is in providing some mechanism for directly gaining feedback regarding false positives for a given behavior while continuing to analyze subjects. If the system was capable of initiating a more careful analysis of recorded video than its real-time methods in order to determine what conditions produced false positives, this new information could be used to dynamically improve the real-time analysis. Something like this is perhaps already implemented, as the article does not go into detail about the nature of the behavioral recognition system.