SCAVAS
Smart CArd Visual Authentication System

Benoît DUCRAY and Graham HILI
Supervisor : Prof. Keith MAYES

ABSTRACT

The explosive growth of the "Consumer Electronics" market as well as of the social networks, along with more natural human-machine interfaces, raises the exchange and storage needs of an impressive amount of information which nature is increasingly varied.
Today, anyone can access in a few clicks to the diversity of resources that makes up this vastness, and make the profit he is capable of according to his own interest.
This context of quasi-chaotic fertility, generates a powerful need to efficiently protect the access to certain resources, reserving them for those who are authorized.

PROBLEM

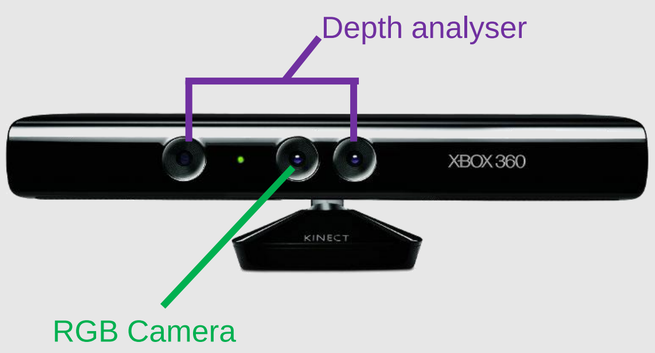
The fact that the objects owned as well as shared secrets are regularly lost, forgotten, stolen or entrusted to others, makes this authentication system no longer considered as reliable nor truly secure.

SOLUTION

We have to make a user authentication system combining both, a high level of safety and a user-friendly interface. The integration in multifactor authentication systems of image-processing-based functions for identifying individuals typically meets this type of requirement.

The Kinect device

- The Kinect uses software created by Rare, a subsidiary of Microsoft Game Studios, and camera technology which understands gestures, from PrimSense. The movements of individuals and objects can be tracked in three dimensions by using an infrared projector and camera fitted with a special microchip. This is known as Light Coding, and uses a variant of image-based 3D reconstruction to provide a 3D scanner system.

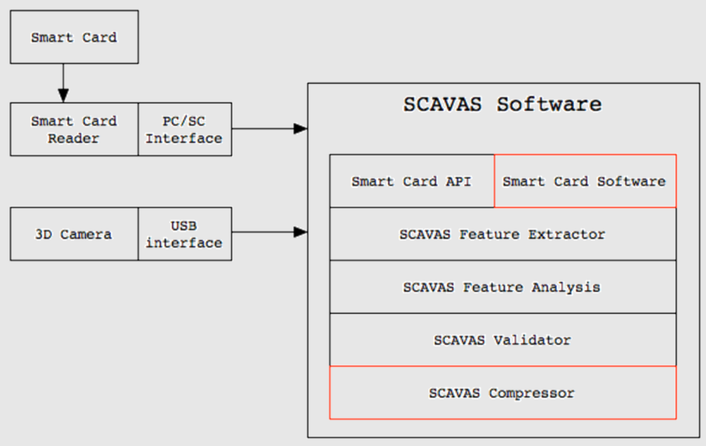


- The depth sensor has an infrared laser projector along with a monochrome CMOS sensor, which enables video data to be captured in 3D under any ambient light conditions. The monochrome depth sensing video stream is in VGA resolution (640 × 480 pixels) with 11 bit depth.

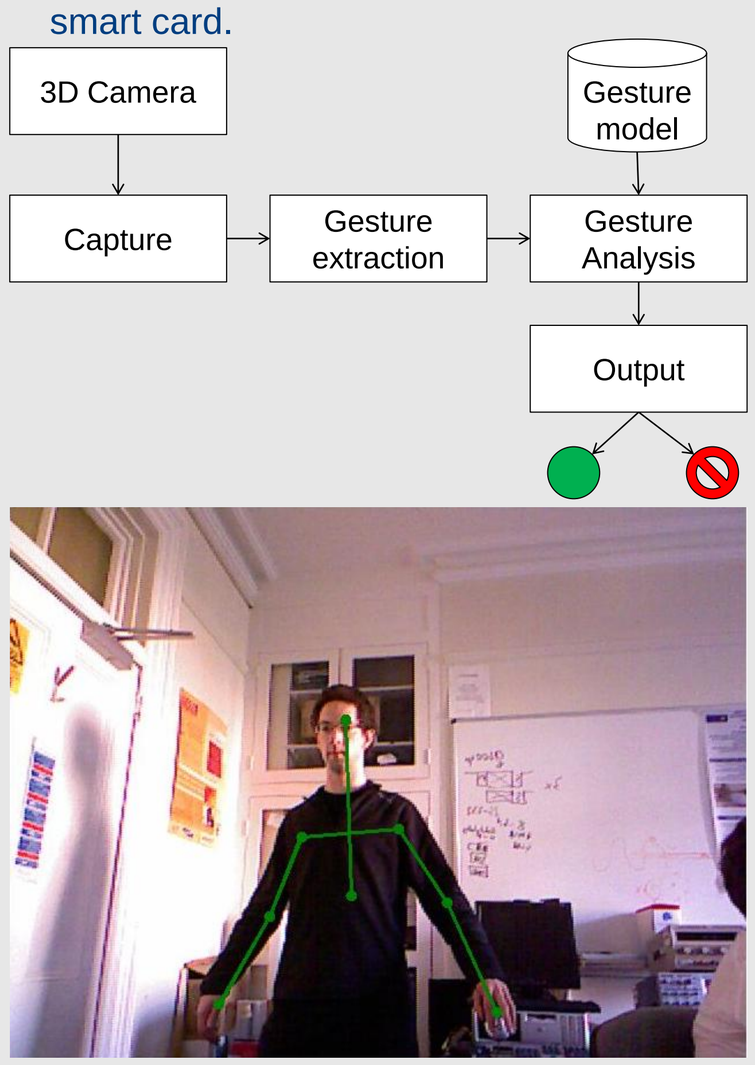
- It has been established that the Kinect sensor outputs video at a frame rate of 30 Hz., and the RGB video stream uses 8-bit VGA resolution (640 × 480 pixels) with a Bayer colour filter.

- The Kinect device is able to track the human body, using depth images it creates as input to its algorithms. Depth images are treated per-pixel to get a segmentation of body parts. The evaluation of each pixel separately avoids a combinatorial search over the different body joints.

SCAVAS



- The architecture of SCAVAS is composed of 3 elements which we can see on the diagram above
- The smart card stores the user's information like name, gesture models...
- The 3D camera, here a Kinect, allows us to extract the user's gesture.
- The software analyses the gesture and compares it to the model stored on the smart card.



- To extract and analyse a gesture, we use 7 body points:
  - the torso
  - both hands
  - elbows
  - shoulders

Dynamic Time Warping

- The dynamic time warping algorithm is used to find an optimal similarity between two sequences, independently of the variation of time or speed between the both sequences. Originally, this algorithm was used in speech recognition and subsequently it has been applied to video, audio, and graphics, because all data we can transform in a linear representation can be analyzed by the dynamic time warping algorithm.

- To use dynamic time warping, we need a cost function or distance function, one of which will give us the error or the distance between our model point and the user's point. It can be calculated by a Euclidian distance.

- The definition of dynamic time warping between two sequences X := (x1, x2, ..., xN) of length N ∈ N and Y := (y1, y2, ..., yM)of length M ∈ N is:

$$\Phi(m; n) = \phi(m; n) + \min(\Phi(m - 1; n - 1); \Phi(m - 1; n); \Phi(m; n - 1))$$

Where:

- Φ (m; n) is a (M+1)X(N+1) matrix
- Φ (0; n) and Φ (m; 0) are initialized with infinity or zero, depending on the application, Φ (0; 0) with zero.
- ϕ (m; n) is the cost function
- The cost between both sequences is contained at Φ (M+1; N+1)

- The analysis with the Dynamic Time Warping algorithm allows us to know if the gesture done by the user is the same as the model on the smart card and it also allow us to have speed information, it is doing quicker or slower as the model, which can increase the security level