Shuowen Shen

004316816

Assignment 10

**Computer Vision -- Open the Eye for Computer Action!**

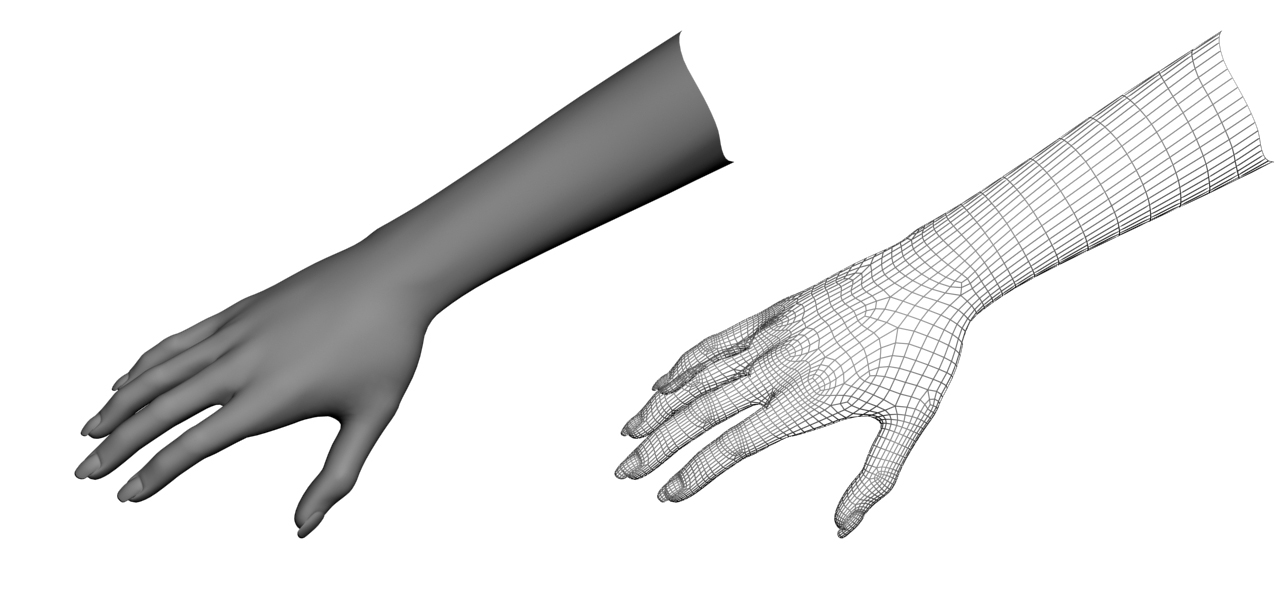
Along with the high-speeded development of computer science, the computer vision has also developed rapidly, which has opened the eyes of the computer. In the fields of the computer vision, there is signal processing, segmentation, 3-D reconstruction, tracking, recognition, etc. I will introduce more about the gesture recognition, which is in the field of recognition, and signal processing with the example of hyperlapse.

To introduce the gesture recognition, I will focus on three aspects – gesture types, algorithms, applications and limitations.

In the process of the gesture recognition, the gestures are usually classified as offline gestures and online gestures. Offline gestures are gestures, “which are manipulated after the user interaction with the object”[1]. One example is to use the gesture to ascertain the condition. The online gestures are “directed manipulation gestures, used to scale or rotate an object.”[1]

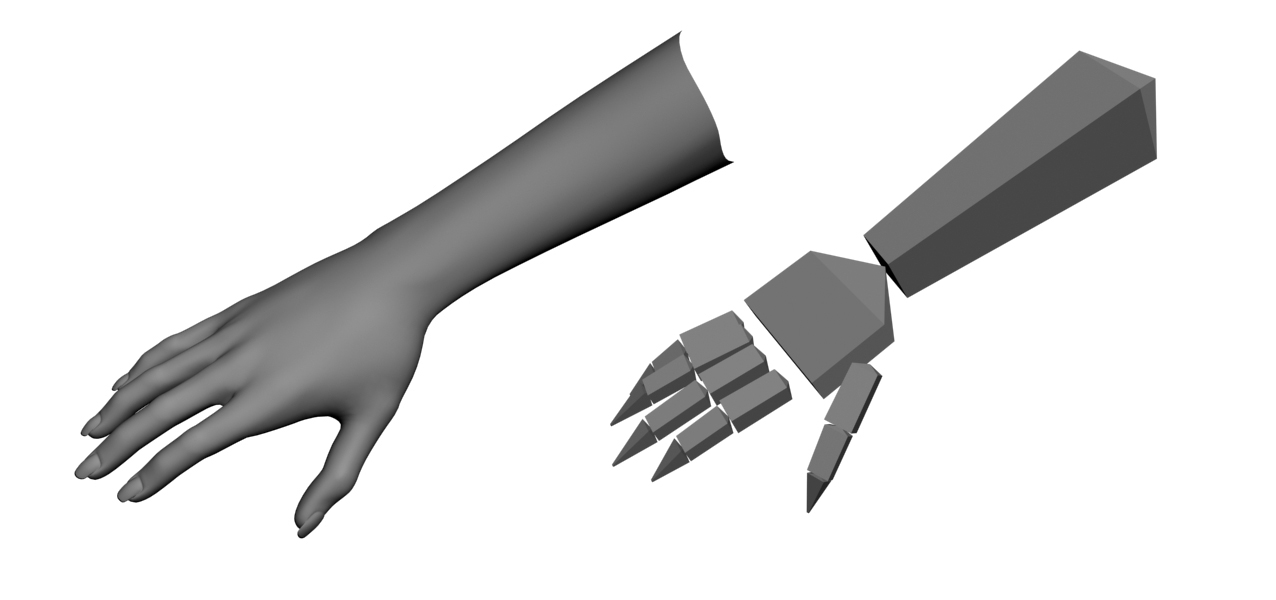
Also, to acquire good models of the gestures, three kinds of algorithms are usually used. While the first algorithm is called 3-D model-based algorithm, the second algorithm is called skeletal-based algorithm, and another is called appearance-based algorithm[2].

For the 3-D model-based algorithm, the models could be either volumetric or skeletal, or the combination of the both. However, usually the 3-D models are used usually in the computation. Also, to make the parameter much simpler, some primitive shapes such as cylinder or triangular are changed with some parameters and used to match some parts of the gestures. This way largely saves the computational space and the computational time. However, since the large scale of the computational intensity for the high-precision volumetric model, this algorithm still costs a lot of time.



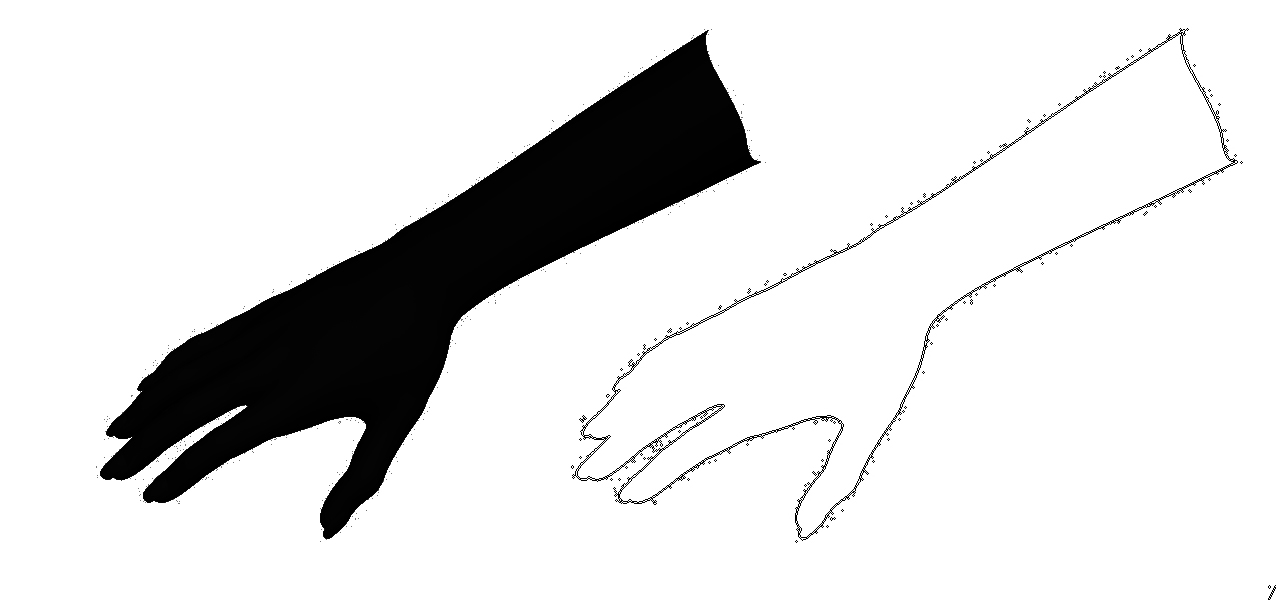
**Figure 1. 3-D model-based algorithm model[1]**

For the skeletal-based algorithm, instead of the primitive 3-D shapes, the joint angle parameter can be used to build the model. Similar to the 3-D model-based algorithm, some primitive models are used to map to the skeletal parts of the human gestures. Different from the 3-D model-based algorithm, the skeletal-based algorithm depends on position and the orientations of different parts of gestures and focus on the relation between these different parts. Since only the key parameters are analyzed, this algorithm is much faster.



**Figure 2. skeletal-based algorithmmodel[1]**

For the appearance-based algorithm, the contour of the gesture is directly acquired by the linear function. Then, the contour is compared with the models in the template and matched with one of the templates. Compared with other two other algorithms, this algorithm has some limitations, since the gesture will differ a lot from different positions and orientations, which makes this 2-D algorithm less precise.



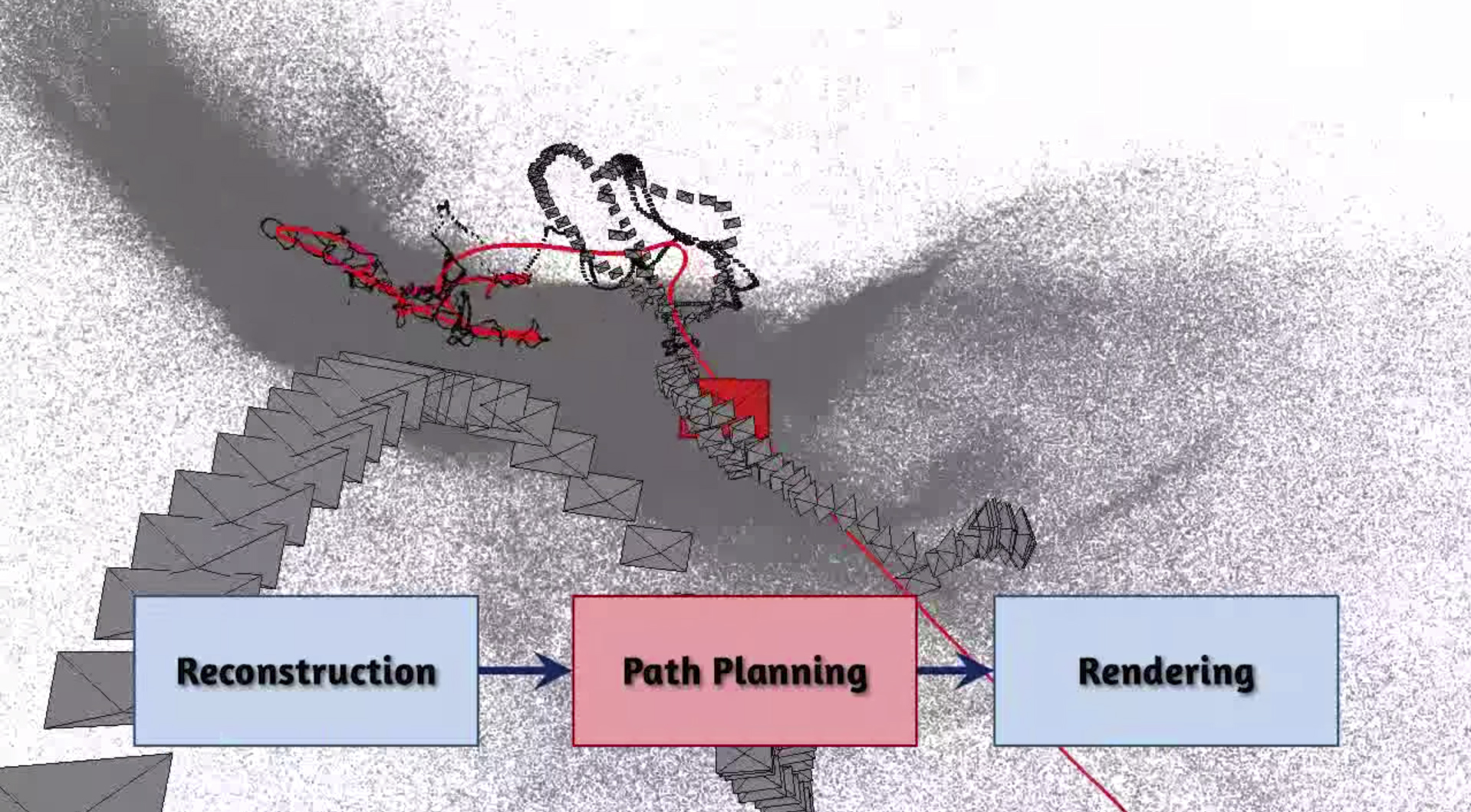
**Figure 3. appearance-based algorithmmodel[1]**

In general, the 3-D model-based algorithm is used most in the gesture recognition along with the depth camera, since this algorithm gives the precise judgment of gesture despite of its intensive time-complexity.

Besides, the gesture application has some practical usages. First, “the gesture recognition can be used to switch television channels without a remote controller”[3]. In this case, this technology provides the “couch potatoes” with a lot of convenience. They can switch the channels with their hands without make the remote controller dirty. Also, They can eat the chips and play the games on the TV at the same time. Also, gesture recognition may improve the safety of the driver. While turning the air condition and driving at the same time may seem to be dangerous, the gesture recognition could let the driver turn the air conditioner and focus on driving at the same time.

However, there are also some limitations of the gesture recognition technology. This technology has a high requirement of the environment. The environment had better not be in the consistent lighting. Also, the result of the gesture recognition will also depend on the different positions and orientations of objects.

In addition to the gesture recognition in the field of the recognition, the signal processing is also an important part of the computer vision. I will talk about one of the part in the signal processing – sampling. It is defined as “The process of converting a continuous-valued discrete-time signal to a digital signal”[4] One of the typical examples is called hyperlapse, “an exposure technique in time-lapse photography, in which the position of the camera is being changed between each exposure in order to create a tracking shot in time-lapse sequences.”[5] First, the continuous signals will be compared together to be estimated the index of the difference. And then the linear modeling is used to find the lines with the least standard deviation, and the corresponding samples are sampled. In this case, the video will be more fluent since some extra information is deleted from the videos.



**Figure 4. hyperlapse algorithmmodel[1]**

This technology is usually used in the filmmaking. Moreover, with this technology, the imagination of the computer vision is also realized. The high-fluency of the videos is just like to remind or image the videos in the mind. It provides the computer with the capability of imaging like the human beings.

Besides, there are also some useful technologies in the field of computer science such as face recognition, non-linear tracking, (which used to position the objects), 3-D reconstruction, etc. With these technologies of computer vision, the computer science will combine with our life more tightly, and provides more services to us.

**Links**

1. Gregor Miller, Sidney Fels and Steve Oldridge A Conceptual Structure for Computer Vision. CCCRV. 2011

2. Johan Bas. Gesture Recognition: Algorithms and Applications, WISE Research Group. 2011

3. Kamakshi S. 5 Useful Applications Of Gesture Technology. Techtree. 2013

4. Boashash, Boualem, ed. (2003). *Time frequency signal analysis and processing a comprehensive reference* (1 ed.). Amsterdam: Elsevier

5. Johannes K, Michael C, Richard S. First –person Hyperlapse. Microsoft Asia Research. 2013

.