Chapter 2
System Configuration and Guidance Principles

2.1 Introduction

When we go to a museum, sometimes we are interested in some exhibitions or we want to know the histories of them. Traditionally, we have to call a person as a guide to introduce them or we have to take a video tour system and choose what we want to listen to. Besides, when we go to a supermarket, we will find that a shopping cart is not easy to control, especially at a corner. And when we enter into a narrow path for buying something, we may lose the shopping cart so that we have to take a new one. It is inconvenient to handle the above-mentioned situations, and more manpower is needed. But if we have an autonomous vehicle equipped with a web camera which can follow people to go everywhere like a guide, it will be very useful for us. The vehicle can provide people with services and furthermore it can also take heavy things for us.

However, in the person following process, because a person may make a fast turn, we let the camera be held by an arm equipped on the vehicle so that the camera can be moved faster. It will keep recording the direction of the person with respect to the vehicle. Then it can make faster responses to a person’s actions and keep the person in the center of the image. The entire hardware equipments and software used in this study are described in Section 2.2.

For person following, this system has four major processes to reach this goal. The first part is the learning process which makes the vehicle to be more intelligent. In Section 2.3, we will introduce principles of the learning strategies. We will describe how to learn some information, such as positions of crossroad points in this environment, reference data for the different people heights, and every area in our environment for human interaction. The second part is the human detection process which can get features for following the target. In Section 2.4, we will describe the major steps which we use to detect a person and get features from images. The third part is the human following process which is used for following a person in environments by the vehicle. And the vehicle can solve some unexpected conditions, such as the case that the following person makes a fast turn at the corner and walks into a narrow path. The proposed principle and process will be introduced in Section 2.5. The fourth part is the human interaction process which is used for the vehicle to provide people with services. And for some applications, the vehicle also can be used as a shopping cart in a shopping mall or a tour guide in a museum. In Section 2.6, we will describe the human interaction principle and process proposed in this study.

2.2 System Configuration

In our system, we use the Pioneer 3, a vehicle made by ActiveMedia Robotics Technologies Inc., as a test bed. The vehicle is equipped with a robotic arm which can reach up to 50 cm .The tip of the arm is used to hold a digital web camera, AXIS210. The camera is IP-based and has a build-in web server of which we can adjust some camera parameters in the graphical user interface of the website, such as image resolution, image format and so on.

In our system, we can control the vehicle by a wire which is used to connect our computer and the vehicle. But the walking range of the vehicle will be limited by this wire. However, we can use wireless computer networks so that we don’t have to connect the vehicle with a wire. In this way we can have wireless communication among the vehicle, the program, and the web camera. By wireless communication, we can capture images into our program and control the vehicle by issuing some commands from our program to the vehicle conveniently.

For the applications of person following, we also equip the vehicle with a steel frame on which a basket is put so that the vehicle becomes a shopping cart for carrying heavy things for the person. The diagram is shown in Figure 2.1.

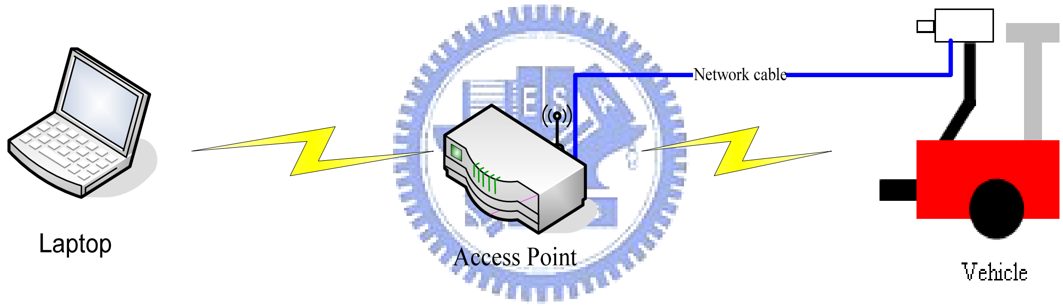

Figure 2.1 Equipment connection situation in this study.

2.2.1 Hardware configuration

The hardware equipments we use in our system include five parts. The first part is a laptop which we use to run our program and this laptop has a remote control system. A kernel program can be executed on the remote control system to control the vehicle by issuing commands to the vehicle. And it also can be used to get the status information from the vehicle and the robot arm.

The second part is the vehicle which has an aluminum body. The size of it is

44cm×38cm×22cm with three wheels of the same diameter of 16.5cm. In the vehicle, there are three 12V batteries which supply the power for the vehicle to run 18-24 hours. The vehicle can reach a forward speed of 160cm per second and a rotation speed of 300 degrees per second. It also has a build-in wireless device and an embedded control system. The embedded control system can be used to control the vehicle to move forward or backward, turn around, or control the degree of the joints of the arm by the user’s commands. The appearance of the vehicle is shown in Figure 2.2.

(a) (b)

Figure 2.2 The vehicle Pioneer3 used in this study. (a) A front view of the vehicle. (b) A side view of the vehicle.

The third part is a digital web camera. Because we want to increase the height of the viewpoint and conduct human following more skillfully, we let the camera be held by the arm which is put on the top platform of the vehicle. This arm has five degrees of freedom. However, the arm has a carry weight restriction, so we choose a camera which meets the weight carrying limit of the arm. But we find that the lighter the camera, the less the capability of the camera. We finally decided to adopt the

AXIS210 web camera which has no panning, tilting, and zooming functions but is weighted for just 305 grams, and by this light camera, we can change the direction of it easily by controlling the arm. The grabbed image has less noise because the signal from the camera is digital and the resolution of the image is 320×240 for processing the image efficiently. The appearance of the web camera is shown in the Figure 2.3.

(a)

(b)

Figure 2.3 The camera AXIS210 used in this study. (a) A front view of the camera. (b) A side view of the camera.

The fourth part is a lamp with an EYE bulb. The lamp has a conditioner which can adjust its voltage to control the intensity of the light. We can also adjust the height of the light stand from 209 cm to 121cm so that we can choose the best lighting condition for our experiment. This lamp uses a kind of blue bulb and the light from this bulb is almost white like a fluorescent lamp which is commonly used in general environments. We also use a digital light meter, TES-1335, for measuring the light power. This light meter can measure the levels of illumination ranging from 0 to 400,000 Lux and 0 to 40, 000 fc. All light equipments are shown in Figure 2.4 and Figure 2.5.

The last part is an access point which conforms to the IEEE 208.11b standard. It is used as a medium to connect the vehicle, the laptop, and the web camera. The web camera is directly connected to an access point by a network cable for transmission of the captured image. Also, we transmit all commands to the vehicle or to the camera by the wireless network. By this access point, we can adjust some camera parameters by the built-in web server and maneuver the vehicle conveniently.

Figure 2.4 The light equipments used in this study. (a) Light stand with an EYE bulb off. (b) Light stand with the EYE bulb on.

Figure 2.5 Digital light meter.

2.2.2 Software configuration

The ActiveMedia Robotics provides an application interface ARIA to control the mobile robot. ARIA is an object-oriented interface which is usable under Linux or Win32 in the C^{++} language and it can dynamically control the velocity, heading, and other navigation settings of the vehicle. We use the ARIA to communicate with the embedded system of the vehicle. And we use the Borland C^{++} Builder as a development tool in our experiments.

2.3 Learning Strategy Principle and Major Steps in Proposed Process

Our system has three major learning strategies. The first part is learning about the reference data and the lengths of the clothes for difference people heights. When the vehicle follows a person, it has to keep a suitable distance to avoid hitting the person. So it has to measure the distance between itself and the person at any time. In our system, we just use one camera and we have to use the top and bottom edges as well as the length of the person's clothes to measure the distance between the person and the vehicle in the real world. For the purpose of keeping a suitable distance, our system has to measure the reference data in advance. For this purpose, we let one person stand 2m in front of the vehicle and control the vehicle to stay at a suitable distance, at which the vehicle can see the top and bottom edges of the person's clothes and will not strike the person. Our system just does this learning process once. Then when the vehicle is used next time for difference people heights, it can measure the distance and keep a suitable one automatically by the triangulation principle. The major steps are shown in Figure 2.6.

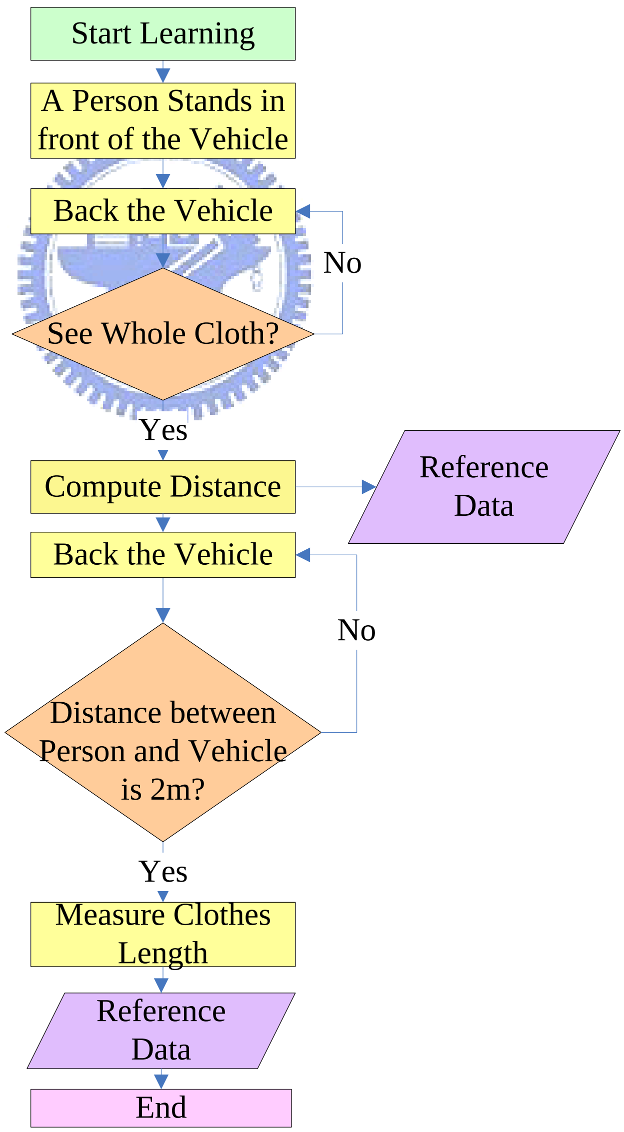

Figure 2.6 An illustration of the learning strategy process to get the reference data.

For measuring the length of the clothes, the person has to stand at a distance of 2 meters in front of the vehicle. And the system will measure the real length of the clothes in the real world by using the angles of the top and bottom edges of the person's clothes in the image. If this length is the same as that recorded in the reference data, the vehicle will enter the subsequent mode. Otherwise, the system will change the viewing distance and record this length for computing the distance in the following process. The major steps are shown in Figure 2.7.

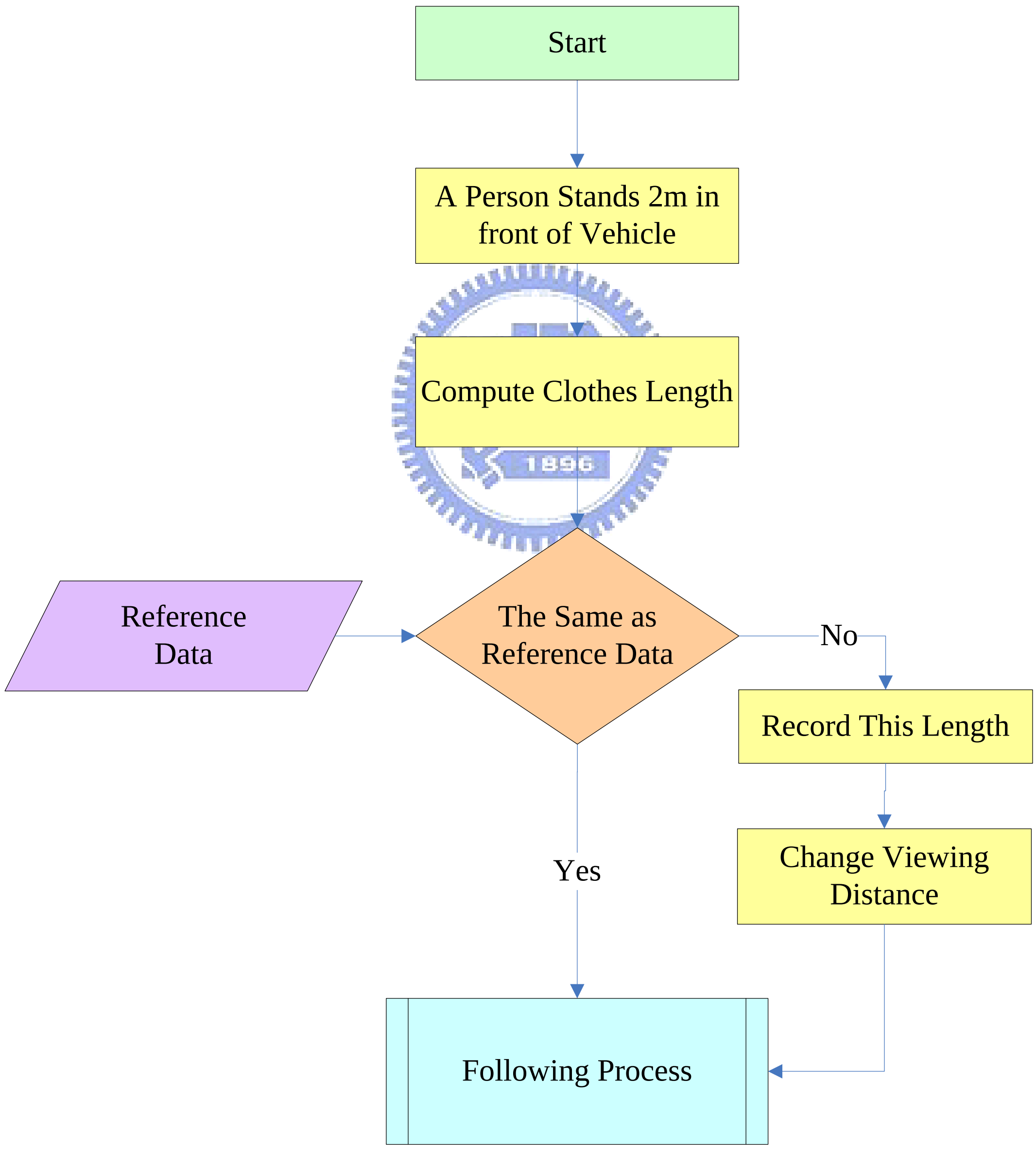

Figure 2.7 An illustration of the learning strategy process for different people heights.

The second part is learning about crossroad points in the environment. When the vehicle follows a person, he/she may make a fast turn. If the vehicle has learned the crossroad points in advance, it can go to the nearest crossroad point and turn to the direction of the disappearing person by the arm information to search the person. The major steps are shown in Figure 2.8.

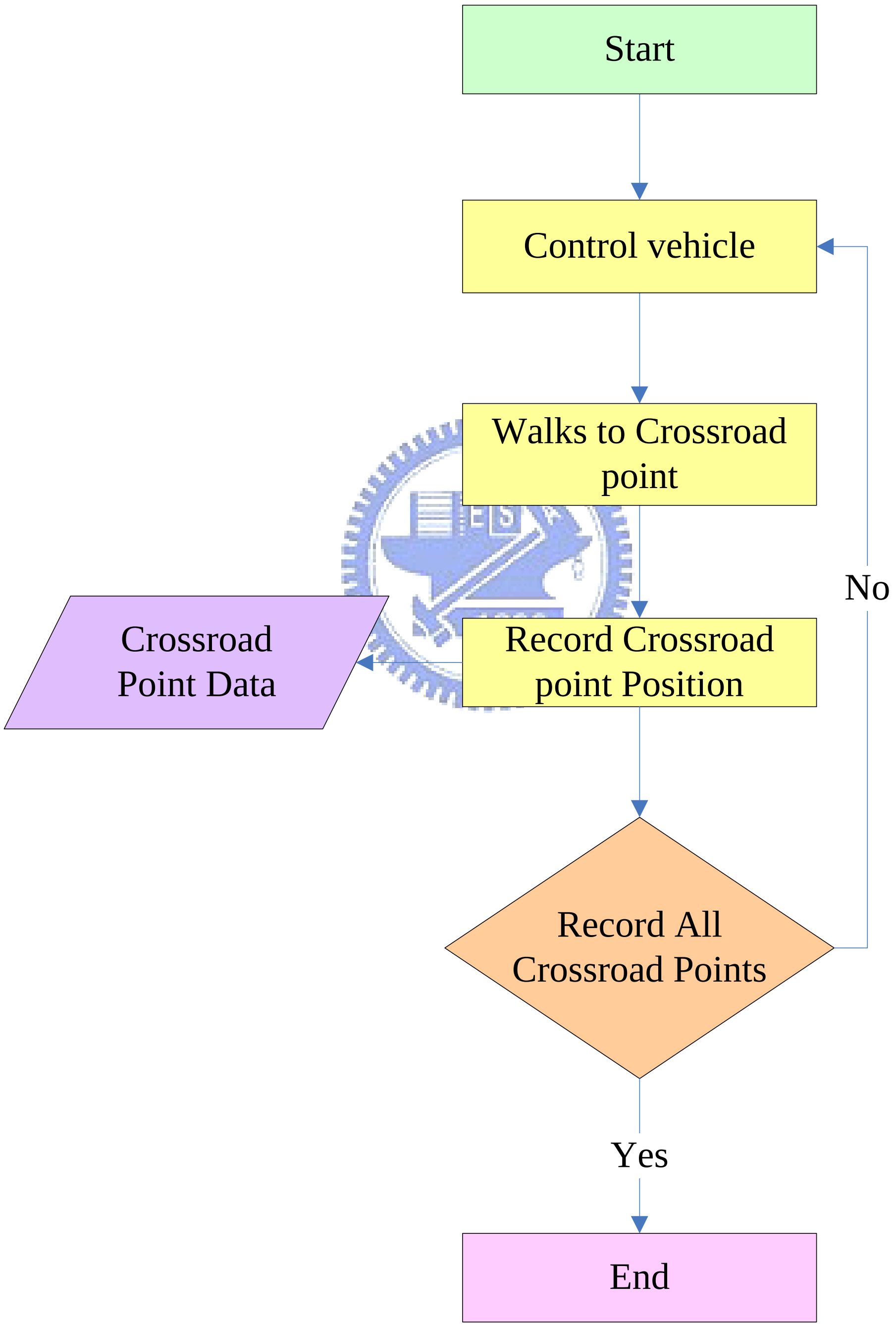

Figure 2.8 An illustration of the learning strategy process for a person making a fast turn.

The last part is learning about the area information in the environment for human interaction. For the person following application, when a person stands in front of an exhibited article, a bookshelf, or a product for a while, the vehicle will consider the person to be interested in this object. Then the vehicle will introduce some information about it, such as the history of this exhibition, the discount of the product, and so on. The vehicle will work like an assistant or a tour guide. The major steps are shown in Figure 2.9.

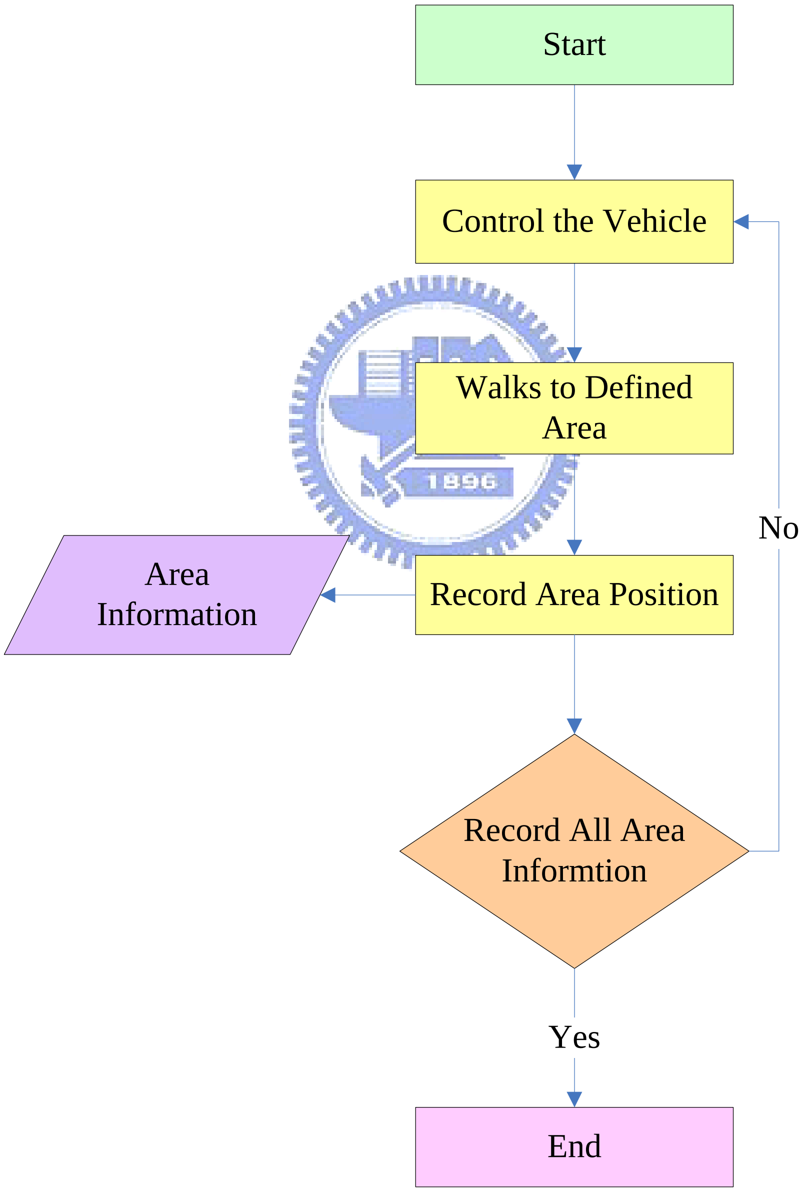

Figure 2.9 An illustration of the learning strategy process for the person following application.

2.4 Human Detection Principle and Major Steps in Proposed Process

Before the vehicle follows a person, the vehicle has to detect the person and get features which can be used to follow this person. The vehicle detects the person by the technique of face identification in this study. When the system gets an image from the camera, we use two kinds of detections first in the mean time. One is *detection of skin-colored ellipses* and the other is *motion detection*. Here the ellipse is taken to be an approximation of the human face shape.

But we will find that the method of detecting skin-colored ellipses is often affected by changes of luminance. In this study, we design a skin-colored model to adjust the parameters of skin-colored ellipses for different intensity values of the light. Besides, the skin-colored ellipse detection method is always affected by non-uniform lighting conditions. So we also design a progressive process to solve this problem. By these measures, it was found in this study that the detection of the human face becomes more accurate. Combining the results of these two kinds of detections, skin-colored ellipse detection and motion detection, the system can define a moving skin color ellipse as a face of a person.

After the person face is detected by the system, the system will extract the information of the clothes and enter the human following mode. Otherwise, the system will stay in the detection mode and continue to detect the followed person. The major steps are shown in Figure 2.10.

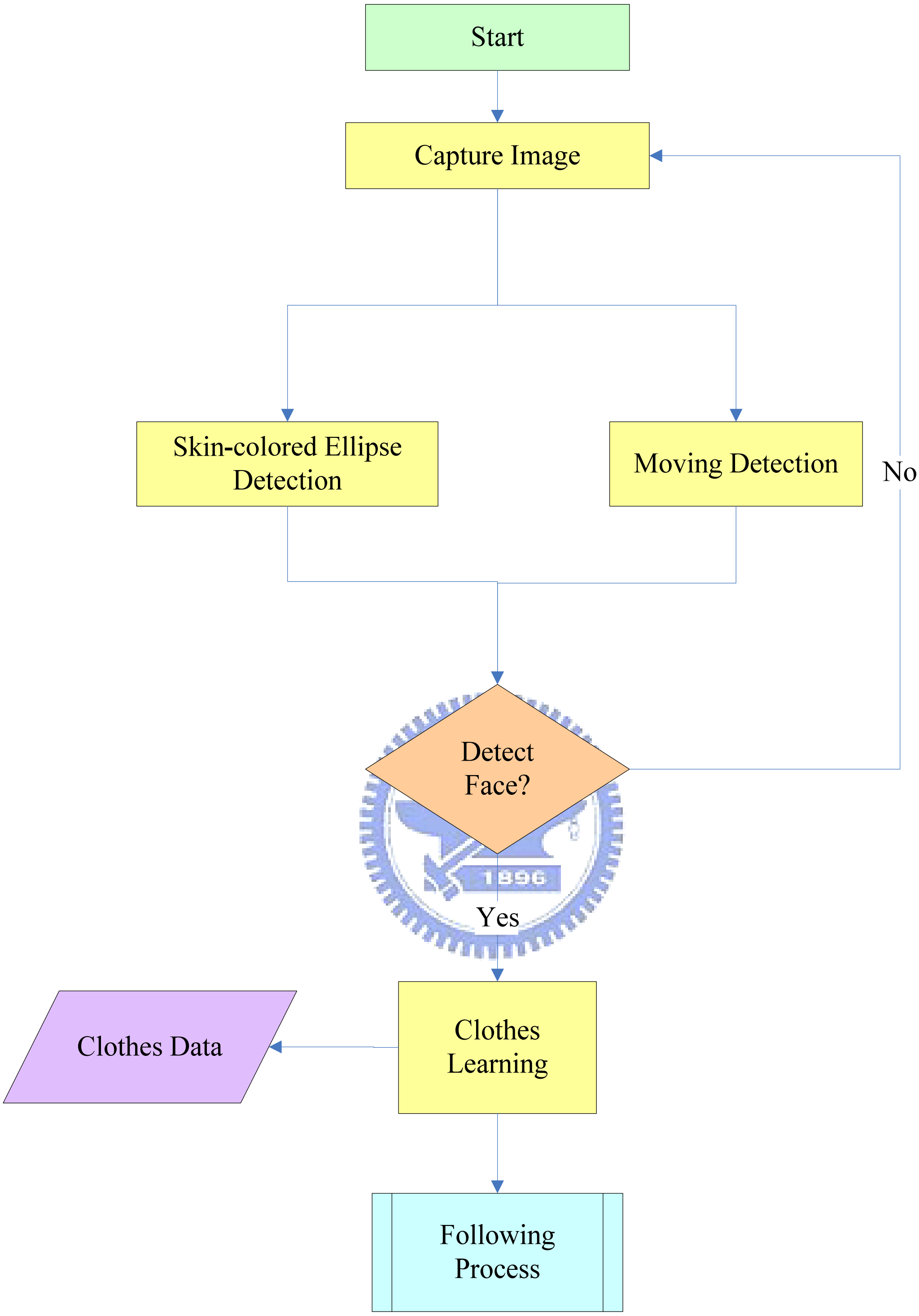

Figure 2.10 An illustration of the human detection process.

2.5 Human Following Principle and Major Steps in Proposed Process

When the vehicle is in the human following mode, it uses the information of the clothes which is obtained from the human detection mode to follow a person. Then it will meet two kinds of situations: the first is that the vehicle can find the person by the clothes feature and the second is that the vehicle cannot find the person because the person makes a fast turn at a corner.

In the first condition, the vehicle will compute the distance between the vehicle and the person. If the person stands far away from the vehicle, the vehicle will go forward to avoid losing the information of the person. If the person is too close to the vehicle, the vehicle will still follow the person by using an area tracking method by 'looking at' some part of the clothes. The vehicle will also give a warning signal to inform the person to be careful. If the person continues to stand at an identical position, the vehicle will enter the human interaction process.

In the second condition, the person makes a fast turn so that the vehicle will not see the person. Then the vehicle cannot follow the person anymore. So in this study, we design a turning mode for this kind of situation. Because the turning speed of the vehicle has a limit, we also use the robotic arm to follow the person in the mean time. Because the arm of the vehicle always follows the person, the vehicle can keep knowing that the person is on which side with respect to the vehicle (left or right). When the vehicle finds out the side on which the disappearing person is, the vehicle will go forward to the closest crossroad point and turns to the correct direction to search the disappearing person. The major steps are shown in Figure 2.11.

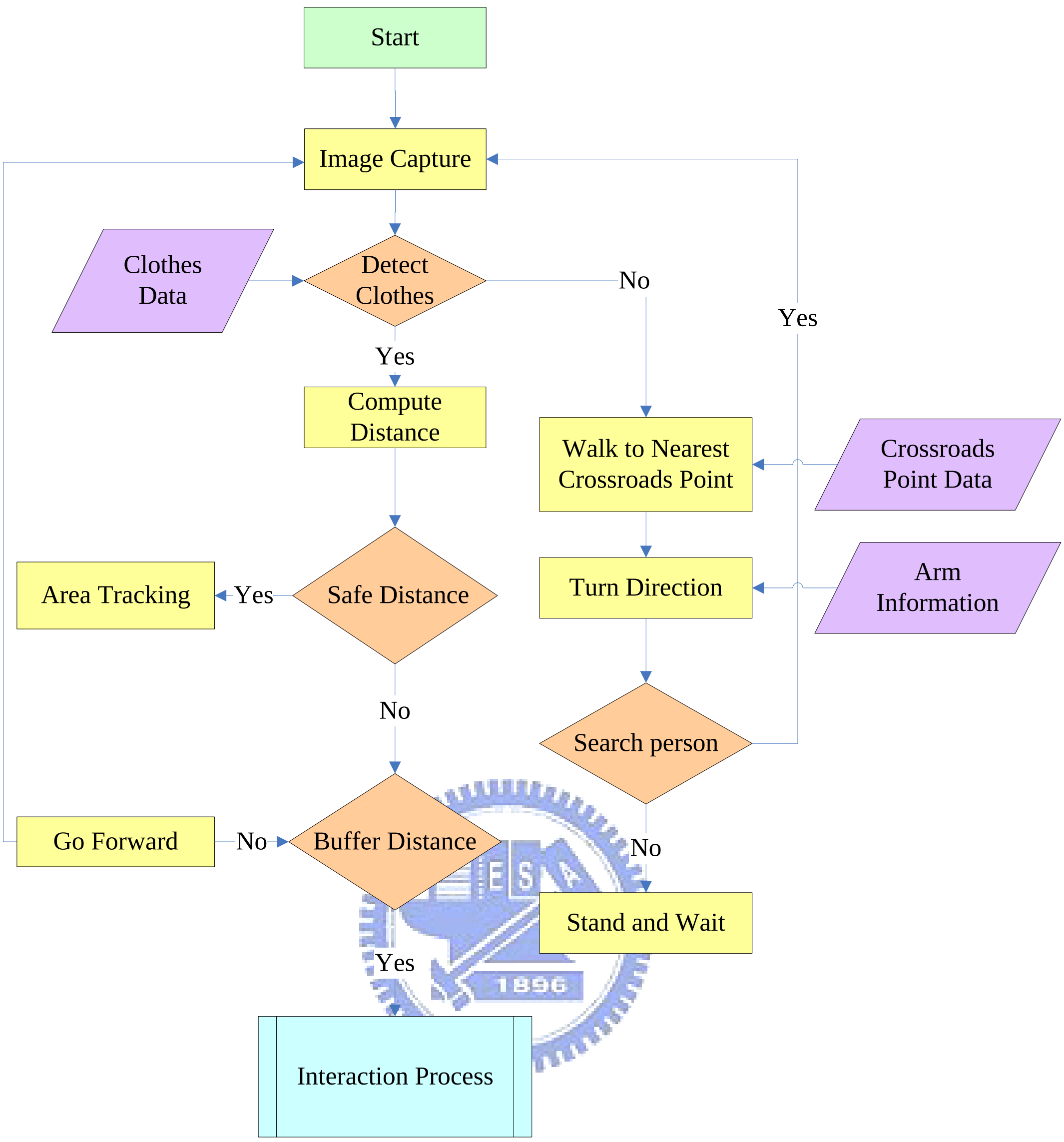

Figure 2.11 An illustration of the human following process.

2.6 Human Interaction Principle and Major Steps in Proposed Process

In the human interaction process, the vehicle is more intelligent and can analyze images to understand the actions of a person. In our study, we design two kinds of human interactions. The first kind is that if the person turns to the left or right direction and stands somewhere for a while, the vehicle will consider that the person is interested in the nearby object. Then the vehicle will make an introduction to it by a recorded audio. The second kind is that if the person faces the vehicle and waves his hands, the vehicle will consider that the person is calling the vehicle, like calling a pet. Then the vehicle will navigate to the person to conduct some works, such as putting heavy things in the basket in the vehicle, like an assistant or a tour guide. The major steps are shown in Figure 2.12.

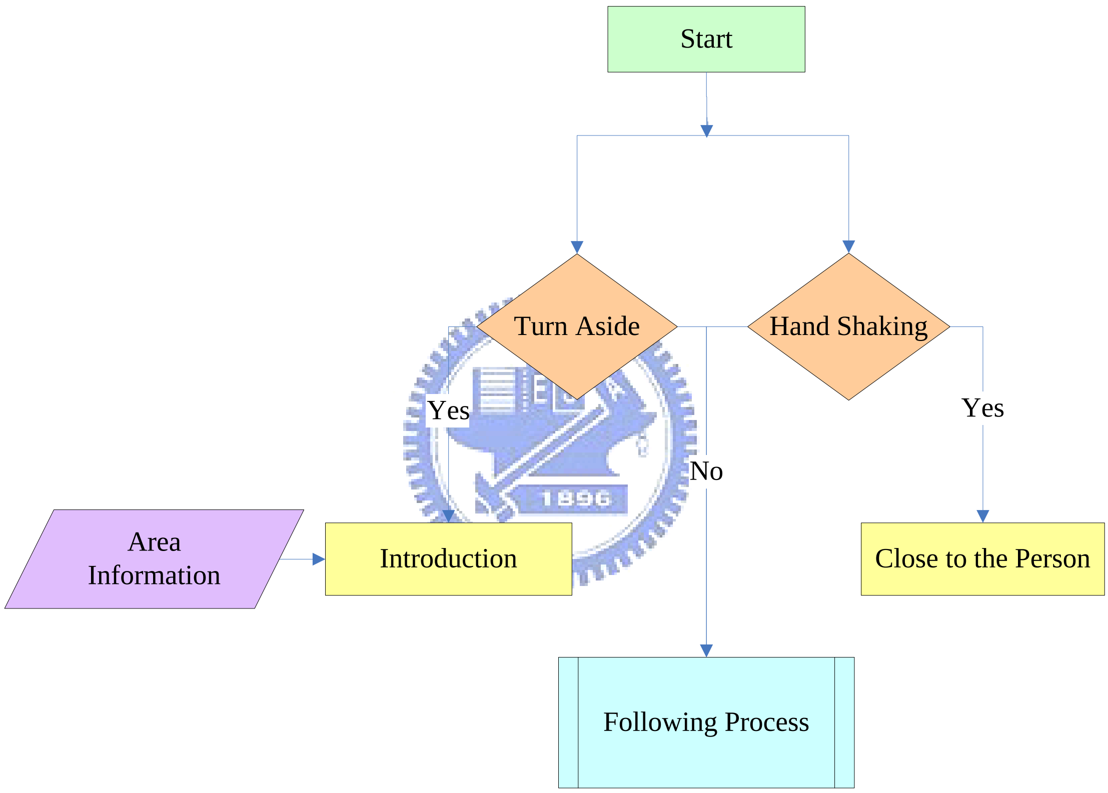

Figure 2.12 An illustration of the human interaction process.