Chapter 1
Introduction

1.1 Motivation

Recently, vision-based autonomous vehicles and mobile robots have been used in more and more human environments to play various roles. They may be utilized to conduct various tasks, such as:

(1) automatic patrolling for security monitoring in indoor or outdoor environments;

(2) keeping company with old people like a nurse to take care of them;

(3) working in dangerous places where people cannot enter or stay;

(4) following or guiding people like assistants, and so on.

For playing the role of a guide by following a person, which we call a *person-following guide* in the sequel, if we use stationary cameras, we have to install many of them and there might be still corners where the cameras cannot cover. But if we use an autonomous vehicle equipped with a video camera, the problem can be solved because the vehicle can follow people to go everywhere like a guide. And there are still other advantages. For example, the vehicle can help people carrying things and interact with them by some actions.

Before the vehicle follows a person, it has to detect the person first. Face detection is a feasible way for such a purpose. However, it is not easy to realize a *robust* face detection function because the taken image might be easily influenced by brightness changes or non-uniform lights in the real world.

When a vehicle follows a person, the vehicle has to keep a suitable distance in order to avoid hitting the person. When the person is too close to the vehicle, it is desired to let the vehicle still follow the person and inform the person to be careful. Even by doing so, some unexpected situations might still be encountered, such as the case that the person makes a fast turn so that the vehicle cannot see the person or the case that the person walks into a narrow path so that the vehicle is apt to strike the wall and difficult to keep following the person.

For the application of person guidance, the vehicle needs the additional functions of making some introductions to the environment or surrounding objects, or interacting with the person to give some responses to the person's actions without the use of voices and touches.

So the goal of this study is to design an intelligent system for person following by the use of a vehicle equipped with a video camera. It is desired to design the system to be capable of distance measurement, human detection and following, as well as interaction with humans. By these functions, more applications of the vehicle can be created, like being used as an autonomous shopping car to help people carrying heavy goods or as a guide in a museum, an exhibition, a tour, and so on.

1.2 Survey of Related Studies

To achieve the mission of person following in indoor environments, the function of human detection is required at first for finding a targeted person in the person following process. Many features have been used to recognize humans, for example, motion, shape, skin color, and so on. By using fixed cameras with stationary backgrounds, the frame differencing method [1] [2] is used frequently to obtain the background while the scene contains no people.

In the study of human skin colors, many methods have been proposed to build a skin color model. Wang and Tsai [3] proposed a method which uses an elliptic skin model to detect human faces by color and shape features in images. The simplest model is to define a region of skin color using the values of C_b and C_r [4]. If the C_b and C_r of a pixel fall into the region which is defined in advance, then this pixel is classified to be of the skin color. Chai and Ngan [5] proposed a method which uses many sets of the values of C_b and C_r to make a *skin-color reference map*. The map is then used to separate skin regions from background. A skin color region defined in the HSV color space to separate skin regions from background was proposed by Kjedsen and Kender [6]. Besides the above methods, clustering which uses Gaussian density functions [7] [8], a mixture of Gaussians used to model skin colors [9] [10] [11], and histogram learning [12] are also popularly-used methods. But for these methods to be applicable, a large amount of training data is required for finding the best classifier.

However, the skin color model, when applied to this study, is often ineffective because of the brightness fluctuations encountered in the environment. The histogram equalization method proposed by Hidai et al. [13] aims to reduce brightness fluctuation in such cases. McKenna et al. [14] used an adaptive color mixture model to estimate the color distribution of an object online and adapt it to accommodate changes in the viewing and lighting conditions to track faces under varying illumination conditions. Recently, several systems using mixed features of color, shape, and motion of the object to achieve human detection and tracking [15] [16] have also been proposed.

Besides, some systems for following a person have been proposed. Ku and Tsai [17] proposed a method that uses sequential pattern recognition to decide the location of the person related to the vehicle and detects the rectangular shape attached on the back of the person to achieve smooth person following. However, the person has to appear in the image all the time and the road has to be wide enough for this method to work. In applications of person following, Kwolek [18] proposed a method that determines the position of a mobile robot by laser readings. The tracking of the human head is done by a particle filter technique using the features of color, depth, gradient, and shape. Hirai and Miroguchi [19] studied person following to provide people with services by the use of a human collaborative robot which tracks the back and shoulder of a person. For helping people carrying heavy things, the system can be developed to be a shopping cart. Kulyukin, Gharpure, and Nicholson [20] proposed an autonomous vehicle for the blind in grocery stores and conducts navigation by using lasers. It is inconvenient because the vehicle must have a laser tool. It also has to paste blue tapes on the floor for the vehicle navigation. So it has to spend more manpower than using vision-based autonomous vehicles which detect and follow a person automatically.

1.3 Overview of Proposed Approach

The goal of this study is to design a vision-based autonomous vehicle system for person following in indoor environments. An overall framework of the proposed system is illustrated in Figure 1.1.

The proposed approach to person following for use in the vision-based autonomous vehicle system includes five major stages. The first stage is the learning process to get reference data for distance measuring; the crossroad-point information for turning situations; and the area information for human interaction.

The second stage is camera calibration for measuring the distance between the person and the vehicle by a so-called *angular mapping* technique [3] in which each point in the image represents a unique light ray from the viewpoint into the camera. By the angular information of the light rays and the height of the camera, we can know the relative distances of targets in images. But the technique is not suitable for different people heights because in this technique the height of the person is limited in a range which must to be defined in advance. So we use the reference data which are obtained in the learning stage and compare the real-time information to measure the distance between the person and the vehicle. Also, the *angular mapping* technique has to use the top and bottom edges of clothes, and so the vehicle has to 'see' the person's clothes entirely. Consequently, in this stage we also change the viewing distance between the vehicle and the person for the vehicle to see the whole clothes of different people heights. Finally, we also propose a technique for the vehicle to follow the person by *area tracking* when the vehicle just 'sees' part of the person's clothes.

The third stage is human skin detection using a model which is adapted to handle the change of luminance. An elliptic skin model to detect human faces by color and shape features in images is proposed by Wang and Tsai [3]. But this model is suitable only in a limited range of luminance. So we propose an improved skin color model which can adjust its elliptic center to adapt to the change of luminance. In establishing the model, we capture face images in different luminance situations by using a lamp with adjustable light strengths. We adjust the lamp in every five scales with a light meter and measure the values of C_b and C_r of the skin colors of the taken human face images. In this way we obtained a skin model curve for this system to improve the face detection effect. The details will be described in Chapter 4.

In addition, we use a progressive method for dealing with none-uniform luminance. In this method, we use blocks which consist of a square region of pixels in the image. The size of these blocks is all the same and the size is from 320×240 to 40×30 until the system detects the human face.

The fourth stage is person following. It not only conducts the basic person following function but also deals with some unexpected situations, such as the case that the person makes a fast turn so that the vehicle cannot see the person or the case that the person walks into a narrow path so that the vehicle will possibly hit the wall and cannot keep following the person.

In order to solve these problems, in this study we design the system to find a disappearing person by learning crossroad-point information first and recording the disappearing direction of the person by using a mechanical arm which is equipped on the autonomous vehicle.

For human interaction or other applications like functioning as a shopping cart, the last stage is human moving direction detection by analyzing the person's shape aspect ratio and the distributions of the skin and the hair colors of the person. Furthermore, we also detect the person's hand movement by motion detection and use the area information of clothes for guidance.

1.4 Contributions

The major contributions of this study are summarized as follows.

(1) A method of computing the distance between a person and the vehicle for different people heights is proposed.

(2) An area tracking method using images including part of clothes for person following is proposed.

(3) A skin detection model method which is adaptive to luminance changes for human face detection is proposed.

(4) A progressive method for human detection in non-uniform luminance is proposed.

(5) A method for detecting a disappearing person who turns fast at a corner and walks into a narrow path is proposed.

(6) A method for detection of human turning directions is proposed.

(7) A method for detection of human hand movement is proposed.

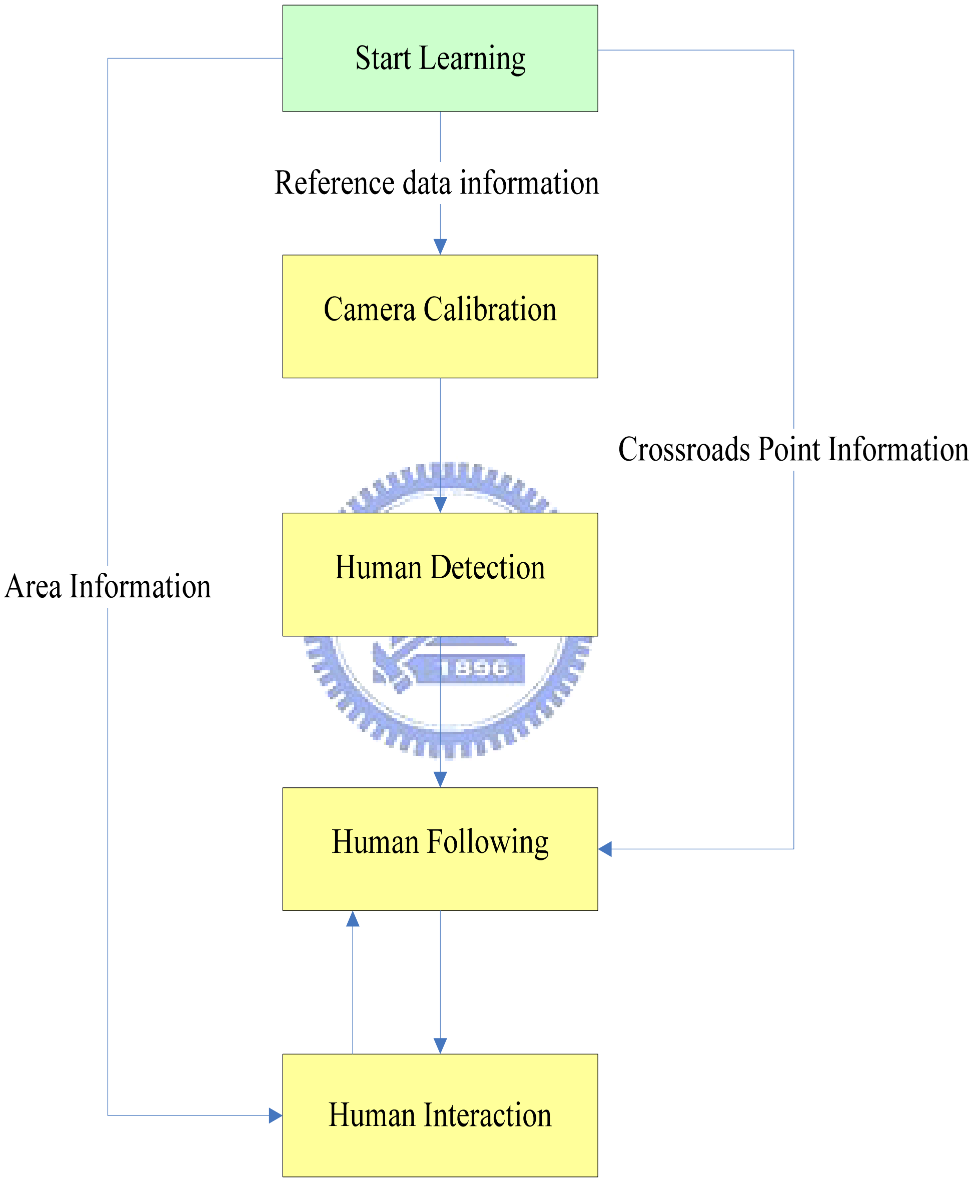

Figure 1.1 A flowchart of proposed system.

1.5 Thesis Organization

The remainder of this thesis is organized as follows. In Chapter 2, we describe the system configuration of the vehicle and the principles of learning, human detection, and human following. In Chapter 3, the proposed the method of using the reference data for difference people heights and the proposed technique of area tracking for the vehicle just seeing part of clothes are described. In Chapter 4, the proposed method for human detection by using the skin color model is described. The proposed techniques for human following are described in Chapter 5. The functions for human direction and hand movement detection are described in Chapter 6. Some satisfactory experimental results are shown in Chapter 7. Finally, some conclusions and suggestions for future works are given in Chapter 8.