Abstract: - This paper presents a new system to estimate the head pose of human in interactive indoor environment that has dynamic illumination change and large working space. The main idea of this system is to suggest a new morphological feature for estimating head angle from Thermal Camera. When a thermal camera shows distribution of face heat, there are obviously some morphological distributions. Applying a threshold to the heat distribution, we also obtain the different morphology image from different head yaw pose. Therefore, we can obtain the morphological shape of heat distribution of face. Through the analysis of this morphological property, the head pose can be estimated. It is simple and significantly invariant from illumination in comparison with other algorithm which adopts normal camera as a sensor. Our system can automatically segment and estimate head pose in a wide range of head motion without manual initialization like other optical flow system. As the result of experiments, we obtained the reliable head orientation data under the real-time performance.

Key-Words: - Gaze direction, Thermal camera, Human robot interaction, Ubiquitous

1 Introduction

Recently the ubiquitous computing which can communicate with the human in anywhere without the user’s recognition of the system, has been noticed as the new paradigm of computerization. In the ubiquitous environments, the computer needs to obtain user’s intention or information with accuracy. For this reason, the Human Computer Interaction (HCI) has occupied an important position. Over the past year, many sensing modalities have been introduced for HCI such as keyboard, mouse, and the haptic device. Although they can acquire user’s intention with accuracy, they have to equip certain device on human’s body or need the skill to control them. This is not sufficient to satisfy those requirements of HCI for ubiquitous computing. At this point, one of the effective sensing modalities that make user’s effort become minimized is the vision system for HCI because it uses cameras which are typical passive sensor and it has reliable performances.

However, it has also the demerit of the computer vision such as considerable amount of image processing [1] and dynamic illumination problem. To overcome these problems, many research results have been proposed.

In vision-based HCI, there are many perception modalities. These can be divided into six categories such as [2]: people tracking, gesture recognition, facial perception, facial expression, eye orientation tracking and head pose tracking. Especially, the head pose tracking has a unique advantage that it is the simple and general approach to obtain user’s geometrical attention. A direction of attention is an intuitive pointing cue for estimating user’s intention. We can simply obtain user’s intention by estimating the head pose. Therefore, the estimating of head pose is simple way to obtain user’s intention.

Many of the previous researches were proposed on this topic. Yang [3], Liang [4] and Haville [5] have developed an optical flow method in which the head pose is tracked by analyzing the change of features between the image frame sequences. However, their methods need the manual feature initialization and they have boost-trapping problem which require several frames to estimating correct data when the typical optical flow system starts. Yang [6] used a skin color as a visual feature of their system. This method might not cover the bad illuminative condition because they used the general monocamera which has a limitation to detect color feature under the bad illuminative condition.

2 Gaze Direction Recognition

Please, leave two blank lines between successive sections as here.
In this section, we propose the entire procedure which estimates head rotation parameter. We suggest the key feature and explain the procedure of our method.

Our head pose estimation process is classified into the following seven steps in detail like Fig. 1. In the step 1 and 2, the thermal morphological feature is derived from thermal camera. Then, the head segmentation from background and head pose estimation process is performed in the step 3-4 and 5-7, respectively. Fig. 1. shows a quick overview of the algorithm.

To obtain precision thermal image, we need a high-quality thermal camera. We used the commercialized thermal cameral which can obtain raw-thermal data. Obviously, the appropriate thermal image range is one of the essential factors for our thermal image based head pose estimation system. We fixed thermal camera range from 10.2 to 38.5 Celsius degree. It obtained from more than 100times try-and-error test. FLIR Thermovision-A20 presents fixed thermal image on real-time. It adopts automatic lens recognition and measurement correction system for research purpose.

However, through the observation of the specific warm region of face, we could find a new meaning of higher thermal image which obtained from applying threshold value to the original thermal image. It has not only the temperature information but also some morphological property about human head posture. In the human head and face, eyes, nose and mouse have a higher temperature compare with other parts like hair, forehead and checks. Fig. 2 shows the morphological feature which obtained from constrained thermal image that emphasizes eyes, nose and mouse of human face under low threshold value. And, the morphological feature has an invariant property for yaw head rotation when the human face can be shown by the camera. We focus on our research to this property. The morphological feature from the constrained threshold value is a main key of our head pose estimation algorithm. We obtain the constrained feature first, then, pose estimation algorithm starts. The detail will be described following section.

3 Experimental Setup and Result
We implement our algorithm using C++ under the Pentium 4 PC with Microsoft Windows environment and FLIR thermal camera connected by IEEE1394 cable on real-time (about 24 frames per second at 320x240 frame resolutions). The camera is fixed vertically on the wall of our experimental studio where is placed in one of the rooms at Korea Electronics of Technology Institute. The camera is hanged at 180cm height from the ground. We set up the system which consists of vertically arranged.

The experimental studio has a complex background and it can control the illumination condition by using 8 fluorescent lamps. The working space has a size which is 6m x 6m x 3m so that the studio could cover every field of camera views and one person can be located in the studio. He or she can stand wearing any clothe and move any anywhere within working space.

To obtain better experimental result, we used several human models that have different appearance in face color, hair style and gender. And target human’s head was fixed against roll and pitch rotational parameters. This makes that we measure only yaw parameter steadily. We collected head pose data that ranged from -90º to 90º for various model appearances and under the different lightning condition.

Fig. 6 shows the result of yaw head pose estimation of different type models and lightning conditions. Our experiments result was acquired from 3 different models that have a dark skin color and
short hair style, middle bright skin color and middle long hair style and bright skin color and long hair style model. All experiments were performed under the 3 different illumination conditions which consist with bright indoor light (230 Lux), middle bright light (110 Lux) and dark illuminating conditions (35 Lux).

As the result of experiments, we obtained the robust orientation parameters against the variation of illumination condition and the appearance of different types. All data could obtain in the real-time (24frames/sec).

4 Conclusion
In this paper, we suggested a new approach of estimating head pose by analyzing morphological feature from the constrained thermal image. We can obtain the morphological shape from the algorithm. Through the analysis of this morphological property, the head pose can be estimated. We show that the experimental result under the different lightning condition and different models. And, we could obtain reliable results. As shown in the experimental results, we showed that our algorithm can be robust against changes of illumination condition and differences of human head appearance. It also has wide working space because we adopted head segmentation method which can extract head in the entire field of view. Additionally, our approach is not based on optical flow. Therefore, it does not require the manual initialization and there is no boost trapping problem. Consequently, our approach is suitable for perception modality of the future ubiquitous and intelligent environments which require all above conditions in the real time performance.

References:


