Contactless Gesture Recognition System Using Proximity Sensors

Heng-Tze Cheng∗†, An Mei Chen†, Ashu Razdan†, Elliot Buller†
∗Carnegie Mellon University †Qualcomm Incorporated

Abstract—In this paper, we present a novel contactless gesture recognition system using proximity sensors. A set of infrared signal feature extraction methods and a decision-tree-based gesture classifier are proposed. The system allows a user to interact with mobile devices using intuitive gestures, without touching the screen or wearing/holding any additional device. Evaluation results show that the system is low-power, and able to recognize 3D gestures with over 98% precision in real time.

I. INTRODUCTION

Gesture-based interfaces provide an intuitive way for users to specify commands and interact with computers [1]. Existing gesture recognition systems can be classified into three types: motion-based, touch-based, and vision-based systems. For motion-based systems [2], [3], a user must hold a mobile device or an external controller to make gestures. Touch-based systems [4], [5] can accurately map the finger/pen positions and moving directions on the touch-screen to different commands. However, 3D gestures are not supported because all possible gestures are confined within the 2D screen surface. While the first two types of system require users to make contact with devices, vision-based systems [1], [6] using camera and computer vision techniques allow users to make intuitive gestures without touching the device. However, vision-based systems are computationally expensive and power-consuming, which are undesirable for resource-limited mobile devices like tablets or mobile phones.

To solve the existing challenges, we present a novel gesture recognition system with the following contributions:

- The design and evaluation of the first contactless gesture recognition system using only infrared proximity sensors.
- The proposed infrared (IR) feature set and classifier for real-time 3D gesture classification.
- Reducing the power consumption of gesture recognition.

The design also reduces the frequency of users’ contact with devices, alleviating the wear and tear of screen surface.

II. SYSTEM DESIGN AND METHODS

A. Proximity Sensor Data Acquisition

For the configuration under study, a proximity sensor consists of two IR LEDs and a IR receiver (see Fig. 1), which are placed underneath a plastic/glass screen surface, surrounded by optical barriers. The LEDs emit IR strobes in turns as two separate channels. When a hand or any object is near, the receiver detects the reflection of the IR light, whose intensity increases as the object distance decreases. The light intensities of the two IR channels are sampled by the firmware at 100Hz.

The time delay $t_D$ is calculated using the cross-correlation value of two discrete signal sequences $f$ and $g$:

$$ t_D = \arg \max_n \sum_{m=-\infty}^{\infty} f^*(m)g(m+n) \quad (1) $$

B. Gesture Recognition Algorithm

The algorithm continuously scans the input IR intensity data and decides if a predefined gesture is observed. First, the data is divided into 50% overlapping frames, 140 ms each. Then, three types of feature are extracted from each frame:

1) Inter-channel Time Delay: The feature measures the pair-wise time delay between the sensor data of two channels, which shows how a hand approaches the IR LEDs at different instants. This corresponds to different moving directions of hands (see Fig. 2 for example). The time delay $t_D$ is calculated by finding the time shift $n$ that yields maximum cross correlation value of two discrete signal sequences $f$ and $g$:

$$ t_D = \arg \max_n \sum_{m=-\infty}^{\infty} f^*(m)g(m+n) \quad (1) $$

Fig. 1: The architecture of the gesture recognition system.

Fig. 2: An example of proximity sensor data and IR features.
Recall of left/right swipe

Recall (%)

100 100 80 20 20 20 60 60 60 80

Ch

Y

0

Right Swipe were not exposed to the users during the data collection. Each user performing each gesture 100 times. To prevent users from adapting to the system over time, the recognition results were not exposed to the users during the data collection.

The dataset consists of 2,000 gesture samples in total, with 5 users, including 1 left-handed and 4 right-handed users. The system is evaluated on a gesture dataset collected from a device. The system is evaluated on a gesture dataset collected from a device.

The system power is dominated by the power consumed by IR LED and the control chip:

\[ P_{\text{LED}} + P_{\text{chip}} = I_{\text{conv}} \cdot T_{\text{prx}} \cdot (I_{\text{LED}} + I_{\text{chip}}) \cdot V_{\text{LED}} \]

where \( T_{\text{prx}} \) is only 0.3 mW (idle) to 20 mW (active, with larger \( T_{\text{prx}} \) when an object is in proximity) [7], much lower than the 200-mW power budget for typical UI of mobile device [8] \( (I_{\text{conv}} \) and \( T_{\text{prx}} \) are conversion frequency and pulse width).

IV. CONCLUSION

We have presented a contactless gesture recognition system that allows users to make gesture inputs without touching, holding, or wearing any device. Using the proposed IR feature set and classifier, the system can recognize 3D gestures with 98% precision. The low power consumption and high recognition accuracy make the system particularly desirable for deployment on resource-limited mobile consumer devices.

REFERENCES