Three-dimensional finger tracking using direct and reflected infrared images

Kensaku Fujii *Jun Shimamura* *Kenichi Arakawa* *Tomohiko Arikawa*

NTT Cyber Space Laboratories, Nippon Telegraph and Telephone Corporation
1-1 Hikarinooka Yokosuka Kanagawa 239-0847, Japan
+81-468-59-2431
fujii.kensaku@lab.ntt.co.jp

ABSTRACT

In this paper, we describe finger detection in three-dimensional (3D) space to more effectively support human computer interaction. To achieve real-time 3D tracking, we use an infrared camera with a structured half-mirror. This configuration allows us to detect direct and reflected infrared images simultaneously. It detects targets reliably and tracks them three-dimensionally based on the reflected symmetry and the epipolar consistency. We also describe experimental results of the 3D tracking accuracy using a Polhemus tracker.

Keywords

real-time 3D tracking, infrared image, catadioptric stereo

INTRODUCTION

Tracking and locating the fingers of a user can play a key role in the field of attentive user interfaces. The use of computer vision to detect fingers and their motion removes the constraints of many existing devices, i.e., glove-based. It is necessary to detect the user's fingers robustly and to track them three-dimensionally in real time for natural interactions.

One of the interesting attempts for reliable finger detection is to use the infrared image. The infrared image does not need to pay attention to the lighting and color conditions. As demonstrated by several researchers, the infrared camera [1] or the infrared illumination [2, 3] can be used for reliable detection. The adaptation for 3D measurement is also proposed, such as Motion Processor [4] using the infrared illumination. It provides a depth calculation based on the infrared intensity of reflection. However, the intensity is variable through the target surface and direction. Due to this instability and the range of the infrared illumination, it is difficult to realize stable and accurate 3D measurement.

We introduce a simple technique known as the catadioptric stereo [5] for 3D measurement using the infrared camera. The catadioptric stereo can capture a stereo view using a single camera and mirror. This scheme has several advantages over traditional two-camera stereo; it avoids the computer vision problems which are connected with two cameras (their technical parameters such as spectral response, gain, and offset, their calibration parameters and the frame synchronization).

This paper describes a new method for real-time 3D finger tracking. Our system exploits an infrared camera with a structured half-mirror that allows us to simultaneously detect and then use direct and reflected infrared images. It is especially advantageous in terms of detection reliability and simplified 3D processing due to the reflected symmetry and the epipolar consistency. Experimental results compared 3D tracking accuracy against a Polhemus tracker show that our approach is stable and accurate.

THREE-DIMENSIONAL TRACKING

Figure 1 shows a schematic illustration of our system. It consists of an infrared camera, a display, a structured half-mirror, and tracking software that runs in real-time on a consumer-end PC. We use an infrared camera (MITSU BISHI IR-M700) to ensure the target detection. The structured half-mirror is used to detect the direct and reflected infrared images simultaneously instead of using two or more cameras. The interactions can be seen through the half-mirror; the user interface is more intuitive.

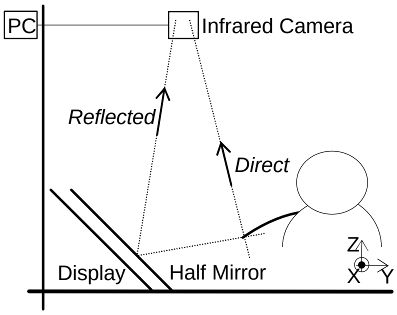

Figure 1: Architecture of 3D tracking system.

We combine the direct and reflected images of fingers to calculate their 3D positions as follow. Only simple thresholding is needed to extract the hand from the background robustly, since the temperature differences are significant. The points which become convex on the contour of the hand region extracted from the direct image are identified as the tracking target. Suppose that the image plane is orthogonal to the Z-axis and the x-axis of the image plane is adjusted to the X-axis, the point corresponding to the target point is located on the same line along the y-axis of the image plane. Due to the epipolar consistency, the corresponding points are detected on the contour extracted from the reflected image. The spatial relationship between the target point $T(X, Y, Z)$ and the corresponding point $T_m(X_m, Y_m, Z_m)$ is the mirror symmetry. That is, $T_m=RAR^{-1}T$, where R (R^{-1}) is the rotational matrix (inverse) by denoting the angle between the Z-axis and the normal to the mirror, and A is the Y-axis symmetry matrix. For perspective projection, the location of T in the image plane is (x_d, y_d), where $x_d=f\,X/Z$, $y_d=f\,Y/Z$, and f (focal length) is a constant. Similarly, the location of T_m in the image plane is (x_r, y_r), where $x_r=f\,X_m/Z_m$ and $y_r=f\,Y_m/Z_m$. The 3D positions are calculated using these equations. The correspondences of tracking targets between successive image frames are also determined to minimize Euclid distance using calculated 3D positions. Thus, our system yields reliable 3D tracking.

EXPERIMENTS

A typical result is shown in Figure 2, where infrared intensity is expressed using gray scale values. The intensity of the reflected image (lower half of 2(a)) is less than that of the directly detected image (upper half of 2(a)) because of the mirror's attenuation. An example of 3D trajectories is shown in Figure 3(a). The current implementation is running almost in real-time, approximately 25 frames per second. 3D tracking has also been used to experiment with immersive 3D graphics control in which natural finger movements are translated to moving the corresponding 3D graphics camera viewpoint and direction (see Figure 3(b)).

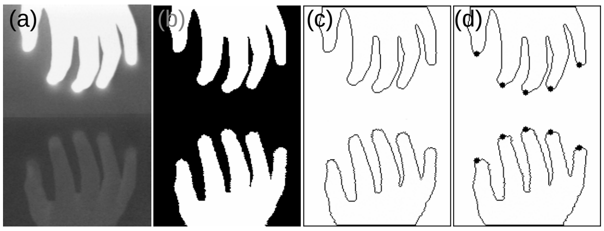

Figure 2: Example of target detection results.

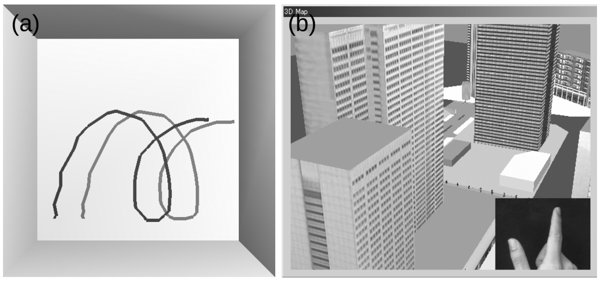

Figure 3: Example of 3D tracking trajectories and 3D tracking used to browse a 3D graphic's model of Shinjuku.

In order to assess the tracking accuracy of our method, we compared its accuracy against a Polhemus tracker, 3SPACE FASTRAK. FASTRAK is a magnetic sensor connected to a system that measures six degrees of spatial freedom and thus can be used for object tracking when tethered to an object. The observed accuracy of FASTRAK is 0.8mm RMS in spatial location. We compared FAST RAK tracking to our 3D tracking using a 320x240 image size (see Figure 4). The coordinate systems of FASTRAK and the camera were carefully aligned prior to testing. The object tracked was pulled three-dimensionally in a set trajectory. The comparison between our method and FASTRAK yielded the good results; the standard deviation of difference in each of X, Y, and Z is 0.71, 0.88, and 1.03 mm, respectively.

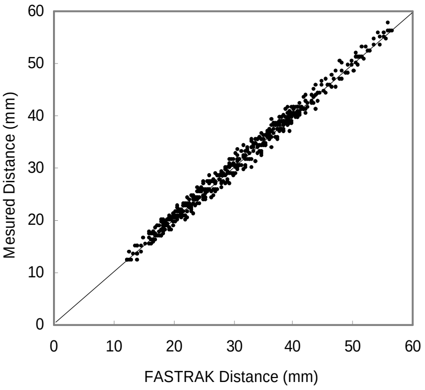

Figure 4: Comparison of Z tracking accuracy using FASTRAK.

CONCLUDING REMARKS

This paper described a real-time 3D finger tracking scheme based on the simultaneous detection of direct and reflected infrared images. It offers precise and robust tracking. Processing calculations are simplified due to the existence of reflection symmetry and epipolar consistency. Work is underway to track other targets and to explore the rich space of intuitive visual control possible with 3D tracking.

REFERENCES

1. Sato, Y., Kobayashi, Y. and Koike, H.: Fast tracking of hands and fingers in infrared images for augmented desk interface. *In Proc. of International Workshop on Face and Gesture Recognition 2000*, 462-467.
2. Rekimoto, J and Matsushita, N: Perceptual Surfaces: Towards a Human and Object Sensitive Interactive Display. *In Proc. of PUI 1997*.
3. Starner, T., Auxier, J., Ashbrook, D. and Gandy, M.: The Gesture Pendant: A Self-illuminating, Wearable, Infrared Computer Vision System for Home Automation Control and Medical Monitoring. *In Proc. of ISWC 2000*.
4. Numazaki, S., Morshita, A., Umeki, N., Ishikawa, M. and Doi, M.: A kinetic and 3D image input device. *In Proc. of CHI 1998*, 237-238.
5. Gluckman, J. and Nayar, S.K.: Rectified Catadioptric Stereo Sensors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(2):224-236, 2002.