



# HHS Public Access

Author manuscript

*Proc IEEE Annu Northeast Bioeng Conf.* Author manuscript; available in PMC 2015 July 15.

Published in final edited form as:

*Proc IEEE Annu Northeast Bioeng Conf.* 2015 April ; 2015: .

## SIFT-Based Indoor Localization for Older Adults Using Wearable Camera

Boxue Zhang\*, Qi Zhao\*, Wenquan Feng\*, Mingui Sun<sup>†,‡</sup>, and Wenyan Jia<sup>‡</sup>

\*School of Electronic and Information Engineering, Beihang University, Beijing, China

<sup>†</sup>Department of Electrical & Computer Engineering, University of Pittsburgh, Pittsburgh, PA 15213, USA

<sup>‡</sup>Department of Neurosurgery, University of Pittsburgh, Pittsburgh, PA 15213, USA

### Abstract

This paper presents an image-based indoor localization system for tracking older individuals' movement at home. In this system, images are acquired at a low frame rate by a miniature camera worn conveniently at the chest position. The correspondence between adjacent frames is first established by matching the SIFT (scale-invariant feature transform) based key points in a pair of images. The location changes of these points are then used to estimate the position of the wearer based on use of the pinhole camera model. A preliminary study conducted in an indoor environment indicates that the location of the wearer can be estimated with an adequate accuracy.

### I. Introduction

A robust and accurate in-home monitoring system is very important for maintaining health and independence of older adults, especially for those with chronic diseases. In such a system, indoor tracking is a fundamental but unsolved problem. A number of sensing technologies have been developed to solve this problem, such as those using the Wireless Fidelity (Wi-Fi), indoor GPS, Radio-frequency identification (RFID), and digital camera [1–3]. Although most of these systems are very helpful, they are often inconvenient to be applied due to operational difficulties and excessive weights/dimensions [4]. In this study, we develop a convenient indoor localization system using a wearable camera and conducted an experiment to validate its performance.

### II. Methods

A wearable camera is used to automatically acquire images at a low frame rate (i.e., one image every few seconds). To estimate wearers movement from one frame to the next, SIFT (scale-invariant feature transform) features are calculated for the key points in both images and the matching points between these two images are extracted. The position changes of the matched points are then used to calculate the moving distance or rotation angle between adjacent frames. In this preliminary study, the movement of the wearer was limited to forward motion only or rotation only, and the size of the room was assumed to be known to provide a solid reference for estimating actual distance from images.

## A. Matching of SIFT descriptor

SIFT is a classical approach to detect and describe local features in images [5]. Positions of key points are located by scale-space extrema of difference-of-Gaussian (DOG) at different scales. A 128-dimensional descriptor, which is invariant to translations, rotations and scaling transformations, is computed for each key point describing its local appearance. For the same key point in two different images, the two SIFT descriptors should be similar, although not the same. By matching the SIFT descriptors of these key points between a pair of images, the position change of each key point can be estimated. The SIFT descriptors for the key points in two images and the matching results are illustrated in Fig.1. It can be seen that the location change of the key points reflects the change of position/orientation of the camera.

## B. Position Localization

By observing the position change of the key points from one frame to the next frame, the points move in radial direction when the camera moves forward, while in parallel when the camera rotates (see Fig.1(c) and (f)). With a pinhole assumption for image acquisition, two projection models are built to study these two cases, shown in Fig. 2.

**1) Forward camera movement**—In Fig. 2(a),  $T$  is the optical center of the camera.  $D$  represents a key point on object plane  $AB$  and  $F$  is the projected points on image plane. When the camera moves forward toward plane  $AB$ , it is equivalent to move plane  $AB$  to plane  $A'B'$ . The original point  $D$  is denoted as  $D'$  on plane  $A'B'$  and its projection point is denoted as  $F'$ .  $D''$  is on the extended line of  $TD'$ .  $|TH|$  represents the focal length  $f$  of the camera, so  $|TH|$  is perpendicular to  $|FH|$ . According to the similarity between  $D''OT$  and  $D'O'T$ , it is easy to obtain:

$$|OO'| = |OT| * (1 - |FH|/|F'H|) \quad (1)$$

where  $|OT|$  is the distance between the camera and the reference plane.

**2) Camera rotation**—The projection point of  $D$  is denoted as  $F$  in the image plane in Fig. 2(b). After camera rotation, the projection point of  $D$  changes to  $F'$ . Because  $ODT$  is similar to  $HFT$ , we have

$$\tan\beta = |OD|/|OT| = |FH|/f,$$

here  $f$  is the focal length of the camera. Similarly

$$\tan\beta' = |O'D|/|O'T| = |F'H|/f,$$

Hence, the rotation angle can be calculated by:

$$\alpha = \beta' - \beta \quad (2)$$

### III. Experimental Results

In this study, an experiment in a classroom was conducted to validate the accuracy of our method. The width and length of the classroom were 11 m and 10 m, respectively. The initial position of the wearer referencing to the walls was assumed to be known. A route was designed to include three segments of walking forward, and two segments of 90° rotation. The images were acquired by a handheld cellphone (iPhone 6 plus) at pre-set locations, as well as during rotation. The focal length of the camera was 4.2mm. 17 full-resolution images were obtained. They were downsampled to 816 × 612 pixels for further analysis. In this experiment, four pairs of key points with large displacements were selected after manually removing the mismatched outliers. The corresponding location/orientation was then computed based on the averaged position change of the four matched key points between each pair of adjacent images. After calculating the moving distance and rotation angle every time the camera moved, the trajectory of the camera was plotted and compared with the pre-defined route, see Fig. 3. In Fig. 3(a), the straight lines represent the moving track. The “\*” points on the straight lines represent the camera location where each image was taking while the “\*” points near the corner represent the camera orientation. Table 1 shows the calculated moving distance and rotation angle corresponding to each image. The comparison with the actual distance and rotation angle is also listed in Table 1. It can be seen that the location can be estimated with an adequate accuracy.

### IV. Conclusion&Discussion

We have described an SIFT-based localization system to monitor older individuals' movement in indoor environment. A preliminary experiment has been conducted and the locations of the wearer can be estimated with an adequate accuracy. However, in the current method, the images were assumed to be produced by a pinhole camera model and the movement of the camera was limited to moving forward, or rotation only. In the experiment, the camera was held to face the front of the wearer to simplify the image processing procedure. In the next step, we will improve the projection model to handle more complicated movements of the wearer and further improve the image processing algorithm to make the method more practical.

### Acknowledgment

This work was supported by the National Institutes of Health Grants No. P30AG024827, U48DP001918, R01CA165255, R21CA172864, the Claude Pepper Center for Aging and the Health Promotion, the Disease Prevention Research Center, University of Pittsburgh, and the China Scholarship Council (201406025003).

### References

1. Otsason, V.; Varshavsky, A.; LaMarca, A.; De Lara, E. Accurate GSM Indoor Localization; Proc. of 7th International Conference on Ubiquitous Computing; September 11–14, 2005; p. 141-158.
2. Farid Z, Nordin R, Ismail M. Recent advances in wireless indoor localization techniques and system. Journal of Computer Networks and Communications. 2013; 2013:185138.
3. Chan M, Esteve D, Escriba C, Campo E. A review of smart homes- present state and future challenges. Comput Methods Programs Biomed. 2008; 91:55–81. [PubMed: 18367286]
4. Tian Y, Hamel WR, Tan J. Accurate human navigation using wearable monocular visual and inertial sensors. IEEE Trans. Instrum. Meas. 2014; 63(1):203–213.

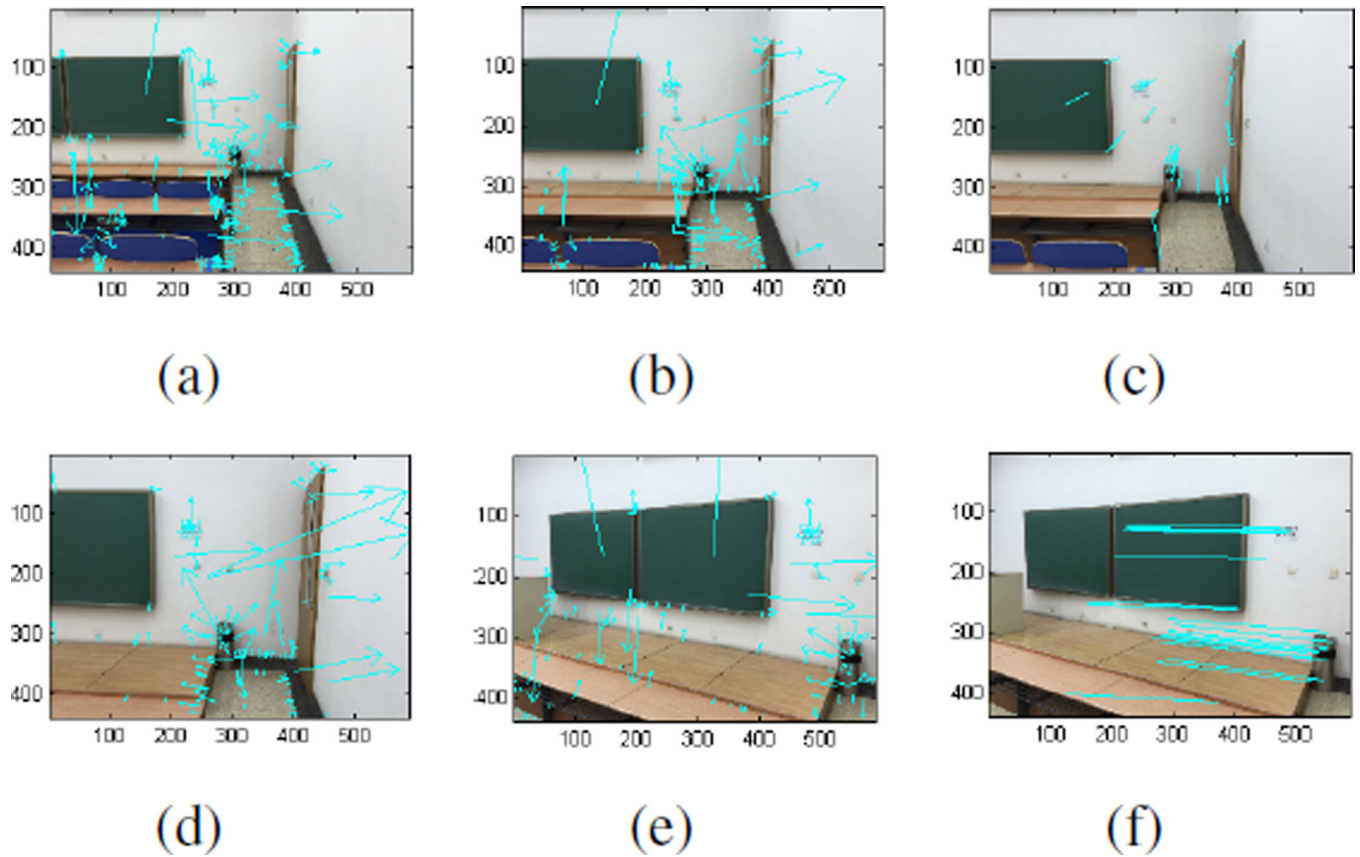
5. Lowe, DG. Object recognition from local scale-invariant features. Proc. of the seventh IEEE international conference on Computer vision; Kerkyra. September 20–27, 1999; p. 1150-1157.

Author Manuscript

Author Manuscript

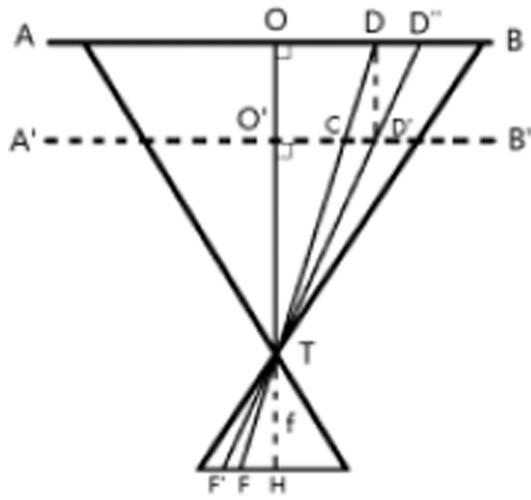
Author Manuscript

Author Manuscript

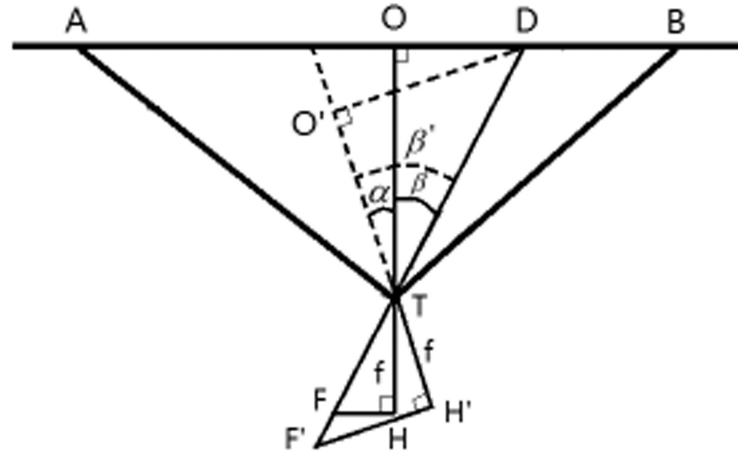


**Fig. 1.**

(a), (b) two adjacent images with SIFT features corresponding to the forward movement of the camera; (c) displacement of the key points from (a) to (b); (d),(e) two pictures with SIFT feature corresponding to the rotation of the camera; (f) displacement of the key points from (d) to (e)

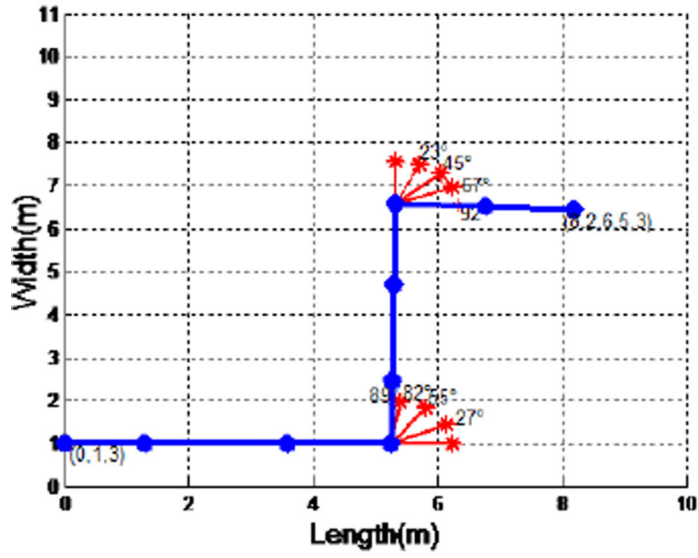


(a)

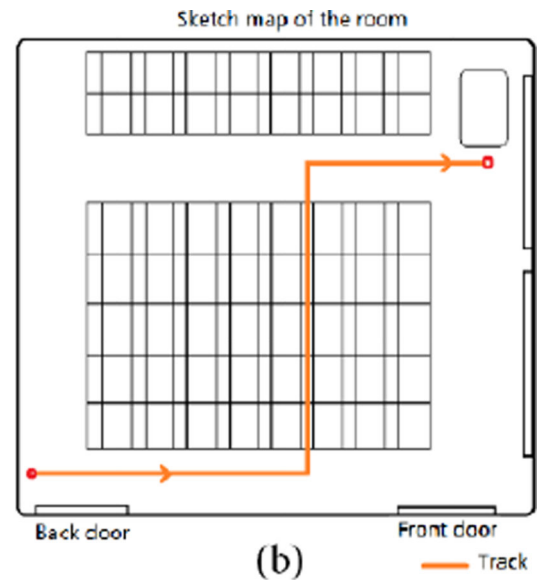


(b)

**Fig. 2.** Projection model when the camera moves forward(a) and rotates(b)



(a)



(b)

Fig. 3. Comparison of the calculated route(a)and the pre-defined route(b)

**Table I**

## Calculated Moving Distances And Rotation Angles

Moving forward	Calculated Distance(m)	Actual Distance(m)	Absolute Error(m)	Relative Error
part1	5.2466	5.00	0.2466	4.93%
part3	5.5749	6.00	-0.4251	-7.08%
part5	2.8858	3.00	-0.1142	-3.80%
Rotation	Calculated Angle(°)	Actual Angle(°)	Absolute Error(°)	Relative Error
part2	89.1890	90.00	-0.811	-0.90%
part4	92.4038	90.00	2.4038	2.67%

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript