Blind User Indoor Environment Mapping System for Self-Localization

S.S.Samsudin, N.Mokhtar, H.Arof, F.Ibrahim
Department of Electrical
University of Malaya
50603 Kuala Lumpur, Malaysia
Email: ctsalbiah@hotmail.com,
norrimamokhtar@um.edu.my, ahamzah@um.edu.my,

M.Iwahashi
Faculty of Engineering
Nagaoka University of Technology
Niigata, 940-2188, Japan
Fatimah@um.edu.my
Email: iwahashi@vos.nagaokaut.ac.jp

Abstract: **This paper described blind user indoor environment mapping system for self-localization via webcam camera. The method used is based on robot navigation systems which capture and build panorama map of the floor. The map is then incrementally update its position as it moves throughout an environment. Image registration techniques are applied to measure translation distance and rotation angle between consecutive frames. We used Phase only correlation (POC) method to matched image differs only by a translation. For rotated image, RI-POC and Radon Transform is used to recovered rotation. The accuracy of these two methods for rotated image is compared for best result.**

Keywords: Mapping, Image registration, computer vision.

I. Introduction

Various efforts have been done to aid the disable community via Human-Machine Interface. Brain signal, eye movement [1], speech signal are among input use for Human-Machine Interface application. As part of the disable community the blind user also need to be assisted for self-localization.

With impairment or failure of vision, blind people must shift their attention to other senses to obtain information about the environment. Those who are totally blind depend entirely on hearing, touch, smell, and the vestibular sense to perceive, interact with, and move about their environments. The traditional white cane is still the most popular device for the blind to preview the area ahead of them as they walk. But it is not suitable to all situations or to all users.

Due to the advanced development of the high-speed computers and sensory devices, it is possible to design sophisticated equipment for vision substitution to help blind user walk safely and independently. Recent research efforts are in progress to produce new navigation systems in which digital video cameras are used as vision sensors.

This paper describes the mapping methods use to localize blind person in indoor environment using webcam camera. We use image registration methods to create a map of the environment. Images taken at different interval of time is matches to make a panorama map of the floor [2].

The algorithms described are used to estimates transformation parameter of the register images misaligned due to translation, rotation and scale [3]. We use Phase only Correlation (POC) based Fourier transform method to calculate translation between two images which also define the motion of the blind.

However POC method is sensitive to the image rotation and only calculates translation [4]. To calculate rotation angle, we use Rotation-Invariant Phase only Correlation (RI-POC) and Radon transform. These two methods are compared to select the best matching method.

II. System Architecture

In order to provide a blind user vision substitution, the system must be able to locate the position of the person and avoid any obstacle near them in real time [5]. The system also must use compact hardware that consumes little power to satisfy the constraints on power consumption of the device. Information of vision has to be transferred to another sensory organ to guide the blind person walk around unknown environment. In this case we use auditory organ to provide sounds which tell about the environment around them. Figure 1 show the flowchart of the system.

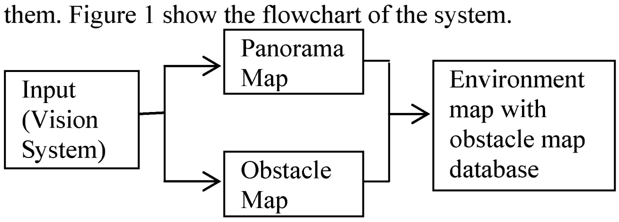


Figure 1. Flowchart of the blind user assistant

We used webcam camera located above the ear and near to the eye of the person to capture the image of the floor. This image will be matches as the blind user walk to create a panorama map of the floor. We use floor map instead of ceiling map asthe previous work [4, 6, 7], because most of the obstacle are locate on the ground. It is necessary for the blind person to avoid any obstacle near them.

All the process of localization and obstacle avoidance will be perform by a processor mounted at the back of the person. The environment map which locates the position and obstacle of the environment will be converted to sound to navigate the blind person safely. However, in this paper we only focus to discuss localization algorithms.

III. Localization

For blind person, way finding is dependent on the ability to remain localized and oriented. Before the disable user can be navigate, localization is necessary for motion planning.

Here we described the algorithms to make a panorama map using POC to matches images. But when the blind user make a rotation as they walk, POC will result error. Hence

978-1-61284-486-2/11/$26.00 ©2011 IEEE

in this case, we used RI-POC and Radon transform to translate the rotated image to translation.

A. Phase only Correlation (POC)

Functions in frequency domain are described in magnitude and phase. But sometimes, phase information is discarded such as in power spectrum where only magnitude in Fourier domain is displayed. However, it is found that in preserving the features of an image pattern, phase information is more important than the magnitude in Fourier domain [8].

Inspired by the above findings, the use of phase only information for image matching[9] and texture pattern have been investigated. It is found that in many application, phase only correlation will give a sharper correlation graph.

Consider two N_1 x N_2 images, $f(n_1, n_2)$ and $g(n_1, n_2)$ where we assume that the index range are $n_1=-M_1 \ldots\ldots M_1 (M_1>0)$ and $n_2=-M_2 \ldots\ldots M_2 (M_2>0)$ for mathematical simplicity, and hence $N_1=2M_1+1$ and $N_2=2M_2+1$ [9]. Let $F(k_1,k_2)$ and $G(k_1,k_2)$ denote the two dimension discrete Fourier transforms(2D DFT) of the two images. $F(k_1,k_2)$ and $G(k_1,k_2)$ are given by

$$F(k_1,k_2) = \sum_{n_1,n_2} f(n_1, n_2)\ W_{N_1}^{K_1 n_1} W_{N_2}^{K_2 n_2}$$

$$= A_F(K_1, K_2) e^{j\theta_F(K_1,K_2)}$$

$$G(k_1,k_2) = \sum_{n_1,n_2} g(n_1, n_2)\ W_{N_1}^{K_1 n_1} W_{N_2}^{K_2 n_2}$$

$$= A_G(K_1, K_2) e^{j\theta_G(K_1,K_2)}$$

Where $k_1 = -M_1 \ldots M_1$, $k_2 = -M_2 \ldots M_2$, $W_{N_1} = \exp(\frac{-j2\pi}{N_1}) W_{N_2} = \exp(\frac{-j2\pi}{N_2})$ and the operator $\sum n_1,n_2$ denotes $\sum_{n_1}^{M_1} = -M_1 \sum_{n_2}^{M_2} = -M_2$. $A_F(K_1, K_2)$ and $A_G(K_1, K_2)$ are amplitude components and $e^{j\theta_F(K_1,K_2)}$ and $e^{j\theta_G(K_1,K_2)}$ are phase components. The cross spectrum $R_{FG}(k_1,k_2)$ between $F(k_1,k_2)$ and $G(k_1,k_2)$ is given by

$$R_{FG}(k_1,k_2) = F(k_1,k_2)\overline{G(K_1, k_2)}$$

$$= A_F(k_1,k_2) A_G(k_1,k_2)\ e^{j\theta(k_1, k_2)}$$

Where $\overline{G(k_1, k_2)}$ denotes the complex conjugate of $G(k_1, k_2)$ and $\theta(k_1, k_2)$ denotes the phase difference $\theta_F(k_1, k_2) - \theta_G(k_1, k_2)$. The ordinary correlation function is given by the two dimension inverse discrete Fourier transform (IDFT) of $R_{FG}(k_1, k_2)$ and is given by

$$r_{fg}(n_1, n_2) = \frac{1}{N_1 N_2} \sum_{k_1,k_2} R_{FG}(n_1, n_2) W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2},$$

Is the 2D inverse Fourier transformsof $R_{FG}(k_1, k_2)$

Where $\sum_{k_1 k_2}$ denotes $\sum_{k_1}^{M_1} = -M_1 \sum_{k_2}^{M_2} = -M_2$. On the other hand, the cross phase spectrum $\widehat{R_{FG}}(k_1, k_2)$ is defined as

$$\widehat{R}_{FG}(k_1, k_2) = \frac{F(k_1, k_2)\overline{G(k_1, k_2)}}{\left|F(k_1, k_2)\overline{G(k_1, k_2)}\right|}$$

$$= e^{j\theta(k_1, k_2)}$$

The phase only correlation (POC) function $\hat{r}_{fg}(n_1, n_2)$ is the 2D IDFT of $\widehat{R}_{FG}(k_1, k_2)$ and is given by

$$\hat{r}_{fg}(n_1, n_2) = \frac{1}{N_1 N_2} \sum_{k_1 k_2} \widehat{R}_{FG}(k_1, k_2) W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2}$$

When $f(n_1, n_2)$ and $g(n_1, n_2)$ are the same image, i.e, $f(n_1, n_2) = g(n_1, n_2)$, the POC function will be given by

$$\hat{r}_{ff}(n_1, n_2) = \frac{1}{N_1 N_2} \sum_{k_1 k_2} W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2}$$

$$= \partial(n_1, n_2)$$

$$= \begin{cases} 1 & \text{if} \quad n_1 = n_2 = 0 \\ 0 & \quad \text{otherwise} \end{cases}$$

The equation above implies that the POC function between two identical images is the kronecker's delta function $\partial(n_1, n_2)$ [9].

In this paper, POC function is used to match images to make a panorama map of the floor. POC have the advantages where it gives a distinct sharp peak when two images are similar compare to ordinary correlation.

Other properties of POC method used for image matching are that it is not influenced by image shift and brightness change and it is highly robust against noise [9].

B. Rotational-invariant Phase only Correlation (RI-POC)

Pattern matching algorithms experience difficulties in judging matching when distinguishing among figures that have been rotated. Rotation-invariant phase only correlation (RI-POC) can measures rotation, translation and scaling between two images.

To estimate translation, rotation and scaling using correlation, assume two given image I_1 and I_2 related by a translation **t**, rotation $\theta_0 \in [0,2\pi)$ and scaling $s > 0$, that is [10]:

$$I_2(x) = I_1(Dx + t)$$

Where $D = s\theta$ and $\theta = \begin{bmatrix} \cos\theta_0 & sin\theta_0 \\ -sin\theta_0 & cos\theta_0 \end{bmatrix}$

In Fourier domain, it holds [10]

$$\widehat{I_2}(k) = \left(\frac{1}{|\Delta|}\right)\widehat{I_1}(k')e^{jk'^T t}$$

Where $k' = D^{-T}k$ and Δ is the determinant of D. Taking the magnitude in both parts and substituting $D^{-T} = \theta/s$, $\Delta = s^2$ gives

$$I_2(k) = \left(\frac{1}{|\Delta|}\right)I_1(K')$$

$$I_2(k) = \left(\frac{1}{S^2}\right)I_1(\theta k/s)$$

Using the log-polar representation yields (ignoring the term $1/s^2$) [10]

$$(k_i) = M_1(k_l - [\log s, \theta_0]^T)$$

From equation above explain that in the log-polar Fourier magnitude domain, the rotation and scaling reduce to 2D translation which is then can be estimated using POC. Figure 2 show the flowchart of RI-POC method.

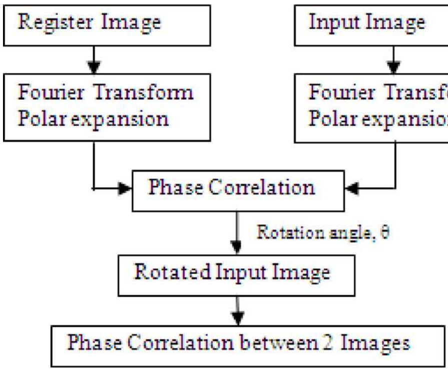


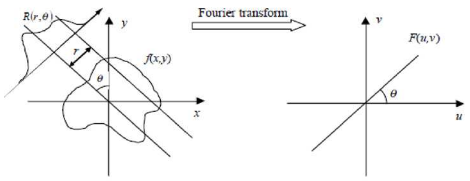
Figure 2. Flowchart for RI-POC

C. Radon Transform

The Radon transform of a 2D function *f(x,y)*, is defined as its linear integral along a line. The integration along a line is defined by a normal distance *r* from the origin and normal angle θ will give corresponding Radon transform point *R(r,θ)*[11]. Mathematically, it can be written as

$$R(r,\theta)[f(x,y)] = \iint_{-\infty}^{\infty} f(x,y)\partial(r - x\cos\theta - y\sin\theta)dxdy$$

Where $-\infty < r < +\infty$, $0 < \theta < \pi$. This transformation is invertible according to the Fourier slice theorem. It states that for 2D image *f(x,y)*, the 1D Fourier transform of the Radon transform along *r*, are the 1D radial samples of the 2D Fourier transform of *f(x,y)* at the corresponding angles [11].



(a) (b)

Figure 3. (a) Radon transform of the image. (b) 1-D Fourier transforms of the projection construct the 2-D Fourier transform of the image.

Radon transform projects the image to the other parameter space. Using the property of Radon transform, rotation of an image by an angle θ_r, cause the Radon transform to be shifted through the same amount [11],

$$f(x\cos\theta_r - y\sin\theta_r, x\sin\theta_r + y\sin\theta_r) \leftrightarrow Rf(r, \theta - \theta_r)$$

In this paper, Radon transform is use to convert rotation to translation. This is the same as RI-POC method to measure the rotation angle between images. After a rotated image is calculated, POC is use matched the translated image [12]. Radon transform is then used to compare the accuracy between these two methods. Figure 4 give the flowchart for radon transform method.

Register Image Input Image

Radon Transform Radon Transform

Phase Correlation

Rotation angle, θ

Rotated Input Image

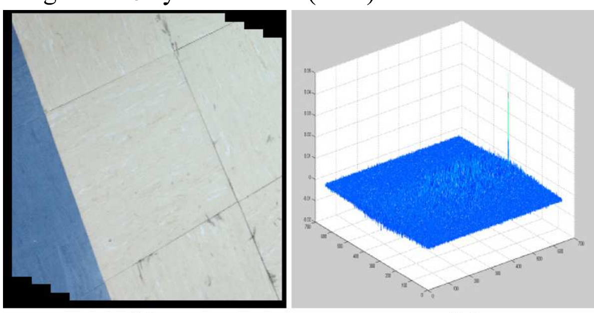
Phase Correlation between 2 Images

Figure 4. Flowchart for Radon Transform

IV. Result and Discussion

In this section, we describe a set of experiment simulation using MATLAB for evaluating matching floor map of the proposed algorithms.

A few sample of corridor floor is capture and analyzed for matching. Figure 5(a) show the result of matching four frame of tile floor image and figure 5(b) show the sharp and distinct correlation peak of matching two frame of image by using Phase only Correlation (POC).



(a) (b)

Figure 5. (a) Panorama map of the floor and (b) correlation graph.

The images match common area of each frame and make a wide view of panorama map of the floor. The height of the peak of the correlation graph shows a good similarity measure for image matching.

As discuss above, POC only calculate register image misaligned due to translation. For rotated image RI-POC and Radon transform method are compared which translate rotation to translation. To compare these two methods, four different floor images is taken and tested at different angle from 0^0 to 180^0. The average error of each floor is plot on the histogram as in figure 6.

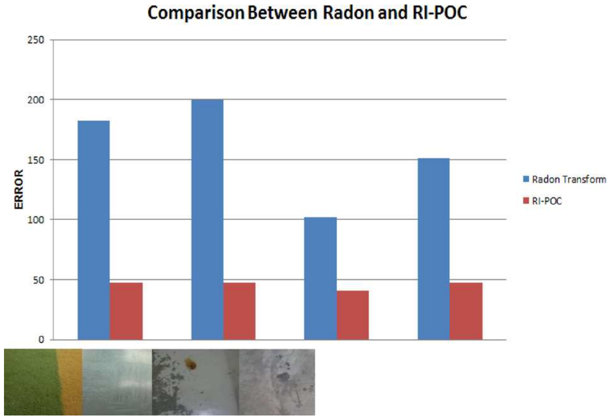


Figure 6. Comparison between Radon and RI-POC method

Radon transform show the largest error compare to RI-POC for the tested image floor. Hence it is proves that Fourier Transform in log polar transform measure rotation better than Radon transform. Moreover RI-POC computes faster than Radon transform. A faster algorithm is needed as the environment map has to be update as the blind walk.

V. Conclusion and Future Research

In this paper we only tested the matching algorithms to be used for blind user assistant for mapping purposes and full system is still under development. From the result show that POC give a sharp correlation graph when two similar common areas is matched. Hence it is suitable to be used on blind user assistant to create a wide view of environment map. We used RI-POC to recover rotated image as it show less error compared to Radon transform.

Acknowledgment

This project is funded by University of Malaya Research Grant (UMRG) RG 025/09AET and Institute of Research Management and Monitoring under PPP grant number PS144-2010A.

References

[1] N.Mokhtar, H.Arof and M.Iwahashi. One Dimensional Image Processing for Eye Tracking using Derivative Dynamic Time Warping. *Scientific Research and Essays Vol.5 (19),pp2947-2952,4 october, 2010.*

[2] Barbara Zitova, Jan Flusser, “Image Regisration Method: a survey”, *Image and Vision Computing 21 (2003) 977-1000.*

[3] George Wolberg, SiavashZokai, “Robust Image Registration Using Log-Polar Transform”, *IEEE conf. on Image Processing, Sep 2000.*

[4] Y. Fukutani, T. Takahashi, M. Iwahashi, T. Kimura, S. S. Salbiah, N. Mokhtar, “Robot Vision Network based on Ceiling Map Sharing”, *The 11th IEEE International Workshop on Advanced Motion Control March 21-24, 2010.*

[5] EdwigePissaloux, “A Characterization of Vision Systems for Blind People Mobility”, *2002 IEEE SMC.*

[6] T.Matsumoto, T.Takahashi, M.Iwahashi, T.Kimura, S.Salbiah and N.Mokhtar. Visual Compensation in Localization of a Robot on a Ceiling Map.*Scientific Reseach and Essays Vol.5(23),4 December, 2010.*

[7] T.Matsumoto, T.Takahashi, M.Iwahashi, T.Kimura, S.Salbiah and N.Mokhtar. Reduction of Data Size for Transmission in Localization of Mobile Robots. *International Journal of the Physical Sciences Vol.5(15),4 December 2010.*

[8] Jia-Zhu Wang, Lawrence E Reinstein, Joseph Hanley, Allen G Meek, “Investigation of a Phase only Correlation Technique for Anatomical Alignment of Portal Image in Radiation Therapy”, *Phys. Med. Biol.* ***41*** *(1996) 1045–1058.*

[9] K. Ito, H. Nakajima, K. Kobayashi,T. Aoki,T. Higuchi, “A Fingerprint Matching Algorithm Using Phase-Only Correlation, *IEICE Trans. Fundamentals, Vol.E87-A,NO.3 March 2004.*

[10] GeorgiosTzimiropoulos and Tania Stathaki, “Robust FFT-Based Scale-Invariant Image Registration”, *4th SEAS DTC Technical Conferences.*

[11] S. S. Xiao, Y. X. Wu, “Rotation-invariant Texture Analysis Using Radon and Fourier Transform”, *Journal of Physics: Conference Series 48 (2006) 1459-119.*

[12] Christopher M. Gifford, “Low-Cost Mobile Robot Localization Using Only a Download-Facing Webcam”, *Technical Report October 13,2009.*