

# How to Prepare a Camera-Ready Paper for Picture Coding Symposium of Japan and Image Media Processing Symposium

Takanori Fujisawa<sup>†</sup>

Masaaki Ikehara<sup>†</sup>

<sup>†</sup>EEE Department, Keio University

**Abstract:** Put your English abstract in 100 words. This paper proposes a novel method to estimate non-integer shift of images based on least squares approximation in the phase region. Conventional methods based on Phase Only Correlation (POC) take correlation between an image and its shifted image, and then estimate the non-integer shift by fitting the model equation. The problem when estimating using POC is that the estimated peak of the fitted model equation may not match the true peak of the POC function. This causes error in non-integer shift estimation. By calculating the phase difference directly in the phase region, the proposed method allows the estimation of sub-pixel shift through least squares approximation. Also by utilizing the characteristics of natural images, the proposed method limits adoption range for least squares approximation. By these improvements, the proposed method achieves high accuracy, and we validate through some examples.

## 1 Length and Typeset Parameters

- up to two pages including figures and tables
- A4 size
- text area width of 180mm
- text area height of 252mm
- baseline skip of 8mm

## 2 Introduction

With the recent advancement of hardware technology, large resolution sensors capable of taking images and videos with huge data sizes have become more common. To utilize these huge data, the needs for technologies like video encoding, signal processing, and object recognition have increased. For example to handle these huge data sized videos, encoding technologies like MPEG-2 and MPEG-4 have been introduced. These technologies encode videos by finding the most similar patch from the frames before. And to find the most similar patch, shift estimation technologies are used. Also, shift estimation methods are very important base technology of object recognition. For example, when manufacturing semiconductors, precise placements of components is very important, and it requires very accurate shift estimation for efficient production. Therefore, the improvements of shift estimations will have a large impact.

Majority of shift estimation methods use phase only correlation of the two input images, which shows sharp peak corresponding to the shift between the images. The sharp peak estimated by POC shows shift on integer accuracy easily. Therefore, main focus of current research has been the estimation of non-integer shift. These include methods which use calculation in frequency domain [1–4], estimation of the true peak by using model equation fitting [5–8], and estimation based on Discrete Cosine Transform called DCT-SPC [9, 10]. Although some methods which use model equation to estimate the true peak of the POC may have high accuracy, they have possibility that the true peak and the estimated peak of the model equation don't match. Therefore, it can be the cause of estimation error.

In this paper, we propose a new method to estimate image shift on non-integer accuracy based on least squares method in phase region. The proposed method estimates shift directly in the frequency domain by calculating the slope of the phase re-

sponse which corresponds to the shift of the images. However, the obtained phase component includes many discontinuity, because of calculation of  $\tan^{-1}$ . These discontinuities prevents us from estimating the slope which corresponds to the shift value. The proposed method unwraps the phase component by subtracting the phase component of the integer shift value from the original phase component. This way, the proposed method obtains a smooth continuous phase component, which allows estimation of the slope corresponding to the decimal shift directly from phase component. Consequently, we avoid the cause of error in the conventional POC based method without using the model equation.

## 3 Conventional method

In this section, we explain the POC based shift estimation. The POC takes the correlation of two images in frequency domain to obtain the peak which corresponds to the shift value between the images.

### 3.1 Phase Only Correlation

First we consider an image  $f(x, y)$  sized  $N_1 \times N_2$  and an image  $g(x, y)$  which is  $(\delta_1, \delta_2)$  shifted image of  $f(x, y)$  in parallel as they are defined below.

$$x \in \{0, 1, 2, \dots, N_1 - 1\} \quad (1)$$

$$y \in \{0, 1, 2, \dots, N_2 - 1\} \quad (2)$$

#### 3.1.1 Phase Only Correlation

First we consider an image  $f(x, y)$  sized  $N_1 \times N_2$  and an image  $g(x, y)$  which is  $(\delta_1, \delta_2)$  shifted image of  $f(x, y)$  in parallel as they are defined below.

$$x \in \{0, 1, 2, \dots, N_1 - 1\} \quad (3)$$

$$y \in \{0, 1, 2, \dots, N_2 - 1\} \quad (4)$$

Also Fourier transforms of the two images  $f(x, y)$  and  $g(x, y)$  are given by

$$\begin{aligned} F(k_1, k_2) &= \sum_{x=0}^{N_1-1} \sum_{y=0}^{N_2-1} f(x, y) e^{-j2\pi(xk_1/N_1 + yk_2/N_2)} \\ &= A(k_1, k_2) e^{j\theta_1(k_1, k_2)} \end{aligned} \quad (5)$$

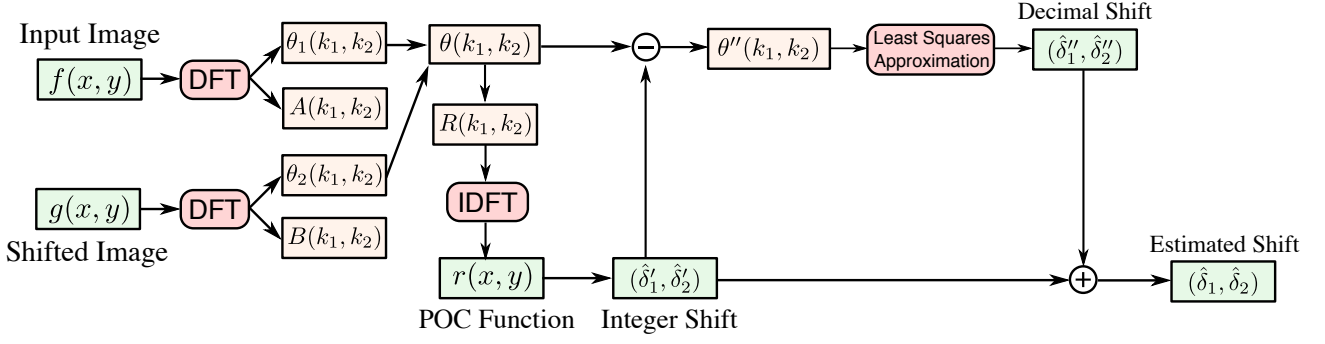


Figure 1: Flow chart of the proposed method

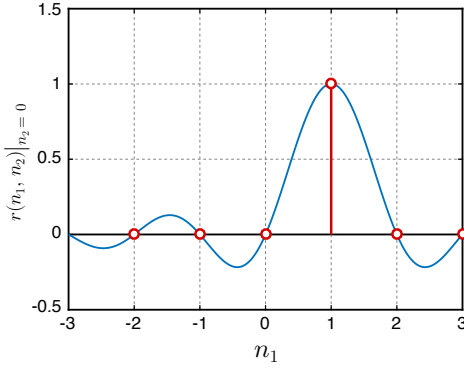


Figure 2: Example of POC function  $r$  with integer shift

$$G(k_1, k_2) = \sum_{x=0}^{N_1-1} \sum_{y=0}^{N_2-1} g(x, y) e^{-j2\pi(xk_1/N_1 + yk_2/N_2)} = B(k_1, k_2) e^{j\theta_2(k_1, k_2)} \quad (6)$$

where  $A(k_1, k_2)$  and  $B(k_1, k_2)$  represent the amplitude components,  $e^{j\theta_i(k_1, k_2)}$  are the phase components of their images, and  $j$  denotes an imaginary unit. Now, a normalized cross power spectrum  $R$  is given by

$$R(k_1, k_2) = \frac{F(k_1, k_2) G^*(k_1, k_2)}{|F(k_1, k_2) G^*(k_1, k_2)|} \quad (7)$$

where  $G^*(k_1, k_2)$  denotes a complex conjugate of  $G(k_1, k_2)$ . Therefore  $R$  can be represented as follows from (5) and (6)

$$R(k_1, k_2) = e^{j\theta(k_1, k_2)} \quad (8)$$

where  $\theta(k_1, k_2) = \theta_1(k_1, k_2) - \theta_2(k_1, k_2)$ , which is the phase difference of the two input images. Then, the POC function  $r(x, y)$  of the two images are given by 2-D inverse discrete Fourier transform of  $R(k_1, k_2)$  by

$$r(x, y) = \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} R(k_1, k_2) e^{j2\pi(k_1 x/N_1 + k_2 y/N_2)} \quad (9)$$

Fig. 2 shows the example of POC with peak indicating integer shift value  $(\delta_1, \delta_2)$ .

### 3.2 Non-integer shift estimation

The peak of the POC function  $r(x, y)$  only shows the integer shift  $(\delta_1', \delta_2')$  of the images. To estimate the non-integer shift

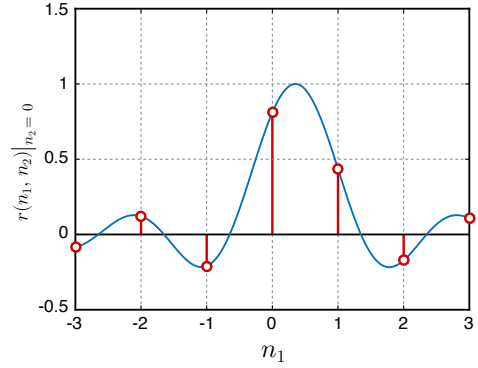


Figure 3: Example of POC function  $r$  with non-integer shift and ideal model equation

of the images, estimating the true peak of the POC function is necessary like in Fig. 3. Conventional methods use the model equation fitting to estimate the true peak. The conventional methods [6, 9, 10] use the model equation shown as follows:

$$r(x, y) \simeq \frac{1}{N_1 N_2} \frac{\sin \pi(x + \delta_1)}{\sin \frac{\pi}{N_1}(x + \delta_1)} \frac{\sin \pi(y + \delta_2)}{\sin \frac{\pi}{N_2}(y + \delta_2)}. \quad (10)$$

### 3.3 Problem with conventional method

In order for the conventional methods to estimate sub-pixel shift, conventional methods need to fit the model equation to the calculated POC function  $r(x, y)$ . However, it is not certain that the estimated peak of the model equation matches the true peak of the POC function  $r(x, y)$ . Also fitting of model equation requires multiple calculations, which leads to larger computational time. By avoiding the fitting of the model equation, the proposed method estimates the shift directly from the phase components of the two images.

## 4 Format Outline

- abstract of about 100 words
- manuscript can be either one-column or two-column.
- put the contact information at the bottom-right of the last page. (You may want to use sample macro `\simplefootnotetext`).

## 5 Font Sizes

Following font sizes are recommended.

**Title** 16pt (`\LARGE`)

Table 1: Encoder Settings.

configuration 1	aa	aaa
configuration 2	bb	bbb
configuration 3	cc	ccc

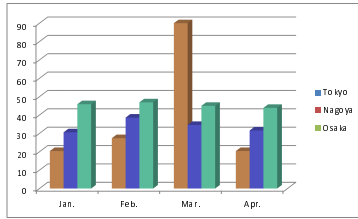


Figure 4: Coding results.

**Author names** 12pt(\large)

**Affiliation and text** 10pt(\normalsize)

## 6 Tables and Figures

Please be aware that your manuscript will be printed in monochrome, halftone dot meshing.

Figures and tables shall be placed at the top or bottom of a column wherever possible. For  $\text{\TeX}$  floats, the placement parameter shall be either [t] or [b]. [h] shall be used sparingly, or not at all. Refer to them with e.g., Fig. 4 and Tab. 1.

## 7 Other $\text{\LaTeX}$ 2 $\epsilon$ -Specific Tips

You may use newtx [9] package etc. to have more beautiful typesetting.

To show URLs, consider the use of url style file with `\url{http://...}`.

You may use hyperref package [?] to build a PDF file that contains hyperlink information.

## 8 Conclusion

2014.05.13 Initial Version

2015.07.15 Revised Version

## References

- [1] M. Balci and H. Foroosh, "Subpixel estimation of shifts directly in the Fourier domain," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 1965–1972, 2006.
- [2] W. S. Hoge, "A subspace identification extension to the phase correlation method [MRI application]," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 277–280, 2003.
- [3] S. L. Kilthau, M. S. Drew, and T. Moller, "Full search content independent block matching based on the fast Fourier transform," in *Proceedings. International Conference on Image Processing*, vol. 1. IEEE, 2002, pp. 1–669.

- [4] B. S. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE transactions on Image processing*, vol. 5, no. 8, pp. 1266–1271, 1996.
- [5] S. Nagashima and T. Aoki, "Improved performance for subpixel image matching based on phase-only correlation," 2004.
- [6] K. Takita, Y. Sasaki, T. Higuchi, and K. Kobayashi, "High-accuracy subpixel image registration based on phase-only correlation," *IEICE transactions on fundamentals of electronics, communications and computer sciences*, vol. 86, no. 8, pp. 1925–1934, 2003.
- [7] K. Kobayashi, H. Nakajima, T. Aoki, M. Kawamata, and T. Higuchi, "Principles of phase only correlation and its applications," *ITE Technical Report*, vol. 20, no. 41, pp. 1–6, 1996.
- [8] T. Dobashi and H. Kiya, "A parallel implementation method of FFT-based full-search block matching algorithms," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2013*. IEEE, 2013, pp. 2644–2648.
- [9] I. Izumi and K. Hitoshi, "Fitting functions for noninteger shift value estimation using DCT sign phase correlation," *IEICE transactions on fundamentals of electronics, communications and computer sciences*, vol. 92, no. 3, pp. 172–181, 2009.
- [10] I. Ito and H. Kiya, "Multiple-peak model fitting function for DCT sign phase correlation with non-integer shift precision," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2009*. IEEE, 2009, pp. 449–452.

---

*Ikehara Laboratory, EEE Department, Keio University  
Kohoku-ku, Yokohama, Kanagawa, 223–8522, Japan  
E-mail: ikehara@tkhm.elec.keio.ac.jp*