Sub-pixel Shift Estimation of Image based on the Least Squares Approximation in Phase Region

Ryo Fujimoto, Takanori Fujisawa, Masaaki Ikehara
EEE Dept., Keio Univ., Yokohama, Kanagawa 223-8522, Japan
Email: {fujimoto,fujisawa,ikehara}@tkhm.elec.keio.ac.jp

***Abstract*—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 a image and its shifted image, and then estimate the non-integer shift by fitting the model equation. The problem with using POC is that the true peak of the POC function may not match the estimated peak of the fitted model equation. This causes error in non-integer shift estimation. By calculating directly in the phase region, the proposed method allows the estimation of decimal shift through least squares approximation. Also by utilizing the characteristics of the natural image, the proposed method limits adoption range for least squares approximation. By these improvements, the proposed method improves the estimation and achieves high accuracy.**

I. 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. With these huge data, the needs for technologies like video encoding, signal processing, and pattern 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. Since the shift estimation is the basic part of many technologies, the improvements will have a large impact.

Major methods of shift estimation uses Phase Only Correlation (POC) of the two input images, which shows sharp peak corresponding to the shift. By using POC, shift estimation on integer accuracy can be done easily. So estimation with non-integer accuracy has gathered attention recently. These include methods which use calculation in frequency domain [1]–[4], estimation of the true peak by using model equation fitting [5]–[8], and DCT coefficients called DCT-SPC [9], [10]. Although 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. As a result, it causes estimation error.

In this paper, we propose a new method to estimate image shift on non-integer accuracy. The proposed method estimates shift directly in the frequency domain by calculating the slope of the phase response. However, the obtained phase component includes many discontinuity, because of calculation of tan^{-1}. This discontinuity 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. Then we obtain a smooth continuous phase component, which allows estimation of the slope corresponding to the decimal shift directly in phase region. Consequently, we avoid the cause of error in the conventional POC based method without using the model equation.

The rest of this paper is organized as follows. In section II, we will explain the conventional POC. Then in section III, we explain the details of the proposed method. Then in section IV, we test the proposed method and compare it to other conventional methods through simulation. Lastly in section V we present conclusion.

II. 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.

A. Phase Only Correlation

First we consider an image $f(x, y)$ sized $M \times N$ and an image $g(x, y)$ which is (δ_1, δ_2) shifted image of $f(x, y)$ in parallel as they are defined below.

$$x \in \{0, 1, 2, \dots, M-1\} \tag{1}$$
$$y \in \{0, 1, 2, \dots, N-1\} \tag{2}$$

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

$$\begin{aligned} F(k_1, k_2) &= \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(xk_1/M + yk_2/N)} \\ &= A(k_1, k_2) e^{j\theta_1(k_1, k_2)} \end{aligned} \tag{3}$$

$$\begin{aligned} G(k_1, k_2) &= \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} g(x, y) e^{-j2\pi(xk_1/M + yk_2/N)} \\ &= B(k_1, k_2) e^{j\theta_2(k_1, k_2)} \end{aligned} \tag{4}$$

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,

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

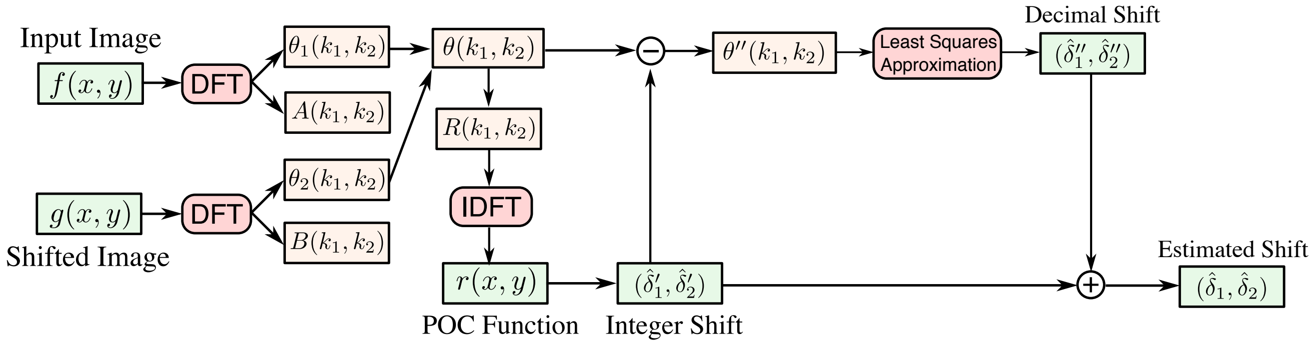

Fig. 1. Flow chart of the proposed method

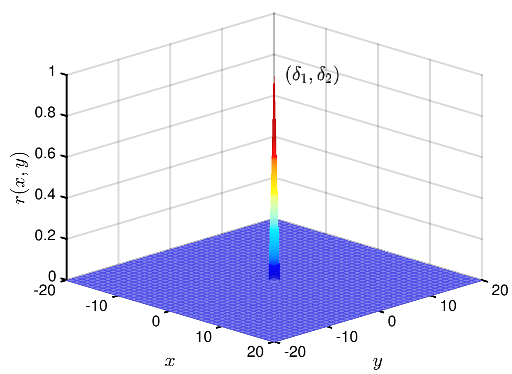

Fig. 2. Example of POC function r with shift value (δ_1,δ_2)

and j denotes an imaginary unit. Now, a normalized cross power spectrum R is given by

$$R(k_1,k_2)=\frac{F(k_1,k_2)G^*(k_1,k_2)}{|F(k_1,k_2)G^*(k_1,k_2)|} \tag{5}$$

where $G^*(k_1,k_2)$ denotes a complex conjugate of $G(k_1,k_2)$. Therefore R can be represented as follows from equation (3) and (4)

$$R(k_1,k_2)=e^{j\theta(k_1,k_2)} \tag{6}$$

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

$$r(x,y)=\frac{1}{MN}\sum_{k_1=0}^{M-1}\sum_{k_2=0}^{N-1}R(k_1,k_2)e^{j2\pi(k_1x/M+k_2y/N)} \tag{7}$$

Fig. 2 shows the example of POC with peak indicating integer shift value (δ_1,δ_2).

B. Non-Integer Shift Estimation

The peak of the POC function $r(x,y)$ only shows the integer shift (δ'_1,δ'_2) of the images. To estimate the non-integer shift of the images, estimating the true peak of the POC function is necessary. Conventional methods use the model equation fitting to estimate the true peak. The model equation used in the conventional methods is as follows.

$$r(x,y)\simeq\frac{1}{MN}\frac{\sin\pi(x+\delta_1)}{\sin\frac{\pi}{M}(x+\delta_1)}\frac{\sin\pi(y+\delta_2)}{\sin\frac{\pi}{N}(y+\delta_2)}. \tag{8}$$

C. Problem with Conventional Method

The conventional methods need to fit the model equation to the calculated POC function $r(x,y)$ to estimate the shift at the non-integer accuracy. However, it is not certain that the estimated peak of the model equation matches the true peak of the POC function $r(x,y)$. By avoiding the use of the model equation fitting, the proposed method estimates the shift directly from the phase components of the two images.

III. Proposed Method

A. Details of Proposed Method

In this section, we explain the proposed method. The flowchart of the proposed method is as shown in Fig. 1. The proposed method estimates the shift between the images, without using the fitting of the model equation. The proposed method uses the phase difference of the two images, $\theta(k_1,k_2)$ from equation (6), then approximating it as a smooth surface which corresponds to the shift of the images.

First, we look at phase difference $\theta(k_1,k_2)$ as follows.

$$\theta(k_1,k_2)=\tan^{-1}\frac{\mathrm{Im}(R(k_1,k_2))}{\mathrm{Re}(R(k_1,k_2))} \tag{9}$$

As the images are shifted in parallel, $\theta(k_1,k_2)$ can be approximated as

$$\theta(k_1,k_2)\simeq ak_1+bk_2 \tag{10}$$

where a and b denote coefficients for slope in each direction of k_1, k_2 axis. We can estimate the pair of coefficients by the following least squares method.

$$\{a,b\}=\underset{\{a,b\}}{\mathrm{argmin}}\sum_{k_1,k_2}\Big(\theta(k_1,k_2)-(ak_1+bk_2)\Big)^2 \tag{11}$$

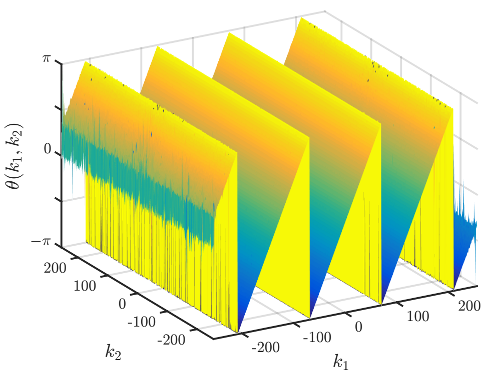

Fig. 3. Phase component before removal of integer shift

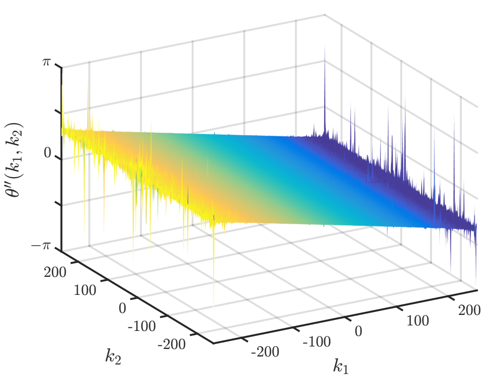

Fig. 4. Phase component with only decimal shift

As a result, we can obtain the estimated true shift $(\hat{\delta}_1, \hat{\delta}_2)$ as

$$\hat{\delta}_1 = \frac{M}{2\pi}a, \quad \hat{\delta}_2 = \frac{N}{2\pi}b. \tag{12}$$

Actually, the phase component θ we obtained at equation (9) has discontinuity shown in Fig. 3. This is due to the value of the phase component over wrapping between range of $-\pi$ to π. To unwrap this, we split the phase component to two separate components, integer and decimal components.

$$\begin{aligned}\theta(k_1, k_2) &\simeq ak_1 + bk_2 \\ &= (a' + a'')k_1 + (b' + b'')k_2\end{aligned} \tag{13}$$

here, a', a'' denote integer and decimal shifts to the k_1 direction. And b', b'' denote integer and decimal shifts to the k_2 direction. By subtracting integer shift from equation (13), we can obtain phase components with only the decimal shift. We use the conventional POC to get the slopes (a', b') of the integer shift. Then we obtain the phase component θ'' with only the decimal shift as follows

$$\begin{aligned}\theta''(k_1, k_2) &= \theta(k_1, k_2) - (a'k_1 + b'k_2) \\ &\simeq a''k_1 + b''k_2.\end{aligned} \tag{14}$$

Equation (14) means the removal of the integer shift estimated from the conventional POC from the original phase components of two images. We can solve the discontinuity

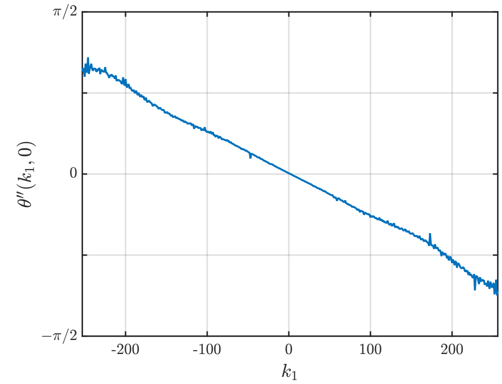

Fig. 5. Slice of phase component

and obtain phase component θ'' with only the decimal shift. Consequently, we apply the next process to obtain the decimal phase component θ'' from R in equation (6), and then apply the least squares approximation as follows.

$$\begin{aligned}e^{j\theta''(k_1,k_2)} &= e^{j(\theta(k_1,k_2) - (a'k_1 + b'k_2))} \\ &= e^{j\theta(k_1,k_2)} / e^{j(a'k_1 + b'k_2)}.\end{aligned} \tag{15}$$

$$\{a'', b''\} = \underset{\{a,b\}}{\mathrm{argmin}} \sum_{k_1,k_2} \left(\theta''(k_1, k_2) - (a''k_1 + b''k_2) \right)^2 \tag{16}$$

The phase component from this process does not cause discontinuity as shown in Fig. 4. As a result, it allows the approximation of the coefficients (a'', b'') to the direction (k_1, k_2), respectively. The proposed method focuses on the axis component of the phase component to reduce the computational cost. Then, the proposed method uses the least squares approximation (16) to estimate the coefficients. Fig. 5 shows the slice of phase component at $k_2 = 0$ of the phase component. By applying the estimated coeffients (a'', b'') to the equation (12), we can obtain the actual decimal shift $(\hat{\delta}''_1, \hat{\delta}''_2)$. We obtain the estimated true shift $(\hat{\delta}_1, \hat{\delta}_2)$ as follows, by adding the integer shift from the conventional POC.

$$\hat{\delta}_1 = \frac{M}{2\pi}a'' + \hat{\delta}'_1 \tag{17}$$

$$\hat{\delta}_2 = \frac{N}{2\pi}b'' + \hat{\delta}'_2 \tag{18}$$

B. Natural Image in Frequency Domain

Natural images tend to have most of its energy in the low frequency area and not so much on the high frequency area. In addition, high frequency area gets highly influenced by noise. Therefore, we limit the range for the least squares approximation to low frequency range to achieve better result. So in the proposed method, we limit the range $\mathcal{K}$ by parameter K_L we use in least squares approximation by $0.1N/2$ from the center of normalized frequency area.

$$\mathcal{K} = \{k_i | -K_L \leq k \leq K_L\} \tag{19}$$

(a) Original Image (b) Shifted Image

Fig. 6. Original image (Left) and non-integer shifted image (Right)

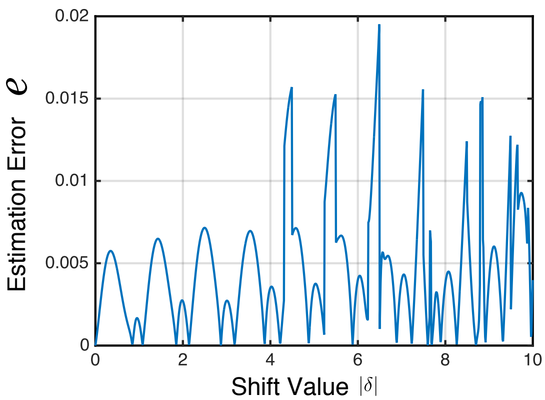

Fig. 7. Relation between estimation error e and shift value$|\delta|$

Finally, we can solve the following two equations to obtain the coeffients which represent the non-integer shifts of the images.

$$a'' = \underset{a}{\operatorname{argmin}} \sum_{k_1 \in \mathcal{K}} \left(\theta''_{k_2=0}(k_1) - ak_1 \right)^2 \tag{20}$$

$$b'' = \underset{b}{\operatorname{argmin}} \sum_{k_2 \in \mathcal{K}} \left(\theta''_{k_1=0}(k_2) - bk_2 \right)^2 \tag{21}$$

where $\theta''_{k_1=0}$ denotes the slice of the phase component at $k_1 = 0$, also same for $\theta''_{k_2=0}$.

IV. Experiments

A. Simulation

First we explain how to create the non-integer shifted images we use for the simulation. For simplicity we explain in 1-D. We consider 1-D signal $f(x)$ and its non-integer shifted signal $g(x)$ and the non-integer shift $|\delta|$. The shifted signal can be obtained by multiplying the non-integer shift $T(K)$ as shown in equation (22) in frequency domain. Therefore, DFT of the non-integer shifted signal $G(k)$ and the non-integer shift $T(k)$ are represented as follows.

$$T(k) = 2\pi k\delta/N \tag{22}$$

$$G(k) = F(k)e^{jT(k)} \tag{23}$$

where $F(k)$ denotes the DFT of original signal $f(x)$. By IDFT of $G(k)$, we can obtain the non-integer shifted signal $g(x)$. Fig. 6 shows an example of the original image and image shifted by (25.5 30.5) pixels.

Also, we applied the Hanning window in order to solve the discontinuity on the ends of the images when we perform DFTs. We use the 2-D Hanning window defined as follows.

$$\omega(x,y) = \frac{1+\cos\frac{\pi(x-M/2)}{M}}{2} \frac{1+\cos\frac{\pi(y-N/2)}{N}}{2} \tag{24}$$

B. Experimental Results

To test our methods, we used the 512×512 sized image named Lena and shifted to x axis direction from 0 to 10 pixels every 0.01 pixel. Then taken an average of the error e and evaluated the results. The error e is the absolute difference of the true shift $\delta = (\delta_1, \delta_2)$ and the estimated shift $\hat{\delta} = (\hat{\delta}_1, \hat{\delta}_2)$ as follows

$$e = |\delta - \hat{\delta}| \tag{25}$$

Also to test the proposed methods resistance to noise, we added noise intensity $\sigma = 0, 5, 10, 15, 20$.

The results are shown in Table I – III. As Table I shows, when the noise intensity σ is 0, the proposed methods estimates at very high accuracy. The proposed method has high accuracy even with noise on the image by limiting the adoption range on Fig. 5 to 0.4 – 0.5. Also, we show the estimation error e on noise $\sigma = 0$, $K_L = 256$ by shift size on Fig. 7. On Table IV, we summarized average errors of the proposed method on other test images by changing the adoption range K_L.

Next we evaluated the proposed method to the conventional POC with fitting the model equation, and DCT-SPC [9], [10]. We summarized the average error on 0–2 and 2–10, every 0.02 pixels and 0.1 pixels respectively. Also to achieve better results for conventional POC, we applied a window function and weighting in frequency domain. As Table II and III show, the proposed method performs with higher accuracy than other conventional methods while shift value is small.

V. Conclusion

In this paper, we proposed a new method to estimate the non-integer shift between two images directly in phase domain without using model equation fitting on the POC function. By using the estimated integer shift from the conventional POC to solve the discontinuity of the original phase component, we obtained the non-integer shift only phase information. Then we estimated the non-integer shift of the two images. In order to achieve better result, we limited the range for least squares approximation to utilize the characteristic of the natural image. Then we achieved high result even with noise on the image. Consequently, our method can perform shift estimation at higher accuracy than conventional methods.

References

[1] M. Balci and H. Foroosh, “Subpixel estimation of shifts directly in the fourier domain,” *Image Processing, IEEE Transactions on*, vol. 15, no. 7, pp. 1965–1972, 2006.

[2] W. S. Hoge, “A subspace identification extension to the phase correlation method [mri application],” *Medical Imaging, IEEE Transactions on*, vol. 22, no. 2, pp. 277–280, 2003.

[3] S. L. Kilthau, M. S. Drew, and T. Möller, “Full search content independent block matching based on the fast fourier transform,” in *Image Processing. 2002. Proceedings. 2002 International Conference on*, vol. 1. IEEE, 2002, pp. I–669.

TABLE I
COMPARISON OF ESTIMATION ERROR e WITH DIFFERENT NOISE LEVEL σ AND ADOPTION RANGE

Noise Level σ	parameter $\mathcal{K}_L$									
	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
0	0.00516	0.005314	0.005143	**0.003655**	0.005431	0.003917	0.007394	0.006581	0.006581	0.010086
5	0.0190	0.018566	0.018260	0.018056	0.020597	**0.014653**	0.020415	0.014881	0.014881	0.019804
10	0.0480	0.046154	0.038080	0.036739	0.034185	0.029610	**0.028724**	0.031265	0.031265	0.040383
15	0.07926	0.074630	0.064279	0.059396	0.045497	0.042581	**0.038240**	0.054100	0.054100	0.054609
20	0.102133	0.094579	0.084080	0.073349	0.070271	0.061955	**0.058595**	0.072366	0.072366	0.069065

TABLE II
COMPARISON OF ESTIMATION ERROR e OF SHIFT VALUE $0 < |\delta| < 2$

	DCT-SPC	POC	Proposed Method
Shift Value	Average	Average	Average
$0.00 \leq |\delta| < 0.50$	0.086346±0.000562	0.024394±0.000342	**0.001409±0.00000037**
$0.50 \leq |\delta| < 1.00$	0.013402±0.000086	0.112840±0.005320	**0.001112±0.00000040**
$1.00 \leq |\delta| < 1.50$	0.006495±0.000340	0.052287±0.002078	**0.001103±0.00000050**
$1.50 \leq |\delta| < 2.00$	0.007844±0.000022	0.023787±0.000139	**0.001265±0.00000121**

TABLE III
COMPARISON OF ESTIMATION ERROR e OF SHIFT VALUE $2 < |\delta| < 10$

	DCT-SPC	POC	Proposed Method
Shift Value	Average	Average	Average
$2.0 \leq |\delta| < 3.9$	0.004466±0.000013	0.100079±0.004047	**0.004012±0.0000010**
$4.0 \leq |\delta| < 5.9$	0.004950±0.000015	0.071758±0.009882	**0.003835±0.0000009**
$6.0 \leq |\delta| < 7.9$	0.004966±0.000008	0.050912±0.006105	**0.002626±0.0000007**
$8.0 \leq |\delta| < 9.9$	**0.004498±0.000013**	0.025563±0.002302	0.006560±0.0000173

TABLE IV
COMPARISON OF ESTIMATION ERROR e WITH DIFFERENT IMAGES AND ADOPTION RANGE

Image	Least Squares Paramerter K_L									
	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
Airplane	0.007280	0.004356	0.002879	0.002136	0.001132	**0.000862**	0.001812	0.003191	0.003711	0.084319
Barbara	0.004771	0.001143	**0.000918**	0.001684	0.001337	0.001328	0.001431	0.005329	0.008243	0.072140
Mandrill	0.003408	0.001215	0.000717	**0.000660**	0.000871	0.001113	0.001303	0.001587	0.003010	0.012648
Milkdrop	0.000644	**0.000567**	0.000729	0.000861	0.002441	0.002590	0.007091	0.006194	0.000822	0.016538

[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 *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. 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 *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*. IEEE, 2009, pp. 449–452.