2023 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)

Maximization of 2D Cross-Correlation Based on Auxiliary Function Method for Image Alignment

Yuma Kinoshita*, Kouei Yamaoka†, and Hitoshi Kiya†
* Tokai University, Kanagawa, Japan
† Tokyo Metropolitan University, Tokyo, Japan

***Abstract*—In this paper, we propose an algorithm to maximize the generalized 2D cross-correlation function for image alignment, a cornerstone task in image signal processing and computer vision. As a fast image alignment method, 2D cross-correlation-based methods are often used in real-time applications. These approaches aim to optimize the 2D cross-correlation function between two images, inferring the displacement between them. Nevertheless, image misalignment commonly necessitates a maximization of the continuous cross-correlation function, and existing iterative and approximation-based approaches exhibit limited computational speed or accuracy. In response to this challenge, this paper proposes a novel algorithm to maximize the generalized 2D cross-correlation function. Inspired by Yamaoka et al.'s work for the 1D case, our approach maximizes a quadratic auxiliary function that bounds the 2D cross-correlation from below, instead of directly maximizing the cross-correlation. The accuracy and computational speed of our method, in comparison to three existing techniques, were assessed via a simulation experiment. Our findings confirm the superiority of the proposed method over a quadratic curve fitting approach in terms of image shift estimation accuracy. Remarkably, our proposed method proved approximately 6.8 and 4 times faster than gradient ascent and golden section search per iteration, respectively, competing with the speed of the non-iterative curve fitting.**

I. Introduction

Image alignment or matching is a fundamental problem in image signal processing and computer vision. Pre-alignment of images are required in many tasks: creating panoramas [1], medical imaging [2], image-based 3D modeling [3], high dynamic range imaging [4], [5], and more. Owing to the necessity, various image alignment methods have been proposed.

There are two main approaches to image alignment: homography-based methods and intensity-based methods. The former method endeavors to estimate a homography matrix between two images utilizing features such as SIFT [6], SURF [7], and A-KAZE [8]. RANSAC [9] is frequently employed to estimate the homography matrix. Contemporary research has explored the use of deep learning for estimating the homography matrix [10], [11]. The latter method typically focuses on estimating optical flow [12]–[15]. Nonetheless, these approaches require a significant amount of time for the estimation of either a homography matrix or optical flow.

Given these limitations, two-dimensional (2D) cross-correlation-based methods [16], [17] are often used for image alignment in real-time applications [5]. These strategies seek to maximize the 2D cross-correlation function between two images to deduce the displacement amongst them. The cross-correlation function for discrete pixel shifts can be computed expeditiously using the fast Fourier transform (FFT), and its maximum is easily found.

However, image misalignment usually manifests sub-pixel-wise necessitating maximization of the continuous cross-correlation function. Regrettably, a closed-form solution to this problem remains undiscovered. Although iterative algorithms and approximation-based solutions have been designed to tackle this challenge, these suffer from limited computational speed or accuracy, respectively. For example, the burst photography showcased in [5] adopts an alignment approach that maximizes the 2D cross-correlation function with subpixel accuracy by using a quadratic curve fitting method. It ensures rapid calculation on smartphones, but the limited accuracy of its estimations necessitates more resolution levels used for alignment in a progressive, coarse-to-fine strategy.

Because of such a situation, we aim to develop a fast and accurate maximization algorithm for the 2D cross-correlation function. For maximizing the generalized 1D cross-correlation function considering not only the standard correlation but also the phase-only correlation [18], Yamaoka et al. have proposed a fast and accurate algorithm based on an auxiliary function method which is also known as the majorization-minimization (MM) algorithm [19]. Following the approach, we derive an algorithm to maximize the generalized 2D cross-correlation function. The proposed method can find a (local) maximizer of the objective function by iteratively maximizing a quadratic auxiliary function that bounds the objective function from below. Since the maximizer of quadratic function can be obtained analytically, the proposed method can efficiently maximize the 2D cross-correlation.

We evaluate the accuracy and computational speed of the proposed method and other three methods, by a simulation experiment. Experimental results demonstrate the superior estimation accuracy of the gradient ascent method (GA) with backtracking, the golden section search (GSS), and the proposed method over a quadratic curve fitting approach for image shift estimation, corroborating the effectiveness of iterative methods including our one. Notably, the proposed method is approximately 6.8 and 4 times faster than GA and GSS per iteration, respectively, and could compete with the non-iterative curve fitting.

II. Problem Statement

In this paper, the focus is on the issue of aligning 2D discrete signals, essentially images, denoted as x and y. Their pixel

979-8-3503-0067-3/23/$31.00 ©2023 IEEE

values at a position $\boldsymbol{p} = (p_1, p_2)^\top$, are represented as $x[\boldsymbol{p}]$ and $y[\boldsymbol{p}]$, respectively, where the superscript $\top$ indicates the transpose of a vector or a matrix. We denote the 2D discrete Fourier transforms of x and y, which are $N \times M$ in size, as $\hat{x}$ and $\hat{y}$, respectively. We assume that $\hat{x}$ and $\hat{y}$ are strictly band-limited, i.e., $\hat{x}(\boldsymbol{\omega}) = \hat{y}(\boldsymbol{\omega}) = 0$ for angular frequency $\boldsymbol{\omega}$ such that at least one element is equal to π. By using the 2D cross-spectrum $\hat{\Phi}_2^{(xy)}(\boldsymbol{\omega}_{kl}) = \hat{x}^*(\boldsymbol{\omega}_{kl})\hat{y}(\boldsymbol{\omega}_{kl})$, the 2D cross-correlation function of x and y can be written as

$$\check{\Phi}_2^{(xy)}[\boldsymbol{p}] = \frac{1}{NM} \sum_{k \in K} \sum_{l \in L} \hat{\Phi}_2^{(xy)}(\boldsymbol{\omega}_{kl}) \exp\left(j \boldsymbol{\omega}_{kl}^\top \boldsymbol{p}\right), \quad (1)$$

where j represents the imaginary unit and $\boldsymbol{\omega}_{kl} = (\omega_k, \omega_l)^\top = (\frac{2\pi k}{N}, \frac{2\pi l}{M})^\top$. The summation range for K and L is given as $K = \{-N/2+1, -N/2+2, \cdots N/2\}$ and $L = \{-M/2+1, -M/2+2, \cdots M/2\}$, respectively.

In applications like image matching, the peak of the cross-correlation function may be sharpened by applying weights to each term in (1). For this reason, we consider the 2D generalized cross-correlation function given as

$$\check{\Phi}_2^{(xy)}[\boldsymbol{p}] = \frac{1}{NM} \sum_{k \in K} \sum_{l \in L} w_{kl} \hat{\Phi}_2^{(xy)}(\boldsymbol{\omega}_{kl}) \exp\left(j \boldsymbol{\omega}_{kl}^\top \boldsymbol{p}\right), \quad (2)$$

where $w_{kl} \in \mathbb{R}^+$ is an arbitrary weight. If the weight w_{kl} is set to 1, it is the same as the standard cross-correlation in (1). If $w_{kl} = |\hat{\Phi}_2^{(xy)}(\boldsymbol{\omega}_{kl})|^{-1}$, it coincides with the phase-only correlation [18], also known as GCC-PHAT [16] in acoustic signal processing.

Usually, (2) is a discrete function calculated with each element of $\boldsymbol{p}$ being an integer. On the other hand, if $\boldsymbol{p}$ is a real vector, $\check{\Phi}_2^{(xy)}$ can be considered as a continuous function. The purpose of this paper is to estimate a real vector $\boldsymbol{p} \in \mathbb{R}^2$ that maximizes the continuous 2D generalized cross-correlation function, that is, to solve the following optimization problem:

$$\tilde{\boldsymbol{p}} = \underset{\boldsymbol{p} \in \mathbb{R}^2}{\arg\max}\, \check{\Phi}_2^{(xy)}(\boldsymbol{p}). \quad (3)$$

III. Maximizing 2D Cross-Corrlation based on Auxiliary Function Method

Yamaoka et al. have already proposed an algorithm to maximize the 1D generalized cross-correlation using an auxiliary function method [19]. In this paper, following Yamaoka's work, we derive an algorithm that maximizes the 2D generalized cross-correlation.

A. Auxiliary Function Method

The auxiliary function method is an iterative algorithm designed for maximizing an objective function [20]–[22]. It maximizes an auxiliary function corresponding to a lower bound of the objective function, rather than directly maximizing the objective function itself.

To employ the auxiliary function method for solving the optimization problem shown in (3), it is necessary to design an auxiliary function $Q(\boldsymbol{p}, \boldsymbol{\theta})$ that satisfies the following conditions:

- For any given vectors $\boldsymbol{p}$ and $\boldsymbol{\theta}$, the inequality $\check{\Phi}_2^{(xy)}(\boldsymbol{p}) \geq Q(\boldsymbol{p}, \boldsymbol{\theta})$ holds true.
- There exist a $\boldsymbol{\theta} = f(\boldsymbol{p})$ satisfying $\check{\Phi}_2^{(xy)}(\boldsymbol{p}) = Q(\boldsymbol{p}, \boldsymbol{\theta})$ for any $\boldsymbol{p}$.

In this context, $\boldsymbol{\theta} \in \mathbb{R}^2$ is referred to as an auxiliary variable. Given the presence of such a $Q(\boldsymbol{p}, \boldsymbol{\theta})$ and an initial estimate $\tilde{\boldsymbol{p}}^{(0)}$, we can assure the subsequent sequence of updates to converge towards a local maximum:

$$\boldsymbol{\theta}^{(i)} = f(\boldsymbol{p}^{(i)}),\ \boldsymbol{p}^{(i+1)} = \underset{\boldsymbol{p} \in \mathbb{R}}{\arg\max}\, Q(\boldsymbol{p}, \boldsymbol{\theta}^{(i)}). \quad (4)$$

Here, i stands for the iteration index.

B. Deriving Auxiliary Function for 2D Correlation

From the conjugate symmetry of $\hat{\Phi}_2^{(xy)}$, we can rewrite $\check{\Phi}_2^{(xy)}$ as a sum of cosines as

$$\check{\Phi}_2^{(xy)}[\boldsymbol{p}] = \frac{1}{NM} \sum_{k \in K^+} \sum_{l \in L} \alpha_{kl} \cos\left(\boldsymbol{\omega}_{kl}^\top \boldsymbol{p} + \varphi_{kl}\right), \quad (5)$$

where $K^+ = \{0, 1, \cdots, N/2\}$,

$$\alpha_{kl} = w_{kl} \beta_{kl} \left|\hat{\Phi}_2^{(xy)}(\boldsymbol{\omega}_{kl})\right|, \quad (6)$$

$$\beta_{kl} = \begin{cases} 0 & \text{if } k = 0 \text{ and } l < 0 \\ 1 & \text{if } k \in \{0, N/2\} \text{ and } l \in \{0, M/2\} \\ 2 & \text{otherwise} \end{cases}, \quad (7)$$

$$\varphi_{kl} = \angle \hat{\Phi}_2^{(xy)}(\boldsymbol{\omega}_{kl}). \quad (8)$$

As shown in Proposition 1 in [19], for any real number θ, the following inequality holds

$$\cos\theta \geq -\frac{1}{2}\frac{\sin\theta_0}{\theta_0}\theta^2 + \left(\cos\theta_0 + \frac{1}{2}\theta_0 \sin\theta_0\right), \quad (9)$$

where $\|\theta_0\| \leq \pi$. Equality holds if and only if $|\theta| = |\theta_0|$ for $\theta_0 < \pi$ and $\theta = (2n+1)\pi, n \in \mathbb{Z}$ for $\theta_0 = \pi$.

Using (9) and the periodicity of $\cos\theta$, we obtain

$$\begin{aligned} &\alpha_{kl} \cos(\boldsymbol{\omega}_{kl}^\top \boldsymbol{p} + \varphi_{kl}) \\ &\geq -\frac{\alpha_{kl}}{2}\frac{\sin\theta_{kl}}{\theta_{kl}} \left(\boldsymbol{\omega}_{kl}^\top \boldsymbol{p} + \varphi_{kl} + 2n_{kl}\pi\right)^2 + C, \end{aligned} \quad (10)$$

where $n_{kl} \in \mathbb{Z}$ is such that $|\boldsymbol{\omega}_{kl}^\top \boldsymbol{p} + \varphi_{kl} + 2n_{kl}\pi| \leq \pi$, and C is a constant term that does not include $\boldsymbol{p}$. Equality holds when

$$\theta_{kl} = \boldsymbol{\omega}_{kl}^\top \boldsymbol{p} + \varphi_{kl} + 2n_{kl}\pi. \quad (11)$$

By applying (10) to each term of (5), an auxiliary function $Q(\boldsymbol{p}, \boldsymbol{\theta})$ of the objective function (5) is derived as follows:

$$\begin{aligned} &\check{\Phi}_2^{(xy)}[\boldsymbol{p}] \geq Q(\boldsymbol{p}, \boldsymbol{\theta}) \\ &= \sum_{k \in K^+} \sum_{l \in L} -\frac{\alpha_{kl}}{2NM}\frac{\sin\theta_{kl}}{\theta_{kl}} \left(\boldsymbol{\omega}_{kl}^\top \boldsymbol{p} + \varphi_{kl} + 2n_{kl}\pi\right)^2 \\ &+ C. \end{aligned} \quad (12)$$

C. Update Rules

Since $Q(\boldsymbol{p}, \boldsymbol{\theta})$ is a quadratic function in terms of $\boldsymbol{p}$, the maximizer of Q is easily obtained by solving an equation

$$\begin{aligned}&\nabla Q(\boldsymbol{p}, \boldsymbol{\theta})\\ &= \sum_{k\in K^+}\sum_{l\in L} -s_{kl}\left(\boldsymbol{\omega}_{kl}^\top \boldsymbol{p} + \varphi_{kl} + 2n_{kl}\pi\right)\boldsymbol{\omega}_{kl}\\ &= \mathbf{0}. \end{aligned} \tag{13}$$

In the above, ∇ indicates the gradient for $\boldsymbol{p}$ and $s_{kl} = \alpha_{kl}\frac{\sin\theta_{kl}}{\theta_{kl}}$. This equation yields, $\mathbf{A}\boldsymbol{p} = \boldsymbol{b}$, defined as:

$$\mathbf{A} = \sum_{k\in K^+}\sum_{l\in L} s_{kl}\boldsymbol{\omega}_{kl}\boldsymbol{\omega}_{kl}^\top, \tag{14}$$

$$\boldsymbol{b} = \sum_{k\in K^+}\sum_{l\in L} s_{kl}\left(-\varphi_{kl} - 2n_{kl}\pi\right)\boldsymbol{\omega}_{kl}. \tag{15}$$

Considering the condition for equality (11), we can substitute $-\varphi_{kl} - 2n_{kl}\pi = \boldsymbol{\omega}_{kl}^\top \boldsymbol{p} - \theta_{kl}$, and obtain the following update rules:

$$n_{kl}^{(i)} \leftarrow \arg\min_{n\in\mathbb{Z}} \left|\boldsymbol{\omega}_{kl}^\top \boldsymbol{p}^{(i)} + \varphi_{kl} + 2n\pi\right|, \tag{16}$$

$$\theta_{kl}^{(i)} \leftarrow \boldsymbol{\omega}_{kl}^\top \boldsymbol{p}^{(i)} + \varphi_{kl} + 2n_{kl}^{(i)}\pi, \tag{17}$$

$$\boldsymbol{p}^{(i+1)} \leftarrow \left(\mathbf{A}^{(i)}\right)^{-1}\boldsymbol{b}^{(i)} = \boldsymbol{p}^{(i)} - \left(\mathbf{A}^{(i)}\right)^{-1}\boldsymbol{v}^{(i)}, \tag{18}$$

where

$$s_{kl}^{(i)} = \alpha_{kl}\frac{\sin\theta_{kl}^{(i)}}{\theta_{kl}^{(i)}}, \tag{19}$$

$$\mathbf{A}^{(i)} = \sum_{k\in K^+}\sum_{l\in L} s_{kl}^{(i)}\boldsymbol{\omega}_{kl}\boldsymbol{\omega}_{kl}^\top, \tag{20}$$

$$\boldsymbol{v}^{(i)} = \sum_{k\in K^+}\sum_{l\in L} s_{kl}^{(i)}\theta_{kl}^{(i)}\boldsymbol{\omega}_{kl}. \tag{21}$$

IV. Simulation

To evaluate the performance of the proposed method in terms of the estimation accuracy and computational cost, we conducted a simulation.

A. Experimental Conditions

In this simulation, we numerically prepared misaligned image pairs from 20 well-known standard test images such as *airplane, barbara, cameraman, mandrill*, and so on[1]. An example of the misaligned image pair is shown in Fig. 1. The size of the original images was 256×256 pixels. To generate shifted images, the original images were shifted both vertically and horizontally, with the cubic interpolation. Each shift amount $\boldsymbol{p}^*$ was randomly selected from the uniform distribution with a range of $[-10, 10]$ pixels. Areas that the image shift causes to fall outside the original image boundaries were cropped. Finally, we added Gaussian noise with a standard deviation of 0.1 to both original and shifted images, independently.

[1]They can be downloaded from https://sipi.usc.edu/database/ and http://www.ess.ic.kanagawa-it.ac.jp/app_images_j.html

(a) Original (b) Shifted

Fig. 1. Example of shifted image (*Pepper*)

TABLE I
Machine spec used for evaluating executing time

Processor	AMD Ryzen 7 5800X (3.8 GHz, 8 cores)
Memory	64 GB
OS	Ubuntu 20.04 LTS
Language	Python 3.9.11
Library	PyTorch 2.0.1

The proposed auxiliary function-based method (Aux.) was compared with three methods: the gradient ascent (GA), the 2D golden section search (GSS) in [23], and a quadratic curve fitting in [5]. GA iteratively updates the maximizer $\tilde{\boldsymbol{p}}$ as $\tilde{\boldsymbol{p}}^{(i+1)} = \tilde{\boldsymbol{p}}^{(i)} + \eta\nabla\check{\Phi}_2^{(xy)}(\tilde{\boldsymbol{p}}^{(i)})$, where the step size η is determined by the backtracking line search following the Armijo rule [24]. GSS is also an iterative algorithm, but it does not require the gradient of the objective function. For GA, GSS, and the proposed method, we set the number of iterations to 50 and the initial estimation $\boldsymbol{p}^{(0)}$ was set to the maximizer of the discrete cross-correlation $\check{\Phi}_2^{(xy)}[\boldsymbol{p}]$ with $w_{kl} = 1$.

The curve fitting can be found an approximate solution by fitting a quadratic curve to the discrete function and finding the maximum of the fitted curve. Although a quadratic curve is fit to the $L2$ distance in [5], we used $\check{\Phi}_2^{(xy)}[\boldsymbol{p}]$ with $w_{kl} = 1$ for the curve fitting similarly to the other methods.

We implemented all methods in Python with PyTorch. The estimation error between $\tilde{\boldsymbol{p}}$ estimated by methods above mentioned and true shift $\boldsymbol{p}^*$ was measured by the root mean square error (RMSE).

The simulation was run on a PC, with a 3.8 GHz processor and a main memory of 64 Gbytes (see Table I). For measuring the executing time, *time.perf_counter()* function in the Python standard library was used, and GPUs were not used in the simulation

B. Results

Table II shows the RMSE between the estimated and true shifts for five images in the test set and the average and standard deviation of the RMSE for all 20 images. From the result, we can see that GA, GSS, and the proposed method achieved the higher estimation accuracy than the curve fitting. For this reason, it is confirmed that the iterative methods are more effective to obtain the accurate shift estimation than the curve fitting.

TABLE II
RMSE BETWEEN ESTIMATED AND TRUE SHIFTS

Image	GA	GSS [23]	Curve fitting [5]	Aux. (ours)
Airplane	0.0158	0.0158	0.0489	0.0158
Barbara	0.0453	0.0453	0.1900	0.0453
Cameraman	0.0139	0.0140	0.0157	0.0140
Pepper	0.1489	0.1489	0.2241	0.1489
Woman	0.0448	0.0454	0.0820	0.0454
Average (20 images)	0.0679	0.0679	0.1202	0.0680
Std. (20 images)	0.0391	0.0391	0.0612	0.0391

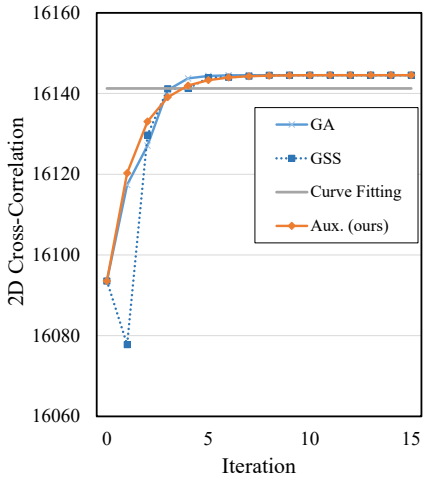

Fig. 2. Values of objective function during iterations for *Pepper*. Since curve fitting is not an iterative method, objective function value is drawn as line

TABLE III
ELAPSED TIME (MS) PER ONE ITERATION FOR ITERATIVE METHODS

Statistic	GA	GSS [23]	Curve fitting [5]	Aux. (ours)
Average	4.793	2.801	4.428	0.705
Std.	1.103	0.097	0.144	0.017

The values of the objective function during iteration are shown in Fig. 2. Here, the input image pair was *Pepper* shown in Fig. 1. From the figure, the objective values for GA and GSS converged mostly to the local maximum after six iterations and eight iterations for the proposed method. The trajectory of the estimated shift $\tilde{\boldsymbol{p}}^{(i)}$ for each method is shown in Figs. 3 to 5. Note that the maximizer of the 2D cross-correlation function $\check{\Phi}_2^{(xy)}(\boldsymbol{p})$ was slightly different from the true shift $\boldsymbol{p}^*$. A reason for this is that the input images were cropped after shifting.

Average elapsed times of the four methods for the 20 images are shown in Table III, where, for the iterative methods, the elapsed time for 50 iterations was measured and the elapsed time per iteration was calculated. As shown in the table, the proposed method is about 6.8 times faster than GA and about 4 times faster than GSS, and if the proposed method converges in about 8 iterations, as shown in Fig. 2, it is competitive with the non-iterative curve fitting.

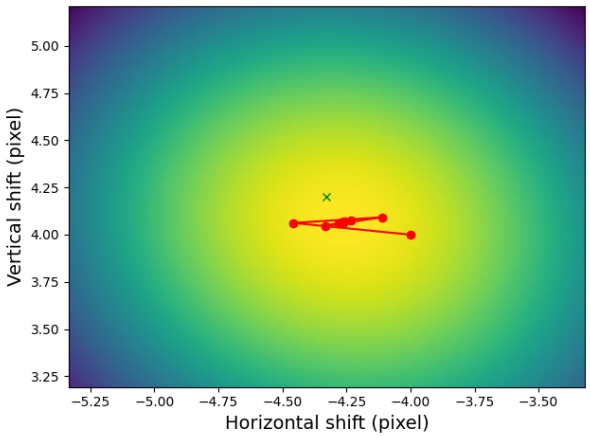

Fig. 3. Trajectories of estimated shifts by GA for image *Pepper* , where red points indicate estimations $\boldsymbol{p}^{(i)}$ and cross indicates true shift $\boldsymbol{p}^*$.

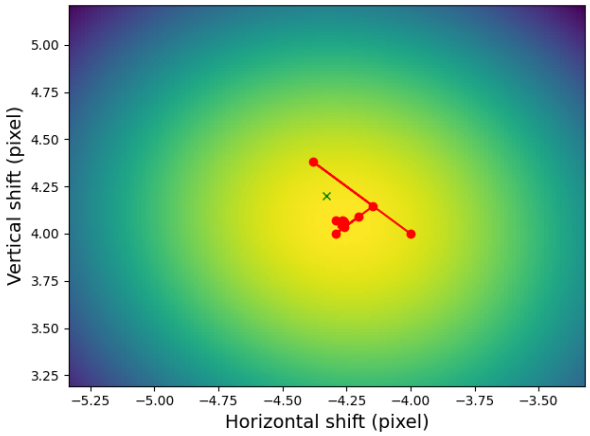

Fig. 4. Trajectories of estimated shifts by GSS [23] for image *Pepper* , where red points indicate estimations $\boldsymbol{p}^{(i)}$ and cross indicates true shift $\boldsymbol{p}^*$.

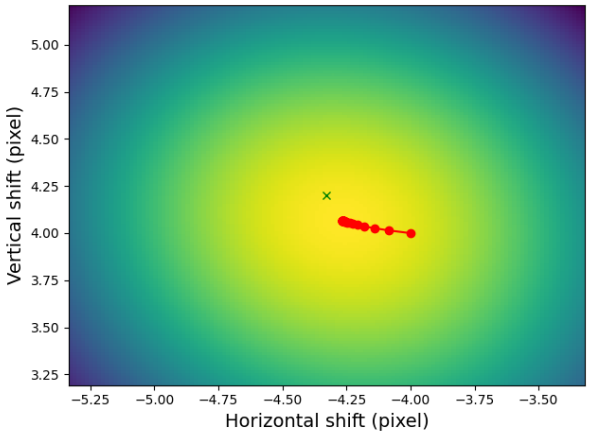

Fig. 5. Trajectories of estimated shifts by Aux. (ours) for image *Pepper* , where red points indicate estimations $\boldsymbol{p}^{(i)}$ and cross indicates true shift $\boldsymbol{p}^*$.

V. CONCLUSIONS

In this paper, we derived a maximization algorithm for the 2D generalized cross-correlation function, on the basis of the

auxiliary function based 1D cross-correlation maximization. The proposed method find the maximizer of the 2D cross-correlation function by iteratively maximizing a quadratic auxiliary function that bounds the 2D cross-correlation function from below. Experimental results showed that the proposed method can estimate the displacement between two images more accurately than a quadratic curve fitting method. In addition, the proposed method is faster than the gradient ascent method the golden section search, and the executing time is competitive with that of the curve fitting method.

REFERENCES

[1] M. Brown and D. G. Lowe, "Automatic Panoramic Image Stitching using Invariant Features," *Int. J. Comput. Vision.*, vol. 74, no. 1, pp. 59–73, Aug. 2007.

[2] G. Haskins, U. Kruger, and P. Yan, "Deep Learning in Medical Image Registration: A Survey," *Machine Vision and Applications*, vol. 31, no. 1-2, p. 8, Feb. 2020.

[3] N. Snavely, S. M. Seitz, and R. Szeliski, "Modeling the World from Internet Photo Collections," *Int J Comput Vis*, vol. 80, no. 2, pp. 189–210, Nov. 2008.

[4] P. Sen, N. K. Kalantari, M. Yaesoubi, S. Darabi, D. B. Goldman, and E. Shechtman, "Robust Patch-Based HDR Reconstruction of Dynamic Scenes," *ACM Trans. Graph.*, vol. 31, no. 6, 203:1–203:11, Nov. 2012.

[5] S. W. Hasinoff, D. Sharlet, R. Geiss, *et al.*, "Burst Photography for High Dynamic Range and Low-Light Imaging on Mobile Cameras," *ACM Trans. Graph.*, vol. 35, no. 6, pp. 1–12, Nov. 2016.

[6] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[7] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," in *Proc. ECCV*, vol. 3951, May 2006, pp. 404–417.

[8] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "KAZE Features," in *Proc. ECCV*, vol. 7577, Oct. 2012, pp. 214–227.

[9] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

[10] S.-Y. Cao, J. Hu, Z. Sheng, and H.-L. Shen, "Iterative Deep Homography Estimation," in *Proc. IEEE/CVF CVPR*, Jun. 2022, pp. 1869–1878.

[11] M. Hong, Y. Lu, N. Ye, C. Lin, Q. Zhao, and S. Liu, "Unsupervised Homography Estimation with Coplanarity-Aware GAN," in *Proc. IEEE/CVF CVPR*, Jun. 2022, pp. 17 642–17 651.

[12] B. D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," in *Proc. DARPA Image Underst. Workshop*, Apr. 1981, pp. 121–130.

[13] A. Dosovitskiy, P. Fischer, E. Ilg, *et al.*, "FlowNet: Learning Optical Flow with Convolutional Networks," in *Proc. IEEE ICCV*, Dec. 2015, pp. 2758–2766.

[14] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox, "FlowNet 2.0: Evolution of Optical Flow Estimation With Deep Networks," in *Proc. IEEE/CVF CVPR*, Jul. 2017, pp. 2462–2470.

[15] H. Jung, Z. Hui, L. Luo, *et al.* "Anyflow: Arbitrary Scale Optical Flow with Implicit Neural Representation." arXiv: 2303 . 16493. (Mar. 29, 2023), [Online]. Available: http://arxiv.org/abs/2303.16493.

[16] C. Knapp and G. Carter, "The Generalized Correlation Method for Estimation of Time Delay," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 24, no. 4, pp. 320–327, Aug. 1976.

[17] K. Takita, T. Higuchi, and K. Kobayashi, "High-Accuracy Subpixel Image Registration Based on Phase-Only Correlation," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. 86, no. 8, pp. 1925–1934, Aug. 1, 2003.

[18] C. D. Kuglin, "The Phase Correlation Image Alignment Method," in *Proc. Int. Conf. Cybern. Soc.*, Sep. 1975, pp. 163–165.

[19] K. Yamaoka, R. Scheibler, N. Ono, and Y. Wakabayashi, "Sub-Sample Time Delay Estimation Via Auxiliary-Function-Based Iterative Updates," in *Proc. IEEE WASPAA*, Oct. 2019, pp. 130–134.

[20] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-Minimization Algorithms in Signal Processing, Communications, and Machine Learning," *IEEE Trans. on Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.

[21] K. Lange, *MM Optimization Algorithms*. Philadelphia: SIAM, Jul. 2016.

[22] D. R. Hunter and K. Lange, "A Tutorial on MM Algorithms," *Am. Stat.*, vol. 58, no. 1, pp. 30–37, Feb. 2004.

[23] C. He, Y. Zheng, and S. Ahalt, "Object Tracking Using the Gabor Wavelet Transform and the Golden Section Algorithm," *IEEE Trans. Multimed.*, vol. 4, no. 4, pp. 528–538, Feb. 2002, ISSN: 1941-0077.

[24] L. Armijo, "Minimization of Functions Having Lipschitz Continuous First Partial Derivatives," *Pacific J. Math.*, vol. 16, no. 1, pp. 1–3, Jan. 1, 1966.