ISTANBUL UNIVERSITY –
JOURNAL OF ELECTRICAL & ELECTRONICS ENGINEERING

YEAR : 2007 (403 - 409)
VOLUME : 7
NUMBER : 2

INVARIANT BANGLA CHARACTER RECOGNITION USING A PROJECTION-SLICE SYNTHETIC-DISCRIMINANT-FUNCTION-BASED ALGORITHM

[1] Md. Waliullah Khan NOMANİ, Syed Mustafa Khelat BARİ, Tanwir Zubayer ISLAM, Mohammed Nazrul ISLAM

Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka 1000, Bangladesh

[1]E-mail : waliullahkhan@eee.buet.ac.bd

ABSTRACT

The projection-slice theorem has been used to create synthetic-discriminant-function based matched filters that are capable of discerning rotation and scale distortions. We propose to use the projection-slice algorithm for invariant Bangla character recognition that is capable of accommodating different handwriting and font-style distortions. The projection-slice algorithm is applied to training images in order to synthesize a composite image to represent each class. The opto-electronic fringe-adjusted joint transform correlator (JTC) technique is then used to correlate the test images with the composite images of each class. This technique is chosen for its superior performance to alternative JTCs and the feasibility of its implementation in the all-optical domain. Simulation results are furnished to prove the effectiveness of the proposed system.

Keywords: : *Joint transform correlator; fringe-adjusted filter; Bangla character recognition; projection-slice synthetic-discriminant-functions.*

1. INTRODUCTION

Real time character recognition by reducing the image processing speed is very much essential for real time applications such as military applications, high-speed character identification and decision-makings. Although there are many effective and efficient ways in offline pattern recognition, the only solution for real time pattern recognition [1] has come through opto-electronic process.

The challenge of Bangla character recognition through computer vision is associated with capturing a two-dimensional (2-D) representation of the input characters and then correlating it with the reference character. Factors such as rotation, variations in scale, font-style and handwriting, hinder successful recognition, and resolving them introduces complexities into the system. Lately, the processing speed and parallelism of optical correlators, combined with training techniques such as synthetic discriminant functions (SDFs) [2] and neural networks [3] have shown potential for providing real-time invariant face recognition.

Received Date:13.05.2007
Accepted Date: 05.12.2007

Recently, Riasati and Abushagur [4] proposed a modification to the SDF algorithm in which the projection slice theorem was applied to create a matched filter pattern recognition system. By creating a matched filter from 1-D Fourier slices, that system was capable of improved pattern recognition over the regular SDF-generated filter. In this paper, we present a system in which the projection-slice SDF technique is applied to the fringe-adjusted joint transform correlator (JTC) to achieve invariant recognition of Bangla characters while accommodating both handwriting and font-style distortions.

In Sec. 2, we describe the underlying fundamentals of the fringe-adjusted JTC, followed by a description of the proposed projection-slice SDF algorithm in Sec. 3. In Sec. 4, the opto-electronic realization of the fringe-adjusted JTC system is presented and the architecture is explained. In Sec. 5, the results from the computer simulation are presented and contrasted against the result obtained using the SDF algorithm. Concluding remarks are made in Sec. 6.

2. FRINGE-ADJUSTED JTC

In joint transform correlator (JTC), let the reference image and the input scene images are given by $r(x, y+y_0)$ and $t_1(x-x_1, y-y_1-y_0)$, $t_2(x-x_2, y-y_2-y_0)$, ……, $t_n(x-x_n, y-y_n-y_0)$ respectively, then the joint image can be expressed as :

$$f(x, y) = r(x, y + y_o) + \sum_{i=1}^{n} t_i (x - x_i, y - y_i - y_o) \tag{1}$$

The joint power spectrum (JPS) of Eqn. (1) is [1]

$$|F(u,v)|^2 = |R(u,v)|^2 + |T(u,v)|^2 + 2\sum_{i=1}^{n} |T_i(u,v)||R(u,v)| \cos[\Phi_{ti}(u,v) - \Phi_r(u,v) - ux_i - vy_i - 2vy_o] \tag{2}$$

where (u, v) is the frequency co-ordinate scaled by $2\pi/f\lambda$, λ is the wave length of collimating light, f is the focal length of the Fourier lens. $R(u, v)$ and $T(u, v)$ are the amplitude spectrums of the reference image and the input object image, respectively. $\Phi_r(u, v)$ and $\Phi_{ti}(u, v)$ are the phase spectrum distributions. Inverse Fourier transform of the JPS in Eqn. (2) produces the correlation output. In Eqn. (2) the first two terms correspond to zero order peaks and the last term corresponds to the correlation between the reference image and the input scene objects.

In classical JTC the JPS expressed by Eqn. (2) is directly fed into a SLM, which is illuminated by a laser. By taking the inverse Fourier transform of the JPS, correlation output is obtained. The correlation output contains broad zero order peak and false alarms between similar target and non-target objects. False alarms can be generated when multiple targets are present in the input scene. To eliminate the false alarms and the other extraneous signals, Fourier plane image subtraction method is employed and the modified JPS is expressed as [1]

$$P(u,v) = |F(u,v)|^2 - |R(u,v)|^2 - |T(u,v)|^2 = 2\sum_{i=1}^{n} |T_i(u,v)||R(u,v)| \times \cos[\Phi_{ti}(u,v) - \Phi_r(u,v) - ux_i - vy_i - 2vy_o] \tag{3}$$

where $|R(u, v)|^2$ and $|T(u, v)|^2$ are the reference only power spectrum and the input only power spectrum respectively. The modified JPS as found in Eqn. (3) is then multiplied by a real valued filter called fringe-adjusted filter (FAF). The FAF can be expressed as [1]

$$H_{FAF}(u,v) = \frac{B(u,v)}{A(u,v) + |R(u,v)|^2} \tag{4}$$

where $B(u, v)$ is used to control the gain of FAF and $A(u, v)$ to eliminate the problem related to poles and to suppress noise or to band limit the signal. After multiplying the JPS with the FAF as found in Eqn. (4), the fringe adjusted JPS is expressed as [1]

$$O(u,v) = H_{FAF}(u,v) \times P(u,v) \tag{5}$$

Inverse Fourier transform of Eqn. (5) produces a pair of sharp delta-like correlation peaks for each target object.

3. PROJECTION-SLICE SYNTHETIC DISCRIMINANT FUNCTION

When a composite image is created using the regular SDF algorithm from various training images, it is generated by a weighted average of the various frequency components of the training images. According to the projection-slice theorem, a 1-D Fourier transform along a projected line in an image corresponds to a slice in the 2-D Fourier transform of that image along the same line. By taking different slices from the training images, a composite image can be created such that the frequency components from different images do not interfere with each other. The final composite image is defined as

$$R_C(u,v) = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} a_n T_n(\omega\cos\Phi_{mn}, \omega\sin\Phi_{mn}) \quad (6)$$

where $R_C(u, v)$ is the composite image, T_n is the n'th training image, a_n is the SDF coefficient corresponding to the n'th training image, and u and v are the frequency variables for the rectangular coordinates, while ω and ϕ are the frequency variables in the polar coordinate system such that

$$u = \omega\cos\Phi$$
$$v = \omega\sin\Phi \quad (7)$$

In Eqn. (6), N is the number of training images and M is the number of slices taken from each image. The slices are taken at angles ϕ_{mn} determined by the expression

$$\Phi_{mn} = \frac{m\pi}{M} + \frac{n\pi}{MN}$$

for n = 0, 1, … …, N-1, m = 0,1,… …, M-1 (8)

By using the composite image $R_C(u, v)$ of Eqn. (6) as a reference image in the fringe-adjusted JTC, the JPS of Eqn. (3)becomes

$$P(u,v) = 2\sum_{i=1}^{n} |T_i(u,v)||R_C(u,v)| \times \cos[\Phi_{ti}(u,v) - \Phi_r(u,v) - ux_i - vy_i - 2vy_o] \quad (9)$$

$$= 2\sum_{i=1}^{n} |T_i(u,v)| \times \left| \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} a_n T_n(\omega\cos\Phi_{mn}, \omega\sin\Phi_{mn}) \right| \times \cos[\Phi_{ti}(u,v) - \Phi_r(u,v) - ux_i - vy_i - 2vy_o]$$

When the test input image $T_i(u, v)$ belongs to the training set, Eqn. (9) becomes [5]

$$P(u,v) = 2\sum_{m=0}^{M-1} a_i |T_i(\omega\cos\Phi_{im}, \omega\sin\Phi_{im})|^2 \times \cos[\Phi_{ti}(u,v) - \Phi_r(u,v) - ux_i - vy_i - 2vy_o] \quad (10)$$

When the JPS of Eqn. (10) is multiplied by the fringe-adjusted filter, the modified JPS becomes [5]

$$O(u,v) = H_{FAF}(u,v) \times P(u,v)$$
$$= \frac{2\sum_{m=0}^{M-1} a_i |T_i(\omega\cos\Phi_{im}, \omega\sin\Phi_{im})|^2 \times \cos[\Phi_{ti}(u,v) - \Phi_r(u,v) - ux_i - vy_i - 2vy_o]}{2\sum_{m=0}^{M-1} |a_i T_i(\omega\cos\Phi_{im}, \omega\sin\Phi_{im})|^2}$$
$$\approx \frac{2}{a_i} \times \cos[\Phi_{ti}(u,v) - \Phi_r(u,v) - ux_i - vy_i - 2vy_o] \quad (11)$$

By applying the inverse Fourier transformation to Eqn. (11), we obtain the correlation output, which contains a distinct narrow peak corresponding to the desired target.

4. OPTO-ELECTRONIC IMPLEMENTATION

The diagram in Fig. 1 depicts the components for opto-electronic implementation of the fringe-adjusted JTC. The fringe-adjusted JTC requires multiplying the JPS by the FAF and applying Fourier-plane image subtraction, and these operations can be implemented digitally on a computer. Following is a step-by-step description of the process.

Input scene images (i.e., test images) are captured through CCD3 and sent to the switching interface. The switching interface acts as a controller between the computer and the optical system. It directs the image from CCD3 to the upper half of the input-plane SLM (as input scene image). It also downloads a composite image from the computer and sends it to the lower half of the input-plane SLM (as reference image). At this point the phase-only SLM will be clear. The input joint image is then Fourier-transformed by lens FL1, and the corresponding JPS is captured by CCD1. The switching interface uploads the captured JPS to the computer.

The beam-splitter (BS) directs the input scene image through lens FL2 so that the input-scene power spectrum is captured by CCD2, which is then uploaded into the computer via the switching interface. To apply the Fourier-image subtraction process, the input-scene power spectrum and the pre-computed reference power spectrum are subtracted from the JPS to yield the modified JPS. The modified JPS is then multiplied by the FAF to yield the modified fringe-adjusted JPS, which constitutes the output of the entire digital processes and is redirected to the input-plane SLM via the switching interface. Since the SLM is incapable of displaying negative values, a phase-only SLM is used to apply a 180-deg phase shift to those pixels that correspond to negative values. Finally, lens FL1 performs an inverse Fourier transformation on to the modified fringe-adjusted JPS to produce the correlation output.

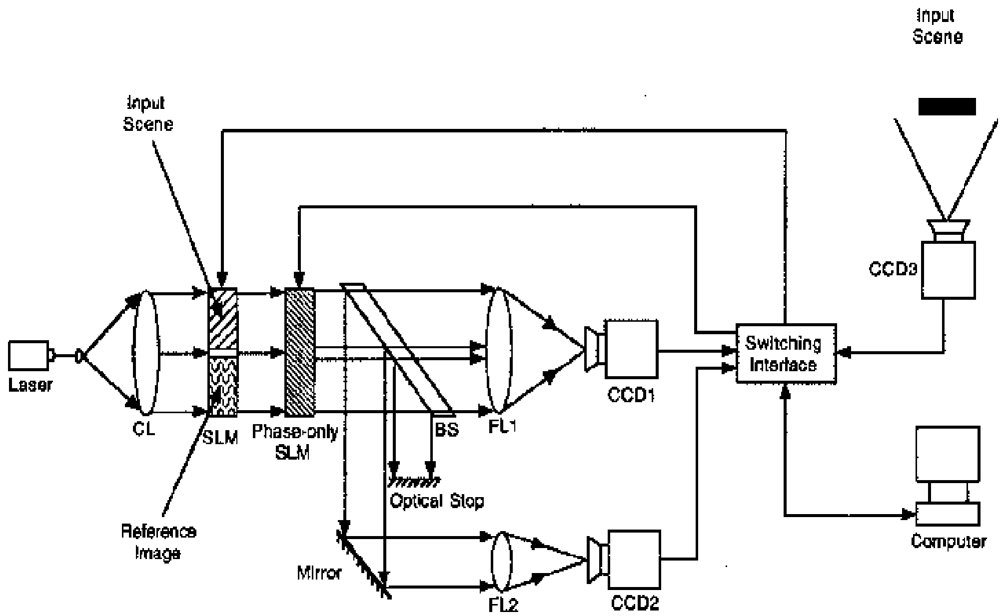

Figure 1: *Opto-electronic implementation setup of the system*

5. SIMULATION RESULTS

To investigate the performance of the proposed scheme, we have used a training set with different distortions of a Bangla character-image. The training set, shown in Fig. 2(a) consists of twenty-three subtraction, and these operations can be implemented digitally on a computer. Following is a step-by-step description of the process.

Input scene images (i.e., test images) are captured through CCD3 and sent to the switching interface. The switching interface acts as a controller between the computer and the optical system. It directs the image from CCD3 to the upper half of the input-plane different variations of handwriting distortions as well as five conventional fonts of different style. Each character has a size of 32×32 pixels. The size of the input and joint image plane is 256×256 pixels. Simulations are performed on a computer using the MATLAB software package. The final composite image is shown in Fig. 2(b). A plot of the normalized correlation peak intensities of different character- images of the training set with the composite image is shown in Fig. 3. The plot in Fig. 3 shows an excellent result in correlation outputs as all the correlation peak intensities converged to an acceptable value and there are very little fluctuations among the correlation outputs. The projection-slice SDF technique avoids Fourier-plane overlaps and hence it is very effective process for composite image formation.

First of all, the performance of a conventional JTC is investigated. A joint image with inputs at the upper half and one-character reference in the lower half is JTC technique alone is not sufficient for detecting invariant Bangla characters.

Finally, the performance of the projection-slice SDF technique-based JTC is investigated. The joint image is shown in Fig. 6. Now, the reference image is the composite image produced by projection-slice SDF technique, shown in the lower half of the joint image. The input scene, placed on the upper half of the joint image, contains two variations of the target character along with several non-targets. The correlation output without FAF is shown in Fig. 7. Fig. 7 depicts that the shown in Fig. 4. The input scene contains an exactly same character as the reference image as well as a character with a variation in font-style. The correlation output of the joint image is shown in Fig. 5. From Fig. 5 it is clear that one pair of delta-like peak have been produced for the input character that is exactly the same as the reference, the style variation has not been detected. Therefore, the conventional correlation output is not satisfactory and large side-lobs are present in the output. The FAF was calculated by Eqn. (4) with $B(u, v)$ set to 1 and $A(u, v)$ set to 0.00001. The correlation output after using FAF is shown in Fig. 8. The output plane contains two pairs of delta-like peaks of almost equal height for the two target inputs. Therefore, this proposed technique successful in detecting Bangla characters of different handwriting and font-style distortions.

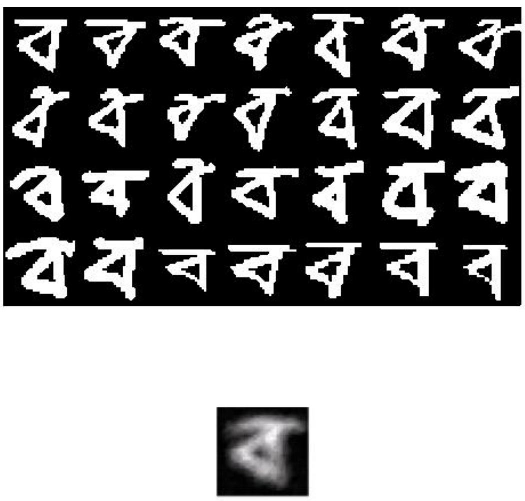

Figure 2.(a) *Training se,* **(b)** *PSDF composite image*

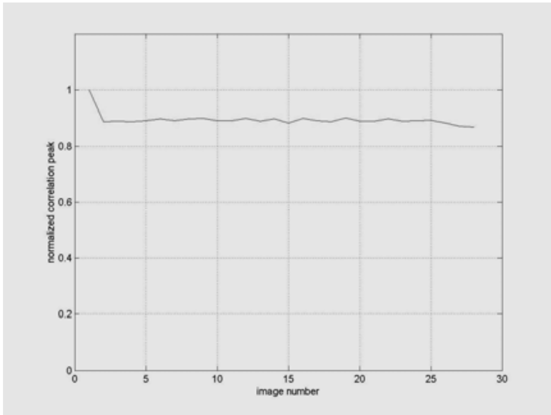

Figure 3. *The correlation peaks versus image number using PSDF composite.*

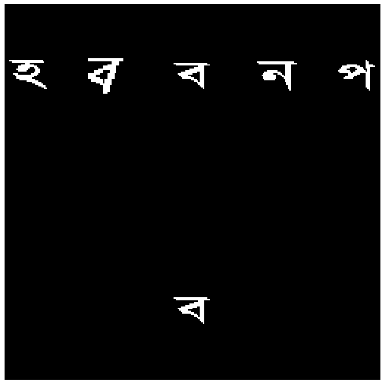

Figure 4. *Joint image.*

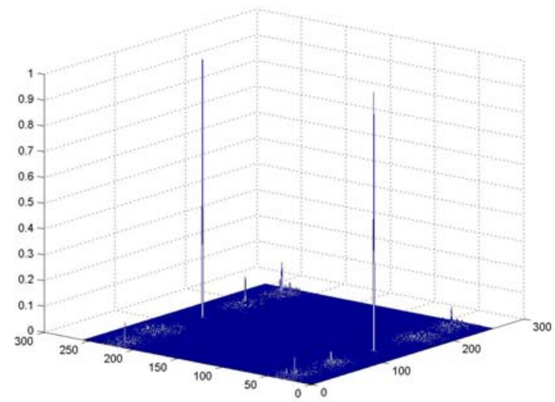

Figure 5. *Correlation output of joint image in fig. 4.*

Figure 6. *Joint image*.

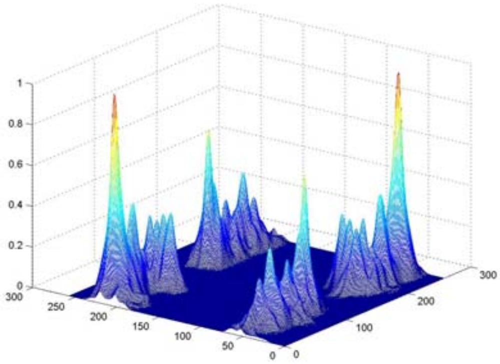

Figure 7. *Correlation output (without FAF) of joint image in fig. 6*

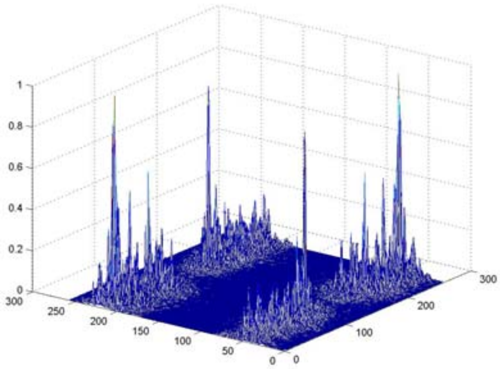

Figure 8. *Correlation output (using FAF) of joint image in fig. 6.*

6. CONCLUSION

In this paper, we have presented a projection-slice SDF-algorithm-based fringe-adjusted JTC technique to overcome various handwriting distortions and font-style variations for Bangla character recognition. We have also shown the use of fringe-adjusted JTC to aid in improving correlation accuracy and processing speed. Opto-electronic representation of the system is also discussed in this paper. Furthermore, the results from computer simulations show that the process works effectively in overcoming various distortions associated with Bnagla character recognition.

REFERENCES

[1] Md. Waliullah Khan Nomani, Syed Mustafa Khelat Bari, Tanwir Zubayer Islam, Md. Rafiqul Haider, and Mohammed Nazrul Islam**,** ''Gain-adjustable non-linear discrimination function for optical pattern recognition,'' in *Proc. ICEECE*, pp. 62-67, December 22-24, Dhaka, Bangladesh.

[2] V. R. Riasati, P. P. Banerjee, M. Abushagur, and K. B. Howell, ''Rotation-invariant synthetic discriminant function for optical pattern recognition,'' *Opt. Eng.* Vol: 39, pp. 1156-1161, 2000.

[3] B. Javidi, J. Li, and Q. Tang, ''Optical implementation of neural net-works for face recognition by use of nonlinear joint transform correlators,'' *Appl. Opt.* Vol: 34, pp. 3950–3962, 1995.

[4] V. R. Riasati and M. Abushagur, ''Projection-slice synthetic discriminant functions for optical pattern recognition,'' *Appl. Opt.* Vol: 36, pp. 3022–3034, 1997.

[5] A. R. Alsamman, ''Ultrafast invariant face recognition,'' PhD Disser-tation, University of Alabama (2001).

Md. Waliullah Khan Nomani received his BS in engineering in 2004 and his MS in engineering in 2006, both in electrical and electronic engineering, from Bangladesh University of Engineering and Technology. He joined the same university in 2004, where he is now a lecturer. He has published his research works in more than 10 refereed journals and conference proceedings. Now he focuses all his attention in optical communication, optical pattern recognition, image processing, wireless communication and solid-state devices & circuits.

Syed Mustafa Khelat Bari received his BS in electrical and engineering in 2004 from Bangladesh University of Engineering and Technology. He is currently pursuing his M.S. degrees in the same department in the same university. His research interests include optical communication, optoelectronics, optical pattern recognition, image processing, and VLSI design.

Tanwir Zubayer Islam received his BS in electrical and engineering in 2004 from Bangladesh University of Engineering and Technology. He is currently pursuing his M.S. degrees in the same department in the same university. He is the team leader of the power IC LTD. Company of Bangladesh. His research interests include optical communication, optoelectronics, optical pattern recognition, image processing, and VLSI design.

Mohammed Nazrul Islam received his BS in engineering in 1991 and his MS in engineering in 1994, both in electrical and electronic engineering, from Bangladesh University of Engineering and Technology. He joined the same university in 1992, where he is now an associate professor. He received his Dr Eng degree in optical communication and optoelectronics from Muroran Institute of Technology, Japan in 1999. Dr. Islam worked as a co-investigator in a research project on optical pattern recognition funded by the National Science Foundation of USA. He is currently working as a visiting researcher in the University of South Alabama in USA. He has published his research works in more than 80 refereed journals and conference proceedings. He was the Chairman of IEEE Bangladesh Section in 2003–2004. His research interests include optical communication, optoelectronics, optical pattern recognition, image processing, and wireless communication.
E-mail: nislam@usouthal.edu