# Positioning of noncooperative objects using joint transform correlation combined with fringe projection

Tobias Haist, Martin Schönleber, and Hans J. Tiziani
Institut für Technische Optik, Universität Stuttgart



***Abstract*— Automated assembly and quality control require reliable systems for the detection of the position and orientation of complicated objects. Correlation methods are well suited, but they are affected by structured backgrounds, varying illumination conditions, and textured or dirty object surfaces. Using fringe projection to exploit the three–dimensional topography of objects, we improve the performance of a nonlinear joint transform correlator. Positioning of noncooperative objects with subpixel accuracy is demonstrated. Additionally, the tilt angle of an arbitrarily shaped object is measured by projecting object–adapted fringes that produce a homogeneous fringe pattern in the image plane. An accuracy of better than one degree is achieved.**

***Index Terms*— image processing, object registration, object detection, correlation, joint transform correlator, fringe projection, triangulation, pattern recognition**

## I. Introduction

MOST automatic methods for quality control and industrial fabrication require exact object positioning. Unfortunately, many “noncooperative” objects, such as free–form objects, are difficult to handle by ordinary digital image processing or by simple distance sensors. Problems occur due to dirt, bright spots, stray light, different labels (like serial numbers), backgrounds, or strong textures. These problems are mainly caused by the two–dimensional object surface. Therefore, it seems to be useful to exploit three–dimensional information about the object (its surface profile) to improve positioning.

Many techniques for obtaining three–dimensional information are available. Ordinary setups for fringe projection (see Fig. 1) with different masks are often used: fixed projected gratings are cheap but have a restricted depth of field. Interferometrically generated fringes extend the field of structured illumination to a volume. Problems due to speckles may occur. Both methods suffer from the restricted adaptability of the fringe pattern. Electrically addressed spatial light modulators (SLMs), used as adaptive fringe masks[1], [2], [3] in combination with multiple images, offer additional merits. The restriction on simple fringes can be circumvented without sacrificing depth of field by using the reconstruction of computergenerated holograms[4]. In some applications low SLM resolution or higher costs might be a problem with this approach. Today’s liquid crystal displays (LCDs) and digital micromirror devices (DMDs) are readily available and the costs of these elements are constantly decreasing while high quality projection of arbitrary patterns is possible. We used twisted nematic liquid crystal devices (Epson ELP–3300 (VGA LCD) and SONY LCX016AL (SVGA LCD)) and a Texas Instruments digital micromirror based VGA–projector for our experiments. Depth of field for fringe projection in these arrangements may be increased by using non–standard apertures[5] that enhance the spatial frequencies of the fringe pattern. The images were detected by a standard CCIR CCD camera.

In this paper, we summarize results obtained by the combination of fringe projection with optical and digital nonlinear joint transform correlators and direct evaluations in the spatial frequency domain. The methods are suitable for the determination of the position, the orientation and the tilt of complicated objects.

In section 2 we discuss the Fourier spectrum of an object illuminated with projected fringes. Section 3 deals with direct evaluations in the Fourier domain and in section 4 we discuss the combination of fringe projection with a nonlinear joint transform correlator.

## II. Fourier spectrum of objects illuminated with projected fringes

The three–dimensional information of the object is obtained by fringe projection. We project sinusoidal fringes with period $p$ along the x–axis onto the object under test (see Fig. 1). Assuming perfect imaging for the projection, the light distribution near the object surface is

$$I(x,y,z) = I_0(x,y,z)\ \left[1+\cos\left\{\frac{2\pi\cos\alpha}{p}(x+z\tan\alpha)\right\}\right], \tag{1}$$

where $\alpha$ denotes the angle between projection and detection and $I_0$ is the mean intensity.

The two–dimensional stray light distribution in the (x,y)–plane at the object surface $z(x,y)$ has the form

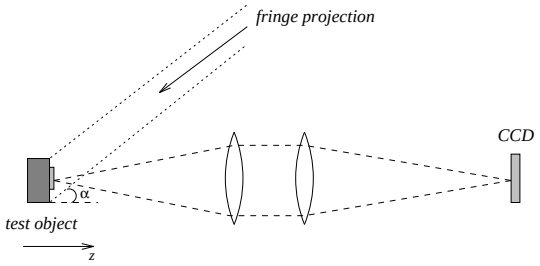



Fig. 1. Setup for fringe projection methods

$$I(x,y) = I_0(x,y)\left[1 + V(x,y)\cos\left\{\frac{2\pi x\cos\alpha}{p} + \Phi(x,y)\right\}\right], \quad (2)$$

where the three–dimensional topography is encoded in the phase

$$\Phi(x,y) = \frac{2\pi z(x,y)\sin\alpha}{p}. \quad (3)$$

In the following, we assume perfect imaging between the object plane and the CCD. Therefore, it is valid to use the object coordinates and Eq. (2) for our further calculations.

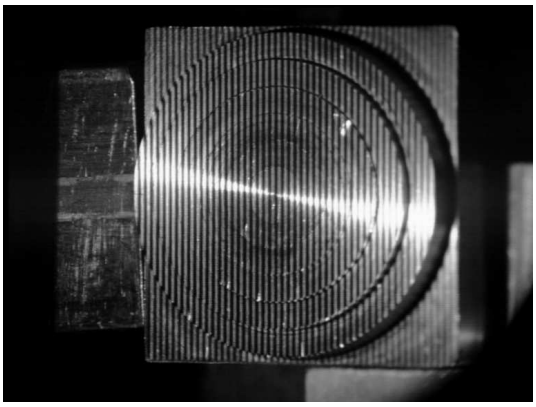


Fig. 2. Object–adapted fringe projection on a circular step object (with a tilt angle of 9 degrees with respect to the reference orientation)

Normally, the phase $\Phi(x,y)$ is calculated from the detected intensities of one or several images in order to obtain the object topography. Here we do not explicitly need this phase information.

We can rewrite Eq. (2) as

$$\begin{aligned} I(x,y) &= I_0 + I_0\, V\, e^{i\Phi} e^{i2\pi\, x\cos(\alpha)/p} && (4)\\ &+ I_0\, V\, e^{-i\Phi} e^{-i2\pi\, x\cos(\alpha)/p} && (5)\\ &:= S_1 + S_2 + S_3, && (6) \end{aligned}$$

denoting definition by ":=".

The Fourier transform $\mathcal{F}$ yields

$$\begin{aligned} s(f_x,f_y) &:= \mathcal{F}[S(x,y)] = \mathcal{F}[S_1] + \mathcal{F}[S_2] + \mathcal{F}[S_3] && (7)\\ &:= s_1 + s_2 + s_3. && (8) \end{aligned}$$

The spectrum of the image is therefore the sum of the pure intensity spectrum $s_1$ and the two spectra $s_2$ and $s_3$ containing the shape information. $s_2$ and $s_3$ are grouped around the positive and negative carrier frequencies of the fringe pattern.

The absolute value of $s$ is not affected by lateral movement of the object but the phases of the three terms $s_i$ change in different ways. Rotation of the object results in a rotation of the three terms $s_i$ around their centers.

## III. Direct evaluation in the Fourier domain

Projecting sinusoidal fringes onto a plane surface and recording of an image under a triangulation angle results in sinusoidal fringes (see Fig. 2). The spatial frequency of the detected fringes of course depends on the tilt of the plane. Tilt evaluation in the Fourier domain is possible by simple means since the fringe pattern results in two peaks (as described in the last section) whose positions correspond to the fringe period and therefore also directly correspond to the tilt of the object. A complicated object with several plane surfaces results in several localized peaks in the Fourier domain.

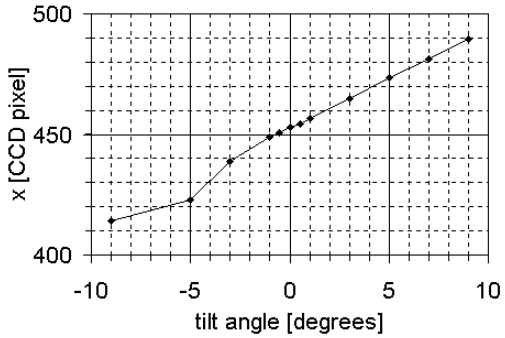



Fig. 3. Fourier peak position vs. object tilt for object–adapted fringe projection

For the determination of the peak positions we use a simple threshold (all values lower than the threshold $T$ are set to zero) followed by normalization and the computation of the center of gravity. The detection of such Fourier peaks has also been used by Will and Pennington for a simple and robust object recognition/detection method[6].

Curved surfaces correspond to broad peaks, which are difficult to evaluate. Sharp peaks are obtained even for complicated objects if object–adapted fringes[1], [2], [3] are used.

The object–adapted fringe mask is computed by ray–tracing by using the computer aided design (CAD) data of the object under test, the position of the object, and the coordinate transforms of the calibrated setup. Projection of this object–adapted mask results in straight fringes on the detecting camera or the Fourier processor (see Fig. 2).

The Fourier transform has been implemented by digital electronics and by optical means. We use a video–addressed twisted–nematic LCD in combination with a HeNe laser (see Fig. III). A simple lens is employed in a converging beam setup which has several advantages compared to a parallel beam arrangement[7]. We reconstruct a 16 times tiled computergenerated Fourier hologram optimized with the intensity based direct binary search algorithm[8] for the adjustment. The tiling results in discrete pixels at the reconstruction that

enable comfortable adjustment. The Fourier transform setup is also used with the correlation method (section 4).

For many practical applications, modern digital signal processors (DSPs) or specialized FFT boards are now often fast enough (video frequency with $512 \times 512$ pixels is possible[9]) to replace coherent optics. Therefore, the Fourier transform is also performed by electronic means via the fast Fourier transform algorithm. We apply additional zero padding for the correlation (see section IV).

In Fig. 3 the Fourier peak position for a circular step object (see Fig. 2), illuminated with object–adapted fringes, is shown in dependence of the object tilt. At $-5°$, an error of unknown origin of 8 CCD pixel is seen. A repeatability of 0.12 degrees has been observed. Apart from the influence of the projection and detection system, the accuracy of the method mainly depends on the object and its correct lateral position. Typically, the accuracy is better than 1 degree.

Instead of directly determining the orientation of an object, we can also use the spectrum of the rotated object for this task. This is possible because the spectrum of a rotated object rotates with the object [10].

Furthermore, all ordinary image processing methods that are used with conventional images (e.g. similarity detection algorithms[11]) can be applied to the spectral representations of the images. This is especially interesting because the power spectrum is shift invariant and therefore useful for object detection and for the determination of object orientations. Unfortunately, the spectra are often less structured than the original images and therefore difficult to evaluate. In addition, the phase information, which is responsible for the position of different frequencies within the image, is lost.

It should be noted that, if — as described above — projected fringes are used to incorporate three–dimensional information, only one part of the spectrum ($s_2$ or $s_3$) is evaluated (see section 2).

## IV. Correlation and fringe projection

Correlation methods are often used to recognize and to obtain the two–dimensional position of objects. Sub–pixel accuracy and immunity against localized errors and noise are achieved. Efficient implementation by optical means or by modern digital hardware is possible. Nonlinear joint Fourier transform correlators[12], [13], [14] seem to be best suited for applications in an industrial environment. Therefore, we concentrated our work on such correlators and used simple local nonlinearities which are easy to implement. Especially kth law modification[13], [14] ($I \rightarrow I^k$ with $k = 0..1$) is interesting because such nonlinearities can be achieved by electronics with lookup tables or optically by optically addressed liquid crystal displays[15]. For our optical correlations it has been sufficient to use appropriate camera shutter speeds in combination with the nonlinear responses of the LCDs and the CCDs.

The two images to be correlated, $G$ and $F$, are written side by side into one spatial light modulator. The joint power spectrum is given by

$$\begin{aligned} i(f_x, f_y) &:= \mathcal{F}[F(x,y) + G(x,y)]^2 \\ &= f(f_x, f_y) f^*(f_x, f_y) + g(f_x, f_y) g^*(f_x, f_y) \\ &\quad + f(f_x, f_y) g^*(f_x, f_y) + g(f_x, f_y) f^*(f_x, f_y), \end{aligned} \tag{9}$$

where we used lower case letters for the transformed images. By Fourier transforming Eq. (10) and taking into consideration that the images $F$ and $G$ are shifted along the x–axis by $+x_0$ and $-x_0$ respectively, we obtain the correlator output

$$\begin{aligned} C(x,y) &= F(x,y) \otimes F(x,y) + G(x,y) \otimes G(x,y) \\ &+ F(x + 2x_0, y) \otimes G(x,y) \\ &+ F(x - 2x_0, y) \otimes G(x,y), \end{aligned} \tag{10}$$

where the operator $\otimes$ denotes correlation.

The first two terms are the autocorrelations of the two images. The third and the fourth term are the crosscorrelations.

If an additional nonlinearity is applied to Eq. (10), the peak sharpness of the correlator output can be improved. However, a simple analytic description is not possible anymore.

For homogeneously illuminated objects in front of uniform backgrounds, we achieved with the optical JTC a mean error of the position of 1/50 of a pixel. The maximum error was 1/5 of a pixel.

Although correlation methods are indeed quite immune against localized errors and noise, problems occur if the errors become too large. Different methods, especially filter optimizations (see below) and more complicated modifications in the Fourier plane, have been used in the past to improve correlator performance. Here we propose another approach. We project fringes onto the object and therefore exploit three–dimensional object information to improve correlator performance.

Instead of using a single image, we record two images with rectangular fringes shifted by half of a period. The resulting images are subtracted from each other and histogram equalized (simple normalization nearly works equally well). The resulting image mainly consists of deformed fringes due to the object topography. Two–dimensional object information, especially object textures, dirt, or illumination effects, is strongly reduced. Very difficult scenes can be handled with this approach. In Fig. 4, the object under test, a power plug, is placed in front of two different and very dominating backgrounds. Obtaining the position of the plug by ordinary image processing methods is very difficult if not impossible (at least if the methods are not trained for this special application). Direct correlation of the images with the nonlinear joint transform correlator (NLJTC) is presented in Fig. 6 (a). Determination of the position of the object is not possible. If we use the described fringe projection difference images (in Fig. 5 a magnified part of one difference image is depicted), we obtain the simply detectable peak shown in Fig. 6 (b).

This example shows how the correlator output can be dramatically improved by using the fringe projection method.

At first, it seems to be clear how this improvement is achieved. However a problem occurs because the projected fringes will remain at fixed positions no matter where the

(a)

(b)

Fig. 4. Power plug in front of different backgrounds

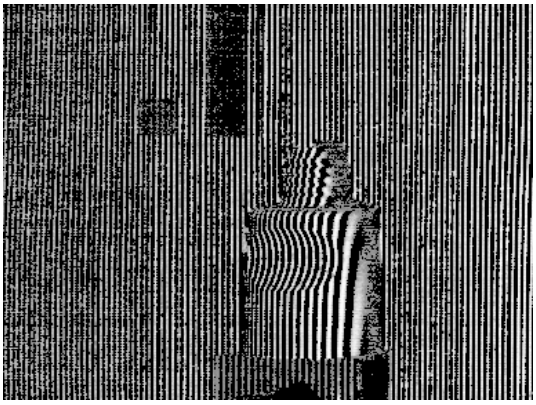
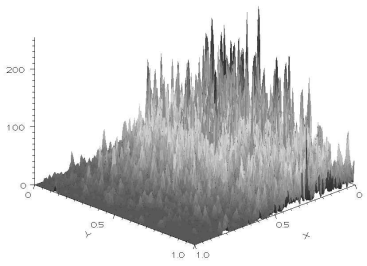


Fig. 5. Magnified part of a difference image

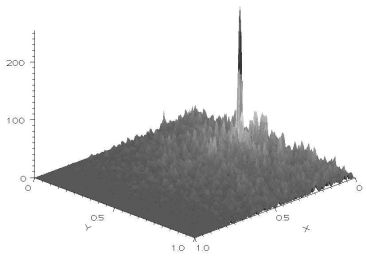
object is situated. Only the fringe deformation changes its position.

Even worse, this deformation sometimes does not move, as one would expect, but rather changes its shape. Consider, for example, a point defect. This defect is not visible in both images because it will be illuminated only in one of them. Therefore, the difference image will be inverted if we shift the defect by half of the fringe projection wavelength. Fortunately, the typical topography of real objects is not dominated by such small features, and there has been enough information available to successfully correlate the images with the objects we have tested.

Another problem remains: correlation might result in multiple peaks if fringe images are used. In order to understand and avoid this, we visualize (one–dimensional) correlation by the process depicted in Fig. 7: while the reference image $R(x)$ is shifted across the input image $F(x)$, each pixel $R(x)$ is multiplied with the corresponding pixel $F(x)$. The sum of all these products is calculated to obtain the correlation. If the shifted reference and the input correctly lie upon each other, maximum correlation is obtained.




(a) correlation of images




(b) correlation with fringe enhancement

Fig. 6. Numerical JTC correlation (a) of the images shown in Fig. 4 and (b) modified correlation using fringe enhancement (see Fig. 5) and additional filtering in the Fourier plane.

If additional high contrast fringes are superposed on the images, we run into problems since these fringes dominate the images. We get strong correlation signals for positions that correspond to fringes lying on each other. Therefore, we obtain two or more correlation peaks instead of one single peak. The positions of the peaks are slightly wrong and the distances between them correspond to the fringe carrier frequency. If and how many additional peaks are visible mainly depends on the object to be positioned, on the projected fringes, and on the position of the object. In most cases, only one additional peak at distinct positions is visible, but simulation of the whole process with simple, circular symmetric step objects resulted in up to five easily detectable additional peaks.

In order to avoid this problem, we have to eliminate the fringes without relinquishing the three–dimensional information of the object. We eliminate the fringes by a simple stop filter for the carrier frequency of the fringes in the Fourier domain since only the fringe deformation is important. There is no additional computational cost since one already needs the transformation into the Fourier domain for performing the correlation.

A certain width of the filter is necessary because the fringe period in the recorded images depends on the tilt of different object surfaces. The width of the filter again depends

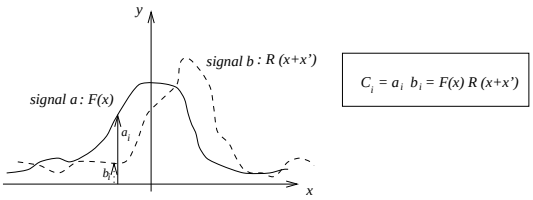
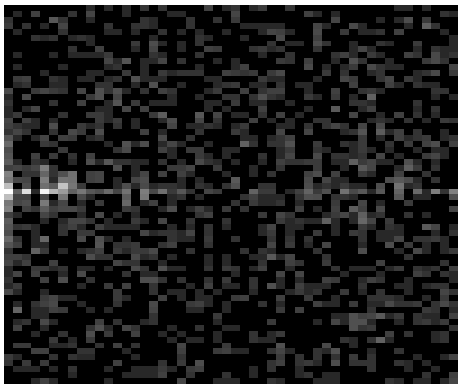


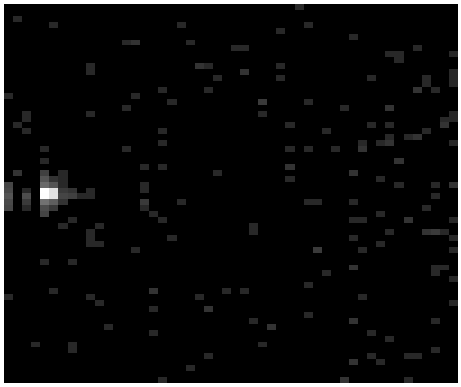

Fig. 7. One–dimensional visualisation of correlation between two signals

on the fringes and the object to be positioned. Automatic optimization should be possible because the whole process (projection, recording, and processing) can be simulated within the computer. An intelligent, nonlinear filter detecting the local fringe structure also might be an interesting alternative for future research.

Another approach for avoiding multiple peaks is to use an optimized reference image that introduces additional nonlinearities that improve correlation results. Various optimization methods have been used in the past to improve nonlinear JTCs and linear correlators[16], [17], [18], [19], [20]. Javidi and Painchaud extended the formalism of synthetic discriminant functions (SDF) to nonlinear JTCs in order to improve the immunity of joint transform correlation against scaling and rotation[14]. Different performance criteria were investigated and an overall improvement of the correlator performance was achieved with nonquantized references.



(a) without optimization



(b) with DBS optimization

Fig. 8. Numerical correlation obtained (a) without optimized reference and (b) with DBS optimized reference ($\gamma = 3$ applied to both images to enhance noise for visualisation)

Since we are interested here in cheap binary SLMs, we decided to use a quite simple but computationally intensive binary optimization method, the direct binary search (DBS) algorithm[8], [21], [22], which is a well known method for the optimization of computergenerated holograms.

Multiple peaks were strongly reduced because the optimization automatically attenuated the periodicity of the reference image. The correlator output performance criterion used was the peak–to–average correlation energy[23] (PACE).

In Fig. 8 (a) and (b) examples of correlator outputs without and with binary optimization of the reference image are shown. Apart from a better overall peak quality and reduced noise, the optimized reference led to a single peak instead of multiple peaks.

We can further improve the correlation results by applying ordinary image processing algorithms for pre– and post–processing. Edge enhancement of the input picture can have some positive (but mostly small) effects[24] . Contrast enhancement by histogram equalization (or even simpler: normalization) always improved correlator performance in our experiments. Simple smoothing of the output was used to reduce binarization noise.

Since we were only interested in obtaining the position of single objects we did not investigate chirp–encoded joint transform correlation[25] as suggested by a reviewer for multi–object detection.

In analogy to the positioning of objects, the object orientation can be detected by different techniques. If the position has already been determined, ordinary image processing algorithms are a good choice. Interesting alternatives are — again — correlation methods. The traditional approach using coordinate transforms[26] has been replaced by a more robust but computationally intensive method[27]. The peak–to–average correlation energy[23] (PACE) has a maximum if the input object has the correct orientation.

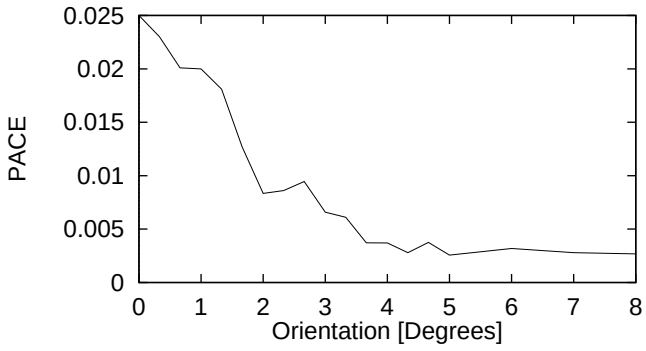



Fig. 9. Peak–to–correlation energy for rotated object (central part of a spoon (high rotational symmetry))

If the object under test is rotated by a few degrees, the PACE dramatically drops. This is not wanted for many applications. Therefore, special filter designs are often used to achieve high PACEs even if the orientation deviates from the reference orientation[28], [29], [14]. Instead of applying such designs, we use the PACE for detecting the object orientation.

A typical experimental result for different orientations is

shown in Fig. 9. The rotational symmetry of the object controls the steepness of the curve.

In order to search for the maximum PACE, many correlations are needed (simple method: correlation with $N$ different reference objects, rotated by $\Delta\phi = 360°/N$). Very fast digital electronics or optical implementation (e.g. rotating prisms or switchable ferroelectric LCDs in combination with random addressable detectors (e.g. CMOS cameras)) are therefore mandatory. An optical and electronical joint transform correlator was demonstrated but not yet optimized with respect to speed.

## V. Conclusions

We have investiged the combination of fringe projection with direct evaluation in the Fourier domain and correlation methods in order to recognize and detect objects that are difficult to handle by ordinary image processing methods. Detection of the position, the orientation, and the tilt of such objects was achieved with subpixel accuracy.

We are grateful for the financial support of the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie.

## References

[1] T. Yatagai and M. Idesawa, “Use of synthethic deformed gratings in moire topography”, *Optics Communications* **20**, pp. 243–245, 1977.

[2] J. Harthong and H. Sahli, “Theory of moiré sensing by means of contour functions”, *Applied Optics* **31**, pp. 1436–1443, 1992.

[3] M. Schönleber and H. Tiziani, “Fast and flexible shape control with adaptive lcd fringe masks” in *Conference on Optical Inspection and Micromeasurement 2*, , G. Gorecki, ed., *Proceedings of the SPIE* **3098**, pp. 35–42, 1997.

[4] T. Haist and H. Tiziani, “Computergenerated holograms from 3d–objects written on twisted–nematic liquid crystal displays”, *Optics Communications* **140**, pp. 299–308, 1997.

[5] D. Williams, ed., *Optical Methods in Engineering Metrology*, ch. 9.2.1, pp. 349–352. Chapman and Hall, 1993.

[6] P. Will and K. Pennington, “Grid coding: a novel technique for image processing”, *Proceedings of the IEEE* **60**, pp. 669–680, 1972.

[7] D. Joyeux and S. Lowenthal, “Optical fourier transform: what is the optimal setup ?”, *Applied Optics* **21**, pp. 4368–4372, 1982.

[8] M. Seldowitz, J. Allebach and D. Sweeney, “Synthesis of digital holograms by direct binary search”, *Applied Optics* **26**, pp. 2788–2798, 1987.

[9] M. Cavadini, M. Wosnitza, M. Thaler and G. Tröster, “A vlsi architecture for real time object detection on high resolution images” in *Signal Processing VIII, Theory and Applications, EUSIPCO-96*, , G. Ramponi, G. Sicuranza, . Carrato and S. Marsi, eds., vol. 3, pp. 1475–1478, 1996.

[10] M. Schönleber and H.-J.Tiziani, “Spatial light modulator used for the fourier transformation of video images”, *Journal of Modern Optics* **40**, pp. 2039–2052, 1993.

[11] D. Barnea and H. Silverman, “A class of algorithms for fast digital image registration”, *IEEE Transactions on Computers* **21**, pp. 179–186, 1972.

[12] D. Jared, K. Johnson and G. Moddel, “Joint transform correlation using an amorphous silicon ferroelectric liquid crystal spatial light modulator” in *Optical Information Processing Systems and Architectures*, , B. Javidi, ed., *Proceedings of the SPIE Vol. 1151* **1151**, pp. 148–153, 1989.

[13] B. Javidi, “Nonlinear joint power spectrum based optical correlation”, *Applied Optics* **28**, pp. 2358–2367, 1989.

[14] B. Javidi and D. Painchaud, “Distortion–invariant pattern recognition with fourier–plane nonlinear filters”, *Applied Optics* **35**, pp. 318–331, 1996.

[15] L. Guibert, G. Keryer, A. Servel, M. Attia, H. MacKenzie, P. Pellat-Finet and J. de Bourgrenet de la Tocnaye, “On–board optical joint transform correlator for real–time road sign recognition”, *Optical Engineering* **34**, pp. 135–143, 1995.

[16] B. Kumar, “Tutorial survey of composite filter design for optical correlators”, *Applied Optics* **31**, pp. 4773–4801, 1992.

[17] M. Taniguchi, K. Matsuoka and Y. Ichioka, “Computer–generated multiple–object discriminant correlation filters: design by simulated annealing”, *Applied Optics* **34**, pp. 1379–1385, 1995.

[18] F. Wyrowski and M. Bernhardt, “Marriage between digital holography and optical pattern recognition” in *Computer and Optically Generated Holographic Optics 4*, , H. Caulfield, R. Johnson and Q. Huang, eds., vol. 1555 of *Proceedings of the SPIE*, pp. 146–153, 1991.

[19] U. Mahlab and J. Shamir, “Genetic algorithm for optical pattern recognition”, *Optics Letters* **16**, pp. 648–650, 1991.

[20] M. Fleischer, U. Mahlab and J. Shamir, “Entropy optimized filter for pattern recognition”, *Applied Optics* **29**, pp. 2091–2098, 1990.

[21] M. I. J. Ding and T. Yatagai, “Design of optimal phase–only filters by direct iterative search”, *Optics Communications* **118**, pp. 90–101, 1995.

[22] J. Ding, M. Itoh and T. Yatagai, “Optimal incoherent correlator for noisy gray–tone image recognition”, *Optics Letters* **20**, pp. 2411–2413, 1995.

[23] B. Kumar, C. Hendrix and W. Shi, “An algorithm for designing phase–only filters with maximally sharp correlation peaks” in *Advances in Optical Information Processing 4*, , D. Pape, ed., vol. 1296 of *Proceedings of the SPIE*, pp. 132–139, 1990.

[24] M. Alam and Y. Gu, “Sobel operator based multiobject joint transform correlation”, *Optik* **100**, pp. 28–32, 1995.

[25] Q. Tang and B. Javidi, “Multiple–object detection with a chirp–encoded joint transform correlator”, *Applied Optics* **32**(26), pp. 5079–5088, 1993.

[26] B. Reddy and B. Chatterji, “An fft–based technique for translation, rotation and scale–invariant image registration”, *IEEE Transactions on Image Processing* **5**, pp. 1266–1271, 1996.

[27] Y. Lee and W. Rhodes, “Feature detection and enhancement by a rotating kernel min–max transformation” in *Optical Information Processing Systems and Architectures*, , B. Javidi, ed., vol. 1151 of *Proceedings of the SPIE*, p. 332, 1989.

[28] Y. Hsu, H. Arsenault and G. April, “Rotation–invariant digital pattern recognition using circular harmonic expansion”, *Applied Optics* **21**, pp. 4012–4015, 1982.

[29] C. Hester and D. Casasent, “Multivariant technique for multiclass pattern recognition”, *Applied Optics* **19**, pp. 1758–1761, 1980.