

SPG MITTEILUNGEN COMMUNICATIONS DE LA SSP

ÜBERSETZUNG - TRANSLATION

“Intelligent” Image Sensors

Bernhard Braunecker, Richard Hauck

Der Artikel erschien im Original auf Deutsch. - The original of this article appeared in German.

This article has been downloaded from:
http://www.sps.ch/fileadmin/articles-pdf/2017/Mitteilungen_ImageSensors.pdf

© see http://www.sps.ch/bottom_menu/impressum/

“Intelligent” Image Sensors

Bernhard Braunecker, Richard Hauck

Introduction

Digitization will quickly pervade all areas of daily life, including the home, the workplace, and the broader social environment. As autonomous microsystems, every household appliance, vehicle, and tool will not only collect and pass on data to superordinated systems in the cloud but will also take actions previously reserved for human beings. Estimates show that useful machine intelligence, as in the field of data interpretation, will in many respects be equal to the human brain in about twenty years. To achieve this goal requires high-performance image sensors. In the following article, we recall former sensor concepts which, thanks to technological progress, could now also be of interest in the context of industrial robots.

In-situ measurement technology

Any action to be performed on a workpiece requires its precise measurement. The object is first identified, and its surface quality is precisely assessed. In addition, detailed information about its current 3D position and orientation is required. Datasets are sent to computers, and special algorithms calculate the commands of the processing machine. Measurement and processing must be carried out quickly and simultaneously for reasons of efficiency, which requires the implementation of workpiece measurement as *in-situ* technology. Large amounts of data must then be transferred at high data rates to the cloud to allow for calculation of the relevant control parameters of the processing machine.

‘Sense and point-and-shoot’ optics

In addition to assembly robots, the use of which is well established, robots are increasingly playing important roles in material processing with the advent of pulse lasers such as femtosecond lasers, which offer oxidation-free surfaces of the highest quality as well as more flexibility, shorter through-

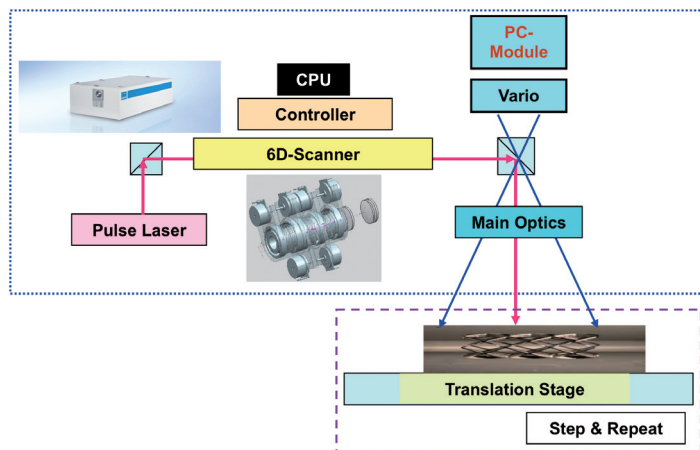


Fig. 1: Modern Sense, Point, and Shoot System: Using ultrashort laser pulses object structures like non-cylindrical microchannels of highest surface quality can be drilled in material bodies like turbine blades or stents. When applying the PC measuring module (Principal Components), already existing object structures can be identified and localized to guide the laser beam across the workpiece by means of a digitally controllable scanner.

put times, and higher accuracy than traditional mechanical tools such as drills and milling cutters. Material processing requires special bidirectional optics for the simultaneous measurement and laser processing of objects (‘Sense and point-and-shoot’ systems). These novel systems offer combinations of high-performance optics, high-speed scanners, and novel sensor modules (Fig. 1).

Optical pupils

The determination of the identity, surface condition, and position of an object is usually carried out in the image plane of an optical system, which corresponds geometrically and physically to the actual object. On the other hand, complete information about an object is also available in the pupil planes (diaphragms) of an optical system, where the light sources are usually imaged. Thus, analyzing wavefronts in the pupils allows one not only to access object information but also to influence it as desired. A mathematical Fourier relationship exists between both images, so each pupil manipulation also affects the object image.¹

In the following section, we present a system through which pupil manipulation allows one to detect single objects out of collectives. By means of special holographic components in the pupils, two of the problems addressed, namely object identification and data reduction, can then be efficiently solved.

Principal Component (PC) method for data collectives

The method of *principal components* (PC) provides highly effective data reduction. It assumes that objects, hereinafter also referred to as characters, detected by a sensor can be identified as components of known collectives. In the simplest case, these might be identified as letters of an alphabet. Then only the collective index, which consists of a few bits, is passed on to the cloud, where the character can be reconstructed unambiguously from stored collective data.

Therefore, if a character consisting of a large number of sampling elements is recognized as part of a collective of N characters, its class index can be described by a binary word consisting of only $K = \log_2(N)$ bits. Then K binary queries, that is, only K filter operations, would suffice: Filter 1 asks from which half of the collective the character originates, filter 2 from which quarter, etc. These principal component filters are thus not determined by individual characters, as with known matched filters, but by the collective.

How can this be quickly and robustly realized in the case of pictorial structures using modern lithographic technologies? In the following section, we recall a hybrid-optical (opto-electronical) method from 1979 through which analog-to-digital conversion for images was developed in detail, which

¹ One generally tries to keep the wavefronts in the pupils as aberration-free as possible, i.e. without any wavefront distortion to obtain diffraction-limited object imaging.

might attract new interest ^{2 3} due to the technical progress achieved over the past years.

Character recognition with spatially incoherent object illumination

We consider a self-luminous or diffuse-reflecting object which is imaged by an optical system on a 2D sensor at the location $\mathbf{x} = \mathbf{x}_{\text{Object}}$, $\mathbf{x} = (x, y)$ as an intensity distribution $I_{\text{Object}}(\mathbf{x})$. If a transparent hologram with the amplitude function $\tilde{u}_{\text{Filter}}$ is inserted into the exit pupil of the optics, the 2D correlation of the intensity functions I_{Object} and I_{Filter} is now obtained in the sensor plane instead of the object I_{Object} ,

$$I_{\text{Sensor}}(\mathbf{x} - \mathbf{x}_{\text{Object}}) = \iint d\mathbf{x}' \{I_{\text{Object}}(\mathbf{x}' + (\mathbf{x} - \mathbf{x}_{\text{Object}})) I_{\text{Filter}}(\mathbf{x}')\} \quad (1)$$

where $I_{\text{Filter}} = |u_{\text{Filter}}|^2$ and $u_{\text{Filter}} = \text{Fourier}(\tilde{u}_{\text{Filter}})$. Considering the intensity distribution at $\mathbf{x}_{\text{Object}}$, (1) reduces to the scalar product of I_{Object} and I_{Filter}

$$I_{\text{Sensor}}(0) = \iint d\mathbf{x}' \{I_{\text{Object}}(\mathbf{x}') I_{\text{Filter}}(\mathbf{x}')\} := I_{\text{Code}} \quad (2)$$

where I_{Code} is the chosen code word of the object.

How are the filter intensities I_{Filter} and the hologram functions $\tilde{u}_{\text{Filter}}$ calculated from these? For this purpose, each of the N object functions I_{Object} is digitized into M sampling points and all $N \times M$ values thus obtained are inserted as columns into a matrix $\mathbf{O}_{M,N}$ or as lines in the transposed matrix $\mathbf{O}_{N,M}^+$. Furthermore, the code words I_{Code} of the length K assigned to the N characters are arranged in a matrix $I_{N,K}$ and thus the linear equation system (3) is obtained from Eq. (2)

$$\mathbf{O}_{N,M}^+ \mathbf{F}_{M,K} = I_{N,K} \text{ mit } I_{N,K} = \text{code word.} \quad (3)$$

From the equation we can deduce the matrix $\mathbf{F}_{M,K}$ of the K filter functions I_{Filter} and calculate it under the condition $M \geq N$ according to

$$\mathbf{F}_{M,K} = \mathbf{O}_{M,N} (\mathbf{O}_{N,M}^+ \mathbf{O}_{M,N})^{-1} I_{N,K} := \mathbf{O}_{M,N}^\# I_{N,K} \quad (4)$$

with the help of $\mathbf{O}_{M,N}^\#$, the generalized inverse of $\mathbf{O}_{M,N}$.

Since $\mathbf{O}_{M,N}^\#$ can become unstable in the case of very similar characters, stochastic distortions $\delta\mathbf{O}$ in the objects would result in strong fluctuations of the code signal values $\delta I_{N,K} = \delta\mathbf{O}_{N,M}^+ \mathbf{O}_{M,N}^\# I_{N,K}$ and thus would deteriorate the signal-to-noise ratios $\{I_{N,K} / \delta I_{N,K}\}$. Therefore, when calculating Eq. (4), the variance

$$\sigma_{K,K}^2 = \langle \delta I_{K,N}^\# \delta I_{N,K} \rangle = \mathbf{F}_{K,M}^\# \Sigma_{M,M}^2 \mathbf{F}_{M,K} \quad (5)$$

has to be simultaneously minimized using the covariance matrix of the input noise

$$\Sigma_{M,M}^2 := \langle \delta\mathbf{O}_{M,N} \delta\mathbf{O}_{N,M}^+ \rangle. \quad (6)$$

In the case of $M > N$, the redundancy in Eq. (4) can be used to increase the robustness of the identification (Info Box II).

² Optical character recognition based on nonredundant correlation measurement, B. Braunecker, R. Hauck, and A. W. Lohmann, *Applied Optics* / Vol.18, No.16 / 15 August 1979

³ Dissertation R. Hauck, Universität Erlangen-Nürnberg, 1980

As a final result, we obtain

$$\mathbf{F}_{M,K} = \Sigma_{M,M}^2 \mathbf{O}_{M,N} (\mathbf{O}_{N,M}^+ \Sigma_{M,M}^2 \mathbf{O}_{M,N})^{-1} I_{N,K} \quad (7)$$

which replaces Eq. (4). Instead of Eq. (5), the minimized signal variance is

$$\sigma_{K,K}^2 = I_{K,N}^\# (\mathbf{O}_{N,M}^+ \Sigma_{M,M}^2 \mathbf{O}_{M,N})^{-1} I_{N,K}. \quad (8)$$

PC coding

In the following section, we analyze the situation with $K = 6$ PC filters. If in Eq. (2) the binary values $[0, 1]$ are assigned to the intensities $I_{N,K}$, one can recognize $N = 2^6 = 64$ different characters. If on the other hand the measurement technique is capable of determining the intensity values $I_{N,K}$ with a resolution better than 50%, ternary code values $[0, 1, 2]$ ⁴ can be used to identify $N = 3^6 = 729$ characters with six filters. Thus, different locations and orientations of a workpiece which the robot must access can be defined as individual objects. With a little effort, even intensity measurements of 25% accuracy should be possible to recognize a character out of a collective of $N = 4^6 = 4'096$ characters. ⁵

Furthermore, it makes sense to evaluate not only the intensity value of the code pixel selected according to Eq. (2) for identification but also the intensity differences to its $L = 8$ nearest neighbors. The code pixel would thus be a kind of pivot point of this mini collective.

Optical design and imaging optics

In Fig. 2, the characters in the object field plane (left) are imaged with a vario optics in an intermediate image plane where either a CMOS sensor (digital version I) or the PC optics module (optical version II) is installed. The versions differ in the way in which Eq. (1) is executed, whether electrical or optical. In both cases, however, the size, position, and image sharpness of the characters must be measured in advance to drive the vario zoom motors in such a way that the characters appear with constant size ⁶ and are best focused in the intermediate image. The vario is telecentric on the image side, as are most sensor optics, so the light beam bundles hit the sensor plane perpendicularly. Then the corresponding entrance pupil into which each object point emits is located in the front focal plane of the vario.

Since the code signals are derived from the measured intensities according to Eq. (2), the illumination intensity must be known in the intermediate image plane. This depends on the illumination profile of the light source, the lambertian radiation characteristic of the diffuse-reflecting objects, and the vignetting of the optics. All three effects depend on the position of a character in the object field, which, however, can be largely precompensated by a gray gradient filter in the intermediate image plane. The remaining intensity variations, such as by changes in illumination, are measured,

⁴ The ternary PC coding was experimentally verified in 2 with x-rays (!).

⁵ In the Kanji-Kentei test (https://en.wikipedia.org/wiki/Kanji_kentei) which is carried out every year by the Japanese government, this number of kanji characters must be identified.

⁶ The control of the vario magnification can be omitted if characters of different size are registered as individual objects in the collective.

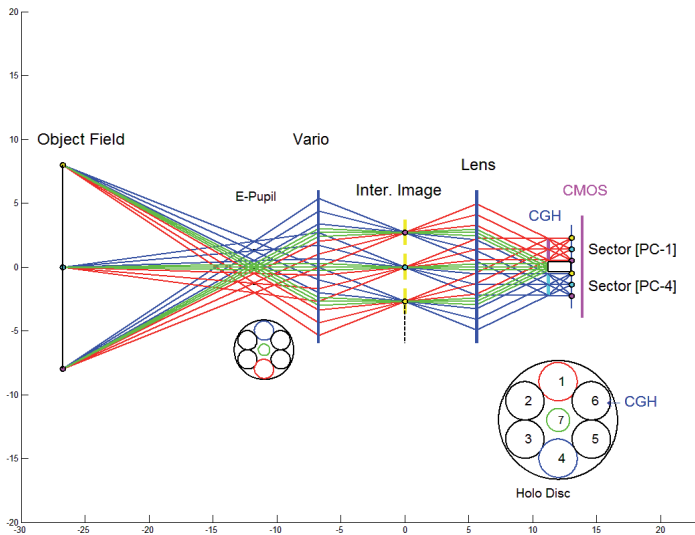


Fig. 2: Beam path through vario and PC module in the case of incoherent illumination. The 6 PC-CGHs are placed in the exit pupil of the single lens (Lens) as single component (Holo disc). Then the entrance pupil of the vario, into which all object points emit, is virtually divided into 6 sub-pupils. More details in fig. 3.

stored in look-up tables, and introduced as field-dependent correcting factors in Eq. (2). These determine the intensity threshold values of the code assignment. Note that these procedures apply to the digital and the analog-optical realization in the same way.

Variant I: Digital realization of character recognition

In this variant, the measured intensity values are directly digitally integrated with the six filter functions $F_{M,K}$ according to Eq. (2).

Variant II: Analog-optical realization of character recognition

Recognition can be carried out optically for fast industrial processes. For this purpose, a special optical module is connected to the vario (Fig. 2). It consists of a single lens of constant focal length f_{Obj} , a carrier disc with six computer-generated holograms (CGHs) in hexagonal arrangement, and a CMOS sensor. All six CGHs described by the transmission function \tilde{u}_{Filter} are calculated as diffractive components including a constant optical power equivalent to the focal length $f_{Lenslet}$ and are lithographically etched on a common disc of fused silica (Info Box III). The CMOS sensor is mounted on the disc at a distance $f_{Lenslet}$.

The PC module is inserted in the vario image plane, which is the front image plane of the PC single-lens. As seen in Fig. 2, the entrance pupil of the vario is divided into six sub-apertures defined by the six CGHs. The light bundle emitted from each object point into the vario entrance pupil is therefore split into six equal light cones, which illuminate the CGHs. The wavefronts of the six bundles act as carriers of the spatial spectrum of the entire object field.⁷

How do we calculate the CGHs from the filter functions $F_{M,K}$? For this purpose, we calculate $u_{M,K} = \sqrt{F_{M,K}}$ and obtain the hologram functions \tilde{u}_{Filter} by means of a Fourier transformation. These can be implemented as spatially modulated dif-

fraction gratings, where the desired filter function usually occurs in one of the two first diffraction orders. However, unwanted higher diffraction orders can appear together with the zero order, which unnecessarily occupy space in the sensor plane and which could possibly interfere with the image configuration of an adjacent hologram (crosstalk).

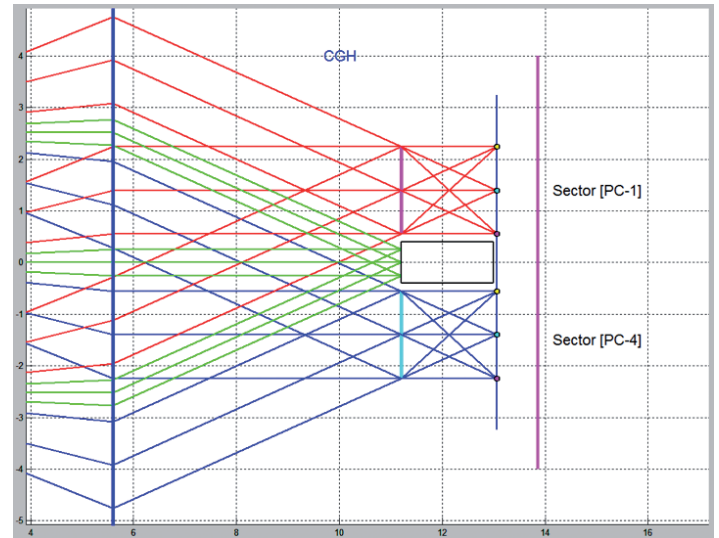


Fig. 3: A section is shown through PC-CGHs No. 1 and No. 4 from fig. 2. Each of the 6 CGH multiplies the object spectrum with its assigned PC filter function. Additionally a focusing power is part of the design function of each CGH, which causes a Fourier transform of the filtered object spectrum and so generates the PC-code signal on the common CMOS sensor. It is further recommended to add prismatic terms to the CGH design in order to shift the resulting code image on the CMOS. Then a rectangular CMOS area can be completely filled. The central black box represents a 45° mirror that directs the unmanipulated light beam of the 7th sub-pupil (green) perpendicular to the paper plane. This light beam is received by a separate sensor module to analyze the image sharpness and the radiometry. It is of further advantage to insert there a hologram to generate a visual frame with calibration marks, which is projected by the lens and the vario back to the object plane to facilitate the object recognition.

Since, however, the phase in the sensor plane can be freely chosen in the case of incoherent illumination, $u_{M,K}$ can be multiplied appropriately by phase factors to approximate a homogeneous amplitude distribution $|\tilde{u}_{Filter}| \sim 1$ by means of special iterative algorithms⁸. The CGHs can then be lithographically realized by the manufacturer as a pure phase structure. These holograms, also called kinoforms, produce the desired filter function u_{Filter} on-axis, thus with high efficiency and without disturbing diffraction orders in the sensor plane.

Consideration of the space bandwidth product M

If a pre-analysis of the N characters reveals that a spatial resolution δx_{Object} is at least required to distinguish between two characters, the characters of size Δx_{Object} would have to be digitized with $M = \{\Delta x_{Object} / \delta x_{Object}\}^2$ sampling points. For example, δx would be the size of the cross-string in the letter Q to clearly distinguish it from the letter O. The optics when illuminated by the light of the wavelength λ should at least resolve spatial details diffracted into the

⁷ In Info Box I is shown that, in the case of spatial coherent illumination, the tilted wavefronts of six laser diodes are the carriers of the object spectrum.

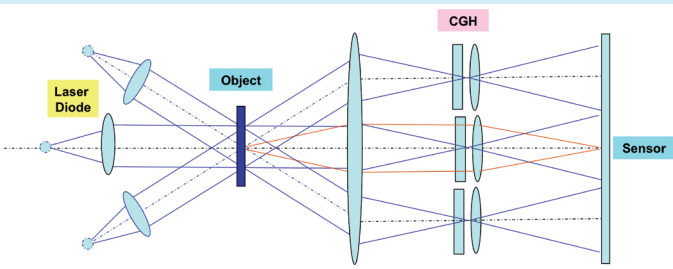
⁸ R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik* 35, 237 (1972)

I: Spatial coherent illumination

In the case that the objects to be identified are recorded on photographic film, one can illuminate the transparency by the tilted wavefronts of K laser collimators. They transfer the spatial object information to the K CGHs. We obtain in the sensor plane the intensity of the 2D correlation of the amplitude functions u_{Object} and u_{Filter}

$$I_{\text{Sensor}}(\mathbf{x} - \mathbf{x}_{\text{Object}}) = \left| \iint d\mathbf{x}' \{u_{\text{Object}}^*(\mathbf{x}' + (\mathbf{x} - \mathbf{x}_{\text{Object}})) u_{\text{Filter}}(\mathbf{x}')\} \right|^2,$$
 which reduces at the location of the purely geometric mapping $\mathbf{x} = \mathbf{x}_{\text{Object}}$ to the 2D integral

$$I_{\text{Sensor}}(0) = \left| \iint d\mathbf{x}' \{u_{\text{Object}}^*(\mathbf{x}') u_{\text{Filter}}(\mathbf{x}')\} \right|^2 := I_{\text{Code}}.$$



\pm first order at the angle θ_{Diff} with sufficient contrast, where $\sin(\theta_{\text{Diff}}) \sim \pm \lambda / \delta x_{\text{Object}}$. In the arrangement of Fig. 2, the numerical aperture $NA = 1/(2 F\text{-number})$ of the vario on its image side is equal to that of the PC lens by construction, and it must be at least $3 \sin(\theta_{\text{Diff}})$ for the $K = 6$ PC subapertures in the hexagonal arrangement shown.

An example may illustrate this: If we select a PC lens with an F-number of 1.0, corresponding to $NA = 0.50$, we obtain $NA_{\text{SA}} \sim NA/3 = 0.167$ for each of the six PC subapertures. Thus, object details $\delta x_{\text{Object}} \sim \lambda / NA_{\text{SA}} = 3.37 \mu\text{m}$ can be resolved by each PC module at a wavelength of $0.56 \mu\text{m}$. If about 50 resolution points per character in one dimension are sufficient to ensure a secure identification,⁹ this would correspond in the intermediate image to a character size $\Delta x_{\text{Object}} = 0.168 \text{ mm}$.

A CMOS array (1/2") with 16 Mpix and $1.12 \mu\text{m}$ pitch is provided as a common sensor for all six subpupils. This results in the imaging ratio $V = 3.37 \mu\text{m} / 1.12 \mu\text{m} = 3 = f_{\text{Obj}} / f_{\text{Lenslet}}$. If one assigns one third of the available 4'000 CMOS pixels to each PC channel, this would cover $\Delta x_{\text{CMOS}} = 1'333 * 1.12 \mu\text{m} = 1.49 \text{ mm}$ on the CMOS sensor or $\Delta x_{\text{FoV}} = 4.47 \text{ mm}$ in the object plane. One could therefore detect $Q = \{4.47 \text{ mm} / 0.168 \text{ mm}\}^2 = 708$ characters simultaneously.

We still can choose the focal length f_{Obj} to determine the size and resolution of the CGH. A $f_{\text{Obj}} = 5.6 \text{ mm}$ would yield $\Delta x_{\text{CGH}} = 2 NA_{\text{SA}} f_{\text{Obj}} = 1.87 \text{ mm}$ with $\delta x_{\text{CGH}} = 1.87 \text{ mm} / 50 = 37 \mu\text{m}$ samples. Details about the fabrication of synthetic holograms are summarized in Info Box III. This clearly shows that our PC application in no way exploits the potential of this fascinating technology.

If PC character recognition is carried out purely digitally, the number of floating point operations to be executed is

⁹ With $M = 50 \times 50 = 2'500$ sampling points, the ternary coding of the 6 PC filters could be applied for a collective with $N = 3^6 = 729$ characters.

Advantages and disadvantages of spatially coherent and incoherent illumination

All K PC channels can be operated independently and simultaneously in both methods. The coherent method has the advantage that complex-valued PC filter functions can be realized as CGHs. The disadvantages are that the objects must be transparent and that the CGH filters must be adjusted very precisely. This is different in the case of the incoherent method, where $\tilde{u}_{\text{Filter}}$ is not the transfer function, but $\tilde{I}_{\text{Filter}} = \tilde{u}_{\text{Filter}}^{(**)} \tilde{u}_{\text{Filter}}$, where $(**)$ denotes the autocorrelation. Then the CGHs can even be moved during the measurement. The disadvantage of the incoherent method is that the algorithm of (4) could also yield negative values of $F_{\text{M,K}}$ that could not be optically realized. This, however, can be avoided either by means of clever assigning the code indices to the characters, or by expressing the calculated F as difference of two positive quantities

$$F = F_{\text{I}} - F_{\text{II}} \text{ with } F_{\text{I}} \geq 0 \text{ and } F_{\text{II}} \geq 0.$$

Consequently, this would require two identical optical measurement channels for each PC filter.

$P = \{Q \times K \times M \times (L+1)\}$ with Q number of characters in the field of view, K number of PC filters, M number of pixels per character, and L number of nearest neighbors. Using the values of our example [$Q=708$, $K=6$, $M=2500$, $L=8$] and assuming a computing time of $t = 2 \text{ ns}$ per floating point operation of a commercially available processor, we obtain $T = P * t = 0.19 \text{ s}$. The equivalent optical parallel process, on the other hand, is independent of P and would last only 0.1 ns , taking into account a distance of 3 cm between the intermediate image plane and the detector.

Summary

Ever-increasing data flows have brought existing communication channels to the limits of their capacities. If additional sensors are to be expected everywhere in the industry and around the clock, a massive increase in data transmission capacity must result. It makes sense to combine the measurement, processing, and transfer of image data as a total process to minimize throughput time. While sensor technology and data processing could merge to create smart sensing solutions, business models must take into account that the value of a sensor is determined solely by the quality of processed information sent to the cloud.

Data reduction should therefore take place immediately after or more ideally during optical data acquisition. Maximum efficiency is achieved if a priori information about the characters or structures to be identified is available so that those characters can be assigned to a collective. The character collective can contain different positions, sizes, orientations, and distortions of the same item as single objects as long as these are meaningfully distinguishable. All characters in the field of view of the optics can then be transformed into code indices as registered in the collective. Only these few bits must be passed electronically to the cloud. The transformation can be performed digitally or optically.

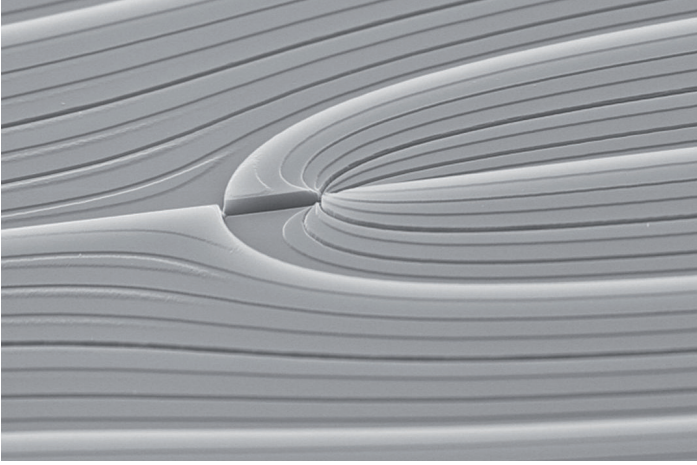


Fig. 4: Optical phase structure of a digital optical element DOE, etched in fused silica.

http://www.iap.uni-jena.de/Micro_+structure+technology.html

The PC concepts were developed by both authors at the beginning of the 1980s at the Physical Institute of the University of Erlangen-Nürnberg. After further common years at the University in Essen, RH joined the company *Audiodev* as head of the development and production of measuring instruments for surface coatings, while BB moved to *Leica Geosystems* in Heerbrugg, where he was responsible for the optics development.

In the optical variant, the light beams emitted by each object character are transmitted to, for example, six synthetic CGHs, which are lithographically etched on a common carrier disc of fused silica. Insertion of the disc into the PC module is not mechanically critical. The PC total module is only several centimeters in size and can easily be added to the back end of modern digital cameras.

II: System optimization

A) PRINCIPLE OF OPTIMIZATION

In the following we denote the solution vectors of the *inhomogeneous* equation system (4) as $\mathbf{E}_{M,K}$ with the variance of the signal values of (5), which has to be minimized. The covariance matrix $\Sigma_{M,M}^2$ of the input characters according to (6) is separable in the case of uncorrelated noise. It is diagonal and thus can be written as $\Sigma_{M,M}^2 = \mathbf{C} \mathbf{C}^+$.

The equation system (4) is redundant in most application cases, which means $M > N$. This can be used to minimize $\sigma_{K,K}^2$ by also solving the *homogeneous* equation

$$\mathbf{O}_{N,M}^+ \mathbf{G}_{M,R} = \mathbf{0}_{N,R} \quad (\text{A-1})$$

with $R = M - N > 0$ and the zero matrix $\mathbf{0}_{N,R}$.

The R filter vectors $\mathbf{G}_{M,R}$ can be combined with arbitrary coefficients $\mathbf{g}_{R,K}$ to $\mathbf{H}_{M,K} = \mathbf{G}_{M,R} \mathbf{g}_{R,K}$, to get the general solution vectors $\mathbf{F}_{M,K} = \mathbf{E}_{M,K} + \mathbf{H}_{M,K}$ of (3).

Note that the $\mathbf{H}_{M,K}$ are orthogonal to $\mathbf{E}_{M,K}$ with $\mathbf{E}_{K,M}^+ \mathbf{H}_{M,K} = \mathbf{f}_{K,N}^+ (\mathbf{O}_{N,M}^+ \mathbf{O}_{M,N})^{\pm 1} \mathbf{O}_{N,M}^+ \mathbf{G}_{M,R} \mathbf{g}_{R,K} = \mathbf{0}_{K,K}$ because of (A-1).

The minimization of $\sigma_{K,K}^2$ leads to

$$\mathbf{H}_{M,K} = -\mu_{M,M} \mathbf{E}_{M,K}$$

$$\text{with } \mu_{M,M} = \mathbf{G}_{M,R} (\mathbf{G}_{R,M}^+ \Sigma_{M,M}^2 \mathbf{G}_{M,R})^{-1} \mathbf{G}_{R,M}^+ \Sigma_{M,M}^2 \quad (\text{A-2a})$$

and thus to

$$\mathbf{F}_{M,K} = \mathbf{E}_{M,K} + \mathbf{H}_{M,K} = (\mathbf{1}_{M,M} - \mu_{M,M}) \mathbf{E}_{M,K} \quad (\text{A-2b})$$

and to

$$\sigma_{K,K}^2 = \mathbf{E}_{K,M}^+ (\mathbf{1}_{M,M} - \mu_{M,M}^+) \Sigma_{M,M}^2 (\mathbf{1}_{M,M} - \mu_{M,M}) \mathbf{E}_{M,K} \quad (\text{A-2c})$$

B) OPTIMIZATION WITH REDUCTION OF THE COMPUTATION EFFORT

The above optimization requires the inversion of the large $R \times R$ matrix in (A-2a). The large numerical effort can be avoided by introducing an auxiliary operator \mathbf{V} .

Step 1: Modification of the objects

The auxiliary operator is applied to all objects. We thus obtain the modified objects $\hat{\mathbf{O}}_{M,N} = \mathbf{V}_{M,M}^+ \mathbf{O}_{M,N}$. $\mathbf{V}_{M,M}^+$ is a diagonal matrix with still arbitrary coefficients.

Step 2: Optimization

Now the same procedure as described in A) is applied to the modified objects $\hat{\mathbf{O}}_{M,N}$. One gets the adequate filters of the objects in analogy

$$\begin{aligned} \hat{\mathbf{E}}_{M,K} &= \hat{\mathbf{O}}_{M,N} (\hat{\mathbf{O}}_{N,M}^+ \hat{\mathbf{O}}_{M,N})^{-1} \mathbf{I}_{N,K} \\ &= \mathbf{V}_{M,M}^+ \mathbf{O}_{M,N} (\mathbf{O}_{N,M}^+ \mathbf{V}_{M,M}^2 \mathbf{O}_{M,N})^{-1} \mathbf{I}_{N,K} \end{aligned}$$

as well as the optimized orthogonal filters

$$\hat{\mathbf{H}}_{M,K} = (\dots) \mathbf{G}_{R,M}^+ \Sigma_{M,M}^2 \hat{\mathbf{E}}_{M,K}$$

with $(\dots) =$ terms from (A-2a).

In a next step the auxiliary operator will be merged with the PC-filters. Then the modified filters $\check{\mathbf{E}}_{M,K} = \mathbf{V}_{M,M} \hat{\mathbf{E}}_{M,K}$ again operate on the original objects:

$$\begin{aligned} \check{\mathbf{E}}_{M,K} &= \mathbf{V}_{M,M} \hat{\mathbf{O}}_{M,N} (\hat{\mathbf{O}}_{N,M}^+ \hat{\mathbf{O}}_{M,N})^{-1} \mathbf{I}_{N,K} \\ &= \mathbf{V}_{M,M}^2 \mathbf{O}_{M,N} (\mathbf{O}_{N,M}^+ \mathbf{V}_{M,M}^2 \mathbf{O}_{M,N})^{-1} \mathbf{I}_{N,K} \end{aligned}$$

as well as the optimized orthogonal filters

$$\begin{aligned} \check{\mathbf{H}}_{M,K} &= (\dots) \mathbf{G}_{R,M}^+ \Sigma_{M,M}^2 \check{\mathbf{E}}_{M,K} = (\dots) \mathbf{G}_{R,M}^+ \Sigma_{M,M}^2 \mathbf{V}_{M,M} \hat{\mathbf{E}}_{M,K} \\ &= (\dots) \mathbf{G}_{R,M}^+ \Sigma_{M,M}^2 \mathbf{V}_{M,M} \hat{\mathbf{O}}_{M,N} (\dots) \\ &= (\dots) \mathbf{G}_{R,M}^+ [\Sigma_{M,M}^2 \mathbf{V}_{M,M}^2] \mathbf{O}_{M,N} (\dots). \end{aligned}$$

Step 3: Special selection of the auxiliary operator

Selecting $\mathbf{V}^2 = \Sigma^2$, leads to the diminishing of

$$\check{\mathbf{H}}_{M,K} = (\dots) \mathbf{G}_{R,M}^+ \mathbf{O}_{M,K} (\dots) = \mathbf{0}_{M,K} \text{ because of (A-1).}$$

Then the vectors $\check{\mathbf{E}}_{M,K}$ are already the optimal solution, but with the advantage that the large $R \times R$ matrix from (A-2a) has not to be inverted. As result of the optimization one gets the terms of (7) and (8)

$\check{\mathbf{E}}_{M,K} = \Sigma_{M,M}^2 \mathbf{O}_{M,N} (\mathbf{O}_{N,M}^+ \Sigma_{M,M}^2 \mathbf{O}_{M,N})^{-1} \mathbf{I}_{N,K}$ instead of (4), and the minimized signal variance

$$\sigma_{K,K}^2 = \mathbf{f}_{K,N}^+ (\mathbf{O}_{N,M}^+ \Sigma_{M,M}^2 \mathbf{O}_{M,N})^{-1} \mathbf{I}_{N,K} \text{ instead of (5).}$$

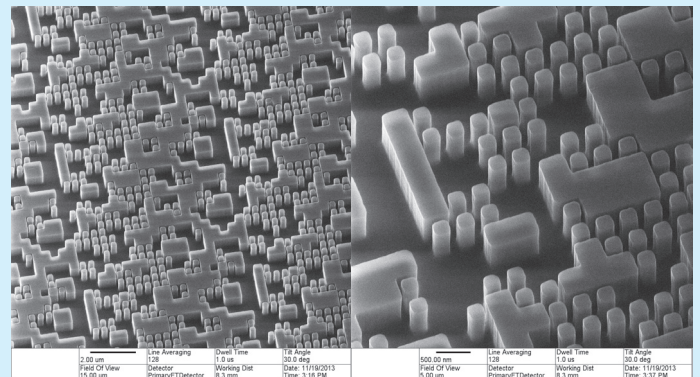
Note, that the inversion of $M \times M$ matrix Σ is trivial due to its diagonality. Obviously $\check{\mathbf{E}}_{M,K}$ fulfills the PC code equation $\mathbf{O}_{N,M}^+ \check{\mathbf{E}}_{M,K} = \mathbf{I}_{N,K}$.

III: Microstructure technology

Fraunhofer High Performance Center for Photonics at Jena
 Friedrich Schiller University / Institute of Applied Physics
 (http://www.iap.uni-jena.de/Micro_+structure+technology.html)

Micro- and nanostructures are well established in optics for many years. They allow to realize more and more important optical properties and functions in a unique way. Two typical examples are optical gratings in spectrometers and computer generated holograms (CGH) for the interferometric testing of freeform optical surfaces. Because of the permanent progress of lithography smaller and more precise structures are applicable to scientific instruments as well as to industrial products. The Fraunhofer Photonics Center at Jena University with its Institute of Applied Physics runs powerful micro- and nanostructuring facilities since years. The structuring process starts with electron beam lithography (SB350 OS, Vistec company) or with a LED-recorder (HIGHFIVE, self-developed). Both devices operate at the leading edge of technology. The SB 350 OS electron beam lithography system works with a „variable shape beam“ or with „cell projection“ on large substrates up to 300 mm edge length and 15 mm thickness. Very large writing speeds are achieved for repetitive

nanostructures. For example to write a grating with a 200 nm period (100 nm line pairs) on an area of 100 mm x 100 mm requires only about 6 hours. Thereby the placement accuracy of the structure is better than 1.5 nm rms, and the accuracy of the grating period can be kept below 5 pm at 10 mm integration length. The structures can be etched in substrates like fused silica, silicon or material stacks. Based on several process steps multi-level profiles can be realized and thus almost any diffractive structure consisting of several phase functions can be produced.



DOE based on effective media, pixel size of diffractive pixels 380 nm, within the pixel a substructure realizes an effective refractive index, material is fused silica.

This article has explored the inherent potential of optical image processing. Thanks to recent progress in optics, lithography, and sensor technology, optical methods are suitable for intelligent and extremely fast analyses of very large amounts of data. Former fears that optical processing would not be practical because of size and high precision requirements are today no longer valid. Miniaturization and accuracy are ensured by modern manufacturing technologies, while installation in commercially available cameras is subject to standard production tolerances. However, the task of

an optical processor must be well defined, since reprogramming, that is, changing CGHs, always takes more time with optics than with a digital computer.

The optical parallel process runs in fractions of nanoseconds, whereas in the case of the digital variant, the computation can last parts of seconds depending on the number of floating point operations to be carried out. Our article should also encourage institutes to reconsider programmed optics. After all, the 21st century is that of the photon!

Addendum to the optimization

In addition to Info Box II we want to show by physical arguments, why the optimization also reduces the redundancy in a collective of N characters. From a heuristic point of view, the auxiliary operator V can be understood as a gray level filter superimposed on all objects. Due to the special choice $V = 1/\Sigma$, where Σ^2 is the input noise covariance of Equation (6), the filter V attenuates the object pixels proportional to their noise level. It is physically reasonable to model the pixel noise as uncorrelated and multiplicative, i.e. black pixels remain black without noise, whereas bright pixels are the more noisy, the brighter they are. Then the diagonal elements of matrix $\Sigma_{M,M}^2$ depend

on the sum over the object intensity functions $O_{m,n}$:

$$\Sigma_{m,m}^2 = \sigma_{\text{Object}}^2 \sum_n O_{m,n}^2 \quad (n = 1 \dots N, m = 1 \dots M),$$

where σ_{Object} describes the amount of noise.

Consider as example an alphabet of bright letters on a dark background. The letters {H, E, F, B} look similar due to the common vertical bar and are difficult to distinguish. The value Σ^2 for the pixels of the bar is increased by the summation over all N letters, which consequently leads to a stronger attenuation by filter $V = 1/\Sigma$. In conclusion, the filter V reduces the noise contributions and so stabilizes the recognition process.