Full Length Article

The ultimate performance of the Rasnik 3-point alignment system

Harry van der Graaf [a,b,*], Alessandro Bertolini [a], Joris van Heijningen [c], Bram Bouwens [d], Nelson de Gaay Fortman [e], Tom van der Reep [f], Lennart Otemann [g]

[a] *Nikhef, Science Park 105, Amsterdam, The Netherlands*
[b] *Delft University of Technology, Mekelweg 1, Delft, The Netherlands*
[c] *Centre for Cosmology, Particle Physics and Phenomenology (CP3), Université catholique de Louvain, Louvain-la-Neuve, Belgium*
[d] *Amsterdam Scientific Instruments, Science Park 106, Amsterdam, The Netherlands*
[e] *AMOLF, Science Park 104, Amsterdam, The Netherlands*
[f] *University of Amsterdam, Nieuwe Achtergracht 129, Amsterdam, The Netherlands[1]*
[g] *Nikhef and TU Delft, The Netherlands[2]*

ARTICLE INFO

ABSTRACT

The Rasnik system is a 3-point optical displacement monitor with sub-nanometer precision. The CCD-Rasnik alignment system was developed in 1993 for monitoring the alignment of the muon chambers of the ATLAS Muon Spectrometer at CERN. Since then, the development has continued as new CMOS imaging pixel chips became available. In this work the system processes and parameters that limit the precision are studied. We conclude that the spatial resolution of Rasnik is only limited by the quantum fluctuations of the photon flux arriving at the pixels of the image sensor. The results of two Rasnik systems are compared to results from simulations, which are in good agreement. The best spatial resolution obtained was $7 \text{ pm}/\sqrt{Hz}$. Finally, some applications of high-precision Rasnik systems are set out.

1. Introduction

The **R**ed **A**lignment **S**ystem **Nik**hef (Rasnik) was originally developed for the alignment of the inner, middle and outer muon chambers of the Muon Spectrometer of the L3 experiment at CERN (1980 - 2001) [1,2]. Units of three position-sensitive detectors, placed at a certain mutual distance, were used to measure the radius-of-curvature of muon tracks, from which the muon momentum can be derived. The precision of the measured muon momentum depends crucially on the alignment of the three chambers. Since it is practically impossible to mechanically constrain the three chambers in their nominally aligned position, the actual alignment is monitored continuously instead. The measured radius-of-curvature is corrected using the data from the alignment system. In this way, the muon momentum measurement is no longer affected by any change in the position of a muon tracking chamber.

Data from the 8000 operational ATLAS Rasnik systems revealed a much better intrinsic performance than the original ATLAS requirement (30 μm). In several studies, a sub-μm spatial resolution was demonstrated [1]. This paper reports the results of MonteCarlo simulations, and the results of two CCD-Rasnik systems, build for this purpose. First,

in Section 2, the basic principles of the CCD-Rasnik system are set out. In Section 3, all error sources are addressed and their contribution is analysed and quantified. Then, in Section 4, the MonteCarlo (MC) image generator is set out, followed by a description of the image analysis routine. By using this routine for analysing MC images, the performance of Rasnik systems can be estimated. The expected performance is compared to the results from the two hardware Rasnik systems in Section 5. Finally, possible improvements for future systems are listed and quantified.

2. Principles

The basic principle of the CCD-Rasnik alignment system is shown in Fig. 1. The **optical axis**, running from the centre of the image sensor through the optical centre of the lens, crosses the mask at a specific point. This point is projected in the centre of the images from the sensor. The image analysis routine yields an x and y coordinate of this crossing point, with the mask pattern defining the coordinate system. Given the known mask pattern and the pitch of the pixels of the image sensor, the scale of the image can be obtained as third output parameter. This scale S equals the ratio of the image and object
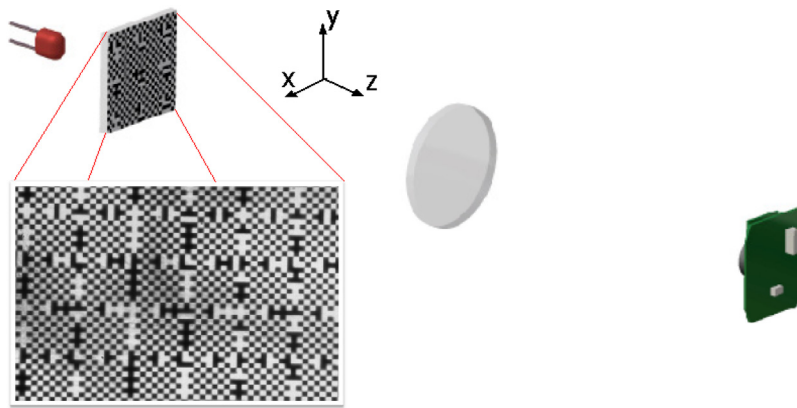
Fig. 1. Principle of the CCD-Rasnik system. An image of the LED back-illuminated coded mask is projected onto the image pixel sensor by means of a positive singlet lens. The sensor is connected to a PC by means of Ethernet or USB. The images are processed, resulting in four output parameters describing the alignment of mask, lens and sensor. If the lens and image sensor are fixed together (forming a camera), Rasnik turns into a system that monitors the 4D position of the coded mask.
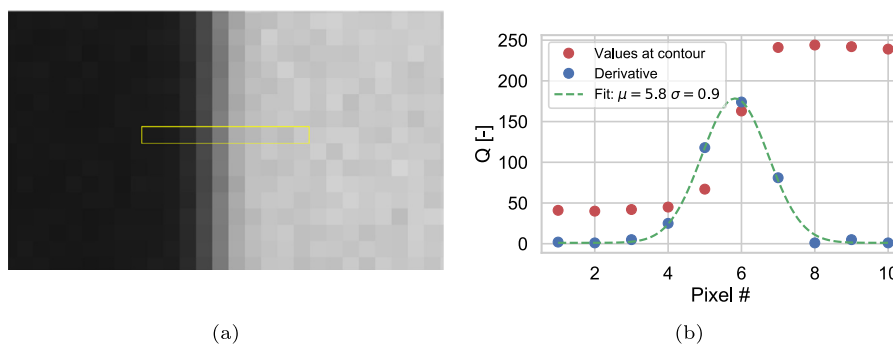*Source:* This figure is reproduced from Ref. [1].

(a)

(b)

Fig. 2. (a) A zoomed-in image of a vertical contour. (b) Red: the light intensity Q_n on a row of $n = 0 - 9$ pixels between values 0–255. Blue: differentiation: $Q_n = Q_{n+1} - Q_{n-1}$.

distances: S = 1 if the lens is in the centre between the mask and the sensor. The fourth output parameter is the differential rotation θ_z of the mask and the sensor around the optical (z) axis.

The pattern of the coded mask

The position of the image on the sensor is defined by black–white contours of the pattern: see Fig. 2(a). The black-to-white transition is confined, in this case, by a row of 10 pixels. The content of the pixels is displayed in the right-hand picture in red dots, forming a Fresnel edge diffraction curve. After differentiation, the top of the Gaussian-like curve through the blue dots is a measure for the contour position. In standard Rasnik systems (pixel pitch of $10\,\mu m$, and sensor dimension of 4 mm × 3 mm), this contour position is determined with a precision of $\sim 0.2\,\mu m$. This measurement is performed 15k times along the total contour length of 100 mm, both in the x and y directions, resulting in a spatial resolution of $\sim$1.5 nm *per image*.

Of all possible patterns, the chessboard pattern has the highest possible contour length in the x and y direction. Note also that the contour length is equal in both directions. The smaller the chess fields, the longer the total contour lengths. Because of diffraction, however, the ratio of the white and black pixel contents decreases with decreasing chess field size, approaching unity for very small chess fields (gray image). This implies that there is one optimal chess field size.

The **range of measurement** of a Rasnik system is determined by the size of the coded mask: if the optical axis crosses a plain chess field coded mask, an analysable image will be generated. For a pure chess field pattern, the x and y response of the image analysis routine versus a mask displacement would be periodical. To break the periodicity, each ninth column and row of chess fields is coded digitally by inverting black and white, setting a valid bit. In this way, 8 bits are available for identifying the 9th, 18th, 27th, etc. row or column, and an arbitrarily large mask (and therefore dynamic range) can be realised.

In this paper the **standard Rasnik system** is referred to, fixing system parameters. It has the lens (aperture 0.2) centred between mask and image sensor, a ChessField size of 100 μm, a sensor pixel size of 10 μm, light wavelength λ of 450 nm, and an image size of 400 × 300 pixels. The demonstrated essential spatial resolution of this system is 2 nm *per image*, in both directions.

3. Error sources

In this chapter the error sources limiting the performance of Rasnik are described in a comprehensive and complete way. The error contributions are analysed or estimated. The quantum nature of light, causing fluctuations in pixel charge signals, is found to be the fundamental limit for spatial resolution. This is confirmed by the results given in Section 5.

3.1. The propagation of light through air

Light rays are bent if the index of refraction n of the ambient medium has a gradient in a direction perpendicular to the direction of the ray. This causes a first order error in the measured image position. In general this error is proportional to the square of the distance between the sensor and the mask. The reduced index of refraction $(n - 1)$ of air is approximately proportional to the air specific density and thus pressure. Temperature gradients, and density fluctuations due to convection cause systematic shifts and variations with time in the measured image position. In addition, these variations may cause image deformation. The effects of density variations in ambient air are analysed in detail in [1].

Temperature gradients in air can be reduced by shielding the light path with, for instance, metal tubes. Another way is to mix the ambient

air by means of a set of fans. This causes, however, variations in air density, reducing the systematic error at the cost of an increase of the noise of the position measurement.

Another solution is to apply a medium with less variation in n. This could take the form of a glass or quartz bar with optical flat surfaces at the ends. These bars could be replaced by vacuum tubes equipped with optically flat windows at the ends. Finally, the complete system can be placed in vacuum, provided that heat dissipation, in particular of the image pixel sensor, is addressed. The vacuum should be such that the mean free path of the air molecules is in the order of the light beam diameter: a pressure of 1×10^{-2} mbar is usually adequate.

3.2. Errors in the coded mask

The image analysis routine yields the x and y coordinate of the image on the sensor. Each of these coordinates is the mean of order 1.5×10^4 measured values of contour positions of equal weight. The deviation of the nominal position of a section of a contour contributes therefore only by a fractional weight to the error in the position measurement. In modern mask technology, standard precision in contour position is better than 6 nm. In standard Rasnik systems, an image contains some 40 and 30 vertical and horizontal contours, respectively. The 6 nm error of a single contour results in a linearity error of order 1 nm in the position measurement.

Over a large range mask, the overall scale of a mask may depend on temperature, but by printing the mask on quartz or Zerodur [3], the scale error can be limited, if necessary. As a result, the linearity error of Rasnik can be as low as 1 nm over an arbitrarily large range. This error can be further reduced by applying state-of-the-art mask making technology [4].

3.3. Image distortion

Optical image distortions cause perfect image points (x, y) to appear at (x', y'). The deviation is usually rotational symmetric around the optical axis. An example is the well known pin-cushion effect. The deviations are constant, only depending on x and y, and therefore do not cause a first order error since the x and y response is the average of 1.5×10^4 (x, y)-points of equal weight. These points are quite evenly distributed over the image area, so identical deviations are constantly present in each image. A third order error in linearity may occur in the case an image edge nearly coincides with a contour: in this case some parts of the contour participate, whereas others do not. This linearity error is periodic with the chess field size and with twice the chess field size. The error can be quantified by making density plots of real data with random x and y, and an empirical correction can be applied.

3.4. Image blurring

Image blurring makes contours vague. The three contributions to image blurring are

- the finite pixel size of the sensor. If a contour crosses a pixel its light content will be between high and low. As a result there will be a gray pixel sandwiched between light and dark pixels: see Fig. 3(b).
- diffraction. This fundamental optical effect is essential for the Rasnik performance. The black–white transition can be described by the Fresnel edge diffraction curve (see Fig. 2(b)). The steepness of the slope is inversely proportional to the wavelength λ of the applied (led) light. The curve is non-symmetric, and a black–white transition is therefore not identical to a white–black transition.
- out-of-focus of sensor. This causes a linear transition from white to black. In practice, systems are well-focused. The effect could be used as fifth output parameter.

3.5. Errors associated with the image pixel sensor

Quantum fluctuations: shot noise. When light falls on the active area of a pixel of the sensor during the shutter open time, the absorbed photons create photo-electrons, charging up the well of the pixel: this charge is proportional to the light fallen onto the pixel. If N photons are absorbed on average, the pixel electron content will fluctuate according a normal distribution with RMS = $\sqrt{N}$ due to the quantum character of light.

In addition, each pixel content will establish electronic noise (due to dark current, bias current) which is, in the first order, equal for all pixels. The contribution of electronic noise is constant and small with respect to the quantum fluctuations, as can be expected for modern pixel sensors. In the following, pixel noise includes electronic shot noise and quantum fluctuations. The amplitude of the fluctuations can therefore be described as the sum of a squared constant term due to the electronic (dark current) noise, and a squared term proportional to the square root of the average pixel content, due to the quantum fluctuations. In Fig. 4 the measured fluctuations are depicted for some pixels chosen at random.

In practice, the applied MER-041-302GM pixel sensor can operate with 8-bit and 12-bit readout. For light levels above 20% of the maximum, the pixel noise is larger than the smallest 8-bit ADC step in pixel content. The 12-bit conversion, which reduces the maximum frame rate by a factor of order 2, does not improve the performance of the pixel sensor in our application.

The S/N ratio of the pixel content is improved when more light is absorbed. There is, however, an upper limit to the light level that CCDs or CMOS image pixels can accept before charge leaks away from the wells. The amount of light, collected during the shutter time, resulting in the maximum ADC output of 255 and 4095 for the 8-bit and 12-bit conversion, respectively, matches the well capacity. This limitation of CCDs and CMOS pixel image sensors can be reduced by applying sensors with a higher frame rate, but the development a special ASIC pixel sensor, capable to handle much higher light levels, would be the preferred solution.

Pixel, row and column geometry. Although image pixel sensors are mechanically very precisely made, thanks to state-of-the-art lithography, some pixel rows or columns may be wider than nominal, for instance. Such a deviation results in an constant image distortion as addressed in Section 3.3: the effect can be discarded.

Pixel response: slope and pedestal, pixel cross talk. For each pixel in the image sensor, the linear relation between pixel content and absorbed light is described by a 'slope' constant and a 'pedestal' constant. Although these constants vary little over all pixels, the variation may have an effect on the spatial resolution, which can be estimated.

A too large pixel content, due to a deviating slope and/or pedestal constant, may virtually 'pull or push' the calculated position of the nearest local contour, in the x and y directions. Since the direction of this virtual 'force' has equal chance to be $-x$ or $+x$, respectively $-y$ or $+y$, the average force will be zero when taken the average of the deviations associated with 10^5 pixels. Pixel cross talk can be treated as a fourth source of image blurring, and is negligible with respect to diffraction.

Pixel Flicker noise. Since the image pixel sensor is a crystalline silicon solid, it may be considered as constant in time in terms of shape and dimension. This is also the case for the coded mask and the lens. A pixel output can be characterised in terms of a slope and pedestal. Flicker noise is due to changes in the depletion layer of pixels causing changes in characterisation and is one of the components of so-called 1/f, or pink, noise. This variation has little effect on the measured image position, as has been argued in the previous item. Even in the case of a common local group behaviour of pixels, the pulling or pushing of contours is randomised. Since 5×10^5 pixels participate, it will be a challenge to measure the practical Flicker noise, cleaned from temperature effects. Equivalently, the drift in Rasnik systems with time should be close to zero.
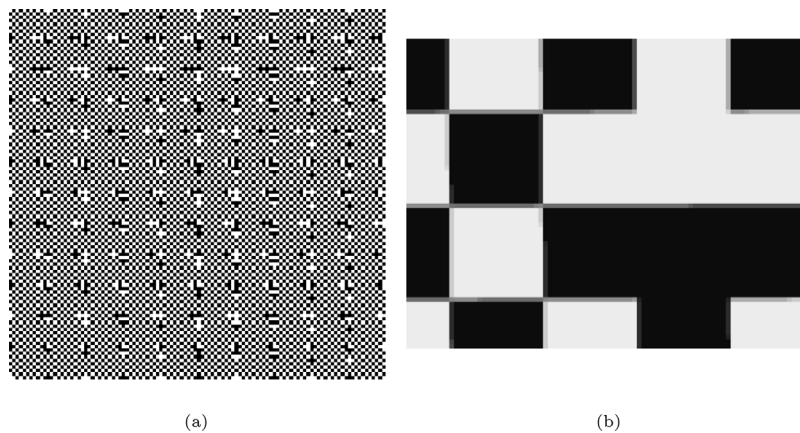
(a)   (b)

Fig. 3. (a) The ChessField pattern. Each 9th row and column includes its ID, providing the coarse position. (b) Gray pixels, located on the edge between light and dark areas.

Table 1
The parameters applied in Section 4.

Parameter	Unit	Description
x	μm	x-coordinate of crossing point of mask and optical axis also output of the SOAP routine
y	μm	y-coordinate as x above
x_{sim}	μm	x-coordinate input for the image generator
y_{sim}	μm	y-coordinate input for the image generator
θ_z	rad	Differential rotation of mask and pixel sensor. For $\theta_z = 0$, the grids of sensor and mask are in line
S	(–)	Image scale, equal to ratio of image distance and object distance
Br	μm	Image blurring constant: sigma (radial) of 2D Gaussian distribution over the pixel area

3.6. Temperature effects

Mask. If the temperature of the mask changes, for instance due to fractional absorption of the light emitted from the back-illuminating led, only the scale of the mask may change. The mask material should be selected in order to meet specifications. Masks can be printed on Zerodur with its thermal expansion coefficient lower than 10^{-9}K^{-1} [3].

Lens, or microscope objective. Due to its transparency, the temperature of a lens will change little. Moreover, this does not cause a variation in the position of the optical centre. There may be a small variation of the focal length, but this does not change the image scale. The image blurring may vary slightly, but this is a negligible third order effect, leaving the output data unaffected.

Image pixel sensor. The silicon chip will expand by $2.6 \times 10^{-6}\,\text{K}^{-1}$ with temperature. State-of-the-art CCDs or CMOS sensor chips dissipate little, while the signal processing circuitry may be positioned elsewhere, including cooling. Thermal expansion causes an error in the measured scale of an image, and can be minimised by gluing the sensor chip, in its centre, at the flat tip of a cooled Zerodur rod. The position of the optical centre of the sensor remains unaffected, and a change of temperature of the sensor has no influence on the x and y data.

Component suspension, arm effects. The position information of mask, lens and sensor is to be transferred to the objects whose alignment is to be monitored. There may be three component holders that may be subject to thermal expansion, and variations appear as drift in the measured x and y coordinates. These variations can be minimised by integrating the Rasnik components carefully in the design of the application.

> **In the following chapters it will be demonstrated that the spatial resolution of Rasnik systems is limited by quantum fluctuations in the applied light.**

4. Image simulations and image analysis

A Rasnik system produces image data files which are sequentially processed by an image analysis program. The output of the analysis program are floats representing x, y and θ_z. A simulation program was developed which generates image data files. The only variation in the image data lies in the pixel content: the fluctuations, set out in Fig. 4, were generated by means of a MonteCarlo (MC) method. The MC generated images were processed by the image analysis program, and the output of the latter was compared to the input parameters of the MC simulation program. In the case of agreement, the performance of both the MC program and the image analysis program is certified. In the case of deviations it may not be clear whether these occur in the image generator or in the analysis program.

The parameters used throughout this section are listed in Table 1.

Image simulations. First, the typical ChessField pattern is generated (Fig. 3(a)). For a given (x, y) as input, a 0 (black) or 1 (white) is returned. In a next step the pixel response is calculated: by sampling an array of points within the pixel area, the fractional pixel content is calculated, relevant in case a contour crosses a pixel area (Fig. 3(b)). If relevant, the inter-pixel dead area can be taken in account by reducing the fiducial pixel area. Then, image blurring is addressed by spreading the pixel content over an array of the neighbouring 11×11 pixels following a 2D Gaussian distribution; this area is large enough to cope with very wide distributions. The width of this distribution is obtained from real image data. The simulated image blurring is shown in Figs. 5(a) and 5(b) for two different Gaussian widths. Finally, fluctuations in the pixel content are applied following the relation $Q_{\text{fluc}} = \sqrt{A + B.Q}$ found in Fig. 4, with $A = 0.26$ and $B = 0.011$. Fig. 6 shows the effect of exaggerated shot noise in the pixel contents.

Image analysis. The SOAP image analysis program for processing the ChessField Rasnik images is set out in detail in [1]. First, the raw image as depicted in Fig. 7(a) is converted into a 2D Fourier image, shown in Fig. 7(b). Here, the phase of each Fourier peak is coded by colour. The positions of the strongest (primary) peaks reveal the frequency and possible rotation of the pattern, while the complex phases indicate a shift. Secondary (harmonic) peaks can be used to refine these parameters, and the nominal positions of the squares is found. Then the intensities near the centres of the squares is evaluated
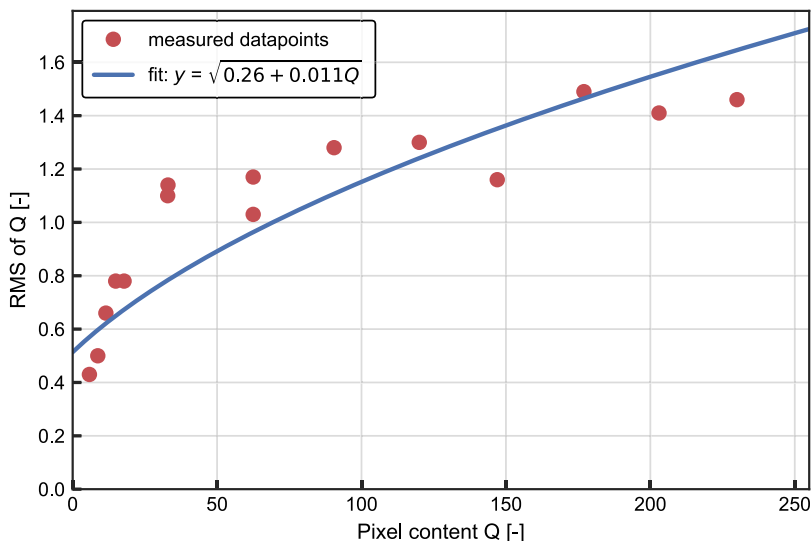
Fig. 4. The measured fluctuations in pixel content versus their average content of some selected pixels with different gray value. For this, the 8-bit setting of the MER-041-302GM image sensor was used, and a set of 256 images was taken. The fluctuations of the pixel content has an electronic component and a component due to quantum fluctuations. The fit parameters are input for the image generator discussed in Section 4.
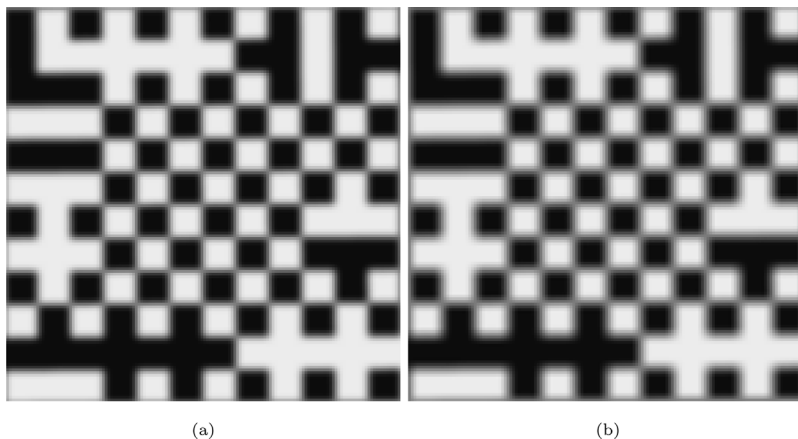

Fig. 5. (a) Image blurring: 2D Gauss width $\sigma = 6.0\,\mu\text{m}$. (b) $\sigma = 12.0\,\mu\text{m}$.
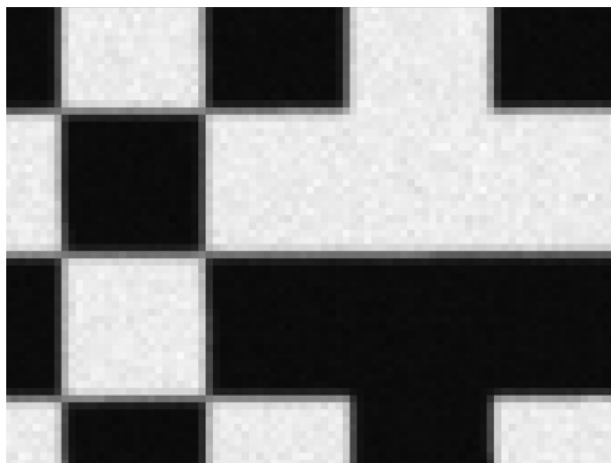
Fig. 6. Exaggerated pixel content noise.

to determine the positions of the code lines, and the values encoded in them.

Verification of the SOAP image analysis program. The performance of the SOAP analysis program has been certified to be better than 1 nm in resolution and 10 nm in linearity [1]. With the image simulation program, pgm-formatted image files are generated: after analysing these by SOAP, the output can be compared to the input parameters of the simulation, revealing systematic and statistical deviations. If significant deviations are found, however, it may not be clear if these originate from the simulation or from the image analysis routine.

In the following, the image data generator uses parameters (i.e. blurring and geometry) of the SolidRas system, described in detail in Section 5.

Spatial resolution. The quality of the SOAP image analysis program, in terms of spatial resolution and linearity, has been verified by the production of a series of SolidRas images, generated at nine (x_{sim}, y_{sim}) positions, with the first (bottom-left) position at $(4000\,\mu\text{m}, 5000\,\mu\text{m})$. The other (x_{sim}, y_{sim}) positions lie on a square 10 by 10 nm grid. The (x, y) output of the SOAP analysis program is shown in Fig. 8 by the blue points. The spatial resolution is of order 1.5 nm, equal for the x and y. The red dots show the effect of reducing the amplitude of the quantum fluctuations by a factor 0.25. The image analysis routine appears to allow sub-nm level precision in spatial resolution.
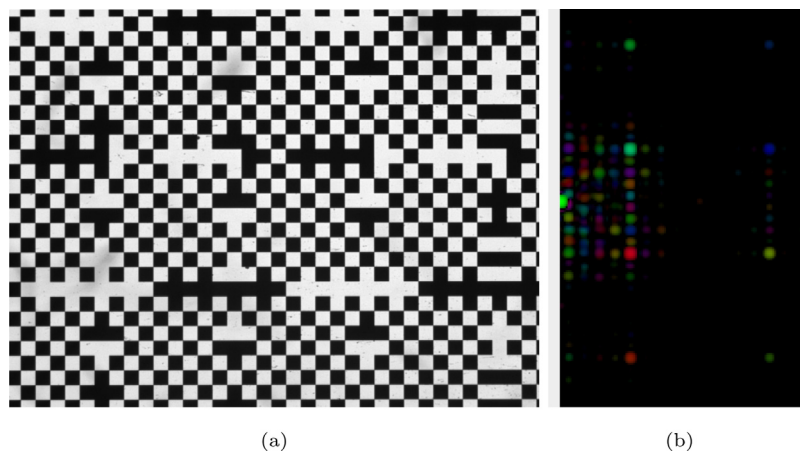
(a)   (b)

Fig. 7. (a) Typical Rasnik image from SolidRas. Here, a ChessField mask of $120\,\mu m$ was applied. (b) Typical SolidRas FFT-converted Rasnik image of Fig. 7(a). The peaks are associated with the periodicity of the image pattern defined by the ChessField side and by the coarse rows and columns at 9 times the ChessField side. All peaks may contribute to the determination of the phase of the contours.
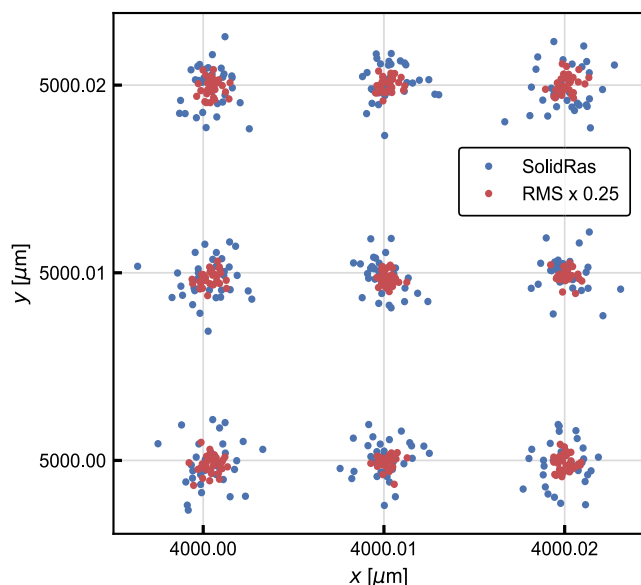

Fig. 8. The (x, y) response of the SOAP image analysis routine for images from a SolidRas MonteCarlo generator. The (x, y) input of the generator was $(4000, 5000)$ for the bottom-left point, and was followed by 8 points with steps of $0.01\,\mu m$ and $0.02\,\mu m$, in x and y, respectively. The non-linearity is noticeable; most likely this is caused by the phase determination of the Fourier peaks. The spread associated with a point is only due to fluctuations in the pixel content. Blue: normal SolidRas parameters; red: pixel quantum fluctuations reduced by factor 0.25.

The linearity of the response. The systematic deviation of the SOAP analysis output (x, y) from the associated generator input values $(x_{\mathrm{sim}}, y_{\mathrm{sim}})$ is clearly visible in Fig. 8. The output for $(x_{\mathrm{sim}}, y_{\mathrm{sim}}) = (4000\,\mu m, 5000\,\mu m)$ is centred around $(x, y) = (4000.001\,\mu m, 4999.999\,\mu m)$ instead of being centred around the input.

The linearity of the SOAP analysis routine is tested by generating images with a step-wise variation in x_{sim}, while keeping y_{sim} fixed, fluctuations in pixel content small, and θ_z set at zero. In Fig. 9 the difference between the generator input x_{sim} and the average analysis output x is plotted against the input x_{sim}. The result of the analysis program deviates up to 35 nm over a range in x_{sim} of 50 nm. This would result in a deviation of the local slope $(\mathrm{d}x/\mathrm{d}x_{\mathrm{sim}})$ of 70 percent or worse. In green, the deviation is shown with $\theta_z = 20\,\mathrm{mrad}$: the linearity error is greatly reduced, and evolved in an offset error of order 3 nm. This suggests that the error is associated with the determination of the (Fourier output) phase. This deviation needs to be studied, and possibly

other algorithms than Fourier conversion must be developed. In blue dots, the linearity error for a range of $2\,\mu m$ and $\theta_z = 2\,\mathrm{mrad}$ is plotted, indicating that the magnitude of the error is limited to 5 nm in practical cases.

Data processing. As readout system for the MER-041 − 302 (Ethernet) image pixel sensor, a Dell OptiPlex 3020 was used with Linux Ubuntu 18.04.4 LTS as operating system. A GPU (GeForce GTX 1050/PCIe/SSE2) card was added to speed up Fourier conversion. With this, a frame rate of 275 Hz was possible, albeit that four pixels in a square were combined into one superpixel. This reduced the time of the raw image data transfer by a factor 4.

5. Results of the SolidRas and MicroRas systems

SolidRas is an assembly of light source, coded mask, bi-convex lens and image pixel sensor, fixed together in a stable and solid unit, see Fig. 10. It provides image data with little variation, and the unit can be suspended as pendulum, reducing the effect of external (seismic, acoustical) vibrations.

The LED (Luxeon StarLed RoyalBlue LXHL-MRRA, $\lambda = 448\,\mathrm{nm}$), and collimating lens (Edmund Optics #63 − 539) form a Köhler light source, providing efficient back-illumination of the coded mask [5]. The ChessField mask has chess field sides of $85\,\mu m$. The projecting lens, identical to the collimating lens, has a focal length of 30 mm. The lens aperture was defined by a diaphragm of 8 mm diameter. As image pixel sensor, the Daheng Mercury MER-041 − 302G was applied with the following specifications: 720×540 pixels, pixel pitch sq. $6.9\,\mu m$, 8-bit deep conversion, max frame rate 302 fps.

The SolidRas unit can be placed in vacuum and suspended as pendulum by two thin wires in order to limit mechanical variations induced by seismic and acoustic vibrations. Data is taken using an internal trigger setting the frame rate. The real-time image analysis yields a data file which includes a time stamp, the x coordinate, the y coordinate, the image scale (∼1), and θ_z.

Welch plots. The performance of Rasnik is analysed by quantifying noise sources. In order to obtain the contribution of various noise sources, a Power Spectral Density (PSD) plot or Amplitude Spectral Density (ASD) plot is desired. For these parameters a Fast Fourier Transform can be executed in order to obtain the ASD. In this paper the Welch's Method [6] is used, in which the data is divided into overlapping windows, calculating the periodogram of each window, followed by taking their average. In this paper Hann's window, a variation of the rectangular window function, is used. The Hann window has an overlap $N_{\mathrm{overlap}} = N_{\mathrm{perwin}}/2$, corresponding to an overlap of 50%. The square root is taken of the output of Welch's PSD estimation to obtain the ASD
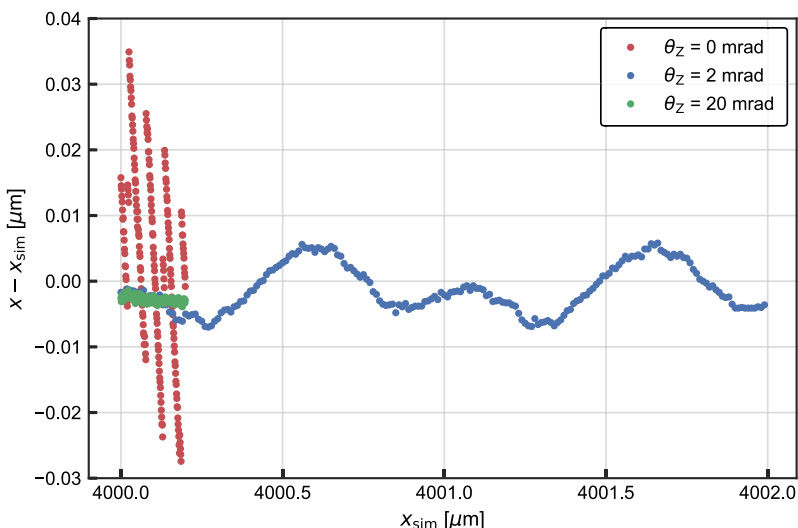
Fig. 9. Linearity error of the combined MonteCarlo image generator and the SOAP image analysis program. The amplitude of the error depends strongly on the differential rotation θ_z of the coded mask and image pixel sensor around the z-axis.
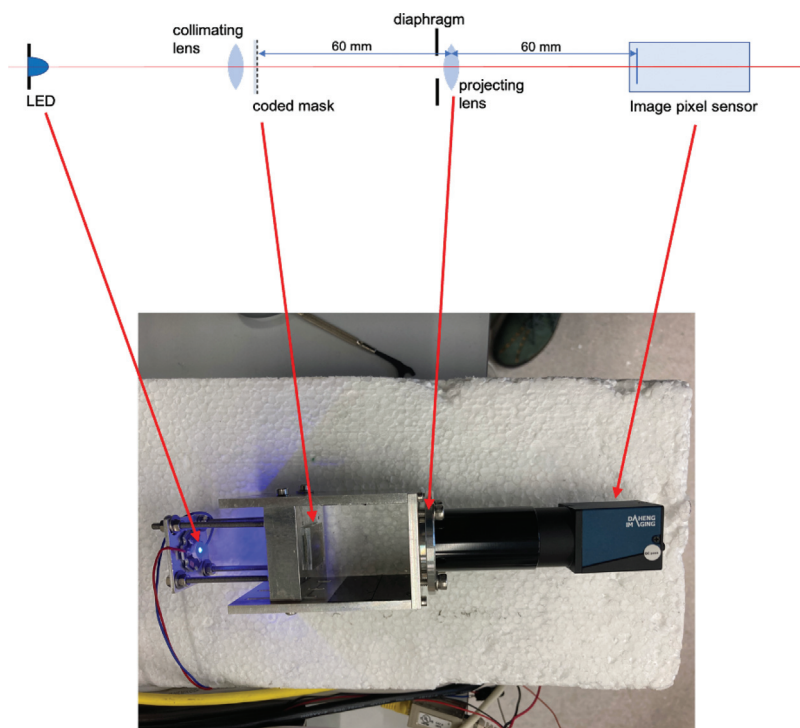

Fig. 10. Layout of the SolidRas system: although quite solid and stiff, the flexibility in the plane of the picture is less than in the normal of the plane due to the U-profile on which the lens and light source are fixed. The eigenfrequencies, appearing in Fig. 15, in the x direction are therefore much higher.

in m/$\sqrt{\text{Hz}}$. The relevant parameters to be considered are: quantity of x and y data, sample frequency (= frame rate) and number of windows. The data sets spanned one hour, and were divided in 5000 windows.

The performance of SolidRas, in open air on a lab table, is shown in Fig. 11. The noise increase at low frequencies can be attributed to variations in the index of refraction of the ambient air. A noise floor of $180\,\text{pm}/\sqrt{\text{Hz}}$ is reached when the intensity of the LED light is adjusted such that the maximum pixel content does not exceed the 8-bit content (255), see the green curve. For the red curve, half the LED current was used, and for the light blue and dark blue curves, again half of the current was applied, respectively. The noise floor level rises a constant factor $\sqrt{2}$ when the light amplitude is reduced by a factor 2. This tallies with the consideration that the data of two images can be superimposed, forming an image with twice the light intensity.

This demonstrates that the spatial resolution of Rasnik is limited by the quantum fluctuations in the pixel contents Q and that other noise sources such as pixel dark current, contribute little.

Next, the influence of the frame rate on the performance of the pendulum-suspended SolidRas is evaluated. The corresponding ASD is shown in Fig. 12: the bottom (red) curve shows the best performance, taken with (practical maximal possible) 275 frames per second (fps). The exposure time was set as high as possible. The current through the LED was adjusted such that the maximum pixel content was just below 255, which is the maximum digital output for the 8-bit setting. The electric power transferred to the led, for that setting, was 300 mW. According to the manufacturer, 20 percent of this power is emitted through photons. Taking into account the optical transmission efficiency, the quantum efficiency of the pixel sensor and the shutter
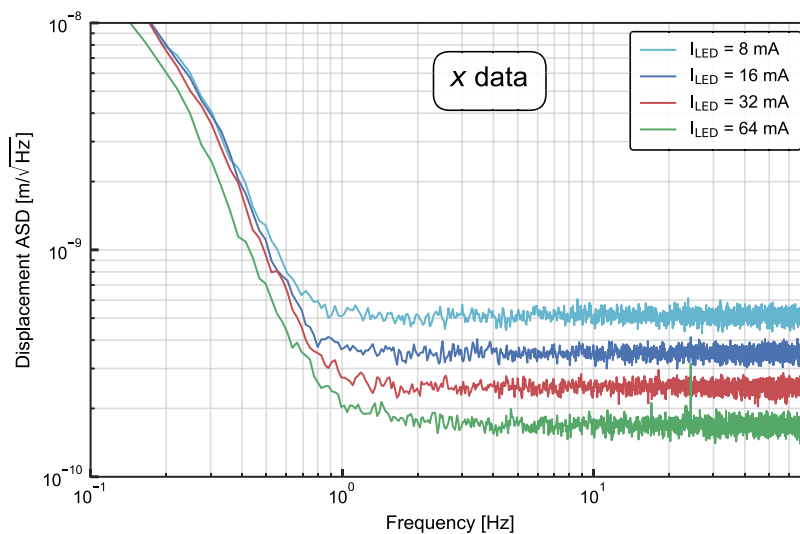
Fig. 11. The noise floor levels for maximum light intensity (bottom curve), and for the LED current reduced by a factor 2, 4 and 8.

time, the number of photons onto a pixel agree with the well depth ($22\mathrm{k}\ e^-$) of this pixel sensor.

The light blue curve shows the result of data taken at a frame rate of 100 Hz while keeping the exposure time unchanged. The spatial resolution *per image* remains unchanged. The noise floor level is increased by a factor of $\sqrt{275/100} = 1.6$ as expected. The light output of the LED was verified to be well-proportional to its current. The dark blue curve, taken at the maximum frame rate of 275 Hz shows the effect of the current, thus light intensity, reduced by a factor 2, as was observed in Fig. 11. This demonstrates again that the fundamental limitation of Rasnik's spatial resolution is due to white noise in the pixel signals, and that the measured x and y have a Gaussian distribution.

The spatial resolution per image has been obtained by the analysis of Monte-Carlo generated images. The settings of the simulation program were adjusted such that the difference between simulated and real images was minimised:

- a pixel content dark region of 25;
- a pixel content light region: 252 (just below the 8-bit maximum of 255);
- a width of Gauss distribution describing image blur: $4.5\,\mu\mathrm{m}$ (see Fig. 2(a) and Section 3.4). This value was obtained after analysing the distributions of pixel row content covering contours (see Fig. 2), and is in good agreement with the expected blurring for a diaphragm of 4 mm in this diffraction-limited case;
- The quantum fluctuations of pixel content Q were taken as $\sigma_Q = \sqrt{0.26 + 0.011Q}$, as found in Fig. 4.

A set of 256 images was generated; for each image, the quantum fluctuation was superimposed onto the individual pixel content. A real and a simulated image is shown in Figs. 13 and 14. After analysis with SOAP, a spatial resolution per image of 3.0 nm, both in x and y, was obtained. For the half-current data, this value was 4.8 nm.

Frame rates above 100 Hz were only possible with reduced data formats: four pixels in a square were combined into one pixel, with 8-bit depth. The simulations result in spatial resolutions per image of 3.3 nm for full LED current, and 5.0 nm for half the LED current, both in x and y. The spatial resolution associated with superpixels is expected to worsen a little since a black–white transition is still covered by 3 to 4 pixels. The noise floor level following from the simulations should equal $3.3\,\mathrm{nm}/\sqrt{275\,\mathrm{Hz}} = 200\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$ in agreement with Fig. 12. This spatial resolution was predicted by Cramer–Rao simulations before [7].

In a next step, the LED was replaced by the UV LightAvenue KE120UYG, reducing the applied wavelength from 448 to 365 nm. In addition, the diameter of the diaphragm was doubled from 4 to 8 mm. The image blurring, and, subsequently, the spatial resolution is expected to be improved by a factor 2, by just the latter change. To test the performance of SolidRas unit with these changes, it was carefully suspended, in a vacuum tank, by means of two thin polyester strings with a length of $\sim 70\,\mathrm{mm}$. The displacement-ASD result is shown in Fig. 15 where a noise floor of $110\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$ is reached. The peaks in the y data in the 20 Hz region can be associated with longitudinal vibrations of the suspension strings. The difference in flexibility of the SolidRas unit in the x and y directions explains the absence of these peaks in the x data: the associated resonance frequencies are much higher and damped more strongly.

The measured noise floor level for the Rasnik output parameter θ_z is $1.0 \times 10^{-6}\,\mathrm{rad}/\sqrt{\mathrm{Hz}}$, in good agreement with simulations.

Performance of the MicroRas system. The layout of the MicroRas system is shown in Figs. 16 and 17. Instead of the projecting lens as in SolidRas, a microscope objective (Newport M-20X) is applied: the image displacement equals 20× the mask displacement, improving the performance. The coded mask ($5\,\mathrm{mm} \times 5\,\mathrm{mm}$) [4] is back-illuminated by means of a glued-on diffuser (Opaline 3 mm) and a LED (Luxeon StarLed RoyalBlue LXHL-MRRA 1W) with $\lambda = 448\,\mathrm{nm}$. The led-mask unit is mounted onto a 3D stage in which x and y can be adjusted manually by means of dial gauges. In addition, the x and z position can be remotely adjusted by means of a piezo-actuator (PiezoSystemJena HPS1000/25 − 15/7): this enables the necessary adjustment of the optimal in-focus position of the mask after applying vacuum, to a maximum of $9\,\mu\mathrm{m}$. The MicroRas system is suspended by two thin 30 mm-long polyester wires attached to the top left and right corners of the base plate.

The performance of the MicroRas system is shown in Fig. 18. Peaks near 3 Hz may be associated with the suspension mechanics. The wide peak at 25 Hz may be associated with eigenfrequencies of the unit itself. The sharp peaks at 25, 50, 75 and 100 Hz may be associated with grid power electro-magnetic signal interference, or with grid power induced Eddy currents deforming the MicroRas base plate. A noise floor level of $7\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$ is reached.

6. Improvements

Led wavelength. The spatial resolution of Rasnik is, to first order, proportional to the wavelength of the applied (non-polarised, non-coherent) LED light. In addition, the optimal ChessField size will be smaller for smaller wavelength, increasing the total contour length in images. The spatial resolution of Rasnik is therefore proportional to $\lambda^{(1/1.5)}$. Since powerful UV leds with smaller wavelength are now
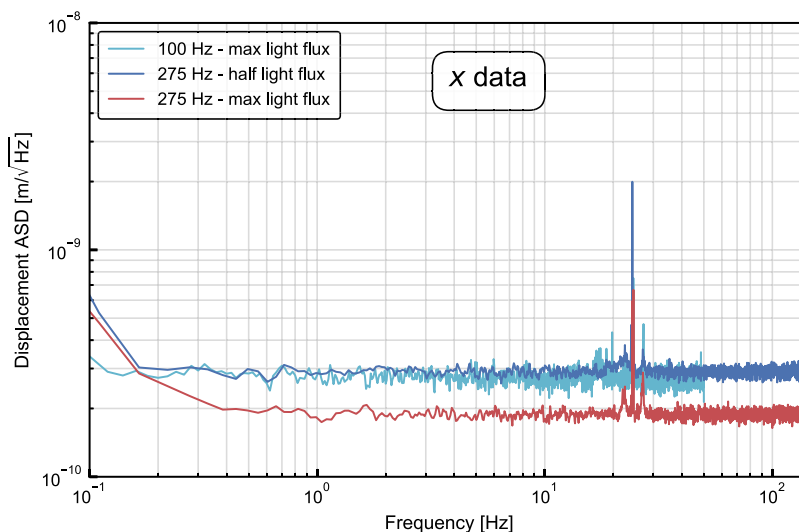
(a)
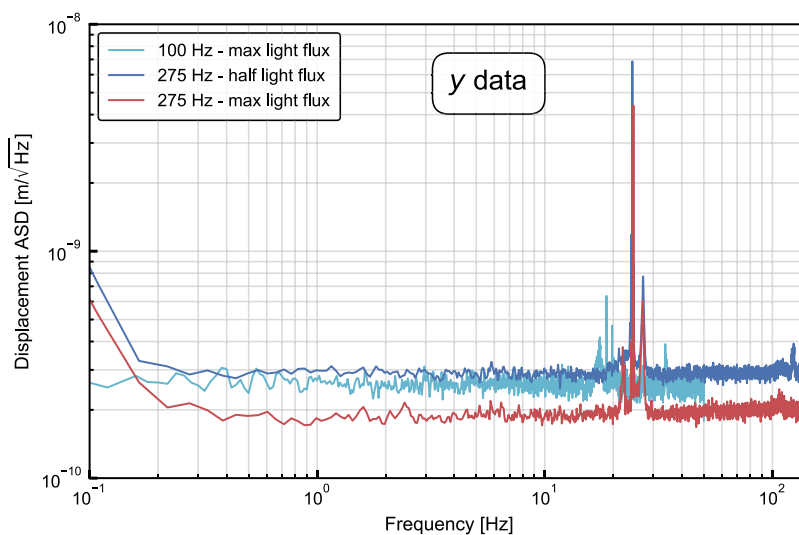
(b)

Fig. 12. Welch plot of a SolidRas data set, taken during a period of 30 min. Red: best performance at maximum practical frame rate (275 Hz). Light blue: frame rate reduced to 100 Hz, maintaining exposure time. Dark blue: frame rate 275 Hz, LED current reduced by factor 2. (a) x data, and (b) y data. Colour figures online.

available, the shortest possible wavelength should be selected. Limits are the transmission of the lens or objective, and the quantum efficiency of the image pixel sensor. The transmission of *reflective* objectives allows a wavelength down to 200 nm, and image pixel sensors (CCDs) are available with sufficient quantum efficiency for this wavelength. The improvement in spatial resolution is expected to be a factor 3 when changing the wavelength from 448 to 200 nm.

Magnification of objective. Objectives with magnification 100× are available. The correct and stable distance between coded mask and the objective is essential and may cause practical problems. If UV light is applied, reflection objectives should be used (magnification up to 40×). The highest Numerical Aperture (NA) should be chosen in order to minimise image blurring. Changing only the magnification from 20× to 40× should improve the spatial resolution by a factor 2.

Diaphragm, aperture. For diffraction-limited images projected by a convex singlet, the image blurring Br = 1.22 λ/NA (μm). For a projecting lens, NA = D/b where b equals the image distance and D the diameter of the lens (or diaphragm). The lens diameter or diaphragm should chosen as large as possible. For very large aperture, spherical

aberration may cause image deformation, but this has no direct effect on the spatial resolution (see Section 3.3).

The image pixel sensor. The spatial resolution of a Rasnik system is inversely proportional to the square root of the surface of the sensor: sensors with a large active surface are advantageous. A multiple of sensors is feasible, provided that the image quality is maintained. Since the *information per unit surface* matters, the pixel pitch is not critical and has a wide optimum. A black–white transition should be covered by minimal 3 pixels: a rule of thumb is that the pixel size should be smaller than 2× the image blur Br. Small pixels however, with large image data files, require much CPU power.

The spatial resolution of Rasnik systems is determined by the quantum fluctuations of the pixel's content. Increasing the light intensity would improve the spatial resolution, but if the maximum of photoelectrons in a pixel well is exceeded, the charge will leak away, and the pixel charge content is no longer proportional to the amount of light on that pixel. Increasing the frame rate shortens the exposure time: with the MER-031−860U3M pixel sensor, up to 860 fps can be recorded, albeit that substantial computing power is required for the image analysis.
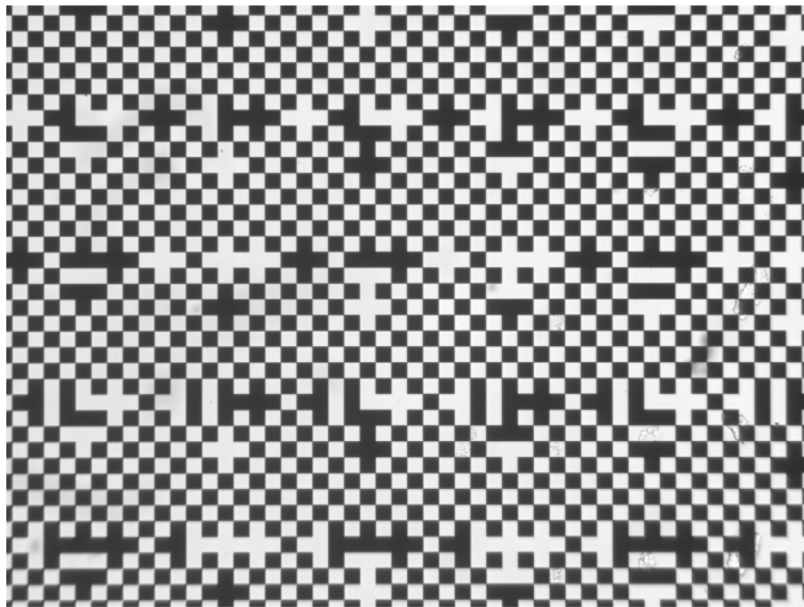
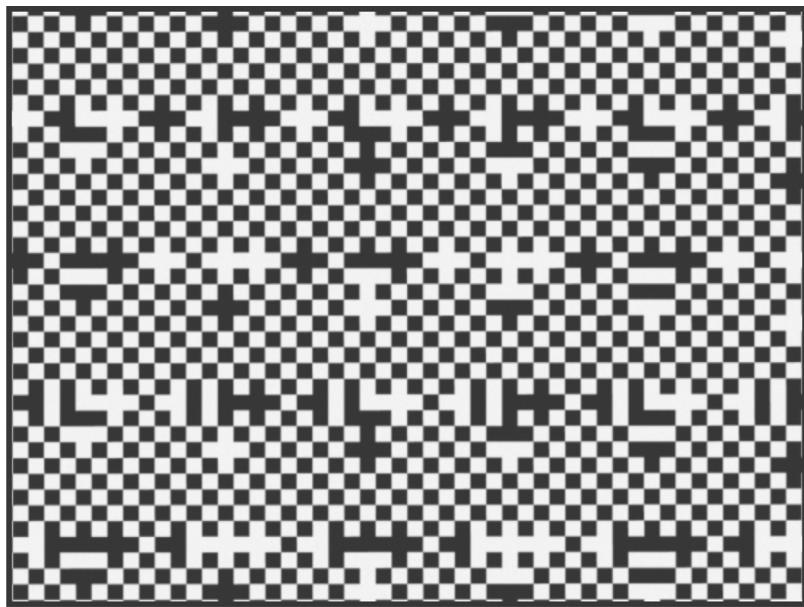**Fig. 13.** Real SolidRas image, displaying some stains and dust particles.

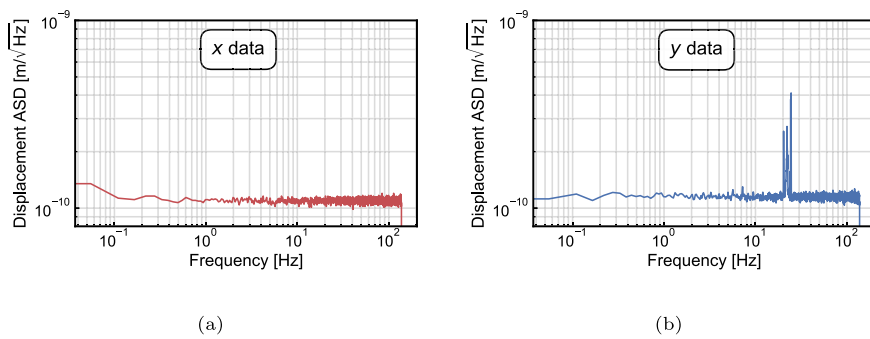**Fig. 14.** The image of Fig. 13 simulated by the Monte Carlo (MC) generator.

(a)

(b)

Fig. 15. Welch plots of continuous SolidRas x and y data. Diaphragm $= 8\,\mathrm{mm}$, $\lambda = 365\,\mathrm{nm}$.
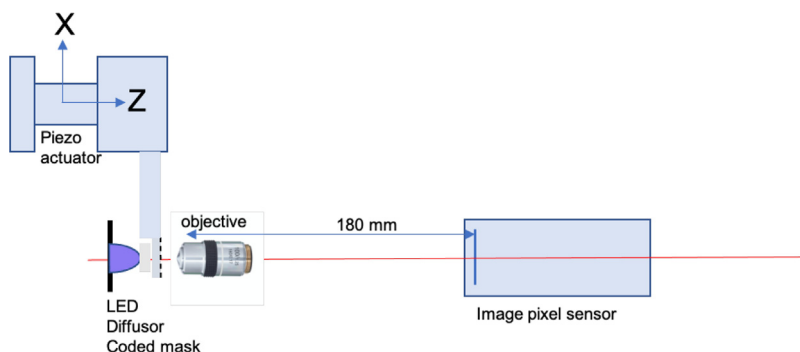
Fig. 16. The layout of the MicroRas system. The mask can be displaced in *x*, *y* and *z* by manual dial gauges and, in addition, in *x* and *z* by means of piezo-actuators.
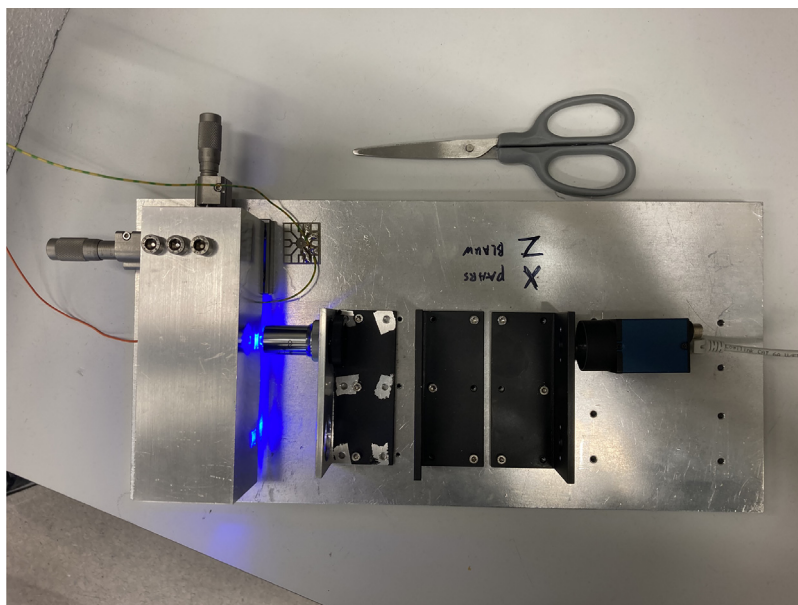

Fig. 17. The MicroRas unit: the MER pixel sensor, microscope objective and mask suspension are bolted firmly onto a common base plate.

The mere increase of frame rate from 275 Hz to 860 Hz should improve the spatial resolution per $\sqrt{\text{Hz}}$ by a factor 1.7. Assuming identical well depth, the shorter exposure time will require an inversely higher LED light flux.

High full-well capacity (FWC). The most important quality factor of state-of-the-art image pixel sensors has been the quantum efficiency which is now approaching the maximum value of 1. There is now an increasing interest in the development of efficient image pixel sensors that, in addition, can handle high light intensities. In [8], a new sensor is described where a capacitor-per-pixel is applied, increasing the FWC. A further improvement lies in the recording of the pixel current instead of the pixel charge.

Data processing. State-of-the-art image pixel sensors communicate by means of USB3. This enables direct access and processing of the image data by any suitable processor. GPUs can (Fourier) process images extremely fast, and multiple GPUs can be controlled by one master processor. Given commercial interest in machine viewing/learning and in automotive (self-driving), further integration of image sensor and image processing can be expected.

With MicroRas, a spatial resolution was reached of $7 \, \text{pm}/\sqrt{\text{Hz}}$. By replacing the Royal Blue LED by a UV led, the Newport M 20× microscope objective by the ThorLabs LMM-40X-UVV-160 reflective objective, and the MER-041 − 302G pixel sensor by the MER-031 − 860U3M, the spatial resolution could improve by a factor 7, bringing the performance of

Rasnik to the level of $1 \, \text{pm}/\sqrt{\text{Hz}}$. With a newly designed ASIC CMOS pixel sensor, a value of $0.2 \, \text{pm}/\sqrt{\text{Hz}}$ should be reachable.

7. Conclusions

Two Rasnik systems were operated while suspended as pendula in vacuum, well-isolated from seismic and acoustical vibrations. The position of the three typical components, namely mask, lens and sensor, were firmly fixed onto a common support. The resulting images were therefore of constant alignment, while the pixel noise is the main variation in the data, revealing the spatial resolution of the system.

The best performance of the SolidRas system is $110 \, \text{pm}/\sqrt{\text{Hz}}$, both in linear displacements *x* and *y*, in good agreement with simulations. For the resolution of rotation θ_z, an RMS value of $1.0 \times 10^{-6} \, \text{rad}/\sqrt{\text{Hz}}$ was measured. The precision of measuring scale factor S was below expectation and needs further analysis.

With MicroRas, the image displacement is multiplied with the magnification of the microscope objective (20×), replacing the lens. A spatial resolution of $7 \, \text{pm}/\sqrt{\text{Hz}}$, both for *x* and *y*, was reached, again in good agreement with simulations. Given the diffraction-limited images, the performance of Rasnik is limited by the quantum fluctuations of the light, absorbed by the pixels.

The demonstrated performance of MicroRas of $7 \, \text{pm}/\sqrt{\text{Hz}}$ can be improved by using a UV led, by replacing the microscope objective
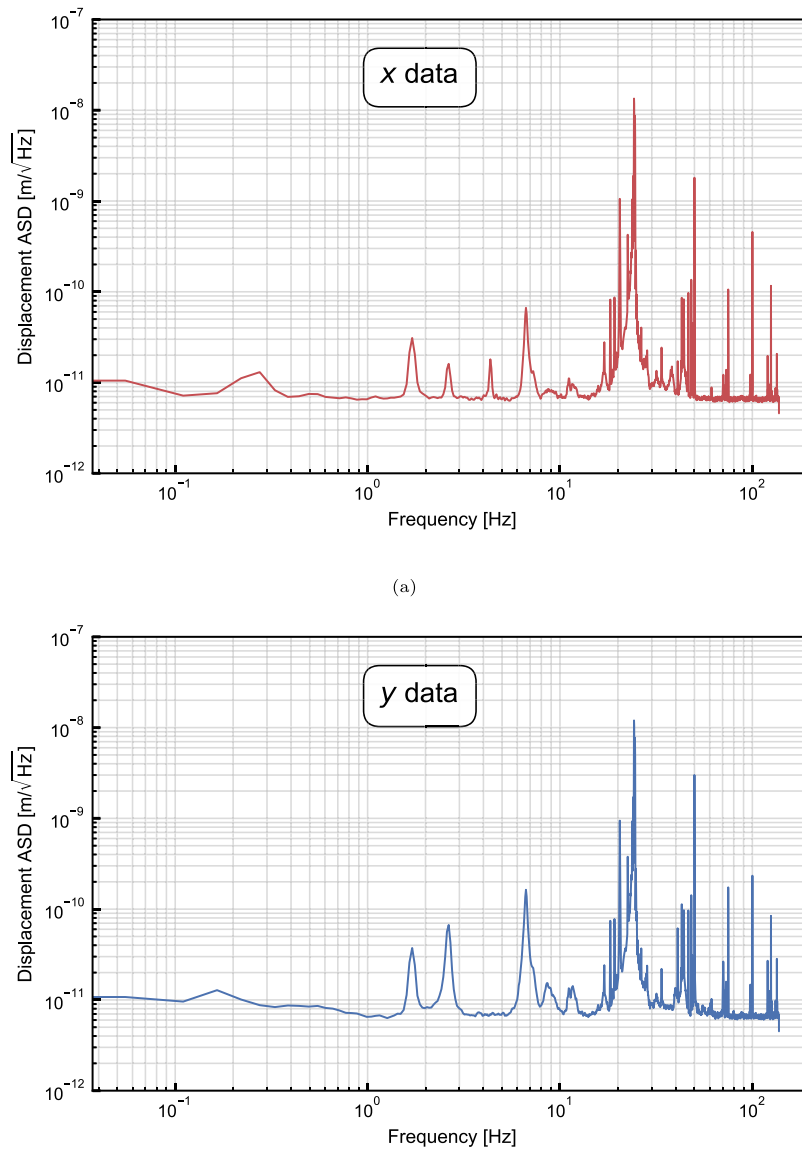
(a)

Fig. 18. Welch plots of continuous x and y data of the MicroRas system. The frame rate is 275 Hz.

by a reflective objective, and using a faster version of the image pixel sensor: a value of $1\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$ should be within reach, with off-the-shelf components. By developing a special image pixel sensor capable to handle a higher light flux, the spatial resolution could be further improved by a substantial factor.

The following features make the Rasnik systems compatible or advantageous with respect to other displacement sensors:

- Rasnik measures a 2D displacement in the plane of a planar object (mask) with respect to the optical axis, perpendicular to the mask plane. The optical axis is defined by the fixed-together lens and optical image sensor unit
- systematical linearity error 50 pm or smaller over arbitrarily large dynamic range, in x and y
- only one precision element: the coded mask, available at low cost thanks to MEMS and IC industry
- no cross coupling between output parameters x, y, S and θ_z; perpendicularity between x and y only defined by the mask
- no calibration required
- straight, analog/digital linear measurement system: no 'lock' control and feedback systems required

- arbitrarily large range of measurement (determined by mask size)
- extremely low Flicker noise
- no special electronics required: a system can be composed of only commercially available CMOS pixel image sensors, CPUs, GPUs and USB or Ethernet3 networking, and COTS optical components
- (very) low cost: 300 - 3000 € excl. network and CPUs
- light pressure onto the mask acts perpendicular to the directions of interest
- system can operate in air, or in vacuum, and in cryogenic environments if case-specific precautions are taken.

The proven spatial resolution in x and y of $7\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$ is not as good as can be obtained with interferometers ($4\,\mathrm{fm}/\sqrt{\mathrm{Hz}}$, see [9]) and SQUIDs (0.1 to $3\,\mathrm{fm}/\sqrt{\mathrm{Hz}}$, see [10]). Another disadvantage of Rasnik is that it requires substantial CPU power, albeit that this is the focus of future (machine viewing, LADAR, self-driving cars) developments.

> **Because of its low cost and simplicity, Rasnik could be well-applied in a large area network of multiple seismic sensors, and in deformation-sensing instrumentation in general.**
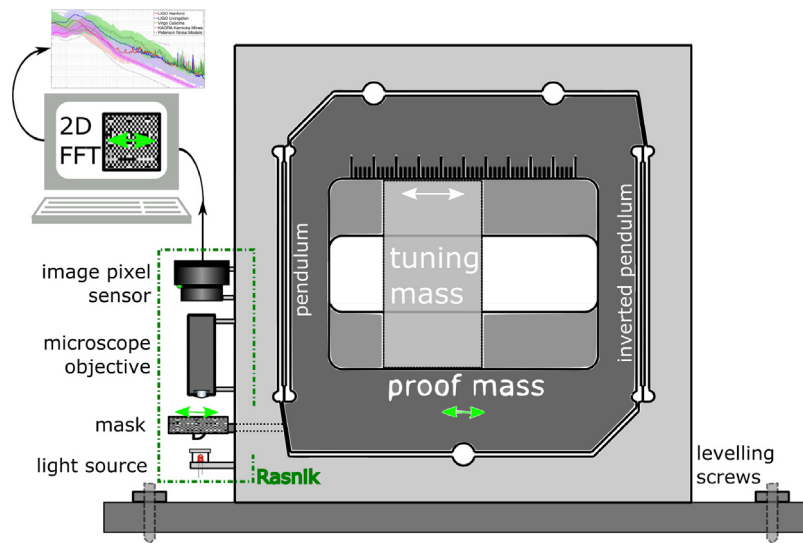
Fig. 19. Accelerometer with a Rasnik readout:the position of the proof mass with respect to the base frame is probed by a Rasnik system.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Appendix. Applications

Seismic sensor: accelerometer. Rasnik could be applied in instruments for registering seismic motions in general [11]. A practical application is shown in Fig. 19. The proof mass is suspended in a Watt's linkage configuration which allows for low frequency tuning and thus increased mechanical sensitivity. The position of the proof mass with respect to the ground-attached frame is recorded by one or more Rasnik systems: each system measures two degrees of displacement freedom. With one Rasnik system, both the horizontal (x and y) acceleration is measured, in absence of cross-coupling between the two. With three Rasnik systems, a 6D inertial sensor can be realised. The Rasnik systems are operational as soon as the proof mass is released, and no feedback systems are required. With the high-precision mask as unique reference, calibration is not needed, omitting calibration hardware. Given the large (mm) range of operation, the systems will continue to operate after high-amplitude events like earthquakes or storms.

The Moon as antenna for gravitational waves. In Ref. [12], a network of seismic sensors is proposed, using the Moon as object of which seismic waves, deformations and vibrations are to be recorded. The relevant information is embedded in the phase *difference* in the seismic signals from multiple sensors distributed over a certain area. By placing Rasnik systems on the Moon's surface spanning a large distance (10–50 km) between the light source and image sensor, the relevant differential displacements can be recorded [13]. In absence of feedback systems, the Lunar Rasnik system would be operational immediately after the installation of the three components including the power- and data systems.

References

[1] M. Beker, et al., The Rasnik 3-point optical alignment system, J. Instrum. 14 (08) (2019) P08010, http://dx.doi.org/10.1088/1748-0221/14/08/P08010.

[2] Patent granted in the Netherlands under number NL1034328, March 5, 2009. Rasnik: Werkwijze en inrichting voor het meten van geringe vervormingen of verplaatsingen van objecten.

[3] M.J. Veins, Fracture toughness and crack growth of Zerodur, NASA Tech. Memo. 14 (4185) (1990).

[4] Photronics. 1 Technology Dr, Bridgend CF31 3LU, United Kingdom. For making this e-beam written mask, special care was taken to avoid aliasing due to round-off errors in the chess field contour positions.

[5] August Köhler, Ein neues beleuchtungsverfahren für mikrophotographische zwecke, Z. Wiss. Mikrosk. Mikrosk. Tech. 10 (4) (1893) P433–P440.

[6] P.D. Welch, The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, IEEE Trans. Audio Electroacoust. (1967).

[7] J.V. van Heijningen, Precision Improvement in Optical Alignment Systems for Linear Colliders (M.Sc. Thesis), Nikhef, Amsterdam and Delft University of Technology, 2012.

[8] Woo-Tae Kim, et al., A high FullWell capacity CMOS image sensor for space applications, Sensors 19 (2019) 1505, http://dx.doi.org/10.3390/s19071505, https://www.mdpi.com/1424-8220/19/7/1505.

[9] J.V. van Heijningen, et al., A novel interferometrically read out inertial sensor for future gravitational wave detectors, IEEE SAS Proc. (2018) 76–80.

[10] Yuki Sato, Richard Packard, DC superconducting quantum interference device based neodymium magnet displacement sensor for superfluid experiments, Rev. Sci. Instrum. 80 (2009) 055102, http://dx.doi.org/10.1063/1.3129942.

[11] Alessandro Bertolini, et al., Mechanical design of a single-axis monolithic accelerometer for advanced seismic attenuation systems, Nucl. Instrum. Meth. (ISSN: 0168-9002) A556 (2) (2006) 616–623, http://dx.doi.org/10.1016/j.nima.2005.10.117, https://www.sciencedirect.com/science/article/pii/S0168900205020930.

[12] Jan Harms, et al., Lunar gravitational-wave antenna, Astrophys. J. 910 (1) (2021) 1, http://dx.doi.org/10.3847/1538-4357/abe5a7.

[13] A. Seryi, Investigation of slow motions of the SLAC linac tunnel, SLAC-PUB-8597 8597 (1) (2000) P06034, arXiv:physics/0008195.