The Rasnik 3-point optical alignment system

Published by IOP Publishing for Sissa Medialab

Received: *March 1, 2019*
Revised: *May 24, 2019*
Accepted: *July 12, 2019*
Published: *August 8, 2019*

The Rasnik 3-point optical alignment system

M. Beker,[a,1] **G. Bobbink,**[a] **B. Bouwens,**[b] **N. Deelen,**[a,2] **P. Duinker,**[a,3] **J. van Eldik,**[a,2] **N. de Gaay Fortman,**[a,c] **R. van der Geer,**[a,4] **H. van der Graaf,**[a,c,3,5] **H. Groenstege,**[a,3] **R. Hart,**[a] **K. Hashemi,**[a,6] **J. van Heijningen,**[a,7] **M. Kea,**[a,8] **J. Koopstra,**[a] **X. Leijtens,**[a,9] **F. Linde,**[a] **J.A. Paradiso,**[d] **H. Tolsma**[a,8] **and M. Woudstra**[a,10]

[a] *Nikhef,*
Science Park 105, 1089 XG Amsterdam, The Netherlands
[b] *Amsterdam Scientific Instruments,*
Science Park 106, 1089 XG Amsterdam, The Netherlands
[c] *Delft University of Technology,*
Lorentzweg 1, 2628 CJ Delft, The Netherlands
[d] *Massachusetts Institute of Technology,*
77 Massachusetts Ave, Cambridge, MA 02139, U.S.A.

E-mail: vdgraaf@nikhef.nl

Abstract: The Rasnik alignment system was developed initially in 1983 for the monitoring of the alignment of the muon chambers of the L_3 Muon Spectrometer at CERN. Since then, the development has continued as new opto-electronic components become available. Rasnik systems are 3-point optical displacement monitors and their precision ranges from below nanometers to several micrometers, depending on the design and requirements of the systems. A result, expressed in the range/precision ratio of 2×10^6, is presented. According to the calculations of the Cramér-Rao limit, and by means of MonteCarlo simulations, a typical Rasnik image should have enough information to reach deep sub-nanometer precision. This paper is an overview of the technological developments and achievements since Rasnik was applied in high energy physics experiments.

Keywords: Detector alignment and calibration methods (lasers, sources, particle-beams); Overall mechanics design (support structures and materials, vibration analysis etc)

[1] Present address: Innoseis B.V., Science Park 105, 1089 XG Amsterdam, The Netherlands.
[2] Present address: CERN, Route de Meyrin, 1211, Geneva, Switzerland.
[3] Retired.
[4] Present address: GoDataDriven, Wibautstraat 202, 1091 GS Amsterdam, The Netherlands.
[5] Corresponding author.
[6] Present address: Brandeis University Boston, 415 South St, Waltham, MA 02453, U.S.A.
[7] Present address: University of Western Australia, 35 Stirling Hwy, Crawley WA 6009, Australia.
[8] Present address: ASML Netherlands B.V., De Run 6501, 5504 DR Veldhoven, The Netherlands.
[9] Present address: Eindhoven University of Technology, De Rondom 70, 5612 AP Eindhoven, The Netherlands.
[10] Present address: TMC, High Tech Campus 96, 5656 AG Eindhoven, The Netherlands.

https://doi.org/10.1088/1748-0221/14/08/P08010

2019 JINST 14 P08010

Contents

2019 JINST 14 P08010

1 Introduction

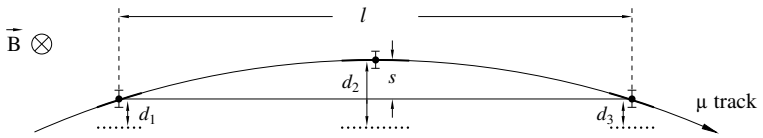


Figure 1. Principle of measuring the momentum of an energetic charged muon particle, moving in a magnetic field. Three muon track positions are measured with respect to the wire positions of three muon chambers (d_1, d_2 and d_3). The sagitta s of the muon track is defined as the deviation, taken at the position of the central detector from a straight (infinite momentum) track. The recorded distances depend on the alignment of the wire planes of the three chambers. The deviation of perfect alignment of the wire planes can also be expressed in terms of a sagitta, which can be applied as off-line correction to the track position data. Then, the chamber positions don't matter any longer. In the L3 experiment, the distance l between the inner and outer chambers was 2.5 m. It should be noted that a translation or a rotation of a detector unit, consisting of a complete set of three chambers, mutually fixed, does not influence the alignment of the wire planes.

High momentum, minimum ionizing charged particles, traversing a gaseous detector placed in a magnetic field, will be curved due to the Lorentz force. At these high momenta the radius-of-curvature R can be determined by measuring the sagitta s, which is the distance between the central track point and the line through the outer measured tracking positions with length l as shown in figure 1 for the L3 muon spectrometer at LEP [1]. For high momenta, R is proportional to the momentum of the particle and inversely proportional to s and can be approximated for small l with respect to R by:

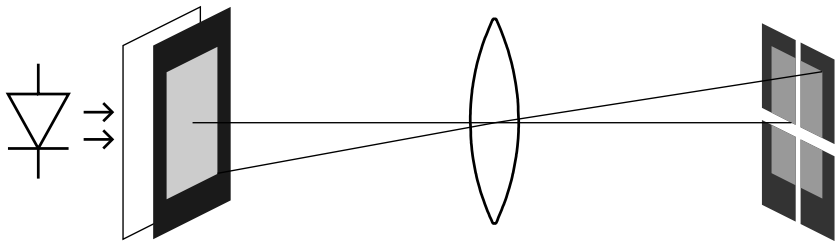
$$R \approx l^2/2s \tag{1.1}$$

The relative error in the measured R and thus momentum is therefore the same as in the measured sagitta. The uncertainty in the momentum measurement of the particle is directly related to the intrinsic precision of the position-sensitive detectors, multiple scattering, and the precision of the relative location of these detectors. Intrinsic position measurement accuracies of order 50 μm are realized with Multi Wire Proportional Chambers (MWPCs) and order 10 μm with silicon strip detectors. The error in the alignment, expressed in terms of sagitta, should be in the order of the spatial resolution of the track detectors, or smaller. Comparable precision and stability of mechanical alignment over distances larger than 1 m is notoriously difficult. Until 1983, accurate alignment could only be achieved by applying a stretched wire as reference, and by a Taylor Hobson telescope [2], but these instruments could not be used as monitors, automatically following actual variations in alignment.

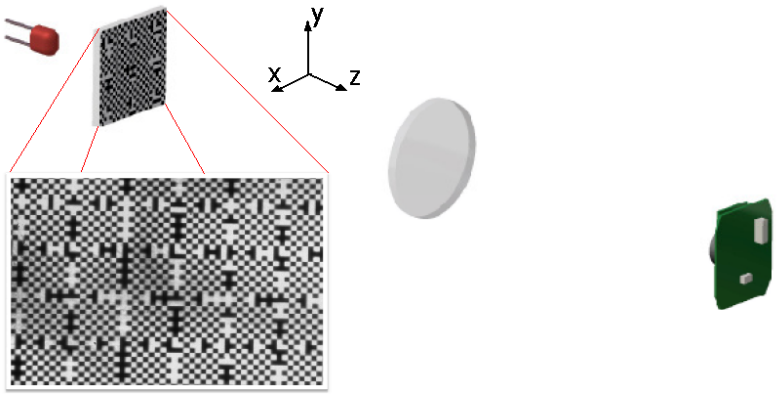
Circa 1983, an opto-electronic alignment monitor system, earlier proposed in NASA publication and subsequent patent [3], was of interest at Draper Laboratories for monitoring the actual geometry of the Space Shuttle's Robotic Arm [4]. The concept of this system, consisting essentially of a light source, a lens, and a four-quadrant light sensor, was prototyped and tested for structural monitoring by Draper [5], and evaluated at Nikhef, where a readout system was developed.

Applying (the then most powerful) red LEDs as light source, a few hundred of these *Red Alignment System NIKhef* (Rasnik) systems were applied in the L3 experiment [1], not only for monitoring the alignment of the tracking detectors, but, in even larger numbers, for the monitoring of the geometry (notably sag) of the muon tracking chambers.

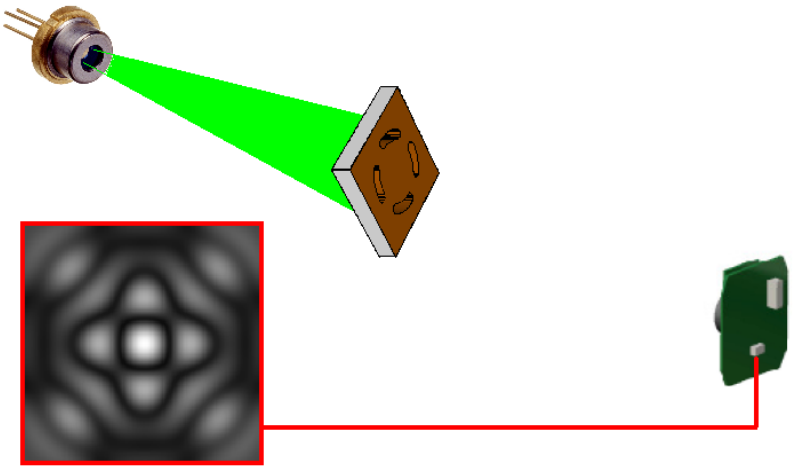
2019 JINST 14 P08010



(a)



(b)



(c)

Figure 2. (a). The 4QD Rasnik schematically. The span of the systems varied between 2 and 5 m. (b). The CCD Rasnik: a LED back illuminates a coded mask and this object is projected onto a pixel sensor using a lens. If any of these three elements move in four degrees of freedom (X, Y, Z and *roll*) the image shifts or rotates and this can be determined using image processing software. The roll (rotZ) of the lens is not determined. The span of the systems have varied between 50 mm and 15 m. (c). The RasDif system: a spherical and monochromatic light source illuminates a zone plate: a diffraction pattern appears on to a pixel sensor. If any of these three elements move transversally (in X and Y), the image shifts and this can be determined using image processing software. The span of the systems were 70, 100 and 140 m.

2019 JINST 14 P08010

In 1990, in concept studies for SSC and LHC experiments, the need for low-cost alignment monitor systems with a large range of operation, in both transversal directions, became apparent. In ATLAS, a 1 TeV muon has a sagitta of about 0.5 mm, so for a momentum resolution of 10 % the sagitta needs to have an uncertainty below 50 μm, in the order of the spatial resolution of a muon detector. At high muon momentum the uncertainty in the sagitta depends mostly on the spatial resolution of the detector layers and the alignment of the detector layers. One can think of measuring the alignment of the layer by switching off the magnetic field, and collecting data from cosmic muons measured by the detectors. Magnetic forces, temperature variations and in-situ interventions on the detector, however, cause displacements. Hence the need for a system that can continuously monitor the alignment.

By replacing the four-quadrant diode by the then available CCD, the precision was greatly improved. The range of the new system is determined by the arbitrary size of the back-illuminated coded mask. Eventually, six thousand Rasnik systems were applied in the ATLAS experiment [6], and in other (also non-HEP) systems.

By fixing three components of a typical Rasnik system to three independently moving parts, the relative displacement of these parts can be monitored to great precision. The Rasnik family has three members, each of them consisting of a (patterned) light source, some projection element e.g. a lens, and a type of pixelised light sensor. In this paper, the 4QD, CCD-Rasnik and RasDif systems, depicted in figure 2a, figure 2b and figure 2c are sequentially addressed.

2 4QD Rasnik

This first version of the Rasnik system was developed for the L_3 experiment [1]. Each octant of the L_3 muon spectrometer had two double vertical Rasnik systems to monitor the alignment between the layers. The system consists of a light source (a red LED with a square mask), a positive lens, and a four quadrant light sensitive photodiode sensor (4QD) as shown in figure 2a. The LED provides a homogeneous square light source, which is projected by the lens onto the 4QD. The position of the square LED image is calculated from the ratios of the measured light intensity on the four quadrants.

The light source for 4QD systems. This light source should be bright, homogeneous, and have a well-defined geometry. In 1980, the largest signals from photodiodes were achieved with LEDs giving red or infra-red light: this is the result of the spectral sensitivity of the photodiode and the power and wavelength of the then available LEDs. The applied large-area LEDs had an opaque layer, and the geometry of a single LED light source was well defined by masking the opaque layer with a transparent square of 1.75 mm x 1.75 mm in the form of a photographic negative. For an upgrade of the L_3 Muon Spectrometer, light sources with a larger area have been realised by combining a light guide and a specially shaped diffuser [7].

The lenses for 4QD systems. See details about projecting lenses for the 4QD Rasnik systems in 3.3.

Image sensor for 4QD systems, acquisition system. A 4×1250 μm 4QD is used for the first Rasnik systems for the L3 experiment: type EG&G UV140BQ-4. For an upgrade, larger custom specific 4QD's were supplied by EG&G. An analog multiplexing read-out system was designed to

2019 JINST 14 P08010

readout the light levels on all quadrant sequentially, eliminating drift in the measured ratio's of light on quadrants within a diode. In order to reduce the noise due to variable background light, notably fluorescent lighting, the LED was modulated with a 4 kHz sinus, and the 'clean' light value from a quadrant was obtained using phase detection in the input stage after the analog multiplexer.

The data acquisition system of the L_3 muon spectrometer was build in CAMAC units and was part of the Slow Control monitor system of the experiment.

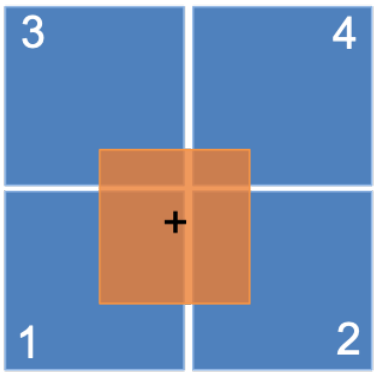


Figure 3. The image projected onto the four-quadrant diode. The digital data is represented by Q_1, Q_2, Q_3 and Q_4, respectively.

Image position analysis. The image, projected onto the four-quadrant diode is shown in figure 3. The relation between the X position of the image on the sensor and the signal ratios of the quadrant is

$$\begin{aligned} Q_{\text{left}} &= Q_1 + Q_3 \text{ and } \mathrm{Q_{right}} = \mathrm{Q_2} + \mathrm{Q_4} \\ X &= \frac{A(Q_{\text{right}} - Q_{\text{left}})}{2(Q_{\text{right}} + Q_{\text{left}})}, \end{aligned} \tag{2.1}$$

where A equals the side of the image square. For the Y coordinate, a similar algorithm is applied. The dynamic range of the systems could be maximised by having A as large as the four quadrants together; this requires an adapted algorithm.

The geometry of the light source determines the value for A, and can be well determined by a photographic mask. The non-homogeneity of the light source affects the linearity of a system, and corrections were applied for the non-sensitive areas between the quadrants. Imaging errors, mainly blurring due to diffraction and out-of-focus, cause systematic errors which can be recorded and corrected for by means of a 2D scan of an operational system. From the data of the scans of all systems, an empirical algorithm was derived, resulting in a maximum systematical deviation, for all systems, of 5 μm. Using this correction table, the typical root-mean-square error was found to be 5 μm for each measurement at a measurement frequency of 1 Hz. The 2D spatial resolution, per measurement, is only affected, in principle, by the noise in the four figures for quadrant light yield, and can be determined easily. This electronic noise depends usually on the integration time constant of filter stages and is closely related to the number of readouts per unit of time.

Since the light intensity on all quadrant is measured with a single ADC channel, the drift (defined as a change in the output data in absence of a change in alignment) of a 4QD Rasnik system

2019 JINST 14 P08010

is extremely low. If measurements are averaged, the resolution of mean values increases with the square root of the number of measurements included. The readout frequency of a Rasnik system matters: the performance in terms of spatial resolution can be expressed in an amplitude spectral density (ASD) with units $\mu m/\sqrt{Hz}$ which decreases towards zero Hz, in the ideal case. The 1/f noise of Rasnik systems is caused by physical changes in the (semiconductor) detecting elements, and by changes in the geometry of components due to gradients in temperature. This will limit eventually the spatial resolution at very low frequencies of variations to be measured.

4QD results of the L_3 muon spectrometer. The 5 m long wires in the muon chambers of L3 were supported, at both ends, by glass plates. In order to limit the sag of the wires, a glass support plate was also placed halfway the wires. Two or three 4QD Rasnik systems were attached to the glass plates, providing redundant information of the alignment. The position of the central plate could be adjusted for the best possible alignment, and data is shown in figure 4. The noise in the systems is dominated by non-straight propagation of light in the ambient air: see (5.1). During the 10 year lifetime of the L3 experiment, the absence of drift in the 4QD systems was established.

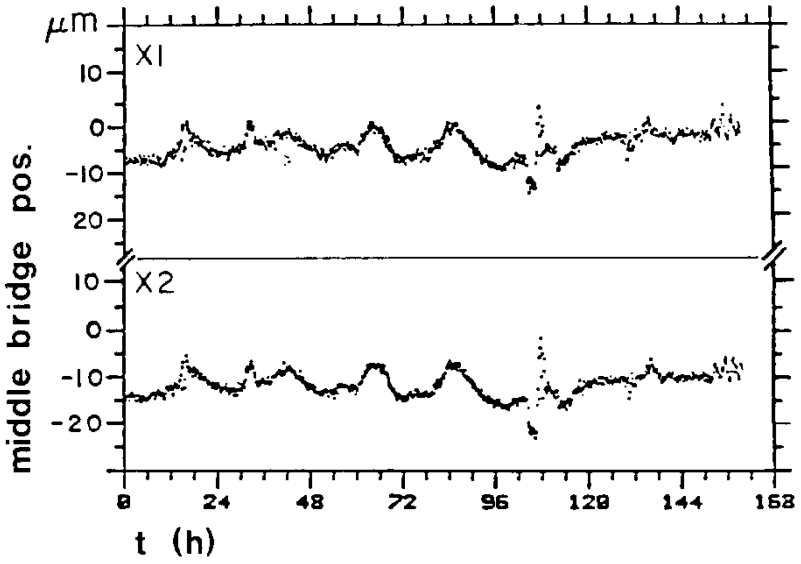


Figure 4. The horizontal position of the middle glass plates with respect to the outer glass plates of the L3 Muon chambers, measured in parallel with two 4QD Rasnik systems with a precision better that 5 μm. The 24 h periodic structure can be well explained by variations in the ambient temperature. The spikes may be explained by people touching, or walking on the chambers, or by sudden changes in gas pressure: the muon chambers acted as balloons, changing their geometry due to pressure variations.

3 CCD-Rasnik

The requirements of the ATLAS muon spectrometer were not met by the 4QD system: for a larger range of measurement, 20 mm square 4QD's would be needed, with complicated calibration [9]. The original NASA engineers who developed the analogous system also understood the need for wider dynamic range in their original alignment system of the late 1970s [3], but their solution was

2019 JINST 14 P08010

to mechanically servo an optical element in front of the photodetector to keep the image centred after deflection, using the correction signal as the alignment measurement. Although active mechanical compensation of detector elements have been considered in high-energy detectors before [10], it introduces prohibitively large additional complication, expense, and accuracy/reliability issues, plus standard motors will have difficulty functioning in the magnetic field in the ATLAS muon region. To solve this, a new system was developed using pixel image sensors that were becoming affordable, and a coded mask on the light source.

The CCD-Rasnik set-up is shown in figure 2b. By applying a back-illuminated coded mask as light source, a section of this mask is projected by the lens onto the CCD. With adequate coding of the mask, the section can be uniquely identified, and the dynamic range of the alignment system is determined by the dimensions of the coded mask. The image data contains information of the X and Y coordinates of the image position relative to the image sensor. The image scale can be determined, and therefore the ratio of the image distance and the object distance. With this, the 3rd (Z) coordinate of either the sensor, lens, or mask, with respect to the other two, can be recorded. In addition, the relative rotation of mask and sensor around the optical (Z) axis follows from the image data. A CCD-Rasnik system thus measures four degrees of freedom. Note that a transversal displacement of the lens results in a twice as large image displacement on the sensor.

It should be noted that, in vacuum, the precision of the measured sagittas, of the measured scale and of the measured rotZ does not depend on the *span* of the system. In practical systems, the precision is limited by density fluctuations in the air.

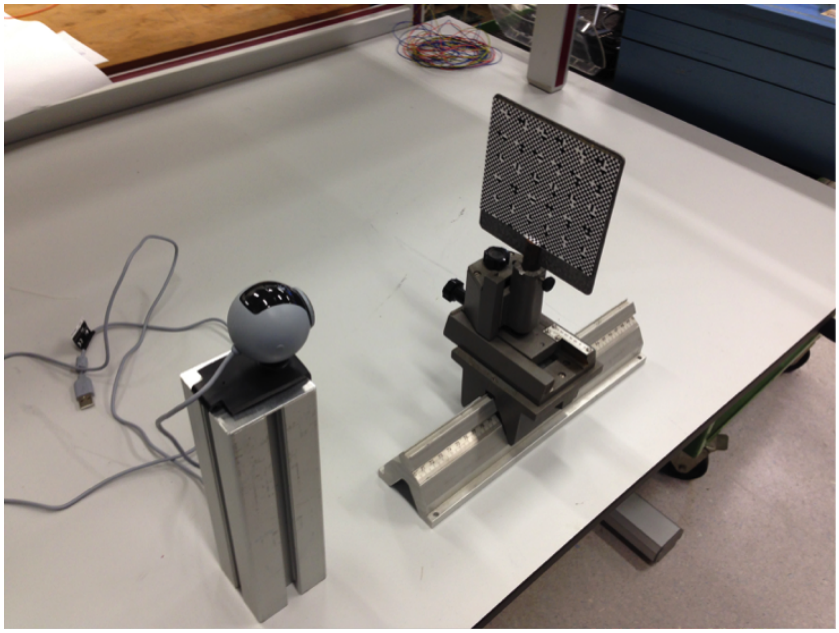
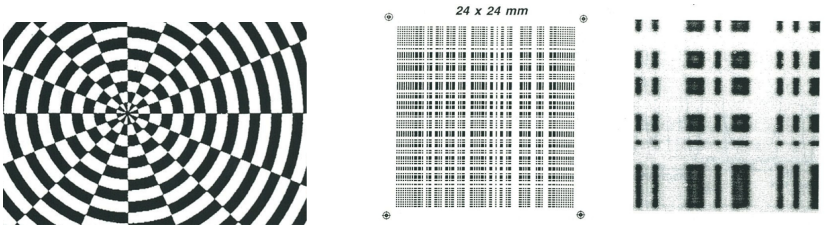


Figure 5. The ChessField pattern printed on paper. With a well-fixed webcam, three translations and the rotation around the optical axis of an object attached to the pattern can be measured with high precision. For instance, with this low-cost set-up, rail sag be monitored during the passage of a train.

2019 JINST 14 P08010

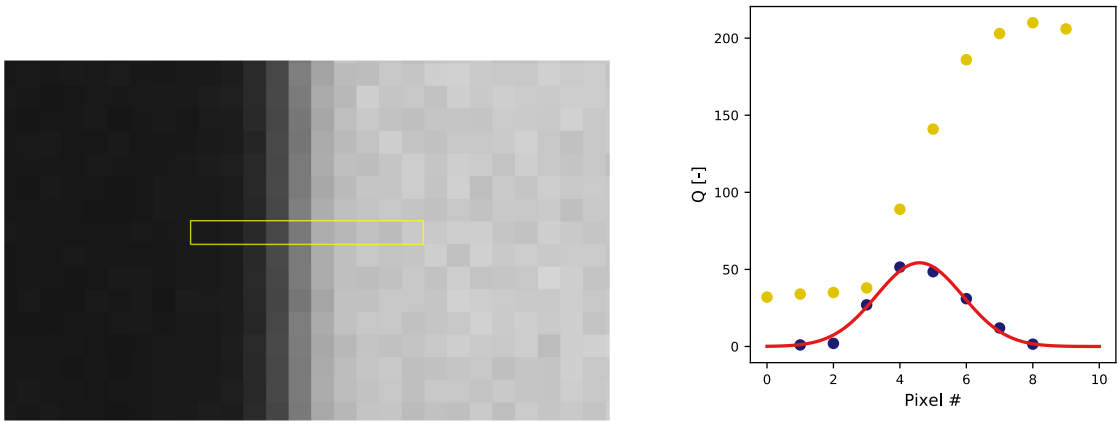


(a) (b)

Figure 6. (a). The *Dart* coded mask. The width of the black/white bands in this mask is 352.7 μm. (b). Left: the double barcode coded mask, and (right) its image from the pixel image sensor.

3.1 The coded masks

As optical object, any light-emitting patterned surface could be applied, in principle. In its simplest form it consists of a black-white pattern, printed on paper, illuminated by daylight as shown in figure 5. With the *Dart* mask as shown in figure 6a, first results were obtained, confirming the wide dynamical range and the high precision of the system: an RMS sagitta of X and Y better than 2 μm was measured [11]. Away from its center, this mask has contours in only one direction, and there the performance of the system would be worse. A better mask was based on the well-known bar code pattern, in which the vertical code is superimposed onto the horizontal code, resulting in a large number of horizontal and vertical contours, shown in figure 6b [12, 13].



(a) (b)

Figure 7. (a). A zoomed-in image of a vertical contour. (b). Yellow: the light Q_n on a row of n = 1-10 pixels. Black: differentiation: $Q_n = Q_{n+1} - Q_{n-1}$.

Figure 7a shows a zoomed-in image of a black-white transition forming a vertical contour. The content of ten pixels selected crossing the contour is shown in figure 7b. With the data from this row of pixels, the contour X position can be calculated, for instance by simply fitting a Gaussian through the curve after differentiation, which is equivalent to fitting an ERF function through the light values of a row of pixels. Fitting a Fraunhofer Edge diffraction curve is closer to reality, but a systematic deviation cancels out since there are equal numbers of black-white and white-black

2019 JINST 14 P08010

contours in an image. The precision of determining the contour position from the info from a row of pixels depends *only* on the width of the black-white transition and the pixel noise (including variations in pixel offset, pixel gain, and inter-pixel crosstalk).

Since any row of pixels crossing a vertical contour can produce a contour position sample, a contour position is sampled many times. On a low-cost webcam CMOS pixel sensor of 3 mm x 2 mm, projected parallel contours should be spaced by at least 100 μm avoiding too weak difference between black and white (diffraction limited image). As a consequence, a total correlated contour length of 60 mm can be reached, both in horizontal and vertical direction. With pixels of 6 μm square, the contour position can be sampled 10^4 times, reducing, in principle, the error in the image X position with a factor 100 with respect to the spatial resolution of a single contour sample. This explains the potentially extremely good spatial resolution of CCD-Rasnik.

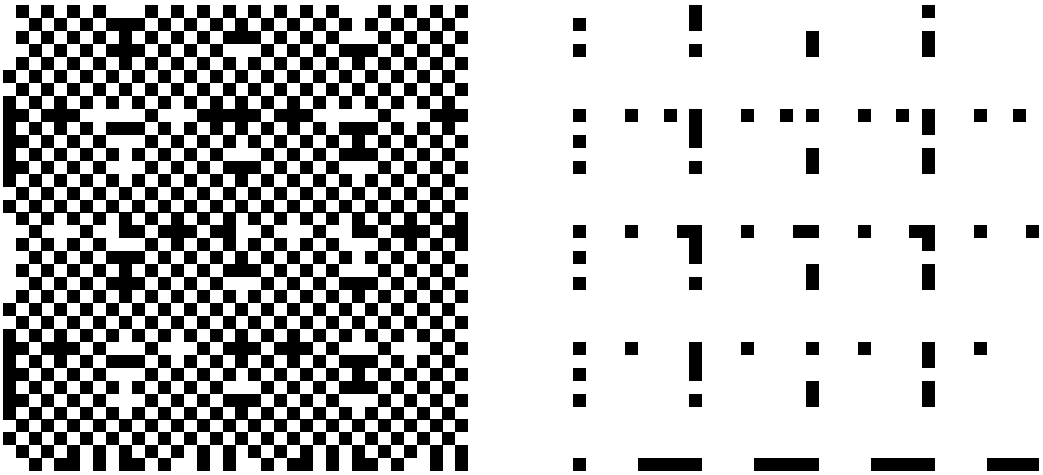


Figure 8. Left: a part of the ChessField mask. Right: the code after an 'XOR' with a continues chess field pattern.

Obviously, the total contour length included in an image should be maximised. A chess field pattern is evidently the best solution for this. An image, however, displaying a section of such a mask, does not allow for determination which section is imaged. The chess fields of the n^{th} row and m^{th} column can be inverted being black or white, providing a digital number for the local row and column. This defines the *coarse* position of the mask section, while the positions of all the vertical and horizontal contours, forming a grid, define the *fine* position of the image on the pixel sensor. The ChessField mask and its coarse coding is shown in figure 8 [14].

For this pattern, the size of a chess field determines the total contour length in an image, and its maximum is set by the Modulation Transfer Function (MTF) of the optical projection by the lens on the pixel sensor. Masks were made with a chess field size of 30, 60, 85, 120, 170, 340, 600 and 1000 μm, always coarse-coded in the 9th row and 9th column. These masks could be selected to match the image scale factor S varying between 0.25 and 4. High-precision masks were obtained from the photo-lithographic industry. Contact copies, maintaining sufficient precision, were used for mass production, costing less than 1 €/cm^2 [16]. Mask dimensions up to 5" square, 3 mm thick, are standard. As mask material lime glass, fused silica, and quartz can be selected, depending on thermal requirements.

2019 JINST 14 P08010

The image analysis routine should find a match between the observed image and a part of the coded mask. In practise, a match is found with the mathematical model of the mask, so the precision by which the contour positions on the coded mask follows this model is essential. In the case of a chess field pattern, the contours form a set of parallel equidistant lines, with one set in the X and the other set in the Y direction. In mask printing technology, a contour is positioned between two smallest surface units (mask pixels) of which one is transparent and one is black. Round-off errors may have caused, in the past, a contour to be off by 0.4 micron, which is large with respect to a potential precision, and should be avoided if high precision is required. In state-of-the-art mask technology, this error is reduced to 50 nm or even smaller. Information provided by the mask manufacturer is key here [16]. Note that the coded mask is the essential and unique component of a CCD-Rasnik system carrying precision. For future applications, using a QR pattern instead of the ChessField pattern could be considered.

3.2 The back-illumination of coded masks

For the back-illumination of the coded mask, two principles are possible: the opaque Lambert radiator [17] as diffusor, illuminated by one or more LEDs, and the application of a condenser lens with one LED in its focal point, see figure 9. The diffusor type is simple, but only a small fraction of the LED light falls onto the projective lens due to its Lambert properties [18]. For the in-plane systems of ATLAS [19], a diffusor (Opaline 3 mm) was applied to create a homogeneous illumination. Here, a Lambert radiator was necessary because of multiple light beams departing from one light source (see (3.6.1)). The diffusor was illuminated with a 3 x 3 array of LEDs (HewlettPackard HSDL-4420 #11) which were selected after radiation hardness tests showing that they could withstand a thermal neutron dose of 10^{15} n/cm^2 [20]. The infra-red LEDs had a wavelength of 875 nm; this light propagates unhindered through colored-glass filter [21] placed in front of the pixel sensor, while disturbing background (day) light is blocked.

For an optical system with finite aperture, the light intensity of an image point is the result of phase contributions departing from the projecting lens. Covering the right-hand half of the lens for instance, results in a significant shift of the image on the sensor. The homogeneity and stability of the back-illumination of the coded mask should therefore be sufficient for the actual application. The total light output of the light source, however, may change: the image position does not change with an overall variation in the light intensity.

Figure 9. The back-illumination of the coded mask. Left: an array of LEDs illuminate a diffusor. Center: the light of a single LED, placed in the focus of a condenser lens, is beamed through the mask. Right: the condenser lens is replaced by a (plastic) Fresnel lens. The distance between the mask and the Fresnel lens should be larger than the depth-of-field of the optical system, to avoid a circular pattern to be projected on the pixel sensor, superimposed on the mask pattern.

2019 JINST 14 P08010

An image, projected by a lens, is blurred due to diffraction: the smearing of contours is inversely proportional to the lens diameter and proportional to the wavelength of the light. The wavelength λ of the LED(s) should be chosen small in order to minimise image blurring due to diffraction, if the ultimate precision is required. Since sub-μm precision is easily reached with larger wavelengths, infra-red LEDs were often applied, enabling shielding the image sensors against daylight. In other cases, the wavelength may be chosen such that the effective focal length of the projecting central lens matches the actual object and image distances of the system. By selecting the wavelength of the LED, the focusing of a Rasnik system with a fixed lens can be adjusted; as an example for BK7 glass, the focal distance varies 2% going from blue to red light [22].

If a light source is only used for one line-of-sight, then its light can be beamed into one direction, and sufficient intensity can be achieved using a single LED placed in the focus of a condenser lens, or a low-cost plastic Fresnel lens: see figure 9. Here, the LED must be carefully 3D-positioned with respect to the lens. The homogeneity of these sources is not as good compared to Lambert radiators.

3.3 The projective lenses for CCD-Rasnik systems

3.3.1 Standard plano-convex singlet spherical lenses

Because of the small aperture of order f/100, the lens of a monochrome CCD- or 4QD-Rasnik system may be a simple low-cost spherical plano-convex singlet [8]. Its focal length equals one fourth of the span l between light source and sensor, assuming equal object and image distances (scale S, being the ratio of the image distance and the object distance s_i/s_o = 1). Lenses with a large diameter favour the light yield on the image pixel sensor, reducing the negative effects of pixel noise and background light, and reduce image blurring due to diffraction. Disadvantages of large diameter lenses are costs, the larger space required for the light beam, and the smaller depth-of-field of the system, possibly imposing a smaller variation in the object and image distances. A large lens mass may introduce particle scattering: if relevant, a plastic planar Fresnel lens could be applied. A large aperture brings non-linearities in the image, e.g. the pincushion effect, which is to be avoided.

In practice, for CCD-Rasnik systems with centred lens, the lens diameter D equals about 1% of the span l between mask and sensor. These systems are *diffraction limited* systems where the image sharpness is determined by the diffraction in the lens - sensor geometry. The diffraction results in a contour sharpness in the order of $\lambda.l/D$ = 50 μm, allowing the projection of at least tens of parallel contours onto a typical webcam pixel sensor area of 3 mm × 2 mm. The dimension of the ChessField squares of the mask should be chosen at least three times larger than the contour sharpness.

If the Z position of the image sensor or the mask can be varied and adjusted, low-cost standard series plano-convex lenses can be applied [23]. If the position of the image sensor and the mask is defined and fixed by the geometry of the object(s) to be monitored, custom specific lenses with a specified focal length must be obtained. The depth-of-field equals +/- 10 mm for standard diffraction-limited systems, setting the tolerance of the focal length for geometry-fixed systems. Note that for lens-centered systems, only the distance between mask and sensor matters: the system is focussed if this distance equals 4x the focal length of the lens. In a focussed system, the lens may be translated along the optical axis by 10 % of the span without losing sharpness.

2019 JINST 14 P08010

If the mask illumination is sufficiently bright, the diameter of the lens can be reduced until the blurring of the image squares by diffraction is barely tolerated by the image analysis. Now the depth of field of the system is maximised, and this maximum depth of field is proportional to the second power of the square width and the first power of the wavelength of the mask illumination.

Since LEDs emit light with a rather peaked wavelength distribution, singlet lenses can be applied: otherwise achromats should be used. Some tuning of the effective focal length is possible by choosing a LED with a specific wavelength.

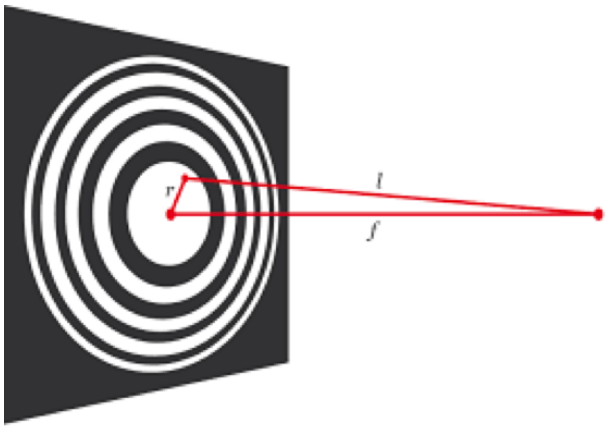


Figure 10. Principle of the zone lens: the image formation is based on interference.

3.3.2 Zone lens

Recently, the commercial availability of standard spherical plano-convex lenses with large diameter and large focal length has seriously dropped [24]. The principle of a zone lens is set out in figure 10. With state-of-the-art mask technology, the printing of patterns, with high-precision, is possible at low cost. In a first attempt to realise a zone lens, the pattern for f = 3750 mm, λ = 589.3 nm was printed on 50 μm thick kapton foil [16]. Although the precision of the pattern was adequate, the quality of the lens was limited due to variations in the thickness of the kapton foil. A 4" square lime glass substrate was $\lambda/4$ flattened at both sides [25], and chrome and photoresist layers were applied [26], resulting in a zone lens with MTF-limited quality: figure 11 shows its resulting image.

Zone lenses are foreseen to be applied in Rasnik systems for ITER-CEA in Cadarache, France.

3.3.3 Microscope objective

In CCD-Rasnik systems the image is magnified by positioning the lens closer to the mask: see figure 2b. A mask displacement, with respect to the lens + image sensor, is also magnified: for a given precision of the measurement of the image position on the image sensor, the precision of the mask displacement scales with the magnification. Singlet lenses cannot be placed close to the object since optical aberrations become relevant. Instead, a microscope objective with large aperture can be used, reducing diffraction while increasing the light strength. With RasMic, direct and linear 2D measurement of object displacements, with precision in the order of 10 pm may be feasible, competing with interferometry-based displacement monitor systems: see figure 17.

2019 JINST 14 P08010

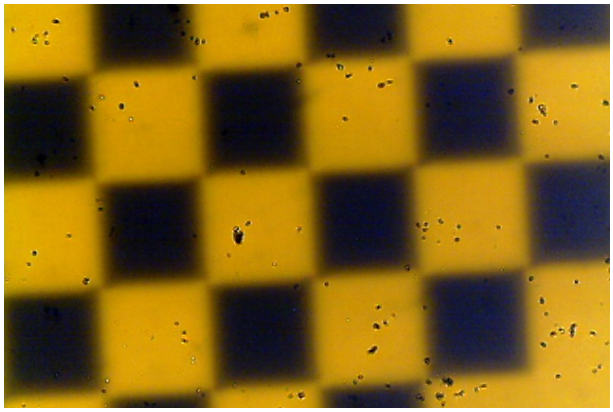


Figure 11. Image of CCD-Rasnik system with zone lens, centered between the 600 μm ChessField coded mask and the Logitec C300 webcam image sensor. The span between the mask and image sensor is 15.00 m. The mask was back-illuminated with an HP Sodium street light. The sharpness of the image is limited by the presence of two wavelengths of this light (589.0 and 589.6 nm). The sharp black spots are due to dust particles on the image sensor; these do not affect the image position analysis.

A disadvantage of applying a microscope objective is that the image scale is not proportional to the object distance: the magnification is optimised for a design value. The image scale, and therefore the Z position of the coded mask, is not measured.

3.4 Image sensor for CCD-Rasnik

For ATLAS, the *RasCam* sensor was developed by Nikhef, and deploys the commercially available radiation hard VLSI Vision VV5430 CMOS sensor [20]. It outputs a standard CCIR composite monochrome video signal allowing for synchronisation, avoiding time (thus X-position) jitter. Newer systems often use a simple USB camera; in particular, the camera board out of a consumer webcam, like the Logitech C250 or C300 are used. Experiments have also been performed with a larger B/W sensor (Videology 24B1.3X) and a Firewire connected unit (Pike F100B) using the Kodak KAI-1020 which offers the possibility of higher frequency measurements (> 100 Hz). In commercial applications where longer cable lengths were required, OEM IP camera units are used, which can be powered over their ethernet cable, like the Wision WS-M5P31/38, which is based on a Aptina MT9P031 sensor. Recently, the Mercury MER-051-120GC is applied, due to its frame rate of 120 Hz. Table 1 lists the geometry of the above mentioned sensors.

In ATLAS the analog signal of the RasCam sensors is multiplexed, so only a few PCs are needed to control and read out several thousands of Rasniks. Here, the Data Translation DT3162 frame grabber was used. Special care was taken to synchronise the video 'line' signal, avoiding jitter in the measured X-position of the image. More recent systems no longer use a separate grabber: instead digital sensors with a USB, Firewire or Ethernet output connection are applied. In particular, the use of power over ethernet (PoE) greatly simplifies the cabling of larger setups using commodity equipment.

The precision of a CCD-Rasnik system is proportional, in principle, to the square root of the total contour length. The images are typically diffraction limited; the features in the image projected on the sensor are typically larger than 80 μm. These features can be sampled by order 5 pixels: as

2019 JINST 14 P08010

Table 1. Dimensions of image sensors applied for CCD-Rasnik and RasDif systems.

Type	pixel count	effective size (mm)	pixel size (μm)
Vision VV5430	384×287	4.66×3.54	12
Logitech C250	640×480	3.84×2.88	6.0
Logitech C300	1280×1024	3.67×2.96	2.9
Videology 24B1.3X	1280×1024	6.66×5.32	5.2
Kodak KAI-1020	1000×1000	7.4×7.4	7.4
Aptina MT9P031	2592×1944	5.70×4.28	2.2
Mercury MER-051	808×608	3.87×2.92	4.8

a consequence the optimal pixel size is in the order of 10 μm square. The surface of the cheapest webcam sensors is sufficient to achieve sub-micron sagitta measurements.

See 5.3 for radiation hardness studies of image sensors.

3.5 Image reconstruction algorithms and measurement accuracy in the CCD-Rasnik system

In the image analysis, image data is eventually converted into values for the image X position, image Y position and, the image scale and rotZ. The first analysis was performed by the Draper group for the coded mask shown in figure 6b, confirming the potential sub-micron precision of the CCD-Rasnik system [12, 13, 27]. At Nikhef, first images were taken using the Dart pattern shown in figure 6a. Such an image was compared, pixel by pixel, with a computer generated image. By systematically varying the input parameters of the computer generated image, the RMS was minimised, resulting in the best possible fit for the X and Y position. With this routine, a spatial resolution was obtained equal to half the pixel size of the image sensor. In 1993, this routine took 40 minutes CPU time per image.

For the ATLAS Rasnik systems, an image analysis algorithm by F. Linde and M. Woudstra was based on finding the local maxima of the absolute value of the horizontal and vertical intensity derivative: see figure 12 [28]. This gave good results for ideal images, but it was quite sensitive for image imperfections, and it was slow. One could overcome these issues by application of image enhancement techniques. After this, FOAM was written, based on a refined 2D Fourier analysis, followed by a novel pattern recognition algorithm to reconstruct the positions and values of the code rows and columns, even in extremely low quality images.

The analysis developed at Brandeis University [29] uses the Fourier transform to obtain the approximate width and height of the squares in the image, after which it fits straight lines to the stripes in the vertical and horizontal derivative images to obtain the location, rotation, and magnification. When the mask is rotated about an axis within its own plane, two distortions of the mask pattern occur: first ‘skew’, which is the change in magnification across the image that arises because the squares seen on one side are farther away than the squares seen on the other side, and second ‘slant’, which is the apparent shrinking of the squares perpendicular to the axis of rotation. The Brandeis analysis accommodates skew and slant in its images, and provides a value for both slant and skew, so that the rotation of the mask about an axis parallel to its surface can be measured: see figure 13.

2019 JINST 14 P08010

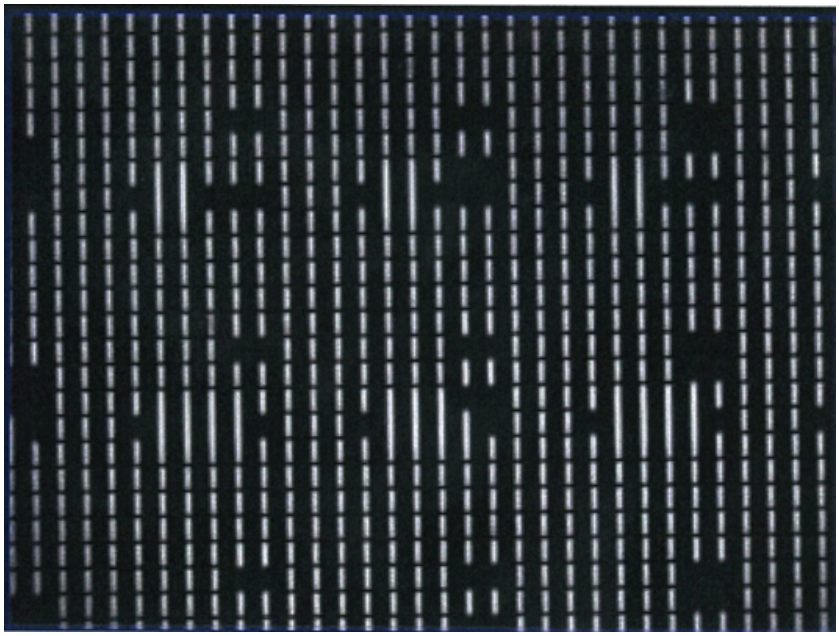


Figure 12. Typical Rasnik image after horizontal differentiation. The light value of a pixel equals the absolute value of the difference between the light values of its left-hand and right-hand neighbour. The apparent edges can be fitted by a set of parallel and equidistant grid lines.

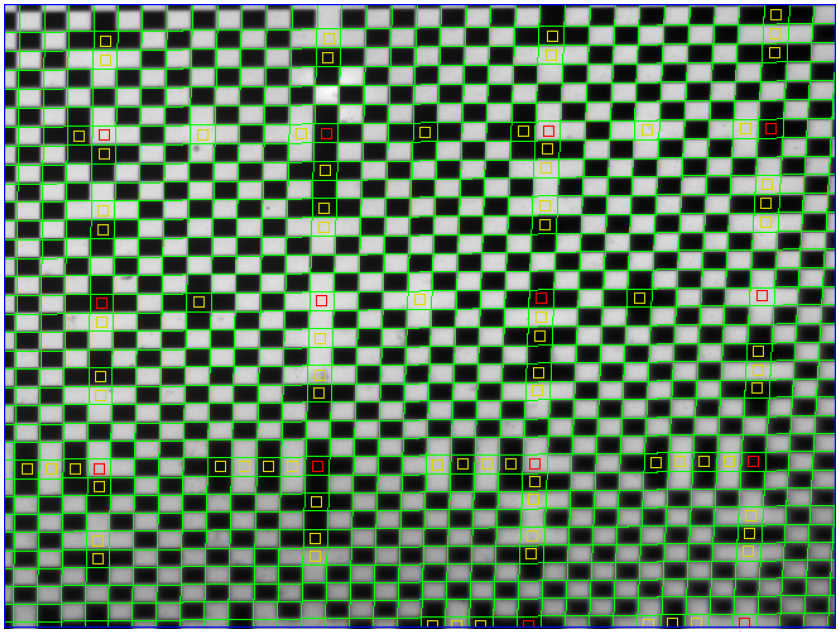


Figure 13. A slanted and skewed Rasnik image with analysis results overlaid. The squares are wider in the horizontal direction than the vertical (slant), and they are smaller on the bottom than the top (skew). Note the marking of code squares with red and yellow rectangles.

In 2008, an image analysis routine was required for the RasDif system (4). Here, the regular ChessField pattern is replaced by the diffraction pattern on the sensor. This routine should be

2019 JINST 14 P08010

independent of the pattern, since the geometry of the pattern remains unchanged, and only translations of the pattern are of interest. For this, a 2D Fourier conversion is convenient: the pattern displacement appears in the phase image. The first image taken functions as a start-and-set image, and translations in X and Y, with respect to the first image, can be calculated. It was shown that the diffraction pattern can be calculated with high precision: such a perfect image, free of pixel noise, can act even better as start image. With this routine, the data of figure 30 was obtained [30].

The performance of 2D Fourier-based FOAM routines was excellent and very fast, and new routines for the ChessField mask of the CCD-Rasnik system were written, based on 2D Fourier conversion [15, 31]. Eventually, routine SOAP was written, based on 2D Fourier analysis, and including new pattern recognition for the coarse code.

The optical axis of a CCD-Rasnik system is defined as the line starting at the centre of the image pixel sensor and going through the optical centre of the lens. The image analysis routines developed for the ChessField mask put out the coordinates of the cross point of the optical axis and the mask, using the coordinate system defined by the mask itself [15]. As a result, the orientation of the mask defines uniquely the Rasnik coordinate system. As long as the optical axis crosses the mask, the coordinates can be given. The (transversal) range of a CCD-Rasnik system is therefore defined by the size of the mask, which may be arbitrarily large. Note that the range, in terms of lens displacement (sagitta), is half of the ranges in terms of mask- or sensor displacements. Since masks can be made with extreme high precision, the transversal displacement of the mask is known directly without calibration. If the pixel pitch of the sensor is known, the transversal displacements of the lens and sensor are also know since the image scale S is known as well. The ChessField mask is the heart of the high-precision alignment system with arbitrarily dynamical range in both transversal directions.

Assuming that the image analysis program is free of systematical errors, the ultimate spatial resolution of a CCD-Rasnik system in vacuum is limited by the image pixel noise, which is determined by the sensor electronics and the fluctuation in the light falling on a pixel. This pixel noise can be obtained from a set of images from a static system where the image position is stable. By means of a Cramér-Rao Lower Bound (CRLB) analysis, the ultimate performance of a CCD-Rasnik system can be determined, approaching a value of 1 nm (in vacuum, using standard LEDs, optics and webcam image sensor). In addition, images have been generated using Monte Carlo simulations in which pixel noise was included. After analysing these images, spatial resolutions of 50 pm seem feasible using pixel image sensors with 12 bit dynamical range, such as the MER-051. This value can only be reached if error sources like pixel cross-talk, image distortions, and non-symmetry in black-white and white-black contours, are sufficiently small. Note that this precision is three orders of magnitude smaller than a pixel side, and two orders of magnitude smaller than one wavelength of the applied light.

In appendix details are set out of the SOAP ChessField image analysis program.

3.6 Results of the CCD-Rasnik system

3.6.1 The alignment of the muon chambers in the ATLAS muon spectrometer

For the muon spectrometer of the ATLAS experiment, almost six thousand CCD-Rasnik systems have been made which are all operational to date: the layout of some of these systems is shown in figure 14. A set of an inner, middle and outer chamber forms a tower with its axis pointing

2019 JINST 14 P08010

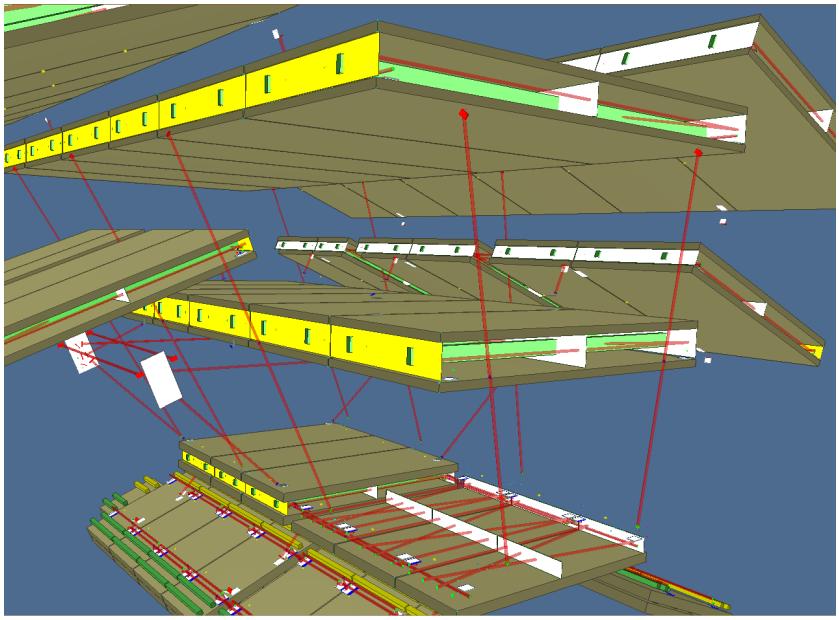


Figure 14. The layout of the InPlane and Projective alignment systems of the ATLAS Muon Spectrometer. The InPlane systems are depicted in detail in figure 15.

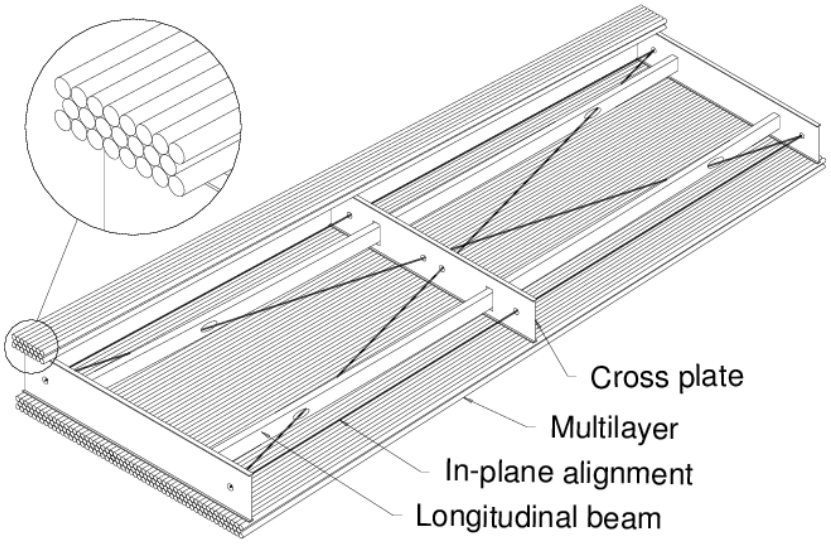


Figure 15. The layout of the InPlane alignment system a an ATLAS muon chamber.

towards the interaction point (from which muon tracks depart). At the four corners of each tower, 'projective' alignment systems are mounted. The optical axis of each projective alignment system points towards the interaction point, so the misalignment of the chambers is measured in the most direct, and therefore effective, direction. The three components of all Projective systems were calibrated in a station at Nikhef, prior to their installation in-situ at CERN on precision mechanical references on the chambers [32].

2019 JINST 14 P08010

Each chamber module is equipped with an 'InPlane' system, consisting of two edge systems and two diagonal systems, which share light sources and image pixel sensors as shown in figure 15. With this, the chamber sag and torsion is monitored. The components of the InPlane systems were mounted during the construction of the chamber in its nominal geometry. As a result, the change in the readout data of the InPlane systems is a direct measure for the deformation of the chamber.

The complete set of data of all alignment systems provided redundancy, such that the performance of the systems could be analysed. All deviations could be explained by mechanical offsets of components. This was confirmed by analysing straight tracks in cosmic ray and collision data with magnetic field off. After taking these offsets into account, variations with time are limited to 5 μm and can be well explained by thermal expansion.

The span of the systems varies between 1 and 6 m. The precision in X and Y, in terms of the relative displacement of the lens, varies between 0.1 and 3 μm, for each measurement, limited by the fluctuation in the density of the ambient air or by pixel noise. From the consistency of the data, the absence of drift of individual systems over a period of 15 years could be established [32–34].

3.6.2 The SmallSpan system

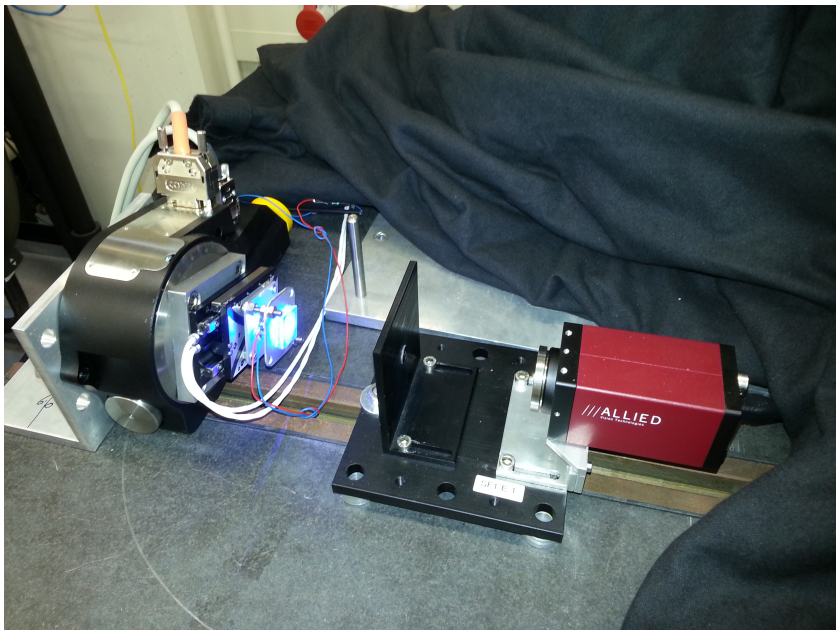


Figure 16. The SmallSpan system: mask (ChessField 85 μm) and back-illumination are mounted on translation/rotation actuators. Lens: plano-convex singlet f = 40 mm, diameter diafragma 2 mm. As pixel image sensor, the Logitech C250 was applied, or the Pike F100B. The distance between mask and sensor was 160 mm; the lens was centred between mask and sensor.

For studying the performance of CCD-Rasnik with webcam-based image sensors, a system has been tested in the SmallSpan set-up: see figure 16. The response of the system was measured for systematic variations in the X position of the mask, and for mask rotations around the optical (Z) axis. The precision of the actuators was a limiting factor in the determination of the system's

2019 JINST 14 P08010

systematical errors. For measuring the X and Y positions of the mask, a precision of 60 nm (RMS) was found, and the image scale (magnification) could be measured with an RMS of 0.2 x 10^{-4}.

3.6.3 The RasMic system

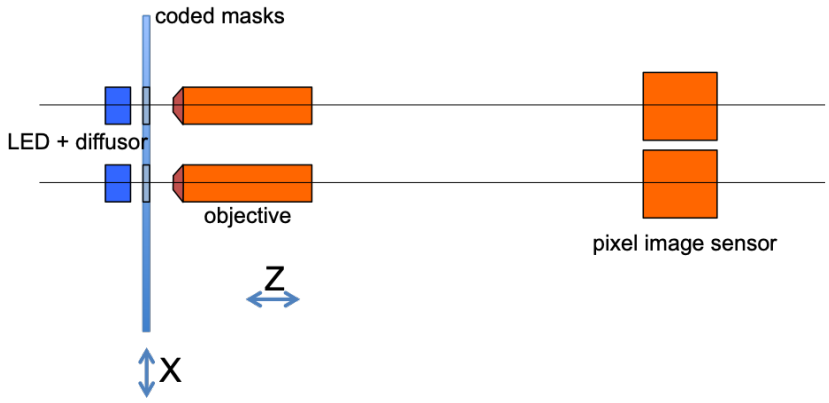


Figure 17. Top view of the RasMic set up. The coded masks are mounted on a common carrier, which can be translated in the X direction by means of a manual dial micrometer gauge or by a piezo actuator (PiezoSystemJena HPSt1000/25-15/7). The coded masks are the objects for two microscope objectives (Newport M-20X), forming 20x enlarged images at 180 mm downstream the objectives. There, the image pixel sensors CAM1 and CAM2 are placed (Logitech C250 or Pike F100B). The coded masks are back-illuminated by blue LED light (StarLED-1W-BL), combined with Opaline diffusors.

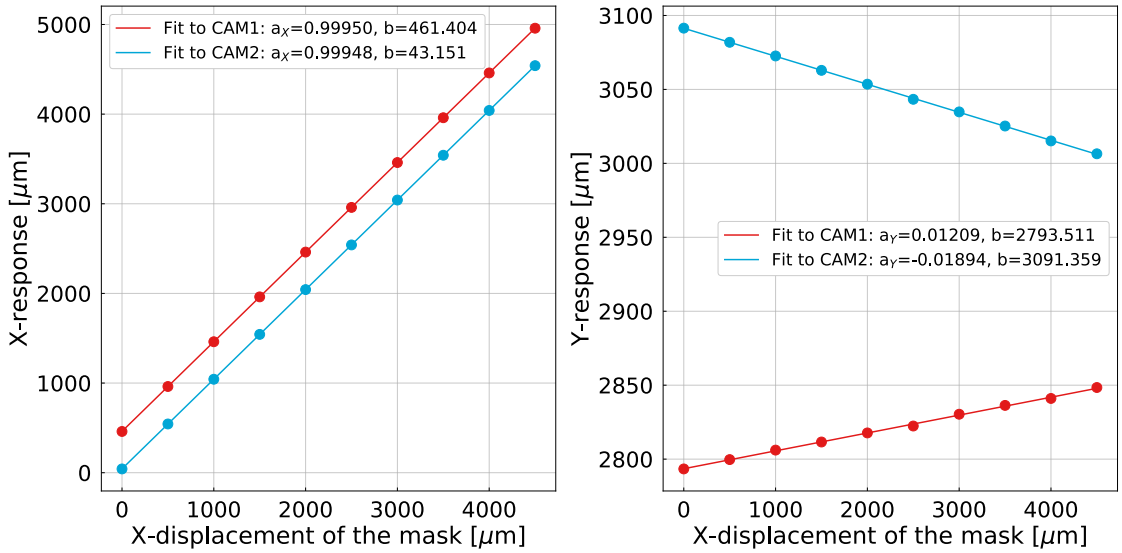


Figure 18. The X and Y response of both RasMic systems for 9 different X-positions of the mask. Since two masks are used, the X and Y offsets of both systems differ. At each position, 20 images were recorded taking 0.6 s. The red and blue dots represent the average of 20 measurements, and least-square straight lines were fitted through these data points. The slope (a) and pedestal (b) parameters of the fitted lines are displayed in the plot.

2019 JINST 14 P08010

Table 2. X measurements and X residuals after straight line fit as shown in figure 18.

CAM1 mean x (μm)	Residuals from fit (μm)	CAM 2 mean x (μm)	Residuals from fit (μm)
460.1	1.3	42.2	1.0
961.7	-0.5	543.6	-0.7
1461.0	-0.1	1042.6	0.0
1961.5	-0.8	1543.0	-0.7
2461.1	-0.7	2042.6	-0.5
2959.5	0.6	2541.0	0.8
3460.1	-0.2	3041.6	0.0
3959.8	-0.1	3541.4	0.0
4459.5	-0.1	4041.0	0.1
4958.5	0.7	4540.8	0.0

Table 3. Slope values from figure 18.

CAM	a_x	a_y	M
1	0.99950	0.01209	0.99957
2	0.99948	-0.01894	0.99966

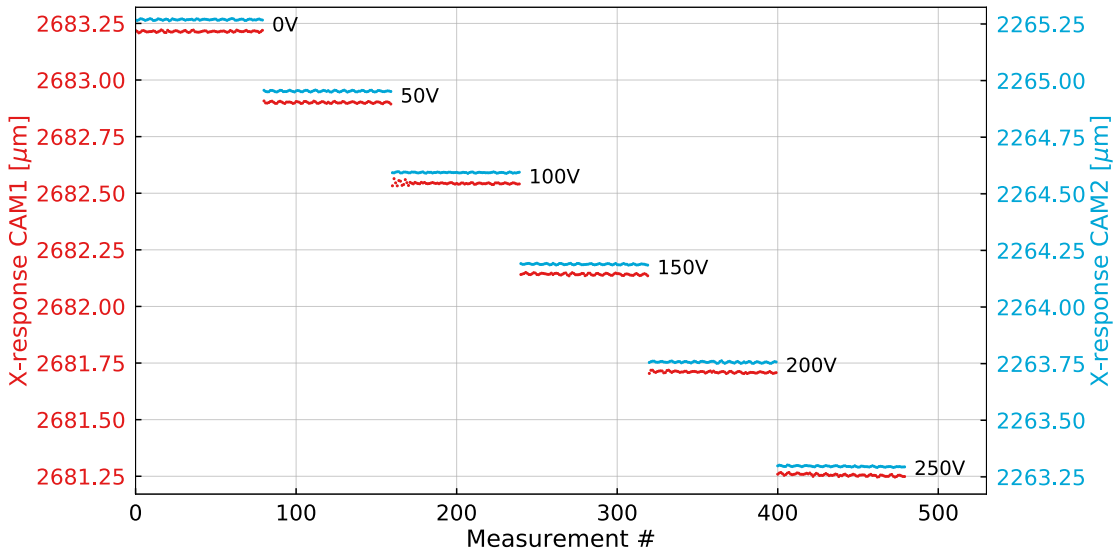
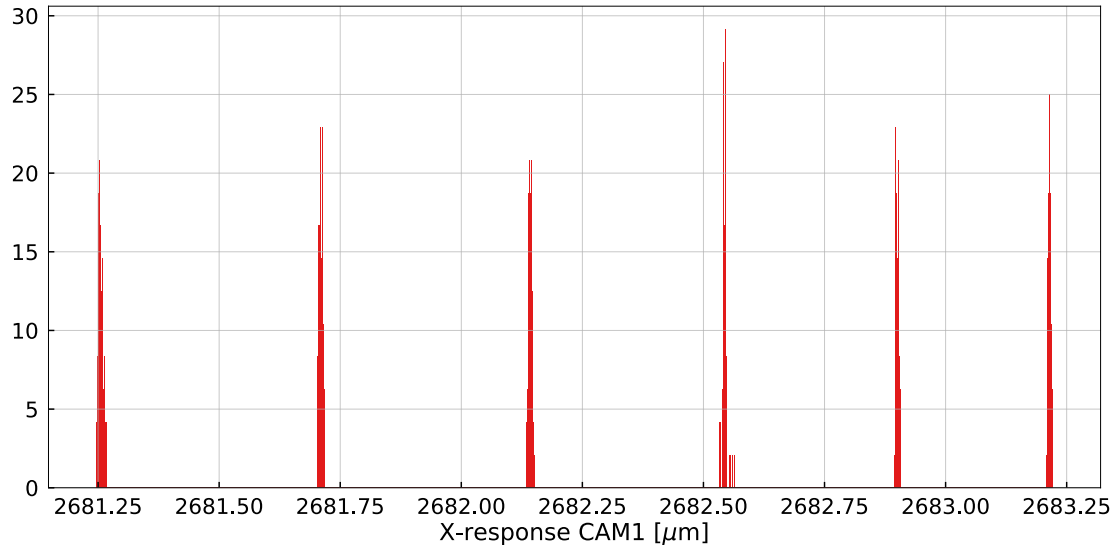
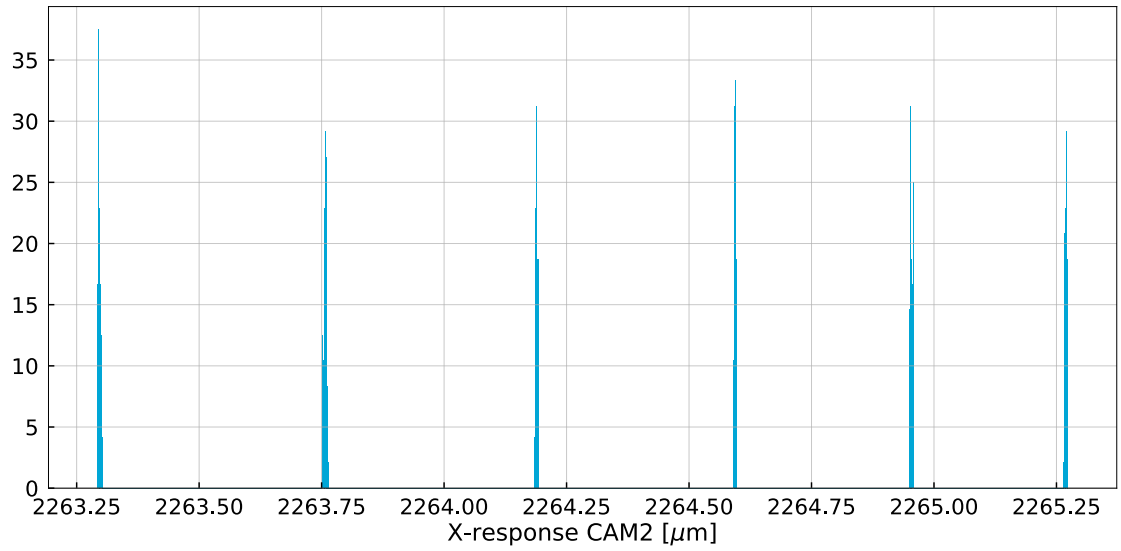


Figure 19. The X-response of the two systems for seven X positions of the coded masks. The positions are associated with the indicated voltage, put on the piezo actuator. For a 50 V change, the actuator moves 0.4 μm. In each position, 80 measurements were recorded, taking 2.4 s. The periodic variations reveal mechanical vibrations.

2019 JINST 14 P08010



(a)



(b)

Figure 20. (a). Histogram of the data of CAM1 of figure 19. The bin size of this histogram is 1 nm. The non-linearity of the piezo actuator appears in the smaller distance between peaks, going from left to right. (b). As for CAM2.

The (ultimate) performance of the CCD-Rasnik system has been studied in the RasMic set up, shown in figure 17. Basically, it consists of two systems (CAM1 and CAM2), coupled in parallel, enabling the comparison of two systems to a common change in the mask positions. For this RasMic system, new ChessField masks were made with field squares of 4, 8, 16, 32 and 64 μm, with the best possible (50 nm) precision [16]. Both systems are mounted on a solid common base plate. The masks of both systems are mounted in a common holder, which can be translated in the X direction, being horizontal and perpendicular to the optical axes. The difference between the X readings of

2019 JINST 14 P08010

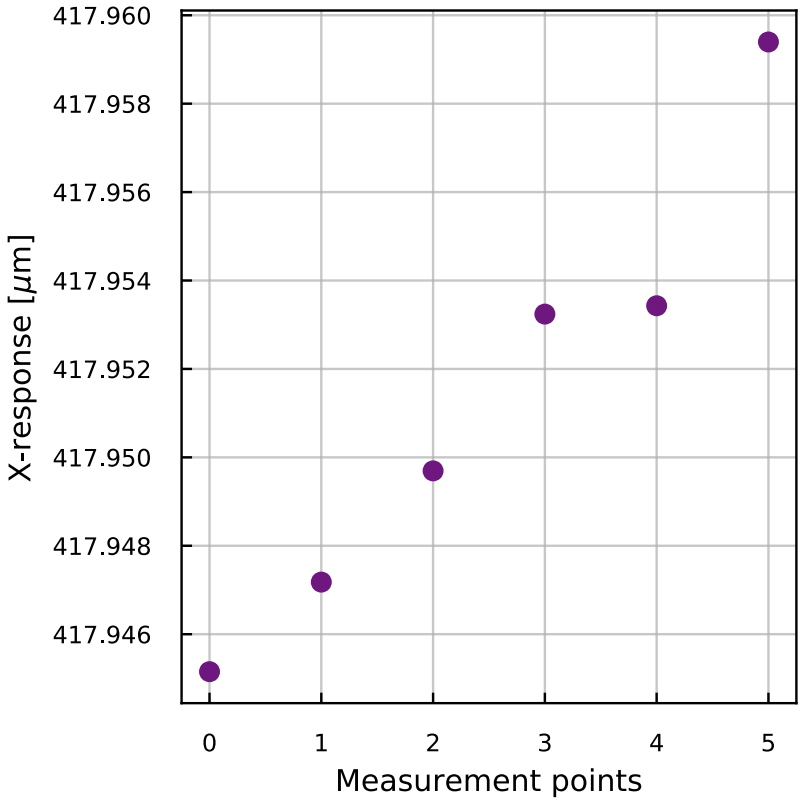


Figure 21. Difference in mean X-responses between the two RasMic systems.The six points of measurement correspond to the six peak positions in figure 20a.

Table 4. Data and differentials as shown in figure 21.

CAM1 mean X (μm)	CAM2 mean X (μm)	CAM1 minus CAM2 mean X (μm)
2683.215	2265.269	417.945
2682.901	2264.954	417.947
2682.544	2264.594	417.950
2682.142	2264.189	417.953
2681.710	2263.757	417.953
2681.256	2263.296	417.959

the two systems should essentially be constant. Deviations in the difference are a measure for the precision of the CCD-Rasnik system.

First, the masks were commonly translated in the X direction by means of a manual micrometer dial gauge, in steps of 500 μm over a range of 4.5 mm. The X and Y response of both systems is shown in figure 18. This data was taken with the sensors' maximum frame rate of 30 Hz in order to minimise mechanical changes with time. In table 2, the results are given, as well as the residuals after a straight-line fit. The X-residuals, for each X position, are quite equal for the two systems, indicating that the residual is essentially due to the manual setting of the dial gauge (10 μm/div).

2019 JINST 14 P08010

The X response of a Rasnik system equals X = X_{act} cos(rotZ), where X_{act} is the action of the dial gauge, and rotZ the angle between direction of the actuator and the X direction of the coded mask. For the Y response, Y = X_{dial} sin(rotZ) holds. The rotZ of both masks will be different. In table 3, the slope values are added quadratically: M = sqrt($a_x^2 + a_y^2$). The values for M represent the proportionality of the mechanical displacement and Rasnik measurement. Since the displacement of the two masks is supposed to be identical, the difference between the two M values is not as small as one might expect from the precision of the scale of the masks, specified to deviate less than 5 x 10^{-6} from unity [16]. The deviation from unity may be explained by the non-perfection of operating the manual dial gauge. Note that the coded mask is essentially the only precision-defining component in CCD-Rasnik systems.

In a second set of measurements, the RasMic systems were placed in vacuum (p < 1 mbar) in order to eliminate possible distortions in the direction of light. This time, the masks were commonly displaced in the X-direction by means of the piezo actuator, set at 0 V to 250 V in steps of 50 V. The X response of the two systems is shown in figure 19. The mask displacement in X may be subject to non-linearity in the actuator response, and to mechanical imperfection of the set-up. The intrinsic precision of the CCD-Rasnik system appears, however, in the difference in the outputted Xs of the two systems. When zoomed in, a small periodic structure is visible: this may be caused by acoustical vibrations. Figure 20a shows the same data in the form of a histogram: note the non-linearity of the piezo actuator, and the resolution of the measurement of the X-position. The difference in X-response between the two systems is shown in figure 21 and in table 4. The increase in difference with X can be explained by a change in temperature of the aluminium mask holder: with a distance between the two masks of 54 mm, a change of temperature of 0.01^{o}C results in a change of distance of 14 nm.

The linearity of a RasMic system is in the order of 2 nm, as can be concluded from table 4 and figure 21. If this precision is combined with the (transversal) range of 4.5 mm of figure 18, a range/precision of 2 x 10^6 is reached.

Figure 22 shows X values taken with the largest possible frame rate of 30 Hz. A period is selected where the systematic change-in-time in the measurements was certified to be much less than 1 nm. Figure 24 shows the typical distribution selected data of CAM1, 1000 measurements taking 34 s. The width of the fitted Gaussian distribution ranges from 0.9 to 2 nm for corresponding samples of X and Y data, taken with CAM1 and CAM2. A Fourier analysis of this data has revealed the presence of mechanical vibrations. This has been confirmed by recent measurements applying faster image sensor, enabling to follow the changes in X and Y data with a higher sample rate. Since Rasnik is a three-point alignment system, the effective image position vibration is the vector sum of three oscillations. This obscures the direct comparison of two parallel coupled systems.

In a next phase, the sampling rate is further improved, and CCD-Rasniks will be made insulated from mechanical and acoustical vibrations in the form of a multistep pendular system [35].

In figure 25, the sigma of the Gauss-fitted distribution is plotted against the level of light, emitted by the two LEDs. This light level is proportional to the current through the LEDs. The spatial resolution of the RasMic system approaches a value of 2 nm which is not improved by a further increase of the light intensity. The spatial resolution, at high light level, is not limited by fluctuations in the pixel content (due to quantum statistics or intrinsic pixel noise). If the actual cause of the limit in resolution is found, a much better spatial resolution may be possible, since

2019 JINST 14 P08010

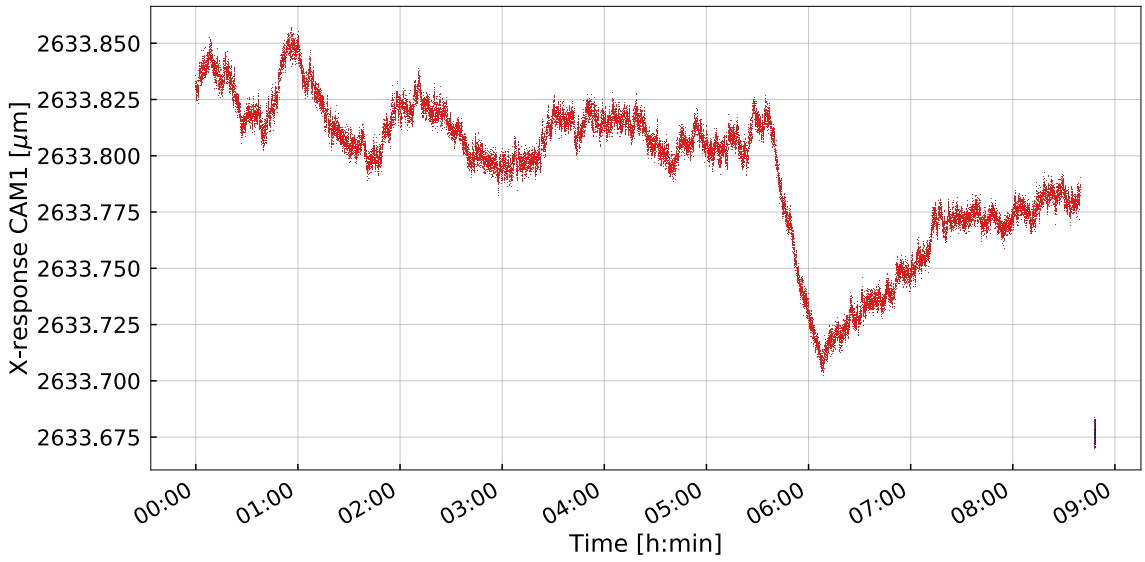


Figure 22. The X value of a data stream from Cam1, taking during a night, storing one image per s. The variation can be well understood in terms of a variation in temperature in RasMic's mechanics. At 8:45, the measurement was stopped. At 8:47, measurements were taken for some seconds, storing 30 images per s.

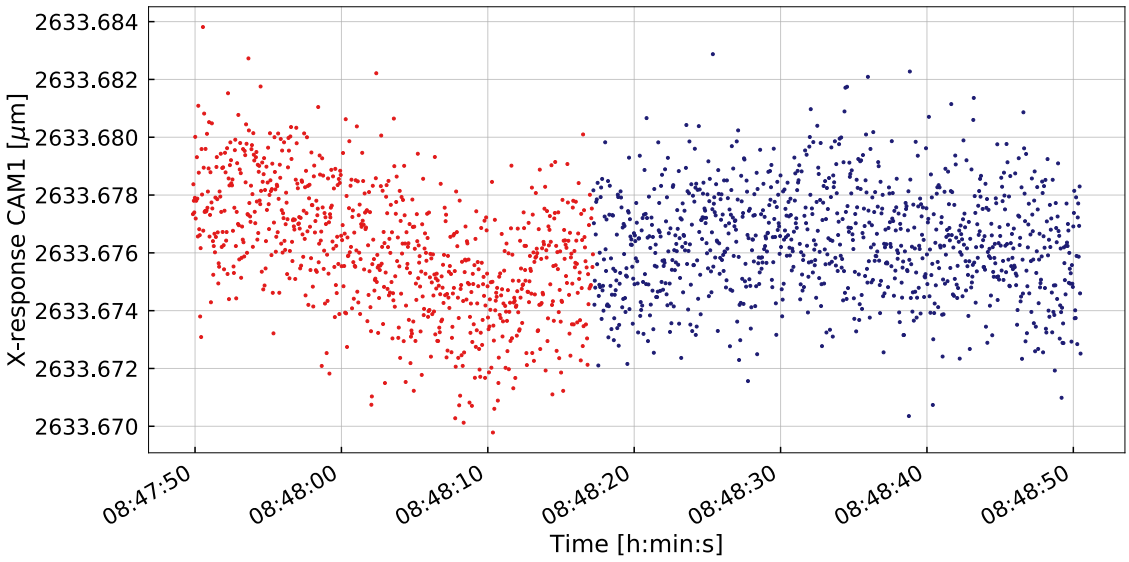


Figure 23. Zoomed-in data of figure 22: the blue selected data is used for the spectrum of figure 24.

the light intensity could be increased significantly: the maximum light level is determined by the thermal expansion of light-exposed components.

3.6.4 The ultimate precision of CCD-Rasnik

The spatial resolution of small-span CCD-Rasnik systems with the lens centred between mask and sensor has been found never to be better than 25 nm, for a variety of image sensors, operating in vacuum. With the RasMic systems mentioned above, a level of 1 nm has been reached: this can be

2019 JINST 14 P08010

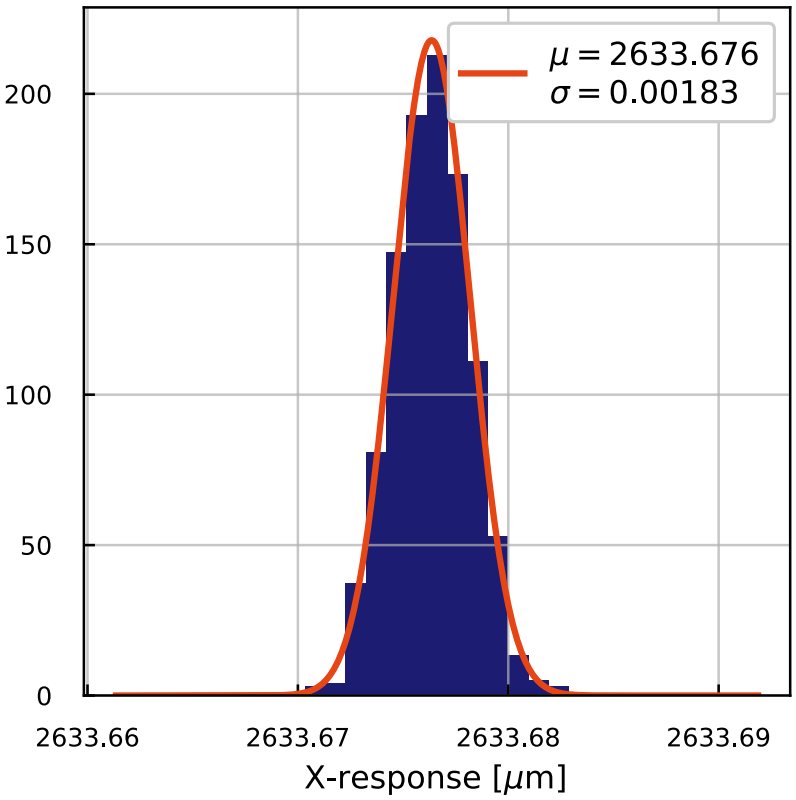


Figure 24. Histogram of the selected blue data of figure 23, in which period the mean value changed little with respect to the spread.

well explained by the scale factor 20 enlargement of the objective. Apparently, the precision in the determination of the image position-on-sensor is limited to 20 nm. This result was obtained with both the simple Logitech webcam image sensor and the much better Pike image sensor, suggesting that disturbances like mechanical vibrations is limiting the value of the ultimate precision. By means of MonteCarlo generated image data, the proper operation of the image analysis routine has been verified to the level of 20 pm (see appendix). By estimating pixel noise, a much better spatial resolution than reached so far is expected. This expectation is also based on Cramér-Rao limits, calculated before [36].

In the following, the precision of the determination of a contour position is analysed. Values (position coordinates and sigma's) are given with respect to the sensor coordinate system, and not in terms of mask displacement. In the RasMic setup, the webcam pixel sensors were replaced by CCDs (Pike F100B). Figure 26 shows a typical RasMic image. At 160 mm behind the objective is the image pixel sensor (Pike F-100 B), magnifying the object with a factor 20. This RasMic system has a measured spatial resolution of 2 - 5 nm, just as in the previous case applying simple webcam image sensors.

In figure 27, pixels closest to a vertical contour, are given a colour varying from red via green to blue, as a function of the local residual. For this, the content of a horizontal row of 11 pixels, symmetrically placed over the contour, is fitted with an ERF function. The residual is defined as

2019 JINST 14 P08010

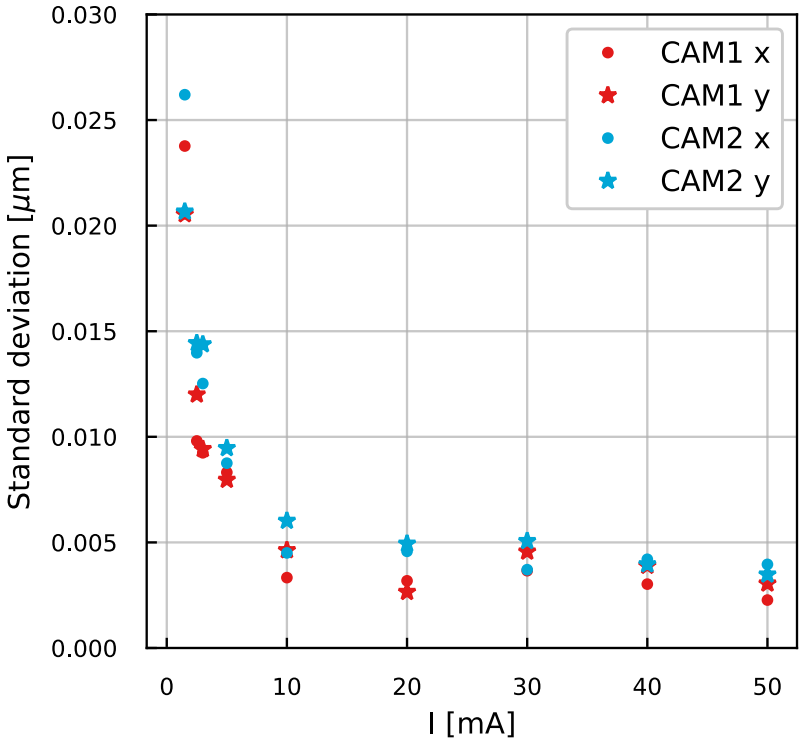


Figure 25. The spatial resolution in X and Y, for both cams, as a function of the current through the LED.

the difference between the bending point's X-position of the ERF function and the local X position of the contour, known from the image analysis. From green indicated contours, the RMS of the contour sample residuals is found to be 0.16 μm. This means that one (vertical) contour X position is sampled with a horizontal row of 11 pixels: compare figure 7a. With about 100 vertical contours on an image of 1000 x 1000 pixels, the total contour could be sampled 100,000 times, in which each of the samples would contribute equally to the X-position of the image. If positioned perfectly, all samples would result in an RMS of an image X position of 0.16 μm / sqrt(100000) = 0.5 nm. In RasMic, this would result in a spatial resolution of the mask position of 25 pm, taking into account the image enlargement with a scale factor 20. The measured spatial resolution of 5 nm is therefore a factor 200 worse than the prediction in the case of an ideal image, where only pixel noise is a limiting factor. This may be explained by the residuals shown in figure 27: the residuals are much larger than their RMS value due to a systematic local and common shifts. The difference in residuals between black-white and white-black contours is clearly visible. This can be explained by fitting the transition with a symmetrical ERF function, where the real dark-light edge transition is a known, non-symmetrical Fresnel/Fraunhofer Edge Diffraction pattern. Since there are equal numbers of samples of black-white and white-black contours, this deviation cancels, in the first order. Other shifts in apparent contour positions can be caused by the presence of local dark or light areas of the coarse code, by errors in the imaging by the objective, by pin-cushion effects, by non-perpendicularity of the objective and the plane of the coded mask, by non-homogeneous back-illumination, by pixel cross-talk by and deviations of the mask pattern geometry. It is not clear

2019 JINST 14 P08010

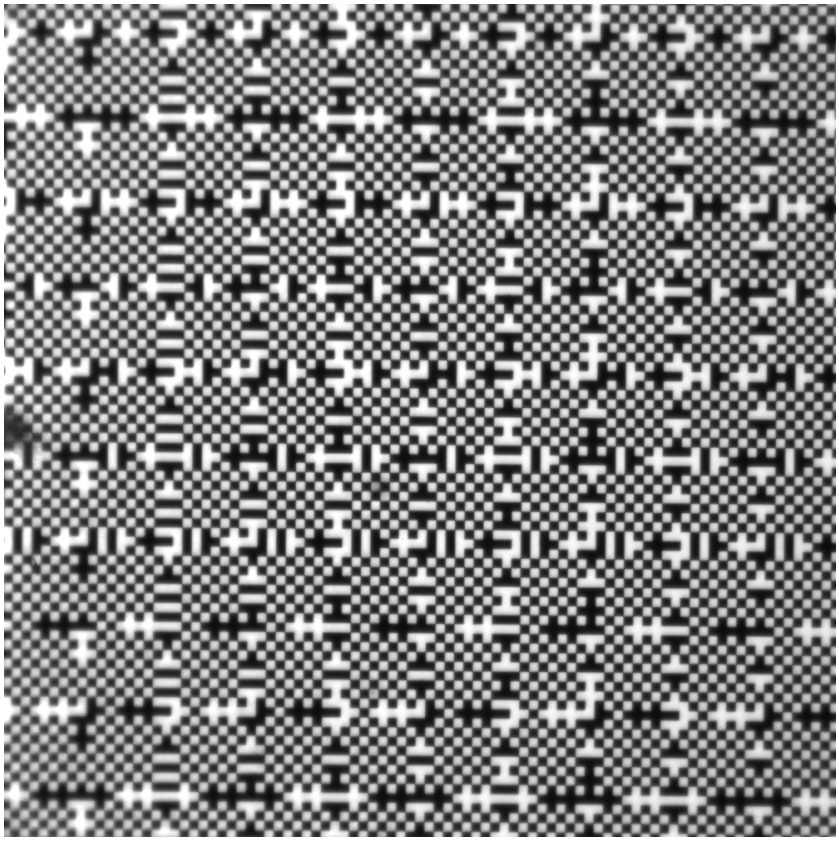


Figure 26. Typical RasMic image when equipped with the Pike image sensor. This pixel sensor covers an area of 7.4 mm x 7.4 mm, including 1000 x 1000 pixels of 7.4 μm sq. One image contains some hundred vertical and horizontal contours with a total contour length of 740 mm each.

how these systematic, but possible stable variations affect the overall RMS of the image position coordinate. There are strong indications in our recent data that mechanical movements (vibrations, or seismic acoustic noise) are limiting the measured spatial resolution. Now, mechanically insulated systems in the form of multiple-mass pendula, placed in vacuum, are under study.

4 RasDif

For the alignment of active beam elements of future CLIC and ILC linear colliders, systems with a span between the image sensor and the mask of a CCD-Rasnik system larger than 20 m are required. The diameter and focal length of the required lens become impractically large, and its replacement by a *zone plate* (further called diffraction plate) was considered. In a RasDif system, the back-illuminated coded mask is replaced by a monochromatic point-like light source generating spherical waves. In figure 2c, the diffraction plate and the resulting diffraction pattern as recorded by the sensor is shown as was used for a test system at CERN.

A diffraction pattern, defined by the geometry of the zone plate, appears on the image sensor, and the position of this pattern on the sensor is a measure for the alignment of three points, as in the CCD-Rasnik system. The image data provides (changes in) the X and Y alignment, with less

2019 JINST 14 P08010

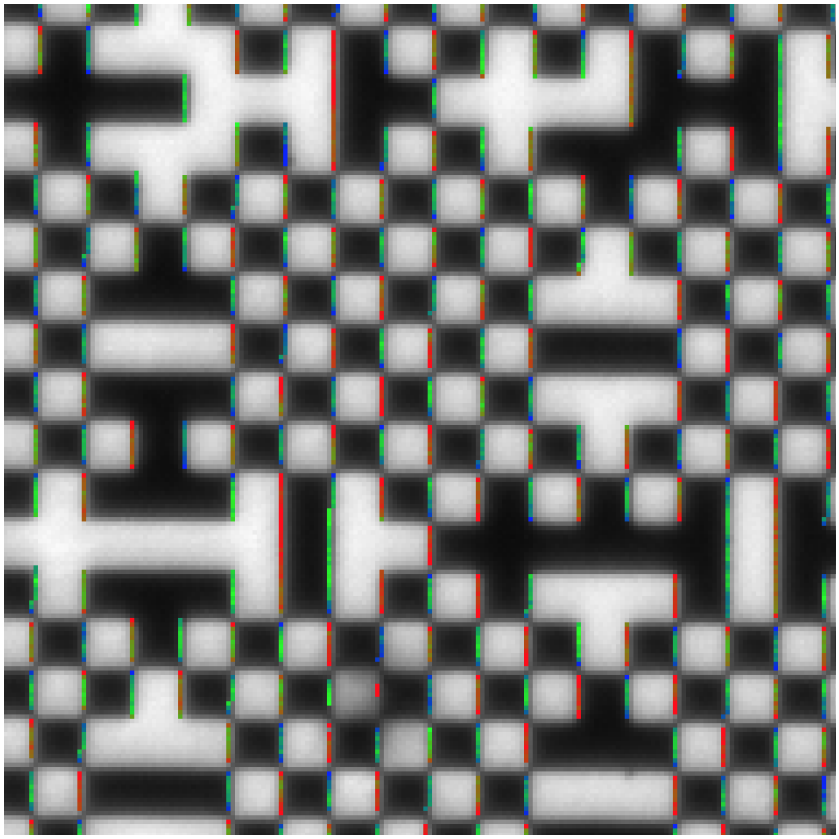


Figure 27. Selected area of RasMic image showing colour-indicated local residual of the fitted contour. The colour gradient red-green-blue corresponds to a residual shift from -3.2 to +3.2 μm. Going from left to right, white-black contours tend towards red, black-white contours tend towards blue.

precision than CCD-Rasnik since a RasDif image contains less information in terms of black-white transitions. A RasDif image varies with the object and image distances s_o and s_i, but the correlation is too complex to provide a useful 3rd (Z) coordinate.

4.1 The light source for RasDif systems

The light source should produce monochromatic spherical waves at the diffraction plate, with the position of the light source in the center of the spheres. The coherence length of the light source should be larger than the span of the system, and therefore lasers are used. The spread in λ of HeNe lasers is sufficiently small: their coherence length exceeds tens of km, and the drift in λ with time is not noticeable, in this system. The divergence of the beam should be such that spherical waves arrive at the diffraction plate, with their apparent point of creation close to the source (Z) position. If this *virtual point of departure* does not correspond to the laser position well enough, a lever arm is created that results in angular-to-translational motion. This results in a systematic error if the laser is rotated around the X or Y axis [36].

For the seismic sensor setup at CERN [37], two stabilized HeNe lasers (Melles Griot 05-LHR-925) were chosen. The position of the exiting beam, in X and Y, relative to the laser tube varied by a micron, exceeding the expected sagitta resolution of 100 nm. This was solved by coupling

2019 JINST 14 P08010

the laser with an optical fibre with out-couplers (Newport M-20X) at the other end. In this set-up, a variation of the beam position at the fibre input only causes a variation in the beam amplitude, not affecting the measurement of alignment. In addition, the beam divergence could be tuned by carefully adjusting the distance between the fibre tip and the out-coupler.

Recently low-cost Diode Pumped Solid State (DPSS) lasers became available: their coherence length exceeds 140 m, but their lifetime and power stability do limit their application.

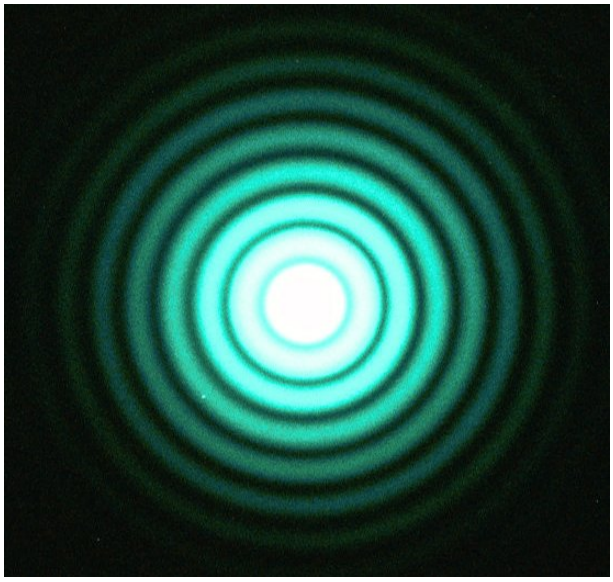


Figure 28. Image of RasDif system if a simple round hole is applied as zone (diffraction) plate.

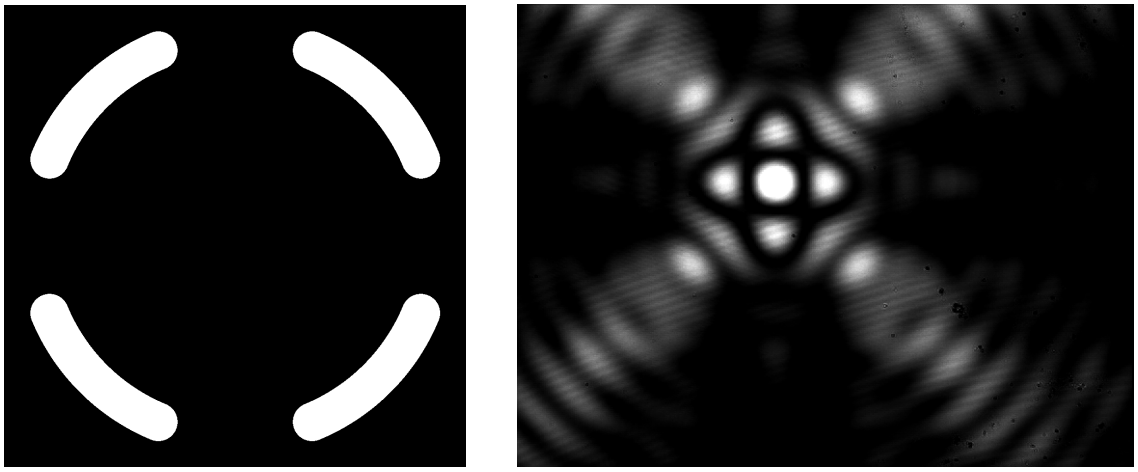
4.2 Diffraction plate for RasDif systems

The performance of RasDif systems is determined by the precision of measuring the (change of the) position of the diffraction pattern on the sensor. Ultimately, this precision is determined by the noise in the image pixels, and can be estimated with the CRLB [36]. In the first RasDif systems, the pattern of transparency of the diffraction plate was a circle. The plate had a simple hole: the diffraction pattern on the sensor is shown in figure 28. The intensity of the light is high in the centre, and decreases rapidly away from the centre. The ring-shaped pattern (a) in figure 29a results in a more evenly distribution of light and has a lower (better) CRLB, as shown in figure 29b. While the ChessField pattern is optimal for the CCD-Rasnik system, the optimal pattern for the diffraction plate is not known at present. Where the range of a CCD-Rasnik system is determined only by the size of the coded mask, the range of operation of a RasDif system depends on the capability of pattern recognition in areas away from the pattern's centre. This may impose more stringent boundary conditions to the design of a diffraction plate.

4.3 RasDif image analysis

For RasDif, an image analysis routine was developed in which the first image taken acts as reference, assuming that only the position of the diffraction pattern on the image sensor may change, while the pattern itself remains explicitly constant. For this, a 2D Fourier conversion of the image is efficient,

2019 JINST 14 P08010



(a) (b)

Figure 29. Panel (a) shows the improved diffraction plate design. Panel (b) shows the RasDif image resulting from this diffraction plate.

since the changes in position appear straight from the change in phase, associated with significant amplitudes in the frequency domain. This algorithm appeared to be very efficient, and it was later applied in the FOAM analysis for the ChessField images of the CCD-Rasnik system [15].

An image from a RasDif system has only two output values: the shifts of the image in X and Y with respect to a *first* image taken. Assuming that there is no variation in the image itself, the phase output of a 2D Fourier conversion includes these two parameters. The pattern of the diffraction plate should be optimised to maximise the values in the frequency domain.

4.4 Results of RasDif

A dual RasDif system with a span of 140 m was set up in the Transfer Tunnel TT1 site at CERN. Both systems were placed in a vacuum tube (dia 159 mm, pressure 10^{-6} mbar). In parallel, a stretched wire alignment system was under test, and by interconnecting the two systems mechanically the output data could be compared [30, 38]. With its very large span, the RasClic system can measure low-frequency seismic motions such as tidal waves [39, 40]. Data from this system is shown in figure 30. This RasClic setup was accidentally destroyed before high-quality data was taken [41], so the best resolution in sagitta obtained is 0.3 μm. The CRLB calculation indicated values in the order of 1 nm [36]. The discrepancy with real data may be explained by image pixel noise or by mechanical vibrations such as seismic motions.

5 Practical Rasnik systems

5.1 Variations in the optical transmission in Air

The precision of all Rasnik systems is limited by variations in the refraction index of the air along the path of the light, affecting the direction of the rays of light that eventually projects the image or pattern onto the image sensor. This causes spread in Rasnik measurements, and possibly a

2019 JINST 14 P08010

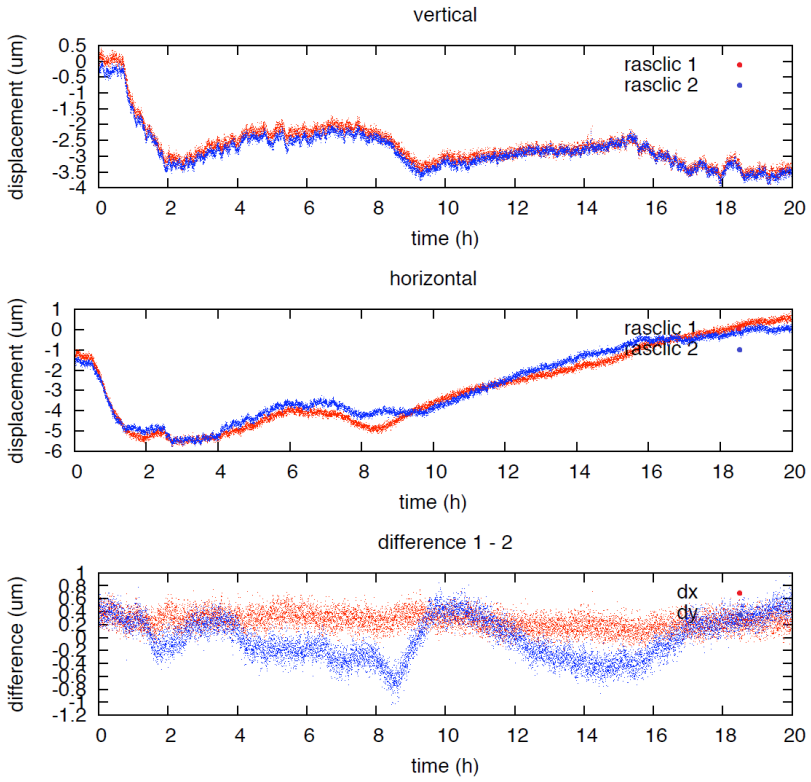


Figure 30. The X- and Y sagitta data of the dual RasClic system. The spatial resolution per measurement equals 0.3 μm RMS. The difference between the two horizontal values may be due to a rotation of the combined zone plate around the optical axis.

systematic shift. This problem can be eliminated by placing the system in vacuum. Other possible ways to reduce the effect are studied below.

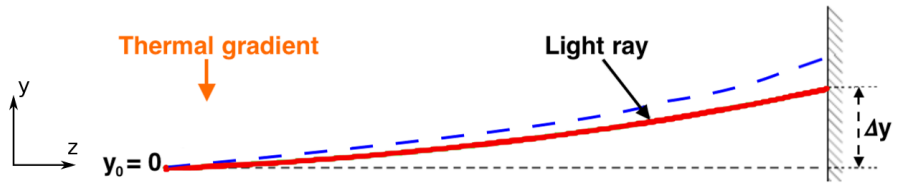


Figure 31. Refractive bending of a red light ray travelling in the z direction subject to a thermal gradient. A dashed blue ray is also shown to show dispersion for shorter wavelengths of light.

In a thermal or density gradient spanned medium, light travels along a region where the refractive index changes continuously in the direction perpendicular to the ray path. This gives for the gradient shown in figure 31 an offset in the opposite direction of the gradient. If only deviations in the vertical (y) direction are considered, the index of refraction can be written as [42]

$$n(y) = n_0(1 - \epsilon y), \tag{5.1}$$

where $n_0\epsilon$ represents the gradient in the refractive index. This gradient finds its origin in thermal gradients and, to less extent, pressure gradients.

2019 JINST 14 P08010

The differential equation of a light ray is $d/ds(n\, d\vec{r}/ds) = \nabla n$ in vector form, where s is the path length and $\vec{r}$ the amount of bending. In the case where small deviations in the y-direction are considered, $s \approx z$ to good approximation. We can approximate $|\vec{r}| \approx y$, and the differential equation yields $d/dz(n\, dy/dz) \approx -n_0\epsilon$. For a paraxial ray starting on the z axis with $y = y_0$, where $n = n_0$, the solution to this differential equation is

$$y \approx y_0 - \epsilon\frac{z^2}{2}. \tag{5.2}$$

The factor ϵ can be found using the Lorenz-Lorentz formula [43]

$$\frac{n^2-1}{n^2+1} = \frac{4\pi}{3}N\alpha_m, \tag{5.3}$$

where N represents the number of molecules per unit volume, and α_m the mean polarizability. The average value for the mean molecular polarizability α_m of air is $2.118 \pm 0.091 \times 10^{-29} m^3$ (95% confidence) [44]. Substituting the molar refractivity $A = (4\pi/3)N_\mathrm{A}\alpha_m$, with N_A Avogadro's number, yields $A = (RT/p)(n^2-1)/(n^2+1)$. For gasses $n^2 \approx 1$ and rearrangement of equation (5.3) yields

$$n - 1 = \frac{3Ap}{RT(n+1)} \approx \frac{3}{2}\frac{A}{R}\frac{p(y)}{T(y)}, \tag{5.4}$$

where R represents the universal gas constant, p the pressure and T the temperature of the gas.The last step in equation (5.4) is made by noting that for gasses $n + 1 \approx 2$.

The pressure and temperature, both dependent on the vertical coordinate, can be expressed to first order as $p(y) \approx p_0 + p'y$ and $T(y) \approx T_0 + T'y$, where p_0 and T_0 represent the conditions at the axis $y = y_0$ and p' and T' represent the gradients in the vertical direction. The ratio $p(y)/T(y)$ can be Taylor expanded to

$$\begin{aligned}\frac{p(y)}{T(y)} &\approx \frac{p_0}{T_0} + y\left[\frac{(T_0+T'y)p' - (p_0+p'y)T'}{(T_0+T'y)^2}\right]_{y=y_0=0} \\ &\approx \frac{p_0}{T_0} + \frac{p'y}{T_0} - \frac{p_0T'y}{T_0^2}.\end{aligned} \tag{5.5}$$

Substituting equation (5.1) and equation (5.5) into equation (5.4) and noting that $n_0 - 1 = (3A/2R)(p_0/T_0)$ yields

$$\begin{aligned}n_0\epsilon y &\approx \frac{3}{2}\frac{A}{R}\left(\frac{p_0T'y}{T_0^2} - \frac{p'y}{T_0}\right) \\ \epsilon &= \frac{1}{n_0}\frac{3}{2}\frac{A}{R}\left(\frac{p_0T'}{T_0^2} - \frac{p'}{T_0}\right).\end{aligned} \tag{5.6}$$

The factor $3A/2R$, can be determined using equation (5.4) using known values for p_0 and T_0. The values of refractive index for air n_0 is dependent on the colour (λ) of the light used [45] and is shown in figure 32 as

$$\begin{aligned}n_0 - 1 = 8.3425\times10^{-5} &+ 2.4061\times10^{-2}\left(130 - \frac{1}{\lambda^2}\right)^{-1} \\ &+ 1.5998\times10^{-4}\left(39 - \frac{1}{\lambda^2}\right)^{-1}.\end{aligned} \tag{5.7}$$

2019 JINST 14 P08010

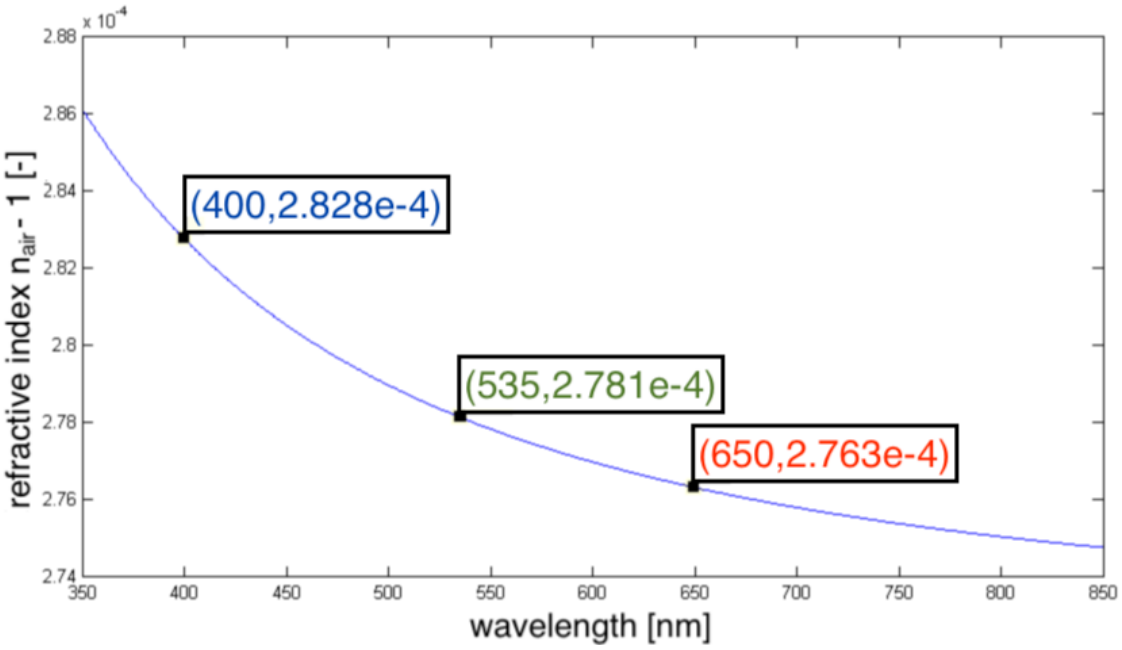


Figure 32. Refractive index of air versus wavelength of light at standard pressure and temperature.

Combining equation (5.2) and equation (5.6) yields the bending y over a distance z subject to temperature and pressure gradients as

$$y \approx y_0 - \frac{1}{n_0}\frac{3}{2}\frac{A}{R}\left(\frac{p_0 T'}{T_0^2} - \frac{p'}{T_0}\right)\frac{z^2}{2}. \tag{5.8}$$

As an example, for green light ($\lambda = 535$ nm) $n_0 - 1 = 2.781 \times 10^{-4}$. Therefore, at $p_0 = 1.0132$ bar and $T_0 = 293$ K the factor $3A/2R$ is equal to 8.042×10^{-7} K/Pa. The pressure gradient for air is typically $-1.15 \times 10^{-4} p_0$/m and its pre-factor in equation (5.6) is typically smaller than the thermal gradient pre-factor, i.e. $1/T_0 < p_0/T_0^2$. It is consequentially a small contributor to refractive bending. Using equation (5.2), ignoring the influence of the pressure gradient, a 1 K/m gradient would bend a beam 7.59 μm over a 4 m path in atmospheric pressure. Sub-micron bending of light over 4 m would therefore require a gradient below the value of 0.132 K/m.

Changes in the local conditions can alter the position of the image, while the elements that are aligned by the system are not moving. The pre-factor of the thermal gradient in equation (5.6) shows that also the pressure can be lowered or around the entire optical alignment system to decrease the refractive bending. This would mean the system would need to be in a separate volume in order to establish the wanted conditions.

Thermal shielding of a light path is possible by means of two coaxial aluminium tubes with insulating material in between, without a forced air flow. A reduction of a typical lab thermal gradient of 1 K/m in the vertical direction to 0.132 K/m by this method was modelled to be possible [36]. Variations in the gradient larger than this value are eliminated and sub-micron refractive bending in a thermally shielded light path of several meters can therefore be expected.

2019 JINST 14 P08010

Local variations in the air density, namely convection, can be reduced by forced mixing the air using fans. This will reduce Rasnik's systematic error, but at the cost of an increase of random fluctuations.

The error due to the non-straightness of light beam propagation can be corrected, in principle, by recording the difference between measurements with different wave length, as shown in figure 31: blue light is affected slightly more than red light.

5.2 Mechanical aspects of alignment: mechanical references and absolute calibration

In 1980, alignment could be measured in situ by means of the Taylor Hobson telescope [2], and by measuring the distances between the three objects and a stretched reference wire, respectively. By means of the then newly available photo diode, the alignment of three objects could be monitored by the 4QD Rasnik system. The L3 muon spectrometer was designed to be stable and precise close to the specifications imposed by the alignment requirements. It has been considered to correct the chamber alignment by means of a feed-back system, activating actuators attached to the middle chambers, but the sufficient linearity and range of the 4QD systems made this unnecessary. The alignment of the L3 muon spectrometer could be verified by taking straight track muon data with magnetic field off [1].

For the ATLAS muon spectrometer, the ultimate design was based on the 'floppy chamber' concept, in which large detector deformations were acceptable, but accurately recorded, providing large, but high-precision corrections in the muon track position data. This was possible thanks the CCD-Rasnik systems with their arbitrarily large range and precision much better than the equivalent spatial resolution of the muon chambers. Here, too, the alignment has been fully confirmed by the analysis of straight muon tracks, taken with zero magnetic field [6].

The only alternative for Rasnik systems is the Stretched Wire system [46]. In 1986, the alignment obtained with the 4QD system was compared to a 6 m long stretched, gold plated tungsten wire, dia 40 μm. The alignment of a 4QD system was transferred to gold plated pins. After positioning the pins by the contact-no contact criterion, the two systems agreed within 2 μm (not published). In 2015, a set up was realised at CERN in which both Stretched Wire sensor units and CCD-Rasnik units could be calibrated. There was an agreement to better than one μm, both in the X and Y direction [47].

System integration. In complex mechanical structures, the distribution of alignment systems (each formed by three components) needs to be carefully analysed in order to optimise their efficiency. A component of a system can act with its same functionality in another system: this is demonstrated in the InPlane systems of the ATLAS Muon chambers, depicted in figure 15. Each image sensor serves two optical axes: the light source of the non-active one is off. Two optical axes depart from each light source. A lens should be mounted to be perpendicular to the optical axis: this component can't be used to serve more systems. A multi-system component, or two components mounted together form a *node*: a multiple of nodes form a 3D network of position-sensitive detectors, capable to provide the required 3D alignment data. It should be stated that deformations such as shear stress and a global change of scale are not recorded by three-point alignment systems. The ATLAS Projective Alignment systems, shown in figure 14 is another example of applying alignment systems in an efficient way: here, the straightness of muon track is directly compared to

2019 JINST 14 P08010

the measured alignment of the track measuring detectors. The three components of a system should be fixed kinematically to the object structure. With three bolts and three spacers, a mask or a sensor can be mounted in its correct (focused) position after selecting the height of the spacers, adapting to the focal length of the lens. When fixing the coded mask, its orientation should be well-controlled since the data from a CCD-Rasnik system is given with respect to the mask coordinate system.

As an alternative, pre-calibrated components can be fixed accurately on mechanical references fixed on the object. For this, three-ball mounts are often applied [32].

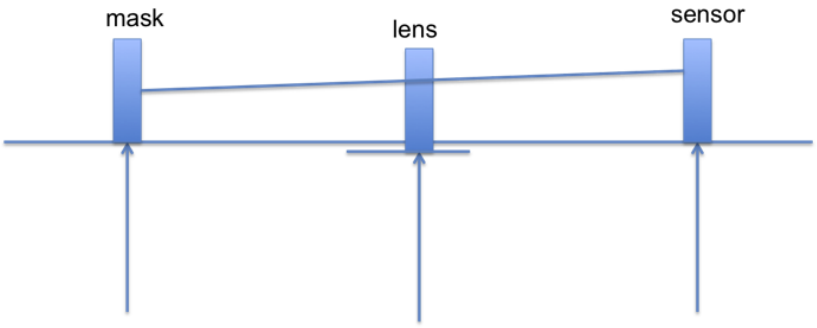


Figure 33. Principle of absolute calibration of a Rasnik system. The components are fixed on gauge blocks with well-defined and identical height. The three blocks are placed on an Alignment Station which defines their vertical position in a stable and accurate way.

In most applications, alignment systems can be integrated during the assembly of the device to be monitored. The Rasnik data displays changes with respect to the moment of fixing the three components onto the object to be monitored. In other cases, the alignment of three individually made objects with reference surfaces need to be monitored. Here, the data of absolute alignment of three Rasnik components is required. This is possible by mounting the components on precision gauge blocks. These blocks are placed in an Alignment Station, shown in figure 33. If all three blocks are rotated 180 ^{o} around their optical axes then the image shift on the sensor equals 4 times the deviation from perfect alignment (sagitta) of the Alignment Station. The Station could now be used to calibrate Rasnik systems with the optical components fixed on normal referenced mounts. The calibration is then known of a complete set of three components. By rotating only one of the three gauge blocks over 180 ^{o}, the position of its optical centre can be measured, and now all mounts of that particular component can be compared. After this is done for all mounts, the calibration of any set of three mounts is known. This principle has been applied for the Projective Alignment Systems for the ATLAS muon spectrometer. Re-calibration of prototype systems, after some years, gave identical calibration values within +/- 2 μm [32]. The principle of obtaining absolute calibration by applying the (high precision) alignment system itself is applicable for other alignment systems.

5.3 Radiation hardness

For the Rasnik systems, applied in the ATLAS experiment, all applied components have been verified to be sufficiently radiation hard. For ATLAS, the *RasCam* sensor was developed at Nikhef which deploy the commercially available VLSI Vision VV5430 CMOS sensor [20]. The diffusor

2019 JINST 14 P08010

was illuminated with a 3 x 3 array of LEDs (HP HSDL-4420 #11) which were selected after radiation hardness tests showing that they could withstand a thermal neutron dose of 10^{15} n/cm^2 [20].

The dominant symptom of radiation damage in LEDs is a decrease in brightness, and in CCDs or CMOS image sensors the sensor dark current will increase. These two effect exacerbate one another, because dim illumination requires a longer exposure, and a longer exposure gives dark current more time to accumulate in the image sensor. In the end-cap alignment system of the ATLAS detector, the HSDL-4400 infra-red LED was used to illuminate our Rasnik masks, and the TC255P CCD as image sensor. Both components were irradiated with gamma rays on two separate occasions. No power was applied to the components during the irradiation, so as to mimic the conditions under which the devices would operate in ATLAS, where these components have power applied for less than 1% of the time. A dose of 400 Gy had no effect upon the optical power output of the HSDL-4400 [48]. A dose of 600 Gy had no significant effect upon the TC255P image sensor. Both components were irradiated with fast neutrons in three separate experiments, with doses up to 10 x 10^{12} 1 MeV eq. n/cm^2 (10 Tn). After a dose of 10 Tn, the HSDL-4400 optical power output had dropped by 90% for the same applied current, while the TC255P dark current increased until the maximum tolerable exposure time is 10 ms. A 10-ms exposure, however, is adequate for all end-cap Rasnik instruments, even when the LEDs lose 90% of their optical output. The tolerance of the end-cap alignment system for neutron radiation is >10 Tn, and for ionizing radiation is >600 Gy. Twelve years after installation no degradation of Rasnik image quality due to radiation damage was observed. The accumulated ionizing and neutron doses are less than 10% of the tolerated doses. Signs of neutron damage in the form of bright pixels in 100-ms exposures taken with the LEDs turned off, however, is observed. By means of these bright pixels, a map is generated of accumulated neutron dose within the end-cap for use by the ATLAS background simulation group.

The ATLAS New Small Wheel alignment system uses the deep-red LXZ1-PA01 and royal-blue LXZ1-PR01 LEDs, and the ICX424E image sensor. These three components were irradiated in our own x-ray and gamma-ray test stands, which we set up for the purpose. No degradation is observed in LED optical output power after 1400 Gy. The ICX424 image sensor, however, barely provides adequate contrast after 1400 Gy [49]. After a 16-Tn neutron dose, the royal-blue LXZ1-PR01 loses less than 1% of its brightness, while the deep-red LXZ1-PA01 loses 75% of its brightness. After 16 Tn, the ICX424 dark current has increased to the point where a 10-ms exposure almost saturates the image pixels. The Rasnik masks are illuminated with royal-blue LXZ1-PR01 LEDs in all regions where the neutron dose is high. The tolerance of the resulting Rasnik instruments is dominated by the vulnerability of the ICX424 to ionizing and neutron radiation [50]. The image sensor will tolerate a combination of 260 Gy and 13 Tn, which is 3.7 times the maximum combined dose any image sensor is expected to receive during the lifetime of the New Small Wheel.

5.4 Applications

Rasnik systems have been originally applied in large quantities in the L_3 and ATLAS experiments at CERN for monitoring the displacement of a point with respect to the (imaginary) center point between two outer points. With these systems, the alignment of the inner, middle and outer muon chambers was measured, providing positional corrections of the measured sagittas of the muon tracks. In addition, Rasnik systems were applied to monitor the sag and torsion of large chambers: see table 5.

2019 JINST 14 P08010

Table 5. The applied Rasnik systems in experiments or HEP experimental sites.

Experiment	Rasnik type	reference
L3	4QD	[1]
ATLAS	CCD-Rasnik	[6]
CMS	CCD-Rasnik	[51]
LHCb	CCD-Rasnik	[52, 53]
HADES	CCD-Rasnik	[54]
CDF	CCD-Rasnik	[34]
CLIC	CCD-Rasnik and RasDif	[41]

Rasnik systems have been applied to monitor small variations in the geometry of large objects, such as bridge decks, pillars and viaducts. The quality of bridges and viaducts can be verified by applying a very large weight, and measure the left-over deformation (hysteresis) after unloading. This method reveals loss of quality of steel-concrete constructions. By fixing the lens and image pixel sensor together, forming essentially a camera, translations of a patterned object can be monitored, relative to the camera. A large image scale can be obtained by using a microscope objective as positive lens. Displacements of tens of picometers should be detectable, opening a new field of applications where now optical interferometers are applied, such as seismometers, pressure sensors, stress sensors, with cost, simplicity, and the absence of 1/f noise as advantages.

Since 2007, the Rasnik systems are commercially available [55].

6 Conclusions

Some 200 4QD-Rasnik systems were applied in the L_3 Muon Spectrometer: during twelve years of operation their precision, stability and absence of drift was demonstrated. With the newly available CCDs and CMOS pixel image sensors, the CCD Rasnik system was developed: for this, a coded mask was developed which is not challenged to date. Some 6000 CCD-Rasnik systems we installed in the ATLAS Muon Spectrometer. These are all operational, to date. The precision of these systems is limited to 0.5 - 5 μm due to density fluctuations in the ambient air. The intrinsic resolution of the sagitta in CCD-Rasnik systems is not depending of the span length of the system and reaches values below 10 nm. The linearity of systems depends only on the precision of the coded mask: values of 2 nm have been demonstrated. The very low level of drift in CCD-Rasnik systems has been demonstrated by the ATLAS Muon Spectrometer alignment systems.

Systems have been commercially applied in several civil engineering projects.

For CCD-Rasnik, an image analysis routine has been developed that has shown to be very reliable and precise in the sub-nm range. This routine process some 200 images per second on graphic cards.

Rasnik offers simple, very low-cost, fast and high-precision displacement or alignment monitoring systems. In high-energy physics, tracking detectors will be more and more precise: Rasnik

2019 JINST 14 P08010

is certainly accurate enough to cope with this, but using track data for obtaining alignment data is a proper alternative.

At present, the ultimate resolution is not known. The performance of Rasnik systems may reach pm precision if the noise of pixels of the image sensor is the limiting factor. If this precision is reached, Rasnik will compete with laser-interferometer displacement sensors because of their straightforward operation, their absence of drift, their arbitrarily large range, and their low cost. For Gravitational Wave experiments, new instruments such as seismic sensors, actuators and deformation monitors are required where Rasnik may well be applicable.

Acknowledgments

We would like to thank Hans Post, Astrid van der Horst, Gerard Faber, Hans Postema, Paul Rewiersma [†], Sandro Sman, Jos Vermeulen, Ulrich Becker, Marion White, Patricia McBride, James Branson, Jim Bensinger, Bianca Jongkind, Arnold Smeulders, Claude Guyot, Philipe Schune, Florian Bauer, Pierre-François Giraud, Helene Mainaud-Durand, Tom Touze, Rinat Faroudinat, Gerrit Brouwer, Oscar van Petten, Berend Munneke and Krista de Roo for their contributions to the Rasnik alignment system.

A Appendix: The SOAP image analysis program for ChessField coded masks

A.1 Performing the FFT

The analysis starts with finding the squares of the chess field pattern. This pattern can be described as the product of two square waves in perpendicular directions. Consequently the Fourier transform is the convolution of the Fourier transform of the square wave in 2 dimensions. A square wave can be described as

$$\begin{aligned} x_{\text{square}}(t) &= \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{\sin\left(2\pi(2k-1)ft\right)}{(2k-1)} \\ &= \frac{4}{\pi}\left(\sin(2\pi ft)+\frac{1}{3}\sin(6\pi ft)+\frac{1}{5}\sin(10\pi ft)+\cdots\right), \end{aligned}$$

so there is a peak at the base frequency f, and peaks at odd multiples of this base frequency. For the 2-dimensional pattern peaks are found at:

$$\mathcal{P}(i,j) = ((2i-1)f_x, (2j-1)f_y) \tag{A.1}$$

for all integer i, j.

Real images will always show some blur: the image observed using a finite size lens is well described by the original diffraction pattern convoluted by an Airy disk pattern, but a more practical approximation is a convolution with a Gaussian. In the Fourier domain this convolution works as a multiplication with a Gaussian, reducing the height of higher order peaks; this significantly reduces the meaning of the peak heights. The limited size of the image taken into the discrete Fourier transform gives rise to additional issues in finding the exact frequency: side bands and a limited frequency resolution.

2019 JINST 14 P08010

A limited part of the square wave looks like it is multiplied by the rectangle function. The Fourier transform of the rectangle function is:

$$\int_{-\infty}^{\infty} \mathrm{rect}(t) \cdot e^{-i2\pi f t}\, dt = \frac{\sin(\pi f)}{\pi f} = \mathrm{sinc}(f), \tag{A.2}$$

so in the Fourier spectrum this convolutes the peaks by a sinc function, which introduces many side bands falling off slowly. Besides that, as $\mathrm{sinc}(f)$ crosses 0 at $f = \pm 1$, the entries on either side of the peak have a very small value.

This makes an interpolation to find the exact peak position very unstable. To remedy this the standard procedure is to multiply the data first by a window function. On the Fourier spectrum this has the effect of a convolution with the window function's Fourier transform, typically resulting in a somewhat wider nearly Gaussian peak with much lower side lobes. Earlier versions of the Rasnik analysis would use a 100 dB Dolph-Chebyshev window, see [56]. Indeed this reduces the side bands tremendously, and makes the peaks stand out clearly. However, the exact peak position is still very difficult to calculate, as most of the signal is typically in just a few FFT bins, and the peak shape is not well-described, as seen by the brown and red points in figure 34. That issue is taken care of by padding the input signal with an equal number of 0 values. This results in twice the number of FFT bins, with a nicely shaped peak distributed over multiple entries, enabling stable calculation of the peak position and amplitude. This is shown in figure 34. Comparison of the precision obtained with different values for the attenuation of the Dolph-Chebyshev window showed that 60 dB was optimal. The performance of this window function is also compared to that of a Hann window function [57]. They are found to be very similar.

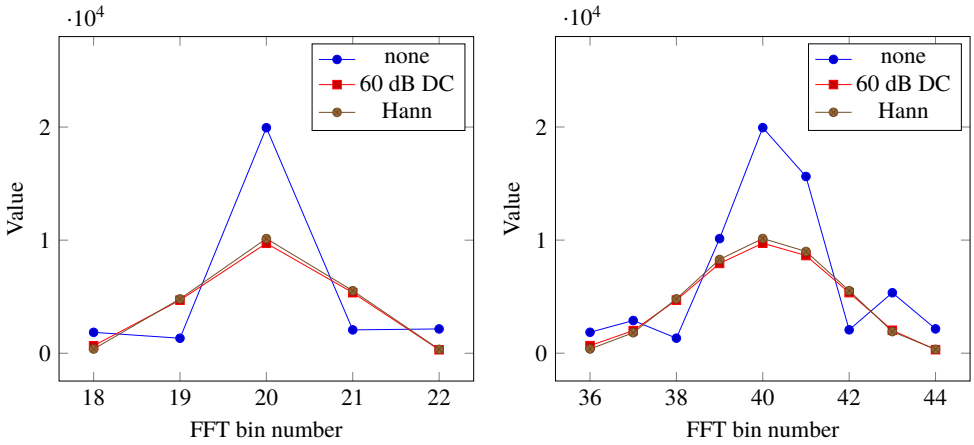
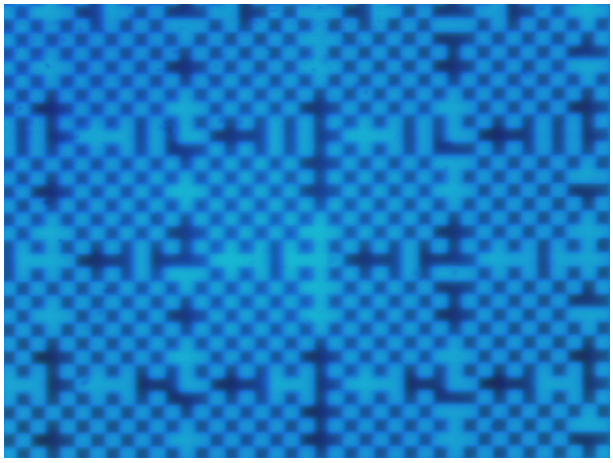


Figure 34. The effect of applying a 60 dB Dolph-Chebyshev or Hann window before the FFT; without and with padding the input signal with 0 values to double the frequency resolution.

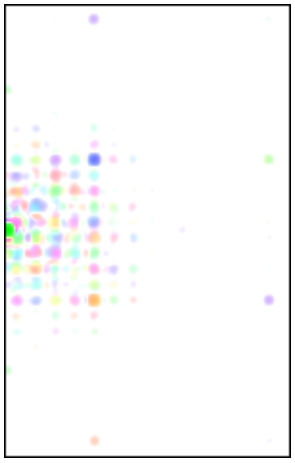
Figure 34 illustrates the effects of the padding and windowing.

Applied on an actually observed image, shown in figure 35a, the FFT diagram presented in figure 35b is obtained. Only the right half of the FFT spectrum is shown: as the input signal is real, the FFT spectrum has a Hermitian symmetry, so the other half is equivalent. On the left edge (so in the middle of the complete spectrum) one can see the green (0,0) peak representing the average value of the image. A blue (1,1) and orange (1,-1) peak correspond to the base frequency of the

2019 JINST 14 P08010



(a) Observed image.



(b) FFT plot.

Figure 35. Image from a demo system and its FFT. The h/v axis represents the h/v frequencies, one pixel per FFT bin; the origin is the green spot half height on the left. The intensity is represented on a logarithmic scale, while the colour represents the complex phase.

chess field pattern, while higher-order peaks are visible further away: near the top a lila (1,3) peak, on the right a green (3,1) and a lila (3,-1) peak, and near the bottom an orange (1,-3) peak. Also visible are many irregular peaks at lower frequencies, caused by the code squares.

A.2 Peak finding and fitting

The resulting FFT spectrum is scanned to locate the positions of the peaks, the entries with a local maximum in the absolute value, or rather its square as that is computationally much cheaper. Effectively this is equivalent to using the power spectrum. As finding the code blocks later requires a minimal number of squares in both dimensions of the picture, a low-frequency cut-off is applied, by skipping part of the array. For the image in figure 35a the algorithm would produce over a hundred peaks; a few dozen of them rejected by the low-frequency cut-off.

The peaks found in the FFT spectrum have a shape dominated by the Fourier transform of the window function. The central part of this is very much like a Gaussian, further away there are side lobes that can not be described easily. For this analysis the position and phase of the peak is required to be found. In 2 dimensions one can take a 3×3 point region, which could even be reduced by dropping the corner points, leaving a plus (+) shaped region.

Finding the position of a peak in one dimension requires an interpolation between 3 points. There are several ways to perform the interpolation, the most obvious choices being a parabolic or Gaussian fit. In this analysis a 2-dimensional Gaussian is fit to the peak, with independent peak widths in the 2 dimensions, no correlation, and no background. As the fit is performed on a very limited number of points, the χ^2 is typically well below 0.1. The actual fit is performed using the Minuit package [58], in particular the FreeHEP Java version of it.

2019 JINST 14 P08010

A.3 Peak pattern fit

These peaks are sorted by amplitude, so the primary peaks should be the first two. The above peak fitting is performed to refine the position of the peaks. In some images astigmatism seems to play a role, causing one of the primary peaks to be much lower than expected, so the algorithm does not simply take the largest two peaks as the primaries. Instead it scans the list for the best candidate pair of peaks. The first peak is required to be among the top 5, and have a height at least 0.4 times the highest; the second peak is in the top 20, having a height at least 0.02 times the highest. Both have a positive x coordinate, the one with the highest y coordinate must be the $\mathcal{P}(1,1)$ peak, the other the $\mathcal{P}(1,-1)$ peak as described in equation (A.1).

The FFT was performed on a different size in the width and height; as a result the height-over-width ratio has to be taken into account when doing calculations on peak coordinates. This could have been avoided by padding the image into a square shape before the FFT, making it a little bit slower. E.g. in the same figure padding the image to the double width and height gives space for double the number of squares, so in the FFT the peaks will then come out close to (42, +-34). If the image would be padded more vertically, so as many squares would fit in both dimensions, the peaks would appear around (42, +-42) but this would make the data array even bigger. For simplicity this ratio isn't mentioned in the expressions below. Additionally, a rotation of the mask in the image will cause a rotation of the FFT spectrum as well. This rotation angle needs to be determined as well.

The vectors $\mathbf{a} = \mathcal{P}(1,1)$ and $\mathbf{b} = \mathcal{P}(1,-1)$ describing the primary peaks are expected to be perpendicular, and have the same length. In that case the function

$$S(\mathbf{a},\mathbf{b}) = (\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2)/(det(\mathbf{b},\mathbf{a})) \tag{A.3}$$

is exactly 2. Deviations in the length ratio or the opening angle produce an increase, such that $\sqrt{(S-2)}$ is found to be a good measure of the deviation. Based on a test set of real border case images, a good maximum value was found to be 0.03. The vector sum/difference of **a** and **b** corresponds to $\mathcal{P}(2,0)$ and $\mathcal{P}(0,2)$, twice the base frequencies, and the expected positions of the first harmonic peaks follow: $\mathcal{P}(1,\pm3)$, $\mathcal{P}(3,\pm1)$ and $\mathcal{P}(3,\pm3)$. These are all combined in a single multi-peak fit to get the most precise values for the (primary) horizontal and vertical frequency and rotation.

This multi-peak fit works remarkable well on a pure chess field pattern, but when the code lines/columns are introduced a subtle problem is observed: depending on where these are in the image, the peaks are shifted in position by approximately 1% of their width. This doesn't seem much, but it is much larger than the error produced in the peak fit for the primary peaks, and it does have an impact on the precision obtained for the magnification and rotation. To account for the uncertainty this causes, we add an additional error estimate of 0.0025 FFT bin for the peak position. This value was found to be optimal when the FFT uses a Hann window. The Dolph-Chebyshev window turned out to be less well-behaved, probably due to the larger values at the edges. This was the final decision point to use the Hann window.

The frequencies obtained from the multi-peak fit represent the number of squares that fit in the width and height of the input for the FFT, which is twice the size of the image due to the padding mentioned earlier, so the observed square sizes follow, and from that the optical image magnification which should be the same in any direction. This is not necessarily true in case of pincushion or barrel distortion, which is more pronounced in the longer dimension of the image.

2019 JINST 14 P08010

A.4 Peak phases

A well known property of Fourier transforms is that a translation of the input function transforms into a linear phase shift in the result. So the complex phase of the FFT values reflect the shift of the image: see [60]. A positive and real lowest order peak means that the function looks like a cosine, so there is a maximum centered around 0. To exploit this, a pixel in the middle of the image is taken as the reference point, so positive real values of the primary peak values would indicate that this pixel is exactly in the middle of a white square of the pattern, and other values reflect the shift of the pattern (modulo 2 times the block size). To obtain the complex phase of a peak, the real and imaginary values are calculated by a parabolic interpolation between the nearest FFT bin and its 4 nearest neighbours.

Now the primary peaks provide a first estimate for these phases by which the pattern is shifted. This is used to calculate the expected phases at the harmonic peaks. The offsets are put into a 2-dimensional linear fit to refine the original estimate. The phase typically has an error in the order of 0.001, which translates to an error of 3000 times less than the square size. The precision obtained in this step dominates the precision of the X and Y coordinates of the final result.

From the peak and phase fit on the FFT we have the horizontal and vertical frequency of the pattern, the rotation angle, and the phases at the reference pixel near the middle of the image. The width of a square is the width of the (padded) image by twice the horizontal frequency from the FFT. To find the middle of the central square, the square width times the phase divided by π is to be subtracted. Same for the vertical dimension. From the position of the central square, the middles of the other squares can be calculated, the deviations from a chess field pattern analysed, giving the positions of the code rows and columns. By extracting the numbers encoded the "coarse" position is found. These calculations give results on the scale of sensor pixels; the pixel size and geometry of the optical system are needed to calculate the actual alignment result.

References

[1] P. Duinker et al., *Some methods and tools for testing and optimizing proportional wire chambers*, *Nucl. Instrum. Meth.* **A 273** (1988) 814.

[2] AMETEK/Taylor-Hobson, *The Taylor-Hobson telescope*.

[3] E.L. Bryant, *Noncontacting method for measuring angular deflection*, cited in NASA Tech Briefs, U.S. patent 4,189,234, 19 February 1980.

[4] W. Toth, D. Araujo, J. Paradiso, H. van der Graaf and H. Groenstege, private communication, Charles Stark Draper Laboratory, Cambridge, MA, U.S.A., October 1983.

[5] W.E. Toth, *Muon detector program: prototype octant construction and evaluation with production phase recommendations*, CSDL-R-1885, Draper Lab, Cambridge, MA, U.S.A., October 1987.

[6] N. van Eldik, *The ATLAS muon spectrometer: calibration and pattern recognition*, Ph.D. thesis, CERN-THESIS-2007-045, Universiteit van Amsterdam, Amsterdam, The Netherlands (2007).

[7] W. Post, *A homogeneous light source for RASNIK*, HMN internship report, Nikhef, The Netherlands, July 1992.

[8] For instance: EKSMA, *the Plano Convex* 110 *series*.

[9] J. Alberdi et al., *Enlarged linearity range of a rasnik system using an integrating sphere*, *Nucl. Instrum. Meth.* **A 408** (1998) 591.

2019 JINST 14 P08010

[10] U. Becker and J.A. Paradiso, *Dynamic compensation of small errors in drift chamber positioning using an optical CCD based system*, *Nucl. Instrum. Meth.* **196** (1982) 381.

[11] H. Dekker et al., *The RASNIK/CCD* 3*-dimensional alignment system*, in *Proceedings of the Third International Workshop on Accelerator Alignment, II*/147, CERN/Annecy, France, 28 September–1 October 1993.

[12] J.A. Paradiso and D.B. Goodwin, *Wide range precision alignment for the GEM muon system*, in *Proceedings of the Third International Workshop on Accelerator Alignment, II*/131, CERN/Annecy, France, 28 September–1 October 1993.

[13] J.A. Paradiso, *New technologies for monitoring the precision alignment of large detector systems*, *Nucl. Instrum. Meth.* **A 386** (1997) 409.

[14] H.L. Groenstege, *The coding of the mask for CCD_Rasnik*, Nikhef ETR 94-10, Nikhef, The Netherlands (1994).

[15] M. Kea, *FOAM: an image analysis routine for the ATLAS barrel muon spectrometer alignment system*, internship report, Nikhef, Amsterdam and Delft University of Technology, The Netherlands (2007).

[16] Photronics, 1 Technology Dr, Bridgend CF31 3LU, U.K.

[17] J.H. Lambert, *Photometria, sive de mensura et gradibus luminus, colorum et umbrae*, in *Lamberts Photometrie*, (1760).

[18] X.J.M. Leijtens, *The electro-optical alignment system for the L3 muon chambers*, Master's thesis, Nikhef and University of Amsterdam, The Netherlands, June 1988.

[19] ATLAS collaboration, *Search for new particles in two-jet final states in* 7 *TeV proton-proton collisions with the ATLAS detector at the LHC*, *Phys. Rev. Lett.* **105** (2010) 161801 [arXiv:1008.2461].

[20] ATLAS collaboration, *The radiation hardness of RASNIK components for the ATLAS muon barrel spectrometer*, ATL-ELEC-2002-002, CERN, Geneva, Switzerland (2002).

[21] Melles Griot Color Glass Filters RG series FSWP-800-50.0-50.0 M.

[22] *Schott BK*7 *datasheet*, https://www.schott.com/advanced_optics/us/abbe_datasheets/schott-datasheet-p-bk7.pdf.

[23] *Standard series low cost plano-convex lenses BK*7, e.g. EKSMA, Edmund Optics, Melles Griot, Thorlabs.

[24] Plano-Convex singlets with a custom-specific focal length are hard to obtain. For the ATLAS Projective Alignment systems, some 600 lenses with specific focal length with tight tolerances were made by Protvino IHEP (Rinat Faroudinat).

[25] United Lens Company, 259 Worcester St, Southbridge, MA 01550, U.S.A.

[26] Nanofilm, 10111 Sweet Valley Drive Valley View, Ohio 44125, U.S.A.

[27] J. Paradiso, *Testing and development of extended range straightness monitor systems*, GEM Collaboration Report (SSCL), GEM-TN-93-331, http://paradiso.media.mit.edu/papers/VSMCloseoutFixed.pdf, May 1994.

[28] M.J. Woudstra, *Precision of the ATLAS muon spectrometer*, Ph.D. thesis, CERN-THESIS-2003-015, Universiteit van Amsterdam, Amsterdam, The Netherlands, (2002).

[29] K. Hashemi, *Rasnik analysis*, technical report, http://alignment.hep.brandeis.edu/Devices/RASNIK/Analysis.html.

2019 JINST 14 P08010

[30] M.G. Beker, *RasClic: a long base, low frequency high precision seismograph*, M.Sc. thesis, Nikhef and Delft University of Technology-TNW, The Netherlands, 24 October 2008.

[31] R. Sandström, *Rasnik image analysis for optical alignment of the ATLAS muon spectrometer*, CERN technical report, (2009).

[32] H. van der Graaf, *The calibration of the ATLAS muon spectrometer projective alignment systems*, ATL-MUON-2004-024, CERN, Geneva, Switzerland (2004).

[33] H. Tolsma, *The honeycomb strip chamber*, Ph.D. thesis, University of Amsterdam, Amsterdam, The Netherlands, April 1996.

[34] D. Goldstein and D. Saltzberg, *The RASNIK real time relative alignment monitor for the CDF inner tracking detectors*, *Nucl. Instrum. Meth.* **A 506** (2003) 92 [hep-ex/0212025].

[35] S. Braccini et al., *Measurement of the seismic attenuation performance of the VIRGO superattenuator*, *Astropart. Phys.* **23** (2005) 557.

[36] J. van Heijningen, *Precision improvement in optical alignment systems for linear colliders*, M.Sc. thesis, Nikhef and Delft University of Technology, The Netherlands (2012).

[37] M. Kea, *RasClic: aligning accelerators, the development of a new optical alignment system*, M.Sc. thesis, Nikhef and Delft University of Technology, The Netherlands (2008).

[38] R. van der Geer, *A diffraction simulation for RasClic*, M.Sc. thesis, Nikhef and Leiden University, The Netherlands (2008).

[39] A. Seryi, *Investigations of slow motions of the SLAC LINAC tunnel*, *eConf* **C 000821** (2000) MOA11 [physics/0008195].

[40] W.B. Herrmannsfeldt, M.J. Lee, J.J. Spranza and K.R. Trigger, *Precision alignment using a system of large rectangular Fresnel lenses*, *Appl. Opt.* **7** (1968) 995.

[41] H. Manaud Durand et al., *RASCLIC: a long baseline* 3*-point alignment system for particle accelerators*, presented at the 10th *International Workshop on Accelerator Alignment*, KEK, Tsukuba, Japan February 2008.

[42] L.V. Griffith et al., *Magnetic axis alignment and the Poisson alignment reference system*, Lawrence Livermore Nat. Lab. Report UCID-21591, U.S.A., January 1989.

[43] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*, 7th ed., section 2.3.3, Cambridge University Press, Cambridge, U.K. (1999).

[44] M.J. Madsen et al., *Measuring the molecular polarizability of air*, *Amer. J. Phys.* **79** (2011) 428 [arXiv:0907.0782].

[45] K.P. Birch and M.J. Downs, *An updated Edlén equation for the refractive index of air*, *Metrologia* **30** (1993) 155.

[46] H. Manaud Durand et al., *Stretched wire offset measurements:* 40 *years of this technique at CERN*, presented at the 10th *International Workshop on Accelerator Alignment*, KEK, Tsukuba, Japan, 11–15 February 2008.

[47] N. Deelen, *Pre-alignment of CLIC: Rasclic versus WPS*, Master thesis, Nikhef and University Utrecht, The Netherlands, August 2015.

[48] K. Hashemi, *Radiation tolerance of end-cap alignment electronics*, Electronics Production Readiness Review, Harvard University, U.S.A. (2003).

[49] K. Hashemi, *Alignment electronics for the new small wheel*, nSW review, (2015).

2019 JINST 14 P08010

[50] M. MacKenzie, *ATLAS NSW alignment system radiation tolerance*, Senior thesis, Brandeis University, U.S.A. (2016).

[51] CMS collaboration, *CMS endcap muon system tests of Rasnik/straight line monitor R, ϕ, Z alignment linking using MPI transparent amorphous silicon optical beam position sensors, laser/LED Z distance sensors and R wire displacement potentiometers*, CERN-CMS-NOTE-1998-076, CERN, Geneva, Switzerland (1998).

[52] LHCb Outer Tracker Group collaboration, *Improved performance of the LHCb outer tracker in LHC run* 2, 2017 *JINST* **12** P11016 [arXiv:1708.00819].

[53] LHCb Outer Tracker Group collaboration, *Performance of the LHCb outer tracker*, 2014 *JINST* **9** P01002 [arXiv:1311.3893].

[54] J.A. Garzoń et al., *Dielectron production in heavy ions collisions: the HADES experiment*, *AIP Conf. Proc.* **495** (1999) 399.

[55] Sensiflex, http://www.sensiflex.eu/, Amsterdam, The Netherlands.

[56] P. Lynch, *The Dolph-Chebyshev window: a simple optimal filter*, *Mon. Weather Rev.* **125** (1997) 655.

[57] F.J. Harris, *On the use of windows for harmonic analysis with the discrete Fourier transform*, *Proc. IEEE* **66** (1978) 51.

[58] F. James and M. Roos, *Minuit: a system for function minimization and analysis of the parameter errors and correlations*, *Comput. Phys. Commun.* **10** (1975) 343.

[59] *FreeHEP Java Minuit*, http://java.freehep.org/freehep-jminuit/project-summary.html.

[60] See for instance: https://en.wikipedia.org/wiki/File:Fourier_transform_-_time_shifted_signal.gif.

2019 JINST 14 P08010