Design for Robustness and Cost Effectiveness: The Case of an Optical Profilometer

Antonio Baldi[1†], Paola Pedone[2], and Daniele Romano[3]

Department of Mechanical Engineering, University of Cagliari, Italy
E-mails: [1]baldi@dimeca.iris.it, [2]ppedone@dimeca.unica.it, [3]romano@dimeca.unica.it

Abstract

The paper presents the design of an optical profilometer, a device used for the reconstruction of the micro-geometry of mechanical parts in applications where high precision is needed. The design is based on Robust Design, a major methodology for quality improvement of engineering systems. Several design solutions, namely different hardware and software setups of the device, are compared in order to select a configuration realising a desired trade-off between performance and cost. The peculiarity of the design strategy is the use of a computer model of the measurement process where the physical part of the process is simulated. This allows for an extensive exploration of the design space, thus opening the way to product innovation.

1. Computer Simulation for Robust Engineering Design

Simulation models of the product and of the related processes are being used at an increasing pace in industry along the product development phase. Not only as a support tool, but sometimes also as a strategic factor to shorten time to market, reduce costs, improve product quality. However, diffusion of simulation in engineering design has not been backed by an evenly fast diffusion of methodologies for managing simulation and increasing its effectiveness. In particular we refer to the ability of the simulation to aid the designer's creativity with the view of making the simulation a standard design tool for continuous product innovation.

The paper proposes a method that fully incorporates computer simulation in the design of measurement instruments. The method is based on Robust Design (Myers and Montgomery, 1995; Nair, 1992), a key statistical methodology for quality improvement that aims at designing products that work as required regardless of production and use conditions. The rationale is to set the design parameters in such a way to track the product performance on a desired target while minimizing its random variability. The latter being a cause of definite user's dissatisfaction. This can be done by studying how the design parameters and the dis-

† Corresponding Author

turbance variables affect the mean and the variability of the performance. This information is obtained by running designed experiments on the product where the experimental factors are the design parameters (control factors) and the disturbance variables (noise factors). A limitation of Robust Design comes from the difficulty and the cost of controlling the noise factors in a physical setup. Conversely, in a computer experiment, i.e. an experiment run on a simulation model of the product, noise factors can be not only controlled, by assigning to them a desired value at each code execution, but even treated as random variables, just as they are in the real system. This significantly extends the potential of the methodology.

Another interesting aspect is that modern systems for measurement and diagnosis are characterized by an intensive use of computer processing of the information. In this way they become more complex, but also more prone to be designed (numerical processing involves a huge amount of design parameters) and to be simulated on the computer (numerical processing itself is a computer code). Interestingly, the design of such systems is naturally a robust design problem. In fact, the two main quality characteristics of a measurement, lack of bias and repeatability, are intimately related to the mean and the dispersion of the measurement result respectively. From the above it is reasonable to forecast that measurement and diagnosis systems will become a major application field for Robust Design and computer simulation. The methodology we propose exploits the synergy between Robust Design and computer simulation. Let a measurement process be represented by the cascade of hardware and software operations applied respectively to analog and digital signals carrying information on a measurand (Figure 1).

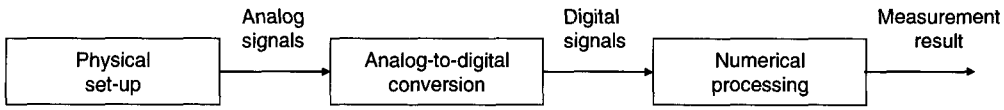


Figure 1. Scheme of a measurement process involving hardware and software treatment

If the analog treatment operated by the physical part of the process can be simulated, the whole process is reproducible on the computer. Hence the robust design of the measurement process can be entirely performed on the computer thus allowing for an extensive exploration of design options at a negligible cost. The method has the potential to deliver a design solution being robust and innovative. Of course the quality of solution critically depends on how accurate the simulation of the hardware operations is. At this regard, a difficult task is the correct simulation of the physical sources of random variation of the signals.

In the next section we introduce the optical profilometer and its measurement process. Then the set of design solutions to be analyzed is described. Section 4 presents the computer experiment adopted for the application of Robust Design. The selection of the best design solutions is shown in section 5. A final discussion concludes the paper.

2. The Optical Profilometer

The optical profilometer is a device used for the reconstruction of the micro-geometry of mechanical parts in applications where high precision is needed. It is based on the physical mechanism of white-light interferometry (WLI).

WLI mainly offers two advantages over the conventional interferometric techniques which use monochromatic sources. First, WLI has a virtually unlimited unambiguous range, whereas conventional interferometric techniques are usually limited to no more than half a wavelength. Second, it allows to perform an optical sectioning of the object being observed due to the short coherence length of the source.

A typical WLI system based on the Linnik interferometer is shown in Figure 2a. It is basically composed of the following components: a white light source (WLS), a beam splitter (BS), a reference mirror (M) moved by a piezo-electric translator (PZT), a testpiece (T), a video camera (CCD) and a displacement sensor (DS). White light, emitted from the source, is split in two beams going to the mirror and the testpiece respectively. After reflection the beams join again producing interference fringes (alternating light and dark bands) whose intensity is a function of the difference of the optical paths of the beams.

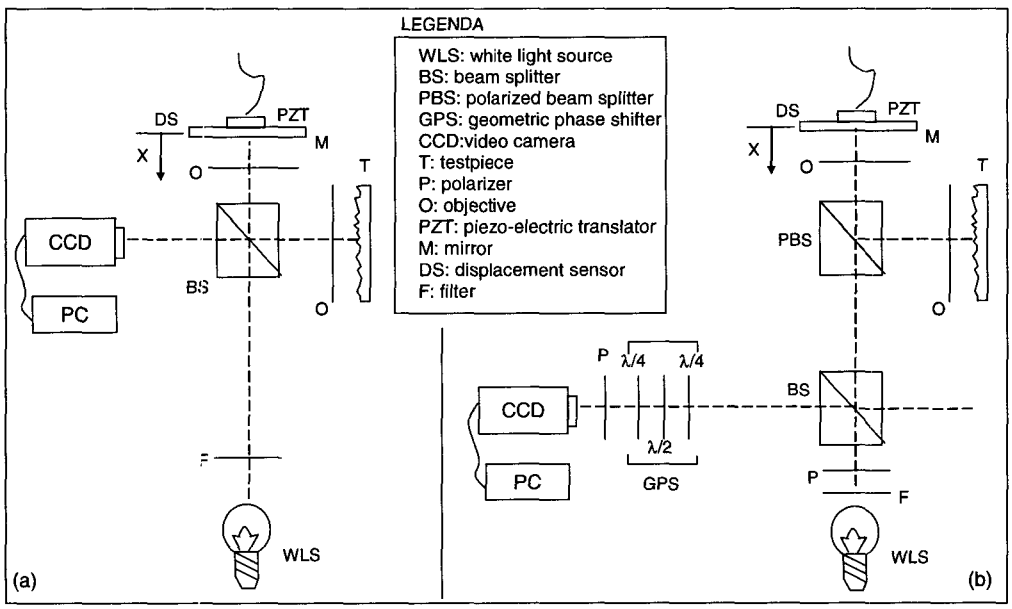


Figure 2. White-light interferometer: left, basic setup. Right: setup with phase shifters

By displacing the mirror along the x-axis, a phase shift, due to the change of the optical path difference of the two beams, is introduced. In such a way the condition of maximum

interference (path difference equals zero) is realized in different areas of the testpiece until the whole surface profile is measured. The outcoming light intensity can be mathematically formulated as:

$$I = I_1 + I_2 + \gamma(x)\cos\left(\frac{2x\pi}{\lambda} + \phi_0\right) \tag{1}$$

where the first two terms I_1 and I_2 are the intensities of the two beams acting independently and the third term represents the interference term. In this latter $\gamma(x)$ is the fringe visibility which corresponds to the envelope of the interference fringes and it is related to the spectral profile of the white light, x is the optical path difference, λ the carrier wavelength and ϕ_0 the beam initial phase.

For a rectangular spectrum the fringe visibility follows a sinc^2 pattern:

$$\gamma(x) = \gamma_0 \ \mathrm{sinc}^2\left(\frac{2x\pi}{L}\right) \text{ where } \quad \mathrm{sinc}(t)\left(\frac{\sin t}{t}\right) \tag{2}$$

where γ_0 is the maximum intensity modulation and L the coherence length of the source.

Figure 3 shows an interferogram representing the variation in intensity at a given point in the image as a function of x, the optical path difference; the dotted line is the fringe envelope.

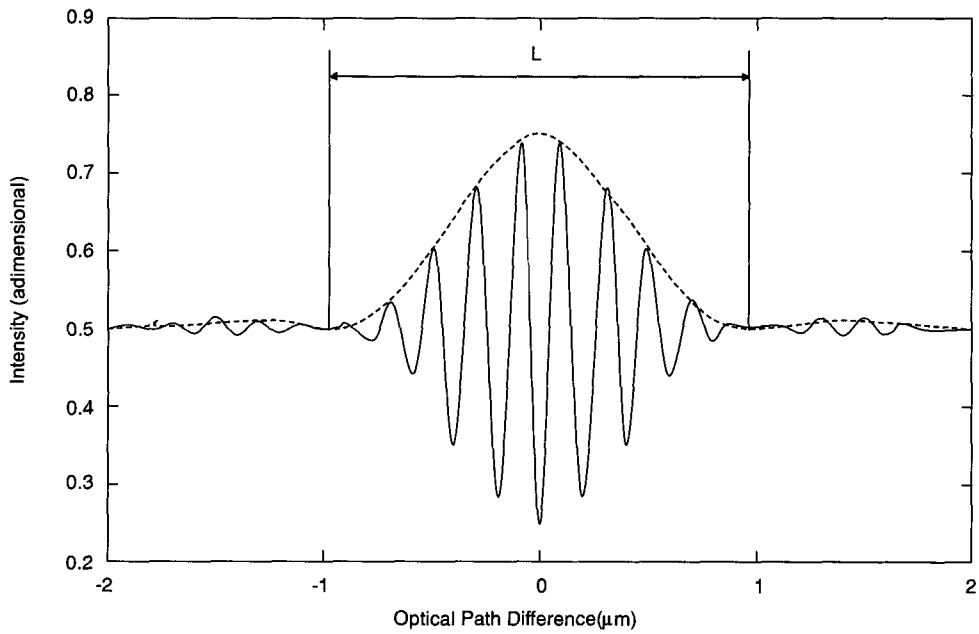


Figure 3. Typical intensity signal (interferogram) at one pixel of the image. The modulation signal is also represented (dotted line)

The evaluation of the interferogram at each pixel of the image provides the information on the surface profile. In fact the abscissa of the maximum modulation, x_M, corresponding to a zero difference of the optical path, is exactly the surface height (relative to the mirror initial location) at the point of the testpiece related to a given pixel in the image.

2.1 Estimation of the Abscissa of Maximum Modulation

Different WLI profilometers are characterized by the method for estimating the abscissa of maximum modulation.

Two are the principal approaches: the first utilizes the intensity function (solid line in Figure 3), that can directly be obtained by using the basic WLI system (Figure 2a). In the second the intensity signal is further processed to obtain the fringe modulation function (dotted line in Figure 3).

To this aim the so called "temporal phase-shifting methods" are used. They introduce an additional, known, phase-shift (a) in the intensity function:

$$I(x) = I_0 + + \gamma(x)\cos\left(\frac{2x\pi}{\lambda} + \phi_0 + a\right) \tag{3}$$

where $I_0 = I_1 + I_2$ is the mean intensity. Since expression (3) involves three unknown, namely I_0, γ_0 and ϕ_0, at least three values of intensity (related to different a values) are needed in order to obtain the corresponding analytical solution. Several algorithms have been developed and reported in literature for estimating the intensity modulation; amongst these two of the most known are considered in this study (Robinson and Reid, 1993; De Groot, 1995).

The phase shift can be mainly induced either by using only a piezo-electric translator which moves the reference mirror or the testpiece, or by adding a geometric phase-shifter (GPS), a system composed of a number of optical components (Figure 2b) (Roy *et al.*, 2001; Dorrìo and Fernandez, 1999).

3. The Design Solutions

Figure 4 schematically shows the three main design solutions of the profilometer. Each solution has one hardware setup and two software setups. The hardware setup includes the basic configuration (Figure 2a) with (solution 3) or without (solutions 1 and 2) the GPS device.

The "Pre-Processing" section contains algorithms to extract intensity modulation from intensity signal. Two of them refer to the hardware setups including GPS (3), while the other is an ad-hoc algorithm developed for configuration 2. Note that there is no pre-processing

step for configuration 1.

In the "Post-processing" section two algorithms are used to estimate the abscissa of maximum modulation: the "weighted centroid" when working with the intensity signal and a non linear curve fitting when working with the modulation signal.

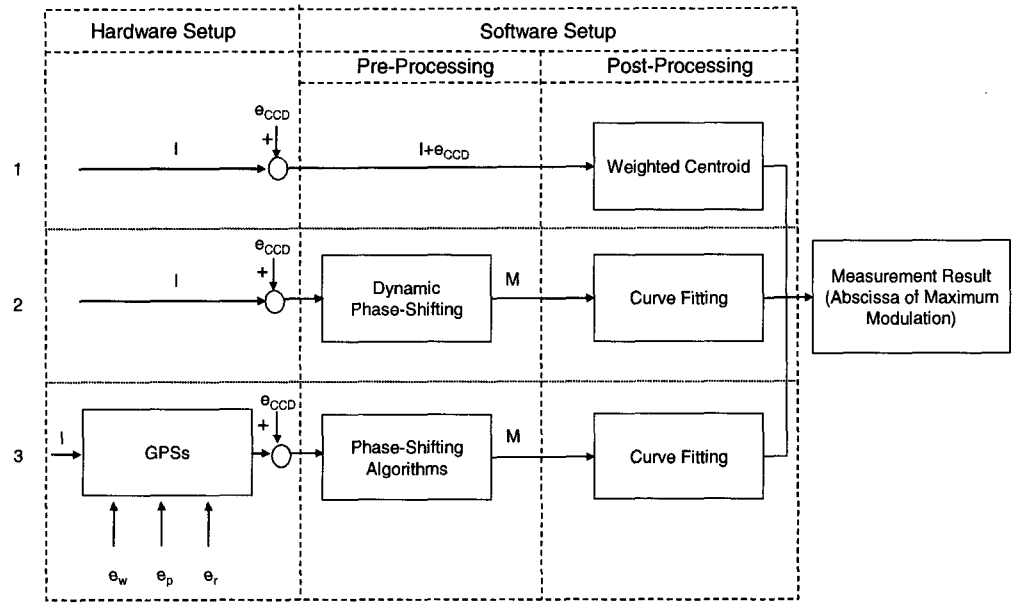


Figure 4. Simplified schemes of the design solutions. The common components are omitted

3.1 The Hardware Setups: The Geometric Phase-Shifters

GPS is mainly composed of retardation plates and polarizers. Two kinds of GPS are considered in the following. The retardation plates produce a phase shift between the beams thanks to their peculiar optical proprieties. They are called "quarter wave plates" ($\lambda/4$) if the thickness of the plate is such that the phase difference is just a quarter of the wavelength; similarly half wave plates ($\lambda/2$) if they induce a half wavelength phase difference.

A very important feature of this kind of components is that the retardation (ϕ) is strongly dependent on the wavelength value (λ) of the incident beam. In fact the birefringency property of the material of the laminae induce a shift proportional to their thickness. The "nominal" phase shift is obtained only when the beam crosses the laminae at the wavelength for which they have been manufactured.

There are additional sources of error represented by the component manufacturing inaccuracies such as the thickness error (e_w) and parallelism error (e_p), and wavelength.

The first GPS configuration is composed of: a quarter wave plate ($\lambda/4$), a rotating half wave plate ($\lambda/2$) and another quarter wave plate ($\lambda/4$). The second GPS has a simplified

structure with only one retardation plate and one rotating polarizer.

The fixed rotation of the half wave plate in the first GPS, or the polarizer in the second one, is controlled via a stepper-motor and affected by a random error (e_r). If the plate is rotated by θ degree, the out-coming beam acquires a phase-shift respectively equal to 4θ in the first GPS and 2θ in the second one.

Since the phase-shift depends on the wavelength of the source, two different optical filters have been considered. Note that a larger bandwidth implies a smaller coherence length on the one hand (making easier identifying the peak of the envelope), a larger phase shift error of GPS on the other hand.

3.2 The Software Setups

3.2.1 Phase-Shifting Algorithms

Although a minimum of three values of intensity data at a single point in the interferogram are needed to solve equation (3), many algorithms with more than three values of intensity data have been proposed in the literature to compute the phase ϕ_0 and the intensity modulation γ_0.

The main reason for increasing the number of data frames considered is to make the algorithm more robust to errors (miscalibration and nonlinearity behaviour of the phase-shifter) by performing a least square fit to a sinusoidal function.

According to this, two well-known algorithms with four- and five-rotations have been selected in this paper.

Configuration 2 requires an ad-hoc algorithm to take into account the dynamic variation of modulation during the translation of the mirror (Baldi *et al.*, 2003).

3.2.2 Estimation of the Abscissa of the Maximum Modulation

The "weighted centroid" is a simple and computationally fast algorithm dealing with the intensity signal. Abscissae of the sampled intensity are "weighted" by the absolute difference between each sampled intensity value I_i and the mean value I_0:

$$x_M = \frac{\sum_k | I_k - I_o | x_k}{\sum_k | I_k - I_o |} \tag{4}$$

where x_k represents the mirror position at the sampled point k.

Since this algorithm is strongly influenced by the tails where x_k is far from x_M, abscissae with small weights are filtered out.

The curve fitting algorithm estimates the abscissa of the maximum of three functions (gaussian, parabolic and $sinc^2$ function) fitted to the modulation points using a non linear least squares procedure.

4. Computer Experiments

Three different computer experiments are run, one for each solution scheme. Table 1 shows the experimental factors, classified according to their type (control, internal and external noise) and the design solution they belong to. Control factors have fixed levels whereas levels of noise factors are randomized according to their distribution. Hence the experiments follow the crossed-array procedure. However, differently from the conventional Taguchi's crossed-array approach, where a few fixed levels are adopted in the outer experiment involving noise factors, here the outer array is fully randomized, mimicking the physical "in process" conditions. This option is interesting to consider when the computer model is not that expensive, in terms of setup and computing time. The advantage of this procedure is its simplicity: sample mean $\overline{x_i}$ and sample variance s_i^2, computed at the i-th treatment, are direct estimates of "in process" mean and variance of the response conditional on that treatment.

Four control factors are common to all the setups. One is the sampling step of the mirror; note that larger steps imply a faster acquisition process at the price of collecting in the region of interest less information. The second is the standard deviation of the random error (e_x) in the x-coordinate of the sampled points; this error is due to imperfect displacement of the reference mirror as caused by errors on the piezo-electric translator (this factor is introduced to allow for the choice of PZTs of different quality). Similarly, another control factor is the standard deviation of the random error (e_{CCD}) due to the video acquisition system. The last control factor is the filter bandwidth. Random errors e_x and e_{CCD} are internal noise factors that are randomized in the experiments according to Gaussian distributions. Two external noise factors are common to all the setups: the scale, which depends on the reflectivity of the testpiece and may change from one point to another on the surface, and the knot centering, which accounts for the not symmetric sampling (around the maximum point) of the intensity envelope. Since their values are not under control during the real measurement process, they are randomized in the simulated process. The distribution parameters are estimated basing on theoretical considerations and/or experimental tests.

The last six factors concern the configuration 3 (Figure 3): e_r, e_w, e_p are the internal noise factors related to the GPS optical system, while phase-shifting algorithm, the type of GPS and fitting function are control factors. The latter is also used in configuration 2.

The computer model works differently depending on the hardware setup and the type of signal considered. If the setup includes the GPS, it generates directly the modulation, otherwise the intensity signal is generated. In the first case suitable simulated tests have been performed to characterize the error distribution of modulation taking into account both the GPS type and the phase-shifting algorithm.

Treatments analyzed for the three design solutions are 54, 164 and 648 respectively.

Experiments are replicated 10^4 times. The longest experiment, the third, took about 7 hours on a high power workstation (AMD Opteron processor 244).

Table 1. Synoptic description of all the factors analysed in the three designed experiments

Factor (units)	Type	Levels	Affected design solutions
Sampling step (μm)	Control	0.1-0.2-0.3	ALL
Filter bandwidth (nm)	Control	50 (1)-100(2)	ALL
Standard deviation of the mirror step (nm)	Control	5-10-15	ALL
Standard deviation of video acquisition system	Control	0.008-0.012-0.016	ALL
Error of the mirror step (e_x)	Internal noise	Randomized : N(0, σ_x^2)	ALL
Error of video acquisition system (e_{CCD})	Internal noise	Randomized : N(0, σ_{CCD}^2)	ALL
Knots centering	External noise	Randomized : U(-0.5, 0.5)	ALL
Scale	External noise	Randomized: N(0.25, 0.078^2)	ALL
GPS	Control	$\lambda/4$-$\lambda/2$-$\lambda/4$ (1)-$\lambda/4$-P (2)	3
Phase-shifting algorithm	Control	4 intensity values-5 intensity values	3
Fitting function	Control	Gaussian (1)-Parabola (2)-$Sinc^2$ (3)	2 and 3
e_w	Internal noise	Randomized : N(0, $\lambda^2/250^2$)	3
e_r	Internal noise	Randomized : N(0, 0.01^2)	3
e_P	Internal noise	Randomized : N(0, $2.4\square10^{-12}$)	3

5. Results

5.1 Analysis of Measurement Bias

Absence of measurement bias, namely a systematic deviation of the measurement result from its true value, is ascertained by testing the null hypothesis $\mu_{x_M}=0$ against the alternative $\mu_{x_M}\neq0$ for each experimental setting of the control factors. Using ordinary t-tests the null hypothesis is rejected at a $1-\alpha$ significance level if

$$\overline{x_i} > z_{1-\alpha} \cdot \frac{S_i}{\sqrt{r}} \tag{5}$$

where subscript i denotes the experimental treatments, r is the number of replications and $z_{1-\alpha}$ the $1-\alpha$ quantile of the standardized Normal distribution; use of this distribution is

justified by the high number of replications ($r = 10^4$). A very low rejection rate of the tests (1.9%, 2.5% and 2.6% for the configurations 1, 2 and 3 respectively with $\alpha = 5\%$) indicates that measurement bias is not a big design concern.

5.2 Analysis of Measurement Uncertainty

The standard deviation of the measurement result is commonly used as a basic indicator of measurement uncertainty. It can be evaluated by looking at the plots of the mean effects and the most significant factorxfactor interactions for the three design configurations. As the same scale is used in all the plots, effects can be easily compared across the configurations. Notice that only two levels of the mirror displacement are considered for the first two design schemes. This is due to the fact that the intermediate level, 0.2 μm, being very close to the period of the intensity curve, causes an unfavourable sampling pattern. This does not affect scheme 3 as the phase shifters provide the mechanism to directly sample the modulation curve.

Interestingly the simplest configuration, the first, also provides the lowest uncertainty. As main effects (all 95% significant) are moderate it works sufficiently well over the entire design space. This can be beneficial as it allows looking for a design possibly balancing measurement quality, cost and speed. Anyway, lower uncertainty is obtained by using small mirror steps and CCD with medium-to-high quality. Filter 2 is slightly preferable. The design is extremely robust with respect to errors in mirror displacement. No particular interaction effect is worth noting.

Configuration 2 is characterized by larger main effects. In principle, this is not necessarily bad as it means that the designer can exert more control on the design. Uncertainty is quite sensitive to the size of mirror displacement and its error and, at a lesser extent, to the CCD error. Therefore, it would appear that low uncertainty could only be achieved with sacrifice of product cost and measurement speed. By the way, precious information comes from the interaction plots in Figure 6. Filter 2 is the key component of this design scheme; not only it makes the device extremely robust against random variations in the mirror displacement and the CCD but also provides a very low uncertainty regardless of the choice of the size of the mirror step and the fitting function. Thus solutions with high quality, low cost and high speed are feasible.

The last configuration exhibits the largest uncertainty. Yet initially it was considered the most promising since it includes more control factors and phase shifters are among the most widespread optical components. Even though the configuration is robust with respect to errors in the mirror step and in the CCD (the effect of σ_{CCD} is not significant at a 95% confidence level) uncertainty is unacceptably high: additional components, the GPS and the polarizers, and a heavier numerical processing, phase shifting and curve fitting algorithms, inflate measurement variability. Opposed to the other configurations, in this case filter 1 is by

far the best option.

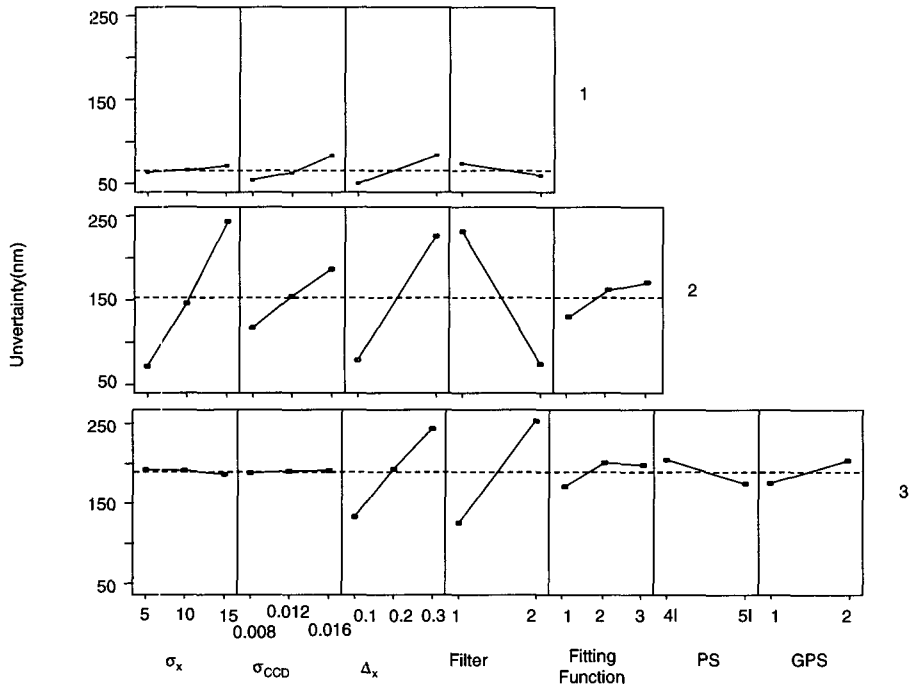


Figure 5. Main effects plots for the three configurations. Standard deviation of the measurement result is used as measurement uncertainty

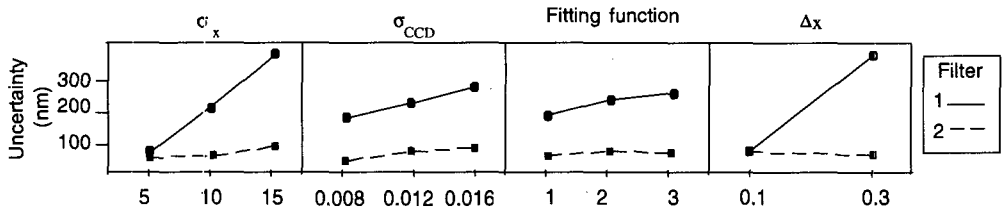


Figure 6. Interaction plots involving the filter as one of the factor for configuration 2

5.3 Overall Assessment of The Design Solutions

A comparison of the overall performance (uncertainty, speed and cost) of the three design schemes is made in Figure 7. From the runs of the computer experiments the most interesting solutions are picked up. The selection criterion looks for those solutions that minimize uncertainty for significant classes of product cost. These best solutions are represented in the cost-uncertainty plane for two levels of measurement speed, corresponding to a mirror step

of 0.1 μm (low speed) and 0.3 μm (high speed). Cost classes depend on the components adopted and their quality level. Main cost drivers, in the decreasing order of monetary impact, are the video camera, the phase-shifters, the piezo-electric translator and the filter (Figure 8). Solution 1 and 2 closely compete for the best performance. Solution 1 is slightly preferable for low measurement speed but is inferior to solution 2 if fast measurements are desired. This is valuable information for the designer. Another useful and unexpected outcome is that, in the case of solution 3, the considerable money spent to use a high quality CCD would have no return at all.

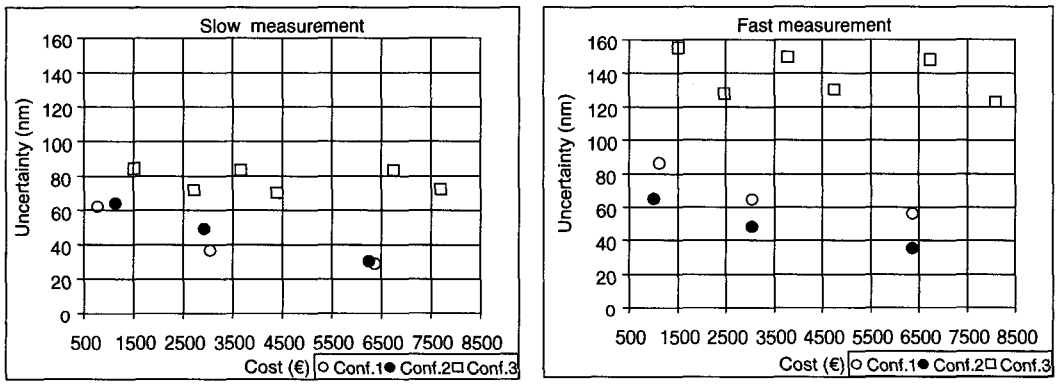
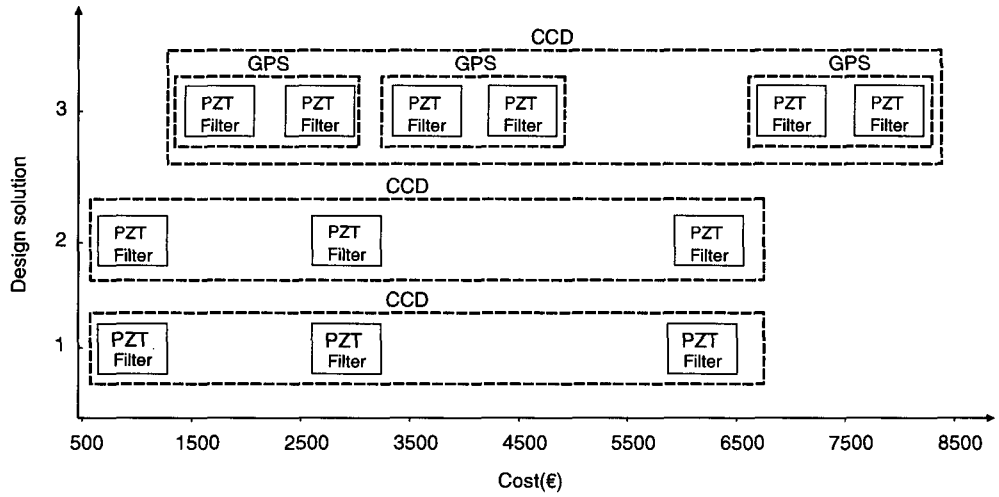


Figure 7. Comparison of the best design solutions for relevant classes of cost (three classes for solutions 1 and 2, six for solution 3). Points refer to the experimental runs minimizing uncertainty within each class. Left: solutions with low speed (0.1 μm mirror step); right: solutions with high speed (0.3 μm)



Source: ThorLabs catalog sales, Vol. 17 (2005)

Figure 8. Meaningful cost classes and their cost drivers

6. Discussion

The findings of the study are of major practical value for designing a profilometer that is able to respond to critical user needs. They specifically focus on product functionality (bias, uncertainty and speed of the measurement) and cost allowing the designer to trade them off. The solution is generated developing one of a number of design ideas initially proposed by the designer himself. The method can be seen as a powerful design facilitator since it is able to compare, on an objective base, different design concepts. But there is something more. The discovery of cause-effect relations between design ideas and product performance may in turn stimulate the designer's intuition so that new design ideas are generated. Hence, this concerted mechanism has the potential not only to improve the product but also to innovate it.

Research in the sector of optical devices is very vital at the moment. New systems and ingenious algorithms are continuously developed, new application fields opened. This is also due to the important role that numerical processing plays in the design of optical systems. Nevertheless there is a lack of systematic approaches to validate and rank the plethora of innovative thinking. New design concepts are often promoted without demonstrating their soundness. One major reason for it is that physical prototypes are normally very costly and difficult to operate. Moreover, in the case of measuring devices, the study of measurement uncertainty usually requires long experimental campaigns. At this regard, the method proposed in the paper offers an interesting way for efficient resources allocation. The paradigm would be to use extensive numerical experiments to comprehensively explore the design space, then a few experimental runs in the lab to confirm the most attractive findings.

A final note concerns the easy of use of the method. In the case-study of the profilometer the analysis is kept to a minimum and results are easily drawn mainly by using graphical tools. This was purposely done to reassure potential users of the method. In fact, most of the times the fundamental results of such a study can be grasped without any specific statistical expertise. However, more sophisticated tools, such as regression analysis and optimisation, might be needed at a later stage when, after the results are confirmed by physical trials, the final optimized design is sought for.

References

1. Baldi, A., Leban, B., Ginesu, F., Jacquot, P.(2003), "Sulla ricostruzione di superfici di interesse meccanico tramite interferometria in luce bianca," *AIAS Conference Proceeding.*
2. De Groot, P.(1995), "Vibration in phase-shifting interferometry," *J. Opt. Soc. Am.*, Vol. 12, pp. 354-365.

3. Dorrio, B., and Fernandez, J.(1999), "Phase-evaluation methods in whole-field optical measurement techniques," M*eas. Sci. Technol.* Vol. 10, pp. R33-R55.
4. Myers, R. H., and Montgomery, D.C.(1995), *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, John Wiley and Sons, New York, NY.
5. Nair, V.N., Ed.(1992), "Taguchi's Parameter Design: A Panel Discussion," *Technometrics*, Vol. 34, No. 2, pp. 127-161.
6. Robinson, D. W., and Reid, G.(1993), *Interferogram Analysis*, IOP Publishing, Bristol and Philadelphia.
7. Romano, D., Varetto, M., Vicario, G.(2004), "A General Framework for Multiresponse Robust Design based on Combined Array," *Journal of Quality Technology*, Vol. 36, No. 1, American Society for Quality, Milwaukee, Wisconsin (USA), pp. 27-37.
8. Roy, M., Svahn, P., Cherel, L., Sheppard, C.(2001), "Geometric phase-shifting for low-coherence interference microscopy," *Optics and Lasers in Engineering,* No. 37, pp. 631-641.
9. ThorLabs catalog sales(2005), Vol. 17.