DESIGN OF EXPERIMENT FOR QUALITATIVE EQUATION DISCOVERY: A COMPARISON

Federico Di Palma, Monica Reggiani, and Paolo Fiorini

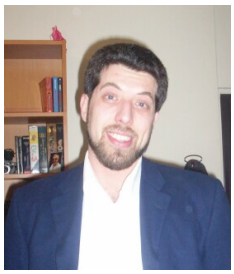
University of Verona, Department of Computer Science
Strada le Grazie 15, Verona, Italy
federico.dipalma@metropolis.sci.univr.it(Federico Di Palma)

Abstract

In the latest years, the research in the field of equation discovery focused from quantitative to qualitative model discovering. This requires a different design of experiment and only a few techniques are currently available to learn qualitative models. Among them random design is still the most adopted for qualitative analysis. This work proposes a methodology to adapt effective experiment design techniques for quantitative discovery to the qualitative field. The proposed methodology can be described as an incremental design technique, where the learning of the qualitative model is a cyclic refinement process. At each cycle, the methodology focuses the new experiment in those areas of the design space which are less covered by the previous experiment and where the current model exhibits a lower performance.
In this work, the proposed methodology is applied to alphabetic optimal and latin hypercube designs. An evaluation of the efficacy of the proposed solution applied to the two design techniques with respect to a random design is proposed within a robotic application. The performance of the design are assessed by means of a new index called E-Index based on the qualitative model extracted from the full design space. This new index is able to distinguish between the performances of the experiment design and of the learning algorithm.

Keywords: Qualitative Equation Discovery, Design of Experiment, Optimal Design, Coverage Design

Presenting Author's Biography



Federico Di Palma received the Laurea degree in Computer Engineering in 2002 and the PhD degree in 2006 from the University of Pavia. He was the recipient of the 2006 Best Doctoral Thesis Award from the IEEE Test Technology Technical Council. He is currently collaborating with the Universities of Pavia and Verona (Italy). His research deals with fault diagnosis for semiconductor manufacturing, neural networks, model predictive control and experimental design.

ISBN 978-3-901608-32-2

1 Introduction

In the latest twenty years, the research in the field of equation discovery focused from quantitative to qualitative models. This growing interest is shown by the existence of several automatic tools for the identification of qualitative models, which have been used in different fields, such as economics [1], medicine [2], and chemistry [3]. In order to build an expressive model for these applications, a preliminary phase to collect data from the real world is necessary. The cost of this expensive phase can be lessened by developing design methods to reduce the number of measurements required to obtain a good qualitative model.

A few design methods have been proposed for qualitative equation discovery, but random design is still the most adopted for qualitative analysis. The aim of this work is to investigate whether effective design techniques for quantitative model discovery problems can be applied to qualitative analysis. A methodology is described and applied to alphabetic optimal [4] and to the latin hypercube [5] experiment designs. The proposed methodology is evaluated in a robotic application where the two design techniques are compared with a random design. Finally, a novel index (*E*-index) is also introduced to assess the performance of each technique.

The paper is organized as follows: next section introduces the proposed methodology. The benchmark is presented in Section 3. The *E*-index is introduced in Section 4 and the comparison results are discussed in Section 5. Finally, a summary of the current work is given. Mathematical details about the derivation of the benchmark model are summarized in the appendix.

2 Methodology

A relation that links a variable $y \in Y$ with a vector variable $x \in X$ can be modeled using two main techniques: qualitative or quantitative. The former approach aims at identifying a function $f : X \mapsto Y$ that is as close as possible to the true relation. The latter is not interested in the numerical evaluation of y, but looks for the qualitative behavior of y when a change on x is observed. Roughly speaking, this model tries to answer the question about what happens to y if x changes. This approach aims at identifying a map $G : X \mapsto Y_b$ where Y_b is the collection of all the qualitative behaviors that the variable y can show, such as "y increases when x decreases".

Among the several qualitative predictive models that maps observation about a system to conclusions about the dependent variable behavior, this work concentrates on qualitative trees [6]. A qualitative tree (Figure 1) is a binary tree with internal nodes called splits and qualitatively constrained functions in the leaves. The splits define a partition of the state space into *regions* with common qualitative behavior of the dependent variable. A leaf represents qualitatively constrains given the *region* defined by the path from the root.

Figure 1 shows a qualitative model for the benchmark proposed in Section 3. The qualitative tree includes two

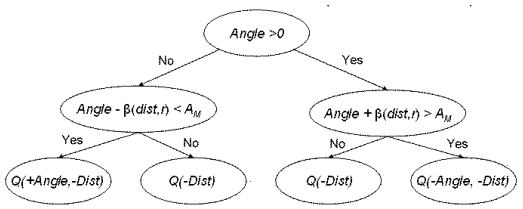


Fig. 1 Qualitative Model for the proposed benchmark.

predictor variables $Angle$ and $Dist$ and defines the following four *regions* (β defined in Equation 3):

$$X_1 := \{(Angle, Dist) : Angle < 0 \wedge Angle - \beta(Dist, f) > A_M\}$$
$$X_2 := \{(Angle, Dist) : Angle < 0 \wedge Angle - \beta(Dist, f) < A_M\}$$
$$X_3 := \{(Angle, Dist) : Angle > 0 \wedge Angle + \beta(Dist, f) < A_M\}$$
$$X_4 := \{(Angle, Dist) : Angle > 0 \wedge Angle + \beta(Dist, f) > A_M\}$$

Each region is then characterized by a qualitatively constrained function. As an example, for region X_1 (leftmost leave of the qualitative tree), $Q(+Angle, -Dist)$ means that the dependent variable is stricktly increasing in its dependence to $Angle$ and stricktly decreasing with $Dist$.

A tree can be "learned" through the identification of the qualitative model based on a collection of N real couple measurements (x, y). The whole data set is usually called *experiment*, while number N is its *size*. The objective of the design of experiment is to carefully choose the values of the predictor variable x within the *design space* X to reduce the size of the experiment required to learn a good model. The set of the chosen x values are usually returned row wise in a *design matrix* D.

This paper presents an experimental design methodology for the determination of qualitative tree models. The methodology can be described as an incremental design technique, where the learning of the qualitative model is a cyclic refinement process (Figure 2). This process involves three agents: the Learner, the Designer, and the Executer. At each cycle, the Learner identifies a qualitative model, called *current model*, from a set of measurements (*current experiment*). Both current model and current experiment are analyzed by the Designer which identifies a new *design matrix*. The Executer will complete the new experiment $E'(D)$ adding to the design matrix the measured dependent variable values. The new data are then added to the previous ones and constitute the new current experiment used by the Learner to build the new model in the next cycle. The overall process is iterated until a suitable model quality is achieved.

To summarize, we can define the design problem as:
At each cycle t, given a design space $X(t)$, a current experiment $E(t)$, and a current model $M(E(t))$, define a design matrix $D(t)$ of N independent variables, that will improve the quality of the next cycle model $M(E(t+1))$, where $E(t+1) = E(t) \bigcup E'(D(t))$.

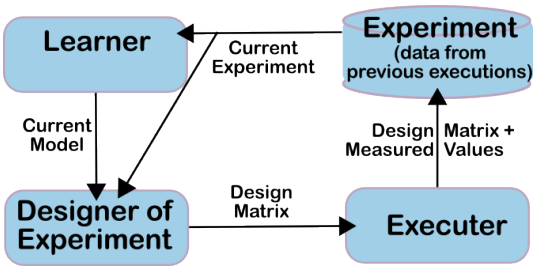


Fig. 2 Learning cycle process for the discovery of qualitative models.

2.1 Design strategy

The design strategy proposed in this paper aims at solving the design problem for qualitative models applying effective design techniques borrowed from the quantitative model discovery research field.

As previously mentioned in this work, the model M (a qualitative tree) partitions the design space X into R regions with common qualitative behavior ($B_r \in Y_b$) of the dependent variable. Each region of the current qualitative model can be evaluated based on the coverage of past experiments (Cov_r) and accuracy of the qualitative behavior B_r (Ac_r). These are two unrelated characteristics of the model: a region could have good accuracy of B_r, i.e. the modeled behavior is coherent with the real collected data of the experiment, and poor coverage, i.e. only a small amount of experimental data are located in the region.

The proposed methodology uses Cov_r and Ac_r to focus the new experiment in those regions which are less covered by the current experiment and exhibits a lower accuracy. Therefore, the size of the experiment in each region, N_r, has an inverse relation with accuracy and coverage measures, i.e. the worst is the model and the less are the information, the larger will be the number of measurements that will be executed in this region. If the two measures are supposed uncorrelated, a linear distribution assigns to each sub-design the size of the experiment according to:

$$\begin{aligned} N_r &= N\frac{1-Cov_r+1-Ac_r}{\sum_{i=1}^{R}(1-Cov_r+1-Ac_r)} = \\ &= N\frac{2-Cov_r-Ac_r}{2R-\sum_{i=1}^{R}(Cov_r+Ac_r)} \end{aligned} \quad (1)$$

with

$$N = \sum N_r$$

Once that the values of N_r are defined for the R regions, a design matrix can be construct applying a standard quantitative design algorithms, such as random, latin hypercube, or lattice sampling (Section 2.2) to the design space X_r of each region. The use of quantitative design algorithms requires an additional assumption as some of them assume to know the structure of the quantitative function of model M while qualitative models only provide a qualitative constrain. This constrain asserts that the sign of the entries of the quantitative model gradient have a constant and known sign within the design space X_r. Several structure of quantitative functions are compatible with this requirement and we decided to adopt a linear structure, the simplest possible model. This solution does not prevent modeling real scenarios with more difficult structures as the complexity can be dealt with through a larger number of regions.

The proposed strategy is summarized in Algorithm 1.

Algorithm 1 Design Methodology

for each Region r **do** // *analysis of the current model*
  Evaluate the coverage Cov_r
  Evaluate the accuracy Ac_r
end for
for each Region r **do** // *design*
  Evaluate the size of the experiment N_r according to equation 1
  Determine the design space X_r
  Apply the design algorithm considering a linear model
end for
Compute the overall design: $D = \bigcup D_r$

The process to evaluate coverage (Cov_r) and accuracy (Ac_r) is not specified in Algorithm 1 because it is strictly dependent on the application at hand. Additional details can be found in Section 3.3 describing how they are computed for the proposed benchmarks.

2.2 Application to qualitative designs

The proposed methodology was used to extend two well known quantitative design techniques, the alphabetic optimal design and the latin hypercube design, to the area of qualitative analysis.

2.2.1 Latin hypercube design

The latin hypercube design [7] is a coverage design algorithm that defines the experiment with the goal of maximizing the number of different levels (values) of each variable.

This algorithm is based on the concept of latin square. In statistical sampling, a square grid containing sample positions is a Latin square if, and only if, there is only one sample in each row and each column. A latin hypercube is the extension of this concept to an arbitrary number of dimensions, whereby each sample is the only one in each axis-aligned hyperplane containing it. The latin hypercube sampling algorithm divides the range of each variable into *M* equally probable intervals. *M* sample points are then placed to satisfy the latin hypercube requirements. The main advantage of this sampling scheme is that the number of samples does not increase with the dimensions, i.e., with the number of variables.

2.2.2 Alphabetic optimal design

The term *alphabetic optimal design* [8] refers to a class of design algorithms widely used for function estimation. These algorithms aim at selecting a design matrix that maximizes the ability of supporting the estimation of a surface with a defined shape. This ability is quantified by many different criteria [4] each one labeled by single capital letters (A, D, I, ...). Roughly speaking, this method aims at selecting a set of independent vector variables whose elements are uncorrelated (orthogonal). The use of low correlated test vectors introduces a small bias as, if two variables are fully correlated (i.e. they always show the same values), it would be impossible to differentiate between them.

In this work we apply the A-optimal design algorithm which adopts the following quality measure:

$$A(E) = trace\,(R'R) \tag{2}$$

where the function $trace(Z)$ indicates the sum of the element on the diagonal of the matrix Z and R is the regressor matrix [9] evaluated using the experiment E. This criterion has several properties: it minimizes the average variance of the parameters and reduces the asphericity of the confidence ellipsoid [10]. In practice, A-optimality design tends to choose values on the border of the candidate set.

2.2.3 Requirements

The considered quantitative design algorithms require two parameters: the size of the designed experiment (N) and a finite design space (X). The former parameter is provided by Equation 1, while the latter is a requisite of the design algorithms. The exact dimension of the design space dependents on the experiment at hand and will be evaluated in Section 3.3 for the proposed benchmark.

2.3 Initialization of the methodology

The methodology reported in Algorithm 1 requires the knowledge of the initial model. A natural approach is to consider the overall design space as a single region and apply the chosen design algorithm from the beginning (i.e. A-optimal design or latin hypercube sampling).

This approach well suites latin hypercube design which is a coverage design, but exhibits weaknesses with optimal design techniques. Alphabetic optimal design applied to linear models chooses design matrix entries in the corner of the design space. The initial qualitative model is therefore strongly biased while it should be rather based on a uniform coverage of the design space. To avoid this drawback, the initial model is built upon a first experiment based on latin hypercube sampling whatever design strategy will hereafter be applied.

3 The benchmark

The benchmark used in this work has been inspired by a robotic learning scenario. This scenario includes a robot moving in an infinite plane with a single object (a sphere). The robot is equipped with a camera with a 1 mm focal length (f) and a viewing angle of $2\pi/3.9$ rad (A_M). The sphere has a radius (r) of 2 cm and its center has the same height of the camera focus. From the sensors the embodiment can extract three features: the distance between the object and the robot (*Dist*), the area of the sphere in the image (*Area*) and the angle between the focal axis and the axis connecting the optical center of the camera and the sphere center (*Angle*). Since the view angle is limited, the sphere may disappear from the robot sight, for example when the robot is turning away from it. The proximity distance sensor is instead always able to measure the *Dist* value. In this benchmark the goal of the learning tool is to identify the model that relates $Area$ values to $Angle$ and $Dist$ sensor readings.

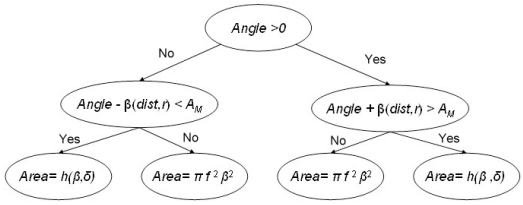


Fig. 3 Quantitative Function for the proposed benchmark.

As shown in Figure 2, the learning process of a qualitative model via an incremental experiment design involves two different agents beside the Designer of the experiment: the Executer and the Learner. The former measures a set of values of the dependent variable $Angle$ based on the design matrix. These data (*current experiment*) will then be used by the Learner to identify the qualitative model.

3.1 The executer

In this first work on this topic, it was decided to use a simulated scenario to provide an error-free contest. This choice was motivated by the possibility to identify the real qualitative model, which is the final goal of the learning process. The difference between this model and the qualitative models proposed by the learner is a simple measure of the performance of our methodology.

The output of the execution phase is produced using quantitative relations among the three variables. The mathematical model[1] is presented in Figure 3 where:

$$h(\beta,\delta) = f^2\left[\beta^2 \arcsin\frac{\delta}{2\beta} + \frac{\pi\beta^2}{2} - \delta\sqrt{\beta^2-\delta^2}\right] \tag{3}$$

with:

$$\beta = \arcsin\left(\frac{r}{r+Dist}\right) \tag{4}$$

$$\delta = 2A_M - \|Angle\| - \beta \tag{5}$$

1. Details about the mathematical model are presented in the Appendix.

The real qualitative model for the benchmark shown in Figure 1 can be obtained from the quantitative model evaluating the sign of the partial derivative of the functions on the leaves of the classification tree in Figure 3.

3.2 The learner

The qualitative tree model is identified using a two-steps learning strategy:

1. *Labeling phase*: the partial derivatives (with respect to *Angle* and *Dist*) are estimated for all the measurements composing the experiment E. The sign of the derivatives is then used to label the vector variables x composing the design matrix.

2. *Rules extraction*: the qualitative tree is identified clustering the current experiment according to the labels of the previous phase.

This strategy is implemented through the machine learning tool Orange [11]: a learning technique based on the Pade algorithm [12] is used for the labeling phase and the CN2 induction algorithm [13] for the construction of the qualitative tree. Figure 4 shows the Orange scheme used in this work and the parameter configuration of each widget. The only variable parameter is the *min instance in leaves* used to limit the number of regions. It has been set to the 7% of the size of the current experiment to limit the number of leaves to fourteen.

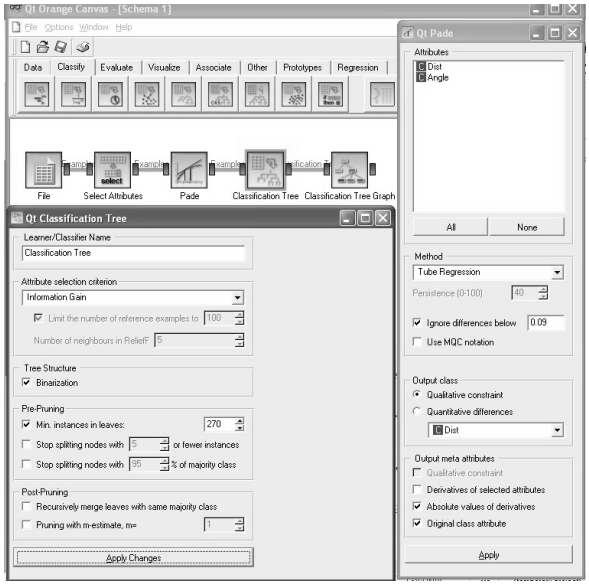


Fig. 4 Orange scheme used for the identification of the qualitative model.

3.3 Application of the Methodology

The use of the general methodology in a specific scenario requires the discretization of the design space (Section 2.2.3) and the definition of the algorithms to compute the accuracy (Ac_r) and coverage (Cov_r) (Section 2.1).

As introduced in Section 2.2.3 both the latin hypercube and the alphabetic optimal designs require a finite design space. In the proposed benchmark the design space is defined by all the possible couples $(Angle, Dist)$. Therefore a design space can be limited by reducing the admitted valued levels of both variables to a finite number. The choice of the levels can be critical because it can prevent the acquisition of meaningful data within a region thus making its correct identification difficult or even impossible. We have identified in 30 a reasonable number of equi-spaced levels. The analysis of the real qualitative model shown in Figure 1 confirms the validity of our choice as it does not create too small regions that would not be correctly identified.

With an error free simulated benchmark, the accuracy of the model within a region (Ac_r) can be estimated as the percentage of labels that match the behavior B_r defined by the qualitative model. As we are working with a finite discrete design space, the coverage (Cov_r) can instead be obtained as:

$$Cov_r = \frac{N_r}{X_r}$$

where N_r is the size of the experiment in the region r and X_r the number of possible couples $(Angle, Dist)$ in the region design space.

4 Performance assessment

A fairly intuitive measurement of performance of the applied design techniques is the "distance" between the identified qualitative model and the real one. However, the identification of the real model is quite complex and the learning tools can fail. Even our rather simple benchmark results in a complex qualitative model and the learning tool is unable to manage its complex constrains failing in identifying the real model.

This deficiency of the learning tool can create some difficulties in assessing the performance of the design techniques. Indeed, when the real qualitative model is not obtained, it becomes quite complex understanding whether this is due to the weakness of the learning tools or of the applied design technique. Therefore, we decided to assess the performance of the design technique in relation to a *reference model*, which is the model obtained by the learning method when fed by the whole design space (Figure 5).

A measurement of the performance of the experiment design is defined by a *E*-index, a new index which refers to the percentage of design space measurements that the current model labels accordingly to the reference model. The use of a reference model instead of the real qualitative one is the novelty introduced by the E-index.

The definition of the E-index requires a finite design space that the learning tools can handle. When the learning tool is unable to analyze the whole design space or when the design space is infinite, the design space is sampled and used to evaluate the reference model. In the proposed benchmark the design space is

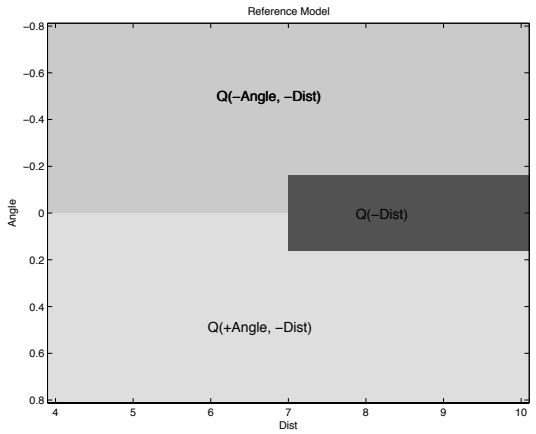


Fig. 5 Reference model for the proposed benchmark, the model returned by the learning algorithm when is fed with the whole design space.

discrete and the learning tool can directly handle it, then the reference model and the E-index can be evaluated without any further action.

5 Design comparison

The proposed methodology (Algorithm 1) applied to the latin hypercube and the alphabetic optimal design has been compared with the random design. The proposed procedure requires the definition of only one parameter: the dimension of the experiment N. At each step, a new model will be identified by adding a new set of N samples to the previous acquired data. N should be small enough to appreciate the performance of the different methods but, at the same time, the number of samples should allow an effective refinement to the current model. In this comparison we identified $N = 90$ (the 10% of the design space) as an acceptable trade off.

Figure 6 shows the evolution of the identified models for the proposed benchmark using a random design (first column), and the proposed methodology using both the latin hypercube (second column) and the alphabetic optimal design (third column). In order to have a quantitative and synthetic evaluation of the distance between the identified models and the reference model, the E-indexes were computed. As shown in Figure 7, when a limited number of samples is used to identify the model, the E-indexes of the proposed methodology is higher than the E-index of random design thus indicating a slightly better performance of latin hypercube and alphabetic optimal design techniques.

6 Conclusions

This paper proposes a methodology to extend some standard design techniques for quantitative model discovery problems to the qualitative analysis. The benefits of the described solution have been evaluated through a comparison between a random design and the proposed methodology using the latin hypercube or the alphabetic optimal design techniques. A new index (E-

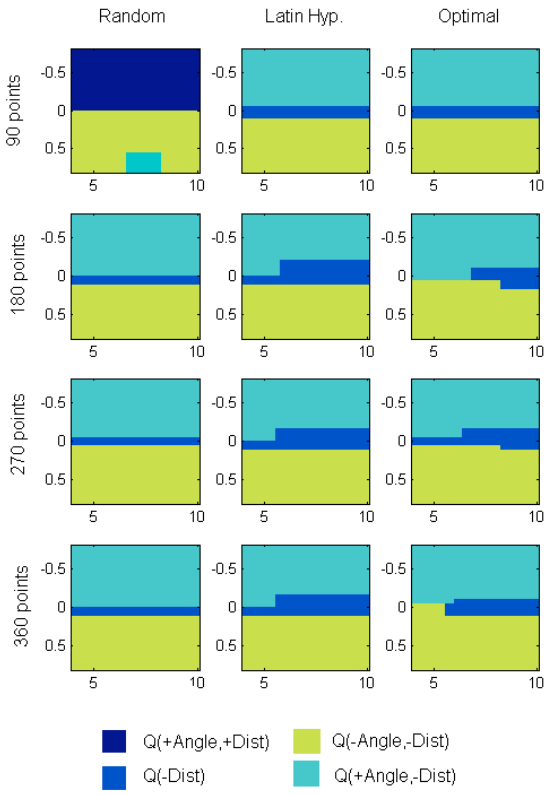


Fig. 6 Evolution of the identified model using a random design (first column), and the methodology reported in Algorithm 1 using both latin hypercube (second column) and the alphabetic optimal (third column) designs.

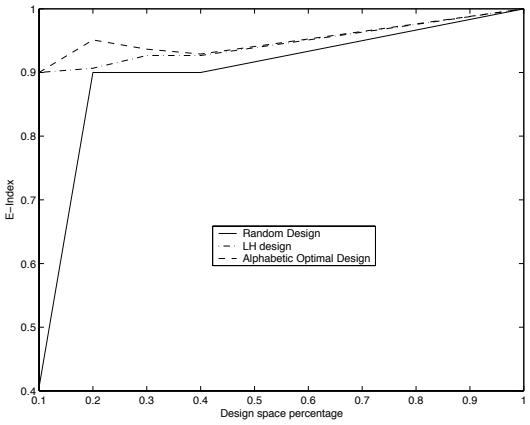


Fig. 7 E-index of the identified models.

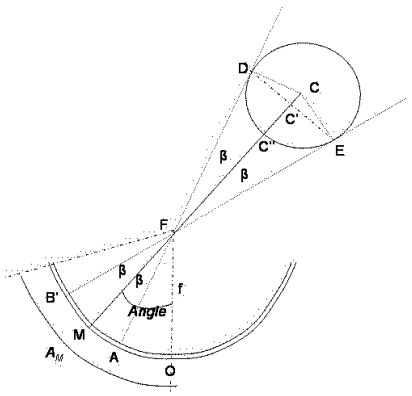


Fig. 8 Bird-view when the sphere is fully visibile.

index) was introduced, able to distinguish between the performance of the experiment design and of the learning algorithm.

The design comparison is limited to a simple experiment and we are aware that additional research is required to assess the performance of the introduced methodology. At the time of writing, further investigations are under analysis. Nevertheless, the early results and the new solution to evaluate the performance of the designer (*E*-index) are already quite promising.

Acknowledgement

The work described in this article has been funded by the European Commission's Sixth Framework Programme under contract no. 029427 as part of the Specific Targeted Research Project *XPERO* ("Robotic Learning by Experimentation").

7 References

[1] Lloyd W. Condra. *Reliability Improvement with Design of Experiments*. CRC Press, 2001.

[2] Karl Thomaseth. Optimal experiment design for assessing plasma clearance of 125i-iothalamate in peritoneal dialysis. In *Proceedings of the 19th International Conference IEEE/EMBS*, pages 2128–2131, Chicago, IL. USA, Oct. 30 - Nov. 2 1997.

[3] Maria Pia Saccomani and Claudio Cobelli. A minimal input-output configuration for a priori identifiability of a compartmental model of leucine metabolism. *IEEE Transactions on Biomedical Engineering*, 40(8):797–803, August 1993.

[4] V. V. Fedorov. *Theory of Optimal Experiments*. Academic Press, Inc., London, UK, 1972.

[5] D. J. C. MacKay. *Bayesian Methods for Adaptive Models*. PhD thesis, California Institute of Technology, 1991.

[6] D. Šuc and I. Bratko. Induction of qualitative trees. In *EMCL '01: Proceedings of the 12th European Conference on Machine Learning*, pages 442–453, London, UK, 2001. Springer-Verlag.

[7] M. Stein. Large sample properties of simulations using latin hypercube sampling. *Tecnometrics*, 29(2):143–151, 1987.

[8] A. I. Schein and L. H. Ungar. A-optimality for active learning of logistic regression classifiers. Technical Report MS-CIS-04-07, The University of Pennsylvania Department of Computer and Information Science, 2004.

[9] L. Ljung. *System Identification, Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ, 1987.

[10] The MathWorks, Natick, MA. *Model-based calibration Toolbox 3 Reference*, 2005.

[11] J. Demšar, B. Zupan, and G. Leban. Orange: From experimental machine learning to interactive data mining. White Paper. Faculty of Computer and Information Science, University of Ljubljana. www.ailab.si/orange, 2004.

[12] J. Žabkar and J. Demšar. Pade. Technical report, Faculty of Computer and Info. Sc., Ljubljana, December 2006.

[13] P. Clark and T. Niblett. The cn2 induction algorithm. *Machine Learning*, 3:262–284, 1989.

Appendix

This section summarizes the mathematical details of the model used to simulate the real scenario in the execution phase. The scenario at hand supposes that the center of the sphere and the focal point of the concave lens are at the same height and that the distance between the robot and the sphere is always available. Moreover, the model considers that the lens has a finite dimension that limits the maximum visible angle to A_M. The model also supposes that the vision sensor does not deform the object. This assumption can be modeled considering a curved image surface with a curvature equal to the focus length.

In this model, two distinct states are possible: full or partial visibility of the sphere.

Full visibility

Figure 8 shows the bird-view when the sphere is fully visibile. In this case, the sensed image includes a full disk whose diameter d is the arc $B'A$ in Figure 8 and is equal to:

$$d = 2f\beta. \tag{6}$$

where the angle β is evaluated as:

$$\beta = \arcsin(\frac{r}{r + Dist}) \tag{7}$$

with r, radius of the sphere. Then, the area is computed as:

$$Area = \pi f^2 \beta^2. \tag{8}$$

It is important to remark that the center of the image M coincides with the projection of the sphere center on the lens.

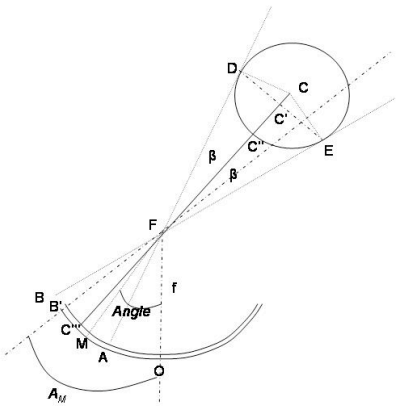


Fig. 9 Bird-view when the sphere is partially occluded.

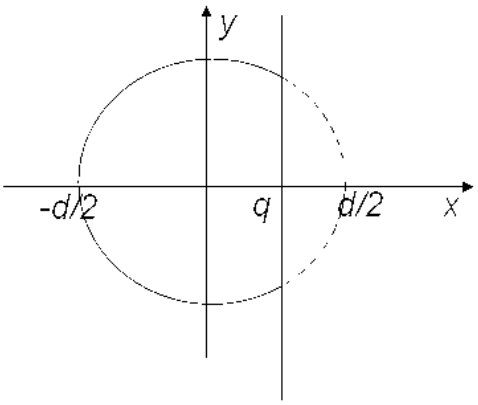


Fig. 10 Image (full line) printed in the image plan from the scenario reported in Figure 9. The image is placed in a Cartesian coordinate system centered onthe center of the full sphere (dot line).

Partial visibility

Figure 9 shows a bird-view when the sphere is partially occluded, i.e. only a partial disk is visible on the lens. The center (M) of this partial disk does no longer coincide with the center of the sphere (C''). The model is in this state when the following equation is not satisfied:

$$\|Angle\| + \beta < A_M. \tag{9}$$

The diameter of the full sphere is (6), while the width d' of the partial disk in the image is the arc $B'A$ in Figure 9. This figure shows that the arc $B'M$ is subtended by the angle $A_M - Angle$, therefore:

$$d' = 2f\,(A_M - \alpha) \tag{10}$$

The computation of the $Area$ is slightly more difficult than in the previous state. Figure 10 shows a partial disk where the Cartesian coordinate system is centered in the center of image of the full sphere. The $Area$ can be evaluated as:

$$\begin{aligned} Area &= 2\int_{-\frac{d}{2}}^{q} \sqrt{\frac{d^2}{4} - x^2}\delta x = \\ &= 2\left[\frac{d^2}{4}\arcsin\frac{2x}{d} + \frac{x}{2}\sqrt{\frac{d^2}{4} - x^2}\right]_{-\frac{d}{2}}^{q} = \\ &= \frac{d^2}{4}\arcsin\frac{2q}{d} - q\sqrt{\frac{d^2}{4} - q^2} + \frac{d^2\pi}{8} \end{aligned} \tag{11}$$

where $q = d' - \frac{d}{2}$. Recalling 11, 6 and 10, the $Area$ value is given by:

$$Area = f^2\left(\beta^2\arcsin\frac{\delta}{\beta} - \delta\sqrt{\beta^2 - \delta^2} + \frac{\pi\beta^2}{2}\right) \tag{12}$$

where

$$\delta = 2A_M - 2\|Angle\| - \beta \tag{13}$$

The model described in this section allows to determine the quantitative tree in Figure 3: the node thresholds are defined using the equation 9 while the leafs include the $Area$ values from equations 12 and 9.

