Hindawi Publishing Corporation
EURASIP Journal on Applied Signal Processing
Volume 2006, Article ID 48012, Pages 1–10
DOI 10.1155/ASP/2006/48012

Design of Experiments for Performance Evaluation and Parameter Tuning of a Road Image Processing Chain

Yves Lucas,[1] Antonio Domingues,[2] Driss Driouchi,[3] and Sylvie Treuillet[4]

[1] *Laboratoire Vision et Robotique, IUT Mesures Physiques, Université d'Orléans, 63 avenue de Lattre, 18020 Bourges cedex, France*
[2] *Laboratoire Vision et Robotique, ENSIB 10 Bd Lahitolle, 18000 Bourges, France*
[3] *Laboratoire de Statistiques Théoriques et Appliquées, Université Pierre & Marie Curie, 175 rue du Chevaleret, 75013 Paris, France*
[4] *Laboratoire Vision et Robotique, Polytech Orléans 12, rue de Blois BP 6744 45067 Orleans, France*

Received 1 March 2005; Revised 20 November 2005; Accepted 28 November 2005

Tuning a complete image processing chain (IPC) is not a straightforward task. The first problem to overcome is the evaluation of the whole process. Until now researchers have focused on the evaluation of single algorithms based on a small number of test images and ad hoc tuning independent of input data. In this paper, we explain how the design of experiments applied on a large image database enables statistical modeling for IPC significant parameter identification. The second problem is then considered: how can we find the relevant tuning and continuously adapt image processing to input data? After the tuning of the IPC on a typical subset of the image database using numerical optimization, we develop an adaptive IPC based on a neural network working on input image descriptors. By testing this approach on an IPC dedicated-to-road obstacle detection, we demonstrate that this experimental methodology and software architecture can ensure continuous efficiency. The reason is simple: the IPC is globally optimized, from a large number of real images and with adaptive processing of input data.

1. ADAPTIVE PROCESSING IN VISION SYSTEMS

Designing an image processing application involves a sequence of low- and medium-level operators (filtering, edge detection and linking, corner detection, region growing, etc.) in order to extract relevant data for decision purposes (pattern recognition, classification, inspection, etc.). At each step of the processing, tuning parameters have a significant influence on the algorithm behavior and the ultimate quality of results. Thanks to the emergence of extremely powerful and low cost processors, artificial vision systems now exist for demanding applications such as video surveillance or car driving where the scene contents are uncontrolled, versatile, and rapidly changing. The automatic tuning of the IPC has to be solved, as the quality of low-level vision processes needs to be continuously preserved to guarantee high-level task robustness.

The first problem to be tackled in order to design adaptive vision systems is the evaluation of image processing tasks. Within the last few years, researchers have proposed rather empirical solutions [1–7]. When confirmed a ground truth is available, it is possible to compare directly this reference to the results obtained by using a specific metric. Sometimes no ground truth exists or data are uncertain and either application experts are needed for qualitative visual assessment or empirical numerical criteria are searched for. All these methods consider only one operator at a time [8–11]. However the separate tuning of each operator rarely leads to an optimal setting of the complete IPC. Moreover, image operators are generally tested on a too small number of test images, sometimes even on artificial noised images, to evaluate algorithm efficiency. This cannot replace a large real image base, for IPC testing. So, how can we evaluate on a great number of images a sequence of image processing operators involving numerous parameters?

A second problem remains unsolved: how to find the relevant tuning and hence how to adapt image processing to maintain a constant quality of results? As real time processing is executed by electronic circuits, this hardware must incorporate programmable facilities so that operator parameters can be modified in real time. Artificial retinas as well as intelligent video cameras already enable the tuning of some acquisition parameters. Concerning the processing parameters, the amount of computing necessary to distinguish the effect on the results of modifying several parameters seems at the first glance dissuasive, as separate images require different parameters. It should be noted that the choice of operators here still appeals to the experimenter but an other research work also examines the possibility of its automation [12, 13].

In this paper, we show how to overcome these problems using an experimental approach combining statistical modeling, numerical optimization, and learning. We illustrate this approach in the case of an IPC dedicated to line extraction for road obstacle detection.

2. METHODOLOGY OVERVIEW

To evaluate a full image processing chain, including a series of low- and medium-level operators with tunable parameters, instead of focusing on single algorithms, we need to adopt a global optimization approach. The first step is the evaluation of the IPC performance, depending on the significant tuning parameters to be identified and on their interactions. The second step is the parameter tuning itself which should enable adaptive image processing. It implies relating input image content to the optimal tuning parameter for each particular image. These two steps are described in the following paragraphs.

2.1. Performance evaluation

Building a specific and exhaustive database for the target application is the preliminary and delicate step to achieve relevant tuning of the IPC. Indeed, this database covering all situations is required during modeling, optimization, and control learning tasks. From a statistical point of view, selected images should reflect the frequency of any image content during the IPC operation and express all its versatility.

A typical subset of this database is then processed by the IPC. Output evaluation is here necessary in off-line mode, for IPC understanding and adjustment. This type of evaluation has been extensively researched even if the studies involve a single algorithm at each time. It remains a critical step, as each IPC is specific and requires its own evaluation criteria. The evaluation can be supported by a ground truth or can be unsupervised when empirical criteria are used instead.

Testing all the tuning parameters on the whole image database would lead to a combinatorial explosion; and moreover a physical model of that IPC could still not be deduced.

As it is necessary to model the influence of the IPC parameters, we decided to build instead a statistical model.

Modeling the parameter influence is carried out through the design of experiments [14]. This is a common tool in engineering practice but has been only recently introduced for machine vision applications [15, 16]. It consists in modeling the effects of simultaneous changes of IPC parameters with a minimum number of trials. In the simplest case, only two modes are allowed for each parameter: a low one and a high one, which means that the parameter bounds need to be carefully set. During the experiments, the IPC is considered as a black box whose factors (X_i tuning parameters) influence the effects (Y_i values of the criteria for output image evaluation) (Figure 1).

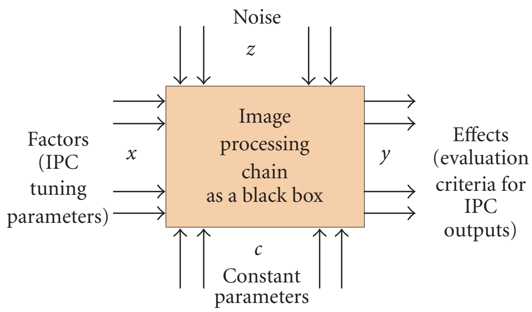


Figure 1: System modeling.

Note that tuning only one parameter at a time can not lead to an optimal setting as some parameters may be interdependent. Hence, the goal is to identify which of the parameters are really significant and their strong interactions with respect to the effects. Generally a polynomial model is adopted, whose coefficients a_{ij} are estimated by least square methods:

$$y = a_0 + a_1x_1 + \cdots + a_kx_k + a_{12}x_1x_2 + \cdots + a_{k-1k}x_{k-1}x_k. \quad (1)$$

The interpretation of the experiments by variance analysis confirms whether the model obtained is really meaningful or not. The amount of computing remains very high as the same trials must be repeated on a large number of test images to obtain statistical evidence. Hence, no optimal tuning is obtained for a given image, only an average tuning for the IPC itself. The parameters influencing significantly the quality of results are identified, and the strong interactions among them are also detected, so that only the latter are considered for further IPC programing tasks.

2.2. Parameter tuning

For each particular test image of the database, the optimal tuning of the IPC parameters still needs to be sought. This is typically an optimization process which still involves the output evaluation. The average tuning obtained previously provides valid initial conditions to the search process and the high and low modes of the significant parameters bound the exploration domain.

To obtain the optimal parameter tuning for the IPC, we look for methods not based on the local gradient computing as it is not available here. The simplex method enables to explore the experimental domain and to reach maxima using a simple cost function to guide the search direction [17]. Experimentally, a figure of $n+1$ points of an n-dimension space is moved and warped through geometric transformations in the parameter space, until a stop condition on the cost function is verified.

This produces a set of test images with optimal tuning parameters. But for real time purposes, the simplex method cannot be used for IPC tuning as it is time consuming. A solution consists in extracting descriptors from input images

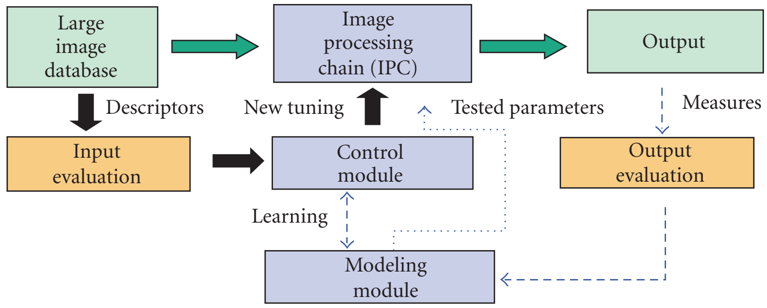


Figure 2: Architecture of an adaptive IPC.

that could be correlated to the optimal tuning parameters of these images. Such descriptors will be calculated also on new incoming images, and we should expect that images with similar descriptors will be processed correctly by the IPC during inline mode, using similar tuning parameters. So, to constitute a learning base, we compute the descriptors of the test images with known optimal tuning parameters.

The selection of relevant descriptors is not an obvious task and implies experimentation. The idea is that such descriptors should extract data which is significant for the tuning parameters of the considered IPC. Input evaluation has been investigated much less than output evaluation. Achieving an adaptive and automatic IPC tuning implies extracting relevant descriptors from input images, that is to say, they are closely related with IPC optimal tuning for each image. Image descriptors also enable the initial dimension of the tuning problem (image size in n^2 pixels) to be lowered, as each image pixel contributes to the tuning. Experimentally, a parameter vector lowers this dimension to the gray-level number ($\approx n$), using a histogram computed over the image.

The last step is the control module programing. This module will compute in real time adapted tuning parameters for new incoming images, using the descriptors of these images.

A neural network is a convenient tool for estimating the complex relation between the input image descriptors and the corresponding values of the tuning parameters. As mentioned previously, the set of test images with optimal tuning parameters constitutes the learning base of this network. Then, if the selected descriptors are relevant for the tuning purpose, the neural network should converge. The other part of the image database is reserved for the test of the neural network. The performance of the tuning will be steadily measured by comparing not the tuning parameters, but the IPC output directly. In particular, we will compare the neural network performance to simplex reference and also to the best trials of the design of experiments.

Finally, after the preceding steps devoted to statistical modeling, numerical optimization, and learning, the IPC is toggled into an operational mode, and the image processing tuning parameters are continuously adapted to the characteristics of new input images. To summarize our approach for IPC tuning, the architecture of an adaptive IPC can be the following (Figure 2).

In the following, we illustrate our approach for IPC tuning on a road image processing chain. This application will also help us to introduce practical details of the methodology. Naturally, input and output image evaluations will be specific to the application, but the methodology is generic.

3. APPLICATION TO A ROAD IMAGE PROCESSING CHAIN

3.1. IPC overview

This application is part of the French PREDIT program and has been integrated in the SPINE project (intelligent passive security) intended to configure an intelligent airbag system in precrash situations. An on-board multisensor system (EEV high speed camera + SICK scanning laser range finder) integrated in a PEUGEOT 406 experimental car classifies potential front obstacles and estimates their collision course in less than 100 ms [18–20]. To respect this drastic real-time constraint, a low and medium image processing has been implemented in the hardware with the support of the MBDA company. It consists of two ASIC circuits [21] embedded with a DSP into an electronic board interfaced with the vehicle CAN bus. As the first tests performed by the industrial car part supplier FAURECIA demonstrated that a static tuning is ineffective against road image variability, an automatic and adaptive tuning based on the approach presented here has been successfully adopted [22]. Eight reconfigurable parameters can be modified at any time: Canny-Deriche filter coefficient (X_1), image amplification coefficient (X_2), edge low and high threshold values (X_3, X_4), the number of elementary automata for contour closing (X_5), polygonal approximation threshold (X_6), little segment elimination threshold (X_7), and the approximation threshold for horizontal and vertical lines (X_8) (Figure 3).

3.2. Output evaluation

The IPC should extract from the image horizontal and vertical lines (Figure 4), which, after perceptual grouping, describe the potential obstacles in front of the experimental vehicle. Then, output evaluation is based on the number, spreading, and length of these segments inside a region of interest (ROI) called W and specified by the scanning laser range finder. We have proposed a quality evaluation criterion

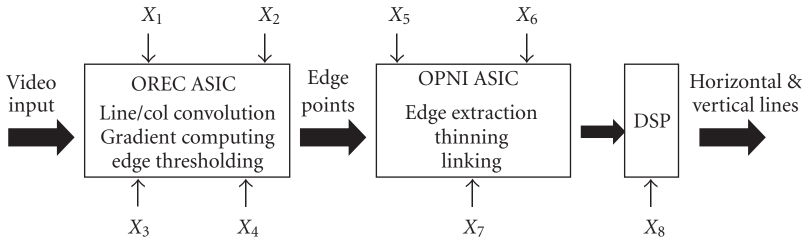
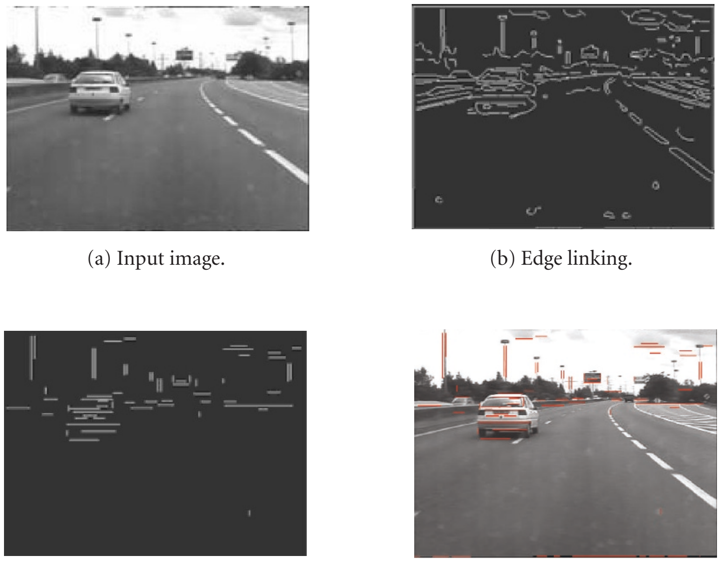


FIGURE 3: Tunable parameters of the road image processing chain.



(a) Input image.

(b) Edge linking.

(c) H/V lines.

(d) Lines over input image.

FIGURE 4: H/V line extraction.

called covering rate, which can be computed for different parameter tunings (Figure 5).

The covering rate r is defined as follows: for each horizontal or vertical S segment, we introduce a rectangular-shaped M_S mask centered on this segment and whose width is proportional to the length of that segment. The shape ratio of the mask is a constant, experimentally tuned on road images, to obtain significant variations of r for different tunings without saturation effects (ROI entirely covered by masks).

For each image pixel (i, j) in W (n_x- and n_y-dimensions), we define a function $f(i, j)$ by

$$\begin{aligned} f(i,j) &= 1 \quad \text{if } \exists S \in W \mid (i,j) \in M_S, \\ f(i,j) &= 0 \quad \text{otherwise.} \end{aligned} \tag{2}$$

The covering rate ($0 \leq r \leq 1$) is then simply given by

$$r = \frac{1}{n_x n_y} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} f(i,j). \tag{3}$$

The higher covering rate is desirable as it indicates that the ROI contains many large and well-distributed segments, which are robust entities for car detection.

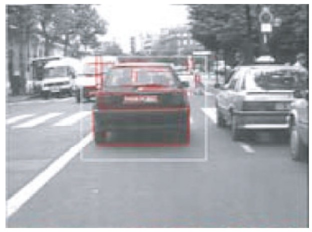
This criterion is dependent on the image content: if only a few segments exist, r cannot reach high scores even after optimal tuning, so r is considered as acceptable when most of the obstacle edges have been well extracted. An intuitive graphical interpretation exists for the covering rate: it is simply the part of the ROI which is covered by the superimposition of the masks associated to the set of segments detected by the IPC; it will be expressed in this paper as a percentage.

3.3. Statistical modeling

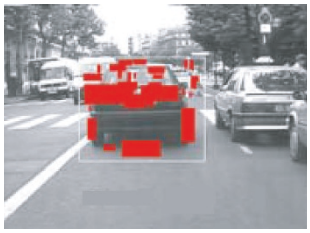
Three experiment designs have been implemented inside the modeling module: $a2^{k-p}$ factorial fractional design with 16 trials [23] to select the really significant parameters, a Rechschaffner design [24] with 37 trials, and finally a quadratic design with 27 trials, by adding an intermediate zero mode to detect nonlinearity. By using two modes for the tuning of each parameter (Table 1), 2^8 different IPC outputs can be compared from any given input image.

A preliminary task consists in specifying for each factor an interval which bounds the experimental domain. During each experimental trial, every factor is set to its low or high mode, depending on -1 or $+1$ value in the normalized

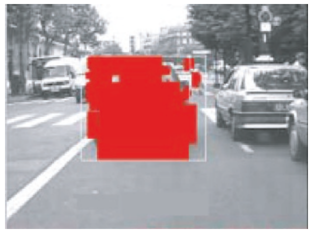
(a) Trial no 1.



(b) Trial no 7.



(c) Trial no 1: covering rate 31.50%.



(d) Trial no 7: covering rate 78.34%.

Figure 5: IPC output evaluation.

Table 1: Modes for all the design of experiments.

Factor	Parameter	Modes	
X_1	Canny-Deriche filter	0.5	1
X_2	Image amplification	33	63
X_3	Edge low threshold	5	15
X_4	Edge high threshold	15	30
X_5	Contour closing	26	30
X_6	Polygonal approximation	5	6
X_7	Little chains threshold	5	10
X_8	Slope threshold	1	3

Table 2: Experiment matrix-fractional factorial 2^{8-3} design: averaged outputs.

Trial	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	r (%)
1	−1	−1	−1	−1	−1	−1	−1	−1	35.535
2	−1	−1	−1	1	1	1	−1	1	40.310
3	−1	−1	1	−1	1	1	1	−1	27.859
4	−1	−1	1	1	−1	−1	1	1	42.436
5	−1	1	−1	−1	1	−1	1	1	47.328
6	−1	1	−1	1	−1	1	1	−1	30.284
7	−1	1	1	−1	−1	1	−1	1	44.034
8	−1	1	1	1	1	−1	−1	−1	37.743
9	1	−1	−1	−1	−1	1	1	1	46.517
10	1	−1	−1	1	1	−1	1	−1	40.469
11	1	−1	1	−1	1	−1	−1	1	50.680
12	1	−1	1	1	−1	1	−1	−1	33.464
13	1	1	−1	−1	1	1	−1	−1	35.169
14	1	1	−1	1	−1	−1	−1	1	49.255
15	1	1	1	−1	−1	−1	1	−1	39.715
16	1	1	1	1	1	1	1	1	44.842

experiment matrix. Therefore, each experiment design of experiments is well defined by its experiment matrix whose line number refers to the number of trials and column number refers to the number of tested parameters. We present below the experiment matrix and the covering rate for the set of trials of the first design of experiments (Table 2).

These designs have been tested on 180 input images selected from a video sequence of over 30 000 city and motorway frames. A statistical model has been deduced and validated by measuring R-Square and Mallow $C(p)$ indicator (Table 3). High R-Square and low $C(p)$ indicate that the number of significant parameters is three (X_1, X_6, X_8). A fourth parameter is not relevant as it does not appreciably improve the R-Square and $C(p)$ values; hence experimental data will not fit better to the model with an additional parameter. The first design of experiments only models the significant parameters without interactions:

$$Y = 51.1965 + 8.65X_1 - 4.08X_6 + 4.31X_8. \tag{4}$$

High module values of the coefficients denote significant parameters as the Y is strongly affected when such parameter toggles from low to high mode. The parameters with low module values are eliminated in the polynomial expression. It is interesting to note that this model is robust to image degradations, as it is not modified when we shift the grey levels of the test images two bit right (darker) or one bit left

Table 3: Significance of the model.

Coef	R-Square	$C(p)$	Factors
1	0.673	49.01	X_8
2	0.827	22.29	X_1, X_8
3	0.938	3.48	X_1, X_6, X_8
4	0.950	3.36	X_1, X_2, X_6, X_8
5	0.956	4.25	X_1, X_2, X_4, X_6, X_8
6	0.960	5.47	$X_1, X_3, X_4, X_6, X_7, X_8$
7	0.961	7.18	$X_1, X_2, X_3, X_4, X_6, X_7, X_8$
8	0.962	9.00	All

(brighter). The coefficients are slightly modified but the signs of the coefficients and the significant parameters remain the same.

We obtain for the left shift:

$$Y = 35.65 + 6.31X_1 - 3.14X_6 + 4.8X_8 \tag{5}$$

and for the right shift:

$$Y = 50.14 + 8.86X_1 - 5.01X_6 + 5.50X_8. \tag{6}$$

In Table 4, we added the internal IPC quality indicators on the 2^{8-3} design results: Y_1 stands for the number of edge points at OREC ASIC output, Y_2 is the average length of linked edge points at OPNI ASIC output, and Y_3 and Y_4 (resp., Y_5 and Y_6) are the number and average length of horizontal (vertical) lines detected at DSP output, respectively. It is clear that a separate tuning of the IPC components does not give optimal results for the whole IPC. Hence, the evaluation criteria for the IPC performance should only be computed at the output.

The second design of experiments (Table 5) displays another polynomial model that extracts the same three significant parameters. As the number of trials is larger, it is possible this time to take the strongest parameter interactions into account (Table 6). There is an interaction between two parameters if the tuning of one of the parameters works differently depending on the tuning of the second one. High module values for the coefficients of X_iX_j products denote strong interaction. Other products are eliminated in the polynomial expression:

$$\begin{aligned} Y &= 40.2 + 2.06X_1 + 0.74X_2 - 2.47X_6 \\ &\quad + 5.30X_8 - 0.92X_1X_2 + 0.95X_6X_8. \end{aligned} \tag{7}$$

Finally, in the third design of experiments (Table 7), only the three significant factors are tuned but a third mode is added to take nonlinear effects into account.

The covering rates obtained for the different trials provide an average tuning for the IPC parameters. This static tuning cannot be optimal for each given input image but it enables initializing the Nelder & Mead optimization algorithm based on the simplex method. This algorithm then computes all the parameter optimal values corresponding to each tested input image.

Table 4: Comparison of internal and output evaluation criteria.

Trial	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	r (%)
1	786	8.49	6.41	27.1	6.02	11.6	35.53
2	749	5.86	6.15	29.2	6.24	13.6	40.31
3	777	6.44	5.23	26.2	3.67	12.2	27.86
4	738	10.1	6.32	30.1	6.01	13.7	42.44
5	887	9.17	8.00	28.9	6.92	14.0	47.33
6	869	6.09	6.08	25.8	3.96	12.0	30.28
7	883	5.13	7.71	27.9	7.52	13.3	44.03
8	868	7.80	7.26	26.5	6.63	11.7	37.74
9	1059	5.61	9.08	27.8	8.07	13.3	46.52
10	1034	8.67	8.77	26.2	7.47	11.4	40.47
11	1048	6.98	9.31	28.6	9.04	13.3	50.68
12	1022	4.79	7.63	25.1	5.62	11.3	33.46
13	1127	3.96	8.88	23.8	6.36	11.2	35.17
14	1104	5.77	10.5	26.2	9.62	13.0	49.25
15	1109	6.84	9.80	24.4	7.43	11.5	39.71
16	1091	4.64	10.0	25.6	8.02	13.2	44.84

3.4. Input evaluation

Before starting the learning of the control module, input descriptors should be computed to characterize input images. The homogeneity histogram [25] of the input image has been selected to take in account regions with uniform shade (e.g., vehicle paintings) as well as homogeneous texture (e.g., road surface) (Figure 6).

The homogeneity measure combines two local criteria: the local contrast σ_{ij} in a $d \times d$ ($d = 5$) window centered on the current pixel (i, j), and a gradient measure e_{ij} in another $t \times t$ ($t = 3$) window:

$$\sigma_{ij} = \sqrt{\frac{1}{d^2} \sum_{p=i-(d-1)/2}^{p=i+(d-1)/2} \sum_{q=j-(d-1)/2}^{q=j+(d-1)/2} \left(g_{pq} - \mu_{ij}\right)^2}, \tag{8}$$

where μ_{ij} is the average of the gray levels computed inside the same window by

$$\mu_{ij} = \frac{1}{d^2} \sum_{p=i-(d-1)/2}^{p=i+(d-1)/2} \sum_{q=j-(d-1)/2}^{q=j+(d-1)/2} g_{pq}. \tag{9}$$

The measure of intensity variations e_{ij} around a pixel (i, j) is computed by Sobel operator:

$$e_{ij} = \sqrt{G_x^2 + G_y^2}, \tag{10}$$

where G_x and G_y are the components of the gradient at pixel (i, j) in x and y directions, respectively.

These measures are normalized using $V_{ij} = \sigma_{ij}/\max \sigma_{ij}$ and $E_{ij} = e_{ij}/\max e_{ij}$. The homogeneity measure is finally

Table 5: Experiment matrix-Rechschaffner design: averaged outputs.

Trial	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	r (%)
1	−1	−1	−1	−1	−1	−1	−1	−1	35.47
2	−1	1	1	1	1	1	1	1	43.50
3	1	−1	1	1	1	1	1	1	45.13
4	1	1	−1	1	1	1	1	1	45.68
5	1	1	1	−1	1	1	1	1	45.01
6	1	1	1	1	−1	1	1	1	44.99
7	1	1	1	1	1	−1	1	1	47.73
8	1	1	1	1	1	1	−1	1	46.53
9	1	1	1	1	1	1	1	−1	33.46
10	1	1	−1	−1	−1	−1	−1	−1	40.99
11	1	−1	1	−1	−1	−1	−1	−1	41.12
12	1	−1	−1	1	−1	−1	−1	−1	40.98
13	1	−1	−1	−1	1	−1	−1	−1	41.69
14	1	−1	−1	−1	−1	1	−1	−1	34.56
15	1	−1	−1	−1	−1	−1	1	−1	40.87
16	1	−1	−1	−1	−1	−1	−1	1	51.06
17	−1	1	1	−1	−1	−1	−1	−1	38.03
18	−1	1	−1	1	−1	−1	−1	−1	37.75
19	−1	1	−1	−1	1	−1	−1	−1	38.19
20	−1	1	−1	−1	−1	1	−1	−1	30.98
21	−1	1	−1	−1	−1	−1	1	−1	37.82
22	−1	1	−1	−1	−1	−1	−1	1	47.78
23	−1	−1	1	1	−1	−1	−1	−1	33.89
24	−1	−1	1	−1	1	−1	−1	−1	35.12
25	−1	−1	1	−1	−1	1	−1	−1	27.49
26	−1	−1	1	−1	−1	−1	1	−1	34.85
27	−1	−1	1	−1	−1	−1	−1	1	43.81
28	−1	−1	−1	1	1	−1	−1	−1	34.42
29	−1	−1	−1	1	−1	1	−1	−1	27.36
30	−1	−1	−1	1	−1	−1	1	−1	34.27
31	−1	−1	−1	1	−1	−1	−1	1	43.29
32	−1	−1	−1	−1	1	1	−1	−1	28.19
33	−1	−1	−1	−1	1	−1	1	−1	35.38
34	−1	−1	−1	−1	1	−1	−1	1	44.66
35	−1	−1	−1	−1	−1	1	1	−1	27.89
36	−1	−1	−1	−1	−1	1	−1	1	41.10
37	−1	−1	−1	−1	−1	−1	1	1	44.26

Table 6: Factor influence and interactions: Rechschaffner design.

X_1	2.06	—	—	—	—	—	—	—
X_2	−0.92	0.74	—	—	—	—	—	—
X_3	−0.05	0.08	−0.23	—	—	—	—	—
X_4	0.07	0.16	0.03	−0.21	—	—	—	—
X_5	−0.04	−0.01	0.06	0.03	0.08	—	—	—
X_6	0.04	0.05	0.01	0.13	0.06	−2.47	—	—
X_7	−0.21	−0.07	0.03	0.03	0.03	0.02	−0.34	—
X_8	−0.04	0.05	−0.11	−0.09	−0.03	0.95	−0.03	5.30
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8

Table 7: Experiment matrix-Quadratic design: averaged outputs.

Trial	X_1	X_6	X_8	r (%)
1	−1	−1	−1	41.22
2	0	−1	−1	44.74
3	+1	−1	−1	45.26
4	−1	0	−1	34.22
5	0	0	−1	38.00
6	+1	0	−1	39.26
7	−1	+1	−1	34.16
8	0	+1	−1	37.90
9	+1	+1	−1	39.10
10	−1	−1	0	47.97
11	0	−1	0	50.96
12	+1	−1	0	51.67
13	−1	0	0	43.23
14	0	0	0	45.56
15	+1	0	0	47.45
16	−1	+1	0	43.10
17	0	+1	0	45.58
18	+1	+1	0	47.33
19	−1	−1	+1	50.40
20	0	−1	+1	53.90
21	+1	−1	+1	55.02
22	−1	0	+1	46.85
23	0	0	+1	51.00
24	+1	0	+1	52.93
25	−1	+1	+1	46.87
26	0	+1	+1	51.13
27	+1	+1	+1	52.83

(a) Input image.

(b) Local contrast image (V_{ij}).

(c) Gradient image (E_{ij}).

(d) Homogeneity image (H_{ij}).

Figure 6: Homogeneity measure.

expressed by

$$H_{ij} = 1 - E_{ij} \cdot V_{ij}. \tag{11}$$

Each pixel (i, j) with a H_{ij} measure verifying $H_{ij} > 0.95$ is taken into account in the histogram computed on the 256 gray levels of the input image.

3.5. IPC control

We have used a simple multilayer perceptron as a control module. It is composed of 256 input neurons (homogeneity histogram levels over the 256 gray levels), 48 hidden neurons (maximum speed convergence during the learning), and output neurons corresponding to the tuning parameters of the IPC. One version of the neural network computes only the significant parameters (NN3) and the other version computes all tuning parameters (NN8).

Table 8: Neural network programing.

Neural network	Parameter MAE (%)	Covering rate Absolute error
	Learning	
NN3	1.4 %	8.06
NN8	0.8 %	3.55
	Test	
NN3	23.7	9.53
NN8	28.6	13.17

Table 9: Comparison of several tuning methods.

	Averaged covering rate (%)	Computing cost
Static	34.84	0
NN8	45.17	Histogram
NN3	49.64	Histogram
Factorial design	50.68	16 trials
Rechs. design	51.06	37 trials
Quadratic design	55.02	27 trials
SPL8	58.34	100 trials
SPL3	59.17	60 trials

During the learning step carried out on 75% of the input images, the decrease of the mean absolute error (MAE) is observed between optimal parameters and those computed by the network (convergence over 400 iterations) (Table 8). It is essential to control on the remaining 25% test images that the tuning parameters computed by the network not only are close enough to the optimal values, but also produce really good results at the IPC output; that is to say, line groups are well detected. We can note that the neural network only based on significant tuning parameters (NN3) is the most robust during the test step although errors are larger during the learning step.

In (Table 9), we compare the output image quality (covering rates) averaged on the set of test images, depending on the tuning process adopted. Eight modes have been tested: a static one (without adaptive tuning, that is to say, an average tuning resulting from the design of experiments), three modes based on the best trial of the design of experiments presented previously, two modes for the neural networks using only significant parameters (NN3) or all tuning parameters (NN8) and two modes for the optimal tuning of significant parameters (SPL3), or all parameters (SPL8) using simplex algorithm.

In static mode, the covering rate is small. When the best trial obtained from a design of experiments is used for the tuning, the results are better. However, this method cannot be applied in real-time situations. The results obtained with the simplex method are naturally optimal but the price for that is the prohibitive time required for the parameter space exploration.

Finally, the neural networks provide high values, especially the 3 output network, with a negligible computing cost ($\approx$ computation of the input image descriptors). We have intentionally mentioned in this table the results obtained for an eight-parameter tuning: we can easily verify that the tuning of the 5 parameters considered little significant by the design of experiments is useless.

4. CONCLUSION

These promising results obtained in the context of an image processing chain (IPC) dedicated to road obstacle detection highlight the interest of the experimental approach for the adaptive tuning of an IPC. The main reasons for this efficiency are simple: unlike previous work, the IPC is globally optimized, from a great number of real test images and by adapting image processing to each input image. We are currently testing this approach on other applications in which the image typology, image processing operators, and data evaluation criteria for inputs as well as outputs are also specific. This should enable us to unify and generalize this methodology for better IPC performance.

ACKNOWLEDGMENT

This research program has been supported by the French PREDIT Program and by Europe FSE grant.

REFERENCES

[1] R. M. Haralick, "Performance characterization protocol in computer vision," in *Proceedings of the ARPA Image Understanding Workshop*, vol. I, pp. 667–673, Monterey, Calif, USA, November 1994.

[2] P. Courtney, N. Thacker, and A. Clark, "Algorithmic modeling for performance evaluation," in *Proceedings of the ECCV Workshop on Performance Characteristics of Vision Algorithms*, p. 13, Cambridge, UK, April 1996.

[3] W. Forstner, "10 pros and cons against performance characterization of vision algorithms," in *Proceedings of the ECCV Workshop on Performance Characteristics of Vision Algorithms*, Cambridge, UK, April 1996.

[4] K. W. Bowyer and P. J. Phillips, *Empirical Evaluation Techniques in Computer Vision*, Wiley-IEEE Computer Society Press, Los Alamitos, Calif, USA, 1998.

[5] P. Meer, B. Matei, and K. Cho, "Input guided performance evaluation," in *Theoretical Foundations of Computer Vision (TFCV '98)*, pp. 115–124, Dagstuhl, Germany, March 1998.

[6] I. T. Phillips and A. K. Chhabra, "Empirical performance evaluation of graphics recognition systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 9, pp. 849–870, 1999.

[7] J. Blanc-Talon and V. Ropert, "Evaluation des chaînes de traitement d'images," *Revue Scientifique et Technique de la Défense*, no. 46, pp. 29–38, 2000.

[8] S. Philipp-Foliguet, *Evaluation de la segmentation*, ETIS, Cergy-Pontoise, France, 2001.

[9] N. Sebe, Q. Tian, E. Loupias, M. S. Lew, and T. S. Huang, "Evaluation of salient point techniques," in *Proceedings of International Conference on Image and Video Retrieval (CIVR '02)*, vol. 2383, pp. 367–377, London, UK, July 2002.

[10] P. L. Rosin and E. Ioannidis, "Evaluation of global image thresholding for change detection," *Pattern Recognition Letters*, vol. 24, no. 14, pp. 2345–2356, 2003.

[11] Y. Yitzhaky and E. Peli, "A method for objective edge detection evaluation and detector parameter selection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 1027–1033, 2003.

[12] V. Ropert, "Proposition d'une architecture de contrôle pour un système de vision," Thèse de l'Université René Descartes (Paris 6), Paris, France, Décembre 2001.

[13] I. Levner and V. Bulitko, "Machine learning for adaptive image interpretation," in *Proceedings of the 16th Innovative Applications of Artificial Intelligence Conference (IAAI '04)*, pp. 870–876, San Jose, Calif, USA, July 2004.

[14] P. Schimmerling, J.-C. Sisson, and A. Zaïdi, *Pratique des Plans d'Expériences*, Lavoisier Tec & Doc, Paris, France, 1998.

[15] S. Treuillet, "Analyse de l'influence des paramètres d'une chaîne de traitements d'images par un plan d'expériences," in *19e colloque GRETSI sur le traitement du signal et des images (GRETSI '03)*, Paris, France, September 2003.

[16] S. Treuillet, D. Driouchi, and P. Ribereau, "Ajustement des paramètres d'une chaîne de traitement d'images par un plan d'expériences fractionnaire 2^{k-p}," *Traitement du Signal*, vol. 21, no. 2, pp. 141–155, 2004.

[17] M. H. Wright, "The nelder-mead simplex method: recent theory and practice," in *Proceedings of the 16th International Symposium on Mathematical Programming (ISMP '97)*, Lausanne, Switzerland, August 1997.

[18] A. Domingues, Y. Lucas, D. Baudrier, and P. Marché, "Détection et suivi d'objets en temps réel par un système embarqué multi capteurs," in *Proceedings of the 18th Symposium GRETSI on Signal and Image Processing (GRETSI '01)*, Toulouse, France, Septembre 2001.

[19] A. Domingues, "Système embarqué multicapteurs pour la détection d'obstacles routiers—Développement du prototype et réglage automatique de la chaîne de traitement d'images," Thèse de l'Université d'Orléans, Orléans, France, Juillet 2004.

[20] Y. Lucas, A. Domingues, M. Boubal, and P. Marché, "Système de vision embarqué pour la détection d'obstacles routiers," *Techniques de l'Ingénieur—Recherche & Innovation*, p. 9, 2005, IN-24.

[21] P. Lamaty, "Opérateurs de niveau intermédiaire pour le traitement temps réel des images," Thèse de Doctorat, Thèse de l'Université de Cergy-Pontoise, Cergy-Pontoise, France, 2000.

[22] Y. Lucas, A. Domingues, D. Driouchi, and P. Marché, "Modeling, evaluation and control of a road image processing chain," in *Proceedings of the 14th Scandinavian Conference on Image Analysis (SCIA '05)*, vol. 3540, pp. 1076–1085, Joensuu, Finland, June 2005.

[23] A. Fries and W. G. Hunter, "Minimum aberration 2^{k-p} designs," *Technometrics*, vol. 22, no. 4, pp. 601–608, 1980.

[24] R. L. Rechtschaffner, "Saturated fractions of 2n and 3n factorial designs," *Technometrics*, vol. 9, pp. 569–575, 1967.

[25] H.-D. Cheng and Y. Sun, "A hierarchical approach to color image segmentation using homogeneity," *IEEE Transactions on Image Processing*, vol. 9, no. 12, pp. 2071–2082, 2000.

Yves Lucas received the Master's degree in discrete mathematics from Lyon 1 University, France, in 1988 and the DEA in computer science and automatic control from the Applied Sciences National Institute of Lyon, France, in 1989. He focused on the field of CAD-based vision system programing and obtained the Ph.D. degree from INSA Lyon, France, in 1993. Then, he joined the Orleans University, France, where he is currently in charge of the Vision Group at the Vision and Robotics laboratory, which is centered on 3D object reconstruction and color image segmentation. His research interests include vision system learning and tuning, as well as pattern recognition and image analysis for medical, industrial, and robotic applications.

Antonio Domingues received the Master's degree in electronic systems for vision and robotics, from Clermont-Ferrand University, France, in 1999. He joined the Vision and Robotics laboratory, Bourges, France, in 2001 and worked in relation with MBDA company on the SPINE project, centered on an embedded road obstacle detection system for intelligent airbag control based on a vision system. He received in 2004 the Ph.D. degree from Orleans University, France, in the field of industrial technology and currently works in a software engineering company in Paris, France.

Driss Driouchi received the Master's degree both in pure mathematics and in mathematical engineering at Paul Sabatier University, Toulouse, France, in 1998 and 1999. He obtained in 2000 the DEA in the field of statistics at Pierre and Marie Curie Paris 6 University, France, where he worked in the team of Professor Paul Deheuvels and received the Ph.D. degree in statistics in 2004. He is currently an Assistant Professor at Mohamed I University, Nador, Morocco. His research interests are in the field of theoretical and practical problems about the design of experiments.

Sylvie Treuillet received the Dipl. Ing. degree in electronic engineering, from the University of Clermont-Ferrand, France, in 1988. She started working as a Research Engineer in a private company and developed an imagery system for chromosomes classification. In 1990, she received a fellowship for a study about multisensory data fusion for obstacle detection and tracking on motorways and obtained the Ph.D. degree in 1993. Since 1993, she has been a Teacher and Researcher in Polytech' Orleans Advanced Engineering School, France. Her research activity is mainly dedicated to the various aspects of image analysis, mainly for 3D object reconstruction and tracking in biomedical or industrial applications.