



DATA DIMENSIONALITY REDUCTION FOR NEURAL BASED CLASSIFICATION OF OPTICAL SURFACES DEFECTS

Matthieu Voiry ^{1&2)}, Kurosh Madani ¹⁾, Véronique Amarger ¹⁾, Joël Bernier ²⁾

¹⁾ Image, Signal and Intelligent Systems Laboratory (LISSI / EA 3956), Senart Institute of Technology, University PARIS XII, Av. Pierre Point, F-77127 Lieusaint, France,
{voiry ; madani ; amarger}@univ-paris12.fr, <http://www.univ-paris12.fr/>

²⁾ SAGEM REOSC
Avenue de la Tour Maury, Saint Pierre du Perray, 91280, France
{mathieu.voiry or joel.bernier}@sagem.com

Abstract: A major step for high-quality optical surfaces faults diagnosis concerns scratches and digs defects characterization in products. This challenging operation is very important since it is directly linked with the produced optical component's quality. A classification phase is mandatory to complete optical devices diagnosis since a number of correctable defects are usually present beside the potential "abiding" ones. Unfortunately relevant data extracted from raw image during defects detection phase are high dimensional. This can have harmful effect on the behaviors of artificial neural networks which are suitable to perform such a challenging classification. Reducing data dimension to a smaller value can decrease the problems related to high dimensionality. In this paper we compare different techniques which permit dimensionality reduction and evaluate their impact on classification tasks performances.

Keywords: Computer Aided Diagnosis Systems (CADS), Artificial Intelligent systems, Industrial applications, Artificial Neural Network, Dimensionality Reduction, Curvilinear Component Analysis (CCA), Curvilinear Distance Analysis (CDA), Self Organizing Maps (SOM).

1. INTRODUCTION

We are involved in fault diagnosis of optical devices in industrial environment. In fact, classification of detected faults is among chief phases for succeeding in such diagnosis. Aesthetic flaws, shaped during different manufacturing steps, could provoke harmful effects on optical devices' functional specificities, as well as on their optical performances by generating undesirable scatter light, which could seriously degrade the expected optical features. Taking into account the above-mentioned points, a reliable diagnosis of these defects in high-quality optical devices becomes a crucial task to ensure products' nominal specification and to enhance the production quality. Moreover, the diagnosis of these defects is strongly motivated by manufacturing process correction requirements in order to guarantee mass production (repetitive) quality with the aim of maintaining acceptable production yield.

Unfortunately, detecting and measuring such defects is still a challenging dilemma in production conditions and the few available automatic control solutions remain ineffective. That's why, in most of

cases, the diagnosis is performed on the basis of a human expert based visual inspection of the whole production. However, this usual solution suffers from several acute restrictions related to human operator's intrinsic limitations (reduced sensitivity for very small defects, detection exhaustiveness alteration due to attentiveness shrinkage, operator's tiredness and weariness due to repetitive nature of fault detection and fault diagnosis tasks).

To overcome these problems we have proposed a detection approach based on Nomarski's microscopy issued imaging [1] [2]. This method provides robust detection and reliable measurement of outward defects, making plausible a fully automatic inspection of optical products. However, the above-mentioned detection process should be completed by an automatic classification system in order to discriminate the "false" defects (correctable defects) from "true" (permanent) ones. In fact, because of industrial environment, a number of correctable defects (like dusts or cleaning marks) are usually present beside the potential abiding defects. That is why the association of a faults' classification system to the aforementioned detection module is a

foremost supply to ensure a reliable diagnosis. In a precedent paper [3], we proposed a method to extract relevant data from raw Nomarski images. In the aim of effectively classify these descriptors, neural network based techniques seem appropriate because they have shown many attractive features in complex pattern recognition and classification tasks [4] [5]. But we are dealing with high dimensional data (13 and more components vectors) so behaviors of a number of these algorithms could be affected. To avoid this problem we are investigating different dimension reduction techniques for achieving better classification (in terms of performance and processing time).

This paper is organized as follows: in the next section, motivations for data dimensionality reduction and also Self Organizing Maps (SOM), Component Analysis (CCA) and Curvilinear Distance Analysis (CDA), three techniques to perform this task are introduced. These techniques have been tested using an experimental protocol presented in Section 3. The Section 4 deals with experiments results: first a comparison of data projections quality and an analysis of their possible impact on classification tasks are carried out. Secondly this impact is studied on a real classification problem involving Multilayer Perceptron artificial neural network, and the obtained results are discussed. The Section 5 concludes this work and gives a number of perspectives.

2. DATA DIMENSIONALITY REDUCTION TECHNIQUES

It can be found in literature, lot of examples using various dimension reduction techniques (linear or not) as a preliminary step before more refined processing: Principal Component Analysis (PCA) [6], Self Organizing Maps (SOM) [7;8], Curvilinear (CCA) [9;10] or (CDA) [11].

Dealing with high-dimensional data indeed poses problems, known as “curse of dimensionality” [10]. First sample number required to reach a predefined level of precision in approximation tasks increases exponentially with dimension. Thus, intuitively, the sample number needed to properly learn high-dimensional data becomes quickly much too large to be collected by real systems, when dimension of data increases. Moreover surprising phenomena appear when working in high dimension [12] : for example, distances variance between vectors remains fixed while its average increases with the space dimension, and Gaussian kernel local properties are also lost. These last points explain that behaviour of a number of artificial neural network algorithms could be affected while dealing with high-dimensional data. Fortunately, most real-world

problem data are located in a manifold of dimension p much smaller than its raw dimension. Reducing data dimensionality to a smaller value can therefore decrease the problems related to high dimension.

2.1 SELF ORGANIZING MAPS

Self-Organizing Map is a classical method originally proposed by Kohonen [13]. The algorithm projects multidimensional feature space into a low-dimensional presentation. Typically a SOM consists of a two dimensional grid of neurons. A vector of features is associated with each neuron. During the training phase, these vectors are tuned to represent the training data under constraint of neighbourhood conservation. Similar data are projected to the same or nearby neurons in the SOM, while different ones are mapped to neurons located further from each other, resulting in a clustering data. Thus SOM is an efficient tool for quantizing the data’s space and projecting this space onto a low-dimensional space, while conserving its topology. SOM is often used in industrial engineering [14], [15] to characterize high-dimensional data or to carry out classification tasks. Unfortunately it suffers of major drawbacks: first the configuration of the topology is static and should be fixed a priori (what is efficient only for little values of projection subspace dimension), moreover the method defines only a discrete nonlinear subspace and finally algorithm is computationally too expensive to be practically applied for projection space dimension higher than three.

2.2 CURVILINEAR COMPONENTS ANALYSIS

The goal of this technique proposed by Demartines [16] is to reproduce the topology of a n -dimension original space in a new p -dimension space (where $p < n$) without fixing any configuration of the topology. To do so, a criterion characterizing the differences between original and projected space topologies is processed:

$$E_{CCA} = \frac{1}{2} \sum_i \sum_{j \neq i} (d_{ij}^n - d_{ij}^p)^2 F(d_{ij}^p) \quad (1)$$

where d_{ij}^n (respectively d_{ij}^p) is the Euclidean distance between vectors x_i and x_j of considered distribution in original space (resp. in projected space), and F is a decreasing function which favors local topology with respect to the global topology. This energy function is minimized by stochastic gradient descent [17]:

$$\forall i \neq j, \Delta x_i^p = \alpha(t) \frac{d_{ij}^n - d_{ij}^p}{d_{ij}^p} u(\lambda(t) - d_{ij}^p) (x_i^p - x_j^p) \quad (2)$$

where $\alpha: \mathbb{R}^+ \rightarrow [0;1]$ is a decreasing function representing a learning parameter, and $\lambda: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is a decreasing function too, representing a neighborhood factor. CCA provides also a similar method to project, in continuous way, new points in the original space onto the projected space, using the knowledge of already projected vectors.

2.3 CURVILINEAR DISTANCE ANALYSIS

Since CCA encounters difficulties with unfolding of very non-linear manifolds, an evolution called CDA has been proposed [18]. It involves curvilinear distances (in order to better approximate geodesic distances on the considered manifold) instead of Euclidean ones. Curvilinear distances are processed in two steps way. First is built a graph between vectors by considering k-NN, ε , or other neighborhood, weighted by Euclidean distance between adjacent nodes. Then the curvilinear distance between two vectors is computed as the minimal distance between these vectors in the graph using Dijkstra's algorithm. Finally the original CCA algorithm is applied using processed curvilinear distances. This algorithm allows dealing with very non-linear manifolds and is much more robust against the choices of α and λ functions.

3. EXPERIMENTAL VALIDATION PROTOCOL

In order to obtain exploitable data for a classification scheme, we first needed to extract relevant information of raw Nomarski's microscopy issued images. We proposed to proceed in two steps [2]: first a detected items' images extraction phase and then an appropriated coding of the extracted images. The image associated to a given detected item is constructed considering a stripe of ten pixels around its pixels. Thus the obtained image gives an isolated (from other items) representation of the defect (e.g. depicts the defect in its immediate environment). Fig. 1 gives four examples of detected items' images using the aforementioned technique. It shows different characteristic items which could be found on optical device in industrial environment. The information contained in such images is highly redundant. Furthermore, the generated images don't have necessarily the same dimension (typically this dimension can turn out to be thousand times as high). That is why these raw data (images) cannot be directly processed and has to be appropriately

encoded.

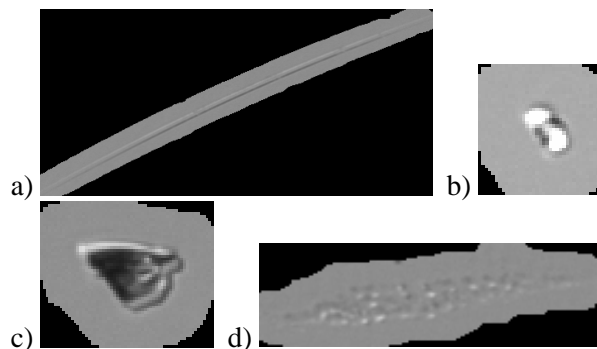


Fig. 1 – Images of characteristic items: a) scratch; b) dig; c) dust; d) cleaning marks.

This is done using a set of Fourier-Mellin transform issued invariants described bellow. The Fourier-Mellin transform of a function $f(r; \theta)$, in polar coordinates, is given by relation (1), with $q \in \mathbb{Z}$, $s = \sigma + ip \in \mathbb{C}$ (see[19]):

$$M_f(q; s) = \int_{r=0}^{\infty} \int_{\theta=0}^{2\pi} r^{s-1} \exp(-iq\theta) f(r; \theta) dr d\theta \quad (3)$$

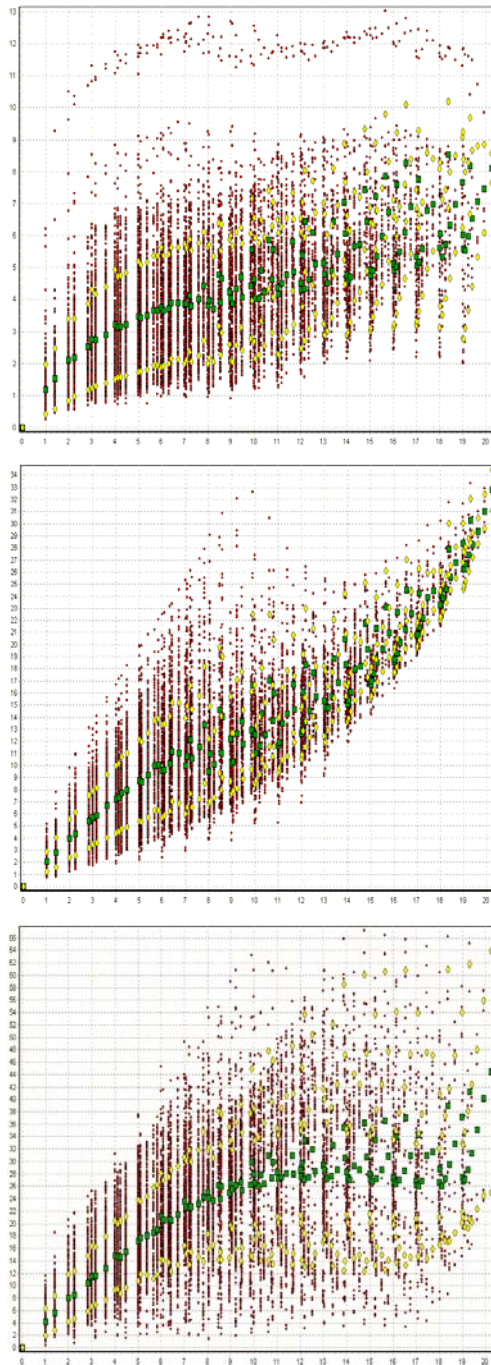
In [20], are proposed a set of features invariant on geometric transformations:

$$I_f(q; s) = M_f(q; s) [M_f(0; \sigma)]_{\sigma}^s [M_f(1; \sigma)]^q |M_f(1; \sigma)|^q \quad (4)$$

In order to validate the above-presented concepts and to provide an industrial prototype, an automatic control system has been realized. It involves an Olympus B52 microscope combined with a Corvus stage (see Fig. 2), which allows scanning an entire optical component. 50x magnification is used, that leads to microscopic 1.77mm x 1.33 mm fields and 1.28 μ m x 1.28 μ m sized pixels.



Fig. 2 – The Olympus microscope.



**Fig. 3 – dy-dx representation of the three obtained SOMs for database B (mean \square and standard deviation \diamond of dx are also represented).
Top: SOM; middle: CCA; bottom: CDA.**

The defects detection methods are implemented in the system in an efficient way, so the prototype permits to check optical devices within execution times in adequacy with production conditions (typically 10 mm x 10 mm surface measured in one minute). These facilities were used to acquire a great number of defects images. These images were coded using Fourier-Mellin transform with $\sigma=1$ and $(q, p) \in \{(q, p) / (q=0; 0 \leq p \leq P) \cup (1 \leq q \leq Q; -P \leq p \leq P)\}$

where $P=1$ and $Q=2$ (see Equation 3 and 4). Such orders of Mellin resp. Fourier spectrum (P resp. Q values) form a first compromise between the size and the quality of the representation and provides a set of 13 features for each item. Three experiments called A, B, C were carried out, using two optical devices. Table 1 shows the different parameters corresponding to these experiments. It's important to note that, in order to avoid false classes learning, items images depicting microscopic field boundaries or two (or more) different defects are discarded from used database. First, since database C is issued from a cleaned device, it's constituted with almost only "permanent" defect. And because database B came from the measurement of the same optical device but without cleaning phase, it's constituted with the same type of "permanent" defects but also with "abiding" ones.

Table 1. Description of the three experiments supplying studied databases.

Database	A	B	C
Optical Device Identifiant	1	2	2
Cleaning	NO	NO	YES
Number of studied microscopic fields	1178	605	529
Correspondant studied area	28	14	12.5
Number of items in the learning database	3865	1910	1544

In the aim of studying structure of space described by database when reducing its dimension, we perform some experiments. First a reduction of dimensionality from 13 (raw dimensionality) to 2 of the database B was performed using SOM, CCA and CDA, in order to compare projection quality of these three techniques. Then the entire database C was projected into the obtained space in order to evaluate the pertinence of dimensionality reduction for discrimination between "correctable" and "abiding" defects. Secondly, a synthetic classification task, involving aforementioned databases and Multilayer Perceptron artificial neural network, was carried out with and without dimensionality reduction phase with the aim to demonstrate usefulness of such pre-processing phase. Finally, we validate the previous results by studying the impact of different dimensionality reduction on a real problem: an expert was asked to define two different real classes of defects and a MLP was used to discriminate between these two classes.

4. EXPERIMENTAL RESULTS AND ANALYSIS

4.1 QUALITY OF PROJECTION

Dimensionality reduction has been performed using the three aforementioned techniques, SOM, ACC and CDA on database B. To compare the results of the three experiments, the 2-D projections issued from CCA and CDA were processed by a SOM, using the same shape of grid (20x8) as in the SOM experiment. An important point is that SOM is just used, in these two past cases, to perform a quantization and not for dimension reduction, since it works on a 2 dimension space. Therefore, we can directly compare dimension reduction ability of the different techniques by comparing these maps with map obtained by applying SOM's algorithm on raw data. The quality evaluation of non-linear projection of the data space onto the neurons grid space is performed by studying, for each pair of neurons, the dx distance between these two neurons in the data space, versus the dy distance between these two neurons in the grid space [21]. For each couple of neurons $(i; j)$ we draw a point $(dy(i, j); dx(i, j))$ where $dx(i, j) = \|\vec{x}_i - \vec{x}_j\|$ and $dy(i, j) = \|\vec{y}_i - \vec{y}_j\|$. \vec{x}_k (resp. \vec{y}_k) is the vector of features corresponding to the k-th neuron in the data space (resp. in the grid space). If the topology of the data space is not well respected, dx is not related to dy and we obtain a diffuse cloud of points. On the contrary, if neurons organization is correct, the drawn points are almost arranged along a straight line.

First, in Fig. 3, cloud of points is more diffuse for SOM than in the case of CCA, and the curve constituted by dx averages for each dy less uniformly monotonic. It reveals the fact that the CCA performs better than SOM, while approximately the same quantity is minimized. The cloud obtained for CDA is quite different because dy is related to curvilinear distance and not Euclidean one. The figure is however the same as for CCA for little dy value, because in these cases Euclidean distance is a good approximate of curvilinear one (and therefore distribution is locally linear).

4.2 ANALYSIS OF POSSIBLE IMPACT ON CLASSIFICATION TASKS

We now consider the database C (only "permanent" defects) and project its items onto the three previously obtained SOMs. We perform also an equivalent experiment on raw data (13-dimension), using k-means algorithm with $k=20 \times 8=160$. Since k-means algorithm has identical behaviour as SOM, except concerning

neighbourhood constraints, it has the same effect on projected items distribution but doesn't allow visual representation. Projected items distribution after SOM (Fig. 3), CCA (Fig. 4) and CDA (Fig. 5) dimension reduction are studied. In these figures, the equalized grey level depicts the number of projected items for each SOM's cell (this number is also reported in the cell).

In table 2 are reported some characteristic values of distributions "homogeneity": entropy and standard deviation of projected items number in each cell; number of empty or quasi-empty cells (less than three projected items). Maps and numerical measurements for SOM and CCA are comparable and therefore these techniques are equivalent for the considered problem. CCA is however easier to perform (no a priori knowledge or difficult choice) and provide more information (continuous projection). CDA offers the same advantages as CCA, but it seems to be more appropriate for pre-processing before classification. Corresponding map depicts indeed more specific "areas" for database C projected defects.

Table 2. Different measurements characterizing the projections distribution of database C items (permanent defects).

Dimensionality reduction technique	None	SOM	CCA	CDA
Standard-deviation of distribution	8.72	5.78	5.72	7.04
Entropy of defects distribution	2.055	2.114	2.121	2.088
Number of empty cells	15	9	5	7
Number of cells with less than 3 defects	30	26	20	32

This intuition is confirmed by numerical measurements: entropy is lower than in SOM and CCA cases (better organization), standard deviation is higher (better contrast between full and empty areas) and there are more quasi-empty cells. We think that this organization is a foremost guarantee for the dimension reduction to allow a better classification. We can also remark that results obtained with CDA are fairly similar as those with raw data; it shows that little information is lost while reducing dimensionality.

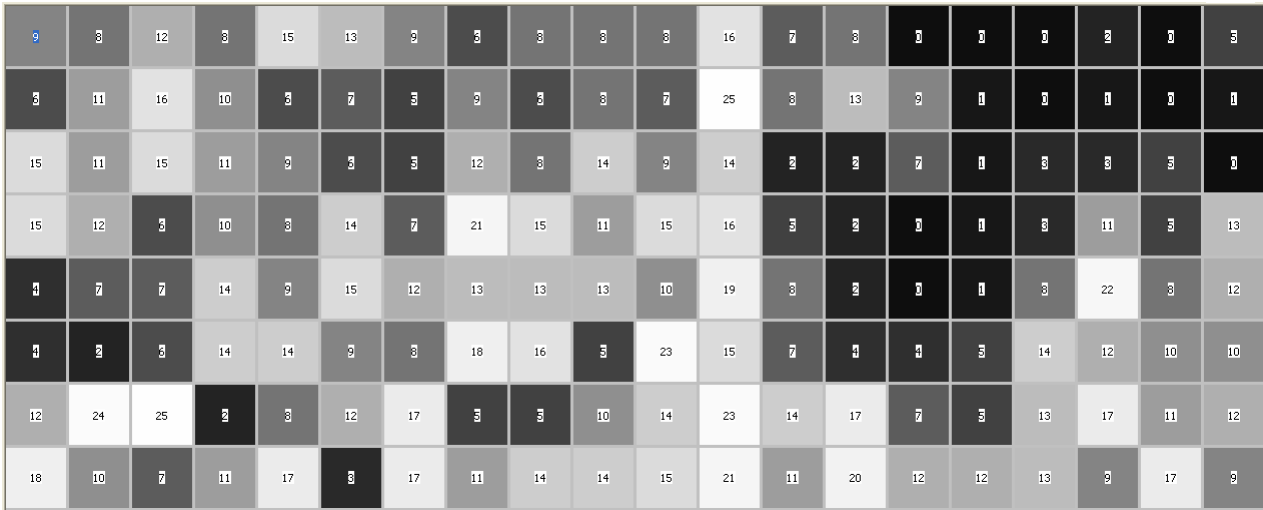


Fig. 4. Distribution of projected items in SOM map. (SOM reduction dimension)

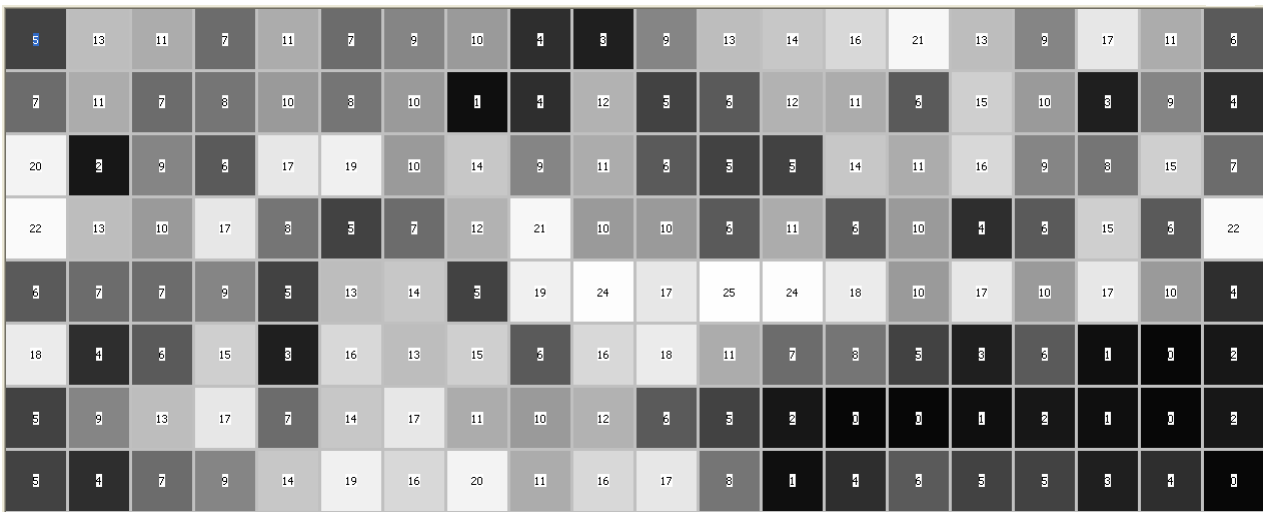


Fig. 5. Distribution of projected items in SOM map. (CCA reduction dimension)

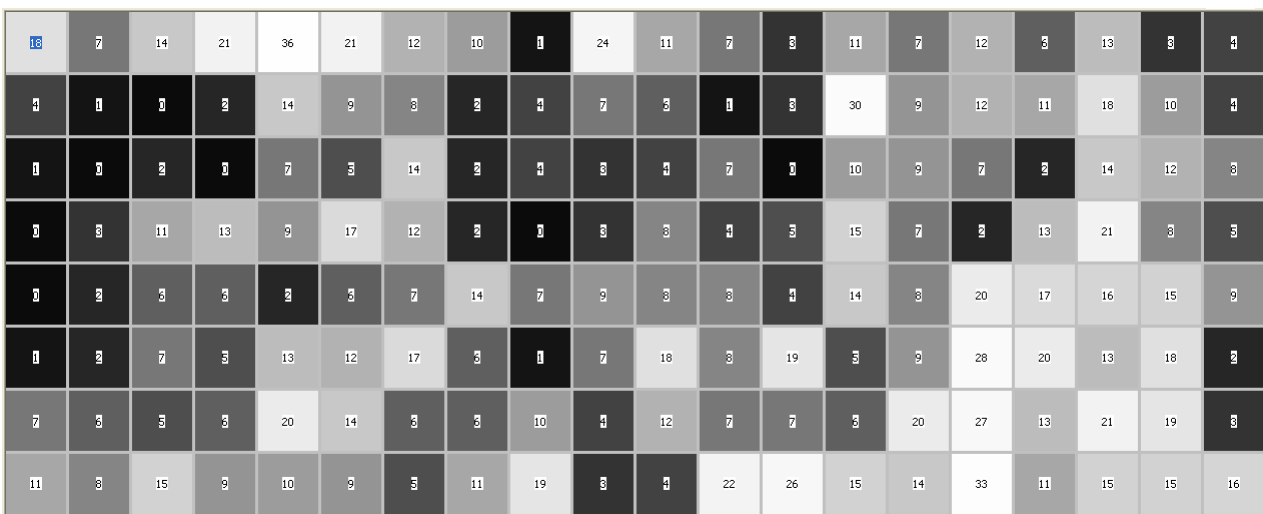


Fig. 6. Distribution of projected items in SOM map. (CDA reduction dimension)

4.3 EXPERIMENTATION ON A REAL PROBLEM

In order to evaluate the benefits of using dimensionality reduction in our diagnosis system, we performed a series of experiments involving a real classification task. Items of the databases A and B were labelled by an expert with two different labels: “dust” (class 1) and “other defects” (class 2). Considering our final goal, this is the most important distinction which the diagnosis system must be able to do. Table 3 shows items repartition between the two defined classes.

Table 3. Description of real classification databases.

Database	Coming from database	Total number of items	Label 1 items	Label 2 items
1	A	3865	275	3590
2	B	1910	184	1726

Using these databases, a number of experiments were carried out, in accordance with a same procedure. It involved a multilayer perceptron with n input neurons, 35 neurons in one hidden layer, and 2 output neurons (n -35-2 MLP). First this artificial neural network was trained for discrimination task between classes 1 and 2, using database B. This training phase used BFGS (Broyden, Fletcher, Goldfarb, and Shanno) with Bayesian regularization algorithm, and was achieved 5 times. Subsequently, the generalization ability of obtained neural network was processed using database A. Since database A and B issued from different optical devices, such generalization results are significant. Following this procedure, 28 different experiments were conducted with the aim of studying the global classification performance and the impact of SOM, ACC and CDA dimensionality reduction on this performance. First experiment used original Fourier-Mellin issued features (13-dimensional), the second used 2D SOM reduced features, and the others used the original features after ACC or CDA n -dimensional space reduction (with n varying between 2 and 13).

Fig. 7 depicts classification global performances (calculated by averaging percentage of well-classified items for the 5 trainings) for the 28 different experiments (for different CDA and CCA issued data dimensionality reduction, for 2D SOM data dimensionality reduction and using raw data). Fig. 8 and Fig. 9 show the class 1 (“dust defects”) and the class 2 (“other defects”) classification performances respectively. In this two figures the standard deviation of the 5 experiments results are

also represented.

4.4 DISCUSSION

First, these experiments show that CDA and CCA generate almost the same performances (CDA is although slightly better) but outperform SOM issued results (see Table 4) for 2D data dimensionality reduction. It is due to the better quality and the continuous nature of the projection provided by these techniques. Thus, it confirms the previous results presented in this paper.

Table 4. Classification performances for the three different techniques, when reducing data dimensionality down to 2.

Dimensionality Reduction Technique	SOM	CCA	CDA
Class 1 Recognition	63.9 %	61.9 %	56.8 %
Class 2 Recognition	91.6 %	96.9 %	97.7 %
Global Performance	89.7 %	94.4 %	94.8 %

Secondly, we can remark (see Table 5) that equivalent performances can be obtained using low-dimensional data instead of unprocessed defects representations (for example using 6-dimensional CDA or 8-dimensional CCA issued representation instead of raw 13-dimensional representation).

Table 5. Classification performances with raw data, CDA and CCA issued data. For CCA and CDA, results are given for the data dimensionality which allows the better global performance.

Used Data	Raw	CCA Issued	CDA Issued
Data Dimensionality	13	8	6
Class 1 Recognition	70.2 %	63.9 %	62.0 %
Class 2 Recognition	97.2 %	97.2 %	97.7 %
Global Performance	95.3 %	94.8 %	95.1 %

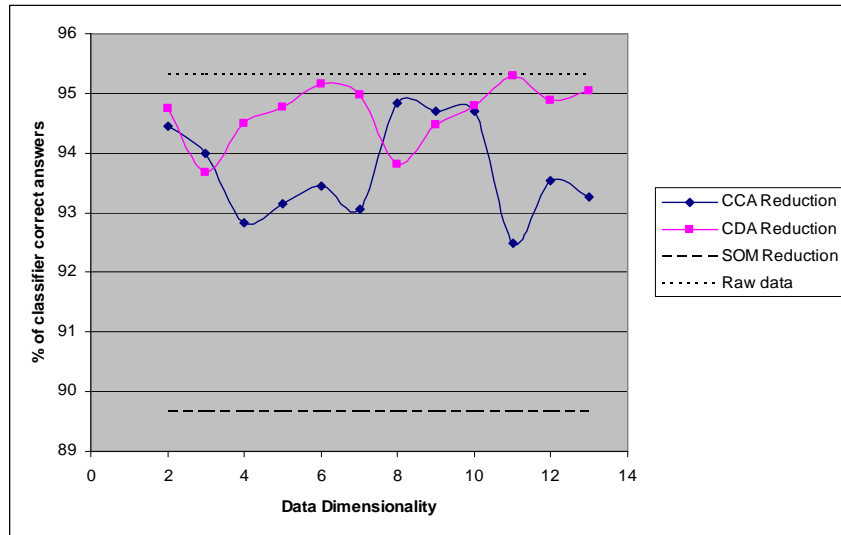


Fig. 7. Classification global performances.

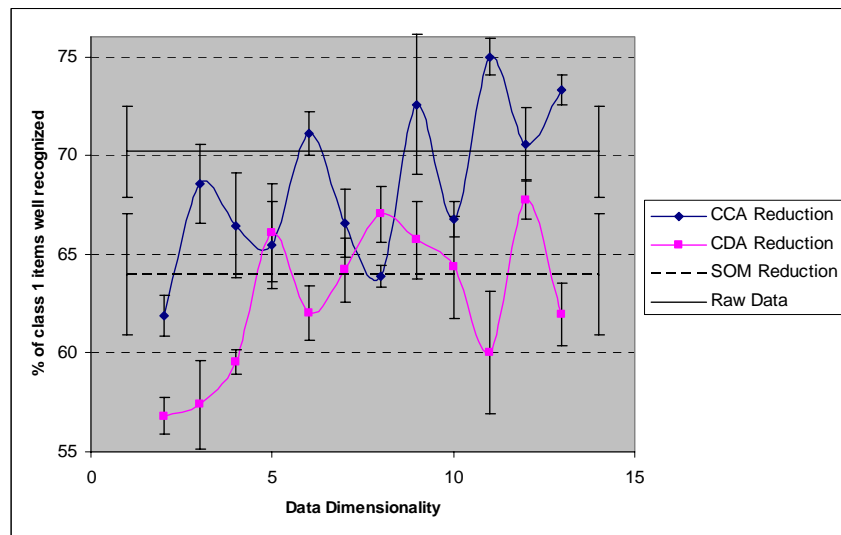


Fig. 8. Class 1 (“dust defects”) classification performances. Standard deviation of the 5 experiments results are also represented.

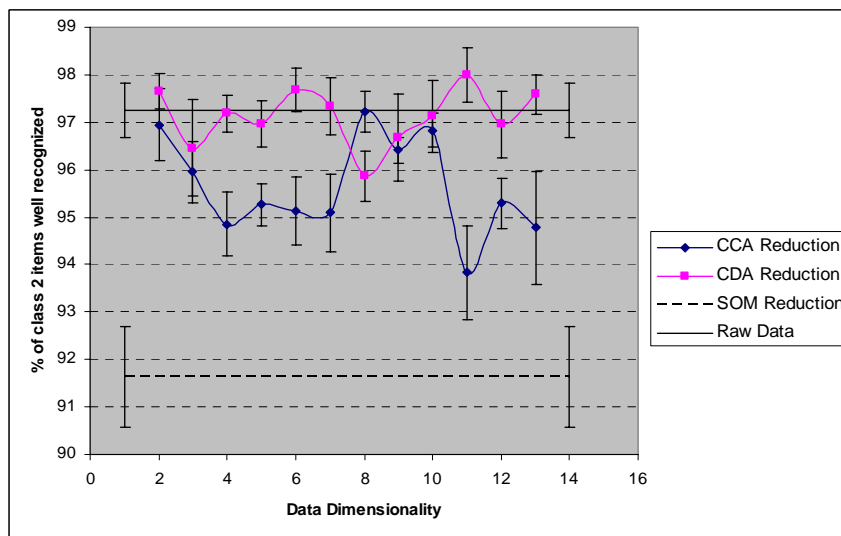


Fig. 9. Class 2 (“other defects”) classification performances. Standard deviation of the 5 experiments results are also represented.

As a consequence neural architecture complexity and therefore processing time can be saved using dimensionality reduction, while keeping performance level. Moreover, obtained scores are satisfactory: about 65% of “dust” defects are well-recognized (this can be enough for aimed application) as well as about 97% of other defects (the few 3% errors can however pose problems because every “permanent” defect has to be reported). Furthermore, we think that this significant performances difference between class 1 and class 2 recognition performances is due to the fact that class 1 is underrepresented in learning database. On another hand, obtained results show that the compromise between these two different performances can be managed by choosing an adequate data final dimension.

Finally, we can see in Fig. 8 and 9 that standard deviation values are relatively small, what’s a sign of the training phase quality. In addition, performance in class 1 recognition is relatively low when reducing data dimensionality down to two (there’s probably no more enough information to correctly classify “dust” defects). But it increases steadily until 5-dimensional CDA issued space. This result agrees with estimated items distribution intrinsic dimension between 5 and 6 (estimation used Grassberger-Procaccia [22] modified algorithm [23]). It’s interesting to note that the global performances are optimal for such a final dimensionality using CDA (about 95% of well-classified defects what is equivalent to performances obtained when using raw data). This level of performance is reached as well with 10 and 11-dimensional CDA reduction and only in these cases. This last point will be studied in greater detail in another paper.

5. CONCLUSION

A reliable diagnosis of aesthetic flaws in high-quality optical devices is a crucial task to ensure products’ nominal specification and to enhance the production quality by studying the impact of the process on such defects. To ensure a reliable diagnosis, an automatic classification system is needed in order to discriminate the “false” defects (correctable defects) from “abiding” (permanent) ones. Unfortunately relevant data extracted from raw Nomarski image during defects detection phase are high dimensional. This can have harmful effect on behaviors of artificial neural networks which are suitable to perform such a challenging classification. Reducing the dimension of the data to a smaller value can decrease the problems related to high dimension. In this paper we have compared different techniques, SOM, CCA and CDA which permit such

dimensionality reduction and evaluated their possible impact on classification tasks involving real industrial data. CDA seems to be the most suitable technique and we have demonstrated its ability to enhance performances (in terms of time and/or well-classified items) in a real classification problem. Next phase of this work will deal with classification tasks involving more classes. We want also use much more Fourier-Mellin invariants, because we think that it would improve classification performance by supplying additional information. CDA based dimensionality reduction technique would in this case be a foremost step to keep reasonable classification system’s complexity.

6. REFERENCES

- [1] M. Voiry, F. Houbre, V. Amarger, and K. Madani. Toward Surface Imperfections Diagnosis Using Optical Microscopy Imaging in Industrial Environment. *Proceedings of the Workshop IAR & ACD 2005*, Mulhouse, France 16-18 November 2005, pp. 139-144.
- [2] M. Voiry, V. Amarger, K. Madani, and F. Houbre. Combining Image Processing and Self Organizing Artificial Neural Network Based Approaches for Industrial Process Faults Clustering. *Proceedings of 13th International Multi-Conference on Advanced Computer Systems (ACS 2006)*, Miedzyzdroje, Poland 18-20 October 2006, pp. 129-138.
- [3] M. Voiry, K. Madani, V. Amarger, and F. Houbre. Toward Automatic Defects Clustering in Industrial Production Process Combining Optical Detection and Unsupervised Artificial Neural Network Techniques. *Proceedings of the 2nd International Workshop on Artificial Neural Networks and Intelligent Information Processing (ANNIP 2006)*, Setúbal, Portugal August 2006, pp. 25-34.
- [4] G. P. Zhang. Neural Networks for Classification: A Survey. *IEEE Trans. on Systems, Man, and Cybernetics - Part C: Applications and Reviews* 30 (4) (2000). p. 451-462.
- [5] M. Egmont-Petersen, D. de Ridder, and H. Handels. Image Processing with Neural Networks - A Review. *Pattern Recognition* 35 (2002). p. 2279-2301.
- [6] P. J. Grother. Karhunen Loève Feature Extraction for Neural Handwritten Character Recognition. *Proceedings of SPIE, vol. 1709, no. Applications of Artificial Neural Networks III*, pp. 155-166, 1992.
- [7] K. Boehm, W. Broll, and M. Sokolewicz. Dynamic Gesture Recognition Using Neural Networks; A Fundament for Advanced

- Interaction Construction. *Proceedings of SPIE, vol. 2177, no. Stereoscopic Displays and Virtual Reality Systems*, pp. 336-346, 1994.
- [8] J. Lampinen and E. Oja. Distortion Tolerant Pattern Recognition Based on Self-Organizing Feature Extraction. *IEEE Trans. On Neural Networks* 6 (3) (1995). p. 539-547.
- [9] S. Buchala, N. Davey, T. M. Gale, and R. J. Frank. Analysis of Linear and Nonlinear Dimensionality Reduction Methods for Gender Classification of Face Images. *International Journal of Systems Science* 14 (36) (2005). p.931-942.
- [10] M. Verleysen. Learning high-dimensional data. *NATO Advanced Research Workshop on Limitations and Future Trends in Neural Computing (LFTNC'2001)*, Siena, Italy 22-24 October 2001, pp.22-24.
- [11] M. Lennon, G. Mercier, M. C. Mouchot, and L. Hubert-Moy. Curvilinear Component Analysis for Nonlinear Dimensionality Reduction of Hyperspectral Images. *Proceedings of SPIE, vol. 4541, Image and Signal Processing for Remote Sensing VII*, pp. 157-168, 2001.
- [12] P. Demartines. *Analyse de Données par Réseaux de Neurones Auto-Organisés*. PhD Thesis Institut National Polytechnique de Grenoble, 1994.
- [13] T. Kohonen. *Self Organizing Maps*. 3rd edition ed. Berlin: Springer, 2001.
- [14] T. Kohonen, E. Oja, O. Simula, A. Visa, and J. Kangas. Engineering Applications of the Self-Organizing Maps. *Proceedings of the IEEE*, vol. 84, no. 10, pp. 1358-1384, Oct.1996.
- [15] J. Heikkonen and J. Lampinen. Building Industrial Applications with Neural Networks. *Proceedings of European Symposium on Intelligent Techniques (ESIT'99)*, Chania, Greece June 1999.
- [16] P. Demartines and J. Héroult. Vector Quantization and Projection Neural Network. *Lecture Notes in Computer Science* 686, International Workshop on Artificial Neural Networks (IWANN'93), p. 328-333.
- [17] P. Demartines and J. Héroult. CCA : "Curvilinear Component Analysis". *Proceedings of 15th workshop GRETSI (GRETSI'95)*, Juan-les-pins, France, 15 September 1995.
- [18] J. A. Lee, A. Lendasse, N. Donckers, and M. Verleysen. A Robust Nonlinear Projection Method. *Proceedings of European Symposium on Artificial Neural Networks (ESANN'2000)*.
- [19] S. Derrode. *Représentation de Formes Planes à Niveaux de Gris par Différentes Approximations de Fourier-Mellin Analytique en vue d'Indexation de Bases d'Images*. Phd Thesis Université de Rennes I, 1999.
- [20] F. Ghorbel. A Complete Invariant Description for Gray Level Images by the Harmonic Analysis Approach. *Pattern Recognition* 15 (1994). p. 1043-1051.
- [21] P. Demartines and F. Blayo. Kohonen Self-Organizing Maps: Is the Normalization Necessary? *Complex Systems* 6 (2) (1992). p. 105-123.
- [22] P. Grassberger and I. Procaccia. Measuring the strangeness of strange attractors. *Physica* 9 (1983). p. 189-208.
- [23] F. Camastra and A. Vinciarelli. Intrinsic Dimension Estimation of Data: An Approach Based on Grassberger-Procaccia's Algorithm. *Neural Processing Letters* 14 (1) (2001). p. 27-34.

Matthieu Voiry graduated in 2002 from ENSEIRB and received his MS degree in 2003 from Bordeaux 2 University.

Currently he is working toward his Ph.D. at Image, Signal and Intelligent Systems Laboratory (LISSI / EA 3956) of PARIS XII University with Prof. Kurosh Madani, in collaboration with SAGEM REOSC Company.

His main research interests are image processing and artificial neural networks in industrial computer aided diagnosis.



Prof. Kurosh Madani Received his Ph.D. degree in Electrical Engineering and Computer Sciences from University PARIS XI, Orsay, France, in 1990. From 1989 to 1990, he worked as assistant professor at Institute of Fundamental Electronics of PARIS XI University. In 1990, he joined Creteil-Senart Institute of Technology of University PARIS XII – Val de Marne, Lieusaint, France, where he worked from 1990 to 1998 as assistant professor. In 1995, he received the DHDR Doctor Habilitate degree (senior research Dr. Hab. degree) from University PARIS XII – Val de Marne. Since 1998 he works as Chair Professor in Electrical Engineering of Senart Institute of Technology of University PARIS XII. From 1992 to 2004 he has been head of Intelligence in Instrumentation and Systems Laboratory (I2S / JE 2353) located at Senart Institute of Technology. Since 2005, he is head of one of the three research groups of Image, Signal and Intelligent Systems Laboratory (LISSI / EA 3956) of PARIS XII University. He has worked on both digital and analog implementation of processors arrays for image processing, electro-optical random number

generation, and both analog and digital ANN implementation.

His current research interests include large ANN structures modeling and implementation, hybrid neural based information processing systems and their software and hardware implementations, design and implementation of real-time neuro-control and neural based fault detection and diagnosis systems. Since 1996 he is a permanent member (elected Academician) of International Informatization Academy. In 1997, he was also elected as Academician of International Academy of Technological Cybernetics.



Dr. Véronique Amarger received her PhD degree in Microelectronics and Computer Science from the University of Paris 7, France, in 1993. Since 1993, she has joined Senart Institute of Technology of PARIS XII – Val de Marne University, one of two Institutes of Technology of this University, where she works as Assistant Professor. She is a Staff Member of Image, Signal and Intelligent Systems Laboratory (LISSI / EA 3956) of PARIS XII University.

Her main research interests concern the field of Bio-inspired Artificial Intelligence, Neural Networks and Computer Aided Diagnosis Systems design and applications.



Joël Bernier graduated from the French Engineering school Ecole Supérieure d'Optique (ESO) in 1979, he worked 15 years at MATRA, now part of EADS group, assigned to project management for Optronics equipment development. Then, he joined the SFIM group and in particular its subsidiary REOSC in 1997 where he was the Astronomy and Laser business unit manager. After the takeover of SFIM by SAGEM SA in 1999, he is in charge of the R&D team of the REOSC Program, where all high performance optics for astronomy, laser, space or lithography are designed and manufactured.