Pattern Anal Applic (2005) 8: 181–187
DOI 10.1007/s10044-005-0254-5

THEORETICAL ADVANCES

Manuele Bicego · Maria Del Rosario Martinez
Vittorio Murino

A supervised data-driven approach for microarray spot quality classification

Published online: 11 June 2005

Abstract In this paper, the problem of classifying the quality of microarray data spots is addressed, using concepts derived from the supervised learning theory. The proposed method, after extracting spots from the microarray image, computes several features, which take into account shape, color and variability. The features are classified using support vector machines, a recent statistical classification technique that is being employed widely. The proposed method does not make any assumptions on the problem and does not require any a priori information. The proposed system has been tested in a real case, for several different parameters' configurations. Experimental results show the effectiveness of the proposed approach, also in comparison with state-of-the-art methods.

Keywords Microarray data · Quality classification · Support vector machines · Pattern recognition

1 Introduction

The recent wide employment of microarray tools in molecular biology and genetics have produced an enormous amount of data, which has to be processed to infer knowledge. Due to the dimension and complexity of those data, automatic tools coming from computer science and data analysis research areas have been successfully employed. In the particular case of microarray, the result of an experiment is an image, therefore, the extraction of information is performed through the analysis of visual patterns present in the image itself. The whole analysis is improved and made more reliable with the introduction of automated and sophisticated techniques coming from the image processing and pattern recognition research area [1].

Most literature concerning the employment of computer science methodologies to microarray data is related to two aspects: the detection of the spots (segmentation spot/background, grid matching, noise suppression [2]) and the analysis of data for classification (data mining) [3, 4]. Nevertheless, there is another interesting and crucial issue, typically disregarded in the current literature: the determination of the quality of the microarray experiment. It is essential to employ a quality control process in microarray experiments since each experiment, even if performed under optimal conditions, could result in several spots whose intensity may vary due to experimental oscillations. It is also crucial to exclude spots whose quality is poor in the early stages of a microarray study because data normalization methods typically involve an estimation phase. Algorithms used in estimation may get confused when number of aberrations is large. Furthermore, microarray data are frequently used by researchers different from the original, and distributing quality certified data could be more efficient and useful than distributing images.

The literature about spot quality control is not vast, only recently some papers appeared, which can be subdivided into two main classes of approaches. In the first class [5, 6, 7], the main goal is to assess the quality of the spot by monitoring and imposing thresholds on one or more features of the spot, supposed to be discriminative and informative for the quality. For example, in [5], a composite characteristic is introduced, based on five qualitative features. Five thresholds are set for these characteristics, mainly derived from a priori knowledge, experience and heuristic considerations. A similar approach has been proposed in [6], employing four spot features. In this case also, there is the need to choose some thresholds driving the quality

M. Bicego (✉)
DEIR, University of Sassari,
Via Torre Tonda 34,
Sassari, 07100, Italy
E-mail: bicego@uniss.it

M. Del Rosario Martinez · V. Murino
DI, University of Verona,
Strada Legrazie 15,
Verona, 37135, Italy
E-mail: vittorio.murino@univr.it

assessment. Finally, in [7], the quality is linked to the uniformity of the signal in the spot.

The second class of techniques to quality detection employs a completely different approach. Instead of using heuristic considerations for setting thresholds on arbitrarily chosen characteristics, the idea is to employ pattern recognition techniques to "learn" how the experts in the microarray field separate good spots from bad spots. This class represents a more interesting kind of approach, since the classifier is derived directly from the data, without the need to determine what the "relevant" features are and how the "reasonable" thresholds have to be set. In this class, using concepts and theories of supervised learning, the algorithms try to mimic and replicate the human decision. The approaches proposed in [8] and [9] belong to this class. The former detects the position of the spot in the image, assessing also the quality using a classifier learned on previously and manually classified examples. The classifier uses as features the intensity and the position. Even if it represents the first tentative of using supervised techniques in this context, this approach is quite poor, using very simple tools. More interesting and well principled is the second approach, employing carefully trained Bayesian Networks [10] to assess the quality of a spot. Several experiments have been carried out in that work, using both continuous and discrete features, showing that Bayesian Networks are an effective tool to solve the problem. Nevertheless, the proposed approach has a great disadvantage: the structure of the Bayesian Network, which describes the interrelationships occurring between the components of the model, should be defined and designed a priori, needing to exactly know such relationships [1]. In other words, this method, similarly to the approaches of the first class, needs a significative part of heuristic-based knowledge in order to set up a determinant part of the method.

In this paper, we investigated the use of an alternative system to classify the quality of a spot, which belongs to the second class of the above-described taxonomy, implementing supervised learning techniques to detect the quality of the spots. The proposed method does not make any assumption on the problem, and does not require any a priori knowledge. Even if making assumptions on the problem could be useful in some cases, we are persuaded that wrong a priori specifications could drastically affect performances of the method. For example, in the Bayesian Network approach of [9], the assumption of an incorrect causality between two factors could lead to a bad model. In our opinion, a reasonable compromise could be the use of a "consensus" approach, assessing the quality of a spot by using more than one single method, at least one a priori knowledge driven and one data driven.

In this paper, we investigated the classification capability of a completely data-driven technique in spot quality classification, namely the support vector machine (SVM) [11, 12]: this tool represents a quite recently introduced supervised technique widely and successfully employed in several areas. This technique has been also used in the microarray data classification context [4], even if its use for quality assessment has never been investigated. Spot quality classification is an appropriate problem to assess the capabilities of SVMs, especially because they show some characteristics that are very suitable for that context, as for example, their intrinsic binary nature and their capability to deal with high-dimensional space.

The SVMs are here customized to tackle the above-mentioned problem, and have been experimentally evaluated on the same data set of [9], so that an analytical comparison is possible. A thorough testing has been carried out on the SVM, analyzing different parameters' configurations. Moreover, the effect of a dimensionality reduction technique, such as principal component analysis (PCA) [13], has been investigated. The best result obtained in our experiments outperforms the best result obtained in [9], showing the effectiveness of our approach.

The rest of the paper is organized as follows. In Sect. 2, a brief introduction to SVM is presented, and the proposed approach is detailed in Sect. 3. Experimental results and discussion are described in Sect. 4, and in Sect. 5 conclusions are drawn and future perspectives are envisaged.

2 Support vector machines

Support vector machines [12] are binary classifiers, extensively employed in recent years in several applications, such as face recognition and authentication [14, 15], object classification [16], textile defects classification [17], robotics [18] and others. They have several appealing characteristics: possibility of fast training [19], accurate classification and at the same time, high performance of generalization, i.e., ability of learn the trend and regularity of the data. Due to lack of space, the SVM

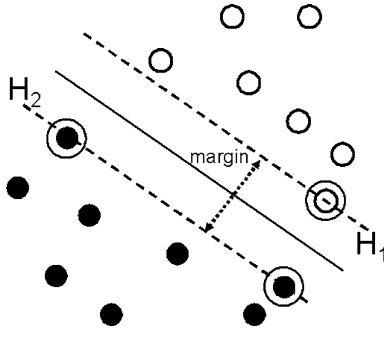

Fig. 1 Geometric interpretation of SVMs. A hyperplane separates black points from white points. The hyperplane is obtained as a linear combination of the circled points, called *support vectors*

[1]The B-Course method used by the authors to infer the structure of the Naive Bayesian Networks merely represents a feature selection step.

methodology is not formally described here, only the intuition is given. Interested readers are referred to [12].

Briefly, SVMs are binary classifiers, able to separate two classes through an optimal hyperplane. The optimal hyperplane is the one maximizing the "margin", defined as the distance between the closest examples of different classes (see Fig. 1).

To obtain a nonlinear decision surface, it is possible to use *kernel functions*, in order to project data in a high-dimensional space, where a hyperplane can more easily separate them. Examples of *kernel* are *polynomial functions*

$$K(x,y) = ((x \cdot y) + 1)^p$$

and the *Gaussian radial basis function*

$$K(x,y) = \exp\left(-\frac{||x-y||^2}{2\sigma^2}\right)$$

It is important to notice that, by the use of this "kernel trick", the non-linear decision surface is obtained in roughly the same amount of time needed to build a linear SVM.

Besides the specific parameter of the chosen kernel, namely σ for the Gaussian radial basis function (RBF) and p for the polynomial, another crucial parameter is the regularization constant C [11]. C controls the tradeoff between complexity of the machine and the number of nonseparable points. A larger C indicates that we assign a high penalty to errors and thus it reduces the number of misclassifications: in this case, we force the classifier to be well trained on the specific training set. In contrast, a smaller C indicates that we neglect some misclassifications to allow a wider margin, typically resulting in a better generalization capability.

3 The proposed approach

Given a microarray experiment, several operations have to be performed before the classification could actually be carried out. In particular, the spot should be localized, i.e., its pixels should be identified, the background should be determined and features can be finally computed. In literature, several approaches have been proposed to solve this problem, mainly based on image processing techniques (see for example [2] or [20]). After features computation, the classification can be performed. Our goal is to investigate the SVM capabilities in classifying the quality of a microarray spot, hence, we do not care about the previous phases, and we start describing our system from the features.

3.1 Features extraction

One of the most critical phases in spot quality determination is the choice of appropriate features. Before describing the approach used to extract them, let us describe what the main problems are that occur in microarray experiments and how they could drive to bad quality spots.

Spot intensity The spot intensity is considered one of the most important features affecting the quality of spot. If a signal is weak, it is impossible to discriminate the actual signal from the background. The principal reason of the weakness of a spot stem in the fact that many genes are physiologically expressed at very low levels. Moreover, there are some experimental factors which may cause low signal intensities such as low amount of DNA in the spot, incomplete hybridization, low sensitivity of the scanner and others.

Spot size Microarray spots are expected to be of roughly equal size, deviations from this fact could represent important cues indicating that something wrong happened. Wrong size spots could be caused by impurities in the printing solution, or by damaged or dirty needles, or by high humidity during printing.

Spot morphology Spots are expected to be circular in shape, so the morphology is important. In particular, the not roundness and the bleeding [2]

Pixel intensity distribution A spot is expected to contain approximately equal amounts of DNA all over its area, so brightness should be uniform. Not uniform area could derive from uneven distribution of the printed DNA in the spot or nonspecific binding.

Starting from these considerations, the features used to classify quality were [9]: the bleeding, the size of the spot, the roundness of the spot, the alignment error, the spot intensity, the intensity of the background and the background noise. Each feature was measured separately using the Cy5- and the Cy3-channels, giving a total of 14 features. The procedure is the following: assuming the bounding box of the spot known, we fitted a Gaussian shaped surface on the spot, considering the color of each pixel as the height of the surface. The fitting has been determined using a standard nonlinear least squares procedure [21]. The fitted Gaussian could be expressed as:

$$f(x,\theta) = A\, \mathrm{e}^{-(\mathbf{x}-\mathbf{m})^{\mathrm{T}}\mathbf{S}(\mathbf{x}-\mathbf{m})} + B$$

where $\mathbf{m}$ is the Gaussian mean, and $\mathbf{S}$ is the covariance matrix, which depends on the two variances σ_1,σ_2 and on the rotation angle ϕ. Six out of the seven features are computed from the estimated Gaussian, more in detail:

- the spot intensity is A;
- the background intensity is B;
- the alignment error is the distance between the center of the Gaussian $\mathbf{m}$ and the center of the bounding box
- the roundness of the spot is the ratio σ_1/σ_2

[2] The bleeding could be defined as the phenomenon in which a spot spreads so much that it is mixed with its neighbors should be carefully avoided.

– the size of the spot is the product $\sigma_1 \sigma_2$
– the background noise is the root mean square error of the fitted Gaussian

Finally, the bleeding is computed by counting the pixel of the spot falling out the estimated Gaussian spot.

The features have been considered as continuous values or have been discretized using the following method: the spot bleeding feature was discretized into two values, using as threshold a constant proportion of 65% of spot pixels. The other features were discretized into three values, so that the 30% of the lowest feature values within the image were labeled with 1, the 30% of the highest features values with 3, and the remaining values with 2.

3.2 Spot quality classification

Once determined the features of the spot, its quality could be assessed. The problem of quality detection is cast into a binary classification problem, where the two classes are "good" (class 1) or "bad" (class −1). The membership of a spot to one class is decided by a supervisedly learned machine, which observes several examples of the problem, with the corresponding labels, and tries to learn the general driving behavior. Therefore, the system is developed in two phases: the training step and the testing step. The former is devoted to the construction of the SVM, given a training set of previously classified spots. This set is crucial, since a bad choice of the training set could invalidate the whole process. The training set should be representative of the process under analysis, should be complete and cover all classes. In our case, the training set is defined as a set of spots whose quality has been judged by three experts in a conservative manner, i.e., discarding those spots for which all experts were not in perfect agreement. Once the training set is given, we have to choose the SVM parameters, as the kernel and the regularization constant C; finally, the SVM is trained. We want to stress the fact that, differently from the approaches previously proposed in this context, this method is completely data-driven, using only the training set in a supervised fashion to set-up the classifier, modeling the relationships among components and decide the goodness or the badness of the microarray spots.

If the number of features is large, a common strategy is to reduce the dimensionality of the resulting space using linear reduction techniques, in order to reduce the impact of the curse of dimensionality problem [22] to the classification. The PCA [13] represents one example of such technique. This method projects the data in the dimension "explaining" the largest amount of variance of the data, and is widely used in the pattern recognition area. In this paper, we tested the effectiveness of such a dimensionality reduction technique when a SVM is applied. In particular, we reduced the data space using PCA, by applying a linear transformation of the data. The transformation matrix is composed by the Eigenvectors corresponding to the dominant Eigenvalues, in order to lose the minimum amount of variance in the transformation. The number of Eigenvectors (which determines the dimensionality of the resulting space) was chosen in order to maintain an adequate amount of variance in the data.

Once trained the SVM, the testing phase consists of the computation of the features of a spot, possibly reducing the dimensionality (by applying the prestored transformation matrix), and, finally, the hyperplane side, the spot belongs to, is determined. If it is in the correct side (class 1), the spot is classified as good; otherwise it is classified as bad.

The SVMs have been chosen for different reasons. The first is that the SVMs are able to automatically determine if a spot is good or not, without the need to find an appropriate threshold owing to the supervised nature of the technique. Moreover, this technique does not require any a priori knowledge on the application domain, making the proposed approach completely "data driven". Another reason is that the problem is binary (spot good or bad), and SVMs are intrinsically binary classifiers (the generalization of SVM to a multiclass system is not a completely solved problem). A further motivation is that SVMs have a very high ability in classifying data in high-dimensionality spaces, like in this case since they are less affected by the curse of dimensionality problem [11]. Moreover, SVM represents one of the most powerful techniques in pattern recognition and their customization to tackle the spot quality classification problem has not been investigated yet. To this aim, results show that microarray-related applications can be greatly improved by this preliminary data analysis. Finally, the SVM methodology proposed in this paper could be straightforwardly extended in order to derive a quality measure, able to label each spot with a confidence score. This could be easily done by considering the quality score as the distance of the point representing the spot from the hyperplane derived from the training procedure [3] This has been already done in other contexts (see for example [23]), and will be investigated by the authors in the future.

4 Results and discussion

The proposed approach has been tested using data obtained from [9], consisting of 155 spots obtained from two different hybridizations (HBL-100 versus BT-474 and MCF7 versus HBL-100). The dataset contains 97 good spots, while the remaining were bad [4] Examples of spots derived from the two hybridizations, after the gridding procedure, are presented in Fig. 2.

[3] In classification, only the sign is used, not the magnitude.
[4] Data, together with experiments' descriptions, data specifications, figures, experts' classifications and labels are available on the web site http://sigwww.cs.tut.fi/TICSP/SpotQuality/.

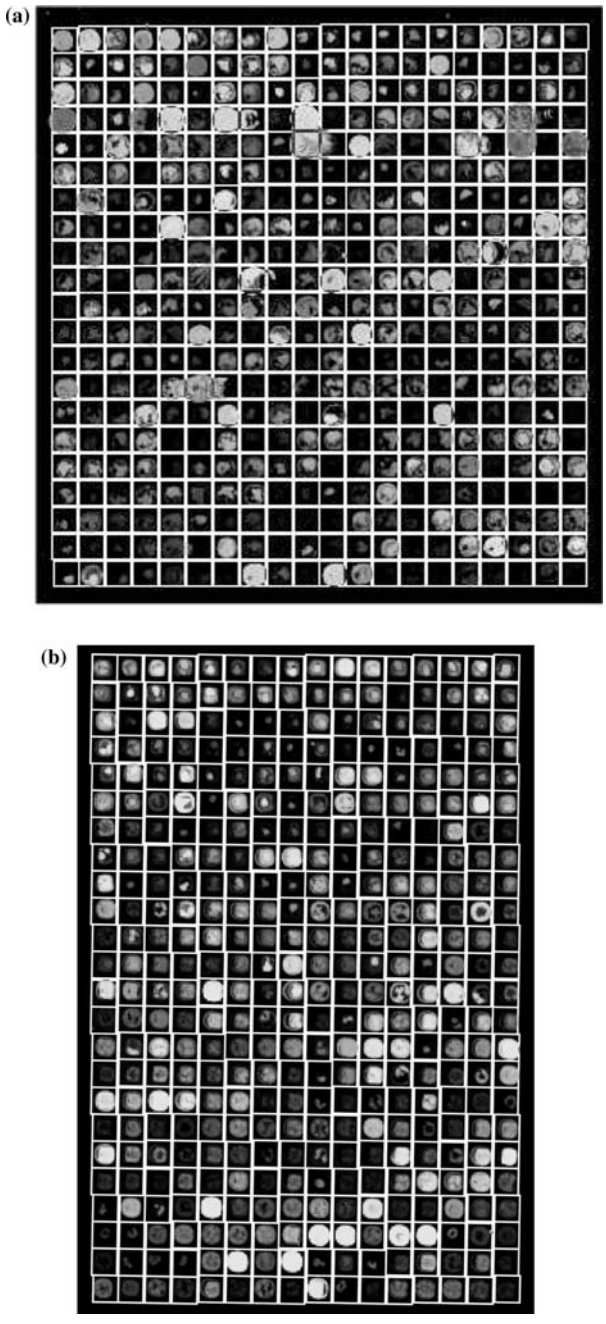

Fig. 2 Spots from the two different hybridizations after the gridding procedure: **a** HBL-100 versus BT-474 and **b** MCF7 versus HBL-100

The subarrays used in the experiments were chosen so that they do not contain any large-scale defects such as scratches that had removed many spots or large areas of strong background. Further, the subarrays were chosen to be of average quality, i.e., they do not only contain mostly good or bad spots. The features have been computed using the strategy proposed in Sect. 3.1.

As in [9], the classification error was determined using Leave One Out Cross Validation [22], a method that separates the training set from the testing set, in order to get a good estimate of the generalization error. We performed several experiments, varying all the SVM parameters, like the kernel and the C parameter. We performed testing using both continuous and discrete features. Moreover, we tested the accuracy with and without the application of the PCA.

All the implementation was done in MATLAB; in particular, we used a publicly available SVM toolbox [24] and the PCA was employed using the MATLAB statistic toolbox.

We tested three kernels, the linear, the polynomial (with the degree ranging from 2 to 8), and the Gaussian RBF (with σ ranging from 2 to 8). For all these configurations, the C parameter from 1 to 96 (step 5) is varied, computing the accuracy error. Due to lack in space, only the best results are reported for each kernel, shown in Table 1. In that table "*PCA2*" means that the accuracy was computed on the space reduced by PCA. We used only two dimensions since with these two the 99% of the total variance of the data is explained. More in detail, each component of the reduction matrix covers a fraction of the total variance of the data. By computing the cumulative variance, we could determine which is the minimum number of components needed to cover a given amount of variance. In our experiments, with only two components, 99% of the variance is covered. In the table also, the number of support vectors (#sv) obtained after the training is shown, giving an idea of the generalization capability of the SVM [25]. The coefficient p indicates the parameter of the polynomial kernel, and σ stands for the parameter of the Gaussian RBF kernel.

By looking at this table, one can notice that the best result is obtained using the polynomial kernel, and with a less number of #sv with respect to the other kernels used. The accuracy is quite satisfactory, reaching almost 95%. Moreover, one can also observe that the use of PCA worsens the classification accuracies: this confirms

Table 1 Best results obtained by the proposed approach using continuous features, varying kernel on PCA-reduced or not reduced space

Kernel	*C*	p/σ	#sv	Accuracy
Linear	86	–	36	90.97%
Polynomial	76	2	27	94.84%
Gaussian RBF	86	2	44	91.61%
Linear (PCA2)	61	–	102	74.84%
Polynomial (PCA2)	31	8	70	82.58%
Gaussian RBF(PCA2)	96	2	109	73.55%

Table 2 Best results obtained by the proposed approach using discretized features, varying kernels and kernel parameters

Kernel	*C*	p/σ	#sv	Accuracy
Linear	6	–	25	96.13%
Polynomial	86	6	13	94.19%
Gaussian RBF	11	2	29	97.42%

Table 3 Leave One Out accuracies of different classifiers on the discrete features set

Method	Accuracy
B-Course (subjective)	96.8%
Pair-wise NB (subjective)	95.5%
NB (subjective)	95.5%
NB (uniform)	94.8%
Decision Tree	91.6%
Neural Networks	90.3%
The proposed approach	97.4%

the fact that the SVMs are more effective in nonstructured spaces, as hypothesized in [14]. In other words, it seems that the SVMs are more effective if the space is coarse and not "preprocessed". Another consideration is that these results confirm the fact that PCA is not always the best choice and especially in not normally distributed classification contexts, other techniques are often more appropriate (one choice could be the Matching Pursuit approach [26]).

The second testing session was performed using the discretized data. As in the previous case, we tested the three kernels, varying the *C* parameter from 1 to 96 (step 5) and computing the accuracy error. Also in this case, only the best results are shown in Table 2, together with the number of support vectors.

From this table, one can notice that the best result is obtained using the kernel Gaussian RBF, which is similar to results typically reported in literature in other application domains [15, 17].

Finally, in order to compare our approach with the state-of-the-art, in Table 3, we summarize our best performance and those of some of the best methods proposed in [9] on the same data set.

For all these classifiers, the Leave One Out Cross Validation error has been computed. The first four entries of the table are different variants of the Bayesian Networks system introduced in [9], where the prior is enclosed in brackets, while the fifth and sixth rows show the results of the standard decision trees and standard Neural Networks, respectively. The last row shows the best result obtained with our approach, which is better than those of all other methods, confirming the effectiveness of the SVM method for this problem. We want to emphasize the fact that we are talking about experiments involving thousands of spots, so the advance with respect to the state-of-the-art, which could be considered quite poor (about 1%), assumes in this context a relevant significance.

5 Conclusions

In this paper, the problem of classifying the quality of microarray spot has been addressed with a complete "data-driven" approach. The method is based on SVMs, a binary classifier which determines a hyperplane able to separate good spots from bad ones. The system has been thoroughly tested in a real experiment, varying the SVM parameters. The use of PCA for dimensionality reduction has also been investigated. Experimental results on real data sets are comparable or better than the state-of-the-art methods, therefore, SVMs seem to be very suitable in spot quality classification. A future issue could be the investigation of the possible derivation of a confidence score from the SVM, in order to give a quality measure to be propagated in all the subsequent analyses.

6 Originality and contributions

The microarray technique has emerged in recent years as an inspection tool, able to contemporarily analyze the expression of several thousands of genes. Nevertheless, several factors could affect a microarray experiment, resulting in bad quality spots that could drastically reduce the significativeness of subsequent analysis, and that should be removed from the data to be processed. On the other side, the large amount of spots resulting from each experiment makes the manual inspection very time consuming. Thus, there is a real need of a reliable system able to automatically detect and discard bad quality spots. In this paper, the problem of classifying the quality of microarray data spots is addressed implementing supervised learning techniques. The key tool of the proposed approach is represented by SVMs, a classification approach introduced in the field of statistical pattern recognition in the 90's. This technique, which presents several appealing intrinsic properties, seems to be really suitable for spot quality classification: nevertheless, its use in this context has never been investigated. Another key feature of the proposed system is that, differently from other approaches proposed in the-state-of-art methods, it does not make any assumption on the problem and does not require any a priori knowledge. The proposed approach has been tested on real data, outperforming other state-of-the-art methods on the same dataset.

Acknowledgements The authors would like to thank very much Dr. Sampsa Hautaniemi of Tampere University of Technology (Finland), for kindly supplied the microarray data and the features used for testing. The authors would like to thank also Dr. S. Barbi and Prof. A. Scarpa of the Department of Pathology of the University of Verona (Italy) for helpful discussions. Finally, the authors would like to thank G.E. Felis for carefully reading the paper.

References

1. Model F, König T, Piepenbrock C, Adorján P (2002) Statistical process control for large scale microarray experiments. Bioinformatics 18:155S–163S
2. Brändle N, Bischof H, Lapp H (2003) Robust dna microarray image analysis. Machine Vision and Applications 15:11–28
3. Valafar F (2002) Pattern recognition techniques in microarray data analysis: A survey. Ann N Y Acad Sci, Special issue on Tech Bioinform Med Inform 980:41–64

4. Mukherjee S (2003) Classifying microarray data using support vector machines. In: Berrar D, Dubitzky W, Granzow M (eds) A Practical Approach to Microarray Data Analysis. Kluwer, pp 166–185
5. Wang X, Ghosh S, Guo S (2001) Quantitative quality control in microarray image processing and data acquisition. Nucleic Acids Res 29:75
6. Chen Y, Kamat V, Dougherty E, Bittner M, Meltzer P, Trent J (2002) Ratio statistics of gene expression levels and applications to microarray data analysis. Bioinformatics 18:1207–1215
7. Brown C, Goodwin P, Sorger P (2001) Image metrics in the statistical analysis of DNA microarray data. Proc Natl Acad Sci USA 98:8944–8949
8. Buhler J, Ideker T, Haynor D (2000) Dapple: improved techniques for finding spots on DNA microarrays. Technical Report UWTR 2000-08-05, Department of Computer Science and Engineering, University of Washington
9. Hautaniemi S, Edgren H, Vesanen P, Wolf M, Järvinen A, Yli-Harja O, Astola J, Kallioniemi O, Monni O (2003) A novel strategy for microarray quality control using Bayesian networks. Bioinformatics 19:2031–2038
10. Heckerman D (1995) A tutorial on learning with bayesian networks. Technical Report MSR-TR-95-06, Microsoft Research Revised November, 1996 - dowloadable from family ftp://ftp.research.microsoft.com/pub/tr/tr-95-06.pdf
11. Vapnik V (1995) The nature of statistical learning theory. Springer, Berlin Heidelberg New York
12. Burges C (1998) A tutorial on support vector machine for pattern recognition. Data Min Knowl Disc 2:121–167
13. Jollife I (1986) Principal component analysis. Springer, Berlin Heidelberg New York
14. Jonsson K, Kittler J, Li YP, Matas J (1993) Support vector machines for face authentication. In: Proceeding of Brit Machine Vision Conf, Nottingham, UK pp 543–553
15. Bicego M, Iacono G, Murino V (2003) Face recognition with multilevel B-splines and support vector machines. In: Proceeding of ACM SIGMM multimedia biometrics methods and applications workshop, pp 17–24
16. Pontil M, Verri A (1998) Support vector machines for 3-D object recognition. IEEE T Pattern Anal Mach Intell 20:637–646
17. Murino V, Bicego M, Rossi I (2004) Statistical classification of raw textile defects. In: Proceedings of IEEE International conference on pattern recognition 4:311–314
18. Castellani A, Botturi D, Bicego M, Fiorini P (2004) Hybrid HMM/SVM model for the analysis and segmentation of teleoperation tasks. In: Proceeding of IEEE International conference on robotics and automation 3:2918–2923
19. Platt JC (1998) Fast training of support vector machines using sequential minimal optimization. Advances in Kernel methods-support vector learning
20. Hartelius K, Carstensen J (2003) Bayesian grid matching. IEEE T Pattern Anal Mach Intell 25:162–173
21. Draper N, Smith H (1998) Applied regression analysis, 3rd edn. Wiley, New York
22. Duda R, Hart P, Stork D (2001) Pattern Classification, 2nd edn. Wiley, New York
23. Bicego M, Grosso E, Tistarelli M (2005) Probabilistic face authentication using hidden markov models. In: Proc of SPIE International Workshop on Biometric Technology for Human Identification
24. Gunn S (1997) Support vector machines for classification and regression. Technical report, Image Speech and Intelligent Systems Research Group, University of Southampton
25. Vapnik V (1998) Statistical Learning Theory. John Wiley, New York
26. Mallat S, Zhang Z (1993) Matching pursuit with time-frequency dictionaries. IEEE T Signal Proces 41:3397–3415