Biocybernetics and Biomedical Engineering
2012, Volume 32, Number 2, pp. 3–16

Highly Robust Statistical Methods in Medical Image Analysis

JAN KALINA*

Center of Biomedical Informatics, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic

Standard multivariate statistical methods in medical applications are too sensitive to the assumption of multivariate normality and the presence of outliers in the data. This paper is devoted to robust statistical methods. In the context of medical image analysis they allow to solve the tasks of face detection and face recognition in a database of images. The results of the robust approaches in image analysis turn out to outperform those obtained with standard methods. Robust methods also have desirable properties appealing for practical applications, including dimension reduction and clear interpretability.

K e y w o r d s: robust statistics, classification, faces, robust image analysis, faces, forensic science

1. Introduction

First methods of multivariate statistical analysis were developed at the beginning of the 20th century for anthropological applications in the physical anthropology. The concepts of correlation analysis, classification analysis and statistical diversity measures (distances) were developed by researchers analyzing anthropological measurements. The most important multivariate statistical methods were proposed by anthropologists and only later they were able to spread to other branches of research. The principal component analysis was proposed by Karl Pearson [1], who was a professor of eugenics and founder of biometrics. The correlation analysis was introduced by [2] for the analysis of anthropometric data [3] in a study of the height of children, which was examined in connection with the length of their bones. The first article on the linear discriminant analysis [4] analyzed craniometry data.

---

* Correspondence to: Jan Kalina, Center of Biomedical Informatics, Institute of Computer Science AS CR, Pod Vodárenskou věží 2, CZ-182 07 Praha 8, Czech Republic, e-mail: kalina@euromise.cz
Received 17 June 2011; Accepted 09 December 2011

The main task of anthropology can be described as a statistical problem of classification analysis, which is based on variability decomposition to the inter-class and intra-class components. Both *face detection* and *face recognition* can be solved by classification analysis, which is a statistical method for group identification. We can understand face detection and face recognition to have antipodal aims. Nevertheless both tasks are of the same type, namely classification analysis. Face detection corresponds to localization of the face in the image and classifies each part of the image as a face or non-face (area not corresponding to a face). Face recognition classifies each face to a particular person of the given database. For both contexts the usual methods of forensic anthropology start by a dimension reduction (principal components, Fourier transform, discrete cosine transform or wavelet transform) and feature extraction to describe the differences among images or their groups and the contribution of variables to these differences [14].

Many papers on physical anthropology are devoted to identification of groups [5]. A paleopathological study of bone shapes by means of the landmark or shift registration was presented by J.O. Ramsay and B.W. Silverman [6]. Various versions of the robust correlation coefficient in the context of medical image analysis were investigated by J. Kalina [7] together with a study of computational aspects of such coefficients. Methods of image analysis are widely used also in forensic anthropology A. Zankl [8] or S. Böhringer et al. [9] studied anthropometric measures in the face, which allowed the diagnosis of genetic diseases based on a craniofacial dysmorphism. Conventional anthropometrics can be used for the description of individuality or deviations from average. It is supplied by a classification procedure assigning each face to the particular genetic syndrome. At the same time anthropometrics can focus on characteristic facial features. Therefore we describe both face detection aiming at typical properties of a given face (common for each face) and also face recognition, which focuses on invididual variability compared to other faces. Applications in forensic stomatology are summarized by J. Zvárová et. al. [10].

Standard methods of image analysis in anthropology or biomedicine (for example analysis of medical ultrasound images) do not have desirable robustness properties, which can be ensured by robust statistical methods. Typical procedures start with a dimension reduction and feature extraction and proceed to a machine learning method (for example neural networks, support vector machines or Bayesian classification), which are revealed by robust mathematical statistics to be non-robust in terms of their vulnerability to outlying values. Individual classifiers are often combined with ad hoc procedures to create complex black box systems with numerous parameters, disabling a simple interpretation. Noise plays a highly influential role in the information extraction from images and the methods are sensitive also to two-dimensional artifacts or to specific assumptions which cannot be verified. The aim of automated methods of digital image analysis in anthropology is the landmark localization. Landmarks are defined as points of correspondence (exactly defined biologically or geometrically) on each object that match between and within populations [11, 12]. Examples of the landmarks include the soft tissue points located on the inner and outer commisure of each eye fissure, the points located at each labial commisure, the midpoints of the vermilion line of the upper and lower lip [13].

Robust statistical methods represent a new paradigm applicable to medical signal and image analysis allowing to obtain a resistant solution with respect to noise in the signal or image, not relying on the assumption of normal distribution of the data, but reliable under atypical situations [15]. While their origin goes back to 1960s [16], only recently they started to penetrate to applications to different fields including biostatistics [17] or microarray image analysis for medical applications [18]. They have not been widely applied to image analysis in anthropology [19]. The breakdown point [20] has become an important statistical measure of sensitivity against noise or outliers in the data. Highly robust methods possess larger values of the breakdown point.

In this paper we study robust statistical methods from the theoretical point of view and apply them to tasks of medical image analysis. In contrary to standard approaches we work with raw data without the usual steps of dimension reduction and feature extraction, which has the aim to describe the differences among groups, to reveal the dimensionality of the separation among groups and the contribution of variables to the separation. The standard transformations in medical image analysis not only lead to a considerable loss of information, but are also sensitive to presence of the outliers in the raw data.

Section 2 of this paper recalls the least weighted squares regression, which is one of robust regression methods with a high breakdown point. Section 3 devoted to face detection by means of template matching works with a database of images of faces, which is further considered also in other sections. Here we compute the templates in a sophisticated procedure to yield the optimal classification rule over the training set of images. Section 4 defines robust procedures for linear and quadratic discrimination analysis, which are based on the idea of implicit weighting of Section 2. Section 5 applies the methods to face detection by means of robust multivariate statistical methods. Section 6 studies face recognition by means of robust measures of symmetry in images. The final discussion in Section 7 summarizes the methods from a unified point of view of robust classification analysis.

2. Least Weighted Squares Regression

This section recalls the least weighted squares (LWS) regression estimator of [21] and summarizes its properties and advantages. In following sections the idea of the LWS to assign weights to the data in an implicit way will be used to define new robust statistical methods. Many robust methods are based on detecting the outliers in the data. Therefore they are influenced by an additional error, which is introduced by the outlier detection. The method may actually rely on a too small percentage of data points extracting too specific properties of the majority of the "good data", which may actually lead to overfitting. On the other hand the LWS regression does not include the outlier detection intrinsically, while the potential outliers are only down-weighted and not trimmed away completely.

Let us consider the linear regression model in the form

$$Y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + e_i, \quad i = 1,\ldots,n, \tag{1}$$

which can be rewritten in the usual matrix notation as $Y = X\beta + e$. The least weighted squares estimator is one of robust estimation methods with a high breakdown point. The LWS estimator down-weights less reliable observations based on the values of squared residuals. The magnitudes of nonnegative weights $w_1, w_2, \ldots, w_n$ must be specified before the computation of the estimator. These are, however, assigned to the particular data points after a permutation, which is determined automatically only during the computation based on the residuals. It is reasonable to choose $w_1, w_2, \ldots, w_n$ as a non-increasing sequence so that the most reliable observations obtain the largest weights, while the outliers with large values of the residuals get small (or zero) weights. One possibility is to choose linearly decreasing weights. The data-adaptive weights of [22] are another choice.

Let us denote the *i*-th order value among the squared residuals for a particular value of the estimate b of the parameter β by $u^2_{(i)}(b)$. The least weighted squares estimator b_{LWS} for model (1) is defined as

$$b_{LWS} = \arg\min \sum_{i=1}^{n} w_i u^2_{(i)}(b),$$

where the minimum is computed over all possible values of b.

The computation of the LWS estimator is intensive and computational aspects are studied by J. Kalina [23]. The least trimmed squares (LTS) regression proposed by P.J. Rousseeuw and A.M. Leroy [24] represents a special case of least weighted squares with weights equal to zero or one only.

The least weighted squares estimator has interesting properties and applications. In particular it is robust for contaminated data sets, where the disturbances are normally distributed random variables contaminated by outliers. At the same time the estimator is reliable for data with normally distributed disturbances without contamination; this reliability is measured in terms of a high efficiency, which compares the asymptotic variability of the estimator relatively to the variability of the least squares estimator. Theoretical properties including the robustness of the LWS estimator are studied by P. Čižek [22], who conjectures that the estimator is a reasonable compromise between the least squares and the least trimmed squares. The breakdown point attains the maximal possible value of (approximately) one half [21]. Diagnostic tools for the disturbances e (random regression errors) are available.

For example tests of heteroscedasticity and autocorrelation of the disturbances can be computed employing the weighted residuals [23]. Such diagnostic tests are asymptotically equivalent with classical tests for least squares regression. The LWS estimator has a small local sensitivity compared to the LTS, which suffers from high sensitivity to small deviations near the center of the data.

3. Face Detection by Robust Template Matching

Template matching is a popular tailor made method for object detection in grey scale images [25]. A template is a model, a typical form, an ideal object. It is placed on every possible position in the image and the similarity is measured between the template and each part of the image, namely the grey value of each pixel of the template is compared with the grey value of the corresponding pixel of the image. M.H. Yang et al. [26] gave references on template matching applied to face detection and face recognition. J. Kalina [27] claimed that the Pearson product-moment correlation coefficient is used as the most common measure of similarity between the image and the template.

Throughout this paper we are working with a database of images coming from the Institute of Human Genetics, University of Duisburg-Essen, Germany (projects BO 1955/2-1 and WU 314/2-1 of the German Research Council). This database contains 212 grey-scale images of the size 192 times 256 pixels, each image corresponding to a different person. The persons are volunteers in the age between 18 and 35 years of German origin. The images were photographed under standardized conditions, the faces have about the same size but are rotated in the plane by small angles and the eyes are not in a perfectly horizontal position in such images. Our aim was finding a robust solution with respect to rotation, occlusion, noise in the image and asymmetry of the face, while allowing for a clear interpretation of the method.

The Institute is working on interesting research problems in anthropology [9] with the aim to examine the connection between the size and shape of facial features and the genetic code and also to visualize a face based only on its biometric measures. Face localization is the first step of any digital processing of the face. The team of researchers in anthropology and genetics uses two approaches to locate 40 landmarks in each face. One possibility is the manual identification, which is carefully and accurately performed by an anthropologist trained in this field. As the second approach the Institute uses an automated method, which bases on the algorithm of [28] based on automatic localization of 40 landmarks in a training set of 83 images of faces. However such approach turns out to be vulnerable to small rotations of the face.

We use the same images as J. Kalina [19], who obtained 100% correct results in automatic localization of the mouth and both eyes in the database of 212 images using the Pearson correlation coefficient and a set of of 7 mouth templates and 6 eye templates, which were used together with their reflections in the mirror symmetry.

In the present work we propose an optimization method to obtain an optimal template for robust template matching. We explain the optimization on the example of locating the mouth. The mouth detection is a classification problem. The training set contains 212 images, in which the mouths are already identified. We aim at improving the classification rule between mouths and other parts of the images. We use the term non-mouth for such parts of the image (in the face or in the hair or background), which do not correspond to the mouth. The result of the procedure is an optimal template applicable for mouth localization in any new images of faces, which are not present in the training database.

In general the aim of classification analysis is to distinguish between the set of mouths and the set of non-mouths, to describe the difference between the two groups, to separate between them. In our case the classification rule is based on measuring some distance (separation measure) between a particular mouth and a particular non-mouth. We define this distance as the ratio of the weighted correlation coefficient between the mouth and the template and the weighted correlation coefficient between the non-mouth and the template. The aim is to improve the worst case over the whole database, in other words to improve the separation measure between the mouth and non-mouth, which have the worst separation measure by means of the template. This corresponds to such a non-mouth, which resembles a real mouth more compared to other non-mouths. We have proposed an algorithm for an approximative solution of such optimization procedure. The template is retained during the computation.

We describe the optimization of one mouth template. We start with a symmetric mouth template of size 27 times 41 pixels. This template was constructed as the mean of the mouths of 10 different persons of the same size 27 times 41 pixels. The weighted Pearson product-moment correlation coefficient of two random vectors $X = (X_1,...,X_n)^T$ and $Y = (Y_1,...,X_n)^T$ defined by

$$r_w(X,Y) = \frac{S_W(X,Y)}{S_W(X)S_W(Y)}$$

is exploited, where $S_W(X, Y)$ is the weighted covariance of X and Y and $S_W^2(X)$ and $S_W^2(Y)$ are weighted variances of X and Y, respectively. The template together with the weighted Pearson correlation coefficient as the similarity measure between the image and the template allows to locate the mouth correctly in 95% of images of the given database.

We use the template as an initial template for the optimization procedure. To avoid overfit we introduce regularization constraints, which do not allow individual weights to exceed a certain upper bound. Such optimal template has prominent lips and lighter areas directly above them. The noise is suppressed in the optimal template compared to the initial one. The optimal template seems harmonic, smooth, without unnatural features and subjectively appealing (based on [29]). The method yields very reliable results in locating the mouth over the whole database of images.

Moreover, the method turns out to be robust with respect to small changes of the size and rotation of the face, noise in the images, non-symmetry or local deformations of the mouth.

Further we optimize the weights for a fixed template. This optimization increases the separation between the mouth and non-mouth in the worst case over the whole database using the optimal template. While the symmetry of the mouth template and corresponding weights is assumed, in the search for eyes we do not assume such symmetry. The solution however tends to degenerate. The largest weights are present in a small number of pixels. Such procedure used too specific properties of the mouths in the training set of images. Therefore to obtain a robust solution we bound the influence of individual pixels. The optimal template together with the optimal set of weights are the output of such procedure, which locates the mouth reliably in 100% of the 212 images. This is a general method not using specific properties of the mouth.

4. Robust Linear and Quadratic Classification Analysis

We propose the minimum weighted covariance determinant (MWCD) estimator, which is a robust multivariate estimator based on the idea of down-weighting of less reliable observations. Then a robust quadratic classification method is proposed based on the MWCD estimator.

The minimum covariance determinant (MCD) estimator is a high-breakdown estimator of multivariate location and scatter [30]. It requires to choose the trimming constant h $(n/2 < h < n)$; while n-h observations are ignored completely, only the h remaining data points are used to compute the estimator. Particularly for the estimation of the multivariate location, the MCD estimator is defined as the trimmed mean with such n-h observations trimmed away, yielding the smallest possible determinant of the trimmed variance matrix. Such trimming involves the complete rejection of n-h data points, while the trimmed mean and the trimmed variance matrix are computed as classical mean and variance matrix using only the h remaining observations.

Let us consider p-dimensional data vectors $X_1, X_2, \ldots, X_n$. We propose to define the minimum weighted covariance determinant (MWCD) estimator as a weighted analogy of the MCD estimator, down-weighting of less reliable data points. It is required to specify the sizes of the weights $w_1, w_2, \ldots, w_n$; for example the linearly decreasing weights [7] can be used. For the fixed weights $w_1, w_2, \ldots, w_n$ the weighted mean $\bar{X}_W$ and the weighted variance matrix

$$S_W = \Sigma_i w_i (X_i - \bar{X}_W)(X_i - \bar{X}_W)^T$$

can be computed. In our case we consider all possible permutations of the weights. We find such permutation of the weights, which yields the minimal determinant of the weighted variance matrix S_W. We define the minimum weighted covariance determinant (MWCD) estimator of location as the weighed mean $\bar{X}_W$ of the data with these optimal weights and the corresponding estimator of the variance matrix is the weighted variance matrix S_W with the optimal weights. The MWCD estimator can be computed by a weighted analogy of the estimator of [30]. In the algorithm, the permutation of the data arranges the data according to the ascending order of the Mahalanobis distances from the robust estimator of the population mean. Therefore observations with a small Mahalanobis distance obtain larger weights.

The MWCD estimator is based on the idea of implicit weighting of all observations, which allows to down-weight less reliable observations. Considering fixed sizes of weights $w_1, ..., w_n$, it estimates the expectation of multivariate data by the weighted mean and the variance matrix by the weighted variance matrix, where the weights are assigned to the particular observations after a permutation. This is determined in order to yield the minimal determinant of the weighted variance matrix. The estimator has a large breakdown point and a large efficiency for normally distributed data. The MCD estimator generalizes the LTS regression to the multivariate context, while the MWCD estimator corresponds to the LWS regression. Similarly with the LTS regression, the MCD estimator suffers from a high local sensitivity. On the other hand the LWS regression has a small local sensitivity [21], which motivates the usage of implicitly weighted estimators such as the MWCD.

Our aim is a robust classification method based on the MWCD estimator. The first robust classification procedure was proposed by M. Hubert and K. van Driessen [31]. A comparison of standard robust approaches to classification analysis is given by V. Todorov and A.M. Pires [32]. We construct a different method based on the idea of implicit weighting, similarly with the least weighted squares regression. While the classification method can be applied to the original data, we will use it on principal components. Similarly to [33] or [34] we compute the robust principal component analysis, which allows us to preserve as much information relevant for the classification as possible while reducing the computational complexity.

The population mean μ_j and the variance matrix $\sum_j$ in the j-th group will be estimated by robust estimators, namely by the MWCD estimator of multivariate location $\bar{X}_{j,MWCD}$ and the variance matrix $S_{j,MWCD}$. This estimation is done separately in each group of data points. The robust estimation of the population characteristics allows us to define the robust quadratic classification method based on the MWCD estimator.

Definition 1. The quadratic MWCD-classification rule assigns a new observation X to the j-th group, if the quadratic classification function

$$Q_j^* = \bar{X}^T_{j,MWCD}\, S^{-1}_{j,MWCD}\, X -$$
$$-\frac{1}{2}\left(\log\left|S_{j,MWCD}\right| + \bar{X}^T_{j,MWCD}\, S^{-1}_{j,MWCD}\, \bar{X}_{j,MWCD} + X^T S^{-1}_{j,MWCD} X\right)$$

is equal to $max\left\{Q_1^*,...,Q_J^*\right\}$.

Robust linear discriminant analysis assumes additionally the same variance matrix in each group.

Definition 2. The linear MWCD-classification rule assigns a new observation X to the j-th group, if the linear classification function

$$L_j^* = -\frac{1}{2}\log|S_{j,MWCD}| - \frac{1}{2}(X - \bar{X}_{j,MWCD})^T S_{MWCD}^{-1}(X - \bar{X}_{j,MWCD})$$

is equal to $max\{L_1^*, ..., L_J^*\}$.

Here the new approach can profit from the properties of the MWCD estimator, inherited from the idea of down-weighting of less reliable data points. The estimator namely turns out to be very robust for highly contaminated data sets and efficient for normal data without contamination, and at the same time robust also with respect to the local sensitivity. Therefore the MWCD estimator overcomes an important drawback of locally sensitive LTS and MCD estimators.

5. Face Detection by Robust Classification Analysis

The quadratic discriminant analysis is a classification method sensitive to the assumption of multivariate normal distribution of the data [35]. It is not feasible for high-dimensional data with a number of variables exceeding the number of observations. Neither the robust methods [36] can be applied to high-dimensional data problems. We propose the robust quadratic discriminant analysis based on the MWCD estimator. This is a natural robust analogy of the classical quadratic discriminant analysis. To obtain a robust method we start the computation by the robust principal component analysis [34] and proceed to the robust classification analysis.

Firstly we use the projection pursuit algorithm, which is a general method of [24] for finding the most informative directions or components for multivariate (high-dimensional) data. Such classification is based on a robust measure of spread of the data, taking into account the outlyingness of each data point. Candidate directions for the principal components are selected by a grid algorithm optimizing such objective function only in a plane, while the subsequent components are added in the later steps. The fast algorithm of [36] for the method is implemented in library *pcaPP* of the *R* software package.

We consider the database of 212 mouths and 212 non-mouths. These mouths were identified manually. In each image we also select manually one non-mouth, which has the largest similarity with the mouth in the same image by means of the weighted correlation coefficient with the template of Section 3 using radial weights.

Firstly we computed 5 main principal components by the projection pursuit algorithm from the mouths and non-mouths and applied the MWCD-classification on the training set of 248 images. The results are 100% correct with linearly decreasing weights for the embedded MWCD estimator. As the method assigns larger weights to reliable or typical data points, we verified that there are again larger weights in the top part of the images, corresponding to the face parts above and aside from the lips. On the other hand the down-weighted data (outliers) are located on the boundary of the mouth or in the bottom part of the images in the area between the mouth and the chin.

6. Face Identification by Robust Correlation Analysis

The original database contains 212 images of faces. In addition we have a new database, which contains two additional images of each of the 212 persons. The aim of face identification is to assign each of the 2*212=424 images to the corresponding image of the original database. The new images are taken at the Institute of Human Genetics under the same conditions. Still some differences between repeatedly photographed faces can be clearly observed. Firstly we recall the robust correlation coefficient.

J.Á. Višek [21] proposed the least weighted squares regression estimator as a general robust statistical method further theoretically studied by the same author [37]. Based on this, J. Kalina [7] proposed the robust correlation coefficient based on the least weighted squares regression. It has the maximal possible value of the breakdown point, which equals the breakdown point of the LWS regression. The following study brings arguments in favor of using robust statistical methods. The robust correlation coefficient is computed as the weighted correlation coefficient with such weights, which are determined automatically by the least weighted squares regression, whereas one variable is considered to be the response and the other variable a regressor. This method localizes the face as the area in the image with the largest level of axial symmetry. Although faces are not entirely symmetric, robust statistical methods are able to handle this problem. Rectangular neighboring areas in the image are examined and the symmetry is measured by the robust correlation coefficient.

Further we localize the face in all images by means of the facial symmetry [7]. Let us consider one particular image from the new database. Let X denote the rectangle covering the left half of the face of width 40 pixels in this image. This will be compared with all 212 images from the reference database. Let us describe the procedure for one particular image from the original database and let Y denote the rectangle covering the left half of the face of width 40 pixels in this image. The problem is reduced to the comparison of X and Y. We use the robust correlation coefficient based on the least weighted squares. The procedure assigns X to such face from the original database, for which the rectangular half of the face Y has the largest robust correlation coefficient with *X*. The method yields results, which are correct over 100% of the new database with 424 images.

For practical purposes it is important for anthropological methods to be robust with respect to occlusion of the image or other special situations. Therefore we perform a study, in which we modify the images from the new database. We modify the images by small occlusion, additional noise, moderate illumination changes or the size of the face is modified slightly. Still 100% of faces from the new database are assigned to the correct face from the original database. The method gives 100% correct results in face identification also if the rectangle with the left half of the face is shifted one pixel aside. Therefore the face identification method based on robust statistics can be considered very reliable in terms of robustness to non-standard situations or shapes.

7. Discussion

While the first methods of multivariate statistics were proposed primarily for anthropological applications, statistics withdrew from this unique connection. However it may bring benefits to reconnect the world of statistics with the world of anthropology. Also current medical applications require new statistical methods compared to the classical methodology, which remains unchanged from the beginning of the 20th century. The new methods should address the new challenges. It is required that reliable statistical methods in medical research are both reliable for high-dimensional data and robust to the presence of outliers and to violations of the assumptions.

However the concept of robustness appears to allow two different interpretations. In statistics robustness is studied as a theoretical tool to ensure resistance of methods against noise, violations of Gaussian normal distribution or other special assumptions. In image analysis robustness is a key property of practical problems, which allows to process reliably also such images, which contain two-dimensional noise, occlusion, different illumination, size or rotation of the objects compared to standard images. Nevertheless the two aspects of robustness are connected and robust statistical methods allow to solve practical problems of anthropology or medical image analysis in a desirable way.

In this paper, apart from theoretical proposals we have also applied robust statistical methods to practical image analysis problems. In Section 3 we apply a method for the template optimization, which allows to solve the problem of face detection in 100% of images in a dataset of 212 images. In Section 5 we apply a robust classification rule based on a new MWCD estimator, which allows to reduce the dimension in images and consequently localizes the face correctly in 100% of images in the given database. Finally Section 6 applies robust correlation analysis to the task of face recognition; this method measures the symmetry of faces in a robust way. In all these three applications the anthropological tasks of face detection or face recognition requires a decomposition of variability. The procedures based on robust statistical methods turn out to possess robustness properties to noise and small variations in the size or rotation of the face and can be interpreted as promising and reliable tools for robust image analysis.

As a future research we intend to implement a 3D image analysis method for forensic analysis of skulls and faces. Geometric morphometrics based on a certain point of landmarks is commonly used to compare a skull of a crime victim with a 2D image of the face of a missing person [25]. Here we propose to objectify the specification of landmarks. The discrimination between the image corresponding to the particular skull and all other images can be maximized by an analogy of the method of Section 3 of this paper. We also plan to work on the problem of identifying assassinated persons by computer anthropological modeling of the original face appearance based on the skull.

Another topic of a future research is an alternative procedure for the image analysis of gene expression data measured by the microarrays technology. Standard automated procedures turn out to be too vulnerable to outlying values and spatial noise or artifacts [18]. This makes the differential expression analysis sensitive to random or systematic errors in the scanned images of fluorescence intensities and robust procedures can again ensure a resistant and reliable solution.

Acknowledgments

The research is supported by project 1M06014 of the Ministry of Education, Youth and Sports of the Czech Republic.

References

1. Pearson K.: On lines and planes of closest fit to systems of points in space. Philosophical Magazine 1901, 2 (6), 559–572.
2. Galton F.: Co-relations and their measurement, chiefly from anthropometric data. Proceedings of the Royal Society 1888, 45, 135–145.
3. Stigler S.M.: Francis Galton's account of the invention of correlation. Statistical Science 1989, 4 (2), 73–79.
4. Fisher R.A.: The use of multiple measurement in taxonomic problems. Annals of Eugenics 1936, 7, 179–188.
5. Van Vark G.N., Howells W.W. (Eds): Multivariate statistical methods in physical anthropology. A review of recent advances and current developments. Reidel Publishing Company, Dordrecht 1984.
6. Ramsay J.O., Silverman B.W.: Applied functional data analysis. Second edition. Springer, New York 2005.

7. Kalina J.: Facial symmetry in robust anthropometrics. Journal of Forensic Sciences. Accepted, to appear in 2012.
8. Zankl A.: Computer-aided anthropometry in the evaluation of dysmorphic children. Pediatrics 2004, 114 (3), 333–336.
9. Böhringer S., Vollmar T., Tasse C., Würtz R.P., Gillessen-Kaesbach G., Horsthemke B., Wieczorek D.: Syndrome identification based on 2D analysis software. European Journal of Human Genetics 2006, 14, 1082–1089.
10. Zvárová J., Hanzlíček P., Nagy M., Přečková P., Zvára K., Seidl L., Bureš V., Šubrt D., Dostálová T., Seydlová M.: Biomedical informatics research for individualized life-long shared healthcare. Biocybern. Biomed. Eng. 2009, 29 (2), 31–41.
11. Bookstein F.L.: Morphometric tools for landmark data. Geometry and biology. Cambridge University Press, Cambridge 1991.
12. Dryden I., Mardia K.: Statistical shape analysis. Wiley, Chichester 1998.
13. Farkas L.: Anthropometry of the head and face. Raven Press, New York 1994.
14. Hastie T., Tibshirani R., Friedman J.: The elements of statistical learning. Springer, New York 2001.
15. Jurečková J., Picek J.: Robust statistical methods with R. Taylor & Francis, Boca Raton 2005.
16. Stigler S.M.: The changing history of robustness. American Statistician 2010, 64 (4), 277–281.
17. Heritier S., Cantoni E., Victoria-Feser M.P., Copt S.: Robust methods in biostatistics. Wiley, New York 2009.
18. Kalina J.: Robust image analysis in the evaluation of gene expression studies. ERCIM News, European Research Consorcium for Informatics and Mathematics 2010, 82, 52.
19. Kalina J.: Robust image analysis of faces for genetic applications. European Journal for Biomedical Informatics 2010, 6 (2): 6–13.
20. Davies P.L., Gather U.: Breakdown and groups. Annals of Statistics 2005, 33 (3), 977–1035.
21. Víšek J.Á.: The least weighted squares I. Bulletin of the Czech Econometric Society 2001, 9 (15), 1–28.
22. Čížek P.: Semiparametrically weighted robust estimation of regression models. Computational Statistics and Data Analysis 2011, 55 (1), 774–788.
23. Kalina J.: Some diagnostic tools in robust econometrics. Acta Universitatis Palackianae Olomucensis, Facultas Rerum Naturalium, Mathematica 2011, 50 (2), 55–67.
24. Rousseeuw P.J., Leroy A.M.: Robust regression and outlier detection. Wiley, New York 1987.
25. Rak R., Matyáš V., Říha Z. (Eds.): Biometrics and human identity in forensic and commercial applications. Grada Publishing, Prague 2008 (in Czech).
26. Yang M.H., Kriegman J., Ahuja N.: Detecting faces in images: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence 2002, 24 (1), 34–58.
27. Kalina J.: Locating landmarks using templates. In Antoch J., Hušková M., Sen P.K. (Eds.): Nonparametrics and robustness in modern statistical inference and time series analysis. Institute of Mathematical Statistics, Beachwood 2010, 113–122.
28. Loos H.S., Wieczorek D., Würtz R.P. Malsburg von der C., Horsthemke B.: Computer-based recognition of dysmorphic faces. European Journal of Human Genetics 2003, 11, 555–560.
29. Enquist M., Arak A.: Symmetry, beauty and evolution. Nature 1994, 372, 169–172.
30. Rousseeuw P.J., van Driessen K.: A fast algorithm for the minimum covariance determinant estimator. Technometrics 1999, 41 (3), 212–223.
31. Hubert M., van Driessen K.: Fast and robust discriminant analysis. Computational Statistics and Data Analysis 2004, 45, 301–320.
32. Todorov V., Pires A.M.: Comparative performance of several robust linear discriminant analysis methods. REVSTAT Statistical Journal 2007, 5, 63–83.
33. Croux C., Haesbroeck G.: Principal component analysis based on robust estimators of the covariance or correlation matrix: influence functions and efficiencies. Biometrika 2000, 87, 603–618.

34. Hubert M., Rousseeuw P.J., Vanden Branden K.: ROBPCA: A new approach to robust principal component analysis. Technometrics 2005, 47, 64–79.
35. Härdle W., Simar L.: Applied multivariate statistical analysis. Second edition. Springer, New York 2007.
36. Croux C., Filzmoser P., Oliveira M.R.: Algorithms for projection-pursuit robust principal component analysis. Chemometrics and Intelligent Laboratory Systems 2007, 87 (2), 218–225.
37. Víšek J.Á.: Consistency of the least weighted squares under heteroscedasticity. Kybernetika 2011, 47 (2), 179–206.