

Geometric morphometry as a new method of differentiation of intrabreed types in horse breeding

*A.V. Borisova**, *A. E. Shemarykin*, *A.A. Datsyshin*, and *G.V. Koroleva*

FSBSI VRI of Horse Breeding, Ryazan region, Rybnovsky district, Divovo settlement, Russia

Abstract. Assessment of type severity and exterior correctness is a necessary breeding measure in breeding work with breeds. Traditionally assessed features demonstrate subjectivism. The assessment of the type severity and the exterior correctness was initially based on the opinion of experts on whether the animal corresponds to a certain type of breed or not. The new method of geometric morphometry, based on comparing objects by their shape without considering differences in size, is of undoubted interest for research aimed at studying the physique features of horses of various intrabreed and breeder types. The key task of the geometric morphometry method is to evaluate the distinctive features of objects in shape, excluding the influence of their differences in size. It is based on the concept of a multidimensional space of forms. The article presents the results of testing the application of the geometric morphometry method for the differentiation of intrabreed horse types of purebred Arabian and performance horse breeds. It is revealed that geometric morphometry within the framework of a theoretical approach can be used as an additional tool to traditional methods of assessing the phenotypic characteristics of horses of various intrabreed types.

1 Introduction

The exterior correctness and the breed type severity is a necessary condition for breeding in horse breeding. At the same time, there are no clear objective criteria for such parameters as the severity of the breed type or the intrabreed (breeding) type. These signs are evaluated visually, and the final decision is based on the expert's judgments. Such an assessment system carries a large share of subjectivity, its quality depends on the specialist professionalism, knowledge of breed characteristics and the requirements imposed on it. Therefore, new methods of assessing the phenotypic characteristics of animals are needed.

Geometric morphometry is a new method of comparing shapes, which is currently actively developing. This is a set of techniques for multidimensional analysis of point coordinates that describe the configuration of an object in a complex [1,2,15,16].

The method of geometric morphometry, widely used in biology and anthropology, allows to compare the shape of objects of various sizes to obtain an average shape of the object under study.

* Corresponding author: vniik63@mail.ru

In modern scientific research, various morphometric methods have been used in determining the breed of honey bees [5] and in studying the intraspecific variability of fish [18].

The application of the geometric morphometry method in animal husbandry is increasingly popular in world practice, where geometric morphometry has been well justified within the framework of the theory of shape analysis and confirmed that it can be combined with traditional methods in this field to characterize the patterns of phenotypic and genetic variability of the shape and size of the lactation curve in dairy cattle [19].

Geometric morphometry is a new approach to comparing objects by their shape, without considering differences in size, and this is confirmed by a large number of publications in this field, especially with its application in horse breeding. Conformational and morphometric characteristics have been studied by foreign scientists in various horse breeds [7,8,9,10,11,12].

Geometric morphometry, which includes more possibilities of interpretation than linear measurements, has opened up a new perspective on veterinary anatomy. It contains useful background information, especially in the field of gender analysis and taxonomy [21].

The method of geometric morphometry has found application for the phenotypic assessment of the exterior of horses based on photography. Modern research aimed at developing methods of morphometric assessment will allow switching to digital methods of controlling the breeding process. This will allow automating research in the field of exterior assessment of horses, creating a system for analyzing and processing photographs to identify exterior features when studying the issue of differentiation of intrabreed types.

The main feature of geometric morphometry is that it standardizes all shapes by translating, rotating and superimposing them according to their average configuration. This process creates new coordinates that can be used for further analysis and do not depend on the actual size, and at the same time describe the shape change [16].

In the domestic practice of horse breeding, research aimed at using the geometric morphometry method to solve breeding problems has recently begun [4,5].

2 Materials and methods of research

The material for the study was digital images of representatives of various intrabreed types in horses of Arab purebred and performance breeds (Vladimir, Soviet performance, Russian performance). In total, 223 photos of stallions and mares of these breeds were processed.

In the TPSdig2 program, 72 marks were applied (Fig. 1) to parts of the horse's body that characterize the main distinctive exterior features of intrabreed (breeding) types.

The alignment of these data was carried out using generalized Procrustean analysis, which removes the effect of size (scaling, rotation, coordinate shift, minimizing the sum of the square between the labels of the same name) and allows to obtain an average model of each type of horse. As a result, both long-distance types were compared with each other and with an average model.

The processing of the obtained data was carried out using statistical analysis: the method of main components, the package "R" or "Statistica".

The tpsUtil program was used to create the primary tps file. In this file, the sequence of calling image files for the tpsDig2 program was formed. The tpsDig2 program was used for marking. The analysis was performed in the MorphoJ program. The analysis was performed using a discriminant function to find differences between groups.

Then two groups of horses of the Russian performance breed and two groups of horses of the Arabian pure breed were formed. The groups were formed by year of birth. According to the Russian performance breed - horses born before 1990 and horses of modern breeding stock. Purebred Arabian horses are winners of the Derby prize at

racetracks from 1946 to 1990 and from 1991 to 2022. These groups were also divided by gender. 84 photographs of Arabian purebred horses and 139 photographs of performance horses of different time periods were taken for processing. The photos were aligned according to the first photo and Procrustean transformations were performed, thereby getting rid of different dimensions. The location of the marks in the Procrustean coordinate system has been clarified.

Next, the covariance matrix was calculated, then the principal component analysis (PCA) was performed. The principal component method (eng. Principal Components Analysis (PCA) is one of the main ways to reduce the dimensionality of data by losing the least amount of information. Principal Component Analysis (PCA) is a method that allows to reduce a large set of variables to several dimensions that represent most of the variations in the data. PCA is calculated by its own decomposition of the sample covariance matrix and represents a rigid rotation of data while preserving procrustean distances between samples. Estimates of the main components are projections of figures onto a low-dimensional space covered by eigenvectors. Invented by K. Pearson (eng. Karl Pearson) in 1901.

Further research was carried out using the Canonical Variation Analysis (CVA) method. Canonical variation analysis (CVA) is a method used to determine the shape features that best distinguish several groups of samples. It is assumed that membership in the group is known a priori. Canonical variation analysis is used to analyze the group structure in multidimensional data. Canonical variation axes are directions in multidimensional space that maximally separate (discretize) predefined interest groups specified in the data. Unlike PCA, canonical variation axes, as a rule, are not orthogonal in the space of the original variables. Canonical variables explain the largest proportion of the difference between two sets of variables [13,14].



Fig. 1. Marking of photo with tags on the example of a Russian performance horse.

The images of horses in the GIMP graphic editor (free software) were marked with reference points, each of which was assigned a letter of the Latin alphabet. Using photographs in the ImageJ program (free software), the lengths of the vector between the reference points (pixels) and angles (degrees) were measured. According to the measurement data, the ratios of individual measurements were calculated among themselves.

3 Results and Discussion

At the first stage of research, after Procrustean alignment and the creation of the "Generate Covariance Matrix", the averaged configuration of the labels and the spread of coordinates of each individual around them were outlined. The data was checked for outlier errors using the author's Python script and the MorphoJ program's Find Outliers tool. The existing errors have been corrected. A visual comparison was made of the differences between the average models of horses of each intrabreed species (Fig. 2), from the general average horse model in our study, using the example of horses of the Soviet performance breed. We performed an analysis of the discriminant function to find differences between groups. Discriminant function analysis (DFA) studies the separation between groups of observations. The groups are known a priori. This analysis showed that there are no significant differences between the groups. At the final stage, after calculating the covariance matrix and conducting principal component analysis (PCA), it was not possible to identify significant differences between the groups in the studied breeds. The first three main components cover almost 80% of the variability. The peculiarities of shape change depending on the intrabreed species have not been revealed.

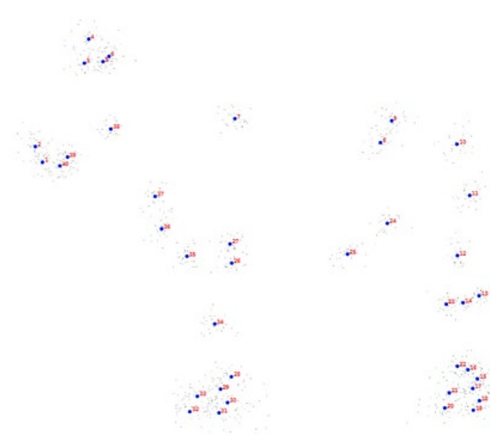


Fig. 2. Visualization of the consensus (averaged) configuration of points and the spread of coordinates of points of each individual around them.

After research by the method of Canonical Variation Analysis (CVA) in Soviet performance breeds of different factory types, the shape change along the CV1 axis showed that the greatest changes were observed at points 21 and 46 (knee joint and shoulder - scapular joint).

The shape change along the CV2 axis showed that the influence of the second canonical variable on the body height and the back length was revealed.

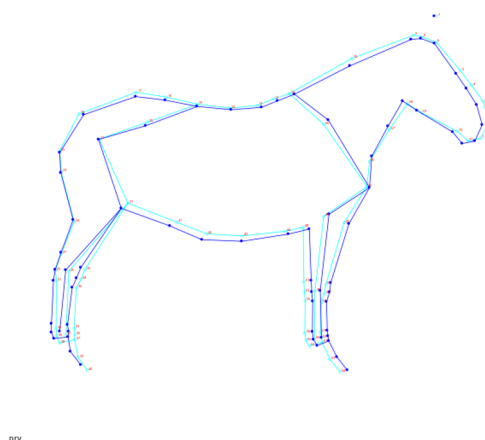


Fig. 3. Sexual dimorphism of the Perevoz type (blue color is an average model of Perevoz type mares, blue color is an average model of Perevoz type stallions).

In the Arabian pure breed, using the method of canonical variables "CVA", it was possible to isolate horses of the Hadban type into a separate cluster. The difference in head size from the average model was revealed, as well as in the average models in the other three groups, there is also a clear difference in the structure of the topline, croup and thigh.

The study of the differences between the average forms in male and female individuals in the studied intrabreed types showed that the sex differences in horses of different factory types are expressed quite clearly, this is most clearly seen in the example of the Perevoz type in the Soviet performance breed (Fig. 3). The most pronounced differences are in the structure of the head, neck, and abdomen. Mares are more "big-bellied", with a neat head. The head of stallions is rougher and larger, the neck is thicker with a pronounced crest, the shoulder blade of stallions is more oblique, unlike mares.

After the formation of two groups for Russian performance and Arab purebred breeds, all data were centered: all variables are replaced by their deviations from the average value. Next, the calculation of the correlation matrix: for each variable, the correlation coefficient with all other variables was calculated. Then the transformation of the original variables into new main components. Using new principal components as variables for further analysis. This method failed to identify significant differences.

In this regard, the use of specialized programs used for image analysis and processing for automatic morphometry in the field of hippometry and exterior assessment of horses is becoming particularly relevant.

Canonical variation analysis (CVA) based on Procrustean coordinates allowed to determine the degree of differences in the types under consideration in shape. In Russian performance horses, the main differences occurred in the first canonical variant (CV1, describes 61.5% variability). In the Arabian pure breed (Fig. 4), the main differences occurred in the first three variables, covering almost 74% of the variability. The graphic images show clear differences between horses tested in the period before 1990 and horses tested in the period 1991-2022.

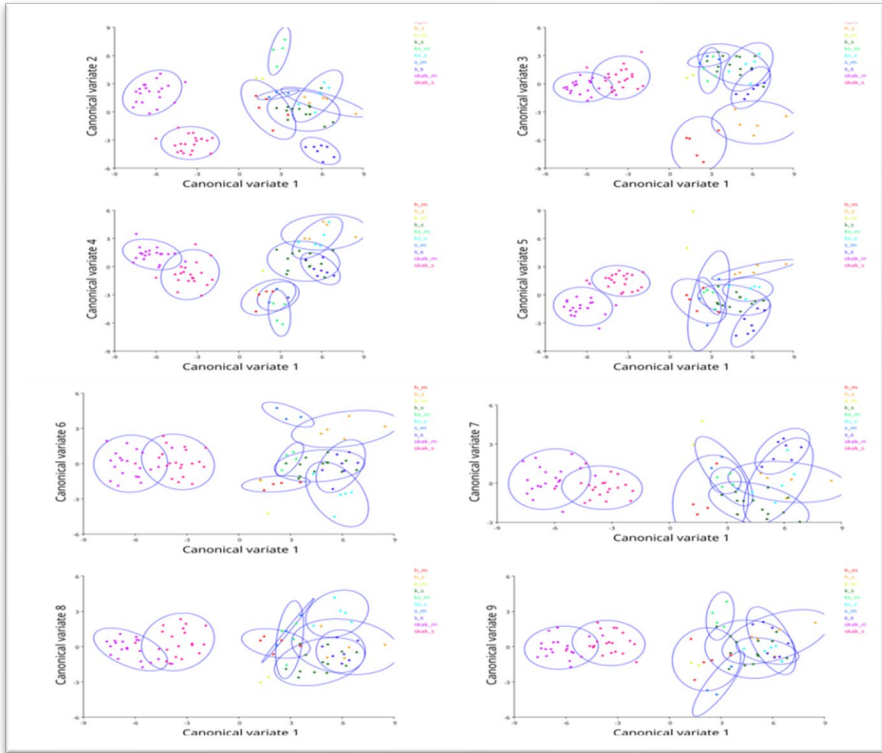


Fig. 4. Analysis by the method of canonical variables (CVA), horses of the purebred Arabian breed.

Reference points were applied to the images of horses in the GIMP graphics editor (free software). Using photographs in the ImageJ program (free software), the lengths of the vector between the reference points (pixels) and angles (degrees) were measured. According to the measurement data, the ratios of individual measurements were calculated among themselves. The calculation was carried out according to the formulas:

- format index (IF) = oblique body length / height at the withers;
- index of legginess (IDn) = length of the front leg / height at the withers;
- Overgrowth index (IP) = croup height / height at the withers;
- index DV/DN = length of top / length of bottom;
- index DZN/DPN = length of the back leg / length of the front leg;
- index DZN/VH = hind leg length / height at the withers;
- index DPN/VH = front leg length / croup height;
- index DZN/VH = hind leg length / croup height
- index DG/DSh = head length/neck length
- index DKr/DK = length of the croup / length of the body

After calculating in the Image J program at certain points, the ratio of individual measurements to each other, there were no major differences between the two different groups in horses of performance breeds. Minor differences were recorded in the ratio of the length of the hind legs to the length of the forelegs, the index of rearrangement, in the second group they are slightly larger. It can be concluded that this is due to the "rich" musculature of the croup in horses of the second type (born before 1990). This confirms the visual differences in Russian performance horses born before 1990 (Table 1)

Table 1. The ratios of individual measurements among themselves in Russian performance horses of different groups.

type	IForm ata	IDlinno nog	IPererr osl	DV/D N	DZN/DP N	DZN/V H	DPN/VK	DZN/VK	DLch/ DSh
1	1.1306	0.5442	1.0194	1.9410	1.12116	0.6095	0.5343	0.6095	0.6095
2	1.1390	0.5469	1.0305	1.9141	1.1310	0.6181	0.53119	0.61815	0.6181

4 Conclusion

1. It turned out that it is better to use automatic placement along the curve describing the animal silhouette to arrange the half-marks.
2. It was not possible to identify distinct shape differences between the groups along the axes of the first three main components.
3. The analysis of the discriminant function failed to identify significant differences between intrabreed types in body shape.
4. The study of the difference between the average configurations in male and female individuals in the Soviet performance breed showed that the gender differences in horses of different types are quite clearly expressed.
5. Procrustean transformations, calculation of the covariance matrix, and principal component analysis (PCA) showed no significant differences between the groups. Based on this, it can be stated that this method is not suitable for fixing differences between types of horses that are clearly visually fixed.
6. Canonical variation analysis (CVA) based on Procrustean coordinates allowed to determine the degree of differences in the types under consideration in shape. The main differences occurred in the first canonical variant in the Russian performance breed CV1, which describes 61.5% variability. In the Arabian pure breed, the main differences occurred in the first three variables, covering almost 74% of the variability.
7. In the Image J program, the ratio of individual measurements between each other did not show any big differences between two different groups in Russian performance horses. Minor differences were recorded in the ratio of the length of the hind legs to the length of the forelegs and the restructuring index, in the second group (horses born before 1990) similar indicators were higher. This confirms the visual differences in Russian performance horses of different birth periods.
8. Geometric morphometry within the framework of a theoretical approach, with certain modifications in the methodology, can be used as an additional tool to traditional methods of assessing the phenotypic characteristics of horses of various intrabreed and breeder types.

References

1. A.G. Vasil'ev, I.A. Vasil'eva, A.O. Shkurikhin, Geometric morphothery: from theory to practice (2018).

2. I.Ja. Pavlinov, N.G. Mikeshina, Journal of General Biology, **63(6)**, 473-493 (2002).

3. L.Ya. Pavlinov, *Geometric morphometrics, a new analytical approach to comparison of digitized images, Information technologies in biodiversity research* (St. Petersburg, 2001) 41-90.

4. A.V. Borisova, A.V. Sanganaeva, A.A. Datsyshin, *Izvestiya of Saint-Petersburg State Agrarian University*, **68(3)**, 103–111, (2022). doi: 10.24412/2078-1318-2022-3-103-111.
5. A.E. Shemarykin, A.A. Datsyshin, G.V. Koroleva, R.A. Khalilov, *Horse breeding and equestrian sports*, **6**, 7-9 (2022.) DOI 10.25727/HS.2022.6.60406. – EDN UQKIAU.
6. A.S. Berezin, *Biomics*, **11(2)**, 167-189 (2019). DOI: 10.31301/2221-6197.bmcs.2019-16 (In Russian)
7. K. Bär, P. Bosch, M. Rensing, U. Bostrel, J. *Equine Veterinary Science*, **78**, 53-59 (2019).
8. F.L. Bookstein, *Morphometric Tools for Landmark Data Geometry and Biology* (Cambridge University Press, 2003).
9. T.Druml, M. Horna, G. Grilz-Seger, M. Dobretsberger, G. Brem, *Arch. Anim. Breed*, **61**, 79-85 (2018).
10. M. Holmström, L.E. Magnusson, *Equine Vet J.* **22**, 186-193 (1990).
11. M.H. Sadek, A.Z. Al-Aboud, A.A. Ashmawy, *J. Anim. Breed Genet.* **123**, 369-377 (2006). Salako
12. B.J. Takaendengana, R.R. Noorc, S. Adiania, *Morphometric Characterization of Minahasa Horse for Breeding and Conservation Purposes* (Media Peternakan, Agustus, 2011) 99-104.
13. P. Zechner, F. Zohman, J. Solkner, I. Bodo, F. Habe, E. Marti, G. Brem, *Livestock Production Science*, **69(2)**, 163-177 (2001).
14. Canonical analysis as a method for studying intergroup variation, Internet resource <https://helpiks.org/7-2422.html>
15. J.M. Birch, *J. Mammal*, **78**, 1187-1198 (1997).
16. N. MacLeod, Phylogenetic signals in morphometric data, *Morphometrics, shape, and phylogeny*, N. MacLeod, P.L. Forey (Taylor & Francis, 2002) 100-138.
17. F. Ruttner, *Biogeography and taxonomy of honey bees* (Springer-Verlag; Berlin, Germany, 1988) doi: 10.1007/978-3-642-72649-1
18. N.V. Gordeeva, O.G. Nanova, *Journal of Ichthyology*, **57(1)**, 29-36 (2017).
19. A.A. Durón-Benítez, J.I. Weller, E. Ezra, *J. Dairy Sci.* **101(12)**, 11132-11142 (2018). doi: 10.3168/jds.2018-15209. Epub 2018 Sep 27.
20. N. Manuta, E. Ozkan, O. Kahvecioğlu, *Veterinariy*, **72(1)**, 15-27 (2023).