Morphable model of quadrupeds skeletons for animating 3D animals

Lionel Reveret, Laurent Favreau, Christine Depraz, Marie-Paule Cani

Eurographics/ACM SIGGRAPH Symposium on Computer Animation (2005)
K. Anjyo, P. Faloutsos (Editors)

Morphable model of quadrupeds skeletons for animating 3D animals

Lionel Reveret, Laurent Favreau, Christine Depraz, Marie-Paule Cani

GRAVIR, INRIA

Abstract

Skeletons are at the core of 3D character animation. The goal of this work is to design a morphable model of 3D skeleton for four footed animals, controlled by a few intuitive parameters. This model enables the automatic generation of an animation skeleton, ready for character rigging, from a few simple measurements performed on the mesh of the quadruped to animate.
Quadruped animals - usually mammals - share similar anatomical structures, but only a skilled animator can easily translate them into a simple skeleton convenient for animation. Our approach for constructing the morphable model thus builds on the statistical learning of reference skeletons designed by an expert animator. This raises the problems of coping with data that includes both translations and rotations, and of avoiding the accumulation of errors due to its hierarchical structure. Our solution relies on a quaternion representation for rotations and the use of a global frame for expressing the skeleton data. We then explore the dimensionality of the space of quadruped skeletons, which yields the extraction of three intuitive parameters for the morphable model, easily measurable on any 3D mesh of a quadruped. We evaluate our method by comparing the predicted skeletons with user-defined ones on one animal example that was not included into the learning database. We finally demonstrate the usability of the morphable skeleton model for animation.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics: Animation]:

1. Introduction

Skeleton construction and articulation placements are the first steps of character rigging. They involve the definition and adjustment of numerous degrees of freedom, namely the 3D position and orientation for each skeleton joint. These complex tasks are usually performed by a skilled animator. Tackling the problem in the case of virtual animals is even more complex than for virtual humans, since less anatomical data is available.

This paper shows that statistical analysis can be applied on a small set of skeleton models built by an expert animator to generate a morphable model of quadrupeds skeletons, easily adaptable to a wide variety of animals.

In the trade-off between fully procedural methods versus data-oriented ones, morphable models have recently become very popular in Computer Graphics. They offer access to high quality data through a practical parametrization that builds on predictive parameters learned from statistical analysis. In [BV99], a morphable models of face 3D shapes and texture is learned from hundreds of accurate laser scans of human subjects. It offers control over intuitive parameters such as age, sex, mood, etc. Similarly a morphable model of body shape has been proposed from laser scans of body shapes [ACP03]. Morphable models outperform simple scaling or FFD-like transformation by allowing to always maintain the result within a plausible space characterized by the learning examples.

For the first time, this paper investigates the generation of a morphable model in the specific case of animation skeletons. This raises the problem of using continuous interpolation over data that represents both rotational angles and limbs lengths. In particular, the parameterization of 3D rotations may present singularities (such as gimbal lock for euler angles) and is not unique (2π-periodicity), which makes its use more difficult in a statistical model. In addition, the morphable model has to take into account values defined in different units (e.g. distances and angles). Finally, to be

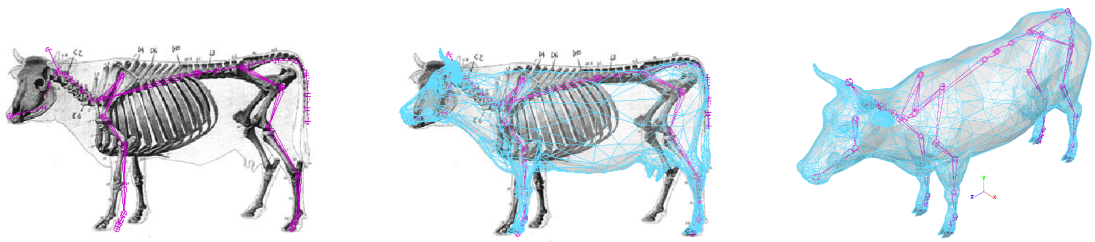

Figure 1: *Constructing learning data - the cow example*

practical, a morphable model should offer appropriate control parameters in the sense of being general enough to offer enough variability in the generated results and specific enough to maintain a good level of intuitive usability. This paper addresses these problems and shows how to build a successful and intuitive morphable model of skeletons for quadrupeds animals.

2. Related work

Jane Wilhems and Allen Van Gelder presented one of the first animation methods for animals [WG97]. Their approach is anatomical and thus relies on an accurate modeling of bones, muscles and skin tissues. We target a different goal, namely the automatic construction of animation skeletons (defined as a hierarchy of joints) dedicated to the efficient animation of animals through a standard geometry attachment method such as smooth skinning. We thus use a statistical analysis over a sample set of animation skeletons rather than anatomical modeling.

Wade and Parent use a medial axis computation to automatically create geometric skeletons inside polygonal meshes of animals [WP00]. The geometric skeleton is cleaned-up and automatically updated so that it can be used for character animation. Our approach is different since we rather learn the placement of joints on reference skeletons, for which an expert animator has placed the joints in accordance with the anatomical data of the associated quadruped. As these examples show, most of the animator-defined joints are not located onto the medial axis of the animal's mesh (see figure 1 for instance). Although less automatic, our approach guarantees that the skeleton we generate for a new animal will be closer to the structure a skilled animator would have built.

A model of animal growth has been proposed by Walter et al. [WF97, WFM01]. This approach covers the aspects of the skeleton, body shape and texture adaptation to represent growth of the animal. We investigate here another source of variation, which is related to change in the skeleton morphology over different species of quadrupeds in our case.

Sumner and Popovic tackle the problem of retargeting the animation of an animal towards another animal's mesh [SP04]. Motion warping is performed directly on triangle meshes, skipping the use of animation skeletons. Our choice is rather to focus on the underlying skeleton structure as the basic component for character animation. This allows us to insert our morphable model into the standard work-flow of 3D character animation.

The theoretical approach in Grochow et al.'s work [GMHP04] is closely related to ours. This work combines a large set of human motion capture data in an elaborated probabilistic model which can generate poses from geometric parameters such as IK handles. In our case, although we target a statistical model controlled by similar geometric parameters, we work on morphological variation between skeletons of animals in rest poses rather than on motion data. In consequence, we are not only trying to capture the variation of joint angles, but also the positions of these joints through the limbs length parameters. The resulting mix of parameter units has raised specific problems for which we discuss solutions in this paper.

3. Learning data set

To be useful, morphable models need accurate data for the learning phase. Laser scans and motion capture systems can provide such data for human body and shape. Obtaining data on animals skeletons is more challenging. As we target ready-to-animate models, we have decided to learn the morphable model on reference skeletons built by an expert animator (15 years experience in professional production).

3.1. Reference skeleton models

We have considered nine four footed animals covering a broad spectrum of morphologies : horse, goat, bear, lion, rat, elephant, cow, dog and pig. The reference skeletons we are using have first been built from anatomical references [EDB56, Cal75]. The latter provide 2D drawings of both the animal's internal and external anatomy, thanks to the outline of the body shape and the structure of the anatomical skeleton for a rest pose. However, these 2D drawings only give side view information. In order to design full 3D skeletons, 3D models of the animal's shape have been used as well. In order to ensure a correct alignment between data during the learning phase, all the skeletons share the same topology in terms of number of articulations and joints hierarchy. We used the standard convention in animation of taking the pelvis articulation as the root of the hierarchy. Each skeleton consists in 58 articulations with 6 degrees of freedom per articulation as they vary in position and orientation for each subject. Figure 1 shows some steps of the skeleton design in the example of a cow.

3.2. Parameterization of the data set

Each skeleton is parameterized as a single observation vector containing position and orientation information for each articulation. One first concern with statistical analysis is to clearly specify which variance will be considered. The metric system of 3D animation is unit less: a mesh may have any unit scale. We decided to get rid of any scaling effect over the data. Consequently, all the skeletons are normalized so that pelvis articulations are at the same location, and the articulation position are uniformly scaled so that the spine column has the same length for every animal. This leaves variability to be explored independently from the size of the animal.

We have also considered two alternatives for features parameterization:

- using euler angles versus quaternions for representing rotations angles
- using local axis coordinates versus global axis coordinates (i.e. world reference), for both translation and rotation values

In the following sections, we show and discuss the impact of these choices. For clarity, we will refer to these conditions as ER for Euler angles rotation, QR for Quaternion rotation, LC for local axis coordinate and GC for global axis coordinates.

Finally, each skeleton is represented by a $58*(3+3) = 348$ scalars vector for ER and $58*(3+4) = 406$ scalars vector for QR. LC and GC are obtained by standard manipulation of transformation matrices, followed by matrix conversion into euler angles or into quaternion space. This leads to 4 conditions to explore: ER×LC, QR×LC, ER×GC, and QR×GC.

When gathering data for all the animals, rotations parameters need a special care for dealing with discontinuities. The statistical analysis will linearly blend values of learning data, providing wrong results if two individual having similar rotation matrices are represented with very different parameters values, such as different 2π factors for Euler angles or opposite signs for quaternions. Moreover, Euler angles near the gimbal loose one degree of freedom leading to many-to-one problems. In each representation, ER or QR, these problems are checked and corrected when the 9 examples of the learning set are gathered.

Finally, we stack all the data in a single $\mathbf{X}$ matrix. We arrange the data in row vectors, each vector being a skeleton. $\mathbf{X}$ is thus a matrix of 9×348 for ER (or 9×406 for QR) scalars.

3.3. Normalization

Mixing data with different units such as rotation and translation parameters raises the issue of normalization. Translation data have been scaled so that the vertebral column has a length of 10. Rotational data are expressed in radians for ER condition and are left in the canonical range of $[-1,+1]$ for the QR condition. The table below reports for each translation and rotational parameters, the highest standard deviation computed over all animals chosen among all the articulations (for the sake of clarity we did not present the standard deviation for each articulation). It gives an idea on a how the learning data are scaled.

	ER×LC	ER×GC	QR×LC	QR×GC
tx	1.1	0.6	1.1	0.6
ty	0.5	3.2	0.5	3.2
tz	0.6	1.6	0.6	1.6
rx	0.1	2.1	0.1	0.2
ry	0.2	1.2	0.1	0.2
rz	0.6	2.0	0.2	0.2
rs	n.a.	n.a.	0.2	0.2

A common practice in statistical analysis consists in dividing each observation parameters by its standard deviation across the learning data set. Choosing a correct normalization can be stated as choosing the right trade-off between a model where no translation data are taken into account (all animals have the same limbs length) and a model where no rotational data are taken into account (all animals have the same articulations). In our case, it was not clear if one should be favored with respect to the other one. We conducted several experiments on normalization values to be applied and we did not report any significant impact on our ultimate goal of constructing a morphable model controlled by geometrical parameters.

4. Exploring dimensionality

Despite each skeleton is parameterized by 348 (or 406) scalars, the data set of 9 examples gives the model a maximum number of 8 linear degrees of freedom plus a mean pose. What we are interested in this section is to give a sense of redundancy between examples. The question is whether this redundancy can be factorized in less control parameters than 8 (similarly to shapes blending, skeletons blending would be a canonical parametrization of a morphable model). We studied this problem by first applying a

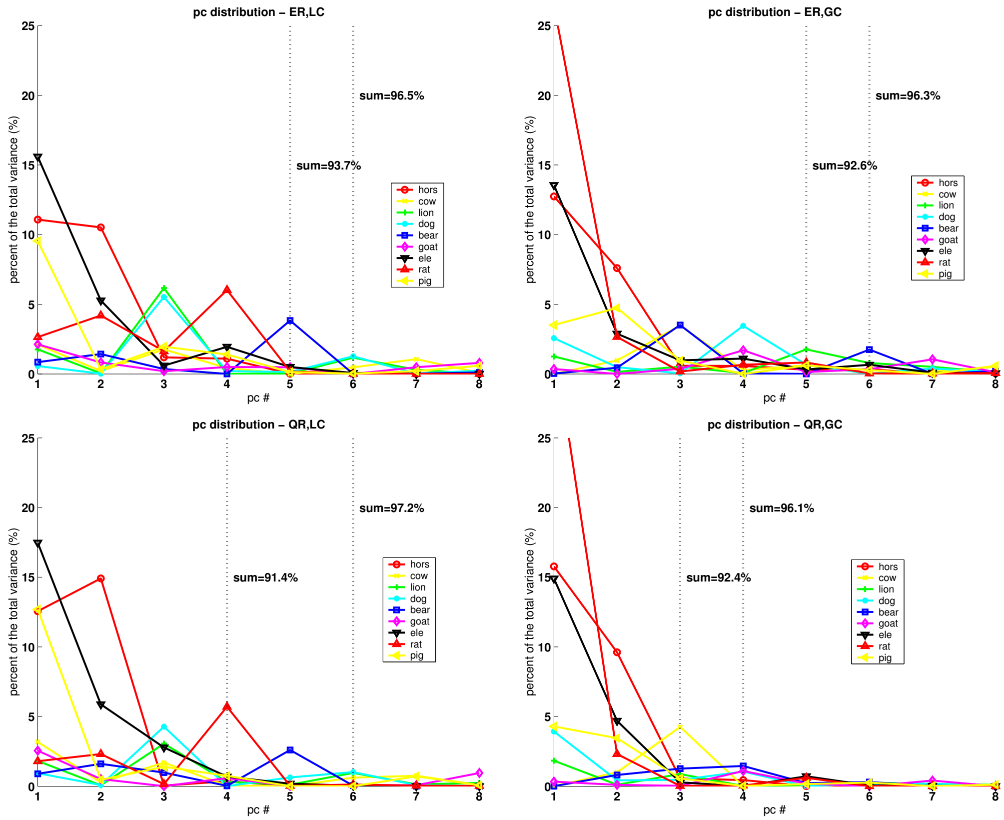

Figure 2: *Principal components distribution in 4 parameterizations*

standard PCA to the data set for each of the 4 conditions $\{ER, QR\} \times \{LC, GC\}$. In figure 2, we show the distribution of the principal components over all the subjects in the 4 features parameterization conditions. Let us detail how this distribution has been estimated.

Usually, PCA results are discussed in terms of the cumulative amount of variance of the data, by adding successive principal component. Figure 2 follows this idea but adds details to analyze results more precisely per individual of the data set. What is represented is computed as follow. PCA consists in computing the p eigenvalues $\mathbf{D}$ and p eigenvectors $\mathbf{E}$ of a data set $\mathbf{X}$ of n elements. $\mathbf{X}$ is centered on the mean skeleton vector. In our case, we have a maximum of $p = 8$ eigenvectors related to non-zero variance. $\mathbf{E}$ stacks eigenvectors in column.

$$\frac{1}{n}\mathbf{X}^t\mathbf{X}\mathbf{E} = \mathbf{E}\mathbf{D} \tag{1}$$

$$\mathbf{E}^t\mathbf{E} = \mathbf{1} \tag{2}$$

$\mathbf{D}$ is a $p \times p$ diagonal matrix where each element gives the total variance of the data explained by each of the p principal components. The total variance of the data set is given by $trace(\mathbf{D})$. Each column of the matrix $\mathbf{P} = \mathbf{X}\mathbf{E}$ gives the projection of all individuals on each eigenvector. Matrix $\mathbf{D}$ can be interpreted as:

$$\mathbf{D} = \frac{1}{n}\mathbf{P}^t\mathbf{P} \tag{3}$$

If we expand $\mathbf{D}$ along the row vectors $\mathbf{P_{i\cdot}}$ of $\mathbf{P}$, we get:

$$\mathbf{D} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{P^t_{i\cdot}}\mathbf{P_{i\cdot}} \tag{4}$$

Finally, as $\mathbf{D}$ is diagonal, we have for each element D_j:

$$D_j = \frac{1}{n}\sum_{i=1}^{n}P_{ij}^2 \tag{5}$$

The variance of each component can thus be decomposed according to the contribution of each one of the n individuals. On figure 2, we plot the squared value of the projection P_{ij}^2 of each individual in function of the *j-th* component, normalized by the total variance $trace(\mathbf{D})$. In addition, we indicate with a vertical line, at which component the cumulative *all-individuals* variance explained by successive components reached 90% and 95% of the total variance.

What we learn with figure 2 is that some parameterization leads to some over-fitting cases. Indeed, especially in the LC cases, the variance of a given component may be mostly related to a single individual. It means that this component is not factorizing information across the data set. The best results are obtained for the GC×QR condition (world coordinates and quaternion parameterization). In this parameterization, few over-fittings occur. Furthermore, 90% of the total variance is already captured using the 3 first components. This result suggests that, under the GC×QR condition, all the data can be efficiently linearly packed into a linear model controlled by only 3 parameters.

A new animation skeleton is simply generated by linear combination of the eigenvectors added to the mean skeleton $\overline{\mathbf{X}}$ of the learning data set:

$$\mathbf{x}(\mathbf{p}) = \overline{\mathbf{X}} + \mathbf{p}\mathbf{E}^{\mathbf{t}} \tag{6}$$

On figure 3, we show the results of the variation of the morphable model along the three first linear modes at minus two/one and plus two/one times the standard deviation of the parameter. The central column corresponds to the mean shape.

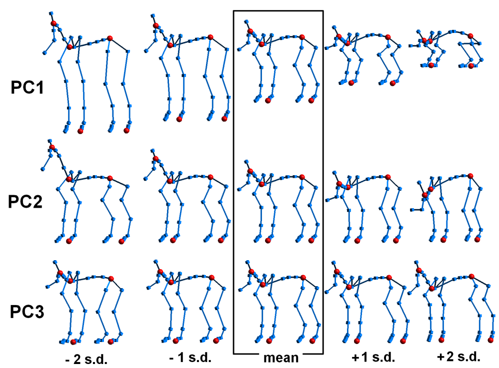

Figure 3: *Three first modes of variation of PCA on skeletons. Surprisingly, these modes are easy to identify as animal height in rest pose, bending of vertebral column and a hoofed* vs *plantigrad parameter.*

Although these parameters can be implicitly interpreted as animal height, bending of vertebral column and a hoofed *vs* plantigrad parameter, they would not be easy to align on a 3D model: their optimal values would be difficult to guess manually for a given mesh. We conclude this section by keeping in mind that 3 parameters might be sufficient. The next section investigates 3 parameters with a more practical usability where the morphable model is controlled explicitly by geometrical parameters.

5. Geometrical parameterization

In [ACP03], Allen et al. mention that even if PCA provides compact parameters, they are not always intuitive to use. Instead, their space of 3D body shapes is controlled with parameters such as weight and height to provide a more usable morphable model. We share the same approach for our morphable model of quadruped skeletons, making the alignment of the skeleton on a an arbitrary mesh much easier to perform. In our case, the control parameters are intuitive geometric values that can be measured on an animal's anatomical skeleton in side view (see figure 4). Indeed, similar geometric measurement can be pointed on the 3D mesh and applied to the morphable model. The generated skeleton will then be aligned with the mesh, granting that the morphable model is stable - in the sense that the geometric control parameters should generate a skeleton having the same geometrical measurement.

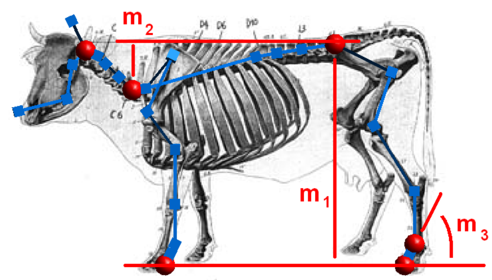

Figure 4: *The three geometrical measurement controlling the morphable model*

Based on the observation of the three main PCA modes, three geometrical parameters have been tested (to be compared with the 9 parameters used in [WF97, WFM01] for animal growth):

- animal height $\mathbf{m_1}$, measured on a mesh as the vertical distance between an estimated location of the pelvis and the floor;
- vertebral column bending $\mathbf{m_2}$ (we remind that vertebral column length is kept constant and serves as a scaling factor) measured as the difference between the animal height (defined as above) and the height of base of its neck; for the sake of clarity, the video shows an arrow from the ground to the neck for this parameter (similarly on the color plate). The true measurement is the one firstly described in this paragraph.
- hoofed *vs* plantigrad parameter, measured by the angle $\mathbf{m_3}$ between floor plane and the line joining the rear foot to the rear ankle.

These parameters are extracted on the reference skeletons by computing the global transformation matrices on pelvis, first neck cervical and rear ankle. As a first model, we perform a simple linear mapping from these three measurement parameters to skeletons rotation and translation. More elaborated mapping could be applied, such as Radial Basis

Functions (RBF) [LCF00] or Scaled Gaussian Process Latent Variable Model(SGPLVM) [GMHP04]. Linear mapping proved to be sufficient in our case. The mapping is done from a 3 scalars measurement vector $\mathbf{m} = [m_1, m_2, m_3]$ to a skeleton vector $\mathbf{x}$, both arranged as row vectors:

$$\mathbf{x}(\mathbf{m}) = \bar{\mathbf{X}} + \mathbf{m}\mathbf{V} \tag{7}$$

The linear model $\mathbf{V}$ is estimated by a least squares fitting between reference skeletons $\mathbf{X}$ and the matrix $\mathbf{M}$ of the three parameters measured on the reference skeletons stacked as row vectors. This leads to:

$$\mathbf{V} = (\mathbf{M}^{\mathbf{t}}\mathbf{M})^{-1}\mathbf{M}^{\mathbf{t}}\mathbf{X} \tag{8}$$

Figure **??** shows the variations of this morphable model along its three control parameters. Note that contrary to the third PCA mode, the third parameter correlates hoofed vs plantigrad with the length of neck (this correlation is indeed existing on the animal examples we provided; we did not investigate yet whether it is valid on all quadruped, which could be true for anatomical reasons). The neck length is itself correlated with the bending of the vertical column, as shown by both the second PCA mode in figure 3 and the second geometric parameter in figure 5.

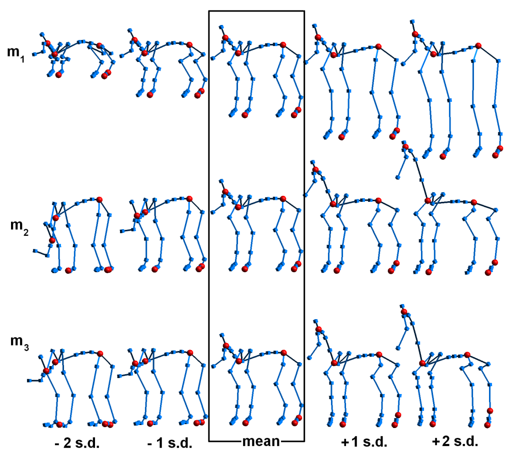

Figure 5: *Three first modes of variation of the morphable model controlled by animal height, vertebral column bending and hoofed* vs *plantigrad parameters (foot angle).*

Any parameters could be used as control parameters of the morphable model. We made some experiments in which the neck length was measured instead of the hoofed vs plantigrad parameter. It results in keeping the neck length constant when modifying the bending of the vertebral column as shown on figure 6.

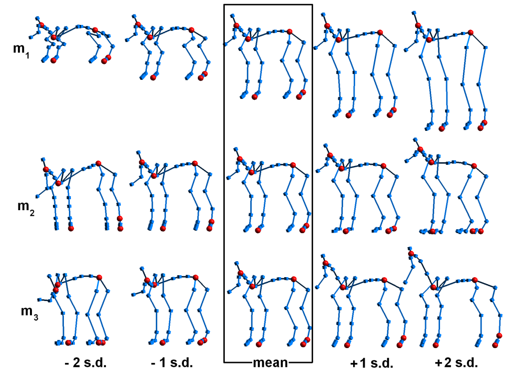

Figure 6: *Three first modes of variation of the morphable model controlled by animal height, vertebral column bending and neck length.*

6. Results and discussion

6.1. Reconstruction of the data-base

We discuss in this section the impact of the 4 conditions $ER, QR \times LC, GC$ on the properties of the morphable models obtained from both PCA and from our geometrically-controlled model. We evaluate the results based on how well the learning examples are correctly reconstructed by the morphable model from their associated input parameters. In the case of the PCA model, these inputs parameters are the projection coefficients of the learning examples onto the eigenvectors. In the case of the morphable controlled by geometrical parameters, these input parameters are the measurements made on the skeletons. As the size of the learning set is small and the number of parameters (three) is inferior to the number of learning examples (nine), non exact matching can be expected in both cases. We separate the evaluation in two parts:

1. error on translations
2. error on rotations

For the error on translations and rotations, all skeletons data, both reference and predicted using one of the 4 conditions, are re-encoded into a common representation to compare the 4 conditions on the same basis: global 4×4 transformation matrices are computed and split into their translation part (fourth column 3×1 vector) and rotation part (first 3×3 sub-matrix). For each individual, the error is estimated as the maximum over all the articulations of the norm of the difference between the reference skeleton and the predicted skeleton, computed on the translation vectors for translation and on rotation matrices for rotation (largest eigenvalue of the SVD is used in this case). Figure 7 (respectively Figure 8) summarizes the results for all the 9 examples under the 4 conditions using the PCA model controlled by the first three principal components (respectively the geometrically-controlled model). They show that the reconstruction of the data set is as good with our geometrically-controlled model than directly using PCA.

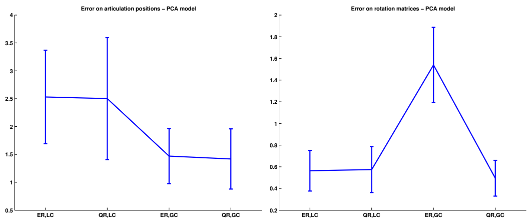

Figure 7: *Mean and standard deviation of the prediction error for the 4 conditions: translation (left), rotation (right) using the PCA-controlled model*

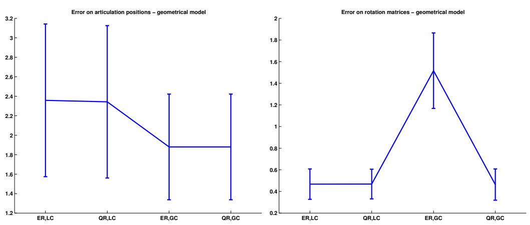

Figure 8: *Mean and standard deviation of the prediction error for the 4 conditions: translation (left), rotation (right) using the geometrically-controlled model*

For each of the models, these results show that the error on translation is smaller with the use of a global frame (GC condition), whatever the representation of rotations. This comes from the fact that in LC conditions, prediction errors are cumulative due to the hierarchical structure of the skeleton. This has maximum effect at limbs ends. The GC consition induces larger variation on rotational data. In this case, quaternions offer more stability than Euler angles. This explain the better results of $GC \times QR$ compared to $GC \times ER$. Figure 9 illustrates these two representations of data in the worst prediction case, which occurs for the elephant. To conclude, these observations suggests that global transformation parameters and quaternions representation are the best choice to build a morphable model of animals skeletons.

6.2. Animation of a new quadruped

In order to validate the whole process – from the skeleton generation to the resulting animation – we applied our model to an animal that was not present in the learning database: a cat. We generated the skeleton using the geometrically-controlled model from measurements taken on a side view of the cat mesh. The cat skin was then rigged to the skeleton using smooth skinning. An example of a final animation is provided in the video. In addition, the video illustrates the capabilities of motion retargeting offered by the morphable model. The same walking animation is applied on the skeleton while the parameters of the morphable model are continuously edited during the sequence.

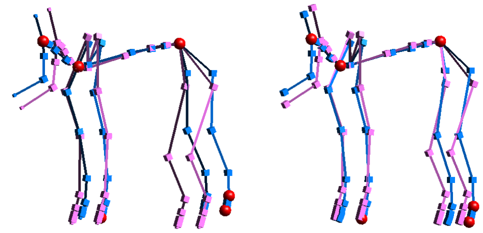

Figure 9: *Comparison between conditions $QR \times LC$ (left) and $QR \times GC$ (right) in the elephant case. The reference is depicted in purple and the predicted skeleton in blue.*

7. Conclusion

This paper has shown that statistical analysis can successfully be used for the automatically generation animation skeletons. We have tackled the problem in the specific case of four footed animals, using some skeletons built by an expert animator as the learning database as well as for validation. Our results have shown that a three dimensional space is sufficient for representing this set of animation skeletons. The resulting morphable skeleton model can easily be fitted to any quadruped by taking three simple measurements on a side view on an animal mesh. We have shown that it yields convincing animation results when combined with standard smooth skinning for geometry attachment. Indeed, expert animators may use this automatically generated skeleton as a starting point and probably reach a higher level of realism through fine tuning dedicated to each animal.

References

[ACP03] ALLEN B., CURLESS B., POPOVIC Z.: The space of all body shapes: reconstruction and parameterization from range scans. *ACM Transactions on Graphics (ACM SIGGRAPH 2003) 22*, 3 (2003), 587–594.

[BV99] BLANZ V., VETTER T.: A morphable model for the synthesis of 3d faces. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1999), ACM Press/Addison-Wesley Publishing Co., pp. 187–194.

[Cal75] CALDERON W. F.: *Animal painting and anatomy*. Dover Publications, Inc., New-York, 1975.

[EDB56] ELLENBERG W., DITTRICH H., BAUM H.: *An atlas of animal anatomy for artists*. Lewis S. Brown (eds), Dover Publications, Inc., New-York, 1956.

[GMHP04] GROCHOW K., MARTIN S. L., HERTZMANN A., POPOVIC Z.: Style-based inverse kinematics. *ACM Transactions on Graphics (ACM SIGGRAPH 2004)* (2004).

[LCF00] LEWIS J., CORDNER M., FONG N.: Pose space deformation: A unified approach to shape interpolation and skeleton-driven defomation. In *Proc. SIGGRAPH'00* (2000), pp. 165–172.

[SP04] SUMNER R. W., POPOVIC J.: Deformation transfer for triangle meshes. *ACM Transactions on Graphics (SIGGRAPH'04) 23*, 3 (August 2004).

[WF97] WALTER M., FOURNIER A.: Growing and animating polygonal models of animals. *Computer Graphics Forum (EUROGRAPHICS'97) 16*, 3 (1997).

[WFM01] WALTER M., FOURNIER A., MENEVEAUX D.: Integrating shape and pattern in mammalian models. In *Proc. SIGGRAPH 2001* (2001), pp. 317–326.

[WG97] WILHELMS J., GELDER A. V.: Anatomically based modeling. In *Computer Graphics (ACM Siggraph Proceedings)* (1997).

[WP00] WADE L., PARENT R. E.: Fast, fully-automated generation of control skeletons for use in animation. In *Proc. of Computer Animation* (2000), pp. 164–169.