Conditional Variability of Statistical Shape Models Based on Surrogate Variables

Rémi Blanc[1], Mauricio Reyes[2], Christof Seiler[1], and Gábor Székely[1]

[1] Computer Vision Laboratory, ETHZ,
Sternwartstrasse 7, 8092 Zürich, Switzerland
{blanc,szekely}@vision.ee.ethz.ch
[2] ARTORG Center for Biomedical Engineering Research, University of Bern,
Stauffacherstrasse 78, 3014 Bern, Switzerland
{christof.seiler,mauricio.reyes}@artorg.unibe.ch

Abstract. We propose to increment a statistical shape model with surrogate variables such as anatomical measurements and patient-related information, allowing conditioning the shape distribution to follow prescribed anatomical constraints. The method is applied to a shape model of the human femur, modeling the joint density of shape and anatomical parameters as a kernel density. Results show that it allows for a fast, intuitive and anatomically meaningful control on the shape deformations and an effective conditioning of the shape distribution, allowing the analysis of the remaining shape variability and relations between shape and anatomy. The approach can be further employed for initializing elastic registration methods such as Active Shape Models, improving their regularization term and reducing the search space for the optimization.

1 Introduction

Statistical Shape Models [1] (SSM) are increasingly used for medical image analysis. For segmentation purposes, they allow making use of prior anatomical knowledge for compensating low contrast and high noise level in the images. SSM are also used for regularizing elastic registration algorithms [2], so that the estimated shape is both anatomically plausible and matches the image information. Recent extensions use them for predicting unobserved parts of a shape [3,4], or even predicting the shape of one organ from the observation of another [5].

The general idea behind SSM is to perform a linear decomposition of the shape variability by defining modes of deformations through various mathematical criteria. The most commonly used, Principal Component Analysis (PCA), estimates orthogonal directions which iteratively maximize the variance. Other linear decompositions have been proposed, based on Principal Factor Analysis (PFA) methods [6], or Independent Component Analysis (ICA) [7,8] which define modes of deformations which are, among other characteristics, more localized and easier to interpret. However, these modes, resulting from purely mathematical criteria, seldom correspond to deformations which have a direct, anatomically meaningful interpretation [6]. [9]

proposed to estimate the remaining variability of a shape model after fixing a part of it, by calculating the conditional distribution under a multivariate normal assumption. However, due to rank deficiency problems, they had to allow for limited random deviations from the observations, too, in order to allow for any remaining variability. Consequently, the conditional distribution they estimated is strongly dependent on the level of tolerated variability. In the context of surgical simulators, [10] employed a non-linear optimization approach for estimating the most probable shape within a SSM, satisfying exact constraints on characteristic dimensions of the uterus. Nevertheless, the remaining shape variability is not explicitly assessed. The approach neither allows to consider generic patient information such as height, weight or age for restricting the shape model, while these certainly impose morphological constraints, e.g. for bony structures.

We propose in this paper to extend linear SSM description in order to provide with intuitive anatomical control over shape deformations, by explicitly integrating such anatomical information within the model. Section 2 recalls basics about SSM and introduces notations employed throughout the paper. In section 3, we present a method for conditioning the shape distribution based on a set of surrogate variables, in particular for non-parametric kernel densities. In section 4, we apply the method to a shape model of the human femur, and evaluate the effectiveness of the conditioning. Conclusions and perspectives are summarized in Section 5.

2   Statistical Shape Models

Statistical shape models aim at describing the natural variability of a shape, e.g. the morphological manifestations of an organ over different individuals or through time. Such models usually rely on a specific parameterization of a set of training shapes $\mathbf{z}_i$, $i \in \{1,...,n\}$, e.g. by a set of d landmarks in correspondence across the various shapes [11], usually lying on the shape's outline. Through this parameterization, the k-dimensional shape, $k \in \{2;3\}$, is stored as a column vector with $p = kd$ elements, and can be viewed as a single point in a p − dimensional space. A collection of different instances of a shape, e.g. the same organ observed for different individuals, then corresponds to a point cloud in the parameterized space. This point cloud contains the information about the shape's variability observed from the available samples and can be analyzed using multivariate statistical techniques [12].

The most classical approaches for shape modeling are linear, in the sense that any shape $\mathbf{z}$ is described as a linear combination of a set of r modes of deformations:

$$\mathbf{z} \approx \mathbf{Ub} + \mathbf{m}, \text{ with } \mathbf{m} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i \text{ and } \mathbf{b} = \mathbf{U}^T(\mathbf{z} - \mathbf{m}) \tag{1}$$

The different existing approach differ in the definition of the modes of deformation $\mathbf{U}$. PCA searches orthogonal modes which iteratively maximize the variance, and can make use of the Singular Value Decomposition to efficiently estimate them. The more general PFA can use other criteria, which usually define orthogonal rotations of the PCA modes. In particular, the varimax rotation [13] intends to cluster the modes of

deformations so that each mode has a large influence over only a limited number of variables, and nearly no impact on the others [6]. ICA includes higher order moments, such as skewness or negentropy [8], in order to find directions in which the shape deformations are as independent as possible. It was shown to perform better than PCA when the shape distribution is significantly non-gaussian [7].

Whichever method is chosen, the subspace of admissible shapes is spanned by the r modes of deformation retained by $\mathbf{U}$, and the $r-$ variate distribution of the model parameters $\mathbf{b}_i$, $i \in \{1,...,n\}$ observed on the training samples allow to define compact models for the shape distribution. In the following, we will restrict to PCA, for which the domain of plausible shapes is usually defined by constraining every parameter to be within a $[-3\sigma;3\sigma]$ interval, σ^2 being the corresponding eigenvalue.

3   Conditioning the Shape Distribution with Surrogate Variables

Though a shape is explicitly represented by its parameterization, using either $\mathbf{z}$ or more conveniently the r parameters $\mathbf{b}$, it is more intuitive to describe a shape through a set of anatomically meaningful measures, such as lengths, angles, curvatures, etc. Those measures can be obtained either through an automatic procedure or defined manually as functions of landmarks of a reference shape, taking advantage of established correspondences allowing a consistent definition across the different samples available. It is also possible to include surrogates which are not directly related to the organ's shape, but provide generic information about the patient, such as its age, height or weight, which is likely to correlate with morphology. This set of surrogate variables $\mathbf{x}_i$, $i \in \{1,...,n\}$ offer an easily interpretable description, and can be used for controlling the predicted shape.

By concatenating the surrogates $\mathbf{x}_i$ to the shape parameters $\mathbf{b}_i$, the set of n training examples form samples from the joint multivariate distribution $p(\mathbf{b},\mathbf{x})$, which can be exploited to condition the shape to follow prescribed anatomical constraints. Indeed, the conditional distribution $p(\mathbf{b}|\mathbf{x}=\mathbf{x}_0)$ of the shape parameters given specific values $\mathbf{x}=\mathbf{x}_0$ for the surrogates, is directly given by Bayes theorem:

$$p(\mathbf{b}|\mathbf{x}=\mathbf{x}_0) = \frac{p(\mathbf{b},\mathbf{x}_0)}{p(\mathbf{x}_0)} \tag{2}$$

where $p(\mathbf{x}_0)$ stands for the marginal probability density of $\mathbf{x}$ evaluated at $\mathbf{x}_0$.

Assuming that $p(\mathbf{b},\mathbf{x})$ follows a multivariate Gaussian distribution, with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$, the conditional distribution $p(\mathbf{b}|\mathbf{x}=\mathbf{x}_0)$ is known [12 p.87] to be also Gaussian, with mean and covariance given by:

$$\begin{aligned}
\boldsymbol{\mu}_{\mathbf{b}|\mathbf{x}_0} &= E\left[\mathbf{b}|\mathbf{x}=\mathbf{x}_0\right] = \boldsymbol{\mu}_{\mathbf{b}} + \boldsymbol{\Sigma}_{\mathbf{bx}}\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}\left(\mathbf{x}_0-\boldsymbol{\mu}_{\mathbf{x}}\right) \\
\boldsymbol{\Sigma}_{\mathbf{bb}|\mathbf{x}_0} &= Cov\left[\mathbf{b}|\mathbf{x}=\mathbf{x}_0\right] = \boldsymbol{\Sigma}_{\mathbf{bb}} - \boldsymbol{\Sigma}_{\mathbf{bx}}\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}\boldsymbol{\Sigma}_{\mathbf{bx}}^{T}
\end{aligned} \tag{3}$$

where the different variables are taken from $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$:

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{b}} \\ \boldsymbol{\mu}_{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} E[\mathbf{b}] \\ E[\mathbf{x}] \end{bmatrix} \text{ and } \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{bb}} & \boldsymbol{\Sigma}_{\mathbf{bx}} \\ \boldsymbol{\Sigma}_{\mathbf{bx}}^{T} & \boldsymbol{\Sigma}_{\mathbf{xx}} \end{bmatrix} = \begin{bmatrix} Cov(\mathbf{b},\mathbf{b}) & Cov(\mathbf{b},\mathbf{x}) \\ Cov(\mathbf{x},\mathbf{b}) & Cov(\mathbf{x},\mathbf{x}) \end{bmatrix} \tag{4}$$

Unless there are more surrogate variables than available samples in the training set, the only reason for $\boldsymbol{\Sigma}_{\mathbf{xx}}$ not to be invertible is that some surrogates are linear combinations of the others, which means that they could be safely ignored.

Nevertheless, the Gaussian assumption may not always be realistic, especially when considering surrogates related to distances or angles. A non-parametric alternative is to rely on kernel density estimation [14], and to model the joint density $p(\mathbf{x}, \mathbf{b})$ as follows, by averaging the contribution of each training sample:

$$p(\mathbf{b}, \mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} K_{\mathbf{H}} \left(\mathbf{b}_i - \mathbf{b}, \mathbf{x}_i - \mathbf{x} \right) \tag{5}$$

where $K_{\mathbf{H}}$ is a normalized multivariate kernel function with bandwidth matrix $\mathbf{H}$. Choosing a Gaussian kernel, $K_{\mathbf{H}}$ is the classical Gaussian density, with zero mean and covariance $\mathbf{W} = \mathbf{H}^2$. For this case it can be shown that the conditional mean and conditional covariance are given by:

$$\boldsymbol{\mu}_{\mathbf{b}|\mathbf{x}_0} = \sum_{i=1}^{n} w_i \boldsymbol{\mu}_{\mathbf{b}|\mathbf{x}_0}^{(i)} \ ; \ \boldsymbol{\Sigma}_{\mathbf{bb}|\mathbf{x}_0} = \sum_{i=1}^{n} w_i \left(\mathbf{W}_{\mathbf{bb}|\mathbf{x}_0} + \left(\boldsymbol{\mu}_{\mathbf{b}|\mathbf{x}_0}^{(i)} - \boldsymbol{\mu}_{\mathbf{b}|\mathbf{x}_0} \right) \left(\boldsymbol{\mu}_{\mathbf{b}|\mathbf{x}_0}^{(i)} - \boldsymbol{\mu}_{\mathbf{b}|\mathbf{x}_0} \right)^T \right) \tag{6}$$

with,

$$w_i = \frac{K_{\mathbf{H}_\mathbf{x}}\left(\mathbf{x}_0 - \mathbf{x}_i\right)}{\sum_{j=1}^{n} K_{\mathbf{H}_\mathbf{x}}\left(\mathbf{x}_0 - \mathbf{x}_j\right)} \ ; \ \boldsymbol{\mu}_{\mathbf{b}|\mathbf{x}_0}^{(i)} = \mathbf{b}_i + \mathbf{W}_{\mathbf{bx}}\mathbf{W}_{\mathbf{xx}}^{-1}\left(\mathbf{x}_0 - \mathbf{x}_i\right) \ ; \ \mathbf{W}_{\mathbf{bb}|\mathbf{x}_0} = \mathbf{W}_{\mathbf{bb}} - \mathbf{W}_{\mathbf{bx}}\mathbf{W}_{\mathbf{xx}}^{-1}\mathbf{W}_{\mathbf{bx}}^T \tag{7}$$

There is a general agreement that the "shape" of the kernel $K_{\mathbf{H}}$ is much less important than the choice of the bandwidth [14 chap.2]. However, full multivariate bandwidth selection is still a difficult issue. A data-driven method has been proposed in [16], where the optimal bandwidth is estimated through a Markov Chain Monte Carlo approach (MCMC). Unfortunately, this approach is still very expensive for high-dimensional distributions. A general rule of thumb for the bandwidth matrix is to choose $\mathbf{H}$ proportional to $\boldsymbol{\Sigma}^{1/2}$ [15 p. 152].

Once the conditional distribution has been calculated, random shapes following the prescribed constraints can be drawn by sampling from $p(\mathbf{b}|\mathbf{x} = \mathbf{x}_0)$. A sample from the conditional distribution can be generated by first selecting i with probability w_i and then drawing a sample from the Gaussian distribution with mean $\boldsymbol{\mu}_{\mathbf{b}|\mathbf{x}_0}^{(i)}$ and covariance $\mathbf{W}_{\mathbf{bb}|\mathbf{x}_0}$.

Such samples are classical shape parameters, and the plausibility of the corresponding shapes can be evaluated by observing their likelihood in the original, unconditional SSM. In the following, we rely on such random samples to evaluate the effectiveness of the conditioning, by comparing the prescribed values of the anatomical parameters with values measured on those shapes.

4 Application to the Human Femur

In this section, we illustrate the concepts developed above on a 3D statistical model of the human femur. The model consists of $n = 170$ femurs obtained from CT scans. The images have been registered using the non-rigid diffeomorphic registration

technique proposed by [17]. A mesh, consisting of $d = 23536$ vertices ($p = 70608$), was extracted from the segmentation of the reference bone, and propagated to the other samples using the correspondences obtained during the volumetric registration step. The SSM is defined by this set of meshes. The parameters $\mathbf{b}_i$, $i \in \{1,...,n\}$ were computed using a PCA, keeping 99,9% of the variance ($r = 144$) in order to preserve the shape distribution as much as possible while reducing the dimensionality of the problem and avoiding numerical instabilities in the computations of the densities.

A set of anatomical surrogate variables has been defined as functions of manually selected vertices of our model, based on [18] and as illustrated on Fig. 1. Thanks to the established correspondences, these measurements $\mathbf{x}_i$, $i \in \{1,...,n\}$ are defined in a reproducible way over shapes from the SSM. Additionally, the age (>21 years), height (between 150 and 181 cm) and weight (between 42 and 140 kg) of each subject was recorded. The gender was also recorded (55% males) but not used in the following.
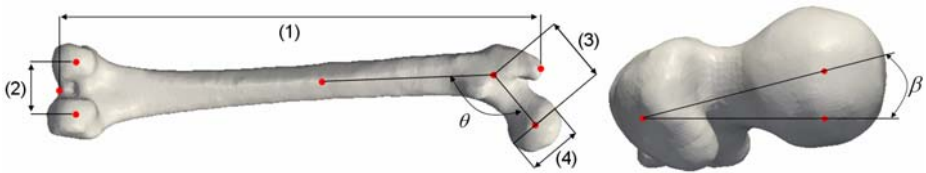
Fig. 1. Reference bone and definition of surrogate variables: femur length (**1**), inter-condyle distance (**2**), neck length (**3**), vertical head diameter (**4**), collo-diaphysal (neck/shaft) angle θ and anteversion angle β. The distances are stored in *mm*, and the angles in *degrees*.

The joint distribution $p(\mathbf{x}, \mathbf{b})$ was modeled through a gaussian kernel density. Using the rule of thumb presented above, the optimal bandwidth matrix was estimated as $\mathbf{W} = 0.59 \, \Sigma$. The retained proportionality factor has been obtained through the optimization of the Kullback-Leibler information criterion, as proposed by [16]. For different combinations of the variables used for conditioning, we estimated the corresponding conditional distribution $p(\mathbf{b}|\mathbf{x} = \mathbf{x}_0)$. Three such combinations used in the following are presented in Table 1. Due to imperfect shape representation through the SSM, a limited number of training shapes, and the fact that the estimation of the conditional distribution does no rely on any known relation between the surrogate variables and the position of the shape vertices, synthetic shapes generated from the conditional distribution may not follow perfectly the prescribed constraints. Consequently, we evaluated the effectiveness of the conditioning by drawing 5000 random shapes from the estimated conditional distribution, and computed the anatomical measurements described on Fig. 1. As can be seen on Table 2, the conditional distributions do not exhibit significant biases and present a low variance, indicating that the conditioning is quite effecitve.

By comparing the traces of the covariance matrices of the joint and conditional distributions for various choices of the conditioning surrogates, their relative influence on the variability of the full shape or of individual morphological parameters can be investigated. This also can provide information about correlations among the surrogates. For example, looking at "Conditions 3", conditioning on the age, weight and height already reduced the shape variability by nearly 50%, especially constraining the femur length while letting the angle β mostly unaffected.

Table 1. Three different choices of the conditioning variables and their values

Surrogates	(1)	(2)	(3)	(4)	θ	β	age	height	weight
Conditions 1	441.75	66.22	72.03	(free)	(free)	9.36	(free)	(free)	(free)
Conditions 2	(free)	63.28	64.52	46.03	115.16	11.07	(free)	(free)	(free)
Conditions 3	(free)	(free)	(free)	(free)	(free)	(free)	40	180	65

Table 2. Effectiveness of the anatomical conditioning. The "% bias" column indicates the observed bias on the prescribed values of the surrogates for the shapes of the conditional distribution. For the variables which were left unconstrained, the observed mean is given instead. The "% var" column indicates the remaining variance of the anatomical measures.

Surrogate variables	Initial distribution		Conditions 1		Conditions 2		Conditions 3	
	mean	std	% bias	% var	% bias	% var	% bias	% var
(1)	426.62	26.63	0.01 %	0.01%	(424.39)	(28.76%)	(458.60)	(32.75%)
(2)	56.68	6.50	-0.07 %	0.61 %	0.06 %	0.59%	(58.85)	(54.27%)
(3)	66.34	5.73	0.09 %	0.51 %	0.04 %	0.24 %	(71.22)	(50.45%)
(4)	46.55	4.88	(50.02)	(30.05%)	0.07 %	0.48 %	(48.79)	(46.98%)
θ	115.08	5.48	(112.33)	(34.35%)	-0.08 %	0.63 %	(115.33)	(58.21%)
β	10.87	2.15	1.33 %	1.34 %	0.14 %	1.32 %	(10.71)	(87.10%)
Remaining shape variability	15.03 %			34.06 %			47.74 %	
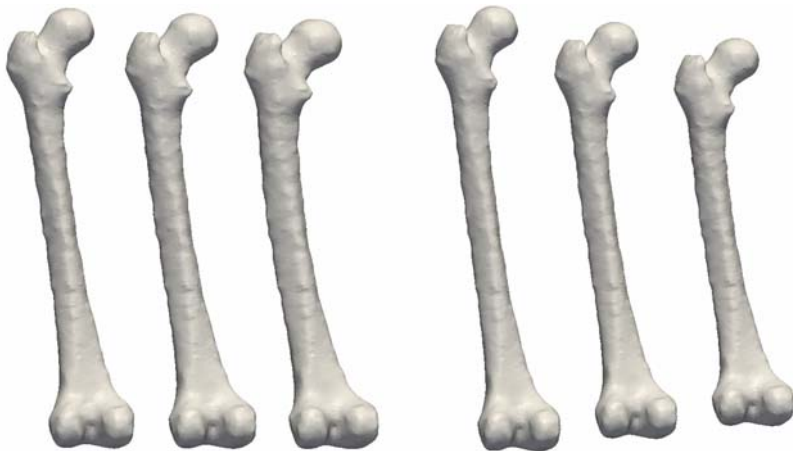
Fig. 2. The three femurs on the left (resp. right) correspond to samples for the Conditional Distribution 1 (resp. 2) defined in Table 1. For both groups, the central bone correspond to the conditional mean, while the bone on the left (resp. right) correspond to variations related to the main mode of the conditional shape covariance, with coefficient -2 (resp. +2).

The conditional covariance matrix $\Sigma_{bb|x_0}$ has been further diagonalized, so as to define orthogonal modes of deformations for the remaining shape variability. For the first two conditional distributions of Table 1, the shape corresponding to the conditional mean and deformations of +/-2 times the standard deviation of the main mode of remaining covariance are represented on Fig. 2. For conditional distribution 1, the femur length remains approximately constant while the remaining variability seems to be mostly related to the unconstrained neck/shaft angle θ. On the other hand, the femur length is clearly dominating the remaining variability of the 2nd distribution, while other parameters do not significantly vary. The obtained shapes were found to be generally plausible – from the initial SSM viewpoint –, as long as the constraints were themselves realistic.

5   Conclusion

In this paper we show how to incorporate explicit anatomical constraints in statistical shape models, in the form of surrogate variables. The conditional shape distribution allows analyzing the remaining shape variability, as well as drawing plausible random shapes following specific anatomical constraints. The method was demonstrated and validated on a statistical shape model of the human femur.

Besides the potential interest for anatomy and morphometry studies, the approach can be further used in conjunction with classical SSM-based registration methods, by providing an initialization of the shape following easily measurable surrogate variables, optimizing the likelihood term used for regularization and reducing the size of the search space. In particular, we plan to adapt the method for SSM of organs subject to respiratory motion, such as the liver, and to rely on the shape distribution conditioned by the phase within the breathing cycle, which could significantly accelerate tracking algorithms.

Finally, in our experiments, we experienced a clear deterioration in both the accuracy of the estimation of the anatomical parameters and in the effectiveness of the conditioning when fewer eigenmodes have been retained. We plan to investigate how this behaviour is influenced if the modes of deformation are defined through PFA or ICA instead of relying on PCA analysis. More generally, this raises the interest for dimensionality reduction techniques which would specifically consider the fidelity of the anatomical representation as a criterion.

References

1. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Burkhardt, H., Neumann, B. (eds.) ECCV 1998. LNCS, vol. 1407, pp. 484–498. Springer, Heidelberg (1998)
2. Kelemen, A., Székely, G., Gerig, G.: Elastic model-based segmentation of 3-D neuroradiological data sets. IEEE Trans. on Medical Imaging 18(10), 828–839 (1999)
3. Rajamani, K.T., Hug, J., Nolte, L.-P., Styner, M.: Bone Morphing with statistical shape models for enhanced visualization. In: Proc. SPIE, vol. 5367, pp. 122–130 (2004)
4. Rao, A., Aljabar, P., Rueckert, D.: Hierarchical statistical shape analysis and prediction of sub-cortical brain structures. Medical Image Analysis 12, 55–68 (2008)

5. Yang, Y.M., Rueckert, D., Bull, A.M.J.: Predicting the shapes of bones at a joint: application to the shoulder. Computer Methods in Biomech. and Biomed. Eng. 11(1), 19–30 (2008)
6. Reyes Aguirre, M., Linguraru, M.G., Marias, K., Ayache, N., Nolte, L.P., Gonzalez Ballester, M.A.: Statistical Shape Analysis via Principal Factor Analysis. In: IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1216–1219 (2007)
7. Üzümcü, M., Frangi, A.F., Sonka, M., Reiber, J.H.C., Lelieveldt, B.: ICA vs. PCA active appearance models: Application to cardiac MR segmentation. In: Ellis, R.E., Peters, T.M. (eds.) MICCAI 2003. LNCS, vol. 2878, pp. 451–458. Springer, Heidelberg (2003)
8. Üzümcü, M., Frangi, A.F., Reiber, J.H.C., Lelieveldt, B.P.F.: Independent Component Analysis in Statistical Shape Models. In: Medical Imaging 2003: Image Processing, Proceedings of SPIE, vol. 5032, pp. 375–383 (2003)
9. Albrecht, T., Knothe, R., Vetter, T.: Modeling the Remaining Flexibility of Partially Fixed Statistical Shape Models. In: Workshop on the Mathematical Foundations of Computational Anatomy, MFCA 2008, New York, USA, September 6 (2008)
10. Sierra, R., Zsemlye, G., Székely, G., Bajka, M.: Generation of variable anatomical models for surgical training simulators. Medical Image Analysis 10, 275–285 (2006)
11. Styner, M.A., Rajamani, K.T., Nolte, L.-P., Zsemlye, G., Székely, G., Taylor, C.J., Davies, R.H.: Evaluation of 3D correspondence methods for model building. In: Taylor, C.J., Noble, J.A. (eds.) IPMI 2003. LNCS, vol. 2732, pp. 63–75. Springer, Heidelberg (2003)
12. Timm, N.H.: Applied Multivariate Statistics. Springer, Heidelberg (2002)
13. Abdi, H.: Factor rotations. In: Lewis-Beck, M., Bryman, A., Futing, T. (eds.) Encyclopedia for research methods for the social sciences, pp. 978–982. Sage, Thousand Oaks (2003)
14. Wand, M.P., Jones, M.C.: Kernel Smoothing. Monographs on Statistics and Applied Probability, vol. 60. Chapman & Hall, Boca Raton (1995)
15. Scott, D.W.: Multivariate Density Estimation: Theory, Practice, and Visualization. J. Wiley & Sons, Chichester (1992)
16. Zhang, X., King, M.L., Hyndman, R.J.: Bandwidth Selection for Multivariate Kernel Density using MCMC. In: Australasian Meetings, p. 120. Econometric Society (2004)
17. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Diffeomorphic Demons: Efficient Non-parametric Image Registration. NeuroImage 45(Suppl. 1), 61–72 (2009)
18. Samaha, A.A., Ivanov, A.V., Haddad, J.J., Kolesnik, A.I., Baydoun, S., Yashina, I.N., Samaha, R.A., Ivanov, D.A.: Biomechanical and system analysis of the human femoral bone: correlation and anatomical approach. Journal of Orthopaedic Surgery and Research 2, 8 (2007)