D. Haase, J. A. Nyakatura, J. Denzler. **Multi-view Active Appearance Models for the X-ray Based Analysis of Avian Bipedal Locomotion**. *Proceedings of the 33rd Annual Symposium of the German Association for Pattern Recognition (DAGM 2011)*. 2011. 11–20.

DOI: 10.1007/978-3-642-23123-0_2.
The original publication is available at www.springerlink.com.

Multi-view Active Appearance Models for the X-ray Based Analysis of Avian Bipedal Locomotion

Daniel Haase[1], John A. Nyakatura[2] and Joachim Denzler[1]

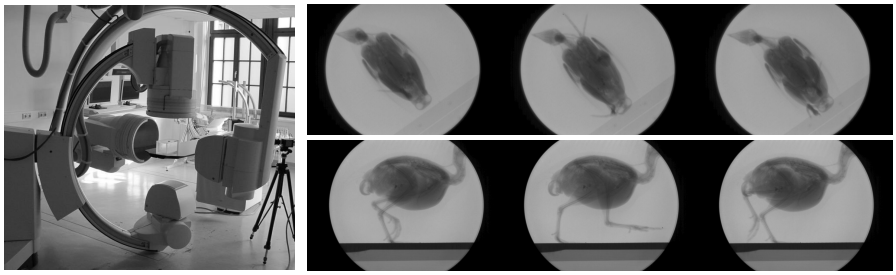
[1]Chair for Computer Vision
[2]Institute of Systematic Zoology and Evolutionary Biology
Friedrich Schiller University of Jena, 07743 Jena, Germany
{daniel.haase,john.nyakatura,joachim.denzler}@uni-jena.de

Abstract. Many fields of research in biology, motion science and robotics depend on the understanding of animal locomotion. Therefore, numerous experiments are performed using high-speed biplanar x-ray acquisition systems which record sequences of walking animals. Until now, the evaluation of these sequences is a very time-consuming task, as human experts have to manually annotate anatomical landmarks in the images. Therefore, an automation of this task at a minimum level of user interaction is worthwhile. However, many difficulties in the data—such as x-ray occlusions or anatomical ambiguities—drastically complicate this problem and require the use of global models. Active Appearance Models (AAMs) are known to be capable of dealing with occlusions, but have problems with ambiguities. We therefore analyze the application of multi-view AAMs in the scenario stated above and show that they can effectively handle uncertainties which can not be dealt with using single-view models. Furthermore, preliminary studies on the tracking performance of human experts indicate that the errors of multi-view AAMs are in the same order of magnitude as in the case of manual tracking.

1 Introduction and Related Work

Understanding animal locomotion is a crucial part of countless problems ranging from the field of biology over motion science to robotics. To name but a few, these problems include gaining a better understanding of evolution [11], the development of mathematical models of locomotion such as the *spring-mass model* [3], or building walking robots. To answer open questions in the field of locomotion research, avian bipedal locomotion provides an appropriate testbed. One reason for the suitability is that bird species exist in countless variations of important locomotion parameters like body mass and limb proportions and exhibit a large range of walking and running speeds.

To gain a profound and detailed insight into terrestrial bird locomotion, many different specimen of various species need to be studied. Nowadays, these studies are often entirely based on high-speed x-ray videography. As opposed to external marker based methods, the key advantage is that all important parts



(a) Acquisition System **(b)** Example Sequence

Fig. 1. (a) Biplanar high-speed x-ray acquisition system (Neurostar®, Siemens AG). (b) Example sequence of a quail (*Coturnix coturnix*) for the dorsoventral (*top row*) and lateral (*bottom row*) camera view acquired with this system.

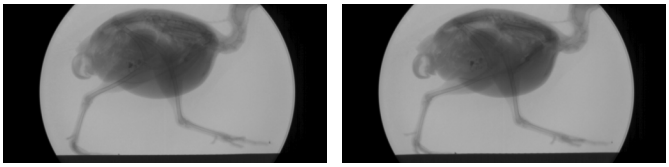
of the locomotor system can be observed directly [6, 11]. A state-of-the-art x-ray acquisition system is shown in Fig 1a. The system consists of two movable x-ray image intensifiers (C-arms) which are positioned around a table and allow for recordings at a high temporal and spatial resolution (1536×1024 pixels at 1 kHz). For the recording of animal locomotion sequences, a non-metallic treadmill is placed on the central table. In Fig. 1b, the locomotion of a quail (*Coturnix coturnix*) acquired using this system is exemplarily shown.

The evaluation of the locomotion sequences is mainly based on anatomical points of interest (*landmarks*), as for instance the *femur* (thighbone), the hip joints or the knee joints. Example landmarks used for a quail are shown in Fig. 3. Until now, the landmarks have to be located manually by human experts. Due to the high temporal resolution, however, this is a highly time-consuming task which has prevented the realization of large-scale studies up to now.

Therefore, there is urgent need to automate the task of anatomical landmark tracking for this application at a minimum of user interaction. At first sight, this might seem to be an easy task, as key point tracking is a well-researched topic in computer vision. Yet, there are several issues which tremendously complicate the procedure. The main problems are the severe and continuously changing occlusions in the x-ray images in consequence of the motion of the animal and the imaging process. This effect causes local image areas around anatomical landmarks to be extremely variable. Thus, local tracking techniques like optical-flow tracking [14], KLT-tracking [1], region-based tracking [13] or SIFT-tracking [16] are rendered impossible [12].

Model-based global approaches, on the other hand, explain each image as a whole and hence are less prone to local disturbances. A prominent example in this context is the registration of a given 3D computer tomography (CT) data set to a 2D image [18, 2, 5]. In our scenario, however, this is a very difficult task, as for each specimen a full-body CT scan plus a skeletal model would be necessary.

Active Appearance Models (AAMs) [7, 10, 8] offer another way of global modeling. They are entirely based on given training images having annotated land-



(a) $t = 0.641\,\mathrm{s}$ **(b)** $t = 0.841\,\mathrm{s}$

Fig. 2. Example for possible anatomical ambiguities. (a) and (b) depict the images 641 ($t = 0.641\,\mathrm{s}$) and 841 ($t = 0.841\,\mathrm{s}$) of a quail sequence, respectively. Both images seem to show the identical pose of the walking bird. However, in the first image, the quail's right leg is ahead of the left leg and in the second image it is vice versa.

marks, and a global model of shape and texture is learnt automatically. The general suitability of AAMs for the present tracking task is shown in [12], where a proof-of-concept is given and the impact of preprocessing methods and the choice of training images are analyzed. Further difficulties of the tracking task at hand which are not already covered in [12] are anatomical ambiguities, especially for parts of the locomotor system. An example for this case is shown in Fig. 2, where two approximately identical images are shown, which however represent opposing states of a walking period. To resolve these ambiguities, either temporal modeling or further context knowledge is necessary. Because one goal is to keep the amount of user-interaction and hence the number of training images small, a temporal model as described in [4] is not applicable. Instead, in the following we analyze the suitability of using both camera views at a time to resolve these uncertainties. For this task, we employ multi-view AAMs [15, 17].

The remainder of this paper is organized as follows. In Sect. 2 we first give a brief overview of basic AAMs and then describe the application of these models for the current tracking task. Thereafter, we describe the adjustments presented in [15, 17] to achieve a multi-view model. We present our experiments and results in Sect. 3. At the end we conclude our findings and discuss future work.

2 Active Appearance Models

Active Appearance Models (AAMs) [7, 10, 8] are well-known statistical models which are used to represent the appearance of objects in digital images. In the following, basic AAMs, their application on locomotion data and the extension on multiple camera views are described.

2.1 Training Step

In the training step of AAMs, the goal is to learn valid appearances of an object based on exemplary images. As the appearance is influenced by both shape and texture, it is necessary to model these two in a combined framework. Thus, the training step consists of building a *shape model*, a *texture model* and a *combined*

model. The training data consist of N training images $\boldsymbol{I}_1, \ldots, \boldsymbol{I}_N$ and M two-dimensional landmarks $\boldsymbol{l}_n = (x_{n,1}, y_{n,1}, \ldots, x_{n,M}, y_{n,M})^{\mathrm{T}}$ for each image $\boldsymbol{I}_n$.

Modeling Shape. The goal in this step is to determine the joint movements of the given landmarks in a statistical manner by using Principle Component Analysis (PCA). As first step, all shapes are aligned with respect to scale, rotation and translation. Then, the landmarks are combined into the matrix $\boldsymbol{L} = (\boldsymbol{l}_1 - \boldsymbol{l}_\mu, \ldots, \boldsymbol{l}_N - \boldsymbol{l}_\mu)$, where $\boldsymbol{l}_\mu = {}^1\!/_N \sum_{n=1}^{N} \boldsymbol{l}_n$ is the *mean shape*. The PCA is applied on $\boldsymbol{L}$, which gives the matrix $\boldsymbol{P}_{\boldsymbol{L}}$ of *shape eigenvectors*. By this means, an arbitrary shape $\boldsymbol{l}'$ can then be described by its *shape parameters* $\boldsymbol{b}_{\boldsymbol{l}'}$ via

$$\boldsymbol{l}' = \boldsymbol{l}_\mu + \boldsymbol{P}_{\boldsymbol{L}} \boldsymbol{b}_{\boldsymbol{l}'}, \quad \text{where} \quad \boldsymbol{b}_{\boldsymbol{l}'} = \boldsymbol{P}_{\boldsymbol{L}}^{\mathrm{T}} \left(\boldsymbol{l}' - \boldsymbol{l}_\mu \right). \tag{1}$$

Modeling Texture. The combined variations of the gray values are analyzed in a similar manner as in the previous step. The object textures of the images $\boldsymbol{I}_n$ are shape-normalized to fit a common reference shape, forming the texture vectors $\boldsymbol{g}_n$. Afterwards, a PCA is applied on the matrix $\boldsymbol{G} = (\boldsymbol{g}_1 - \boldsymbol{g}_\mu, \ldots, \boldsymbol{g}_N - \boldsymbol{g}_\mu)$, where $\boldsymbol{g}_\mu = {}^1\!/_N \sum_{n=1}^{N} \boldsymbol{g}_n$ is called the *mean texture*. The result are the *texture eigenvectors* $\boldsymbol{P}_{\boldsymbol{G}}$, which can be used to represent an arbitrary texture $\boldsymbol{g}'$ by its *texture parameters* $\boldsymbol{b}_{\boldsymbol{g}'}$ by means of

$$\boldsymbol{g}' = \boldsymbol{g}_\mu + \boldsymbol{P}_{\boldsymbol{G}} \boldsymbol{b}_{\boldsymbol{g}'}, \quad \text{where} \quad \boldsymbol{b}_{\boldsymbol{g}'} = \boldsymbol{P}_{\boldsymbol{G}}^{\mathrm{T}} \left(\boldsymbol{g}' - \boldsymbol{g}_\mu \right). \tag{2}$$

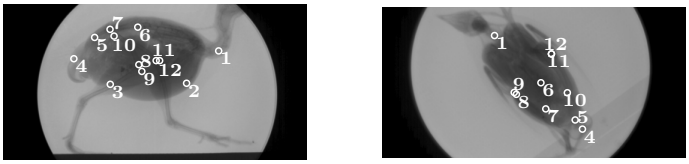
Modeling Appearance. In the third sub-step, the shape parameters $\boldsymbol{b}_{\boldsymbol{l}_n}$ and the texture parameters $\boldsymbol{b}_{\boldsymbol{g}_n}$ are concatenated into a new vector $\boldsymbol{c}_n = (w\boldsymbol{b}_{\boldsymbol{l}_n}^{\mathrm{T}}, \boldsymbol{b}_{\boldsymbol{g}_n}^{\mathrm{T}})^{\mathrm{T}}$ for each training image $\boldsymbol{I}_n$. Here, $w \in \mathbb{R}$ is a scaling factor to account for the different units of shape and intensity. Then, a final PCA is applied on $\boldsymbol{C} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_N)$, which yields the matrix $\boldsymbol{P}_{\boldsymbol{C}}$ of *appearance eigenvectors*. Each object instance with shape parameters $\boldsymbol{b}_{\boldsymbol{l}'}$ and texture parameters $\boldsymbol{b}_{\boldsymbol{g}'}$ can then be described by its *appearance parameters* $\boldsymbol{b}_{\boldsymbol{c}'}$ via

$$\boldsymbol{c}' = (w\boldsymbol{b}_{\boldsymbol{l}'}^{\mathrm{T}}, \boldsymbol{b}_{\boldsymbol{g}'}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{P}_{\boldsymbol{C}} \boldsymbol{b}_{\boldsymbol{c}'}, \quad \text{where} \quad \boldsymbol{b}_{\boldsymbol{c}'} = \boldsymbol{P}_{\boldsymbol{C}}^{\mathrm{T}} \boldsymbol{c}'. \tag{3}$$

By restricting $\boldsymbol{P}_{\boldsymbol{C}}$ on the leading eigenvectors with a certain amount of the total variance, the number of model parameters can be reduced dramatically.

2.2 Model Fitting

To fit a trained model on new data, the necessary parameter updates $\delta\boldsymbol{c}$ are predicted based on the texture difference $\delta\boldsymbol{g}$ between model and image. For this purpose, a linear model $\delta\boldsymbol{c} = \boldsymbol{R}\delta\boldsymbol{g}$ is used. The coefficients $\boldsymbol{R}$ are estimated using multivariate regression by systematically displacing the known model parameters of the training images. For a previously unseen image, the AAM can then be fitted by iteratively adapting the model parameters according to $\boldsymbol{R}$ and $\delta\boldsymbol{g}$.



(a) Lateral View **(b)** Dorsoventral View

(1) 5^{th} *vertebra*
(2) furcula (wishbone)
(3) caudal carina(rear breastbone)
(4) pygostyle (pearson's nose)
(5) pelvis (*caudal*)
(6) pelvis (*cranial*)
(7) left *acetabulum*
(8) left *femur* (thighbone)
(9) left *crista cnemialis*
(10) right *acetabulum*
(11) right *femur* (thighbone)
(12) right *crista cnemialis*

(c) Associations between Landmarks and Anatomical Structures

Fig. 3. Overview of the anatomical landmarks used in the two camera views of the employed quail data set.

2.3 Application on Locomotion Data

For the application of AAMs to the task of landmark tracking in locomotion data, several issues have to be considered. A fundamental question is concerned with the selection of training images and the resulting scope of the model. Possible options range from generic inter-species bird models over specimen-specific models to models for each individual locomotion sequence. Due to differences in anatomy, annotated landmarks and the experimental setup between multiple recordings, we concentrate on sequence-based AAMs.

In this case, annotated images taken from the sequence to be analyzed are used for training. Note that this is an important difference compared to standard AAMs which are usually trained on a set of independent object instances (*e.g.* a face database in the context of face modeling). As a consequence, the resulting AAM becomes a basic locomotion model which expresses the dynamic variation of the landmarks over time. An example for this effect on the quail data set (see Fig. 3) is shown in Fig. 4. It depicts the influence of the first and second shape parameters for a lateral and a dorsoventral AAM. It can clearly be seen that for both models the first shape parameter governs the movement of the thigh bones. The second parameter mainly represents the typical cervical movement of the quail which occurs during locomotion.

Another area of concern for this application is the huge shape non-stationarity (*cf.* [9]) which is induced by the movement of the landmarks during locomotion. As at least a certain degree of shape stationarity is assumed for AAMs, currently only the torso, the knee joint and the hip joint landmarks (see Fig. 3) are considered for automated tracking. In general, simply including the toe landmarks, for instance, will lead to a drastically decreased tracking performance.

More details on the topics described above can be found in [12].

2.4 Multi-View Model

The extension of AAMs on multiple camera views is presented by [15, 17] in the context of medical image analysis. If we denote the number of camera views to be

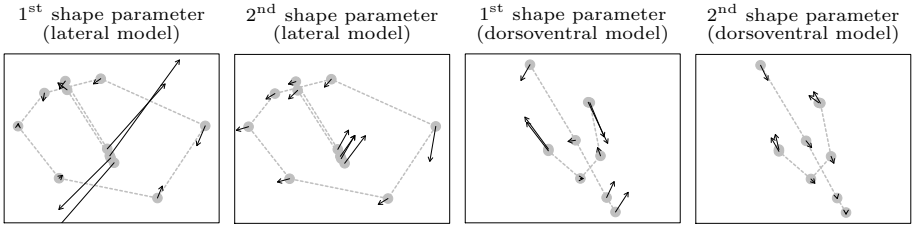


Fig. 4. Influence of the shape parameters of the lateral and dorsoventral shape models. The arrows indicate the movement of the landmarks for positive parameter values. For negative values, the orientation of the arrows is the other way around. The shown landmarks are described in Fig. 3.

modeled by K, then the n^{th} training example consists of the images $\boldsymbol{I}_n^{(1)}, \ldots, \boldsymbol{I}_n^{(K)}$ and the landmarks $\boldsymbol{l}_n^{(1)}, \ldots, \boldsymbol{l}_n^{(K)}$ in these images. As first step, the landmarks $\boldsymbol{l}_n^{(k)}$ of all training examples are aligned camera-wise just like in the single-view case. Then, all training images have to be shape-normalized, however still independently for each camera view, yielding $\boldsymbol{g}_n^{(k)}$. The main idea is then to simply concatenate the landmark vectors and the texture vectors of the camera views for each training example $1 \leq n \leq N$ in the sense of

$$\boldsymbol{l}_n = \left({\boldsymbol{l}_n^{(1)}}^{\mathrm{T}}, \ldots, {\boldsymbol{l}_n^{(K)}}^{\mathrm{T}} \right)^{\mathrm{T}} \quad \text{and} \quad \boldsymbol{g}_n = \left({\boldsymbol{g}_n^{(1)}}^{\mathrm{T}}, \ldots, {\boldsymbol{g}_n^{(K)}}^{\mathrm{T}} \right)^{\mathrm{T}}. \tag{4}$$

In this way, each training example actually consisting of multiple shapes and textures can effectively be reduced to just one landmark vector $\boldsymbol{l}_n$ and one texture vector $\boldsymbol{g}_n$. The subsequent steps exactly follow those from the case of the single-view model, and the multiple views are modeled implicitly.

3 Experiments and Results

In the following we experimentally analyze the benefits of multi-view AAMs in comparison to single-view AAMs for the present task of anatomical landmark tracking. As the main goal is to achieve sound tracking results at a minimum of user interaction, the essential questions to be answered are:

- Can multi-view AAMs substantially resolve anatomical ambiguities which can not be overcome using single-view models?
- Is a reduction of the amount of training data possible with multi-view AAMs?
- How do these models perform compared to manually tracked landmarks?

To answer these questions, all experiments were conducted on a real data set for which comprehensive ground-truth landmark positions are available. This data set shows the locomotion of a quail from two camera views (*lateral* and *dorsoventral*, see Figs. 1 and 3) and has a length of 2245 images (2.245 s recorded at 1 kHz). As a rescaling of the original images (1536×1024) to a size of 25%

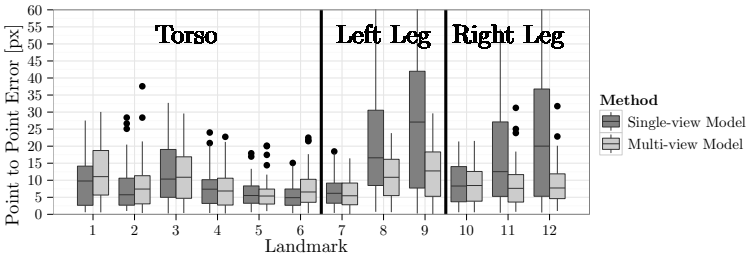


Fig. 5. Comparison of the tracking results between a single-view and a multi-view AAM for the landmarks of the lateral camera view. Using a multi-view model, anatomical ambiguities can be substantially resolved. As a result, the tracking quality of the knee landmarks 8, 9, 11 and 12 is drastically improved.

does not lead to a substantial loss of tracking quality [12], all experiments were conducted on the smaller versions for performance reasons. The evaluations, however, were performed with respect to the original image size in any case. Ground-truth landmark positions obtained from experts are available for 81 frames evenly spread over the entire sequence and allow a systematic evaluation. The quantitative evaluation of the results is based on the Euclidian distance between tracked and ground-truth landmark, which is known as *point to point error* [19].

3.1 Resolving Anatomical Ambiguities

In the course of this paper we presented an example for anatomical ambiguities which can arise in locomotion sequences (see Fig. 2). Furthermore, we stated that these uncertainties can not be resolved using single-view AAMs and that the application of multi-view AAMs is inevitable. To support this hypothesis, we trained two single-view AAMs on the lateral and dorsoventral view of the data set and compared the results with an—in other respects identical—multi-view AAM. For training, 15 images from one walking period at the end of the sequence were selected. Due to the ambiguities described above, the single-view model of the lateral view has severe problems of locating the knee landmarks correctly and even occasionally mixes the landmarks for the left and right knee up. The comparison of single-view and multi-view model is given in Fig. 5. It can clearly be seen that the multi-view model drastically improves the tracking results of the knee landmarks (8, 9, 11, 12, see Fig. 3). Small errors, as, for instance, indicated by the 25% quartiles, are reduced observably, however, the major enhancements are present in the larger error regions. The median error for the right knee landmark (12) is, for instance, reduced from 20 px in the single-view case to 8 px in the multi-view case.

For the torso landmarks, however, no substantial improvement is observed. This result can be explained by the fact that the torso landmarks usually have a low ambiguity due to low interference with parts of the locomotor system.

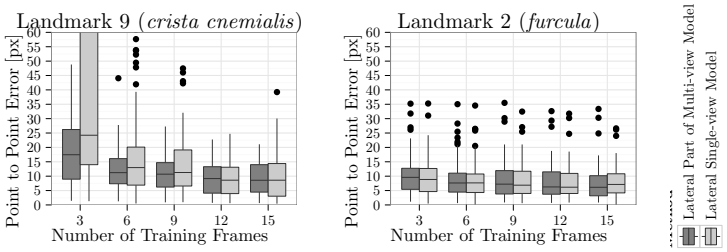


Fig. 6. Influence of the reduction of training data on single-view and multi-view AAMs. The results are shown for (a) an exemplary knee landmark (*crista cnemialis*) and (b) an exemplary torso landmark (*furcula*).

3.2 Reduction of the Amount of Training Data

One very important goal is to keep the human interaction spent for landmark labeling at a minimum to allow for a large amount of data to be processed. Therefore, the amount of training images is an important factor. However, less training images usually cause greater uncertainties and hence greater errors during tracking. As discussed in the last subsection, the multi-view model is capable of reducing uncertainties. For this reason, an interesting question is whether multi-view AAMs can be used to decrease the necessary amount of training data.

To answer this question, we compared the tracking results of single-view and multi-view AAMs with identical parameters for varying numbers of training frames. The frames were chosen from the third walking period of the quail in the middle of the sequence. As in the last subsection, the results vary considerably between torso landmarks and landmarks of the locomotor system. In Figs. 6a and 6b, example results for both cases are shown. The former depicts the case for a knee landmark (landmark 9, left *crista cnemialis*, see Fig. 3). Here, it can be seen that the errors of both the single- and the multi-view model increase as the amount of training frames is reduced. However, the errors for the single-view model rise much more rapidly. In the case of the torso landmark, the results remain approximately constant for both the single- and the multi-view models. Again, this can be explained by the low ambiguity of this kind of landmarks.

Above results indicate that multi-view models can be used to decrease the necessary amount of training frames. While the uncertainty of the torso landmarks can not be decreased substantially as they are not subject to anatomical ambiguities, the uncertainty of the locomotion landmarks can be reduced.

3.3 Comparison to Manual Landmark Tracking

Tracking Time. For the multi-view model presented in Subsec. 3.1, a total time of 38.20 min (15.21 min training, 22.99 min tracking) was required. As the sequence has a total number of 2245 images per camera view, this corresponds

to a time of 0.51 s per image. Human experts, on the contrary, usually need at least about 45 s per image, which results in speed-up factors greater than 90.

Accuracy and Precision. To allow for a meaningful comparison between automated tracking results and manual tracking, currently a large-scale study on the accuracy and precision of human experts is in progress. Here, four experts are to label one and the same locomotion sequence, three times each, and independently of one another. Unfortunately, not all results are available to date.

Yet, first comparisons between multiple labelings of two experts for the given data set indicate that the typical human errors are in the range of about 0.5 px (min.), 5.5 px (1st quartile), 9 px (median), 14 px (3rd quartile) and 40 px (max.). Taking these preliminary results into account, we can state that the errors of the multi-view AAM shown in Fig. 5 are in the same order of magnitude as the manual errors.

4 Conclusions and Further Work

In this work we analyzed the benefits of multi-view Active Appearance Models for the application of anatomical landmark tracking in biplanar x-ray locomotion sequences. We showed that multi-view models perform substantially better than comparable single-view models in situations of high uncertainty, *e.g.* for frames with anatomical ambiguities. Furthermore, we compared single-view and multi-view models for varying amounts of training data and demonstrated that the latter can be used to reduce the necessary amount of user labeled training images. Finally we stated that, based on preliminary studies, the performance of multi-view AAMs is in the same order of magnitude as in the case of manual tracking.

An interesting point for future work is to expand the presented approach on landmark configurations with a substantially larger non-stationarity, as for example shapes including toe landmarks. Also, local refinement methods could be analyzed in order to obtain an even more accurate adaptation to the anatomical structures. The preliminary studies on the precision of manual tracking should be continued to enable a more profound comparison to automated methods.

Acknowledgements

The authors would like to thank Alexander Stößel from the Institute of Systematic Zoology at the Friedrich Schiller University of Jena for providing the labeled quail dataset.

This research was supported by grant DE 735/8-1 of the German Research Foundation (DFG).

References

1. Baker, S., Matthews, I.: Lucas-kanade 20 years on: A unifying framework. Int. J. Comput. Vision 56(3), 221–255 (2004)

2. Bey, M.J., Zauel, R., Brock, S.K., Tashman, S.: Validation of a new model-based tracking technique for measuring three-dimensional, in vivo glenohumeral joint kinematics. J. Biomech. Eng. 128, 604–609 (2006)
3. Blickhan, R.: The spring-mass model for running and hopping. J. Biomech. 22(11-12), 1217–1227 (1989)
4. Bosch, J.G., Mitchell, S.C., Lelieveldt, B.P.F., Nijland, F., Kamp, O., Sonka, M., Reiber, J.H.C.: Automatic segmentation of echocardiographic sequences by active appearance motion models. IEEE T. Med. Imaging 21(11), 1374–1383 (2002)
5. Brainerd, E.L., Gatesy, S.M., Baier, D.B., Hedrick, T.L.: A method for accurate 3D reconstruction of skeletal morphology and movement with CTX imaging. Comp. Biochem. Physiol. 146, 119 (2007)
6. Brainerd, E.L., Baier, D.B., Gatesy, S.M., Hedrick, T.L., Metzger, K.A., Gilbert, S.L., Crisco, J.J.: X-ray reconstruction of moving morphology (XROMM): Precision, accuracy and applications in comparative biomechanics research. J. Exp. Zool. A 313A(5), 262–279 (2010)
7. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Burkhardt, H., Neumann, B. (eds.) Proceedings of the 5th European Conference on Computer Vision. LNCS, vol. 1407, pp. 484–498. Springer (1998)
8. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE T. Pattern Anal. 23(6), 681–685 (2001)
9. Das, S., Vaswani, N.: Nonstationary shape activities: Dynamic models for landmark shape change and applications. IEEE T. Pattern Anal. 32(4), 579–592 (2010)
10. Edwards, G.J., Cootes, T.F., Taylor, C.J.: Face recognition using active appearance models. In: Burkhardt, H., Neumann, B. (eds.) Proceedings of the 5th European Conference on Computer Vision. LNCS, vol. 1407, pp. 581–595. Springer (1998)
11. Gatesy, S.M.: Guineafowl hind limb function. I: Cineradiographic analysis and speed effects. J. Morphol. 240(2), 1097–4687 (1999)
12. Haase, D., Denzler, J.: Anatomical landmark tracking for the analysis of animal locomotion in x-ray videos using active appearance models. In: Heyden, A., Kahl, F. (eds.) Proceedings of the 17th Scandinavian Conference on Image Analysis. pp. 604–615. No. 6688 in LNCS, Springer (2011)
13. Hager, G.D., Belhumeur, P.N.: Efficient region tracking with parametric models of geometry and illumination. IEEE T. Pattern Anal. 20(10), 1025–1039 (1998)
14. Horn, B.K.P., Schunck, B.G.: Determining optical flow. Artif. Intell. 17(1-3), 185–203 (1981)
15. Lelieveldt, B., Üzümcü, M., van der Geest, R., Reiber, J., Sonka, M.: Multi-view active appearance models for consistent segmentation of multiple standard views. International Congress Series 1256, 1141–1146 (2003)
16. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60(2), 91–110 (2004)
17. Oost, E., Koning, G., Sonka, M., Oemrawsingh, P.V., Reiber, J.H.C., Lelieveldt, B.P.F.: Automated contour detection in x-ray left ventricular angiograms using multiview active appearance models and dynamic programming. IEEE T. Med. Imaging 25(9), 1158–1171 (2006)
18. Rohlfing, T., Denzler, J., Gräßl, C., Russakoff, D.B., Jr., C.R.M.: Markerless real-time 3-d target region tracking by motion backprojection from projection images. IEEE T. Med. Imaging 24(11), 1455–1468 (2005)
19. Stegmann, M.B.: Active Appearance Models: Theory, Extensions and Cases. Master's thesis, Technical University of Denmark, DTU (2000)