FACE RECOGNITION FOR GREAT APES: IDENTIFICATION OF PRIMATES IN VIDEOS

Alexander Loos and *Talat Anand Mohan Kalyanasundaram*

Fraunhofer Institute for Digital Media Technology IDMT
Audio-Visual Systems
Ehrenbergstraße 31, 98693 Ilmenau, Germany

ABSTRACT

Due to the ongoing biodiversity crisis, many species including great apes such as chimpanzees or gorillas are threatened and need to be protected. To overcome the catastrophic decline of biodiversity, biologists recently started to use remote cameras for wildlife monitoring. However, the manual analysis of the resulting image and video material is extremely tedious, time consuming, and highly cost intensive.

To overcome the burden of time-consuming routine work we studied and proposed novel approaches for automatic chimpanzee identification in our previous work. Starting from the assumption that humans and our closest relatives share similar facial properties, algorithms for human face recognition were adapted and extended for this purpose. However, the proposed algorithms were designed to recognize chimpanzee individuals in still images only. In this paper we extend these ideas towards chimpanzee identification in video sequences. Thus, a novel frame weighting approach is presented which significantly improves the system's accuracy.

***Index Terms*—** Animal Biometrics, Face Recognition

1. INTRODUCTION AND MOTIVATION

Due to the ongoing biodiversity crisis, many species are on the brink of extinction. Primates like chimpanzees or gorillas are hit by the crisis and belong to a species that is severely endangered. Walsh *et al.* [1] for instance reported a decrease of ape populations in western equatorial Africa by more than a half between 1983 and 2000. Similar conclusions were drawn by Campbell *et al.* in [2]. They observed a 90% decrease of chimpanzee sleeping nests in Côte d'Ivoire between 1990 and 2007. Those agitating results demonstrate the urgent need to intensify close surveillance of this threatened species in order to protect the remaining populations. However, effectively protecting animals requires good knowledge of existing populations and fluctuations of population sizes over time. Non-invasive monitoring techniques using autonomous recording devices is therefore tremendously increasing [3]. However, the collected data often needs to be evaluated manually which is a time and resource consuming task. Consequently, there is a high demand for automated algorithms to analyze remotely gathered video recordings. Starting from the assumption that humans and great apes share similar properties of the face we explored face detection and recognition techniques, originally developed for human identification, to recognize chimpanzees in wildlife footage in our previous work [4, 5, 6, 7]. However, the proposed framework was only capable to automatically detect and identify chimpanzee faces in still images. In this paper we extend these ideas to face detection, tracking, and identification in video sequences. A novel frame-weighting approach is proposed which implicitly exhibits temporal information of video recordings. The proposed approach is thoroughly tested and evaluated on self-established realistic real-world video datasets of free-living and captive chimpanzee individuals. It is shown that the proposed approach performs better than recognition on a single frame basis as well as a uniform weighting scheme.

2. PREVIOUS WORK

Based on the assumption that humans and our closest relatives share similar properties of the face, we suggested to adapt and extend face recognition techniques for the identification of great apes. In [4] we showed that state-of-the-art face recognition techniques are capable to also identify chimpanzees and gorillas. Based on these results we significantly improved the performance of the proposed system by using Gabor features in combination with Locality Preserving Projections (LPP) for dimensionality reduction in [5]. A Sparse Representation Classification (SRC) scheme was used to assign identities to the facial images. Although the obtained results were very promising, the accuracy of the system decreased significantly if non-frontal face images were used for testing. We later combined face and facial feature detection as well as face recognition and presented a completely automated identification system for chimpanzees in [6]. In order to increase the robustness of the proposed framework against difficult lighting situations, pose, partial occlusion, and the vast number of occurring expressions we additionally combined the results of global and local features using a decision fusion scheme [7].

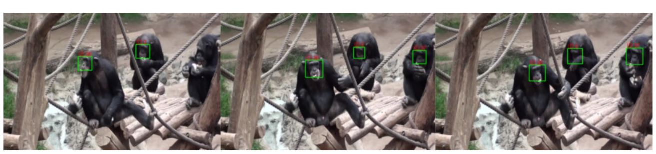
In summary, our proposed framework consists of three main parts: *Face and Facial Feature Detection* using a state-of-the-art rigid object detector proposed in [8], *Face Alignment* using an affine transform based on detected coordinates of both eye and the mouth-center, and *Face Recognition* based on global and local feature extraction, feature space transformation, classification, and decision fusion [6, 7]. It is known from the literature that different features tend to misclassify different patterns. Thus, first global Gabor-based features are extracted from the aligned face. After feature space transformation using LPP, a SRC-based classification scheme, originally proposed by Wright *et al.* in [9], is utilized for identification using global features. Additionally, SURF descriptors are extracted on six facial fiducial points around both eyes and the nose in order to exhibit individually unique permanent wrinkle patterns of the chimpanzee's face. A final local feature vector is subsequently constructed by concatenating the resulting local descriptors. An SVM with RBF-kernel is trained and utilized to predict the individual's identity using local features. The results of the proposed local and global face recognition pipeline are subsequently fused by taking the confidences of both classification schemes into account. For details of the developed Primate Recognition Framework (PRF) the interested reader is referred to [6, 7].

3. IDENTIFICATION OF PRIMATES IN VIDEOS

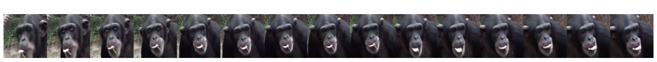
One key-step for face recognition in video is to simultaneously track multiple detected faces. Once each tracked target is assigned to a face-track, faces are first aligned as done for still images. Subsequently, modules for quality assessment are applied which analyze parameters such as pose and visual quality in order to select the frames best suited for recognition. Each selected and aligned face is subsequently identified by means of the algorithms we presented in our previous work [6, 7] and were outlined in the previous section. To further enhance the system's accuracy, a frame-weighting approach is proposed which implicitly exhibits the temporal information in video recordings.

Tracking by Continuous Detection Once faces were detected they have to be tracked through the video sequence. Therefore, unique object-IDs are assigned to each detected face which are maintained for the subsequent frames. This procedure results in a so called face-track, a collection of faces from one single individual in various appearances. Within the proposed framework, a face and facial feature detection and tracking library named SHORE, developed by Küblbeck and Ernst in [10] and extended in [8] to detect and track faces of great apes, is utilized. They suggested a tracking by continuous detection approach in [10] to overcome the deficiencies of pure tracking algorithms. Each frame is processed with a fast and accurate real-time face detector. A motion model is then applied to connect the detections of subsequent frames. In order to estimate the current state of a tracked face from the detection results, a linear Kalman filter is applied. For details about the face detection and tracking procedure the interested reader is referred to [10, 8].

Figure 1 shows an excerpt of a video sequence gathered in the zoo of Leipzig, Germany, and an extracted face-track for one of the individuals present in the video. As can be seen, not all frames are equally well suited for automatic identification since the proposed recognition framework is optimized for full-frontal faces. Hence, the frames which are best suited for recognition are automatically selected in a second step which involves assessment of various parameters regarding visual quality of a face.



(a)



(b)

Fig. 1. (a) Three selected frames of a video gathered in the zoo of Leipzig, Germany. The detection and tracking results by SHORE are superimposed. (b) An extracted face-track of one of the three individuals in the video sequence.

Best Frame Selection This section describes algorithms for full-automatic selection of the frames of a face-track which are best suited for identification. The most important cues to achieve high recognition accuracies in practical applications are the pose of a face and its visual quality. Both parameters are analyzed automatically.

Automatic head pose estimation has been an active research topic for decades. For the application presented in this thesis an exact estimation of pitch, roll, and yaw angles in 3D space is not required. Since the proposed face recognition pipeline is designed for frontal face images, the objective of the proposed pose estimation module is to automatically select those frames within a face-track that contain faces in a near-frontal pose. The following head pose estimation technique is proposed which performs sufficiently well for the task at hand. First, detected and tracked faces are aligned by applying an affine transformation based on the detected eye and mouth coordinates and subsequently converted to gray-scale. As holistic face representation we extract Gabor-based features as described in our previous work [6] which are later used as global features for the recognition part and can therefore be reused again for identification. A PCA is subsequently applied to project the resulting high-dimensional feature vectors into a smaller dimensional subspace of size 100. For classification, an SVM with RBF kernel was trained on a set of frontal and non-frontal primate faces. The two user-defined parameters (C, γ) were optimized during training using a grid-search in combination with a 5-fold cross-validation on the training set. The resulting model was finally saved and is applied in the test case. The confidence measure of the SVM prediction is utilized as a weighting factor for every face within a face-track. Hence, the *Pose-Quality* is a number between 0 and 1. The higher the value, the more frontal is the pose of the considered face.

Besides the head pose, other parameters regarding the visual quality of the facial image are crucial for accurate identification. Thus, we further propose a number of lightweight software modules which estimate the visual quality of an image. For instance, the low-pass filtering characteristic of a blurred image might not contain sufficient detailed information important for accurate identification. Our blur estimation approach is influenced by ideas of Liu *et al.* [11]. The detection of blurred images is done by analyzing its *Local Power Spectrum Slope*. First, the power spectrum's slope of the global image α_g is calculated. Then the input image is divided into 3×3 blocks and the slope of the power spectrum is calculated separately for each resulting region. The final metric for each patch p is then given by $\eta_p = \frac{\alpha_p - \alpha_g}{\alpha_g}$,where α_p is the slope of the power spectrum of patch p. The overall blur measure is finally given by averaging the η_p's of all patches. After normalization, a *Blur-Quality* measure in the range $[0, 1]$ can be used.

Besides pose and image blur, overexposure and underexposure are other important photometric factors for visual quality which should be taken into account. Keeping in mind that the proposed system should work in real-time, simple histogram-based statistics are assessed in order to estimate the lighting conditions of a facial image. Due to space constraints in this paper the interested reader is referred to our previous work [12]. The *Lighting-Quality* is again a real number scaled between 0 and 1. All quality measures are finally combined in a multiplicative fashion in order to sort the images of an extracted face-track according to the overall *Face-Quality*.

Frame Weighting As outlined above, video acquisition in natural habitats of great apes often leads to large quality variations between frames. Thus, recognition in a single frame might often not lead to the desired result. To overcome this issue, a novel frame-weighting approach is proposed which combines individual frame-based classifications into a single score per face-track and hence penalize uncertain frames. Based on the face-quality modules outlined above, the F frames with the highest quality are selected, aligned, and classified according to proposed image face recognition approach described in our previous work. The following confidence measures can subsequently be derived for every classification: First, three measures for the recognition pipeline using global features and SRC are proposed:

(1) Minimal Residual: The minimal residual between the test sample $\boldsymbol{t}$ and the matrix of training samples $\boldsymbol{A}$ defined as

$$\arg\min_i r_i(\boldsymbol{t}) \quad \text{with} \quad r_i(\boldsymbol{t}) = \|\boldsymbol{t} - \boldsymbol{A}(\boldsymbol{\delta}_i \odot \hat{\boldsymbol{p}}_1)\|_2 \qquad (1)$$

is chosen as the first confidence measure, where $\boldsymbol{\delta}_i$ is the characteristic function of class i, $\hat{\boldsymbol{p}}_1$ is the sparse coefficient vector obtained by ℓ_1-norm minimization, and $\odot$ represents the Hadamard-Schur product also know as the element-wise product. The smaller the minimal residual, the more confident is the classifier.

(2) Distance between the first and second residual: In case of misclassification, the difference of the minimal and second smallest residual is usually smaller than in the correct case. Hence, the absolute difference of the smallest two residuals Δr is used as second confidence measure. A confident classification should have a high Δr, while for incorrect classifications the difference of the two smallest residuals is rather small.

(3) Sparse Concentration Index (SCI): Besides the minimal residual, Wright *et al.* [9] propose to utilize the sparsity of the vector $\hat{\boldsymbol{p}}_1$ as an additional confidence measure of SRC. Therefore, they introduced a measure called SCI which is defined as

$$\text{SCI}(\hat{\boldsymbol{p}}_1) = \frac{C \cdot \max_i \left(\|\boldsymbol{\delta}_i \odot \hat{\boldsymbol{p}}_1\|_1\right) / \|\hat{\boldsymbol{p}}_1\|_1 - 1}{C - 1} \quad \in [0, 1], \qquad (2)$$

where C is the number of classes. The lager the SCI, the sparser the vector $\hat{\boldsymbol{p}}_1$ which in turn is a measure for the confidence of the classifier.

Secondly, two confidence measures for the recognition pipeline using local features in combination with an SVM are taken into account:

(4) SVM Probability: The probability estimates of LibSVM [13, 14] are used as confidence measures for identification based on local features. While for SRC the minimal residual determines the class affiliation, for SVM the test sample is assigned to the class with the maximum probability.

(5) Difference of the Two Highest Probabilities: As the difference of the two smallest residuals can be used as additional confidence measure for SRC-based classification, the difference between the two largest probabilities is utilized for classification via SVM.

All five measures are scaled between 0 and 1 and are subsequently concatenated into a confidence vector $\boldsymbol{v}$. The goal of the proposed frame-weighting approach is to estimate the probability that the classification of frame f was correct. A weighting factor can then be assigned to each classified frame. The weighted confidences of the selected frames are then aggregated in order to obtain a final prediction. Thus, frame-weighting is divided into a training and test phase:

Training: First a 20-fold Monte-Carlo cross-validation on the training set is applied in order to construct the confidence-vectors for each correct and incorrect classification as explained above. This procedure results in two clusters of correct and incorrect classifications in 5-dimensional space. Figure 2(a) depicts a scatter plot of the resulting clusters. Misclassifications are mostly located in one particular corner while the cluster of correct classifications is more scattered. However, it can be seen that correctly classified samples can be separated from misclassifications quite well, since only a few incorrectly identified samples are located in the area of the majority of correct classifications. In order to estimate the probability of a correct classification, the next step is to calculate the Mahalanobis-distances [15] of each sample to the cluster of correct and to the cluster of false classifications. In previous experiments

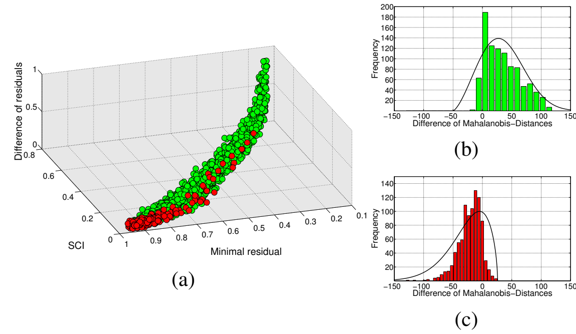


Fig. 2. Figure (a) depicts clusters of the proposed confidence vectors of correct classifications (green samples) and incorrect classifications (red samples). Note that for illustration purposes only the first three dimensions of the 5-D space are plotted. Figures (b) and (c) illustrate the histograms of the difference of Mahalanobis-distances $\Delta d_M(\boldsymbol{t})$ for correct and incorrect classifications, respectively. The fitted Extreme-Value Distributions are superimposed in black.

we noticed that the difference of Mahalanobis-distances $\Delta d_M(\boldsymbol{t})$ is more distinctive than the distance to a good or bad cluster itself. The difference of Mahalanobis-distances is given by

$$\Delta d_M(\boldsymbol{t}) = d_M^{IC}(\boldsymbol{t}, \boldsymbol{O}_{IC}) - d_M^{C}(\boldsymbol{t}, \boldsymbol{O}_C), \qquad (3)$$

where $d_M^{IC}(\boldsymbol{t}, \boldsymbol{O}_{IC})$ is the Mahalanobis-distance of $\boldsymbol{t}$ to the cluster of incorrect classifications $\boldsymbol{O}_{IC}$ and $d_M^{C}(\boldsymbol{t}, \boldsymbol{O}_C)$ represents the Mahalanobis-distance of the test sample to the cluster of correct classifications. The calculation of $\Delta d_M(\boldsymbol{t})$ is repeated in a leave-one-out fashion, i.e. every correctly and incorrectly classified sample takes the role of the test vector $\boldsymbol{t}$ once. Figures 2(b) and 2(c) show the obtained histograms of $\Delta d_M(\boldsymbol{t})$ for correct and incorrect classifications, respectively.

The next step of the training phase is to apply Maximum-Likelihood-Estimation (MLE) in order to fit an Extreme Value Distribution (EVD) [16] to the observed data. A detailed derivation of the subsequent equations can be found in [17]. The Probability-Density-Function (PDF) given an extreme-valued random variable x is defined as

$$P(x) = \lambda \exp\left[-\lambda(x - \mu) - \mathrm{e}^{-\lambda(x-\mu)}\right], \qquad (4)$$

where μ and λ are location and scale parameters, respectively. The likelihood of drawing N samples x_i from an EVD with parameters μ and λ can thus be written as

$$P(x_1 \dots x_N | \lambda, \mu) = \prod_{i=1}^{N} \lambda \exp\left[-\lambda(x_i - \mu) - \mathrm{e}^{-\lambda(x_i-\mu)}\right]. \quad (5)$$

Maximizing Equation 5 with respect to λ and μ yields the maximum likelihood estimation of both parameters. The *Newton-Raphson-Algorithm* [18, 19] can be applied in order to estimate the two parameters $\hat{\lambda}$ and $\hat{\mu}$. The fitted EVDs of the difference of Mahalanobis-distances of correct and incorrect classifications are superimposed in Figures 2(b) and 2(c), respectively.

Test: After parameter estimation using MLE, the fitted EVDs can be utilized in the test phase to calculate the probability that the classification in frame f was correct and weight the resulting prediction accordingly.

The weighting-factor w_f of frame f is given by the Bayes' theorem. Let $P(\Delta d_M|C)$ be the probability of the difference of Mahalanobis-distances given a correct classification, then the weighting-factor $w_f = P(C|\Delta d_M)$ is given by

$$P(C|\Delta d_M) = \frac{P(\Delta d_M|C)P(C)}{P(\Delta d_M|C)P(C) + P(\Delta d_M|IC)(1 - P(C))},$$

where $P(C)$ is the probability of a correct classification and $P(\Delta d_M|IC)$ is the probability of the difference of Mahalanobis-distances given an incorrect classification which can be estimated from the PDF of the EVD of incorrect classifications fitted during training. The correct classification probability $P(C)$ is taken from the mean accuracy of the proposed system after 20-fold crossvalidation applied on the training set as explained above. Moreover, $P(\Delta d_M|C)$ is also calculated from the estimated PDF of correct classifications. Once the weighting factor w_f has been calculated for every frame, the frame-weighting procedure is as follows: Let $\boldsymbol{s}_f$ be the score vector for classifying the face in frame f, then the cumulative score vector $\boldsymbol{s}_c \in \mathbb{R}^{C\times 1}$ is defined as the weighted average of the score-vectors for all selected frames $f = 1 \dots F$

$$\boldsymbol{s}_c = \frac{1}{F}\sum_{f=1}^{F} w_f \cdot \boldsymbol{s}_f. \tag{6}$$

Once all frames have been processed, the index of the maximum element of $\boldsymbol{s}_c$ denotes the final prediction of the current face-track.

4. EXPERIMENTS AND RESULTS

Datasets and Experiment Design: Video footage of captured as well as free-living chimpanzees was recorded at the zoo of Leipzig, Germany and the Taï National Park, Côte d'Ivoire, Africa. The entire image datasets used for evaluation of our previous approaches were used for training in this paper. Note that image and video datasets were gathered independently and thus are completely separated. All video and image datasets can be purchased over our project website *www.saisbeco.com* for benchmark purposes. Hereafter, the video datasets are referred to as *ChimpZoo-Video* and *ChimpTaï-Video*, respectively. Statistics about both video datasets can be found in Table 1.

Dataset	Videos	Face Tracks	Frames per Track	Individuals in Database
ChimpZoo-Video	14	264	$1 - 818$	24
ChimpTaï-Video	11	198	$1 - 1149$	49

Table 1. Statistics about the video datasets used for experimentation.

Based on the observation that false-positive detections usually cannot be tracked and thus the resulting face-tracks are extremely short, the minimum length of a face-track is set to 10 frames. All tracks below that threshold are automatically classified as *unknown* and are not further processed. This procedure correctly eliminates 91.80% and 95.18% of all false-positive detections for the ChimpZoo-Video and the ChimpTaï-Video dataset, respectively.

Results: Four different approaches are compared:
(1) First Frame: The applied face detection library SHORE was trained on full-frontal faces with moderate pose offsets. Hence, it can be assumed that the first frame of a face-track contains a face with full-frontal pose. Based on this assumption, the first approach identifies great apes solely in the first frame of a face-track.
(2) Best Frame: With the help of the facial quality estimation modules proposed, the identity of a face-track is predicted solely based on the frame with the best estimated quality.
(3) Uniform Weighting: The $F = 10$ frames with the best facial quality are selected for identification. The prediction of each frame is weighted equally, i.e. every prediction contributes the same to the final result.
(4) Frame Weighting: Finally, the proposed frame weighting scheme is applied. Again, the $F = 10$ frames with the best facial quality are used for subsequent recognition. The decisions for all analyzed frames are aggregated using the proposed frame weighting approach.

Cum. Acc. [%]	ChimpZoo-Video			
	Frame Weighting	Uniform	Best Frame	First Frame
Rank-1	**70.94**	63.55	56.16	53.60
Rank-2	**80.30**	72.42	68.48	63.45
Rank-3	**84.24**	73.90	71.44	68.38
	ChimpTaï-Video			
Rank-1	**67.40**	61.33	58.01	55.53
Rank-2	**75.69**	68.51	67.95	62.10
Rank-3	**78.45**	72.32	71.82	67.68

Table 2. Obtained results for the ChimpZoo-Video and the ChimpTaï-Video dataset.

Table 2 lists the obtained results for all four approaches. To investigate the robustness of each approach, the results including rank-2 and rank-3 are depicted. For both datasets recognition solely based on the first frame performs worst. Obviously, SHORE is capable of accurately detecting primate faces even under difficult conditions which hamper performance of subsequent recognition. By applying the proposed quality estimation modules and performing recognition on the frame with the best visual quality results in a higher accuracy for both datasets. This is particularly obvious for the rank-2 recognition rate which improves by 5% for both datasets. Thus, first sorting the frames according to their visual quality can increase the accuracy of facial identification quite significantly. The applied uniform weighting scheme performed significantly better than the previous two single-frame-based approaches. Hence, taking recognition results of multiple frames into account seems to improve the system's performance significantly because identification is not dependent on a single frame with potentially bad visual quality. However, the predictions of each frame are weighted equally and the confidences of classification in each frame are not taken into account. Substantial improvements were achieved by the proposed frame-weighting approach with regard to the previous three approaches. The rank-1 accuracy could be improved by more than 6% on both datasets. The rank-3 accuracy for the ChimpZoo-Video dataset is thus 84.24% and 78.45% for the ChimpTaï-Video dataset.

5. CONCLUSION

In this paper we substantially extended our previous work on full-automatic identification of great apes from still images to recognition in video sequences. Therefor, we utilized a state-of-the art face detection and tracking library to create one face-track per detected chimpanzee face. Since not all frames are equally well suited for subsequent recognition we proposed several software modules to asses the visual quality of a face. After selecting the best frames, we applied a novel frame-weighting paradigm which was shown to outperform approaches based on recognition of a single frame and a uniform weighting scheme.

6. REFERENCES

[1] P. D. Walsh, K. A. A. Abernethy, and M Bermejo, "Catastrophic Ape Decline in Western Equatorial Africa," *Nature*, vol. 422, pp. 611–614, 2003.

[2] G. Campbell, H. Kühl, P.N. Kouame, and C. Boesch, "Alarming Decline of West African Chimpanzees in Côte d'Ivoire," *Current Biology*, vol. 18, no. 19, pp. R904–R905, 2008.

[3] J. M. Rowcliffe and C. Carbone, "Surveys using Camera Traps: Are We Looking to a Brighter Future?," *Animal Conservation*, vol. 11, no. 3, pp. 185–186, 2008.

[4] Alexander Loos, Martin Pfitzer, and Laura Aporius, "Identification of Great Apes Using Face Recognition," in *European Signal Processing Conference (EUSIPCO)*, Barcelona, Spain, 2011.

[5] Alexander Loos, "Identification of Great Apes Using Gabor Features and Locality Preserving Projections," in *ACM International Workshop on Multimedia Analysis for Ecological Data (MAED) in conjunction with ACM Multimedia*, Nara, Japan, 2012.

[6] Alexander Loos and Andreas Ernst, "An Automated Chimpanzee Identification System Using Face Detection and Recognition," *EURASIP Journal on Image and Video Processing: Special Issue on Animal Behaviour Understanding in Image Sequences*, vol. 2013, no. 49, 2013.

[7] Alexander Loos, "Chimpanzee Identification Using Global and Local Features," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, 2013.

[8] A. Ernst and C. Küblbeck, "Fast Face Detection and Species Cassification of African Great Apes," in *International Conference on Advanced Video and Signal-Based Surveillance (AVSS)*, Klagenfurt, Austria, Aug. 2011, pp. 279–284, IEEE.

[9] John Wright, Allen Y Yang, Arvind Ganesh, S Shankar Sastry, and Yi Ma, "Robust Face Recognition via Sparse Representation," *IEEE Transactions on Parallel and Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–27, 2009.

[10] C. Küblbeck and A. Ernst, "Face detection and tracking in video sequences using the modified census transformation," *Image and Vision Computing*, vol. 24, no. 6, pp. 564–572, 2006.

[11] R. Liu, Z. Li, and J. Jia, "Image Partial Blur Detection and Classification," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.

[12] M. Graef, *Entwicklung und Implementierung von Modulen zur Erkennung technischer Bildfehler*, Bachelor Thesis, Hochschule für Technik, Wirtschaft und Kultur Leipzig, 2011.

[13] Chih-Chung Chang and Chih-Jen Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011.

[14] T.-F. Wu, C.-J. Lin, and R.C. Weng, "Probability estimates for multi-class classification by pairwise coupling," *Journal of Machine Learning Research*, vol. 5, pp. 975–1005, 2004.

[15] P.C. Mahalanobis, "On the Generalized Distance in Statistics," *Proceedings of the National Institute of Sciences of India*, vol. 2, no. 1, pp. 49–55, 1936.

[16] S. Kotz and S. Nadarajah, *Extreme Value Distributions: Theory and Applications*, Imperial College Press, 2000.

[17] S.R. Eddy, "Maximum Likelihood Fitting of Extreme Value Distributions," Tech. Rep., Department of Genetics, Washington School of Medicine, 1997.

[18] T.J. Ypma, "Historical Development of the Newton-Raphson Method," *Society for Industrial and Applied Mathematics (SIAM) Review*, vol. 37, no. 4, pp. 531–551, 1995.

[19] K. Atkinson, *An Introduction to Numerical Analysis*, John Wiley and Sons, Inc., New York, 1989.