A WEIGHTING STRATEGY FOR ACTIVE SHAPE MODELS

Alma Eguizabal and Peter J. Schreier

Signal and System Theory Group, University of Paderborn, Germany
e-mail: {alma.eguizabal, peter.schreier}@sst.upb.de

ABSTRACT

Active Shape Models (ASM) are an iterative segmentation technique to find a landmark-based contour of an object. In each iteration, a least-squares fit of a plausible shape to some detected target landmarks is determined. Finding these targets is a critical step: some landmarks are more reliably detected than others, and some landmarks may not be within the field of view of their detectors. To add robustness while preserving simplicity at the same time, a generalized least-squares approach can be used, where a weighting matrix incorporates reliability information about the landmarks. We propose a strategy to choose this matrix, based on the covariance of empirically determined residuals of the fit. We perform a further step to determine whether the target landmarks are within the range of their detectors. We evaluate our strategy on fluoroscopic X-ray images to segment the femur. We show that our technique outperforms the standard ASM as well as other more heuristic weighted least-squares strategies.

Index Terms— Active Shape Models, Generalized Least-Squares, fluoroscopic X-rays.

1. INTRODUCTION

Active Shape Models (ASM) [1] are a segmentation technique widely used in computer vision and image processing. They find an object contour parameterized by a set of landmarks using an iterative least-squares fit that minimizes the distance between a plausible object shape and detected target landmarks. ASM adapt to many kinds of shapes and imaging modalities. However, they do require images with reasonably good contrast in order to find the object contours. An example of a challenging scenario is intraoperative fluoroscopic X-ray imaging, which produces low-quality images due to the low X-ray dose. More robust alternatives to ASM exist (e.g., [2]), but they are typically more computationally expensive. In this paper, we show that we can keep the simplicity of ASM and still improve robustness by replacing the least-squares procedure with a generalized least-squares (GLS) approach. The idea of GLS is to weight individual landmarks according to their reliability: reliably identified landmarks should be trusted more than less reliable landmarks. This raises the question of how to identify the right weighting strategy.

This question has already been addressed before: in [3], where directional regularization is proposed; in [4], which suggests to use robust parameter estimation; in [5] and [6], where the weights change in every iteration depending on a score of the target detectors; in [7], which proposes a pose-invariant metric; and in [8], which measures the reliability of the landmarks based on models of local appearance. Many of these weighting strategies are heuristic and also prone to over-fitting. Some of them also add computational complexity to the ASM algorithm.

Our contribution is a strategy for choosing the weights, which has a theoretical justification and still keeps the simplicity of the ASM algorithm. Our proposed strategy measures the reliability of the target landmarks based on the covariance of the residuals of the fit obtained from training data. Additionally, we perform a test to determine whether a landmark is valid, i.e., whether it is within the field of view of its target detector, and we incorporate this test into the weighting matrix.

Our program for this paper is as follows. Section 2 summarizes the ASM algorithm and introduces the GLS approach. In Section 3 we discuss our approach for choosing the weights in the GLS optimization. We also introduce our test for determining whether a landmark is valid. In Section 4, we evaluate the performance of our technique based on a leave-one-out test. Finally, conclusions are drawn in Section 5.

2. PROBLEM FORMULATION

We use ASM to segment an object contour in a 2D image. This segmentation places a set of N landmarks on the contour of the object of interest. The variability of the position of these landmarks for a particular object is modeled in terms of pose and shape: pose refers to any rotation, scale and translation; shape refers to the variability remaining after transforming the landmarks into coordinates with common pose. We use M training images with manually placed landmarks to learn the shape variability. We model the landmarks in a complex space $\mathbb{C}$, where the pose parameters are defined by a complex affine transformation: the landmarks in the mth training image are complex vectors $\boldsymbol{\kappa}_m \in \mathbb{C}^{N \times 1}$, whose real and imaginary parts correspond to the coordinates of the two-dimensional Euclidean space.

To model the shape variability, ASM use a Point Distribution Model (PDM). To build the PDM, we perform a Procrustes alignment [9] to find the coordinate space of the common pose. The pose-aligned vectors are $\mathbf{k}_m = r_m \boldsymbol{\kappa}_m + \mathbf{1} t_m$, where $\{r_m, t_m\} \in \mathbb{C}$ are the pose parameters. Let us define an invertible function $V : \mathbb{C}^{N \times 1} \to \mathbb{R}^{2N \times 1}$ as $V(\mathbf{k}) = [\mathrm{Re}(\mathbf{k})^T, \mathrm{Im}(\mathbf{k})^T]^T$, which provides the corresponding real description. We also define the vector $\mathbf{x}_m = V(\mathbf{k}_m) \in \mathbb{R}^{2N \times 1}$ and the matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M] \in \mathbb{R}^{2N \times M}$ that contains the M aligned training vectors.

We then compute the sample mean of these vectors as $\hat{\boldsymbol{\mu}}_\mathbf{x} = \frac{1}{M} \sum_{m=1}^M \mathbf{x}_m$ as well as the sample covariance matrix $\hat{\mathbf{S}}_\mathbf{x} = \frac{1}{M-1}(\mathbf{X}\mathbf{X}^T - \hat{\boldsymbol{\mu}}_\mathbf{x}\hat{\boldsymbol{\mu}}_\mathbf{x}^T)$. After decomposing $\hat{\mathbf{S}}_\mathbf{x}$ into its eigenvectors $\mathbf{U}$ and eigenvalues $\boldsymbol{\Lambda}$, i.e. $\hat{\mathbf{S}}_\mathbf{x} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$, we keep the p most significant eigenvectors in a matrix $\mathbf{P} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p] \in \mathbb{R}^{2N \times p}$. Consequently, $\mathbf{x}_m$ can be linearly approximated as $\mathbf{x}_m \approx \hat{\boldsymbol{\mu}}_\mathbf{x} + \mathbf{P}\mathbf{b}_m$, and therefore any vector of landmarks can be approximated as

$$\boldsymbol{\kappa}_m \approx \frac{1}{r_m}\left(V^{-1}(\hat{\boldsymbol{\mu}}_\mathbf{x} + \mathbf{P}\mathbf{b}_m) - \mathbf{1}t_m\right), \tag{1}$$

where $\mathbf{b}_m \in \mathbb{R}^{p \times 1}$ is the shape parameter vector of $\boldsymbol{\kappa}_m$, and $\{r_m, t_m\}$ its pose parameters.

We now briefly describe the iteration process of ASM. Let us assume that $\boldsymbol{\kappa}^{(i)}$ is the resulting vector of landmarks after iteration i. Each iteration first examines local regions around each landmark in $\boldsymbol{\kappa}^{(i-1)}$ in order to detect new target landmarks $\tilde{\boldsymbol{\kappa}}^{(i)}$. However, $\tilde{\boldsymbol{\kappa}}^{(i)}$ may not describe a plausible shape or an accurate pose. In order to find the closest plausible shape and pose, that is, the landmarks in vector $\boldsymbol{\kappa}^{(i)}$, the following least-squares problem is solved:

$$\min_{r^{(i)}, t^{(i)}, \mathbf{b}^{(i)} \in \mathbb{B}} ||\boldsymbol{\kappa}^{(i)} - \tilde{\boldsymbol{\kappa}}^{(i)}||^2 \,, \qquad (2)$$

where $\boldsymbol{\kappa}^{(i)} = \frac{1}{r^{(i)}}[V^{-1}(\hat{\boldsymbol{\mu}}_\mathbf{x} + \mathbf{P}\mathbf{b}^{(i)}) - \mathbf{1}t^{(i)}]$, that is, $r^{(i)}, t^{(i)}$ are the pose parameters and $\mathbf{b}^{(i)}$ is the shape parameter vector. To enforce shape plausibility, $\mathbf{b}^{(i)}$ must be contained in $\mathbb{B} = \{\mathbf{b} \in \mathbb{R}^{p \times 1} : \mathbf{b}^T \boldsymbol{\Lambda}_p^{-1} \mathbf{b} \le \xi\}$, where $\boldsymbol{\Lambda}_p$ is a diagonal matrix with the p most significant eigenvalues corresponding to the matrix $\mathbf{P}$, and ξ is a suitable threshold. The expression in (2) is an ordinary least-squares problem, where every component of the difference between the vectors $\boldsymbol{\kappa}^{(i)}$ and $\tilde{\boldsymbol{\kappa}}^{(i)}$ has the same impact on the minimization. These components, i.e. the landmark positions, can be weighted according to their reliability in a generalized least-squares (GLS) problem:

$$\min_{r^{(i)}, t^{(i)}, \mathbf{b}^{(i)} \in \mathbb{B}} (\boldsymbol{\kappa}^{(i)} - \tilde{\boldsymbol{\kappa}}^{(i)})^H \mathbf{W}(\boldsymbol{\kappa}^{(i)} - \tilde{\boldsymbol{\kappa}}^{(i)}) \,, \qquad (3)$$

where $\mathbf{W}$ is the matrix of weights that controls the importance of individual landmarks in the optimization. We will consider both diagonal and nondiagonal weight matrices.

3. PROPOSED SOLUTION

As described in Section 2, in each ASM iteration we need to find the new targets $\tilde{\boldsymbol{\kappa}}^{(i)}$. We determine these by means of a detector $\mathcal{T}(\boldsymbol{\kappa}^{(i-1)}) = \tilde{\boldsymbol{\kappa}}^{(i)}$, which explores the local regions around the previous landmarks $\boldsymbol{\kappa}^{(i-1)}$. We assume that this detector searches for the best match of the gray-level profile around each landmark as measured by the Mahalanobis distance [1]. However, our technique can be generalized to other strategies as well. Section 3.1 and 3.2 describe how to choose the weighting matrix $\mathbf{W}$ based on the empirically determined residual errors.

A complicating factor is that the true landmarks $\boldsymbol{\kappa}^\star$ are not necessarily within the local regions explored by the detector $\mathcal{T}(\boldsymbol{\kappa}^{(i-1)})$, i.e. they are not within its field of view. This is the case when the true landmarks are occluded or out of alignment (especially during the first iterations of the ASM algorithm). We therefore incorporate into our strategy a test of whether or not the true landmark is visible to the detector. If the test determines that a landmark is not visible, then this landmark is excluded by setting the corresponding entry in the weighting matrix $\mathbf{W}$ to zero. This is described in Section 3.3.

3.1. GLS as a Maximum Likelihood (ML) problem

Let us first assume the true landmarks $\boldsymbol{\kappa}^\star$ are within the field of view of the detector $\mathcal{T}$ in iteration i, i.e. when determining $\tilde{\boldsymbol{\kappa}}^{(i)}$. We define the residual error vector as $\boldsymbol{\epsilon}^{(i)} = \boldsymbol{\kappa}^{(i)} - \tilde{\boldsymbol{\kappa}}^{(i)}$. We also assume that $\boldsymbol{\epsilon}^{(i)}$ is complex normal with zero mean and covariance matrix $\mathbf{R}$. Thus, the likelihood of the pose and shape parameters $\{r^{(i)}, t^{(i)}, \mathbf{b}^{(i)}\}$ given the error $\boldsymbol{\epsilon}^{(i)}$ is

$$\mathcal{L}(r^{(i)}, t^{(i)}, \mathbf{b}^{(i)}) = \frac{1}{\pi^N |\mathbf{R}|} \exp\left(-\boldsymbol{\epsilon}^{(i)H} \mathbf{R}^{-1} \boldsymbol{\epsilon}^{(i)}\right) \,, \qquad (4)$$

where ^{H} denotes Hermitian transpose. Given these conditions, the ML estimation of the parameters $\{r^{(i)}, t^{(i)}, \mathbf{b}^{(i)}\}$ is equivalent to a GLS problem [10]:

$$\max_{r^{(i)}, t^{(i)}, \mathbf{b}^{(i)}} \mathcal{L}(r^{(i)}, t^{(i)}, \mathbf{b}^{(i)}) = \min_{r^{(i)}, t^{(i)}, \mathbf{b}^{(i)}} \boldsymbol{\epsilon}^{(i)H} \mathbf{R}^{-1} \boldsymbol{\epsilon}^{(i)} \,, \quad (5)$$

where the constraints on $\mathbf{b}^{(i)}$ imposed in (3) affect the solution, but do not alter the equality in (5). Therefore, if we assume that the residual vector is normally distributed, finding the weighting matrix $\mathbf{W}$ in (3) is equivalent to estimating the covariance matrix of the residuals, i.e. we set $\mathbf{W} = \mathbf{R}^{-1}$. If the training dataset for determining $\mathbf{R}$ is small, it is typically preferable to constrain $\mathbf{W}$ to be diagonal in order to avoid overfitting. This is equivalent to the assumption of statistical independence between the residuals of each landmark.

3.2. Empirical determination of the residual errors

We need to estimate $\mathbf{R}$ based on a set of residual errors that are empirically determined from training images. We first simulate the search of the target landmarks using a detector $\mathcal{T}'$. We split our available training data set containing $2S$ images into two subsets of equal size. With the first subset we train $\mathcal{T}'$; with the second subset we measure the residual errors when employing $\mathcal{T}'$. For this, we simulate the detection of target landmarks, $\mathcal{T}'(\boldsymbol{\kappa}^{(i-1)})$, for each of the S images in the second subset, assuming that the true landmarks of each image are within the field of view of the detector $\mathcal{T}'$. This is achieved by setting $\boldsymbol{\kappa}^{(i-1)} = \boldsymbol{\kappa}_s^\star + \boldsymbol{\delta}$, where $\boldsymbol{\delta} \in \mathbb{C}^N$ is a vector of translations small enough so that the true landmarks of the sth image, $\boldsymbol{\kappa}_s^\star$, remain within the field of view of $\mathcal{T}'(\boldsymbol{\kappa}^{(i-1)})$. We employ the detector that searches for the best match of a gray-level profile as in [1]; hence we determine $\boldsymbol{\delta}$ to place $\boldsymbol{\kappa}^{(i-1)}$ on the line perpendicular to the true object contour, with translations small enough so that $\boldsymbol{\kappa}_s^\star$ is sampled by $\mathcal{T}'(\boldsymbol{\kappa}^{(i-1)})$.

For each of the S images, we perform $\mathcal{T}'(\boldsymbol{\kappa}_s^\star + \boldsymbol{\delta}) = \hat{\boldsymbol{\kappa}}_s$ and measure $\hat{\boldsymbol{\epsilon}}_s = \boldsymbol{\kappa}_s^\star - \hat{\boldsymbol{\kappa}}_s$, which is the residual error obtained from training sample s. To determine the sample covariance matrix of the residuals we assemble the error matrix $\hat{\mathbf{E}} = [\hat{\boldsymbol{\epsilon}}_1, \ldots, \hat{\boldsymbol{\epsilon}}_S] \in \mathbb{C}^{N \times S}$, which contains the residuals from all training samples. The sample covariance matrix is then

$$\hat{\mathbf{R}} = \frac{1}{S} \hat{\mathbf{E}} \hat{\mathbf{E}}^H. \qquad (6)$$

We note that the so determined $\hat{\mathbf{R}}$ does not depend on the iteration i of the ASM algorithm.

3.3. Testing whether a target landmark is valid

The sample covariance matrix of the residual errors $\hat{\mathbf{R}}$ is determined under the assumption that the true landmarks $\boldsymbol{\kappa}^\star$ are within the field of view of the detectors $\mathcal{T}(\boldsymbol{\kappa}^{(i-1)})$. If the true landmarks are not visible to the detector, the determined target landmarks are certainly incorrect. In such a case, the target landmarks are not valid. In order to test whether a target landmark in $\tilde{\boldsymbol{\kappa}}^{(i)}$ returned by the ASM algorithm is valid, we run a hypothesis test based on the metric that is used to determine the target landmark. In our case, this metric is a Mahalanobis distance $d_n^{(i)}$ that measures the distance between the observed and the modeled local appearance. Let us denote by $\mathbf{g}_n^{(i)} \in \mathbb{R}^{\ell \times 1}$ the vector containing ℓ intensity values representing the appearance observed by detector $\mathcal{T}(\boldsymbol{\kappa}^{(i-1)})$ to determine landmark n. The null hypothesis $\mathcal{H}_0$ is that the target landmark, which corresponds to the point that minimizes $d_n^{(i)}$, is valid.

We assume $\mathbf{g}_n^{(i)} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{g}_n}, \mathbf{S}_{\mathbf{g}_n})$. Thus, the Mahalanobis distance $d_n^{(i)} = (\mathbf{g}_n^{(i)} - \boldsymbol{\mu}_{\mathbf{g}_n})^T \mathbf{S}_{\mathbf{g}_n}^{-1}(\mathbf{g}_n^{(i)} - \boldsymbol{\mu}_{\mathbf{g}_n})$ follows a chi-squared distribution with ℓ degrees of freedom under the null hypothesis:

$$\mathcal{H}_0 : d_n^{(i)} \sim \chi_\ell^2 \tag{7}$$

Therefore, for each iteration i and each landmark n, we perform a chi-squared test of $d_n^{(i)}$: if the null hypothesis is not rejected, then we set an indicator variable $f_n^{(i)} = 1$; otherwise $f_n^{(i)} = 0$. We can fix a false alarm rate for this test based on the χ_ℓ^2 distribution. The idea of this approach is to work only with target landmarks whose observed gray-level profiles are close enough to the model profiles.

This test can be incorporated into the weighting matrix $\mathbf{W}$. For each iteration i, we define a diagonal matrix $\mathbf{F}^{(i)} = \mathrm{diag}(f_1^{(i)}, \ldots, f_N^{(i)})$. The weighting matrix $\mathbf{W}^{(i)}$ is then obtained as

$$\mathbf{W}^{(i)} = \mathbf{F}^{(i)} \hat{\mathbf{R}}^{-1} \mathbf{F}^{(i)}. \tag{8}$$

4. RESULTS AND DISCUSSION

To evaluate the performance of our proposed method we choose a challenging segmentation task: the contour of the femur in fluoroscopic (low-dose) X-ray images. The quality of this image modality is low and thus the performance of the standard ASM algorithm suffers. Our database contains 350 gray-scale images that have been acquired during surgeries treating hip fractures in an approximate anterior-posterior orientation. These images show the upper part of the femur and part of the hip. They belong to different surgical interventions and C-arm devices. The image sizes range between 450x450 to 510x510 pixels. We have a ground truth consisting of a manually segmented femur contour and landmarks for every image. We show one of these images in Fig. 1. As these images come from surgeries treating hip fractures, they contain an intermedullary nail and a screw, which further complicate the segmentation task with occlusions.
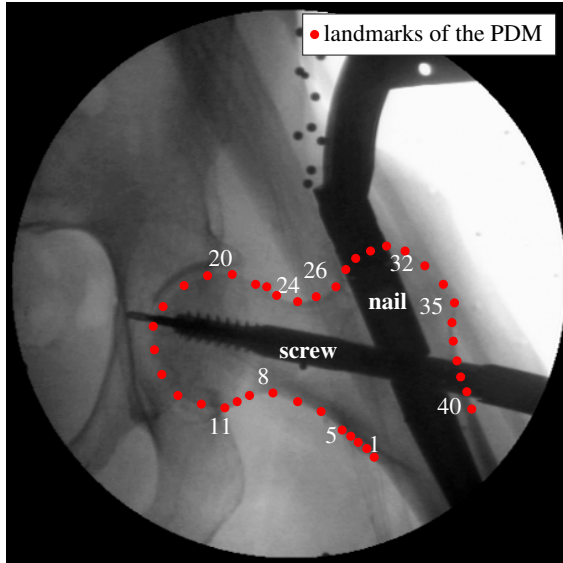

Fig. 1. Landmarks of a PDM (1 to 40) of the femur in a fluroscopic X-ray image in anterior-posterior orientation.

4.1. Leave-one-out test

We evaluate the performance of our method using a leave-one-out test. For all the images in the database we "leave one image out" to test and keep the remaining as training images. The process is repeated for every available image and the results are averaged over all test images. For each leave-one-out test iteration t, we obtain $\hat{\mathbf{P}}_{(t)}$, $\hat{\mathbf{\Lambda}}_{(t)}$ and detector $\hat{\mathcal{T}}_{(t)}$ from the training images. To perform the chi-squared test, for every landmark n we also obtain $\hat{\boldsymbol{\mu}}_{\mathbf{g}_n(t)}$ and $\hat{\mathbf{S}}_{\mathbf{g}_n(t)}$. We determine the sample covariance of the residuals $\hat{\mathbf{R}}_{(t)}$ as described in Section 3.2.

We performed this leave-one-out test for all considered strategies. The same conditions (initial guess, shape parameters, number of ASM iterations...) were used for all, only changing the definition of the weighting matrix $\mathbf{W}$. For each strategy, we ran 100 ASM iterations on each image. We then measured the squared distance of every resulting landmark to the ground truth contour. We show the root mean square error (RMSE) for each PDM landmark in Fig. 2.

4.2. Evaluation of the proposed method

We implemented two different versions of our approach: one where we use the full matrix $\mathbf{W}$, and one where we restrict it to be diagonal. We compared these to the following strategies:

1. the standard ASM [1] without weighting, i.e., $\mathbf{W} = \mathbf{I}$.

2. $\mathbf{W}$ as a diagonal matrix with diagonal elements $1/\hat{d}_n$, as proposed by [5].

3. $\mathbf{W}$ as a diagonal matrix with diagonal elements $1/(1 + \mathrm{trace}(\mathbf{S}_{\mathbf{g}_n}))$, as proposed by [8].

As shown in Fig. 2, our strategy outperforms all the other strategies on average, in particular when $\mathbf{W}$ is diagonal, i.e. assuming independent residuals (this could be due to the small sample size available). The improvement our strategy provides over other strategies is especially significant in the area of the femoral head (between landmark 11 and 20, with $\sim$30% improvement achieved). This is the most challenging area of segmentation since there is overlap from other hip bones, and contour edges are particularly weak. Also, for the surgical procedure from which we obtained these images, this is an important region of interest [11].

As an example, we show in Fig. 3 the landmarks found after 100 ASM iterations for our proposed weighting strategy with diagonal $\mathbf{W}$ compared to the standard ASM, for a particular image in the database. Our strategy follows the real contour much more closely.

We also further investigated the hypothesis test described in Section 3.3: We measured the norm d_n after the simulation of several ASM iterations. Then we collected the observed $\hat{d}_n$ values either as $\hat{d}_n | f_n = 1$, if the true point was valid (null hypothesis), or $\hat{d}_n | f_n = 0$ otherwise. We computed normalized histograms of $\hat{d}_n$, as shown in Fig. 4, to see how well a chi-squared distribution fits $\hat{d}_n | f_n = 1$. The approximation is fairly good, although it does not account for a strong tail in the histogram of $\hat{d}_n | f_n = 1$. We set a probability of false alarm, i.e. the probability of classifying a valid landmark as invalid, of 10% based on the approximating χ_ℓ^2 distribution. This covers $\sim$75% of the values from $\hat{d}_n | f_n = 1$, but only $\sim$30% from $\hat{d}_n | f_n = 0$. We suggest to set a relatively low false alarm threshold, which prevents the least-squares solution to be implausible. If too many landmarks are regarded as invalid and excluded, the problem may become underdetermined, and solutions may be unreliable.
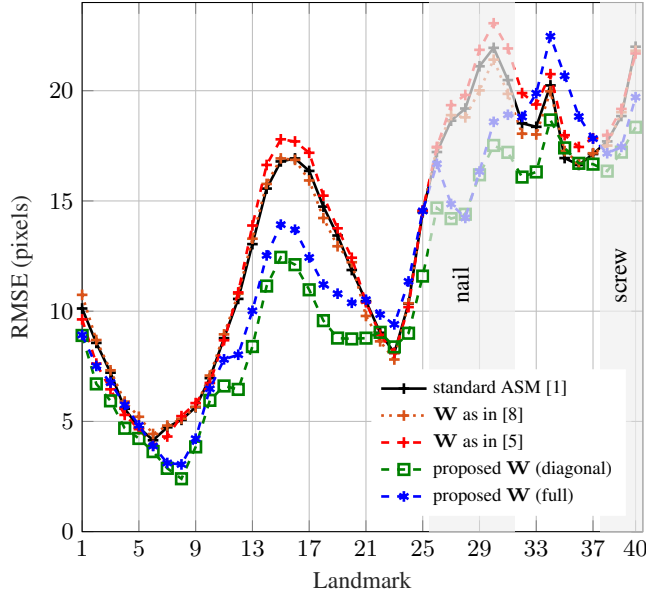
Fig. 2. RMSE of our proposed strategy compared with [8], [5] and [1]. We mask in gray the areas where the surgical implants (nail and screw) most likely occlude the bone contour.
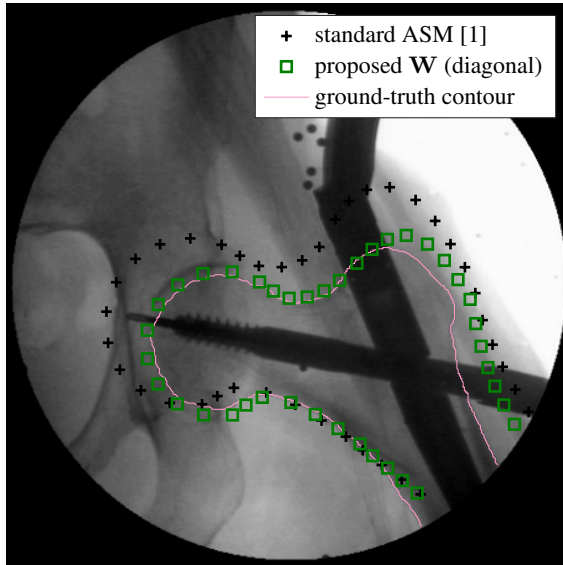

Fig. 3. One example of a segmented femur, comparing our technique with the standard ASM algorithm (without weighting).

5. CONCLUSIONS

We proposed a GLS instead of an ordinary least-squares fit as a straightforward and simple strategy to add robustness to the ASM algorithm. The weights in the GLS fit should not be selected only assuming heuristic confidence metrics of the landmarks, since these do not generalize and are prone to over-fitting. We proposed to choose the weights based on empirically determined errors of the least-squares fit, where the weighting matrix $\mathbf{W}$ is determined as the inverse of the sample covariance of the residuals. This can be inter-
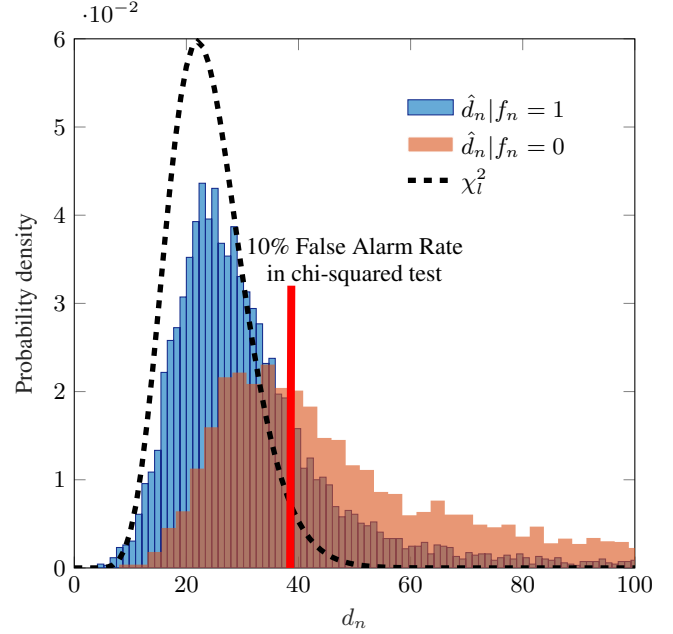

Fig. 4. Normalized histograms of the distances $\hat{d}_n|f_n = 1$ and $\hat{d}_n|f_n = 0$ and the approximating chi-squared distribution for comparison.

preted as a Maximum Likelihood solution of the least-squares problem. Additionally we used a chi-squared test to identify target landmarks that are likely to be incorrect and thus exclude them from the fit. Therefore, we combine prior knowledge about the performance of the target detectors, $\hat{\mathbf{R}}$, with additional information about the unreachable landmarks (either occluded or out of alignment), $\mathbf{F}$. We tested the strategy in fluoroscopic X-ray images of the femur taken in actual surgeries. We showed that our strategy outperforms the standard ASM as well as other weighting strategies. While our approach was based on a particular metric for identifying target landmarks (the Mahalanobis distance), our idea should generalize to other metrics as well.

6. REFERENCES

[1] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham, "Active shape models–their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38 – 59, 1995.

[2] C. Lindner, S. Thiagarajah, J. M. Wilkinson, The arcOGEN Consortium, G. A. Wallis, and T. F. Cootes, "Fully automatic segmentation of the proximal femur using random forest regression voting," *IEEE Transactions on Medical Imaging*, vol. 32, no. 8, pp. 1462–1472, Aug 2013.

[3] A. Hill, T.F. Cootes, and C.J. Taylor, "Active shape models and the shape approximation problem," *Image and Vision Computing*, vol. 14, no. 8, pp. 601 – 607, 1996.

[4] M. Rogers and J. Graham, "Robust active shape model search," in *Proc. of the 7th European Conference on Computer Vision*, A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, Eds. 2002, pp. 517–530, Springer.

[5] M. Zhao, S. Z. Li, C. Chen, and J. Bu, "Shape evaluation for

weighted active shape models," in *Proc. of the Asian Conference on Computer Vision*, 2004, pp. 1074–1079.

[6] G. Ruiz, E. Ramon, J. Garca, M. A. G. Ballester, and F. M. Sukno, "Weighted regularized ASM for face alignment," in *Proc. of the IEEE International Conference on Image Processing (ICIP)*, Sept 2016, pp. 2906–2910.

[7] K. Lekadir, R. Merrifield, and G. Z. Yang, "Outlier detection and handling for robust 3-d active shape models search," *IEEE Transactions on Medical Imaging*, vol. 26, no. 2, pp. 212–222, Feb 2007.

[8] X. Yang, J. Cerrolaza, C. Duan, Q. Zhao, J. Murnick, N. Safdar, R. Avery, and M. G. Linguraru, "Weighted partitioned active shape model for optic pathway segmentation in MRI," in *Proc. of Clinical Image-Based Procedures. Translational Research in Medical Imaging*. 2014, pp. 109–117, Springer.

[9] F. L. Bookstein, "Landmark methods for forms without landmarks: morphometrics of group differences in outline shape," *Medical Image Analysis*, vol. 1, no. 3, pp. 225 – 243, 1997.

[10] A. Charnes, E. L. Frome, and P. L. Yu, "The equivalence of generalized least squares and maximum likelihood estimates in the exponential family," *Journal of the American Statistical Association*, vol. 71, no. 353, pp. 169–171, 1976.

[11] L. Wang, M. Kohnen, O. Friman, and H. K. Hahn, "Fast automated segmentation of femoral heads in fluoroscopic x-ray images," in *Proc. of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2011, pp. 984–988.