Effect of Segmentation Errors on 3D-to-2D Registration of Implant Models in X-ray Images

Mohamed R. Mahfouz, PhD[1,2]

William A Hoff, PhD[1,2]

[1]Rocky Mountain Musculoskeletal Research Laboratory

Denver, CO

[2]Colorado School of Mines

Golden, CO

Send Correspondence To: Mohamed R. Mahfouz
RMMRL
2425 So Colorado Blvd., #280
Denver, CO 80222
Phone: (303)996-1270
Fax: (303) 759-2316
e-mail: mmahfouz@rmmrl.org

Abstract

In many biomedical applications, it is desirable to estimate the three dimensional (3D) position and orientation (pose) of a metallic rigid object (such as a knee or hip implant) from its projection in a two dimensional (2D) X-ray image. If the geometry of the object is known, as well as the details of the image formation process, then the pose of the object with respect to the sensor can be determined. A common method for 3D-to-2D registration is to first segment the silhouette contour from the X-ray image; that is, identify all points in the image that belong to the 2D silhouette and not to the background. This segmentation step is then followed by a search for the 3D pose that will best match the observed contour with a predicted contour.

Although the silhouette of a metallic object is often clearly visible in an X-ray image, adjacent tissue and occlusions can make the exact location of the silhouette contour difficult to determine in places. Occlusion can occur when another object (such as another implant component) partially blocks the view of the object of interest. In this paper, we argue that common methods for segmentation can produce errors in the location of the 2D contour, and hence errors in the resulting 3D estimate of the pose. We show, on a typical fluoroscopy image of a knee implant component, that interactive and automatic methods for segmentation result in segmented contours that vary significantly. We show how the variability in the 2D contours (quantified by two different metrics) corresponds to variability in the 3D poses. Finally, we illustrate how traditional segmentation methods can fail completely in the (not uncommon) cases of images with occlusion.

1 Introduction

In many biomedical applications, it is desirable to estimate the three dimensional (3D) position and orientation (pose) of a metallic object from its projection in a two dimensional (2D) X-ray image. For example, Figure 1 shows a knee implant consisting of two metallic components, and an X-ray fluoroscopy image of the knee implant within the body. Note that only the silhouettes of the implant components are visible.

In some applications, only a single 2D image is possible to obtain. For example, in "in vivo" kinematic analysis of knee movement, X-ray fluoroscopy is used to image the motion of a knee implant during dynamic, weight-bearing activities such as gait, stair step, and chair rise. With single-plane fluoroscopy, the patient is allowed free motion in the plane between the X-ray source and the image intensifier (Figure 2). It has been shown that one can achieve errors of one degree or less in rotation and one mm or less in translation parallel to the image plane (Banks and Hodge, 1996; Mahfouz et al., 2002).

Although fluoroscopy images are only two-dimensional images, we can recover all six degrees-of-freedom of the pose if we have an accurate geometric model of the object (implant component). This is based on the fact that, given the model of the object and the model of the image formation process, the appearance of the object in the image can be predicted. The fluoroscope can be modeled by a perspective projection image formation model (Figure 3). This model treats the fluoroscope as consisting of an X-ray point source and a planar phosphor screen upon which the image is formed. Although image distortion and non-uniform scaling can occur, these can be compensated for by careful calibration.

The predicted image of the model is dependent on all six degrees of freedom of its pose (*i.e.*, three translational and three rotational degrees of freedom). For example, moving the object away from the sensor reduces the size of the silhouette in the image. Rotating the object changes the shape of the image silhouette. The silhouette is uniquely determined by the pose in a unique manner (unless the object is symmetrical, in which case situations may arise where the pose is ambiguous). By searching the space of possible poses, one can find the pose for which the predicted image best matches the actual image of the object.

A common method for 3D-to-2D registration is to first segment the silhouette contour from the X-ray image; that is, identify all points in the image that belong to the 2D silhouette and not to the background. This segmentation step is then followed by a search for the 3D pose that will best match the observed contour with a predicted contour.

Although the silhouette of a metallic object is often clearly visible in an X-ray image, adjacent tissue such as bone can reduce the contrast between the object and background, and make the exact location of the silhouette contour difficult to determine. For example, in Figure 4 the edge between the object and background is not sharp, but rises gradually over a short distance. The actual location of the boundary could be anywhere in this range, and it would be difficult to determine the exact location based on local intensity information alone. Unfortunately, an error in the estimated location of the contour can lead to significant errors in the estimated pose of the object.

Occlusions are another situation where the location of the boundary is difficult to determine. This can occur when the projection of the implant overlaps the other leg (Figure 5, left), or another implant (Figure 5, right). In the overlap area, the boundary is completely obscured in the image.

Typical edge segmentation methods attempt to identify discrete image points that belong to the silhouette of the object. Unfortunately, it is difficult to correctly identify these points based on local intensity information alone, as the above examples illustrate. Misplacing the location of the contour can lead to errors in the resulting estimate of the object pose. Essentially, early segmentation of the contour is an irreversible and premature decision that is made based on limited information. It does not use information about the 3D shape of the object. A possibly better approach is to delay segmentation and integrate it with the process of pose estimation.

In this paper, we argue that common methods for segmentation can produce errors in the location of the 2D contour, and hence errors in the resulting 3D estimate of the pose. We show, on a typical fluoroscopy image of a knee implant component, that interactive and automatic methods for segmentation result in segmented contours that can vary significantly. We show how the variability in the 2D contours (quantified by two different metrics) corresponds to variability in the 3D poses. Finally, we illustrate how traditional segmentation methods can fail completely in the (not uncommon) cases of images with occlusion.

2 Registration Methods

A common approach for 3D-to-2D registration is a hypothesize-and-test process (Figure 6), whereby features are extracted from the image, then matched to a predicted set of features

computed from the model in a hypothesized pose (Grimson, 1990). Another term for this is "recognition by alignment" (Shapiro and Stockman, 2001). Based on the quality of the match, the pose is refined using an optimization or search algorithm and the process is iterated until convergence or failure.

The following sections describe key elements of the object recognition process; namely, feature detection, matching, and optimization (or search).

2.1 Feature Detection

Perhaps the simplest method of segmenting the object from the background is to threshold the image; that is, assign all pixels with intensity less than a fixed value to the class of "object", and the others to "background". In the case of implant components, this works because the silhouette of a dense metallic object is darker than soft tissue in fluoroscopy images. Figure 7 (left) shows a binary image obtained by thresholding the image in Figure 1. The white region is the segmented femoral component, where other segmented regions (such as that corresponding to the tibial component) have been manually deleted. Figure 7 (center) shows the resulting contour.

This method, used by Banks (1992) and Hoff *et al*., (1998) can result in localized errors in the silhouette. Therefore, a person can manually edit the silhouette to make it better fit the observed image intensity boundary. Using a drawing program such as Adobe Photoshop™, the person views the silhouette as an overlay on the input image. The person can then erase small protrusions in the silhouette, or add small regions, to make the silhouette better fit the observed image. Figure 7 (right) shows the contour after editing of the silhouette in this way. The contour

qualitatively looks much better. However, this method can result in significant variation from person to person, as shown in the next section.

A more automated method to identify edges is the method of deformable models, or "active contours" (Kass et al., 1987). Also known as "snakes", this method requires the person to place an initial coarse estimate of the contour in the image. Then the contour automatically deforms itself to align to the image data (usually to the image points with high gradient magnitude), while minimizing the stretching and bending energies of the contour. For example, Figure 8 shows the result of fitting a snake to the femoral component in Figure 1, using the implementation of Xu and Prince, (1998).

Snakes have been used to extract contours from CT slices and fluoroscopy images (Gueziec et al., 1998). However, as shown in Section 3.2, snakes may have difficulty accurately locating the contour in areas where the curvature is high, or where there is low contrast between the object and background (as in Figure 4).

The above methods result in a complete segmented contour of the object. Other edge detection methods do not explicitly create closed contours, nor do they distinguish object edge points from background (or noise) edge points. The Canny edge detector (Canny, 1986) has been used to extract contours from fluoroscopy images (Zuffi et al., 1999). As shown in Figure 9, the Canny operator does a good job of finding the contour around the femoral component, although it may have difficulty in low contrast areas. However, it does not create a closed contour (there are small gaps) and it does not identify points explicitly belonging to the femoral contour.

Additional processing would still be necessary to identify the contour points for the object of interest.

Rather than performing feature detection to explicitly segmenting the contour, an alternative is to match the image values directly to a predicted image of the object. A predicted image is generated of the object in a hypothesized pose, and the pixel values are compared directly to the values in the actual input image, without trying to pre-segment the object from the image. For X-rays, the predicted images are known as "digitally reconstructed radiographs" (DRRs). With this approach, a 3-D volumetric model, rather than a surface model, can be used. Researchers have matched 3-D volumetric models derived from CT, MRI, or PET data to static X-rays (Lemieux et al., 1994) or fluoroscopy images (Penney et al., 1998; Weese et al., 1997). Although most past approaches have focused on CT-derived volumetric models, direct image comparison methods can also be used for surface models (*i.e.*, implants) (Mahfouz et al., 2001; Sarojak et al., 1999). With this technique, it is not necessary to segment a complete and valid contour prior to registration, and no human intervention is required. However, it is necessary to have either a good initial estimate of the pose, or else a robust optimization algorithm to perform the search.

2.2 Matching

Given a set of features (such as a silhouette contour) extracted from the image, it is necessary to match those features to a set of expected features derived from the model. A variety of similarity measures, or metrics, have been used.

One simple measure compares the binary image of the segmented silhouette with a binary image of the predicted silhouette. The predicted silhouette is created by rendering the object model in a hypothesized pose. A measure of the difference between the images is the count of the pixels that are different between the two silhouettes (Hoff et al., 1998). An alternative method was used by Banks et al., (1996), who represented the silhouette contour as a set of Fourier descriptors. The measure of the difference between the measured silhouette and the predicted silhouette was the Euclidean distance between the corresponding vectors of Fourier descriptors. In both these methods, a complete closed contour or silhouette of the object of interest must be available.

Another metric is the distance between predicted contour points and actual contour points in the image. For example, Feldmar et al., (1997) measures the sum of the squared errors between points on the projected model silhouette and the identified edge contour. This approach is also known as "chamfer" matching, and has been widely used in the computer vision field e.g., (Basri and Weinshall, 1996). An alternative is to measure the distances in three dimensions (Gueziec et al., 1998; Lavallee and Szeliski, 1995; Zuffi et al., 1999). Here, the measured points on the contour in the image correspond to 3-D rays, projecting outward from the sensor to infinity. The 3-D distances between these rays and the closest point on the surface of the object are recorded. In these methods, only a limited number of points on the contour are needed, rather than the complete closed contour. However, the points must belong to the object of interest, rather than to the background, or else the metric will be adversely affected (however, it is possible to mitigate the effect of outliers somewhat by using weighting methods such as robust estimators).

Another class of similarity measures does not require any pre-segmentation of the object contour, but rather operates directly on the image intensities. As described in the previous section, a DRR is generated to create a predicted image of the model. This image is compared directly to the original image using measures such as cross correlation or pattern intensity (Lemieux et al, 1994; Mahfouz et al., 2001; Penney et al., 1998; Sarojok et al., 1999;Weese et al., 1997). Some methods also incorporate image gradients in the similarity measure.

2.3 Search Methods

The type of feature detector used (if any) affects the choice of the search method that is used to find the pose of the object of interest. In general, a feature detector that produces incomplete and spurious data will require a more sophisticated search algorithm. For example, a simple look-up table method (Banks and Hodge, 1996; Hoff et al, 1998) can be used to estimate the pose of the object when a complete, closed contour can be extracted from the input image. If the contour is incomplete (but no spurious data is present), a least-squares algorithm (Lavallee and Szeliski, 1995) can be used.

If the input data has spurious points, or if no segmentation is done at all to identify the points belonging to the object of interest, then a more sophisticated search algorithm must be used. A hierarchical or multiresolution approach can be used (Hellier et al., 2001), as well as robust estimators Feldmar et al., 1997). In general, these methods require a close initial estimate of the pose to be able to converge to the correct solution. Global search methods such as simulated annealing (Mahfouz et al, 2001, Sarojok et al, 1999) or genetic algorithms (Matsopoulous et al., 1999; Rouet et al., 2000) are less sensitive to the initial starting conditions, but these can be very slow.

3 Experiments

We illustrate how different methods of segmentation can result in errors in the resulting pose after registration, using the fluoroscopy image in Figure 1. This image was taken by a VF-2000 fluoroscope from Radiographic and Data Solutions, Inc. (Minneapolis, MN). Images were captured using a progressive scan video camera and subsequently digitized to 8 bits and 640x480 pixels using a frame grabber attached to a PC. The image intensifier had a diameter of 12 inches, and the principal distance was approximately 1200 mm.

For comparison, we also give results from a 3D-to-2D registration method (Mahfouz et al, 2001, Sarajok et al, 1999) that does not perform explicit segmentation; but instead performs direct image-to-image matching. The similarity measure for matching the predicted X-ray image and the actual input X-ray image is similar to a cross-correlation between the two images. Cross correlation is also performed on an edge-enhanced version of the original image and the predicted X-ray image, and the two scores are combined. The resulting total matching score produces a distinct minimum when the CAD model is exactly aligned with the image of the implant in the input x-ray image (Figure 10). Using exhaustive search, the optimal pose of the model (in terms of this similarity measure) was found. This pose was used as the "gold standard" for subsequent comparisons.

3.1 Human Segmentation

The first segmentation method was simple thresholding followed by human editing, as described in Section 2.1. The fluoroscopy image was first processed with a 3x3 median filter to reduce noise. A threshold level was then chosen interactively to give good visual segmentation of the

femoral component. The binary region belonging to the femoral component was saved and all other regions discarded (Figure 7). Twelve silhouettes were created by twelve different human volunteers. The resulting silhouettes showed some minor variations, particularly in places where the contrast between object and background was low. The following subsections describe the variability of the silhouettes using 2D image metrics (Section 3.1.1) and 3D pose metrics (Section 3.1.2).

3.1.1 Analysis Using 2D Image Metrics

We first analyzed the difference between the human-generated silhouettes and the ideal ("gold standard") silhouette using 2D image metrics. For illustration, we used the same 2D image metrics to compare the ideal silhouette to a set of synthetic displaced silhouettes, created by rendering the model in a range of poses around the ideal pose. The displaced silhouettes were generated by varying the X,Y angles (*i.e.*, the out-of-plane rotation angles) of the model by ±2º, at increments of 1º, for a total of 25 images[1].

Two image metrics were used to compare the silhouette contours. The first metric, the image difference metric, was the count of the pixels that were different between two binary silhouettes, as described in Section 2.2. The second metric, the Fourier descriptor metric, was the Euclidean distance between the corresponding vectors of Fourier descriptors, also described in Section 2.2. The difference values between each contour and the ideal silhouette were sorted in increasing order and plotted. The results for the two metrics are shown in Figure 11 and Figure 12.

[1] Although the model has a total of six degrees of freedom (three translational and three rotational), we limited the variation to the out-of-plane angles, since these affect the shape of the silhouette.

Although the exact order of the silhouettes varies, it is evident that the silhouettes created by human editing vary significantly from each other and from the true silhouette. The distance of the "best" human edited silhouette from the true pose is about the same as the distance of a synthetic silhouette that has approximately 1º error in X and Y angles. The "worst" human edited silhouette corresponds to a synthetic silhouette that has approximately 2º error in X and Y angles. If sub-degree accuracy is desired, the variation in human editing may make this method of segmentation unreliable.

3.1.2 Analysis Using 3D Pose Metrics

The analysis in the preceding section shows that human-generated silhouettes can vary significantly from each other and from the true silhouette, as measured using 2D image metrics. To analyze the variability of the resulting 3D poses after registration, we need to actually register the human-generated silhouettes to the model. This requires a search to find the predicted silhouette that most closely matches the input.

We applied the 3D-to-2D registration method described in (Mahfouz et al, 2001, Sarojok et al, 1999) to register the human-generated silhouettes to the model, and recorded the difference between the resulting poses and the true pose. The RMS (root mean square) values of the errors are given in Table 1. The table also shows the pose errors from registering the model directly to the image without any segmentation. This registration was repeated 30 times from different initial starting poses, and allowed to run to completion to find the best match. (The variation is

due to the fact that the search algorithm may converge upon a local minimum.) The results show that the error in the resulting pose is much higher if segmentation is done prior to registration.

3.2 Semi-automatic Segmentation

Two semi-automatic methods of segmenting contours were analyzed. The first method was the method of active contours (snakes). The image was smoothed slightly with a Gaussian filter with a standard deviation of 0.5 pixels. An initial coarse estimate of the contour was manually placed around the femoral component (Figure 8, left), and the snake was allowed to converge to its final pose (Figure 8, right). The second method was the Canny edge detector (Canny, 1986) (Figure 9). The edge points were manually edited to remove all points not belonging to the femoral component silhouette. In addition, small gaps (up to 2-3 pixels long) in the contour were manually filled so that the result was a complete closed contour.

To analyze the variability of the segmented contours, we added noise to the image in Figure 1 prior to segmentation. Random Gaussian distributed noise (μ=0, σ=20) was added to the image[2] (Figure 13). Five noisy images were generated in this way, and segmentation performed for each one.

The same two image metrics of Section 3.1 were used to compare the silhouette contours. The result for the image difference metric is shown in Figure 14, and the result for the Fourier descriptor metric is shown in Figure 15. These results show that noise causes silhouettes

[2] This is much higher that the observed measured pixel intensity noise on our fluoroscope and image capture equipment, which is approximately 3 gray levels.

generated by semi-automated segmentation to vary significantly from each other and from the true silhouette, as measured using 2D image metrics.

To analyze the variability of the resulting 3D poses, we registered the segmented silhouettes to the model, and recorded the difference between the resulting poses and the true pose. The RMS values of the errors are given in Table 2. The table also shows the pose errors from registering the model directly to the noisy images without any segmentation. This registration was repeated 30 times from different initial starting poses, and allowed to run to completion to find the best match. The results show that the resulting poses of the segmented contours are more variable than if segmentation is done prior to registration. However, the variability of the auto-segmented contours is not as great as the human-segmented contours.

3.3 Failures in Automatic Edge Detection

In many images, automatic edge detectors such as the Canny operator are not able to extract a complete closed contour, due to occlusion and noise. For example, Figure 16 (left) shows an image where the projection of the implant overlaps the other leg. In this case, the left side of the femoral component silhouette is completely lost. Figure 16 (right) shows the result of applying the Canny operator to this image. Here, a low threshold was used in order to bring out as much of the silhouette contour as possible. Note the complete loss of the contour along the left side.

In cases such as these, if a complete contour were needed for registration, manual editing would be necessary to complete the contour. However, given the demonstrated variability in manual editing, the results would probably not be reliable.

4 Discussion

A common method for 3D-to-2D registration is to first segment the silhouette contour from the X-ray image; that is, identify all points in the image that belong to the 2D silhouette and not to the background (Banks and Hodge, 1996; Banks, 1992; Hoff et al., 1998). However, errors in segmenting the silhouette contour of an object from an X-ray image can lead to errors in the estimated pose of the object. Adjacent tissue and occlusions can make the exact location of the silhouette contour difficult to determine in places. We have shown that human editing can lead to wide variation among the contours. We also implemented two additional semi-automated segmentation methods – the active contour (snake) method and the Canny edge detector. We compared contours obtained from segmentation with a set of predicted contours, generated by rendering the model at a series of known poses. For comparison, we also implemented a registration method that directly matches the model to the image, without segmentation.

The results showed that the contours obtained from human segmentation vary significantly. Using 2D image metrics, the contours differed from the true contour by an amount corresponding to up to 2 degrees of rotation. After registration to the segmented contours, RMS errors of approximately 0.85 degrees were observed in the Y angle, as opposed to only 0.15 degrees when no segmentation was done.

Similarly, the contours obtained from semi-automated segmentation (Canny and snakes) also varied significantly when noise was added to the image. After registration to the segmented contours, the RMS errors in the Y angle were 0.84 degrees and 0.61 for the Canny and snake contours, respectively. However, when no segmentation was done the Y angle error was only

0.34 degrees. Finally, we showed that segmentation fails completely in the (not uncommon) cases of images with occlusion. In these cases, it is not possible to obtain a complete contour of the object.

Our results suggest that registration might be more accurate if no *a priori* image segmentation is done, since noise can introduce irreversible errors in the segmented contour that affect the resulting pose accuracy. A possibly better approach is to match the object model directly to the image, which effectively integrates the process of segmentation with pose estimation. This also has the advantage of using information about the 3D shape of the object.

5 Acknowledgements

The authors wish to acknowledge the support of the Rocky Mountain Musculoskeletal Research Laboratory in this work. We also thank the anonymous reviewers of this paper for helpful comments and suggestions, which improved the content and presentation.

Tables

Table 1 RMS (root mean square) pose errors, after registration to (a) human-segmented contours, and (b) directly to the image without segmentation.

	Translation error (mm)			*Rotation error (degrees)*		
Method	*X*	*Y*	*Z*	θ_X	θ_Y	θ_Z
(a) Registration to human segmented contours	0.322	0.178	6.588	0.291	0.845	0.530
(b) Registration directly to image (no segmentation)	0.080	0.044	1.367	0.145	0.157	0.338

Table 2 RMS (root mean square) pose errors, after registration to (a) Canny-segmented contours, (b) snake-segmented contours, and (c) directly to the image without segmentation.

	Translation error (mm)			*Rotation error (degrees)*		
Method	*X*	*Y*	*Z*	θ_X	θ_Y	θ_Z
(a) Registration to Canny-segmented contours	0.196	0.220	2.752	0.336	0.837	0.264
(b) Registration to snake segmented contours	0.499	0.279	5.191	0.477	0.609	0.800
(c) Registration directly to image (no segmentation)	0.140	0.098	3.164	0.179	0.343	0.537

6 Legends

Figure 1 (Left) Artificial knee implant, with tibial and femoral components. (Right) "In vivo" fluoroscopy image of knee implant.

Figure 2 Single plane fluoroscopy allows the patient to perform deep knee bends and other activities unimpaired.

Figure 3 (Left) The imaging model for the fluoroscope is perspective projection. X-rays are emitted from a point source (F), pass through the object (such as point G), and strike the image plane (point Q). (Right) A graphical presentation of the imaging model. The silhouette of the model can be predicted and compared to the observed silhouette in the image.

Figure 4 The boundary between the object (implant component) and background (bone) is gradual, as shown by the intensity profile along the path indicated by the white line.

Figure 5 Occlusion occurs when the projection of the implant component overlaps the other leg in the image (left) or another implant (right).

Figure 6 Hypothesize-and-test model-based object recognition paradigm for 3D-to-2D registration.

Figure 7 (Left) The segmented femoral silhouette from Figure 1, obtained by thresholding. (Center) The silhouette contour, shown as an overlay on the original image. (Right) The contour after human editing.

Figure 8 (Left) Initial manual placement of the silhouette contour. (Right) The deformed contour after automatic processing.

Figure 9 Result of edge detection using the Canny operator.

Figure 10 The matching score achieves a distinct minimum when the femur model was exactly aligned with the image of the implant in Figure 1. To create this plot, the pose of the model was varied in two of its degrees of freedom (rotation about the X and Y axes).

Figure 11 Using the image difference metric, the differences between each contour and the ideal silhouette are calculated, and plotted in increasing order.

Figure 12 Using the Fourier descriptor metric, the differences between each contour and the ideal silhouette are calculated, and plotted in increasing order.

Figure 13 The test fluoroscopy image with a large amount of added noise.

Figure 14 Using the image difference metric, the differences between each contour and the ideal silhouette for the noisy images are calculated, and plotted in increasing order.

Figure 15 Using the Fourier descriptor metric, the differences between each contour and the ideal silhouette for the noisy images are calculated, and plotted in increasing order.

Figure 16 (Left) An image in which one leg is in front of the other leg, resulting in occlusion of part of the femoral implant. (Right) The result of applying the Canny operator, using $\sigma = 2.0$, a "high" threshold of 0.025, and a low threshold of 0.01.

7 References

1. Banks, S.A. and Hodge, W.A. 1996. Accurate measurement of three-dimensional knee replacement kinematics using single-plane fluoroscopy. IEEE Trans Biomed Eng, **43**(6): p. 638-49.

2. Banks, S.A., 1992. Model based 3-D kinematic estimation from 2-D perspective silhouettes: application with total knee prostheses. Massachusetts Institute of Technology: Cambridge, MA.**5**

3. Basri, R. and Weinshall, D. 1996. Distance metric between 3D models and 2D images for recognition and classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, **18**(4): p. 465-79.

4. Canny, J., 1986. A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, **PAMI-8**(6): p. 679-698.

5. Feldmar, J., Ayache, N., Betting, F. 1997. 3D-2D projective registration of free-form curves and surfaces. Computer Vision and Image Understanding, **65**(3): p. 403-424.

6. Grimson, W.E.L., Object Recognition by Computer. 1990. Cambridge, Massachusetts: MIT Press.

7. Gueziec, A., et al., 1998. Anatomy-based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot. IEEE Trans. Medical Imaging, **17**(5): p. 715-728.

8. Hellier, P., et al., 2001. Hierarchical estimation of a dense deformation field for 3-D robust registration. IEEE Trans. Medical Imaging, **20**(5): p. 388-402.

9. Hoff, W.A., et al., 1998. Three-dimensional determination of femoral-tibial contact positions under "in-vivo" conditions using fluoroscopy. Clinical Biomechanics, **13**(7): p. 455-472.

10. Kass, M., Witkin, A., Terzopoulos, D. 1987. Snakes: active contour models. Int. J. Computer Vision, **1**(4): p. 321-331.

11. Lavallee, S. and Szeliski, R. 1995. Recovering the position and orientation of free-form objects from image contours using 3D distance maps. IEEE Trans. Pattern Analysis and Machine Intelligence, **17**(4): p. 378-390.

12. Lemieux, L., et al., 1994.,A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. Med. Phys., **21**(11): p. 1749-1760.

13. Mahfouz, M., et al., 2001. A Robust Method for Registration of Three-Dimensional Knee Implant Models to Two-Dimensional Fluoroscopy Images. submitted to IEEE Trans. Medical Imaging,.

14. Mahfouz, M., et al. 2002. Verification Of Three-Dimensional Joint Kinematics Determined Using Fluoroscopy: An Error Analysis. in 48th Annual Meeting Orthopaedic Research Society - Poster. Dallas, TX.

15. Matsopoulos, G., et al., 1999. Automatic retinal image registration scheme using global optimization techniques. IEEE Trans. Information Technology in Biomedicine, **3**(1): p. 47-60.

16. Penney, G., et al., 1998. A comparison of similarity measures for use in 2-D-3-D medical image registration. IEEE Trans. on Medical Imaging, **17**(4): p. 586-595.

17. Rouet, J., Jacq, J., Roux, C., 2000. Genetic algorithms for a robust 3-D MR-CT registration. IEEE Trans. Information Technology in Biomedicine, **4**(2): p. 126-136.

18. Sarojak, M.E., et al. 1999. An interactive system for kinematic analysis of artificial joint implants. in 36th Rocky Mountain Bioengineering Symposium. Copper Mountain, Colorado.

19. Shapiro, L. and Stockman, G. 2001. Computer Vision. Upper Saddle River, New Jersey: Prentice-Hall.

20. Weese, J., et al., 1997. Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery. IEEE Trans. Information Technology in Biomedicine, 1 **1**(4): p. 284-293.

21. Xu, C. and Prince J.L. 1998. , Snakes, Shapes, and Gradient Vector Flow. IEEE Transactions on Image Processing,. **7**(3): p. 359-369.

22. Zuffi, S., et al., 1999. A model-based method for the reconstruction of total knee replacement kinematics. IEEE Trans Med Imaging, **18**(10): p. 981-91.

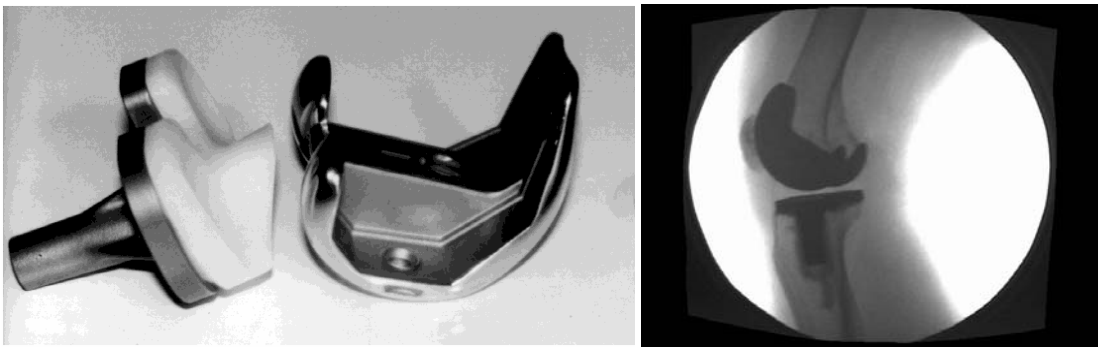

Figure 1

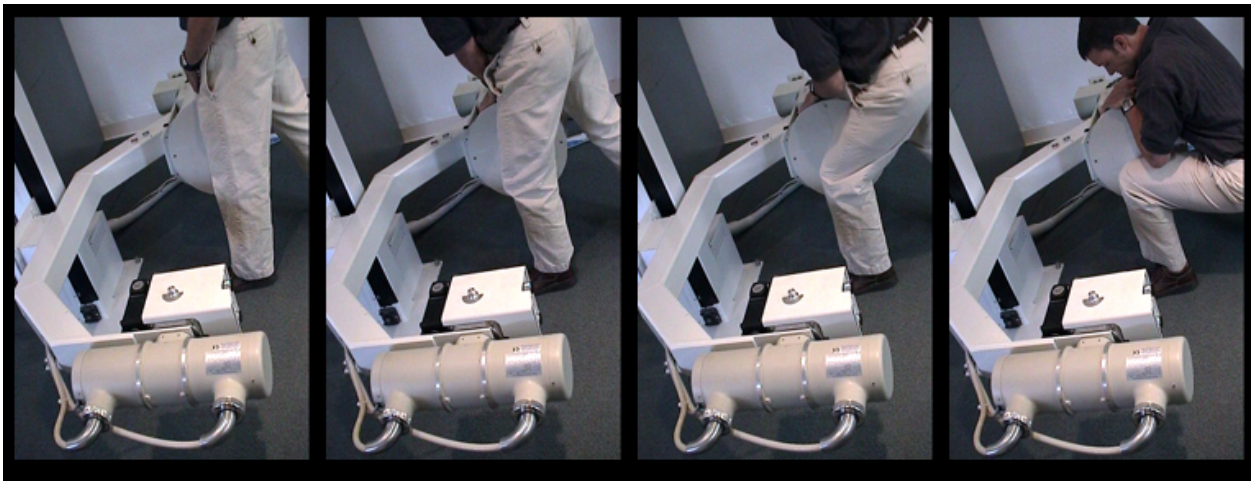

Figure 2

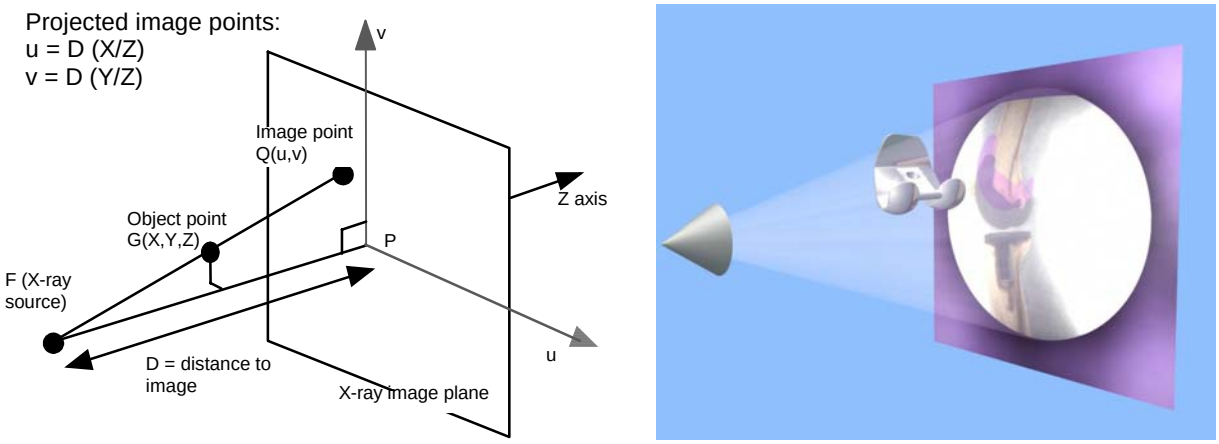
Projected image points:
u = D (X/Z)
v = D (Y/Z)
v
Image point
Q(u,v)
Z axis
Object point
G(X,Y,Z)
P
F (X-ray
source)
u
D = distance to
image
X-ray image plane

Figure 3

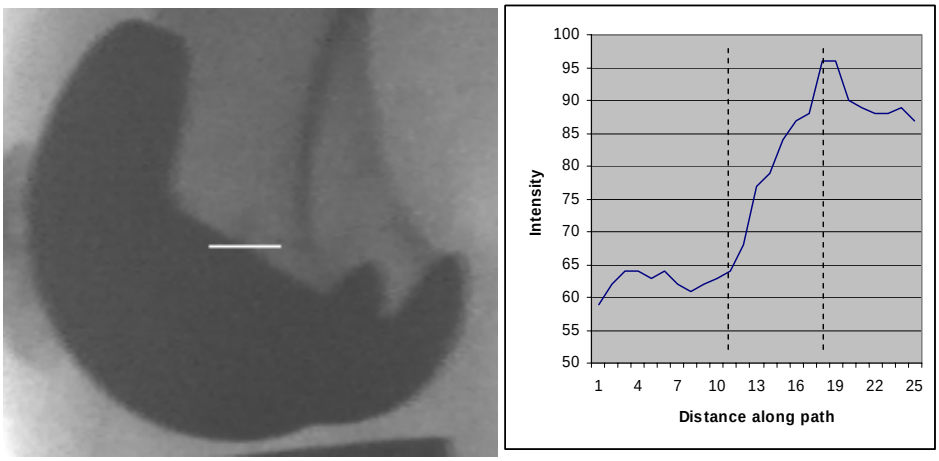
100
95
90
85
80
75
70
65
60
55
50
Intensity
1
4
7
10
13
16
19
22
25
Distance along path

Figure 4

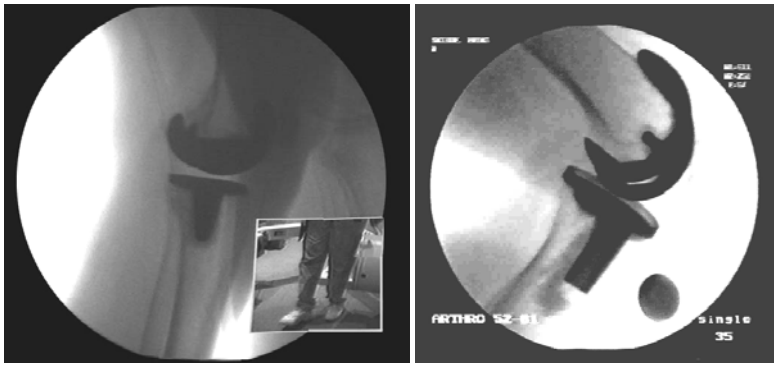

Figure 5

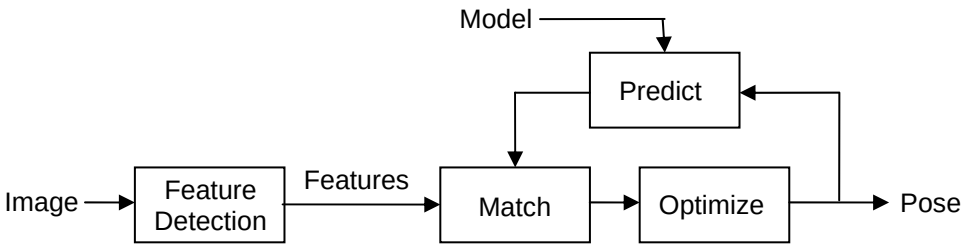
Model
Predict
Image
Feature Detection
Features
Match
Optimize
Pose

Figure 6

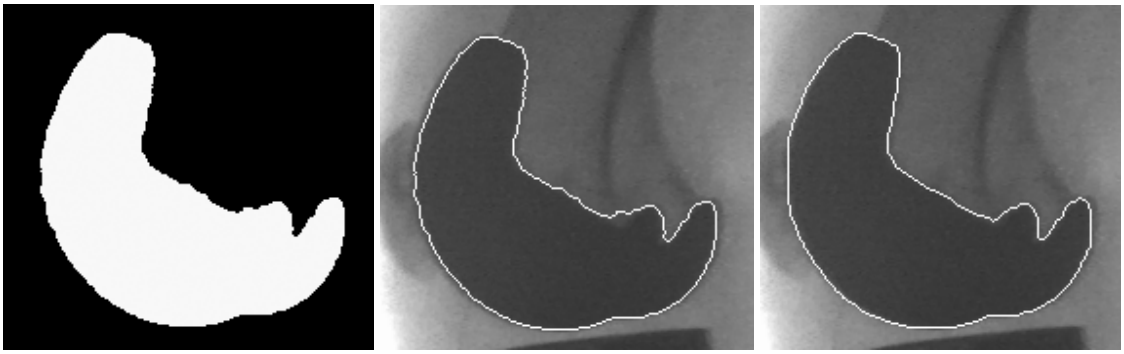

Figure 7

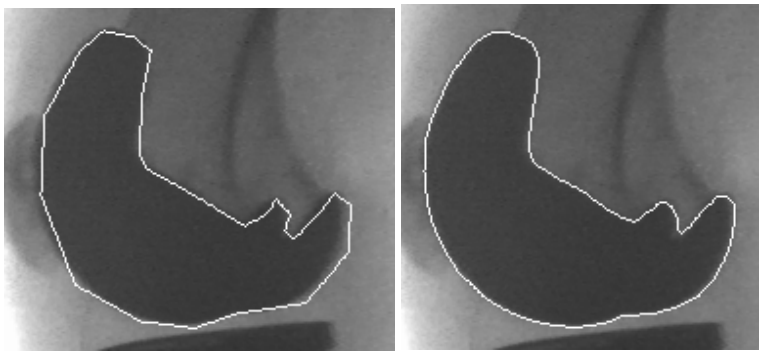

Figure 8 .

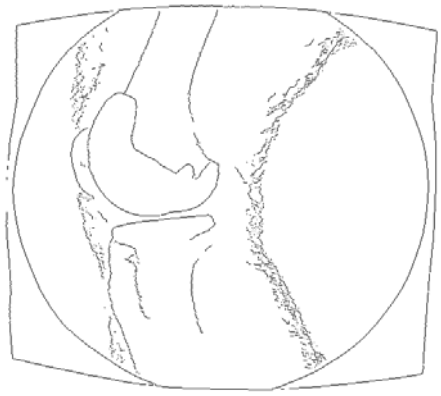

Figure 9

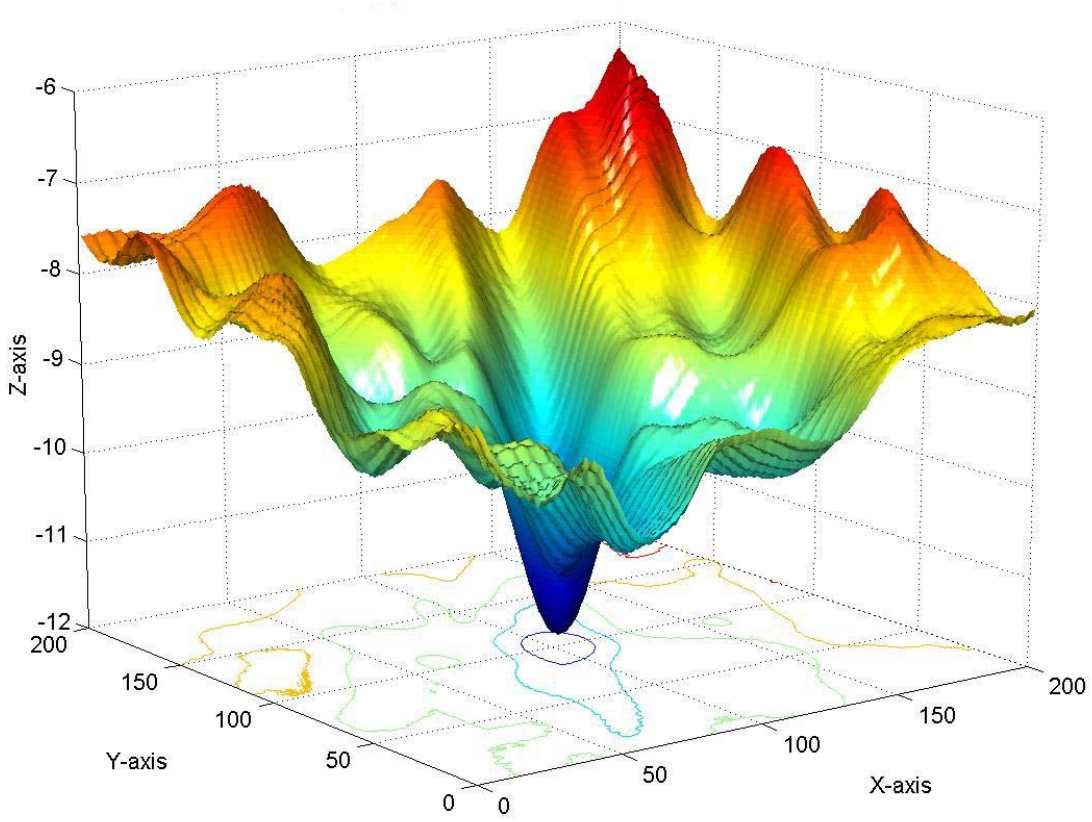

Figure 10

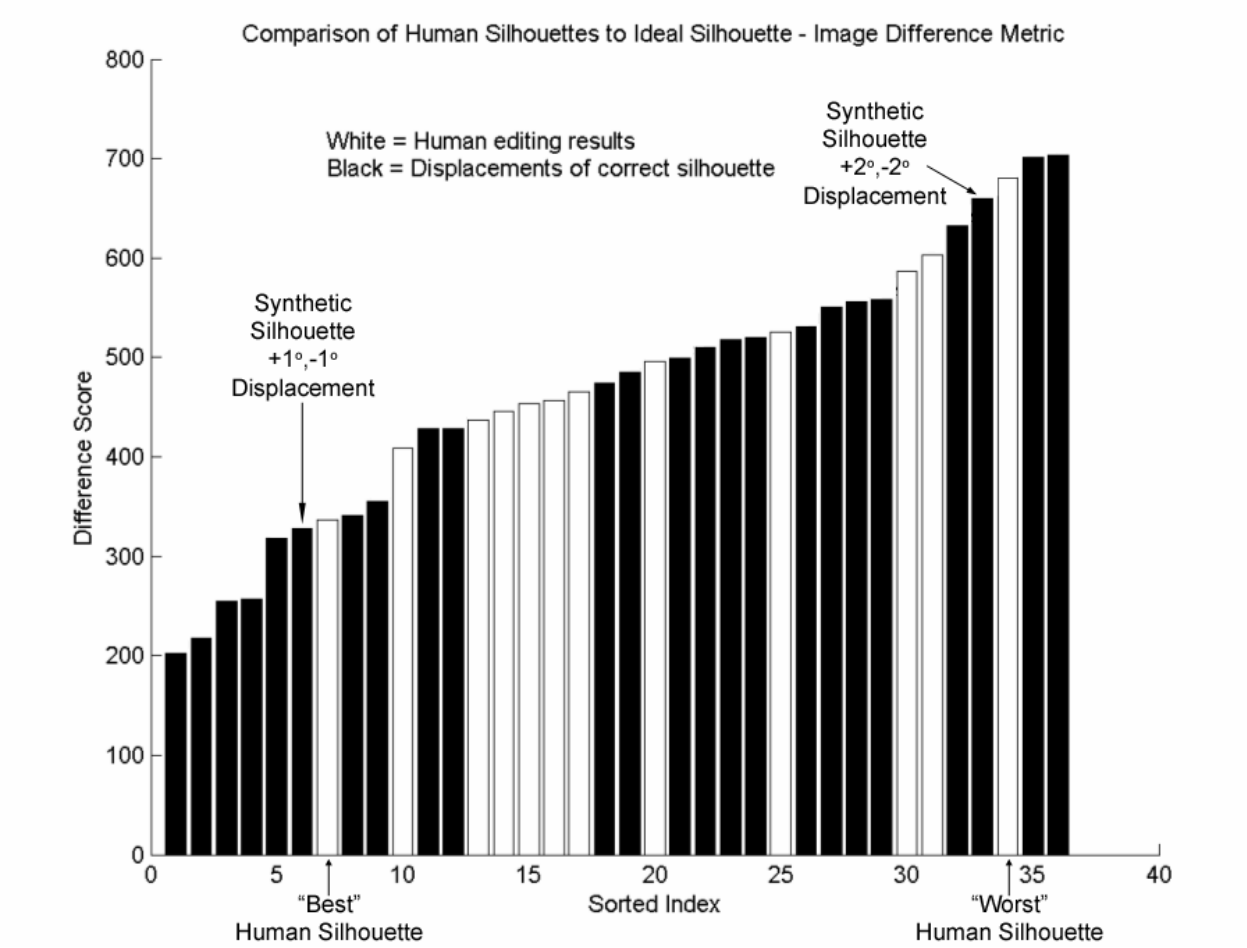
Comparison of Human Silhouettes to Ideal Silhouette - Image Difference Metric
White = Human editing results
Black = Displacements of correct silhouette
Synthetic Silhouette +2°,-2° Displacement
Synthetic Silhouette +1°,-1° Displacement
Difference Score
Sorted Index
"Best" Human Silhouette
"Worst" Human Silhouette
0
100
200
300
400
500
600
700
800
0
5
10
15
20
25
30
35
40

Figure 11

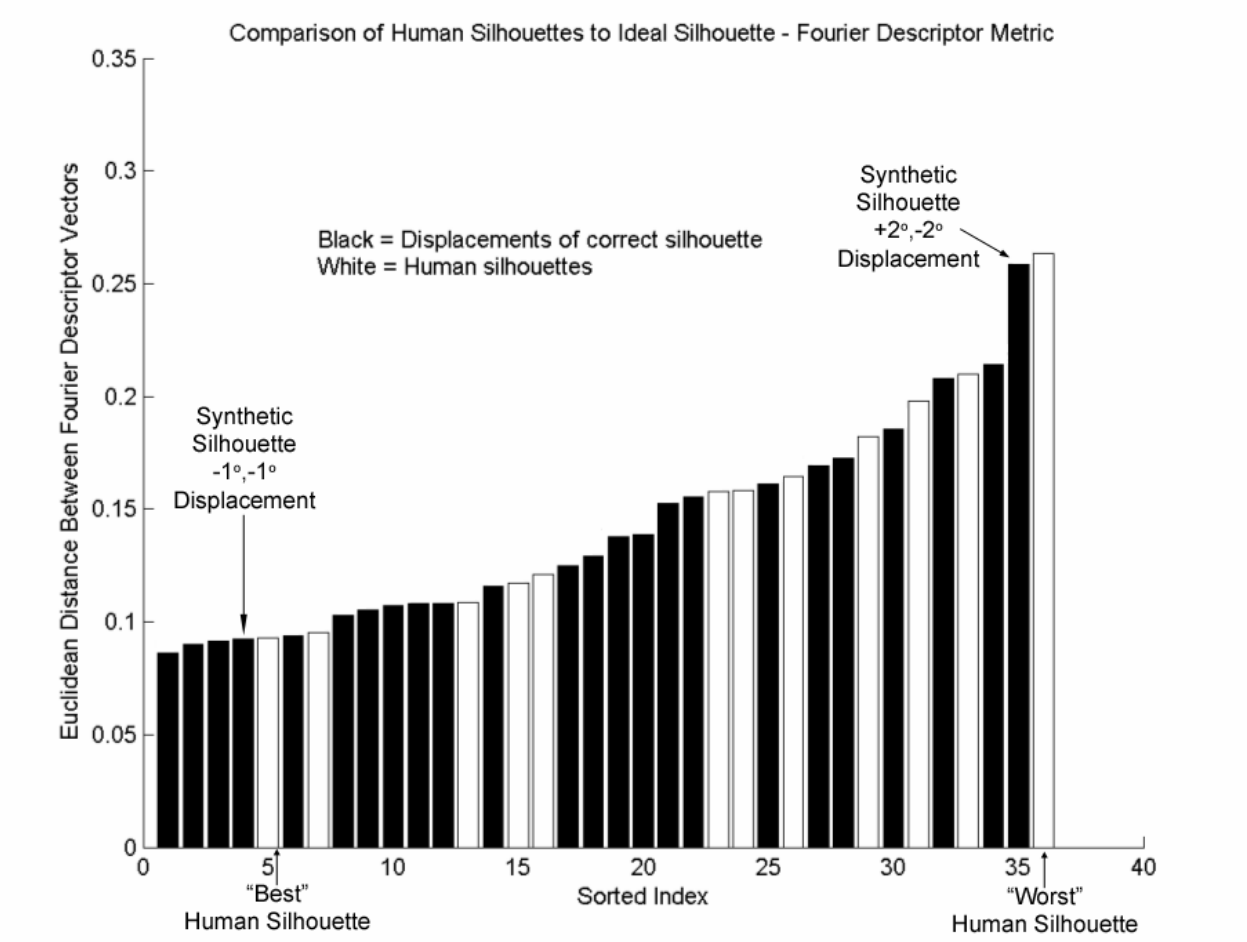
Comparison of Human Silhouettes to Ideal Silhouette - Fourier Descriptor Metric
Euclidean Distance Between Fourier Descriptor Vectors
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30
35
40
Sorted Index
Black = Displacements of correct silhouette
White = Human silhouettes
Synthetic Silhouette +2°,-2° Displacement
Synthetic Silhouette -1°,-1° Displacement
"Best" Human Silhouette
"Worst" Human Silhouette

Figure 12

Figure 13

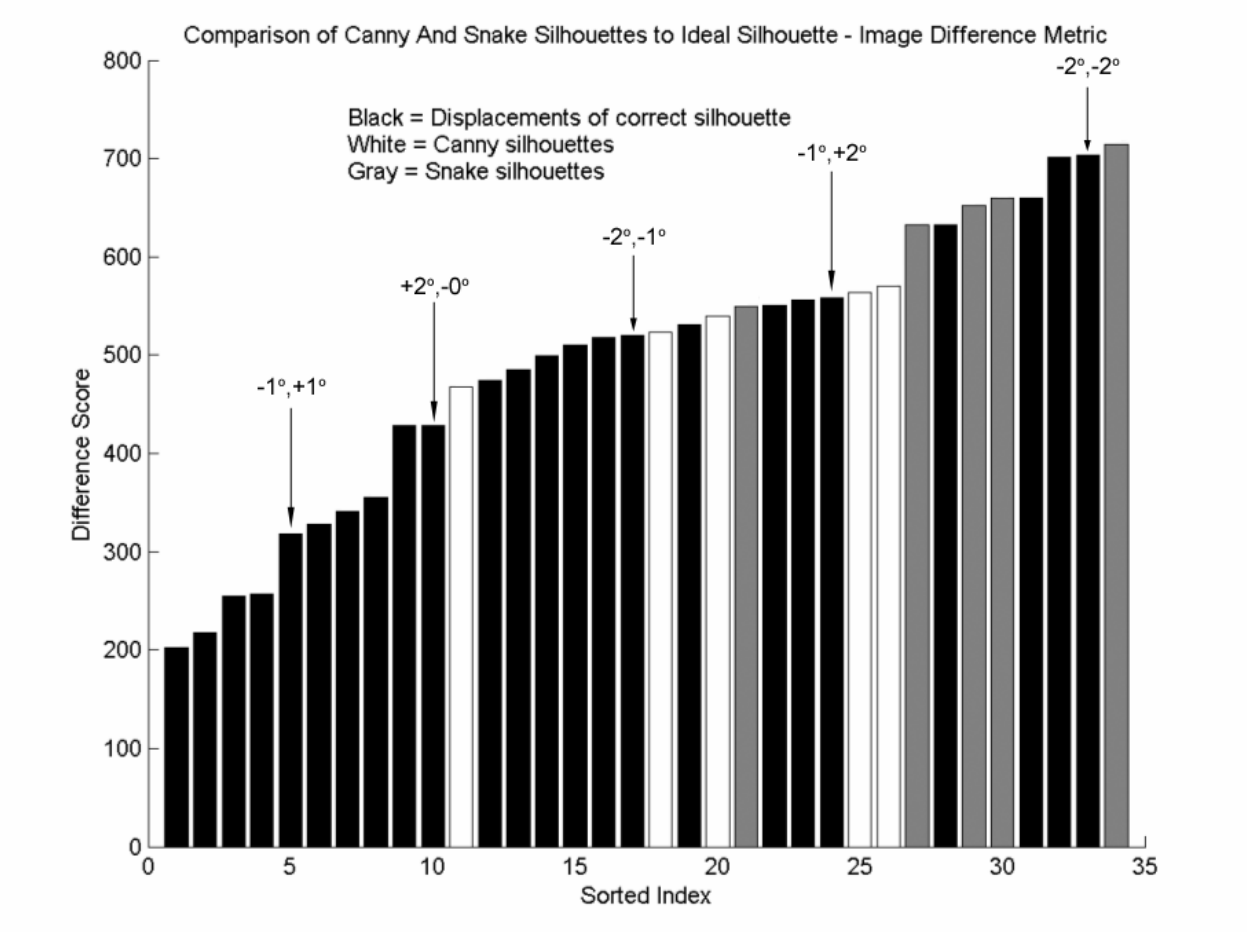
Comparison of Canny And Snake Silhouettes to Ideal Silhouette - Image Difference Metric
Black = Displacements of correct silhouette
White = Canny silhouettes
Gray = Snake silhouettes
-1°,+1°
+2°,-0°
-2°,-1°
-1°,+2°
-2°,-2°
Difference Score
Sorted Index
0
100
200
300
400
500
600
700
800
0
5
10
15
20
25
30
35

Figure 14

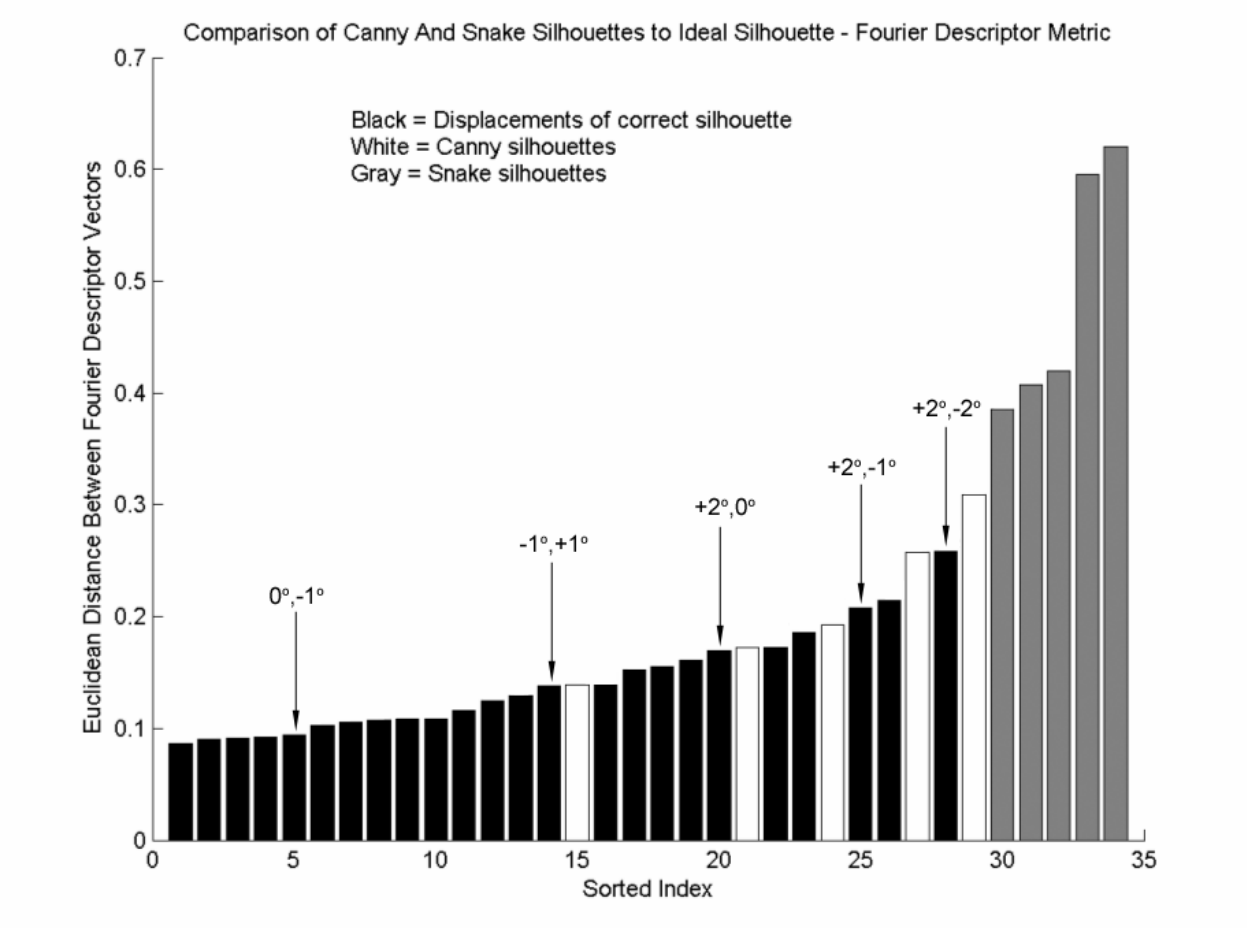
Comparison of Canny And Snake Silhouettes to Ideal Silhouette - Fourier Descriptor Metric
Black = Displacements of correct silhouette
White = Canny silhouettes
Gray = Snake silhouettes
Euclidean Distance Between Fourier Descriptor Vectors
Sorted Index
0°,-1°
-1°,+1°
+2°,0°
+2°,-1°
+2°,-2°
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
5
10
15
20
25
30
35

Figure 15

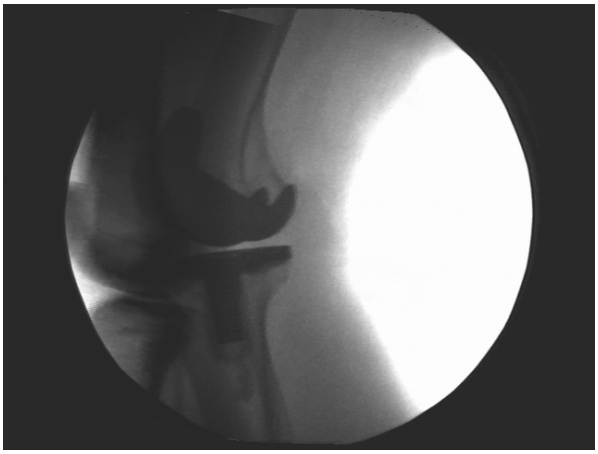

Figure 16