AN AUTOMATIC POSE RECOGNITION SYSTEM FOR KNEE IMPLANTS

By

AVANTIKA VARDHAN

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2009

To Saraswathi, the source of all learning

ACKNOWLEDGMENTS

I would like to express my sincere thanks to my advisor, Dr. Scott Banks, for his unfailing enthusiasm and support during my thesis. This research work would not have been possible without his guidance and encouragement. I am very grateful to Dr. Kenneth Slatton and Dr. Amauri Arroyo for being a part of my committee. I thank my lab members, particularly Mu Shang, John Yamakoski, and Dr. Mikashima, who have helped me in several ways. I want to acknowledge my friends Shivakeshavan Giridharan, Venkat Jayaraman, Anirudh Sudarshan and Adithya Rajan for helping me obtain computing resources. I am grateful to all the members of my extended family, who have always believed in me and encouraged me in academic pursuits. Finally, I am deeply thankful to my parents for supporting my endeavors in innumerable ways and being a great source of strength during every stage of my thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science

AN AUTOMATIC POSE RECOGNITION SYSTEM FOR KNEE IMPLANTS

By

Avantika Vardhan

May 2009

Chair: Kenneth Slatton
Major: Electrical and Computer Engineering

The problem of the registration of two dimensional fluoroscopic and X ray images with respect to a three dimensional joint implant model has been solved using a variety of techniques. In the case of knee joints, this problem is very important because accurate pose determination results in better design of implants and useful data for clinical studies. Current joint implant registration techniques require an initial model placement by a human operator. This work attempts to find an initial pose with minimal to no user interaction and hence make the pose recognition system fully automated.

This study uses a two level approach to the pose identification problem. An object recognition system has been designed which can find the Regions of Interest (ROI). The Regions of Interest in this case are the femur and tibia boundaries, which are found using a cascaded implementation of image processing and image segmentation blocks. The resulting candidate regions are then passed through an image recognition system and the final femur and tibia regions are identified. AdaBoost (adaptive boosting) performed with Haar rectangular features and a neural network classifier used on SIFT (Scale Invariant Feature Transform) generated features have been implemented for object recognition.

In the second stage of pose recognition a SIFT based classifier is used. The position and angular orientations of the image are found by matching corresponding features of the 2D image to a library of 3D models. The 3D models are generated by a computer which uses limiting biomechanical constraints. The nearest neighbor approach is used for feature matching between the image sets. It is shown that using this method the model achieved a pose matching accuracy within the acceptable range for a correct optimized result and is relatively invariant to model type, illumination and noise factors.

CHAPTER 1
INTRODUCTION

Motivation

Total Knee Arthroplasty, commonly known as knee replacement has been frequently used to treat patients with severe knee damage. In TKA, a man made surface of metal and plastic is used to replace damaged bone and cartilage. In the period between 1979 and 2002, the number of TKA procedures performed almost increased by an eightfold factor [1]. With this rapid increase, the design of better joint implants is of great importance. The determination of the exact pose of knee implants during load bearing activities and motion gives us valuable information which could help in improving current design technology and in overcoming problems such as the premature failure of implants. In addition it could assist in clinical studies, analysis of effectiveness of new treatments, development of better rehabilitation strategies, and the development of new surgical techniques [2],[3].

In order to study the in-vivo kinematics of joint function, the current techniques used for joint implant matching depend on a manually generated initial pose. In this case, the accuracy of the final result depends to a large extent on the user's input pose. Our system, which would automate finding the pose of the model, would hence eliminate errors caused by the observers and by differences between observers [4]. In addition, it would speed up the process and ensure repeatability of pose matching results.

The approach we have taken to tackle the problem of pose recognition is two fold. Initially, object recognition is performed to find the femoral and tibial regions of interest. Prior to recognition, several image processing techniques are applied to remove the high level of noise present in fluoroscopic and x- ray images. We have obtained the different candidate ROI's (Regions of Interest) using a normalized cut vertex segmentation algorithm.

These candidate regions, once passed through an image classifier are then classified as either the femur or tibia. The best matches for the femur and tibia are chosen for further processing. Two kinds of image recognition approaches are used. First, an AdaBoost (adaptive boosting) classifier is used along with a large set of rectangular Haar Features, which compute the difference in pixel intensities over various regions of the image [5].Secondly, SIFT (Scale Invariant Feature Transform) generated features are used along with a three layer Perceptron model. The SIFT key points are then used to build a machine learning algorithm using a neural network classifier.

Once the femoral and tibial regions are found, the pose determination algorithm is applied to match the exact pose along the 6 DOF (degrees of freedom). SIFT extracts the dominant key points using a DoG (Difference of Gaussian) function by finding the local peaks and dominant feature orientations [6]. The use of SIFT ensures that key points generated remain invariant to changes in scale, illumination and projective distortions. A large model image library is generated which consists of projections of the model provided by manufacturers. The images in the library have been generated for various rotations along all the X, Y and Z axes. Once a set of SIFT key points are generated for the 2 dimensional image and the library of models, matching of the key points is then performed using a nearest neighbor search [6]. Coarse to fine library matching is used, in which higher levels of pose accuracy are found with each correct image to model match.

In addition to finding the right orientation, the position information is found by performing a matching of the location of the origins of the key point descriptors. Once the translations and rotations are found, currently available methods for 2D to 3D joint image registration can be applied to find the accurate pose.

Software for the final optimization steps are already available in an open source format known as Joint track. This implementation uses the Mahfouz metric to compute the cost function of an image match and then performs simulated annealing to obtain an optimized result [7].

Figure 1-1. Joint track GUI: final optimized result of image registration

Related Work

Various methods have been used previously to tackle the problem of 2D to 3D image registration for knee kinematics. These methods have been developed for the registration of bone images to CT (computed tomography) and DRR (Digitally Reconstructed Radiographs).

The second application of this technique is for the registration of an implant image with CAD models or drawings of knee implants provided by the manufacturers.

Among the several approaches that exist to perform this registration, commonly used ones have been intensity based and contour based matching algorithms [4]. In the contour matching approach, the edge or exterior surface of the object was compared with the silhouette of the model. The pixel points close to the contour were assigned a weight which was inversely proportional to their distance to the contours. Contour matching was then performed by taking the cross correlation between the edge images belonging to both the model and the image .The intensity matching done consisted of taking a cross correlation between the silhouette and the radiograph and in essence finding the average pixel intensity across the area of overlap between the two. Combinations of both contour and intensity matching were used to produce minima when exact matching of model and image was obtained [4]. Although this technique is relatively robust and gives a close optimization match, it requires user supervision at several levels and manual pose estimation.

Another method for improved registration of bone images uses bony reference points located at certain anatomical landmarks of the tibia and femur for image fitting [8]. In addition, an anatomical coordinate system was found for the femur and tibia based on landmarks such as the medial and lateral points of the bone. A set of 3D models were generated and then fitted with the images by applying deformations in such a way that the surface shape and reference points match between the two. By extending this technique to joint implants, the TKA implants were also fitted with respect to the 3 D mechanical axis of the bones. The advantage of this technique is that the component pose was found in relation to the entire lower extremity of the bone and was not affected by changes in x-ray source position or exposure parameters.

In template matching procedures, the implant model was matched with a library of generated model templates over a wide range of rotation and translation parameters. The use of

Fourier descriptors found by the application of the DFT (Discrete Fourier Transform) operation was used to summarize contour information of a joint image during in vivo kinematics [9]. These normalized Fourier descriptors provided a highly compressed form of the contour data which can be visualized as a number of periodic functions. However, the use of Fourier descriptors requires complete segmentation of the image and the presence of a completely closed contour. Therefore, in noisy images and occluded components, this method is much less efficient. Another template matching algorithm involved the use of differences in template areas as opposed to contours [10]. Although this overcomes the limitations of generating closed contours, it nevertheless requires a complete segmentation of components.

Apart from template based techniques, the alignment of the component with a projection from a camera using a distance metric as a cost function which was optimized using a simple nonlinear minimization approach [11]. The fitted object satisfied the tangent condition between the camera projection lines and the model surface if the Euclidean distance between each point of the projection rays and each point of the model was zero. However, the use of distance maps is considerably slow and also requires highly intensive computations. The success of this technique also depends largely on the resolution of the distance map and on obtaining perfect match between the geometry of the CAD models and the implants used.

Since the highest level of inaccuracy is present in the depth position in the plane perpendicular to the image, new techniques have focused on improving the value of this coordinate. Two orthogonal images have been used in order to ensure that the depth accuracy from a single view can be enhanced by the in-plane translation in an orthogonal view [12].

Apart from the non-invasive techniques outlined above, a number of invasive methods have also been found to perform registration for TKA. For instance, bone markers or tantalum

beads were used and point and intensity based registration was performed with these points as references [13]. Such methods involve painful surgical procedures and pose the risk of infection.

The technique which we implement overcomes many problems faced by the above mentioned algorithms. A rough segmentation which yields the ROI is required, hence eliminating the need for a very accurate segmentation step. Since the SIFT features are scale invariant, a single model library can be used with reasonable accuracy over several image datasets, hence reducing the computational time for generating a model library. The limitations which exist in the current technique include the need for very good noise removal and adaptive Thresholding steps. The results produced using this technique depends to a large extent on the effectiveness of these image pre processing steps.

The primary purpose of our registration algorithm is to replace the manual intervention of the operator in finding the initial pose of the implant. Hence although it can yield a reasonable initial guess it is beyond the scope of this algorithm to gain a very accurately optimized final pose result.

CHAPTER 2
METHODS

Overview

Our method of finding the initial guess is composed of first finding an initial boundary for the femur and tibia and later performing pose recognition on these images. The bounded regions are passed to an object recognition algorithm which then classifies them into a femoral or tibial implant. This is done to ensure that cross-matching of SIFT key points between the femur and tibia doesn't take place, especially since they have some similar components.

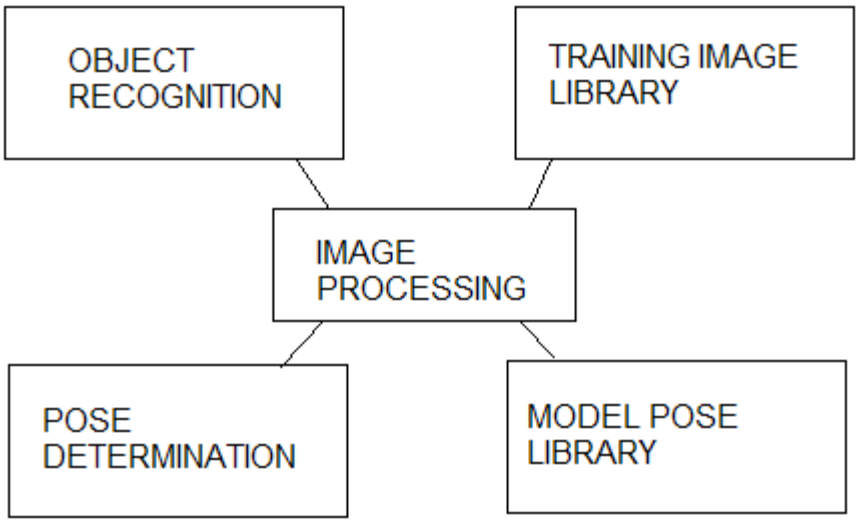


Figure 2-1. Initial pose estimation block set

In addition, finding the exact region of interest eliminates the need for a highly accurate segmentation algorithm. Once roughly segmented, the classified regions matched with a pose determination algorithm based on SIFT key points. The pose is initially matched for all in plane and out of plane rotations and later registration is performed to ensure that position is also accurately matched. As a result, the initial estimate of pose is generated which will automate the optimization procedure involved in finding the implant pose.

Object Recognition

Algorithm flow: Object Recognition consists of extracting the femoral and tibial regions before the pose determination algorithm is applied. Therefore, the effectiveness of the pose determination algorithm depends to a large extent on finding the correct boundaries of the femur and tibia. Since a rough segmentation is sufficient for the later stages, it is merely important to find the approximate bounding box of the two implants. In cases where there is an overlap between the implants, a few seconds of manual intervention would ensure that a line is drawn dividing the image into the femur and tibia regions. Prior to applying the object recognition algorithm, the images used are subjected to several preprocessing steps. In particular, the challenges posed to object recognition such as illumination changes, occlusions and noise must be effectively removed.

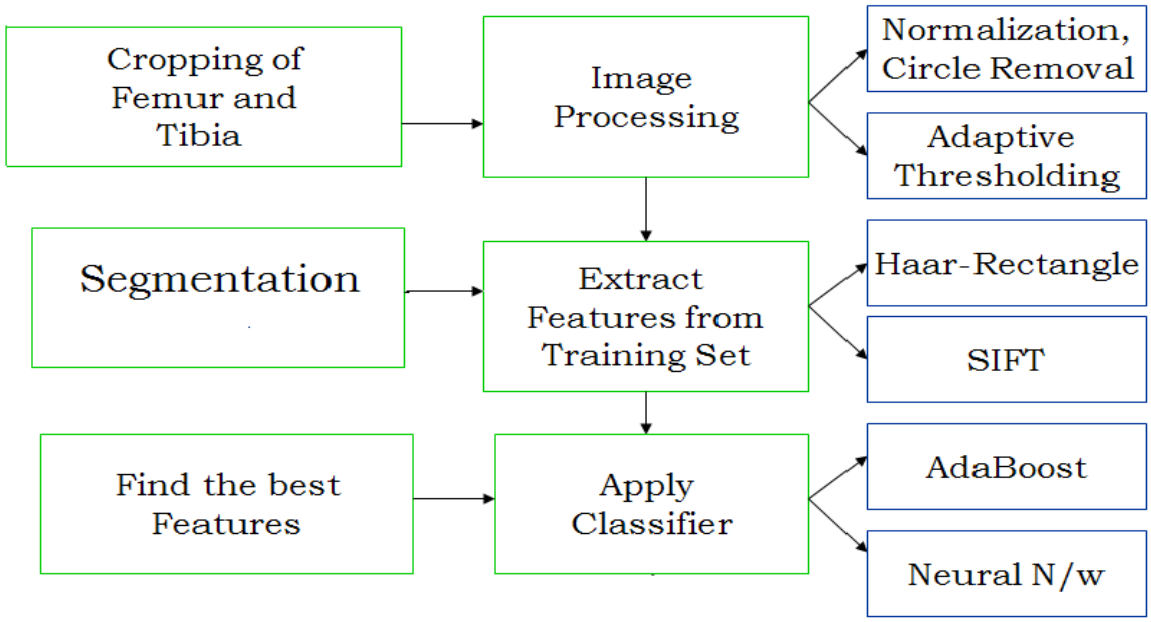


Figure 2-2. The algorithm flow of the object recognition code

Image Preprocessing: Before applying any object recognition technique, first the images have to be suitably processed to form the training and testing datasets. Initially a Hough transform with a curvature range is applied to detect the circular shape which frames

fluoroscopic images. A simple normalization procedure is performed to create a uniform range and remove variability in illumination. The adaptive thresholding algorithm has then been applied to all images in order to remove noise and extract only the essential components. As opposed to setting a global threshold the image is divided into several windows and in each of these an adaptive threshold value is found [14], [15]. The use of an adaptive threshold is especially effective since the intensity values of the implants are not constant or within a fixed range but vary in their pixel values over different image regions. Once this Thresholding is performed, the surrounding bone regions and soft tissues are simultaneously removed and only the required components are retained.

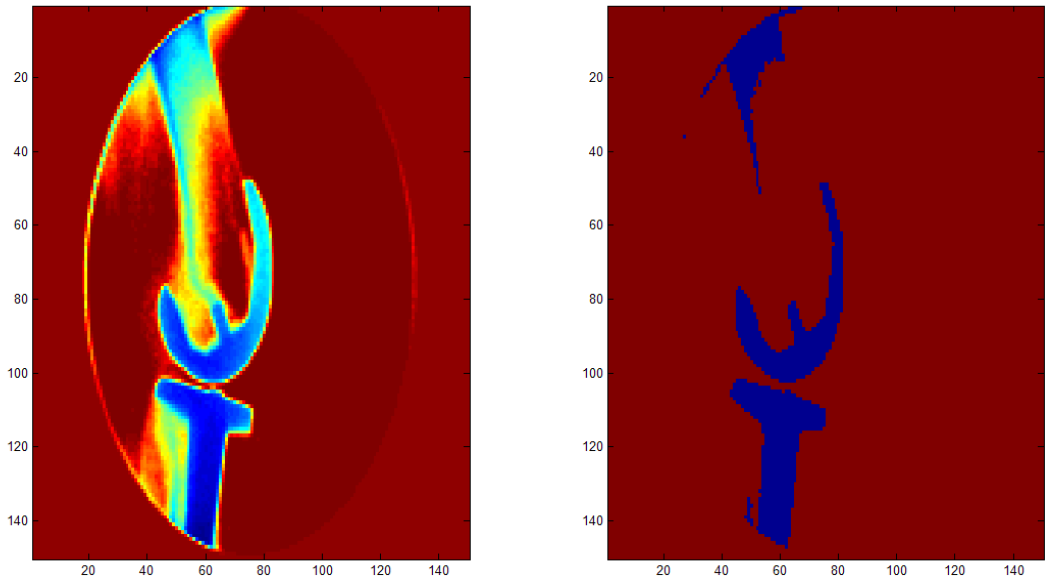


Figure 2-3. An example of the adaptive thresholding technique

Normalized-cut vertex Segmentation: Once image preprocessing has been performed, a sub-image or ROI (region of interest) which contains only the relevant components which are necessary is extracted. On this sub-image, N-cut segmentation is applied to find the approximate boundaries of the femur and tibia. In cases where a clear match is not found by the segmentation technique, a Hough transform or a bounding line can be used to divide the image into the femoral

and tibial regions. The N-cut segmentation is a technique which treats the problem of segmentation as one similar to clustering. Each pixel in the image is considered a node and pixel pairs are considered edges which are assigned weights. The value of the edge weights correspond to the degree of similarity between the edges. In completely disconnected edges, this similarity measure has a value of zero.

$$\text{Cut (A, B)} = \sum_{i \in A,\, j \in B} wij \qquad (2\text{-}1)$$

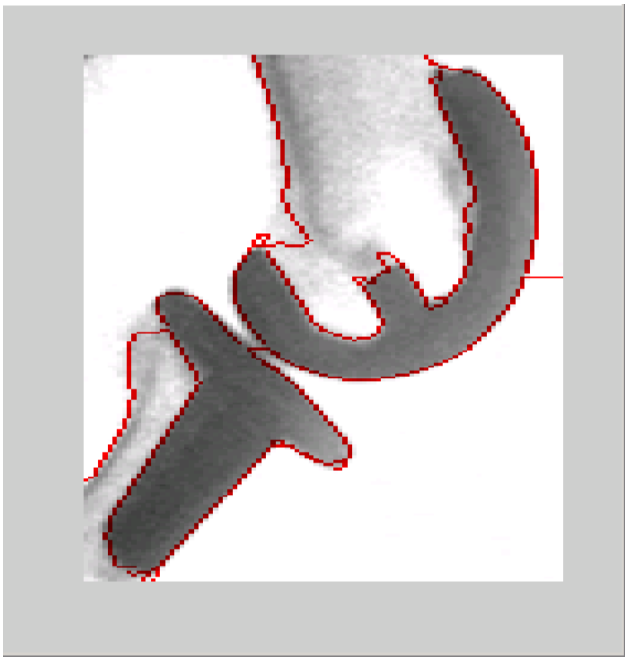


Figure 2-4. An example of the N-Cut segmentation applied on a sub-image

Each possible cut hence divides the image into two segments A and B (Equation 2-1) and results in a summation of weights of each edge between nodes in the two regions. Therefore, an optimal partition minimizes the value of the Cut metric in Equation 2-1. Each partition is then checked for stability and recursive segmentation is performed if necessary [16]. The number of cuts and the number of segments in an image can be specified by the user. As shown in Figure 2-4, the femur and tibia regions segmented within the sub-image show a high degree of accuracy. The main advantages of using the N-cut segmentation method are that it is highly texture

sensitive, which is a useful property while dealing with distinguishing implant regions from bone, soft tissue and surrounding noise. A bounding box surrounding the femoral and tibial regions is now extracted and these regions are sent to a classifier.

AdaBoost for object recognition: AdaBoost (Adaptive Boosting) is an algorithm which is popularly applied in the fields of face and object recognition. The underlying principle of AdaBoost is the process of combining several weak classifiers with different weights to result in a strong classifier [18]. The features extracted from AdaBoost are usually in the order of hundreds of thousands which are then passed through a Boosting cascade which extracts the best features.

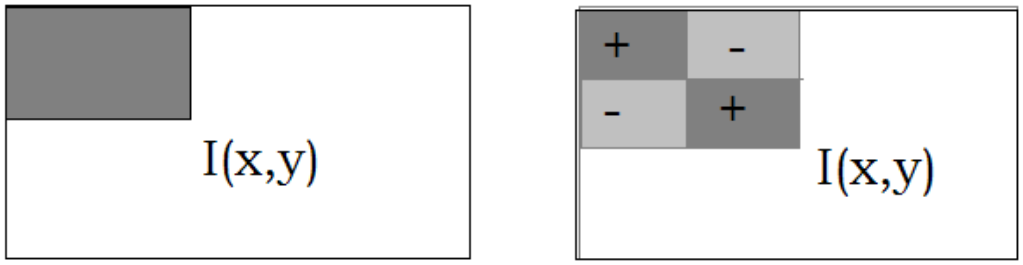
Figure 2-5. Edge and line rectangular Haar features

The AdaBoost program in this case uses rectangular edge and line Haar features which can be computed at very high speeds. These features, which add and subtract combinations of rectangular regions, exploit the inherent symmetry present in the implant components. Haar features are computed over all possible varying window sizes and shapes to form the complete feature set.

A method which is used to speed up computing times is the Integral Image technique. Every pixel of the integral image consists of the summation of all the previous pixels up to that point.

$$I(x, y) = \sum_{i=1}^{x} \sum_{j=1}^{y} \mathrm{Im}(i, j) \tag{2-2}$$

Therefore, all the Haar features which need to be computed can be easily generated using the integral image as illustrated in Figure 2-6.



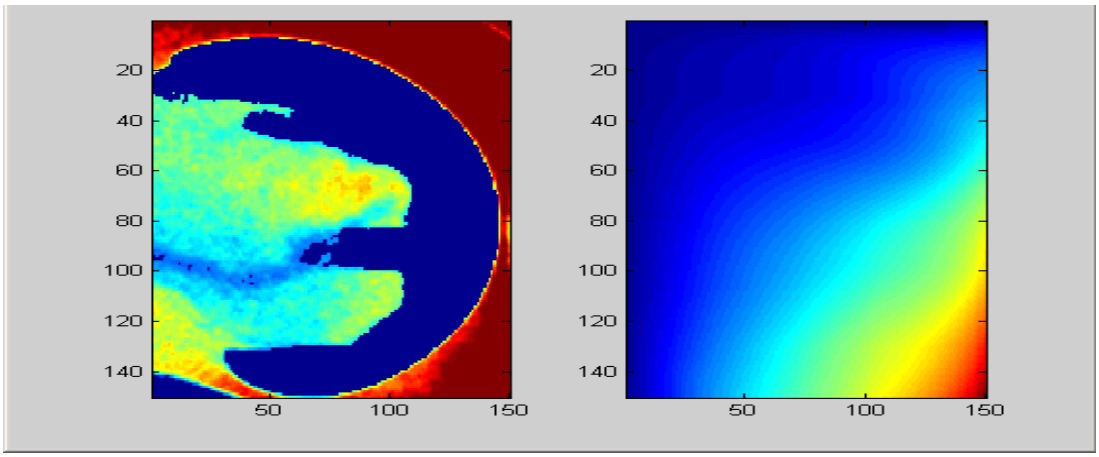
A B

Figure 2-6. Integral images and Haar features. A) Integral image. B) Haar feature generation

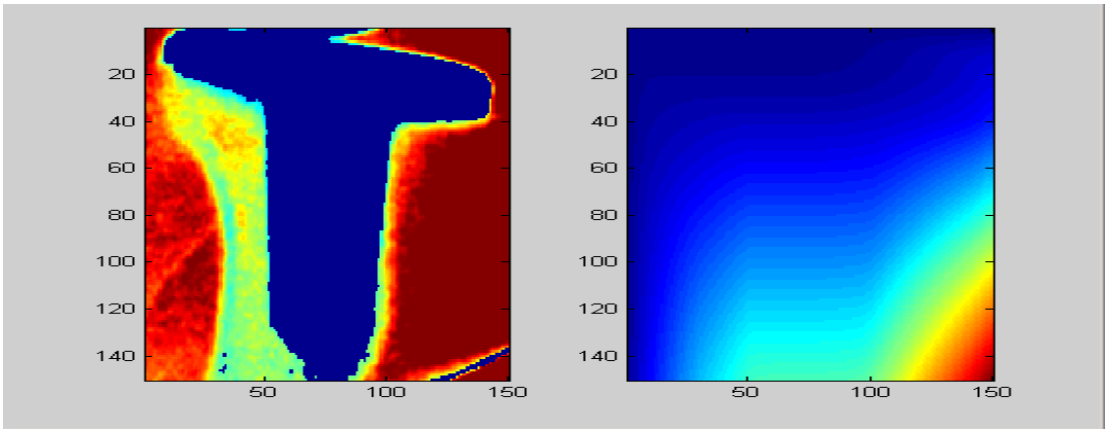
$$F(x) = \sum_{t=1}^{T} \alpha_t h_t \qquad (2\text{-}3)$$

Once features are generated, the AdaBoost classifier undergoes several cycles with each cycle corresponding to a weak classifier chosen for building the final strong classifier, as shown by Equation 2-3. Although in the initial stage of boosting each sample is given a uniform weight, with each iteration of AdaBoost the misclassified samples assume greater weight before the subsequent cycle [19]. The advantages of AdaBoost lie in extremely fast computation, a tendency to work even with highly skewed or biased datasets and robustness with noise. The weak classifier employed is a simple model of a one-layer perceptron consisting of the parameters parity, error and threshold value.

Some other positive attributes of AdaBoost include the facts that the result converges to the log likelihood ratio, it can be generalized and the chances of over fitting are low. Its main drawbacks are that it requires a large training set and uses a suboptimal Greedy learning algorithm for convergence. AdaBoost has also been used recently in multi-view face recognition implementations by using training images which have undergone several rotations [18].



A B



C D

Figure 2-7. Integral images A) Image of Femur, B) Integral Image of Femur, C) Image of Tibia and D) Integral Image of Tibia.

A diverse database covering all possibilities is used as a training dataset in order to make AdaBoost rotation invariant and to accommodate for the changes in view point and model rotation. Since the rotations of the femur and tibia are not over an extremely large range detection is easily possible despite varying angles.

AdaBoost cycle implementation: The following steps illustrate an implementation of AdaBoost.

Given that,

$$(x_1, y_1), \ldots, (x_m, y_m); x_i \in \chi, y_i \in \{-1. + 1\} \quad (2\text{-}4)$$

Where, $(x_1, y_1), \ldots, (x_m, y_m)$ in Equation 2-4 refers to images $x_1, \ldots, x_m$ and $y_1, \ldots, y_m$ refers to the image labels (femur or tibia)

The weights are first initialized as,

$$D_t(i) = 1/m \quad (2\text{-}5)$$

In Equation 2-5 'm' stands for the total number of images in the dataset.

For t=1,., T (where T stands for the total number of Haar Features)

$$h_t = \arg\min \varepsilon_j = \sum_{i=1}^{m} D_t(i)[y_i \neq h(x_i)] \quad (2\text{-}6)$$

If the error, $\varepsilon_t > 1/2$ then the loop is discontinued

Set the weight factor $\alpha_t = 1/2 \log\left(\frac{1-\varepsilon t}{\varepsilon t}\right)$ (2-7)

The weights for each of the 'i' samples is then updated as given by Equation 2-8

$$D_{t+1}(i) = \frac{D_t(i)\exp(-\alpha_t y_t h_t(x_i)}{Z_t} \quad (2\text{-}8)$$

In Equation 2-8, Zt refers to a normalization factor.

Finally, the strong classifier is given by the weighted sum of weak classifiers

$$H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right) \quad (2\text{-}9)$$

SIFT classifier for object recognition: In cases where the dataset might not be large enough or there are very wide variations in rotation and viewpoints of objects, a SIFT classifier can be used instead of AdaBoost to ensure that object recognition is performed. The SIFT algorithm generates key points which remain invariant to several external factors such as illumination, rotation and camera viewpoints while staying partially invariant to affine transformations and the presence of occlusions and noise. Once the SIFT key points were generated, the magnitude and orientation of these key points were used as a feature set to

implement a three layer perceptron based neural network classifier. This text will delve into SIFT features in greater detail in the next section and illustrate exactly how pattern matching occurs. The performance of SIFT and AdaBoost was found to be comparable for the dataset used without any major differences in parameters such as computational time or error rate.

Pose Determination

The central problem which our implementation attempts to solve is that of arriving at a good approximate initial pose to replace the user's manually generated guess. Therefore, once the two rectangular regions representing the femur and tibia have been extracted, pose matching is performed between these regions and a generated library of the projections of the CAD model. While currently popular techniques employ contour matching and edge detection [4], these result in several local minima and are error prone when the image illumination or object model changes. Therefore, we have used the SIFT key points as the primary parameter for matching between the model pose library and the implant of interest. The model library consists of projections of the manufacturer's model under certain biomechanical constraints. Since the most important parameters for 2D to 3D registration are the out of plane rotations, the library consists of variations over the entire range of these out of plane rotations at a resolution of 5 degrees. The model library further computes the range of in plane rotation at a resolution of 10 degrees. Once the matching of the SIFT key points yields an estimated orientation, conventional 2 D image registration techniques can be applied to arrive at a complete initial pose which includes all 6 DOF's (Degrees of Freedom).

SIFT classifier for pose recognition: SIFT Features are based on the generation of key points from images such that they are invariant to several external factors such as illumination, viewpoint, noise and occlusions [6]. Since we seek to replace the function of a human operator with an automated algorithm, it is of interest to note that SIFT features have been shown to

correspond to the biological basis of vision in studies of the primate visual cortex [20]. According to recent studies the use of SIFT Features for correspondence matching has been evaluated as superior to almost all other currently existing techniques [21].

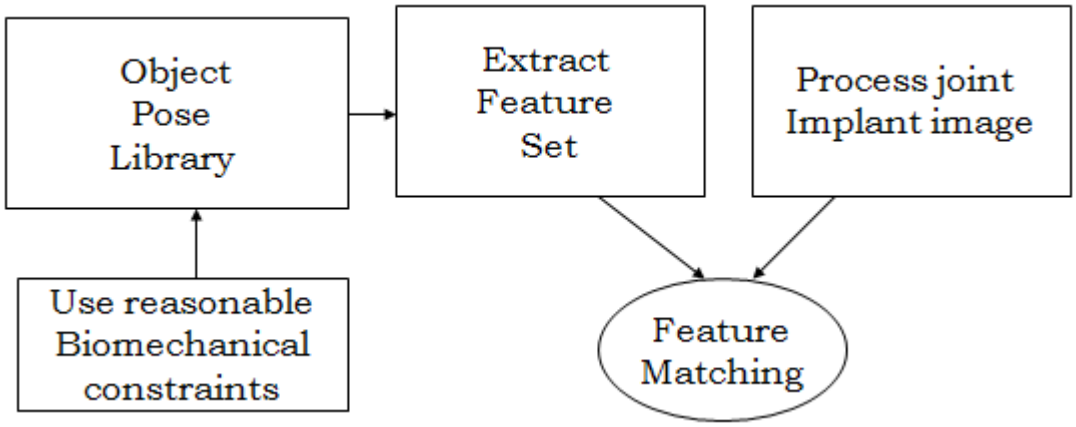


Figure 2-8. SIFT Feature matching process

SIFT key point generation: SIFT key points are generated by finding the local peaks for DoG (Difference of Gaussian) images. The steps involved in finding SIFT key points are:

1. Scale – Space Extrema Generation
2. Key point Localization
3. Orientation Assignment
4. Key point Descriptor

Scale Space Extrema are used to find the locations and scales which remain constant for several views of the given object. This is found by convolving the image with a scale space operator which is generally a Gaussian, as shown in Equation 2-10.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{2-10}$$

As these images are scaled, for instance by a factor ‘k’, the Difference of Gaussians between different scales of the image are computed as given in Equation 2-11

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{2-11}$$

The local maxima and minima of the function D is found by generally comparing it with 8 nearest neighbors of the same scale and 9 nearest neighbors at one scale higher and lower. This procedure ensures that the features are consistent even for progressively blurred images.

The next step consists of testing that each of the key points is above a certain fixed threshold to ensure that low contrast regions are not used for generating key points. This is because low contrast regions usually correspond to noisy portions of images and not the actual features. In addition, adjacent key points are also tested for principle curvature in order to reject edge points.

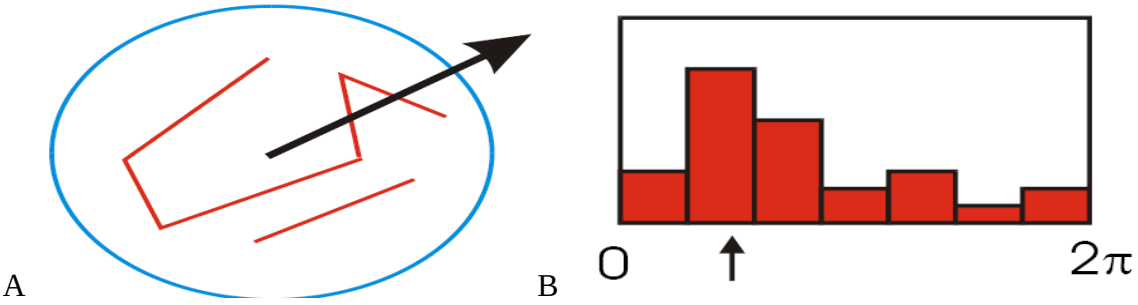


Figure 2-9. Illustration of orientation values of SIFT key points. A) Various gradients over different parts of a local patch, B) Histogram of the gradient orientations. (Source www.open.ou.nl/Touw/TouW-dagen%202008/TouW-dag%208-3-2008/touw2.ppt. Last accessed March, 2009).

After thresholding, each key point is assigned a magnitude and orientation based on the local image properties as given in Equation 2-12 and Equation 2-13. The gradient histograms of sample points are used to generate an orientation histogram for a region. As indicated in the Figure 2-10, the angle corresponding to the peak of that histogram is used to create a key point. The magnitude is given by,

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y)^2 + (L(x,y+1) - L(x,y-1))^2} \quad (2\text{-}12)$$

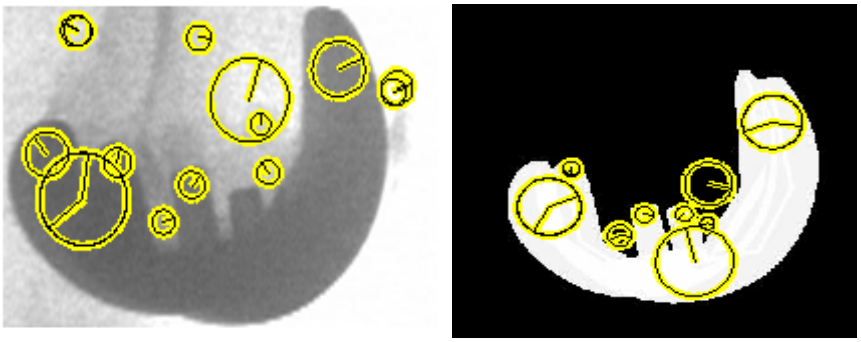
The orientation is given by,

$$\theta(x, y) = \tan^{-1}\left((L(x, y+1) - L(x, y-1)) \Big/ (L(x+1, y) - L(x-1, y)) \right) \quad (2\text{-}13)$$

SIFT key point matching and Library generation: With the application of reasonable biomechanical constraints, the projection of a CAD model of the femoral and tibial implants is obtained to form an extensive image library. SIFT key points are generated for each model pose available in the library and a close matching of the set of features for the model library with respect to the original image features is done [23]. The matching is performed by first calculating the distance function between every combination of possible descriptor pairs between the library model and image. Instead of merely using the results with the minimum value of the distance metric, we first obtain those pairs of images with the largest number of matched key points. In the second stage we find the pair corresponding to the minimum distance metric within that subset.

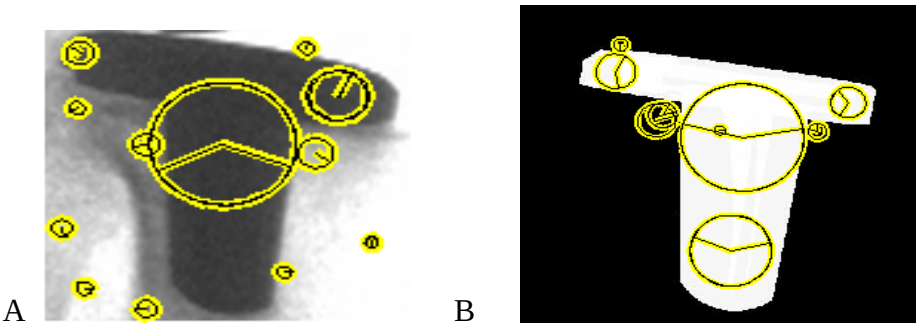
A coarse to fine registration can be performed by changing the resolution and angular spacing between adjacent library entries once an approximate match is found. However, since the pose results which have been generated are not used as a final result this is usually not necessary.

A salient feature of using SIFT is scale invariance, as a result of which several types of image can be matched with a single model library for pose matching in the first stage. As image libraries with finer accuracies are used, the matching error also decreases to an acceptable value in the range of 2 to 5 degrees [4]. In a typical case, the number of features which are generally matched range from 3 to 25. In a counter intuitive argument, it was observed that more features being matched does not always correspond to the best matching result. The best results were those which had a large number of feature matches as well as a low value for the distance between the corresponding features.



A B

Figure 2-10. Femur matching A) Femur key points B) Matched model with key points



A B

Figure 2-11. Tibia matching A) Tibia key points B) Matched model with key points

CHAPTER 3
RESULTS

Results of Object Recognition

The results for the object recognition section was found for both the AdaBoost as well as the SIFT based neural network algorithm. A training data set of 1400 images consisting of 700 femur and 700 tibia images were used to train AdaBoost. The classifier underwent 25 cycles and 25 weak classifiers were combined. The computational time for AdaBoost training on a high speed Intel Quad Core Processor was 20 minutes and the computational time for testing was less than 5 minutes. The testing data set consisted of 50 images of the femur and tibia in a random proportion. The training data set included noisy samples to test the efficiency of the program.

In the case of the SIFT algorithm, a training data set of 1600 images consisting of 800 femur and 800 tibia images were used. However, it was observed that a training data set of just 50 images was sufficient to train the neural network to achieve a detection error rate of less than 10%.

Table 3-1. Object recognition

Classifier	Training images	Test images	Error rate
AdaBoost	1400 images, 700 femur, 700 tibia	50 images	8 %
Neural N/w	1600 images 800 femur, 800 tibia	50 images	6 %

It is noted that matching is currently not accurate in the case of X-ray images in comparison with fluoroscopic images due to the presence of bone components in X-rays.

Results of Pose Determination

Pose determination results were compiled using a library of 936 femur models and 2364 tibia models. Biomechanical constraints which were applied to the CAD models such that most

model poses were created in the library. Initially the resolution for the out of plane rotations were maintained at 5 degrees and the resolution for in plane rotation was 10 degrees. Once a match of reasonable accuracy has been found, a finer match can be found by creating a new image library of 5x5x5 images.

The Figures 3-1 illustrate examples of pose matching in the femur and tibia. The number of key points generated for the femur and tibia ranged between 5 and 15. In a good match the minimum value of distance between the key points found was 200.

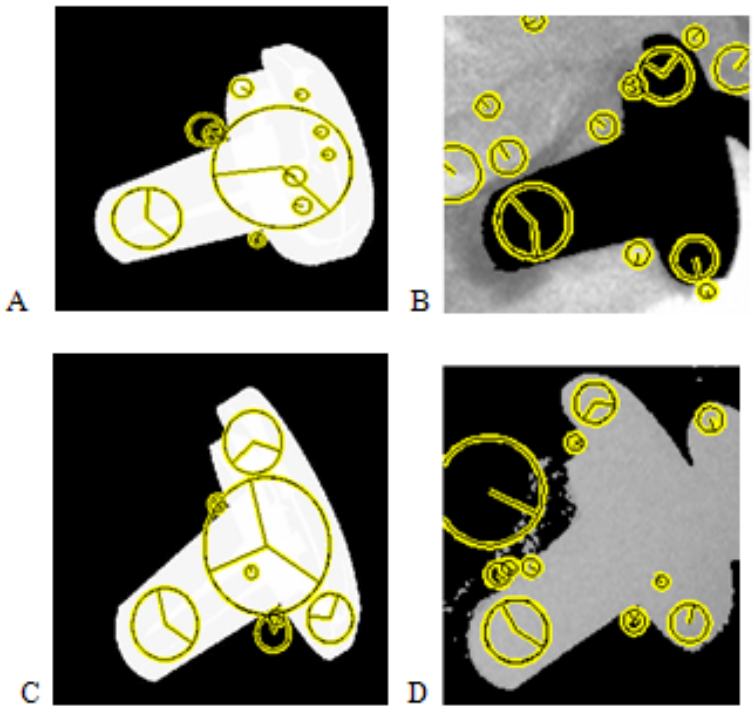


Figure 3-1. Tibia pose A), C) Tibia Models B), D) Tibia image matching

It was observed that errors were primarily in the in-plane rotation which is easy to correct in comparison with the out-of-plane rotations. Most of the images which underwent proper thresholding and noise removal showed a highly accurate pose result.

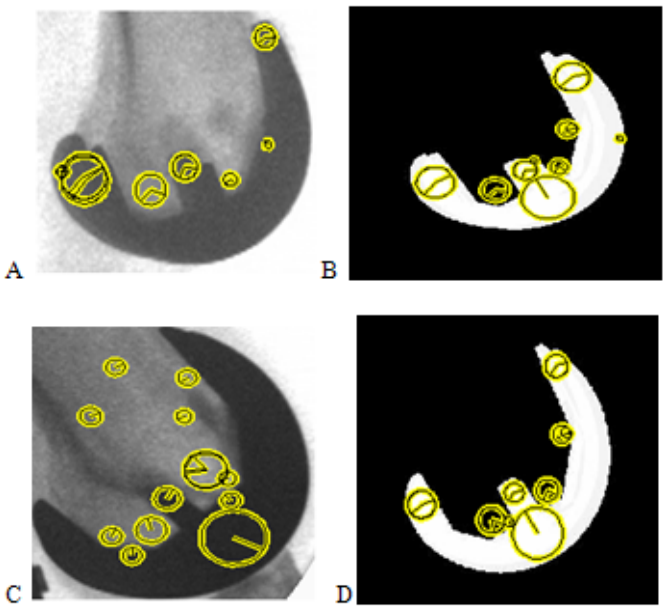


Figure 3-2. Femur pose A), C) Femur images, B), D) Matched femur models

It is noted that matching is currently not accurate in the case of X-ray images in comparison with fluoroscopic images. This could be due to the presence of several soft tissues and bone components which are imaged more clearly in the case of X-rays.

The tabulated results for the x, y and z axes orientations have been shown in Table 3-2. The results were calculated for 15 images while the subject was performing Kneel, Lunge and Stair motions. It can be noted here that since more emphasis was placed on obtaining an accurate out of plane rotation, the number of in plane rotation models were fewer and hence resulted in lower accuracy of the in plane rotation around the x axis. It is observed that the mean of the deviation of the predicted value from the actual results are within the acceptable range of 5 degrees. In the case of in plane rotations, the lower accuracy can be easily compensated by increased library size or further angular refinements.

Table 3-2. Femur Orientation Determination

Axis of rotation	Mean error	Variance of error
X	1.5	1.5714
Y	4.0909	2.2909
Z	2.4	0.7774

Table 3-3. Tibia Orientation Determination

Axis of rotation	Mean error	Variance of error
X	1.167	1.5714
Y	4.16	1.75
Z	3.12	2.18

Discussion

The problem of pose estimation has been tackled in several invasive and non invasive methods. While usually contour or intensity based techniques are commonly employed, the SIFT based technique aims to instantly extract the portions of the contour which correspond to image key points. By matching these key points which are invariant to numerous factors such as occlusions, noise, and illumination we can arrive at an accurate initial pose estimate. Once this is complete, the model in the resulting pose is then matched using currently available software to get an accurate pose.

In the case of the SIFT transform the robustness to noise is not as high as required for medical images. As a result the effectiveness of the SIFT transform depends to a large extent on the successful implementation of image processing techniques such as noise filtering and thresholding. Hence, finding a closer segmentation for the femur and tibia especially in the images in which they intersect or overlap is very important.

A common drawback of most template based matching schemes is the generation of a large image library which is computationally expensive and time consuming. However, in the case of SIFT a single image library works for matching most images which actually use different models. This can be attributed to the fact that it is scale invariant and is used to train large datasets of objects with a wide variation such as cars, vehicles and other objects.

By generating an automatic initial pose, manual errors generated by different observers are overcome. Repeatability of results is ensured and expertise on the part of the user is not needed

for successful operation. In possible future improvements, the initial pose of a sequence of images taken from a video can be computed by using the value of the final pose of the previous image, thus reducing the number of computations. In the future, this algorithm could also be used for the registration of bones and implants in other regions of the body such as the shoulder.

LIST OF REFERENCES

[1] Centers for Disease Control and Prevention. National Center for Health Statistics. , Data Warehouse on Trends in Health and Aging. National Hospital Discharge Survey.Health Data Interactive. www.cdc.gov/nchs/hdi.htm. [Date of access March 2009].

[2] Banks, S. A., Harman, M.K., Bellemans, J., Hodge, W.A., 2003, "Making sense of knee Arthroplasty kinematics: news you can use," Am. J. Bone Joint Surg., 85(A), pp. 64 - 72.

[3] Yamazaki, T., Watanabe, T., Nakajima, Y., Sugamoto, K., Tomita, T., Yoshikawa, H., Tamura, S., 2004, "Improvement of depth position in 2-D/3-D registration of knee implants using single-plane fluoroscopy. IEEE Trans on Medical Imaging," 23(5), pp. 602-612.

[4] Mahfouz, M.R., Hoff, W.A., Komistek, R.D., Dennis, D.A., 2003, "A robust method for registration of three-dimensional knee implant models to two-dimensional fluoroscopy images," IEEE Trans on Med Imaging, 22 (12), pp. 1561-1574.

[5] Viola, P., Jones, M., 2001, "Rapid object detection using a boosted cascade of simple features," IEEE Comp Soc Conf on Computer Vision and Pattern Recognition 2001, vol.1. IEEE, pp. 511-518.

[6] Lowe, D.G., 1999, "Object recognition from local scale-invariant features," Proceedings of the Seventh IEEE International Conference on Computer Vision 1999, vol.2. IEEE, pp.1150-1157.

[7] Shang, M., 2007, "Jointtrack: An Open-Source, Easily Expandable Program For Skeletal Kinematic Measurement Using Model-Image Registration," Master's Thesis, University of Florida Dept of Mechanical and Aerospace Eng, pp 4-24.

[8] Sato, T., Koga, Y., Omori, G., 2004, "Three-dimensional lower extremity alignment assessment system: Application to evaluation of component position after total knee Arthroplasty," The J. of Arthroplasty, 19 (5), pp. 620-628.

[9] Banks, S.A., Hodge, W.A., 1996, " Accurate measurement of three-dimensional knee replacement kinematics using single-plane fluoroscopy," IEEE Trans. Biomed. Eng. , 43 (6), pp.638-649.

[10] Komistek, R. D., Dennis, D.A., Hoff, W.A., Gabriel, S.M., and Walker, S.A., 2002. "Three-dimensional determination of femoral-tibial contact positions under in vivo conditions using fluoroscopy," The J. Arthroplasty, 17 (2), pp. 209-216.

[11] Zuffi, S., Leardini, A., Catani, F., Fantozzi, S., and Cappello, A., 1999. "A Model-Based Method for the Reconstruction of Total Knee Replacement Kinematics," IEEE Trans. On Med.Imaging, 18 (10), pp. 981-991.

[12] Li, G., Wuerz, T. H., and DeFrate, L. E., 2004, "Feasibility of Using Orthogonal Fluoroscopic Images to Measure In Vivo Joint Kinematics. J. Biomech. Engg," 126(2), pp. 314- 318.

[13] Tang, T. S. Y., MacIntyre, N. J., Gill, H. S., Fellows, R. A., Hill, N. A., Wilson, D. R., and Ellis, R. E., 2004, "Accurate assessment of patellar tracking using fiducial and intensity-based fluoroscopic techniques. In: Medical Image Analysis, Medical Image Computing and Computer-Assisted Intervention – MICCAI," 8(3), pp. 343-351

[14] Jain, A., 1986, Fundamentals of digital image processing. Prentice-Hall, pp. 408.

[15] Chow, C.K., and Kaneko, T., 1972, **"**Automatic Boundary Detection of the Left Ventricle from Cineangiograms," Comp. Biomed. Res., 5, pp. 388-410.

[16] Shi, J., Malik, J., 2000, "Normalized cuts and image segmentation," IEEE Trans. Pattern Analysis and Machine Intelligence, 22 (8), pp. 888-905.

[17] Viola, P., Jones, M.,2002, "Fast and robust classification using asymmetric adaboost and a detector cascade," Advances in Neural Information Processing Systems

[18] Wu, B., Ai, H., Huang, C., Lao, S., 2004, "Fast rotation invariant multi-view face detection based on real Adaboost," Proc. of Sixth IEEE Intl. Conf. on Automatic Face and Gesture Recognition. vol., no., pp. 79-84.

[19] Sochman, J., and Matas, J., Center for Machine Perception, Czech Technical University, http://cmp.felk.cvut.cz, [Date of access March 2009].

[20] Serre, T., Wolf, L., Poggio, T., 2005, "Object recognition with features inspired by visual Cortex," IEEE Comp Soc. Conf. on Comp. Vision and Pattern Rec," 2, pp.994-1000.

[21] Mikolajczyk, K., Schmid, C., 2003, "A performance evaluation of Local Descriptors. IEEE Comp. Soc. Conf. on Comp. Vision and Pattern Rec.," 2, pp. 257-264.

[22] Gordon, I. , and Lowe, D.G., 2006, "What and where: 3D object recognition with accurate pose," Springer-Verlag Toward Category-Level Object Recognition , Berlin, Germany, pp. 67-82.

[23] Vedaldi, A., and Fulkerson, B., 2008, "The VL Feat Open and Portable Computer Vision Library," has been used for SIFT Key point Generation.

BIOGRAPHICAL SKETCH

Avantika Vardhan was born in Mumbai in 1986. She lived in Chennai, India for 21 years and completed her bachelor's degree in electronics and communication engineering from Sri Ramaswamy Memorial University. She pursued her master's in the Department of Electrical and Computer Engineering at University of Florida. Her research interests lie in the field of computer vision and image processing with medical applications.