

Wavelet Compressed PCA Models for Real-Time Image Registration in Augmented Reality Applications

Christopher Cooper
College of Engineering
North Carolina State University,
Raleigh, NC, 27695

John Cooper
Department of Chemistry and Biochemistry
Old Dominion University,
Norfolk, VA, 23529

Kent Wise
SGS Inc.
The Woodlands, TX, 77381

Makarand Deo*
Department of Engineering
Norfolk State University,
Norfolk, VA, 23504

Abstract—The use of augmented reality (AR) has shown great promise in enhancing medical training and diagnostics via interactive simulations. This paper presents a novel method to perform accurate and inexpensive image registration (IR) utilizing a pre-constructed database of reference objects in conjunction with a principal component analysis (PCA) model. In addition, a wavelet compression algorithm is utilized to enhance the speed of the registration process. The proposed method is used to perform registration of a virtual 3D heart model based on tracking of an asymmetric reference object. The results indicate that the accuracy of the method is dependent upon the extent of asymmetry of the reference object which required inclusion of higher order principal components in the model. A key advantage of the presented IR technique is the absence of a restart mechanism required by the existing approaches while allowing up to six orders of magnitude compression of the modeled image space. The results demonstrate that the method is computationally inexpensive and thus suitable for real-time augmented reality implementation.

Keywords—Image Registration; Principal Component Analysis; Wavelet Compression; Augmented Reality; Image Classification

I. INTRODUCTION

The utilization of augmented reality (AR) in the medical field provides multiple opportunities to enhance the access to and effectiveness of patient-specific medical information [1][2]. Using real-time AR systems allows the overlay, manipulation, and visualization of the various types of medical images acquired by MRI and tomography procedures (e.g., tissue, charge density, blood flow, etc.)[3][4]. Hence AR-based visualization techniques have been increasingly employed in safer medical practices for better understanding and accurate diagnostics. Creating an interactive 3D virtual model containing multiple dimensions of information, which can be manipulated and visualized in concert, provides immediate opportunities for high-quality medical training. Furthermore, the advanced AR-guided medical procedures have the potential to decrease the invasiveness and increase the safety and accuracy of a surgery by enhancing a surgeon's ability to

utilize medical imagery during the operation [1][5]. The first step to achieving these goals however, is a robust and real-time registration of high resolution images [6]. This paper presents a novel registration method which is accurate and computationally inexpensive.

Image registration (IR) is the process of aligning two similar images, taken at different times or by different sensors, in order to correctly overlay an independent image [7]. IR techniques typically fall into two categories: feature-based [8] and intensity-based [9]. The former method relies on the detection and successful tracking of distinct image features, such as lines, corners, and contours, while the latter method determines a transformation using all of the image data. Each of these techniques relies on an optimization component, which determines the optimal spatial transformation, and a similarity metric, which compares the resemblance of the transformed scene image and the model image [7, 10]. Spatial transformations can be either rigid or non-rigid. Rigid transformations are composed of translation and rotation in three directions, for a total of six degrees of freedom. Non-rigid transformations account for these changes as well as those in the actual structure or anatomy of the object [11].

Initial IR techniques such as the iterative closest point (ICP) algorithm have produced incorrect transformations due to incoming image noise and prealignment errors [12]. One optimization approach to increase convergence range and avoid erroneous local optima is the use of hierarchical multi-scale, however down-sampling of images often suppresses key differences, leading to an absence of distinctive features in similar objects [11]. In response to these errors, evolutionary computation (EC) has been used to help alleviate the complex problems of image processing, most noticeably the need for a good initial estimation of the transformation. These models, included in the broader field of metaheuristics, rely on computational models of evolutionary processes to create populations of solutions [13]. One such example is the scatter search (SS) technique, which is a metaheuristic-based method

This work was supported in part by American Heart Association (AHA) Scientist Development Grant No. 12SDG11480010.

used in both feature-based and intensity-based methods [10]. Use of this technique provides noticeable advantages in the accuracy of transformations and eliminates prealignment error. Nevertheless, even IR techniques with metaheuristics rely on a restart mechanism when transformations become low quality. This is a result of the refinement process, which optimizes the previous transform in order to produce the new spatial transform for the incoming model image. Since optimizing a low-quality transformation is unlikely to produce a high-quality transformation, it is necessary to restart the algorithm and acquire a new initial transformation [10]. Recently, more advanced IR techniques have been proposed based on Speeded up Robust Features (SURF), optical flow method, and marker-free IR method [14][15]. However, these existing methods require extensive computations to achieve real-time IR which is a major concern that limits their use in real-time AR systems.

In this paper, a novel IR technique which is capable of achieving higher accuracy with substantially reduced computational time is presented. This was achieved by creating a database of compressed vectors from reference images of an object at all possible viewing angles and then constructing a corresponding principal component analysis (PCA) model prior to image registration. The proposed approach offers multiple benefits over existing methods. There is no need for an initial estimation or camera calibration, and furthermore, since the model operates independently for each incoming frame, there is no need for a restart mechanism. A systematic performance analysis of the proposed method is presented in this paper.

II. METHODS

A. Creation of Virtual Object Database

Current IR methods involve taking an existing image and transforming it in real time. With the use of high speed flash storage, it is now feasible and competitive to eliminate this transformation step, and simply recall previously generated high-resolution images. In this study, the virtual 2D images used in the registration process were generated from 3D imaging data prior to the real time registration process. A detailed 3D model of canine heart anatomy, derived from high resolution diffusion tensor magnetic resonance imaging (DTMRI) was used as the virtual object. A Virtual Object Database (VOD) was created by manipulating the orientation of the virtual heart model as a function of three orthogonal angles of rotation. Theoretically, there are an infinite number of possible orientations and hence 2D image views; however the ability to distinguish between similar orientations (2D images) decays as the change in an angle approaches zero degrees. Therefore, it is possible to represent an object within a pre-defined angular resolution with a limited number of image orientations. The degree of resolution, however, directly impacts the number of images needed in the VOD of 2D images. A total of 22,104 images were generated that fully represented all non-degenerate object orientations at 10° angle increments. At 5° resolution, a total of 186,624 images were

generated. Since the virtual object was implemented in digital form, a high-resolution database was programmatically created using MATLAB software.

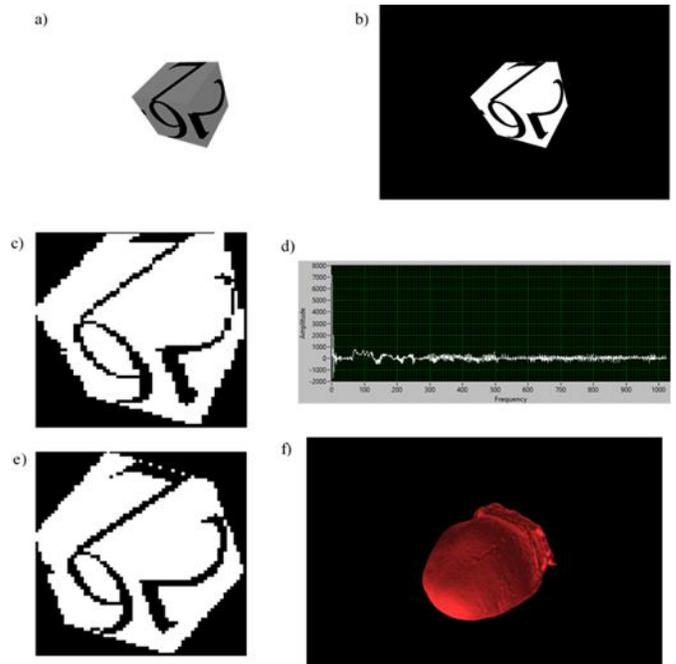


Fig. 1. This flow diagram shows the process of identifying the orientation of a reference object in the reference image and using it to correctly place a virtual heart image. The steps are: a) placement of 3D reference object in the scene, b) acquisition and processing of the reference 2D image, c) scaling of 2D image, d) wavelet transformation and compression of the 2D image, e) prediction of best match using PCA model, and f) registration and display of the appropriate virtual object image. A unique feature of this method is the absence of a restart mechanism

B. Creation of Reference Image Database

The first step in the proposed augmented reality method (shown in Figure 1) was registering an appropriate reference object in the scene to the appropriate 2D image from the VOD. Since the VOD was predetermined, the process was simplified as it only requires knowledge of the rotation of the reference object, its location within the scene, and the requisite scaling. In order to facilitate accurate determination of the three angles of rotation, the reference object must be appropriately designed. Although previous work has shown that this theoretically requires a reference object which has a minimum of four non-planar points [16], the results of this paper (described in Section 4) demonstrate that the accurate determination of the three orthogonal angles of rotation is highly dependent upon the asymmetry of the object such that as the asymmetrical complexity increases, the model accuracy increases as well.

Five reference objects (Objects I-V) with varying degrees of distinctive asymmetry (Figure 2) were created using a CAD software to study their effectiveness for precise object tracking based on optimal number of PCs required.

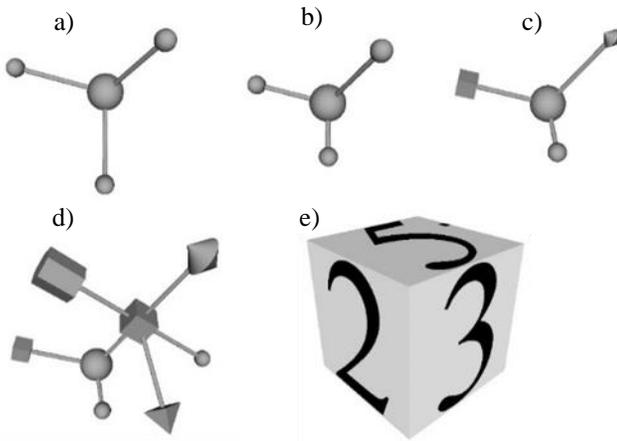


Fig. 2. The reference objects shown above are used to represent an increase in asymmetry for the reference object. The reference objects are a) Object I: Symmetric Spheres, b) Object II: Asymmetric Spheres, c) Object III: Three Shapes, d) Object IV: Multishape, and e) Object V: Dice

The first three reference objects possess only three distinct axes of asymmetry. Object I (Fig. 2a) contains bonds of equal length, resulting in a high degree of symmetry around the axes of rotation. It is worth noting that these symmetry elements are present in the 3D space of the real object. Object II (Fig. 2b) is similar, but all of the bond lengths are unique. In 3D space, this change results in the loss of all symmetry elements except for Identity. In Object III (Fig. 2c), two of the spheres have been replaced by unique shapes (a cube and a top). Object IV (Fig. 2d) possesses five axes extending from a cube where one axis has a highly complex structure attached to the end of the axis. Finally, Object V (Fig. 2e), a dice, has six perpendicular axes of asymmetry, created by a distinct number on each face. The numbers, which range from 2 to 7, were selected for their asymmetry, and hence the numbers 0, 1, and 8 were avoided since they possess symmetry elements other than the Identity.

The 3D reference objects were created as “.stl” files using Solid Edge, and reference images of the various rotations of each object were obtained in MATLAB in 10° increments for each of the three axes. The reference objects were also printed using a 3D printer and were used to acquire test images using a web camera at various object rotations and camera-to-object distances. Creating the reference object as a CAD file allows the Reference Image Database (RID) to be programmatically generated in the same manner as the VOD, while the ability to 3D print the object allows the generation of a physical reference object with high similarity to the RID. This approach ensures a high degree of correlation between the computer-generated images in the RID and the images of the reference object in the scenes that are acquired in real time.

C. Creation of Compressed Wavelet Vector Representation (CWVR) Database

In order to minimize the number of real-time calculations required for image registration, the resolution of the images in the RID was lowered and then the images were compressed further using a wavelet transform. This allowed each image in the RID to be represented by a compressed wavelet vector which was orders of magnitude smaller in number of pixels. These compressed wavelet vectors were arranged into a

reference database which was used to construct the PCA model and to carry-out real-time calculations. Construction of the compressed wavelet vector reference (CWVR) database involved two steps: i) reference image scaling and ii) wavelet transformation and compression.

1) Image Scaling

Although the CAD file used to generate the RID provides high resolution, the results show that a lower resolution allows for faster processing while still allowing for sufficient information to maintain PCA model accuracy. Hence the resolution of the images in the RID was maintained at 64x64 pixels. Due to the calibration-free approach of the proposed model, there is no incoming information about the distance between the reference object and the camera. As such, a scene with a large camera-to-object separation will display a small reference object (low pixel resolution), while a scene with a small camera-to-object separation will display a larger one (high pixel resolution). In each case, however, the orientation of the reference object remains unchanged. Hence, if the object is appropriately scaled, and the scaled image possesses sufficient resolution, the numerical values of three distinct rotation angles can be determined from the PCA model. However, to preserve the scaling required for image registration, it is necessary that both the reference object images in the RID and in the scene be scaled in a similar manner. To achieve this, the incoming scene image was either up-sampled or down-sampled to the same resolution as the RID.

In both cases (scene and RID), the image was restricted to only the contents of a rectangular bounding box using the topmost, leftmost, rightmost, and bottommost points of the reference object. The horizontal scaling (S_h) required for image registration was defined as:

$$S_h = \frac{(i_R - i_L)}{n} \quad (1)$$

where i_R is the index of the rightmost side of the bounding box of the scene image, i_L is the index of the leftmost side of the bounding box of the scene image, and n is the horizontal pixel resolution in the RID. Similarly, the vertical scale (S_v) was defined as:

$$S_v = \frac{(i_T - i_B)}{n} \quad (2)$$

where i_T is the index of the topmost side of the bounding box of the scene image, i_B is the index of the bottommost side of the bounding box of the scene image, and n is the vertical pixel resolution in the RID (same as horizontal).

2) Wavelet Transform and Compression

Each scaled image was subsequently stripped into a single vector by unfolding the rows of the image. A wavelet transform with four wavelets and scaling functions [17] was applied to the resulting vector. The Daubechies family of orthogonal wavelets was chosen due to their extensive use in

data compression. Specifically, the Daubechies wavelet filter with 8 taps and 4 vanishing moments was selected (also referred to as a D8 or db4 referring to the $N=2A$ relationship between the number of taps, N , and the number of vanishing moments, A) as a reasonable compromise between image resolution and compression efficiency. The advantage of the wavelet transform is that it preserves both the frequency and the position information of the image vector (i.e., the function is not translationally invariant as is the case with most Fourier transform methods) [18]. Moreover, the use of discrete wavelet transform (DWT) is computationally efficient. Due to the nature of the reference object image, the dominant and requisite information was contained almost completely in the low frequency wavelet coefficients. Thus, the final step of compression involved truncating the high frequency wavelet coefficients to achieve image compression while preserving the lowest 1024 coefficients that retain essential information needed to determine the object orientation. As shown in Figure 3, the original input sample image (Panel A) and the reconstructed compressed image after wavelet transform (Panel B) were almost identical, but the latter required 75% less data for creation and storage. A slight blurring of sharp edges in the compressed images due to the loss of the high frequency components is evident in the figure but this had an insignificant impact on the accuracy of the PCA model.

D. Constructing the Principal Component Analysis (PCA) Model

Once the CWVR database was created, it was then used to construct a PCA Model [19]. The use of a PCA model provided two distinct advantages. First, it allowed the CWVR for all images of the reference object to be described using a multi-dimensional eigenvector space that accounted for the greatest variance of the underlying data structure. Second, it enabled each CWVR to be mapped into the eigenvector space by using a single scalar value (eigenvalue or score) for each eigenvector. Since the number of eigenvectors (more commonly referred to as principal components) required to account for the majority of variance was considerably less than the length of the CWVR, a significant further compression of the data dimensionality was achieved. For example, if the CWVR contained 1024 data points, and the variance was described by 10 principal components, then a compression of 100-fold was achieved since each CWVR could now be represented by only 10 eigenvalues within the space of the PCA model.

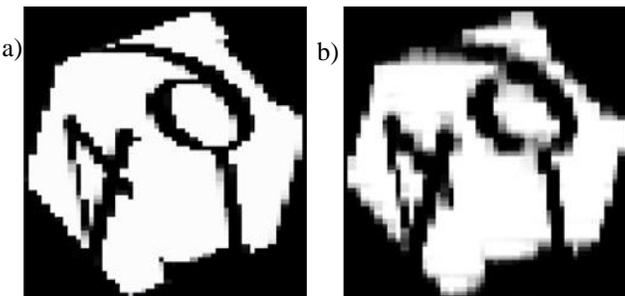


Fig. 3. Two images a) before wavelet compression and b) after wavelet compression are shown. The differences in the images are minimal, however, the amount of information required to reconstruct the image after compression is significantly smaller, and is used as the basis of the CWVR database

The PCA model was constructed by creating a data matrix X , where each row of the matrix corresponds to a CWVR, and the number of rows is equal to the number of reference images. The data matrix was then decomposed using a singular value decomposition algorithm:

$$X = U\Sigma V^T \tag{3}$$

where U is an $n \times m$ orthonormal matrix, Σ is an $m \times m$ eigenvalue matrix with all zero off-diagonal elements, V^T is an orthonormal $m \times m$ matrix, n is the number of reference images and m is the length of the CWVR. Each row of matrix V^T is an eigenvector or principal component (PC), thus the resulting number of principal components is equal to the length of the CWVR. Since V^T is an orthonormal matrix, all of the PC row vectors are orthogonal and define a multivariate space containing the compressed images. The coordinates of each compressed image within this multivariate space was given by the rows of a scores matrix, S , which is simply:

$$S = U\Sigma \tag{4}$$

where S is a $n \times m$ matrix. Thus the scores reflect where each sample lies on the PC axes. However, since the PCs were sorted in decreasing order of variance, the majority of the variance was described by the first few PCs and the higher order PCs were dominated mostly by noise. This allowed the PCA model to be constructed by truncating to the number of columns (k) in the scores matrix, S , and the loadings matrix, V :

$$X' = S'(V')^T \tag{5}$$

Where X' is an approximation of the compressed image data, and S' and V' have $k \times m$ dimension ($k \ll n$). T indicates the transpose. Figure 4 illustrates the matrix order reduction obtained while solving the Eqn. 5.

E. Image Registration

Since this was a rigid image registration method, there were six degrees of freedom which had to be determined in order to accurately display a virtual object in the scene image. These included three degrees of freedom in translation and three degrees of freedom in rotation.

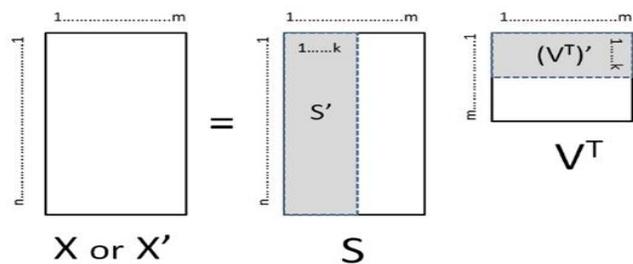


Fig. 4. The equation for creation of the PCA model is shown where n corresponds to the number of images in the PCA model and m corresponds to the length of the compressed wavelet vector representation (CWVR) of each image. The solid polygons correspond to the data matrix X , the scores matrix S , and the loadings (principal component) matrix V^T . The dashed polygon boxes correspond to the retained columns (k) of S and V which are used to construct an approximation (X') of the data matrix to generate the PCA model

The translation that occurred between two images was determined by using the information acquired during the image scaling process. By comparing the geometric mean of the bounding box surrounding the reference object between the model and scene image, the translation on the two axes (which are parallel to the sides of the image) was determined. Translation along the third axis was accounted for by scaling the virtual object image according to the scale of the incoming reference object image. This was performed using the horizontal and vertical scale ratios given by Eqn. 1 and Eqn. 2, respectively.

The rotation around three orthogonal angles was simultaneously determined by using the PCA model as discussed earlier. An image of the reference object in an incoming scene was processed in the same way as that of the CWVR database. The PCs of the unknown object orientation were then calculated using the resulting CWVR (y) and the reference object PCA Model (i.e., the truncated matrix $(V')^T$) by solving Eqn. 6 for the scores vector s :

$$(V')^T s = y \quad (6)$$

Since the PCs are orthogonal, $[(V')^T]^{-1} = V'$. This yielded a trivial solution for determining the scores for unknown compressed images:

$$s_{new} = y_{new} V' \quad (7)$$

where s_{new} corresponds to the scores vector containing the coordinates of a new compressed image in PCA space and y_{new} is the CWVR for the new image. The significance of Eqn. 7 is that regardless of the size of the image database used to create the PCA model, the model of the image space was described by a $n \times k$ matrix, where k is the number of PCs and n is the length of the CWVR. Thus for an image database at 10 degree resolution (total 22,104 images) where each original image was defined by 1024x 1024 pixels ($n = 1024$), a PCA model containing 22 principal components ($k = 22$) yielded a compression of the database by six orders of magnitude.

Although it is possible for the new image to contain an exact match with the scores of the database, this probability was limited by the angular resolution of the database. For this reason, the best match was taken as the nearest neighbor (smallest distance) in PC model space, where the distance from nearest neighbors was calculated as the sum of the squares of the score differences:

$$d = \sum_{PC=1}^k (s_{PC}^{new} - s_{PC}^{model})^2 \quad (8)$$

where d is the distance, k is the maximum number of PCs in the model, and s is the score for a particular PC for the new image being predicted and a neighboring images contained in the PCA model. Since each vector in the CWVR database was indexed to the angles of rotation of its reference image, identifying the best match also identified the corresponding values for rotation about three different axes.

Each reference object was tested using PCA Models constructed using a PC space ranging from one to 50 PCs in order to determine the optimal number of PCs for that object. The optimum number of PCs was chosen as the minimum PC number where adding an additional PC did not yield a statistically significant decrease in the standard error of the model predictions. This was determined by Malinowski's F-test [20]. Malinowski's F-test assumes that the sum of the eigenvalues can be decomposed into either parts which are significant or noise, and that the significant eigenvalues provide an estimate of the true number of principal components needed. If there are a maximum of p possible principal components (the minimum of either the number of samples or the number of variables), then the F-statistic for the s^{th} eigenvalue (λ_s) is:

$$F_s = \frac{\lambda_s}{\sum_{j=s+1}^p \lambda_j / (p - s)} \quad (9)$$

and the maximum PC is taken as that having the minimum F_s value.

The best match for an acquired image was restricted to a reference object image within the CWVR database (10° angle increments) and as such, an angle error of less than or equal to $\pm 5^\circ$ is considered accurate. Thus for an incoming sample image with a 45° rotation on a particular axis, a correct match is either 40° or 50° in that same angle. Figure 5 illustrates one example of the largest error for an incoming scene image with X-axis rotation = 45° , Y-axis rotation = 315° , Z-axis rotation = 0° (henceforth written as: $45^\circ, 315^\circ, 0^\circ$) (panel a) and the corresponding best RID correct match through the PCA

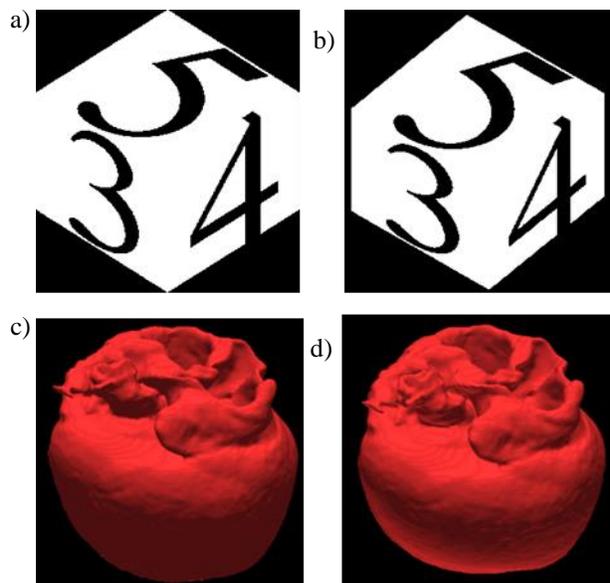


Fig. 5. Shown is an example of the largest error for a correct match given the selected angular resolution (10° increments) of the reference image database. The images shown are a) the scene image ($45^\circ, 315^\circ, 0^\circ$), and b) the best RID match through PCA analysis ($40^\circ, 310^\circ, 0^\circ$). The corresponding registered heart images are shown in c) and d). As seen, the difference between the two Object V images and their respective heart images is barely perceptible, but could be further reduced by using a higher angulation resolution (i.e. 5° increments)

analysis (40°, 310°, 0°). As seen, the difference between the two Object V images and their respective heart images is barely perceptible, but could be further reduced by using a higher angulation resolution (i.e. 5° increments).

The performance of the object matching algorithm was assessed by percent error (δ) calculated as:

$$\delta = \frac{i}{a} \times 100 \quad (10)$$

Where i was the number of incorrect angle matches over a certain tolerance of error for a given PCA model and a was the total number of angles (or three times the number of images tested).

III. RESULTS

The percent standard errors for each reference object model as a function of the number of PCs are plotted in Figure 6. Each model reaches an apparent minimum percent error after reaching an optimum number of PCs, implying that the additional PCs are no longer providing significant additional information. For all objects, the error dropped significantly for the initial PCs after which the accuracy does not improve noticeably for higher order PCs. It was observed that the accuracy improved with extent of asymmetry in the objects with the most symmetric object (Object I) giving the highest error and the most asymmetric object (Object V) giving the lowest error regardless of the number of PCs included.

Table I lists the number of optimal PCs needed for each type of reference object. It was observe that as the asymmetry of the reference object increases, the number of PCs required to describe the variance of the CWVR database increases. Thus the highly symmetric object (Object I) required only 9 PCs, while the most asymmetric object (Object V) required 29 PCs. The table also provides percent errors for each reference object PCA model at the optimum PC number. The percent errors are broken down into three categories: 15° (the percent of incorrect predictions exceeding an error of 15°), 10° (percent of incorrect predictions exceeding an error of 10°), and 5° (percent of incorrect predictions exceeding an error of 5°). Also shown for each object PCA model is the root mean square of the distances between the reference and sample images where an error less than or equal to 5° is considered accurate. As can be observed, the results showed a significant increase in both accuracy and the optimum number of PCs as the complexity of the asymmetry in the reference objects increased.

The errors shown in Table I and Figure 6 both correspond to errors in the predicted angle vs. the actual angle of rotation. However, since each object possesses 3 orthogonal angles of rotation, it is possible for each image mismatch to correspond to one or more improper angles. The actual number of incorrect images and incorrect angles for each of the 672 tested samples is given in Table II.

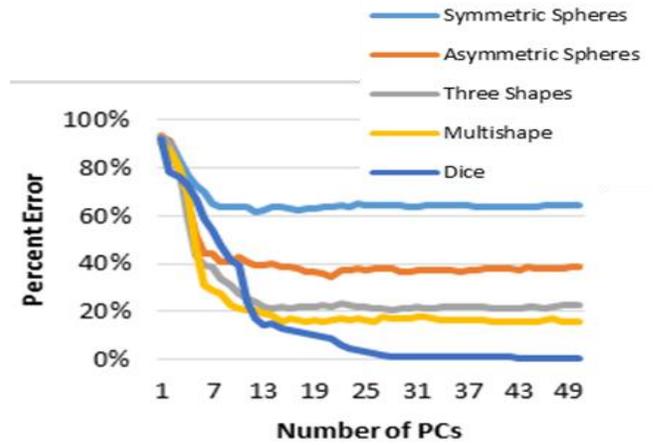


Fig. 6. The graph shows the percent standard error of the PCA model for each reference object as a function of number of retained PCs. Initially the error decreases rapidly with each addition of a PC, however, after a certain number of PC's the resulting decrease in error becomes minimal, implying that the inclusion of further PC's is no longer providing significant information to the PCA model. Reference objects with higher degrees of asymmetry yield correspondingly lower errors once the minimal number of PCs is reached

TABLE I. EFFECT OF REFERENCE OBJECT ASYMMETRY ON PCA OPTIMAL PC# AND PCA MODEL ERROR

Object	Optimal Number of PCs	Percent Error (5°)	Percent Error (10°)	Percent Error (15°)	Average RMS
Object I	9	63.79%	61.71%	56.35%	62.9964
Object II	12	39.63%	30.26%	28.47%	31.9440
Object III	14	21.58%	11.86%	9.62%	12.4449
Object IV	24	16.82%	8.48%	5.26%	6.7919
Object V	29	0.79%	0.00%	0.00%	0.1022

TABLE II. THE NUMBER OF INCORRECT IMAGE AND ANGLE PREDICTIONS (AT 5°) WHEN PREDICTING THE 671 UNIQUE SAMPLES

Object	Optimal Number of PCs	Incorrect Images	Incorrect Angles	Average
Object I	9	498	1286	2.582
Object II	12	330	799	2.421
Object III	14	224	435	1.942
Object IV	24	189	339	1.794
Object V	29	9	16	1.778

As the asymmetry increased for the first four objects, the number of image mismatches decreased monotonically from 498 to 189. For the Object V PCA model, however, there was a precipitous decrease to only 9 incorrect images. Table II also gives the average number of incorrect angles per sample, and shows that as the average number of incorrect angles in each image mismatch decreased with an increase in asymmetry in the objects. One example of an image mismatch for Object II is shown in Figure 7. Panel a) corresponds to a test sample (60°, 50°, 210°) and Panel b) corresponds to the PCA predicted match (140°, 230°, 90°).

As can be seen, although the object possesses no symmetry in 3D space, there still exist combinations of distinct angle rotations which yield 2D views which are nearly identical. This predicted mismatch problem was further exacerbated upon image compression. Figure 8 shows one example for Object III after compression and scaling. The image with rotations $(10^\circ, 180^\circ, 80^\circ)$ (Panel A) was wrongly recognized as $(0^\circ, 0^\circ, 40^\circ)$ (Panel B), leading to an error of 230° . At the lower resolution, the image differences approach the noise limit.

For Object V, the highest error mismatch is shown in Figure 9 (reconstructed from their respective CWVRs). Despite the low resolution, the images are clearly distinguishable with the rotations of $(90^\circ, 90^\circ, 285^\circ)$ in Panel A, and that of $(100^\circ, 80^\circ, 290^\circ)$ in Panel B. The relatively small angle error combined with the *visually distinguishable features* of the two images suggests that the cause of the mismatch was rooted in the PC model itself. Indeed, a plot of the % variance attributable to each PC (Figure 10) indicated that the higher order PCs for object V constituted a significant amount of variance (Figure 10, inset) when compared to that of more symmetric object (Object II). This is consistent with the finding that more PCs are required to effectively describe the modeling space of objects with greater asymmetry (Figure 6). It also suggests that using an F-test may not be a reliable

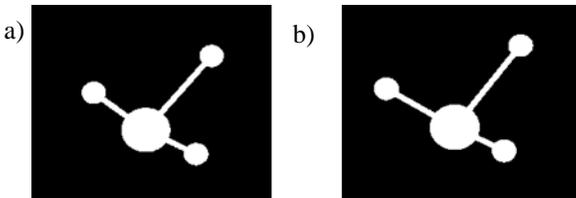


Fig. 7. An image mismatch resulting from PCA model of the Asymmetrical Spheres reference object. The images are shown at high resolution prior to compression. The left image (x-axis rotation = 60° , y-axis rotation = 50° , z-axis rotation 210°) is very similar, even at high resolution, to the right database image (x-axis rotation = 140° , y-axis rotation = 230° , z-axis rotation = 90°); however the angles are significantly different. This is a result of the loss of information as one moves from 3D to 2D space

method to determine the optimum number of PCs since the percentage of incorrectly identified object images beyond 29 PCs is relatively small (1.34%; 9 out of 672 images). When the Object V PC model is expanded to 45 PCs, the number of image mismatches dropped to 6 (0.89% incorrect) and no mismatches at 200 PCs. Thus the error in the Object V PCA models appears to originate in the ability of the PCA model to describe the asymmetry and not in the lack of distinguishable asymmetry. Since the truncation of PCs for a given model is part of the data compression, it suggests the need to balance compression in an effort to maintain the asymmetric properties required for accuracy. This is not noteworthy in case of more symmetric objects since the error remains high even for a significantly larger number of PCs (see Figure 6). Although a higher degree of asymmetry requires a higher number of PCs to maintain accuracy, the overall compression is still significant. For example, the original sample image contains over 1 million pixels, and conversion to a 64×64 image results into 4096 pixels. It is further reduced to 1024 pixels after wavelet compression and truncation. Thus even when using

200 PCs to represent the image, provides an additional factor of 5 in data compression and yields a total data compression of over 5000 per image. With regards to the image database space ($22,103 \text{ images} \times 1024 \times 1024 = 23 \text{ billion coordinates}$), the PCA model compression (200 PCs \times 1024 length of CWV = 204,000 coordinates) still results in five orders of magnitude compression.

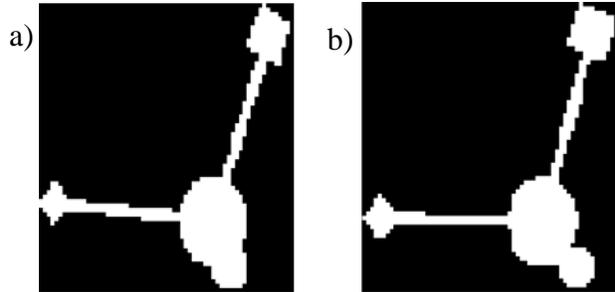


Fig. 8. The figure shows an example of a PCA predicted image-mismatch for the Object III PCA model. The images are shown after scaling and image compression. The differences between the top image ($10^\circ, 180^\circ, 80^\circ$) and the bottom database image ($0^\circ, 0^\circ, 40^\circ$) are almost indistinguishable at the lower resolution, where they approach the noise limit

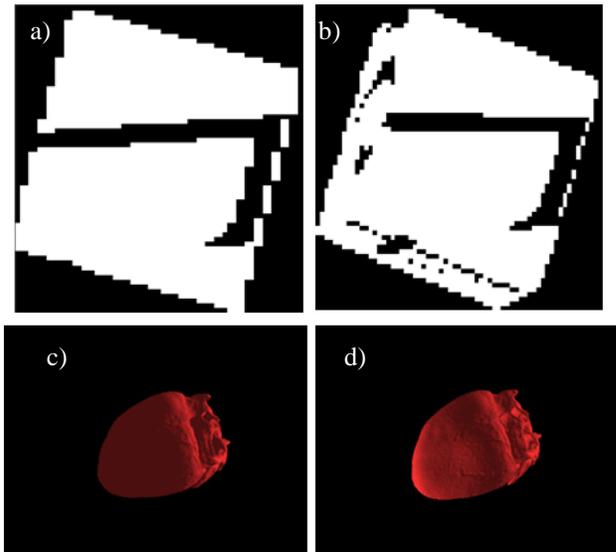


Fig. 9. An example of the maximum error obtained using a PCA model with 29 retained principal components for Object V is shown. The images are a) the sample image ($90^\circ, 90^\circ, 285^\circ$), b) the best match ($100^\circ, 80^\circ, 290^\circ$), c) the heart image corresponding to the sample image, and d) the heart image corresponding to the best match. Since the images are visually distinctive, it can be reasoned that the error lies in the PC model itself and indicates the need for additional PCs if more accuracy is required

The PCA models for each object were constructed using 22,103 images (i.e., 22,103 vectors in the CWVR database). Figure 11 shows the PCA space defined by these images at various orders of PCs for Object V (Fig. 11d-e) and Object II (Fig. 11a-c). The panels A and D show contents of the images extracted in the first three principal components (X, Y and Z axes). As can be seen, the plot for Object II (Panel a) exhibits a great deal of structure. Indeed, if the points in the plot are considered as a solid object, a 3-fold improper axis of rotation exists (S_3). For the Object V plot (Panel d), the defined space is less structured, and although not spherical, is significantly

more homogeneous. As higher PC numbers are used to define the axes, the spread in the PC scores decreased for both objects, as expected.

IV. DISCUSSION

The paper presents a novel image registration method using pre-processed database of compressed image vectors spanning all possible image rotations and scaling. This method uses a combination of discrete wavelet transform to compress the images without losing any valuable information and principal component analysis to construct an accurate estimation model. This approach significantly reduced the computations and enabled real-time processing for seamless medical augmented reality applications.

The computational benefits of this approach are achieved by utilizing additional computational time prior to image registration for processing already acquired reference image database using a reference object with distinct and complex asymmetric properties. By acquiring reference object images

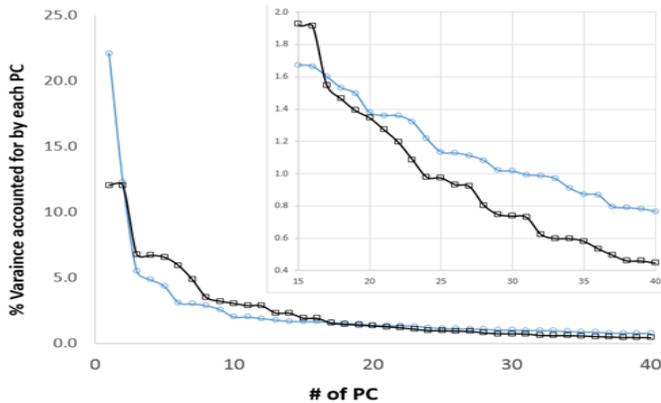


Fig. 10. The graph shows the relationship between % variance and the retained PC number for Object V and Object II. The graph inset shown in the top right shows that the Object V PCA model still contains a significant amount of variance at higher order PCs when compared to the PCA models of the more symmetric reference objects. This indicates that more PCs are required to accurately reflect the higher degree of asymmetry and is supported by the observation that inclusion of 200 PCs in the Object V PCA model eliminates all mismatches

from multiple viewpoints, a comprehensive model can be developed using principal component analysis which accurately matches an incoming image whose angles of rotation (with respect to the viewer) are not known, with a corresponding reference image. Since the reference image is indexed to the angles of rotation, image registration is straightforward and only requires scaling and positioning. Alternatively, these indexed angles can be used to determine an appropriate transform of the 3D virtual heart model into 2D image space. The advantages of the former are - 1) the bulk of the computations (generating high resolution heart images, building the reference image database, and constructing the PCA models) are carried out only once and in advance so that the real-time IR is computationally inexpensive; and 2) the current trend in computing storage is the use of high speed flash interfaced to the CPU via a high speed bus (e.g. PCIe) which allows extremely fast image recall. Thus the IR time of this algorithm remains superior to the existing method while

offering the advantage of a higher resolution rendering upon registration.

The robust nature of the model presented in this paper is created due to the use of a predefined, asymmetric reference object which is present in the incoming image. Unlike other methods, where tracked features are chosen in real time [8,9,21], in this method, the tracked features are predetermined. This significantly enhances the speed and accuracy of image registration at the cost of creating a more rigid technique which requires the presence and visibility of a specific reference object. Traditional IR methods rely on the creation of mathematical transformations to track features in scene images which are cumbersome to use with higher resolution images, and require "good" features which can be easily tracked [22] [23]. Intensity-based IR is one way to bypass this requirement, but these methods still employ a mathematical transformation, which ultimately increases the amount of real-time computation necessary for image registration [9]. Intensity-based IR methods also require a computationally expensive restart mechanism to obtain an optimal transform instead of trying to refine a bad transformation [10]. The proposed method, however, does not refine previous transforms, eliminating the need for a restart mechanism altogether.

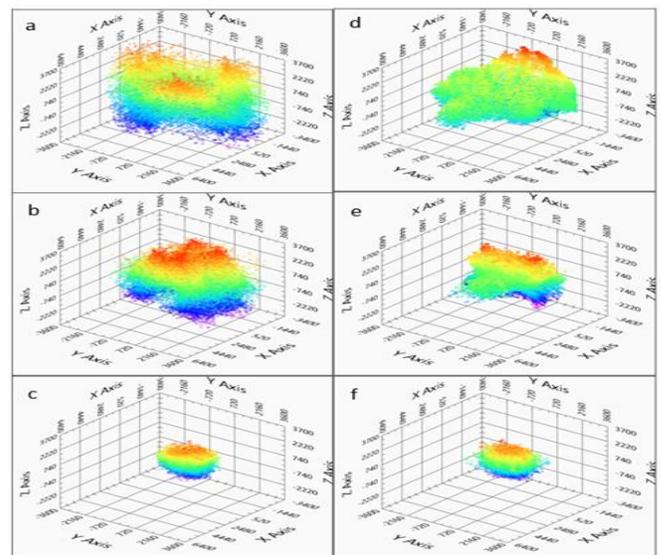


Fig. 11. The figure shows the PC space of the PCA model for Object V (right) and Object II (left) using different combinations of PCs for the axes of the space. The top images show the scores for the 1st (x-axis), 2nd (y-axis), and 3rd (z-axis) PCs. The middle images show the scores for the 4th, 5th, and 6th PCs; and the bottom images show the scores of 27th, 28th, and 29th PCs. A three-fold axis of improper rotation (S_3) can be seen in the Object II PCA space for the first 3 PCs (top left), while the similar PCA space for Object V is more homogenous (top right)

The significant difference of the proposed method compared to other IR techniques makes it difficult to draw a direct comparison. Nevertheless shape-based image retrieval techniques have often utilized PCA in order to reduce data dimensionality and decrease computation time. Image retrieval methods have demonstrated that more complex shapes are easier to use with PCA analysis [24], and the principal component descriptors are preferable to other methods of image identification and retrieval [25]. Content based image

retrieval (CBIR) is a popular technique that utilizes such methods to search and retrieve images from large databases [26]. Typically employed to manage large volumes of digital images, this technique is similar to the proposed method which is repurposed as an IR method.

When uniform PCA space is encountered in a model, it is common to convert from a PCA classification method (as used in this paper) to a principal component regression (PCR) method [27] in order to derive more quantitative results (e.g., the ability to quantitatively interpolate between the modeled angles). Although it is beyond the scope of the present work, construction of a PCR model for this implementation of augmented reality would involve the multivariate regression of the scores of a PCA model against the angles of rotation. This could also be accomplished with a partial least squares model using a PLS-2 algorithm [28]. The more homogenous clustering for the Object V PCA model would suggest a greater likelihood of success of such quantitative modeling when using a higher degree of reference object asymmetry. This further exemplifies the importance of introducing complexity into the asymmetry of reference objects.

V. CONCLUSIONS

This paper presents a novel image registration technique involving image compression and PCA modeling based on the use of reference objects with complex asymmetry. The design provides a method to eliminate the real-time computational costs of performing geometric transforms and by using PCA classification, operates without the need for a restart mechanism. The method was validated using 672 object images to test a PCA model created from 22,103 reference images. The asymmetry of the reference objects was found to highly correlate with the accuracy of the image registration. In particular, for highly asymmetric objects, the accuracy was predominantly dependent upon the inclusion of enough principal component vectors to accurately describe the asymmetry of the objects described in the PCA model space. For higher symmetry objects, the inclusion of higher PC order models had little to no impact on the accuracy. Future studies could further investigate the properties and uses of complex asymmetry to enhance the accuracy of image registration methods. Since the majority of image processing in this method is done prior to the real-time process, and the data compression resulted in significant reduction in memory requirements, the proposed method is well suited for real-time medical augmented reality applications.

REFERENCES

- [1] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. "Recent advances in augmented reality." *Computer Graphics and Applications*, IEEE, vol. 21, no. 6, pp. 34-47, Nov, 2001.
- [2] Van Krevelen, D., & Poelman, R. "A survey of augmented reality technologies, applications and limitations," *International Journal of Virtual Reality*, vol. 9, no. 2, pp. 1-19, Jun, 2010.
- [3] Fritz, J., Paweena, U., Ungi, T., Flammang, A. J., Fichtinger, G., Iordachita, I. I., & Carrino, J. A. "Augmented reality visualisation using an image overlay system for MR-guided interventions: technical performance of spine injection procedures in human cadavers at 1.5 Tesla." *European radiology*, vol. 23, no. 1, pp. 235-245, Jul, 2013.
- [4] Azagury, D., Ryou, M., Shaikh, S., San José Estépar, R., Lengyel, B., Jagadeesan, J., . . . Thompson, C. "Real-time computed tomography-based augmented reality for natural orifice transluminal endoscopic surgery navigation," *British Journal of Surgery*, vol. 99, no. 9, pp. 1246-1253, Sep, 2012.
- [5] Nakamoto, M., Ukimura, O., Faber, K., & Gill, I. S. "Current progress on augmented reality visualization in endoscopic surgery," *Current opinion in urology*, vol. 22, no. 2, pp. 121-126, Mar, 2012.
- [6] Liao, H., Inomata, T., Sakuma, I., & Dohi, T. "3-D augmented reality for MRI-guided surgery using integral videography autostereoscopic zimage overlay," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no.6, pp. 1476-1486, Jun, 2010.
- [7] Zitova, B., & Flusser, J. "Image registration methods: a survey," *Image and vision computing*, vol. 21, no. 11, pp. 977-1000, Jun, 2003.
- [8] Reddy, B. S., & Chatterji, B. N. "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE transactions on image processing*, vol. 5, no. 8, pp. 1266-1271, Aug, 1996.
- [9] Kim, J., & Fessler, J. A. "Intensity-based image registration using robust correlation coefficients," *Medical Imaging, IEEE Transactions on*, vol. 23, no. 11, pp. 1430-1444, Nov, 2004.
- [10] Valsecchi, A., Damas, S., Santamaría, J., & Marrakchi-Kacem, L. "Intensity-based image registration using scatter search," *Artificial intelligence in medicine*, vol. 60, no. 3, pp. 151-163, Feb, 2014.
- [11] Markelj, P., Tomaževič, D., Likar, B., & Pernuš, F. "A review of 3D/2D registration methods for image-guided interventions," *Medical image analysis*, vol. 16, no. 3, pp. 642-661, Apr, 2010.
- [12] Besl, P. J., & McKay, N. D. "Method for registration of 3-D shapes," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 14, no. 2, pp. 239-256, Feb, 1992.
- [13] Glover, F., & Kochenberger, G. A. "Handbook of metaheuristics," Springer Science & Business Media, 2003.
- [14] Li, H., Qi, M., & Wu, Y. "A Real-Time Registration Method of Augmented Reality Based on Surf and Optical Flow," *Journal of Theoretical and Applied Information Technology*, vol. 42, no. 2, pp. 281-286, Aug, 2012.
- [15] Wang, J., Suenaga, H., Hoshi, K., Yang, L., Kobayashi, E., Sakuma, I., & Liao, H. "Augmented Reality Navigation with Automatic Marker-Free Image Registration Using 3D Image Overlay for Dental Surgery," *IEEE Trans Biomed Eng.*, vol. 61, no. 4, pp. 1295-1304, Apr, 2014.
- [16] Vallino, J. R. "Interactive augmented reality," University of Rochester, 1998.
- [17] Mohamed, M., & Deriche, M. "An Approach for ECG Feature Extraction using Daubechies 4 (DB4) Wavelet," *International Journal of Computer Applications*, vol. 96, no. 12, pp. 36-41, Jun, 2014.
- [18] Jensen, A., & Cour-Harbo, A. "Ripples in mathematics: the discrete wavelet transform," Springer, 2001.
- [19] Wold, S., Esbensen, K., & Geladi, P. "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1, pp. 37-52, Aug, 1987.
- [20] Malinowski, E. R. "Statistical F-tests for abstract factor analysis and target testing," *Journal of Chemometrics*, vol. 3, no. 1, pp. 49-60, Jan, 1989.
- [21] Jin, H., Favaro, P., & Soatto, S. "Real-time feature tracking and outlier rejection with changes in illumination," *Computer Vision*, 2001. *Proceedings. Eighth IEEE International Conference on*, pp. 684-689, Jul, 2001.
- [22] Lucas, B. D., & Kanade, T. "An iterative image registration technique with an application to stereo vision," *IJCAI*, vol. 81, pp. 674-679, Aug, 1981.
- [23] Shi, J., & Tomasi, C. "Good features to track," *Computer Vision and Pattern Recognition, Proceedings IEEE Computer Society Conference on*, pp. 593-600, Jun, 1994.
- [24] Tavoli, R., & Mahmoudi, F. "PCA-Based Relevance Feedback in Document Image Retrieval," *International Journal of Computer Science Issues*, vol. 9, no. 4, Jul, 2012.
- [25] Wang, B., & Bangham, J. A. "PCA based shape descriptors for shape retrieval and the evaluations," *Computational Intelligence and Security, International Conference on*, pp. 1401-1406, Nov, 2006.

- [26] Jain, A., Muthuganapathy, R., & Ramani, K. "Content-based image retrieval using shape and depth from an engineering database," *Advances in Visual Computing*, Springer, pp.255-264, Nov, 2007.
- [27] Geladi, P., & Esbensen, K. "Regression on multivariate images: principal component regression for modeling, prediction and visual diagnostic tools," *Journal of Chemometrics*, vol. 5, no. 2, pp. 97-111, Mar, 2005.
- [28] Wold, S. "Nonlinear partial least squares modelling II. Spline inner relation," *Chemometrics and intelligent laboratory systems*, vol. 14, no. 1, pp. 71-84, Apr, 1992.