Published in final edited form as:
Med Image Anal. 2013 December ; 17(8): . doi:10.1016/j.media.2013.06.002.

Explicit shape descriptors: Novel morphologic features for histopathology classification

Rachel Sparks[a,*] and **Anant Madabhushi**[b,*]

[a]Rutgers University, Department of Biomedical Engineering, 599 Taylor Road, Piscataway, NJ, USA

[b]Case Western Reserve University, Department of Biomedical Engineering, 10900 Euclid Ave, Cleveland, OH, USA

Abstract

Object morphology, defined as shape and size characteristics, observed on medical imagery is often an important marker for disease presence and/or aggressiveness. In the context of prostate cancer histopathology, gland morphology is an integral component of the Gleason grading system which enables discrimination between low and high grade disease. However, clinicians are often unable to distinguish between subtle differences in object morphology, as evidenced by high inter-observer variability in Gleason grading. Boundary-based morphologic descriptors, such as the variance in the distance from points on the boundary of an object to its center, may not have the requisite discriminability to separate objects with subtle shape differences. In this paper, we present a set of novel explicit shape descriptors (ESDs) which are capable of distinguishing subtle shape differences between prostate glands of intermediate Gleason grades (grades 3 and 4) on prostate cancer histopathology. Calculation of ESDs involves: (1) representing object morphology using an explicit shape model (e.g. medial axis); (2) aligning the shape models via a non-rigid registration scheme with a diffeomorphic constraint and quantifying shape model dissimilarity; and (3) applying a non-linear dimensionality reduction scheme (e.g. Graph Embedding) to learn a low dimensional projection encoding the shape differences between objects. ESDs are hence the principal eigenvectors in the reduced embedding space. In this work we demonstrate that ESDs in conjunction with a Support Vector Machine classifier are able to correctly distinguish between 888 prostate glands corresponding to different Gleason grades (benign, grade 3, or grade 4) of prostate cancer from 58 needle biopsy specimens with a maximum accuracy of 0.89 and corresponding area under the receiver operating characteristic curve of 0.78.

Keywords

Shape descriptors; Prostate histopathology; Gleason grading; Manifold learning; Diffeomorphic registration

1. Introduction

The morphology of anatomical objects, defined as shape and size characteristics, on medical imagery is often an important cue to determine disease presence and/or disease aggressiveness (Jobst et al., 1994; Epstein et al., 2005; Rangayyan and Nguyen, 2007;

[*]Corresponding authors. Tel.: +1 216368 3502 (R. Sparks). rsparks@eden.rutgers.edu (R. Sparks), axm788@case.edu (A. Madabhushi).

Tabesh et al., 2007; Gerardin et al., 2009; Gorczowski et al., 2010; Doyle et al., 2012a; Looi et al., 2011; Pachauri et al., 2011). One application where object morphology is important is in the Gleason grading of prostate cancer (CaP) which utilizes the appearance of nuclei and glands on histopathology (Epstein et al., 2005; Tabesh et al., 2007; Monaco et al., 2010; Doyle et al., 2012a).

CaP appearance, as seen on needle core biopsies, is classified according to Gleason grade (from grade 1 to grade 5). Gleason grade (Gleason, 1966) has been suggested as one of the most important predictors of disease aggressiveness with higher Gleason grade patterns being typically associated with more aggressive disease (Epstein et al., 2005). Correctly identifying Gleason grade patterns is critical for determining the appropriate treatment strategy for a patient with CaP (Madabhushi et al., 2011). Pathologists typically analyze architectural features, the arrangement and morphology of glands and nuclei within the tissue, in order to predict the corresponding Gleason grade (Epstein et al., 2005). In low grade cancers, prostate tissue has a coherent spatial architecture with distinct gland lumen surrounded by cell nuclei. For higher Gleason grade patterns, gland structure begins to breakdown with gland lumen becoming indistinct and crowded with a large density and concentration of cell nuclei. Fig. 1 displays examples of prostate glands identified as (a) benign, (b) Gleason grade 3, and (c) grade 4 with lumen (red) and nuclear (blue) boundaries of the glands segmented. Distinguishing intermediate Gleason grade patterns is a difficult task, previous studies having reported an inter-observer agreement between pathologists as low as 0.47–0.64 (reflecting low to moderate agreement) (Allsbrook et al., 2001). Low inter-observer agreement reflects the difficulty in distinguishing between objects with very subtle shape differences (e.g. gland appearance between Gleason grade 3 and grade 4 patterns). The low inter-observer agreement for identifying intermediate Gleason grades (Allsbrook et al., 2001) on prostate histopathology reflects the need for developing quantitative, reproducible computer-extracted morphologic descriptors to complement human observers in distinguishing subtle differences in object morphology.

The rest of this paper is organized as follows. In Section 2 we discuss previous work in the context of quantitative attributes for describing object morphology. In Section 3 we present our methodology for capturing object morphology via a shape model and Section 4 describes our methodology to quantify differences between shape models and derive our novel explicit shape descriptors (ESDs). Section 5 describes the experimental design and the results of experimental evaluations. Finally, in Section 6 we present concluding remarks.

2. Previous work quantifying object morphology and novel contributions

Computer-aided image analysis systems that quantify differences in object morphology may reduce inter-observer variability in applications where object morphology is of diagnostic and/or prognostic significance (Rangayyan and Nguyen, 2007; Tabesh et al., 2007; Doyle et al., 2012a; Yang et al., 2009; Georgiou et al., 2007).

Several boundary-based shape descriptors have been previously presented (Rangayyan and Nguyen, 2007; Tabesh et al., 2007; Yang et al., 2009; Georgiou et al., 2007) to extract specific characteristics from an object margin, determined to be important for a specific task. However, these descriptors typically quantify a single, specific shape characteristic. Boundary-based shape descriptors include fractal dimension (Rangayyan and Nguyen, 2007), a measure of the self similarity between the object and its parts; measures of contour variation including symmetry (Yang et al., 2009); and wavelet parameters of the object boundary (Georgiou et al., 2007). Additionally, other measures have included how close to circular an object is (Street et al., 1993) or how quickly the contour varies (Liney et al., 2006). Such descriptors provide a single global measure of object morphology, and hence

may be unable to distinguish between objects with subtle, local shape differences. Fig. 2 illustrates this problem with boundary-based descriptors, where prostate glands with distinct morphology yield similar values in terms of contour smoothness and compactness (Agner et al., 2011). Note that for the glands shown in Fig. 2, there may exist other boundary-based shape descriptors that can accurately distinguish between the glands. However, determining the boundary-based shape descriptors that can best distinguish between the prostate glands shown in Fig. 2 requires *a priori* information about the domain and classification task.

Alternatively, descriptors that explicitly model object shape, and are hence able to recapitulate the original shape of an object, provide an alternative approach to assessing differences in object shape (Bailey and Srinath, 1996; Belongie et al., 2002; Blum, 1967; Brechböhler et al., 1995; Cootes et al., 1994; Kauppinen et al., 1995; Mokhtarian and Abbasi, 2002; Persoon and Fu, 1977; Reuter et al., 2006; Rueda et al., 2010; Shinagawa et al., 1991; Sun et al., 2009; Zhang, 1994). These descriptors may be used in conjunction with an appropriate similarity metric to quantify differences between shape model representations. Such descriptors typically do not require *a priori* information about the domain and classification task to accurately determine subtle morphologic differences between objects.

Development of model-based descriptors is an active area of research and many approaches to modeling the shape of an object have been presented (Bailey and Srinath, 1996; Belongie et al., 2002; Blum, 1967; Brechböhler et al., 1995; Cootes et al., 1994; Kauppinen et al., 1995; Mokhtarian and Abbasi, 2002; Persoon and Fu, 1977; Reuter et al., 2006; Rueda et al., 2010; Shinagawa et al., 1991; Sun et al., 2009; Zhang, 1994). Point Distribution Models (PDM) describe shape as a collection of points on the surface of an object such as in Active Shape Models (ASMs) (Cootes et al., 1994) or Shape Context (Belongie et al., 2002). Moment descriptors such as geometric moments (Bailey and Srinath, 1996) or Zernike Moments (Zhang, 1994) describe the distribution of pixels contained within the object and hence provide a global measure of object morphology that is robust to subtle, local changes in the contour of an object. Fourier Descriptors (FDs) (Persoon and Fu, 1977; Kauppinen et al., 1995) or Curvature Scale Space (Mokhtarian and Abbasi, 2002; Rueda et al., 2010) describe the shape of the object according to the frequency components contained in the contour and are hence sensitive to subtle changes in the contour of an object; Spherical Harmonics (SPHARM) (Brechböhler et al., 1995) describe the spherical basis functions contained in the surface of an object. The Laplace–Beltrami shape descriptors (Reuter et al., 2006) and variants, such as Heat Kernels (Sun et al., 2009), describe the components of some generalized function contained within the object, for instance the Laplace–Beltrami descriptor is defined as the eigenvalues of the Laplacian operator for an object. Additionally, shape models may extract a graph representation of the object, such as the medial axis shape model (MASM) (Blum, 1967) or Reeb graphs (Shinagawa et al., 1991), these methods represent the shape of an object using local symmetry to determine a skeletal representation of the object. The MASM describes object morphology as a medial axis, points within an object that are equidistant from two or more locations on the surface of an object, and hence capture the local symmetry of an object (Blum, 1967).

Despite the wide variety of shape modeling approaches available, relatively few methods have been applied to medical imaging applications. PDMs have been proposed in the context of segmentation for cardiac, prostate, and other organs on medical imagery (Cootes et al., 1994). PDMs have been applied to anatomical structures such as brain hippocampii in conjunction with dimensionality reduction (DR) methods to identify meaningful, local changes to the shape (Sjöstrand et al., 2007; Alcantara et al., 2009). However, obtaining point correspondence is a difficult task for glands on prostate histopathology, as meaningful substructures are difficult to identify. Heat Kernels have been utilized in conjunction with

hippocampal morphology to accurately distinguish between normal patients and those with Alzheimer's disease utilizing (Pachauri et al., 2011). SPHARMs have been also employed to identify Alzheimer's disease utilizing hippocampal morphology (Gerardin et al., 2009). Similar methods have utilized SPHARMs to analyze the atrophy of other brain structures such as the putamen in patients diagnosed with Parkinson's disease (Looi et al., 2011). However no method to directly apply either SPHARMs or Heat Kernels to $2D$ shapes as been presented, although a method has been presented to map a $2D$ object into a $3D$ space and then compute SPHARMs (Sajjanhar et al., 2009). FDs have been applied to the analysis of breast lesions on mammography (Georgiou et al., 2007). Additionally, recent work has shown applications of FDs to distinguishing between normal and abnormal rotator cuffs on MRI (van Kaick et al., 2010).

The MASM, also referred to as m-rep, has been applied to medical imaging tasks including segmentation (Pizer et al., 1999; Yushkevich et al., 2006), registration (Pizer et al., 1999), and calculation of shape statistics over a population of objects (Fletcher et al., 2004). Recently, the m-rep framework has been used to guide disease classification by exploiting morphologic differences between brain substructures in order to distinguish autistic from normal patients (Gorczowski et al., 2010). MASMs have also been applied to the detection of aneurysms (Lauric et al., 2010) as well as the detection of coronary artery stenosis (Xu et al., 2012). The MASM is able to represent a wide range of object morphologies over multiple dimensions and is able to detect and represent subtle differences between objects. The MASM is the shape descriptor we chose to utilize in this work.

Due to the wide application of the MASM to many object recognition and classification tasks, several approaches to efficiently calculate the MASM have been developed (Siddiqi et al., 1999; Hassouna and Farag, 2009; Arcelli et al., 2011; Bai et al., 2007; Macrini et al., 2011; Ward and Hamarneh, 2010). Blum presented the idea of a "grassfire" approach to calculating the MASM, the idea being if a fire was set to the boundary of the object the medial axis would be where the propagating flames meet. Hassouna and Farag (2009) applied a similar technique to extract the MASM for $3D$ objects. Similarly, Siddiqi et al. (1999) presented the idea of shock graphs, calculated by considering the gradient of the level set function for an object. Alternative approaches to calculate the MASM involve iteratively thinning the surface of the object (Arcelli et al., 2011). Some approaches have incorporated pruning, removal of extraneous regions of the axis from the MASM. For instance Dynamic Contour Evolution (DCE), involves first calculating the MASM and then pruning regions on the MASM that lead to small partitions of the contour (Bai et al., 2007). Pruning of the MASM can be performed by removing branches which have similar locations and directions as nearby branches (Macrini et al., 2011). A groupwise medial axis transform was presented in (Ward and Hamarneh, 2010) to retain those branches that are consistent across a set of objects while removing those branches that are not preserved across the group.

Calculating similarity between MASMs can be difficult as there is not always a clear correspondence between regions on medial axes. Several approaches have been presented to overcome this problem (Siddiqi et al., 1999; Macrini et al., 2011; Ward and Hamarneh, 2010; Chang and Kimia, 2011; Sebastian et al., 2004; Bai and Latecki, 2008; Goh, 2008; Ruberto, 2004). For instance comparing MASMs may be formalized as a graph matching problem, where the MASM is broken into a set of medial axis branches; shape similarity is calculated as the summation of similarity between matched branches (Siddiqi et al., 1999). This method can be extended to utilize additional attributes about each branch, such as branch curvature, when calculating MASM similarity (Ruberto, 2004; Macrini et al., 2011). The graph matching problem between MASMs has also been previously formulated for $3D$ structures, where the matching is between hypergraphs which may include $2D$ medial surfaces (Chang and Kimia, 2011). Approaches have also been presented to hierarchically

match and merge regions to quantify MASM similarity (Goh, 2008). Sebastian et al. (2004) introduced the concept of edit-distance, where MASM similarity is calculated as a combination of three possible edits: branch matching, branch removal, and branch addition; where each edit is assigned a similarity value.Bai and Latecki (2008) considered the similarity of paths between end nodes (locations on the MASM that represent where a branch end) to determine similarity. Alternatively, a method to fit a MASM to an object using only likely, predetermined MASM templates have been presented (Yushkevich et al., 2006); MASM similarity is calculated as a difference of distance transforms (Ward and Hamarneh, 2010).

Researchers have recently begun using non-linear DR (NLDR) schemes in conjunction with shape descriptors (Bai et al., 2010; Egozi et al., 2010; Qian et al., 2010). Such methods extract a small set of features which describe the variation in morphology between different objects. For instance, Bai et al. (2010) presented a semi-supervised framework, Graph Transduction, to learn a set of discriminating shape descriptors for content-based image retrieval (CBIR) applications. However, the set of shape descriptors learned is dependent on the query object, suggesting that it may not be naturally extensible to classification problems. A *k*-nearest neighbor approach to finding object similarity in the high dimensional shape space was presented in Egozi et al. (2010). Qian et al. (2010) presented a NLDR scheme to determine relevant morphologic differences between vertebrae, exploiting the definition of the Procrustes shape space. However, their methodology is only applicable to objects represented by PDMs. Nasreddine et al. (2010) utilized a geodesic distance function to distinguish between contours for pairs of objects. The use of NLDR schemes typically improves the ability to measure similarity in object morphology compared to the Euclidean distance in CBIR applications (Bai et al., 2010; Egozi et al., 2010; Qian et al., 2010).

In this paper, we present a method to quantify differences in object morphology. We extract a set of explicit shape descriptors (ESDs) to quantitatively represent object morphology. ESDs calculation involves: (a) utilizing the MASM to quantify object morphology (Blum, 1967), and (b) obtaining a small set of ESDs via the unsupervised NLDR scheme Graph Embedding (GE) (Shi and Malik, 2000). Although each of the steps involved in ESD calculation has been previously reported in the computer vision literature (Blum, 1967; Shi and Malik, 2000), the combination of a shape model (MASM) with a NLDR scheme (GE) represents a novel combination of two well documented methods which allows for extraction of a concise set of shape descriptors that enable discrimination between objects with subtle shape differences.

Fig. 3 presents an overview of our framework and the constituent modules. The novelty of our methodology is twofold. (1) In order to overcome the correspondence issues between MASMs, a diffeomorphic based similarity (DBS) measure (Sparks and Madabhushi, 2010) is used to register MASMs and determine point correspondence between the registered MASMs. A non-rigid registration scheme is used to ensure that MASMs with large topological differences can be registered into a common coordinate frame, while preserving the relationships between regions on the MASM. (2) GE is utilized to obtain a low dimensional representation of morphologic features which can then be employed in conjunction with a classifier (e.g. Support Vector machines (SVMs)) for shape classification. GE allows for ESDs to be defined for a set of unlabeled objects utilizing only a shape dissimilarity matrix. GE utilizes an eigenvalue decomposition of the shape dissimilarity matrix to find a small set of shape descriptors which optimally preserve pairwise relationships between objects. ESDs are defined as the eigenvectors of the shape space obtained from GE.

3. Shape representation model

3.1. Notation

An image scene $\mathscr{C} = (C, \xi)$ is defined by the d-dimensional grid of voxel locations C and an object contour ξ. $c \in C$ represents a voxel defined by a d-dimensional vector that describes its location in C. ξ partitions C into a foreground region Ω_f, belonging to an object of interest, and a background region Ω_k such that $\Omega_k \bigcup \Omega_f = C$. Table 1 describes the notation and symbols that appear frequently in this paper.

3.2. Medial axis shape model construction

The MASM was employed to concisely and explicitly describe the local morphology of an object described by the contour ξ (Blum, 1967). We define the MASM as $\mathscr{M} = (M, v_1, v_2)$ where $M \in C$ is a set of voxels on the medial axis and $v_1(m)$, $v_2(m)$: $m \in M$ are two functions defining vectors to the first and second closest points on ξ. To find M, we calculate the gradient magnitude squared of the signed distance function defined as,

$$\widehat{f}(c) = \sum_{i=1}^{d} \left(\frac{\partial f(c)}{\partial X_i} \right)^2, \quad (1)$$

where $f(c)$ is the signed distance function evaluated over $c \in C$ and $\frac{\partial(\cdot)}{\partial X_i}$ is the partial gradient along X_i corresponding to the ith direction. For a d-dimensional image X_i is found for all $i \in \{1, \ldots, d\}$. We use $\widehat{f}(c)$ to define the medial atoms as $M = \left\{ m : m \in C, \widehat{f}(m) < \tau \right\}$. Empirically, we determined that $\tau = 0.05 \text{argmax}_{c \in C} \left(\widehat{f}(c) \right)$ yields a well defined medial axis. Section 5.3 describes the approach we employed to assess whether a MASM was able to accurately and quantitatively describe the morphology of a given object. The use of τ helps avoid spurious branches on the MASM.

The surface vector functions $v_1(m)$ and $v_2(m)$: $m \in M$, are calculated as $v_1(m) = \widehat{p}_1(m) - m$ and $v_2(m) = \widehat{p}_2(m) - m$, where $\widehat{p}_1(m)$ and $\widehat{p}_2(m)$ represent the two closest points on ξ to m. For m with more than two closest points on the ξ, we chose $\widehat{p}_1$ and $\widehat{p}_2$ that maximize the angle between $v_1(m)$ and $v_2(m)$.

4. Framework for quantifying shape differences

Dissimilarity between a set of N MASMs $\mathbf{M} = \{\mathscr{M}_1, \ldots, \mathscr{M}_N\}$ is quantified by comparing differences in corresponding medial atoms for each pair of MASMs. The parameters of a MASM, $\mathscr{M}_a : a \in \{1, \ldots, N\}$, include M_a, $v_{1,a}$, and $v_{2,a}$ (defined in Section 3.2). We determine correspondence between all $m_a \in M_a$ and all $m_b \in M_b$: $b \neq a$, $b \in \{1, \ldots, N\}$. By this procedure we attempted to implicitly determine correspondence between the surface vector functions $v_1(m_a)$ and $v_1(m_b)$ as well as $v_2(m_a)$ and $v_2(m_b)$. To determine correspondence a two step registration was performed by: (1) affine registration of M_a onto M_b, followed by (2) a diffeomorphic registration of M_a onto M_b. These steps are described in more detail in Sections 4.1 and 4.2, respectively. The affine registration provides a rough alignment of corresponding regions on the MASMs and is necessary for the initialization of the diffeomorphic registration process. Once M_a and M_b have been accurately registered, medial atom correspondence between M_a and M_b is determined and used to calculate the dissimilarity between $\mathscr{M}_a$ and $\mathscr{M}_b$. Section 4.3 describes the calculation of MASM dissimilarity.

4.1. Medial atom affine registration

Affine registration between M_a and M_b is determined by applying the Iterative Closest Point (ICP) algorithm (Zhang, 1994). Point correspondence between M_a and M_b is determined by,

$$(\hat{m}_a, \hat{m}_b) = \underset{\hat{m}_a \in M_a, \hat{m}_b \in M_b}{\operatorname{argmin}} \|m_a - m_b\|, \quad (2)$$

where $\hat{m}_a \in M_a$ and $\hat{m}_b \in M_b$ are the set of corresponding points on M_a and M_b. An affine transformation T^{ab} is found by minimizing the following function,

$$T^{ab} = \operatorname{argmin}_{T^{ab}} \left(\sum_{\hat{m}_a, \hat{m}_b} \|\hat{m}_b - T^{ab}(\hat{m}_a)\| \right). \quad (3)$$

Estimation of point correspondences (Eq. (2)) and the affine transformation (Eq. (3)) are iteratively applied to M_a until point correspondences, $(\hat{m}_a, \hat{m}_b)$, remain unchanged between iterations. The resulting medial axis $\widehat{M}_a = T^{ab}(M_a)$ is affinely registered to M_b.

4.2. Medial atom-based non-rigid registration

For two sets of medial atoms $\widehat{M}_a$ and M_b, which are registered as described in Section 4.1, a diffeomorphic registration is then applied to further align $\widehat{M}_a$ and M_b. We utilized a variation of the diffeomorphic registration method proposed by Guo et al. (2006). A brief overview of our non-rigid registration is as follows: (1) corresponding locations between $\widehat{M}_a$ and M_b are estimated using a deterministic annealing K-means clustering algorithm (Rose et al., 1990) (described in greater detail in Section 4.2.1); (2) A diffeomorphic transformation over the image space C is calculated to minimize the distance between the estimated corresponding locations on $\widehat{M}_a$ and M_b. A diffeomorphic transformation was used to ensure a continuous and differentiable transformation field so that the underlying relationship between all $\hat{m}_a \in \widehat{M}_a$ and all $m_b \in M_b$ are preserved. In Section 4.2.2 we discuss the calculation of the diffeomorphic transformation in greater detail.

4.2.1. Correspondence estimation

Individual medial atom correspondence between $\widehat{M}_a$ and M_b may be difficult to determine accurately. We therefore determine corresponding locations on $\widehat{M}_a$ and M_b via a deterministic annealing K-means clustering algorithm (Rose et al., 1990). We define a set of K cluster centroids at the jth iteration of our registration method as $g_a^{k,j} : k \in \{1, \ldots, K\}$ for all $m_a \in \widehat{M}_a$. Similarly, for M_b we define a set of cluster centroids $g_b^{k,j} : k \in \{1, \ldots, K\}$. The cluster centroids are initialized such that $g_a^{k,0}$ and $g_b^{k,0}$ represent corresponding points on M_a and M_b. As the clusters are used to determine correspondence between MASMs, K is constrained to be the same for both $\widehat{M}_a$ and M_b.

We estimate the probability $P\left(m_a \middle| g_a^{k,j}, \sigma_j\right)$ of a medial atom $\hat{m}_a \in \widehat{M}_a$ belonging to the $g_a^{k,j}$ cluster as:

$$P\left(m_a|g_a^{k,j},\sigma_j\right)=\frac{e^{-\sigma_j\|m_a-g_a^{k,j}\|^2}}{\sum_{k=1}^{K}e^{-\sigma_j\|m_a-g_a^{k,j}\|^2}}. \quad (4)$$

Similarly, the probability of $m_b \in M_b$ belonging to the cluster $g_b^{k,j}$ is given as $P\left(m_b\left|g_b^{k,j},\sigma_j\right.\right)$. The term $e^{-\sigma_j}\|m_a-g_a^{k,j}\|^2$ assigns higher values to medial atoms near the centroid $g_a^{k,j}$ and lower values to medial atoms farther away; σ_j determines which medial atoms are considered near and far from the centroid. Convergence of the membership function is enforced by setting $\sigma_j = (\zeta)^j\sigma_0$ where $\zeta > 1$, hence at each iteration the clustering algorithm considers a smaller region to be near the cluster centroid. Therefore at each step in the algorithm fewer medial atoms have a non-zero probability of belonging to the cluster defined by the centroid $g_a^{k,j}$. Ultimately when j is very large, each medial atom is assigned membership (a non-zero probability) to one cluster centroid. The initial weighting term is set as $\frac{1}{\sigma_0}=\max_{m_a\in\widehat{M}_a}\|m_a-\mu_a\|+\max_{m_b\in M_b}\|m_b-\mu_b\|$.

Cluster centroids are updated according to the probability of all $m_a \in \widehat{M}_a$ belonging to the cluster $g_a^{k,j}$. The cluster $g_a^{k,j}$ is updated by the equation,

$$\overline{1+\sum_{m_a\in M_a}P\left(m_a\left|g_a^{k,j},\sigma_j\right.\right)}. \quad (5)$$

The term $g_b^{k,j}$ defines a centroid on M_b that corresponds to the centroid $g_a^{k,j}$. By taking the average of the two locations $g_a^{k,j}$ and $g_b^{k,j}$ in Eq. (5), we ensure that the $g_a^{k,j+1}$ remains in a location on M_a which is proximal to $g_b^{k,j}$ located on M_b. The centroid $g_b^{k,j+1}$ is determined by a similar equation.

4.2.2. Correspondence registration

The goal of correspondence registration is to find a diffeomorphic transformation Q which best maps the cluster centroids $g_a^{k,j}$ onto $g_b^{k,j}$ for $k \in \{1, \ldots, K\}$. Q is defined as $\{Q(t): t \in \{0, \ldots, t_{\max}\}\}$ where $Q^k(0) = g_a^{k,j}$ and $Q^k(t_{\max}) = g_b^{k,j}$. Hence Q will enable alignment of $g_a^{k,j}$ to $g_b^{k,j}$. Similar to Twining et al. (2002), we use a linear piecewise approximation to solve the energy minimization function,

$$\widehat{Q}=\underset{\widehat{Q}}{\operatorname{argmin}}\sum_{k=1}^{K}\sum_{t=0}^{t_{\max}}\omega(t)\cdot\left(\sum_{\eta=1}^{K}\omega^{\eta}(t)\,G\left(Q^{\eta}(t),Q^{k}(t)\right)\right), \quad (6)$$

where the kernel function G is defined as Green's function: $G(\alpha, \beta) = -(\alpha - \beta)^2\log(\alpha - \beta)^2$ (Bookstein, 1989). Green's function ensures that Q will be smoothly varying over C.

To solve Eq. (6), an optimization of Q and the variables $\omega(t)$ and $\omega^{\eta}(t)$ can be found in an iterative fashion (Twining et al., 2002). This optimization is performed by first holding $\omega(t)$ and $\omega^{\eta}(t)$ constant and using gradient descent to find the optimal Q and then repeating the procedure with Q held constant.

Both the correspondence estimation (Eqs. (4) and (5)) and correspondence registration (Eq. (6)) are iterated until a user defined threshold, Y, is reached by the annealing parameter σ_j. The overall algorithm is as follows:

Algorithm 1

DiffReg

Input: $\hat{M}_a$, M_b

Output: $\tilde{M}a$

begin

1. Initialize σ_j, $g_a^{k,j}$, $g_b^{k,j}$
2. *while* $\sigma_j < Y$
3. Update $P\left(m_a \mid g_a^{k,j}\right)$, $P\left(m_b \mid g_b^{k,j}\right)$ by Eq. (4).
4. Update $g_a^{k,j}$, $g_b^{k,j}$ by Eq. (5).
5. Update Q by Eq. (6).
6. $\tilde{M}a = Q\left(\hat{M}_a\right)$
7. *end*

end

The two sets of cluster centroids are initialized to be equal $g_a^{k,0}=g_b^{k,0}$ and located at $g_a^{k,0}=\frac{\mu_a+\mu_b}{2}+\epsilon$, where ϵ is a random variable with a very small value ($\epsilon \approx 10^{-1}$). The term ϵ is added to ensure that each pair of corresponding cluster centroid $g_a^{k,0}, g_b^{k,0}$ has a unique starting location, and hence will diverge from the other cluster pairs as σ_j increases.

4.3. Medial atom correspondence and shape dissimilarity

Given two medial axes $\tilde{M}_a$ and M_b registered into a common coordinate frame, we determine point correspondence between $\tilde{M}_a$ and M_b as,

$$(\hat{u},\hat{v}) = \underset{\hat{u}\in\tilde{M}_a,\hat{v}\in M_b}{\operatorname{argmin}} \|u-v\|. \quad (7)$$

The set of corresponding medial atoms, $(\hat{u},\hat{v})$ determined via Eq. (7) are then used to calculate dissimilarity between $\tilde{M}_a$ and M_b as,

$$A_{ab}=\sum_{\hat{u},\hat{v}} \left(\|\hat{u}-\hat{v}\|+\|v_{1,a}(\hat{u})-v_{1,b}(\hat{v})\|+\|v_{2,a}(\hat{u})-v_{2,b}(\hat{v})\|\right). \quad (8)$$

For all $\mathscr{M} \in \mathbf{M}$, a dissimilarity matrix $A \in \mathbb{R}^{N\times N}$ is constructed such that it represents a N-dimensional space corresponding to morphologic dissimilarity between all $\mathscr{M} \in \mathbf{M}$.

4.4. Feature extraction via non-linear dimensionality reduction

GE (Shi and Malik, 2000) is applied to the dissimilarity matrix A, which measures the dissimilarity between all all $\mathscr{M} \in \mathbf{M}$, to yield a set of ESD features in a low dimensional

space. Specifically, a n-dimensional embedding is learned from the N-dimensional matrix A, where $n \ll N$. The ESD features for all $\mathscr{M} \in \mathbf{M}$ are defined as $\mathbf{y}' = [y_1, \ldots, y_N]$ where $\mathbf{y}$ represents the top n eigenvectors for the shape space defined by A. $\mathbf{y}$ can be found by minimizing the pairwise reconstruction error measured as,

$$\mathbf{y}=\underset{\mathbf{y}}{\operatorname{argmin}}\sum_{a=1}^{N}\sum_{b=1}^{N}\|y_a - y_b\|^2 W_{ab}, \quad (9)$$

where $W_{ab} = e^{-A_{ab}/c}$. The term γ is used to normalize A and is dataset specific. The value of γ was determined empirically as described in Section 5.6. Eq. (9) can be rewritten as,

$$\underset{\mathbf{y}}{\operatorname{argmin}}\sum_{a=1}^{N}\sum_{b=1}^{N}\left(y_a^2+y_b^2 - 2y_a y_b\right) W_{ab}. \quad (10)$$

Assuming $W_{ab} = W_{ba}$, which will be true since $A_{ab} = A_{ba}$, then Eq. (10) reduces to,

$$\underset{\mathbf{y}}{\operatorname{argmin}}2\sum_{a=1}^{N}\sum_{b=1}^{N}\left(y_a^2 - y_a y_b\right) W_{ab}. \quad (11)$$

The minimization problem can be further simplified by introducing a diagonal matrix defined as $D_{aa} = \sum_b W_{ab}$, making the minimization problem,

$$\underset{\mathbf{y}}{\operatorname{argmin}}\ 2\mathbf{y}'\left(D - W\right)\mathbf{y}. \quad (12)$$

Eq. (12) is equivalent to the minimum eigenvalue decomposition equation,

$$W\mathbf{y}=\lambda D\mathbf{y}, \quad (13)$$

where the top n eigenvalues in λ correspond to the n eigenvectors $\mathbf{y}$ and the top n ESD features. The top n eigenvalues correspond to the projection of the matrix A into the space $\mathbb{R}^{N\times n}$ such that the pairwise distances between the elements in A, and hence the pairwise distances between objects, are preserved. Furthermore as the eigenvectors $\mathbf{y}$ are orthonormal to each other, each additional feature provides independent information on the shape space represented by A.

4.5. Support vector machine classification

A SVM classifier (Cortes and Vapnik, 1995) can be trained using $\mathbf{y}'$, to learn the optimal hyperplane which separates $\mathbf{M}$ into the classes referenced by the label set $L(\mathbf{M}) \in \{-1, 1\}$.

The SVM classifier utilizes y'_a to determine the distance to the hyperplane $\Phi(\mathscr{M}_a)$. SVM classifiers are typically used to generate a hard class decision where $\Phi(\mathscr{M}_a) < 0$ corresponds to assigning class label of −1 to $\mathscr{M}_a$. However, a pseudo-threshold can be generated by varying the decision boundary.

Given a specific decision boundary, ϱ, if $\Phi(\mathscr{M}_a) < \varrho$ and $L(\mathscr{M}_a) = -1$ then $\mathscr{M}_a$ is identified as a true negative (TN); if $\Phi(\mathscr{M}_a) < \varrho$ and $L(\mathscr{M}_a) = 1$ then $\mathscr{M}_a$ is identified as a false negative (FN); if $\Phi(\mathscr{M}_a) > \varrho$ and $L(\mathscr{M}_a) = 1$ then $\mathscr{M}_a$ is identified as false positive (FP); and if $\Phi(\mathscr{M}_a) > \varrho$ and $L(\mathscr{M}_a) = 1$ then $\mathscr{M}_a$ identified as a true positive (TP). The number TN ($N_{TN,\varrho}$), FN ($N_{FP,\varrho}$), FP ($N_{FP,\varrho}$),and TP ($N_{TP,\varrho}$) are calculated over a range of ϱ.

For each ϱ, sensitivity (SN_ϱ), specificity (SP_ϱ), and classifier accuracy (CA_ϱ) can be calculated as,

$$SN_\varrho = \frac{N_{TP,\varrho}}{N_{TP,\varrho}+N_{FN,\varrho}}, \quad SP_\varrho = \frac{N_{TN,\varrho}}{N_{TN,\varrho}+N_{FP,\varrho}}, \quad \text{and } CA_\varrho = \frac{N_{TP,\varrho}+N_{TN,\varrho}}{N}$$

where N is the total number of objects in the database. By plotting SN_ϱ versus $1 - SP_\varrho$ over a range of ϱ a Receiver Operating Characteristic (ROC) curve representing the trade-off between SN and SP for a give feature set is obtained. Area under the ROC curve (AUC) is calculated for each ROC curve.

For our experiments, the SVM classifier was employed with a radial basis function. For the dataset evaluated for our experiments there were three prostate histopathology classes. To evaluate the prostate histopathology dataset using a SVM the classification task was divided into 4 pairwise classification tasks. The training and evaluation involved a randomized 3-fold cross-validation scheme where at each iteration, 2/3 of the dataset was used for training of the classifier, while always maintaining class balance. The remaining 1/3 of the dataset was used for independent testing of the SVM classifier. Training and testing sets were selected such that the training and testing sets never concurrently contained images from the same patient. The cross-validation procedure was repeated 5 times where, at each iteration, the training and testing sets were selected randomly always ensuring that there was no overlap between the training and testing sets in terms of patients.

5. Experimental design and results

5.1. Dataset description

5.1.1. Synthetic super quadratic ellipsoids—Super quadratic ellipsoids represent a class of $3D$ objects with a closed topology such that the shape of a super quadratic ellipsoid is fully determined by the 5 parameters: α_1, α_2, α_3, ϵ_1, and ϵ_2 (Franklin and Barr, 1981). Super quadratic ellipsoids with similar shapes were generated by carefully modulating the model parameters so that the discriminability of the ESD features in a synthetic setting could be evaluated. The boundary of a super quadratic ellipsoid is defined as,

$$S(x_1, x_2, x_3) = \begin{cases} x_1 = \alpha_1 \cos(\theta) \cos(\psi)^{\epsilon_1} \\ x_2 = \alpha_2 \sin(\theta) \cos(\psi)^{\epsilon_2} \\ x_3 = \alpha_3 \cos(\theta) \sin(\psi), \end{cases} \quad (14)$$

where $\theta \in \{-\pi, \ldots, \pi\}$ and $\psi \in \{-\pi/2, \ldots, \pi/2\}$ (Franklin and Barr, 1981). For the purpose of this study the following parameters were combined: $\epsilon_1, \epsilon_2 \in \{0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4\}$ and $\alpha_1, \alpha_2, \alpha_3 \in \{0.7, 0.8, 0.9, 1\}$ resulting in 4096 possible shape combinations. Note however, that several shapes will be scaled versions of each other (e.g. $\epsilon_1 = \epsilon_2 = 0.5$, $\alpha_1 = \alpha_2 = \alpha_3 = 0.7$ and $\epsilon_1 = \epsilon_2 = 0.5$, $\alpha 1 = \alpha_2 = \alpha_3 = 0.8$ will be the same shape with only a scale difference). Two parameters, ϵ_1 and ϵ_2, are referred to as shape parameters and were used to control the concavity/convexity of the super quadratic ellipsoids. The other three parameters, α_1, α_2, and α_3, are scaling parameters and determine the length, width, and depth of the object respectively. By selecting parameters that are close together, we obtain a set of objects with subtle shape differences.

5.1.2. Prostate histopathology—Prostate tissue biopsy cores obtained from 58 patient studies were stained with Hemotoxylin and Eosin (H & E) and digitized using a ScanScope CS™ whole-slide scanning system at 40× optical magnification. An expert pathologist

selected regions of interests on the digitized biopsy image, to obtain a total of 102 regions. The expert pathologist then classified each region as benign (24 regions), Gleason grade 3 (67 regions), or Gleason grade 4 (11 regions). Every gland contained within each region was segmented by a human expert to obtain lumen and nuclear boundaries (Fig. 1). Glands which did not contain either a nuclear or lumen boundary, or where the contour was not fully contained within the region were removed from the study, resulting in a total of 888 glands containing both lumen and nuclear boundary segmentations. These glands were distributed across the three classes: benign ($N = 93$), Gleason grade 3 ($N = 748$), and Gleason grade 4 ($N = 47$).

5.2. Features for comparison against ESDs

Our novel ESD features are compared against three morphologic feature sets: Boundary-based features (referred to as Boundary) (Agner et al., 2011), FDs (Persoon and Fu, 1977), and a MASM path similarity measure (referred to as Path) (Bai and Latecki, 2008). Below, we briefly describe the calculation of each of these feature sets.

5.2.1. Boundary-based features

The Boundary feature set consists of 6 morphologic features that have been previously used with computer-aided diagnostic (CAD) systems for determining Gleason grade using prostate gland morphology (Tabesh et al., 2007). The formulation for each of the shape features is presented in Table 2 and reflects the (a) circularity of an object (area overlap ratio, compactness), (b) how much does object contour vary with respect to the shape of a circle (normalized average radial distance ratio, standard deviation of distance ratio, variance of distance ratio), and (c) how quickly does the object contour change (smoothness).

5.2.2. Fourier descriptors

The FD feature set comprised the first 50 frequency components calculated from the contour of an object ξ (Persoon and Fu, 1977). The frequency of the contour was calculated as follows: a set of ordered points around the contour $p^{(j)} \in \xi$, $j \in \{1, \ldots, J\}$ were found. The magnitude of the points was calculated as $\rho^j = \|p^j - \bar{p}\|$ where $\bar{p} = \frac{1}{|\xi|}\sum_{p \in \xi} p$. The Fourier transform of ρ is calculated and is used to derive the first 50 frequency components of each contour.

5.2.3. Path features

We calculated an alternative shape dissimilarity matrix, Z, using a path-based measure of MASM dissimilarity previously presented by Bai and Latecki (2008). This MASM similarity measure has been demonstrated to perform similarly or better than edit-distances on shock graphs (Sebastian et al., 2004). Given two MASMs $\mathcal{M}_a$ and $\mathcal{M}_b$, a set of medial atoms $m'_a \in M_a$ and $m'_b \in M_b$ comprising the end nodes of the medial axes are identified. End nodes are defined as those medial atoms with only one neighbor on the medial axis. A path between pairs of end nodes $m_a^{'w} \in M_a$ and $m_a^{'z} \in M_a$ is defined as $\nu\left(m_a^{'w}, m_a^{'z}\right)$. Similarity, $\nu\left(m_b^{'\widehat{w}}, m_b^{'\widehat{z}}\right)$ is defined for end node pairs in M_b. The dissimilarity between M_a and M_b is then found by,

$$Z(a,b) = \min_{w,z,\widehat{w},\widehat{z}} \left[\Delta\left(\nu\left(m_a^{'w}, m_a^{'z}\right), \nu\left(m_b^{'\widehat{w}}, m_b^{'\widehat{z}}\right)\right)\right]. \quad (15)$$

The function Δ gives a measure of path similarity which is defined as the summation over the radius and path length (Bai and Latecki, 2008). The dissimilarity matrix Z is a high dimensional representation of the shape space, and as with our dissimilarity matrix A, GE was employed to return the top n eigenvectors.

5.3. Experiment 1: ability of MASM to capture object morphology

We tested the hypothesis that the MASM can accurately represent the morphology of a wide variety of shapes. We assumed that MASM reconstruction accuracy reflects the accuracy of the MASM to describe object shape. To evaluate this quantitatively, we reconstructed all of the objects in each dataset. For each dataset we varied the number of medial atoms contained in the MASM and for each set of medial atoms we reconstructed the object and denoted it as Ω_r. Ω_r is determined as a set of pixels belonging to an object given $\mathscr{M}$ such that,

$$\Omega_r = \{c : \|m - c\| < r(m), m \in M, c \in C\} \quad (16)$$

where the function $r(m)$ is defined by the equation,

$$r(m) = \frac{\| (m - v_1(m)) \| + \| (m - v_2(m)) \|}{2}. \quad (17)$$

For each Ω_r, we measured how close it is to Ω_f using the edge based measures-(a) mean absolute distance (MAD) and (b) Hausdorff distance (Huttenlocher et al., 1993); and the area based measures-(a) Dice's coefficient (DICE) (Dice, 1945) and (b) Positive Predictive Value (PPV). MAD describes on average the extent of variation between the ground truth shape contour and the reconstructed shape and is formally defined as,

$$\frac{1}{|p|} \sum_{p \in \Omega_f} \min_{o \in \Omega_r} \|p - o\|. \quad (18)$$

Hausdorff distance (Huttenlocher et al., 1993) measures the performance of the worst case disparities between two shapes and is defined as,

$$\max_{p \in \Omega_f} \left(\min_{o \in \Omega_r} \|p - o\| \right). \quad (19)$$

DICE (Dice, 1945) is a measure of overlap between two shapes, in this case it reflects the extent of overlap between the reconstructed shape and the ground truth shape and is defined as,

$$\frac{|\Omega_f \cap \Omega_r|}{|\Omega_f| + |\Omega_r|}. \quad (20)$$

PPV in this case is used to evaluate the proportion of pixels in the reconstructed shape accurately identified as belonging to the foreground of the object and is defined as,

$$\frac{\Omega_r \cap \Omega_f}{|\Omega_r|}. \quad (21)$$

For each dataset the fewest number of medial atoms that achieved high DICE and PPV were selected to represent all objects in the database for calculation of ESDs. Fig. 4 illustrates an example of the medial atom evaluation performed for the prostate histopathology dataset, with a red cross-displayed at the optimal number of medial atoms for representing prostate

morphology. The object reconstruction accuracy for the optimal number of medial atoms is reported in Table 3. As seen in Table 3 the synthetic dataset which contains $3D$ objects needs more medial atoms to more accurately represent the morphology compared to the prostate histopathology dataset which contains $2D$ objects. This is to be expected since $3D$ objects can have more complex shapes compared to their $2D$ counterparts.

5.4. Experiment 2: registration evaluation

In this experiment the diffeomorphic registration algorithm presented in Section 4.2 was evaluated in terms of its ability to (a) recover a large range of non-linear deformations applied to MASMs and (b) determine accurate correspondences between medial atoms on M_a and M_b. We conducted a total of 200 experiments in which 20 randomly chosen MASMs and 10 deformation fields were considered. A deformation field denoted as T' was generated by varying the type and magnitude of the deformation applied. We applied T' to the image space C containing $\mathscr{M}_a$ and then used our diffeomorphic registration algorithm to approximate the inverse transformation T^{-1}. We then computed the mean residual error as $\frac{1}{|M_a|}\sum_{m\in M_a}=\|m - T^{-1}\left(T'(m)\right)\|$, where $|M_a|$ is the number of medial atoms contained in $\mathscr{M}_a$.

Mean residual error was on average 1.09 ± 0.24 pixels, where accurate cluster centroid correspondence between MASMs allowed for close alignment between MASMs. The worst case registration for a given MASM was 4.94 ± 2.22 pixels. For this specific case the MASM had several branches so that the cluster centroids on the original MASM and the deformed MASM failed to correspond to the equivalent locations on the original MASM. Incorrect correspondence determination between the cluster centroids may cause the diffeomorphic registration to be unable to approximate T^{-1}. In these 10 cases, on average 2.8 ± 0.3 cluster centroids did not have correct correspondence. This effect was only seen in 10 of 200 MASMs and only during the application of large deformations; deformations that had a magnitude greater than 10% of the area of Ω_f, the foreground region of the object of interest.

5.5. Experiment 3: distinguishing between super quadratic ellipsoids with differing shape parameters

We constructed a set of 4096 super quadratic ellipsoids, denoted by S, to evaluate the ability of our ESD features to represent subtle shape variations between objects (see Section 5.1.1). Dissimilarity between the pair S_a and S_b for known sets of shape parameters was measured as $\Pi(S_a, S_b) = \sum_{\varpi}\| \varpi_a - \varpi_b\| : \varpi \in \{\alpha_1, \alpha_2, \alpha_3, \epsilon_1, \epsilon_2\}, a, b \in \{1, \ldots, N\}$. $\Pi(S_a, S_b)$ represents the total shape dissimilarity.

The correlation between a set of ESD features $\mathbf{y'} = [y_1, \ldots, y_N]$ for a set of N objects and a set of known shape differences Π was calculated as follows. We first define the distance between two objects in the ESD feature space as, $\Xi(S_a, S_b) = \|y_a - y_b\|$. We then calculate Pearson's correlation coefficient between Π and Ξ (Rodgers and Nicewander, 1988). This allows us to quantitatively evaluate the ability for the ESD feature space, represented by $\Xi(S_a, S_b)$, to reflect known shape differences (Π).

Pearson's correlation coefficient was determined to be R = 0.82, demonstrating a strong correlation between the know shape parameters and the underlying ESD feature space. Fig. 5 displays all objects in the first 2 dimensions of the ESD feature space with representative S displayed in Fig. 5b–i. Note that a curvilinear manifold that contains the subspace of these shapes is clearly visible. The first dimension of the ESD feature space correlates to changes in ϵ_1, corresponding to the red line. The second dimension of the ESD feature space

correlates to ϵ_2, corresponding to the blue line. Note that similar super quadratic ellipsoids are embedded adjacent to each other in the feature space while dissimilar super quadratic ellipsoids are embedded far apart. These results on synthetic data suggest that ESDs are able to differentiate between subtle changes in shape.

5.6. Experiment 4: Gleason grading of prostate histopathology

We evaluated the ability of four feature sets (Boundary, FD, Path, ESD) to accurately distinguish between Gleason grade 3 (G3), grade 4 (G4), and benign (BE) prostate glands as seen on histopathology using a SVM classifier. As this is a multiclass problem we evaluated SVM classifiers for the following 4 pairwise classification tasks: BE versus other (G3 and G4), G3 versus other (BE and G4), G4 versus other (BE and G3), and G3 versus G4. SVM training and evaluation was performed as described in Section 4.5.

The ESD and Path feature sets were evaluated over $1 \leq n \leq 30$ and $1 \leq \gamma \leq 2000$ for each classification problem (results not shown). SVM CAs and corresponding n and γ for each features set is shown in Table 4a. For all classification problems, $n = 4$ was empirically determined to yield the consistently best results in the ESD feature space. A narrow range of $1 \leq \gamma \leq 2$ was identified as yielding the best performance in the ESD feature space, however γ can be adjusted to obtain better performance for a specific classification task. In contrast, Path had a wide range of n and γ which yielded high CA and AUC values. For all classification problems considered, the ESD features outperformed the Boundary, FD, and Path feature sets. ROC curves for each classification task are displayed in Fig. 6 and corresponding AUC values are reported in Table 4b.

Fig. 7 displays the first 2 ESD features of the 888 glands; representative glands from three different classes are shown along with their corresponding locations in the ESD feature space. Misclassified glands, shown in the far right row often display characteristics very similar to glands of other Gleason patterns. Consequently, some misclassifications may be attributed to glands displaying atypical attributes.

6. Concluding remarks

In this paper, we presented Explicit Shape Descriptors (ESDs) for use in quantifying morphologic differences between prostate glands on histopathology. ESDs are calculated by (a) fitting a medial axis shape model (MASM), (b) calculating diffeomorphic based similarity (DBS), and (c) applying Graph Embedding (GE) to the shape dissimilarity matrix to find a set of ESDs. The individual modules contained in our ESD calculations (e.g. MASM (Blum, 1967), GE (Shi and Malik, 2000)) have been previously described, however, our methodology represents a novel integration of each of these methods in order to describe object morphology with a concise set of features. ESDs were able to distinguish between subtle differences in super quadratic ellipsoids and were also able to distinguish between prostate glands on histopathology with subtle morphologic differences with a maximum accuracy of 89% for 888 prostate glands acquired from 58 patient needle core biopsies.

Our framework offers distinct advantages compared to previously reported methods which combine shape models with NLDR. Our approach, unlike that of previously presented work (Bai et al., 2010; Egozi et al., 2010), is unique in that it can be applied to data where class information is not known. ESDs are derived from the eigenvector decomposition of the high dimensional shape dissimilarity matrix and hence are guaranteed to be of low dimensionality while simultaneously preserving pairwise class relationships between objects. For the datasets considered in this work, ESDs were better able to accurately capture morphologic differences between objects compared to other classes of feature descriptors including

boundary-based features (Tabesh et al., 2007), Fourier Descriptors (FDs) (Persoon and Fu, 1977), and MASM path similarity (Bai and Latecki, 2008).

One limitation of our current framework is that GE and the subsequent SVM classifier training must be recalculated when new, never before seen objects are introduced into the dataset. Hence, GE learns the low dimensional embeddings from all samples, training and testing, included in the dataset. However, GE is an unsupervised NLDR approach and therefore no label information is incorporated into learning the low dimensional space. Recalculating the low dimensional embeddings when new samples are introduced into the dataset, commonly referred to as the out-of-sample problem, is a well known limitation of NLDR schemes such as GE (Bengio et al., 2003). However, methods to overcome this problem have been previously presented (Bengio et al., 2003; Fowlkes et al., 2004). Future work will involve incorporating and evaluating these methods in the context of ESD feature extraction.

GE, in addition to other NLDR schemes, is known to be sensitive to noise and outliers in the original high dimensional space (Belkin et al., 2006; Chang and Yeung, 2006). Hence the accuracy of the preceding modules of our ESD framework, (a) fitting a MASM to an object and (b) calculating DBS between MASMs, are important to ensure that the low dimensional space found by GE preserves shape differences between objects. We evaluated the performance of each of these modules independently to assess their contribution to the overall performance of ESDs.

The results from Experiment 1 (Section 5.3) show that the MASM is able to accurately model the shapes of glands as seen on prostate histopathology. However, a relatively minor portion of glands (approximately 20% of misclassified glands) may be misclassified due to an inability of the MASM to capture subtle differences in shape (results not shown). Hence, future work may involve calculating the MASM using more advance methods, for example employing methods that preserve branches that best describe object morphology (Bai et al., 2007).

The results from Experiment 2 (Section 5.4) demonstrate that DBS is able to accurately determine correspondence between a wide variety of MASMs over a range of deformations. However, determining MASM correspondence is a difficult task. Furthermore, evaluation of the registration step demonstrated that in approximately 40% of misclassified glands the inability to correctly determine medial atom correspondence, and hence accurately determine shape dissimilarity, may be responsible for misclassification (results not shown). Future work in improving the registration and shape dissimilarity calculation may result in a more accurate algorithm to detect subtle shape differences between objects. For instance, a variety of learning algorithms could be employed to learn a weight for each term in the dissimilarity measure (Eq. (8)) (Chen et al., 2009). By weighting each term according to its importance in detecting subtle shape differences we would expect better separability between objects of different classes.

Finally, gland misclassification may be caused by GE being unable to accurately preserve the relationships between samples when projecting the high dimensional space into a lower dimensional space. We found roughly 25% of misclassified glands may be on account of classes relationships not being preserved during this step (results not shown). Future work will evaluate other NLDR algorithms (e.g. Locally Linear Embedding (LLE) (Roweis and Saul, 2000), Isomaps (Tenenbaum et al., 2000)) within our framework. Additionally, the use of class information when learning the low dimensional ESD features may further improve classification accuracy. Another direction of future work will be to consider integrating

semi-supervised NLDR approaches (Zhao, 2006) for learning the low dimensional ESD features.

For approximately 15% of glands that were misclassified we were unable to determine which step in our algorithm may be responsible for the misclassification. Visual inspection of a subset of these glands showed that some misclassifications may be attributed to glands displaying atypical shape attributes. In future work we will integrate the ESD features with other types of histologic image attributes, such as the shape and arrangement of nuclei (Doyle et al., 2012a) or texture (Doyle et al., 2011; Doyle et al., 2012b), to build CAD systems for Gleason grading of prostate histopathology that may be better able to classify glands with atypical shape attributes.

We have obtained preliminary results that show that ESDs are able to distinguish subtle differences of object morphology in a wide variety of applications, for instance distinguishing benign and malignant lesions on breast MRI (Sparks and Madabhushi, 2010). However, future work is necessary to validate these results in a more comprehensive manner.

Acknowledgments

This work was made possible via grants from the National Cancer Institute under award numbers R01CA13653501, R01CA14077201, and R21CA167811; the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under award number R43EB01519901; and the Department of Defense (W81XWH-677 11-1-0179). We would like to thank Dr. J.E. Tomaszewski from the University at Buffalo School of Medicine and Biomedical Sciences as well as Drs. M.D. Feldman and N. Shih from the Hospital of the University of Pennsylvania for providing prostate histology imagery and corresponding annotations.

References

Agner SC, Soman S, Libfeld E, McDonald M, Thomas K, Englander S, Rosen M, Chin D, Nosher J, Madabhushi A. Textural kinetics: a novel dynamic contrast enhanced (DCE)–MRI feature for breast lesion classification. Journal of Digital Imaging. 2011; 24:446–463. [PubMed: 20508965]

Alcantara DA, Carmichael O, Harcourt-Smith W, Sterner K, Frost S, Dutton R, Thompson P, Delson E, Amenta N. Exploration of shape variation using localized components analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2009; 31:1510–1516. [PubMed: 19542583]

Allsbrook WC, Mangold KA, Johnson MH, Lane RB, Lane CG, Epstein JI. Interobserver reproducibility of Gleason grading of prostatic carcinoma: general pathologist. Human Pathology. 2001; 32:81–88. [PubMed: 11172299]

Arcelli C, di Baja G, Serino L. Distance-driven skeletonization in voxel images. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2011; 33:709–720. [PubMed: 20714012]

Bai X, Latecki LJ. Path similarity skeleton graph matching. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2008; 30:1282–1292. [PubMed: 18550909]

Bai X, Latecki LJ, Liu WY. Skeleton pruning by contour partitioning with discrete curve evolution. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2007; 29:449–462. [PubMed: 17224615]

Bai X, Yang X, Latecki LJ, Liu W, Tu Z. Learning context-sensitive shape similarity by graph transduction. IEEE Transactions on Pattern Analysis and Machine Learning. 2010; 32:861–874.

Bailey R, Srinath M. Orthogonal moment features for use with parametric and non-parametric classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1996; 18:389–399.

Belkin M, Niyogi P, Sindhwani V. Manifold regularization: a geometric framework for learning from labeled and unlabeled examples. Journal of Machine Learning Research. 2006; 7:2399–2434.

Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002; 24:509–522.

Bengio Y, Paiement JF, Vincent P, Delalleau O, Roux NL, Ouimet M. Outof-sample extensions for LLE, Isomap, MDS, Eigenmaps, and Spectral Clustering. Advances in Neural Information Processing Systems. 2003:177–184.

Blum H. A transformation for extracting new descriptors of shape. Proc. of Symposium on Models for the Perception of Speech and Visual Form. 1967:362–380.

Bookstein FL. Principle warps: thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Learning. 1989; 11:567–585.

Brechböhler C, Gerig G, Kübler O. Parametrization of closed surfaces for 3-d shape description. Computer Vision and Image Understanding. 1995; 61:154–170.

Chang H, Yeung D. Robust locally linear embedding. Pattern Recognition. 2006; 39:1053–1065.

Chang MC, Kimia BB. Measuring 3d shape similarity by graph-based matching of the medial scaffolds. Computer Vision and Image Understanding. 2011; 115:707–720.

Chen L, McAuley J, Feris R, Caetano T, Turk M. Shape classification through structured learning of matching measures. Computer Vision and Pattern Recognition. 2009:365–372. 2009.

Cootes TF, Hill A, Taylor CJ, Haslam J. Use of active shape models for locating structure in medical images. Image and Vision Computing. 1994; 12:355–365.

Cortes C, Vapnik V. Support-vector networks. Machine Learning. 1995; 20:273–297.

Dice LR. Measures of the amount of ecologic association between species. Ecology. 1945; 26:297–302.

Doyle S, Feldman M, Shih N, Tomaszewski J, Madabhushi A. Cascaded discrimination of normal, abnormal, and confounder classes in histopathology: Gleason grading of prostate cancer. BMC Bioinformatics. 2012a; 13:282. [PubMed: 23110677]

Doyle S, Feldman M, Tomaszewski J, Madabhushi A. A boosted bayesian multi-resolution classifier for prostate cancer detection from digitized needle biopsies. IEEE Transactions on Biomedical Engineering. 2012b; 59:1205–1218. [PubMed: 20570758]

Doyle S, Monaco J, Feldman M, Tomaszewski J, Madabhushi A. An active learning based classification strategy for the minority class problem: application to histopathology annotation. BMC Bioinformatics. 2011; 12:424. [PubMed: 22034914]

Egozi A, Keller Y, Guterman H. Improving shape retrieval by spectral matching and meta similarity. IEEE Transactions on Image Processing. 2010; 19:1319–1327. [PubMed: 20071259]

Epstein JI, Allsbrook WC, Amin MB, Egevad LL. The 2005 international society of urological pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. American Journal of Surgical Pathology. 2005; 29:1228–1242. [PubMed: 16096414]

Fletcher PT, Lu C, Pizer SM, Joshi S. Principal geodesic analysis for the study of nonlinear statistics of shape. IEEE Transactions on Medical Imaging. 2004; 23:995–1005. [PubMed: 15338733]

Fowlkes C, Belongie S, Chung F, Malik J. Spectral grouping using the Nyström method. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2004; 28:214–225. [PubMed: 15376896]

Franklin WR, Barr AH. Faster calculation of superquadric shapes. IEEE Computer Graphics and Applications. 1981; 1:41–47.

Georgiou H, Mavroforakis M, Dimitropoulos N, Cavouras D, Theodoridis S. Multi-scaled morphological features for the characterization of mammographic masses using statistical classification schemes. Artificial Intelligence in Medicine. 2007; 41:39–55. [PubMed: 17714924]

Gerardin E, Chételat G, Chupin M, Cuingnet R, Desgranges B, Kim HS, Niethammer M, Dubois B, Lehéricy S, Garnero L, Eustache F, Colliot O. The Alzheimer's Disease Neuroimaging Initiative, 2009. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. NeuroImage. 47:1476–1486. [PubMed: 19463957]

Gleason DF. Classification of prostatic carcinomas. Cancer Chemotherapy Reports. 1966; 50:125–128.

Goh WB. Strategies for shape matching using skeletons. Computer Vision and Image Understanding. 2008; 110:326–345.

Gorczowski K, Styner M, Jeong JY, Marron JS, Piven J, Hazlett HC, Pizer SM, Gerig G. Multi-object analysis of volume, pose, and shape using statistical discrimination. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2010; 32:652–661. [PubMed: 20224121]

Guo H, Rangarajan A, Joshi S. Diffeomorphic point matching. Handbook of Mathematical Models in Computer Vision. 2006:205–219.

Hassouna M, Farag A. Variational curve skeletons using gradient vector flow. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2009; 31:2257–2274. [PubMed: 19834145]

Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the Hausdorff distance. IEEE Transaction on Pattern Analysis and Machine Intelligence. 1993; 15:850–863.

Jobst KA, Smith AD, Szatmari M, Esiri MM, Jaskowski A, Hindley N, McDonald B, Molyneux AJ. Rapidly progressing atrophy of medial temporal lobe in Alzheimer's disease. Lancet. 1994; 343:829–830. [PubMed: 7908080]

van Kaick O, Hamarneh G, Ward AD, Schweitzer M, Zhang H. Learning fourier descriptors for computer-aided diagnosis of the supraspinatus. Academic Radiology. 2010; 17:1040–1049. [PubMed: 20605487]

Kauppinen H, Seppanen T, Pietikainen M. An experimental comparison of autoregressive and fourier-based descriptors in 2d shape classification. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1995; 17:201–207.

Lauric A, Miller E, Frisken S, Malek AM. Automated detection of intracranial aneurysms based on parent vessel 3d analysis. Medical Image Analysis. 2010; 14:149–159. [PubMed: 20004607]

Liney GP, Sreenivas M, Gibbs P, Garcia-Alvarez R, Turnbull LW. Breast lesion analysis of shape technique: semiautomated vs. manual morphological description. Journal of Magnetic Resonance Imaging. 2006; 23:493–498. [PubMed: 16523479]

Looi JC, Macfarlane MD, Walterfang M, Styner M, Velakoulis D, Ltt J, van Westen D, Nilsson C. Morphometric analysis of subcortical structures in progressive supranuclear palsy: in vivo evidence of neostriatal and mesencephalic atrophy. Psychiatry Research: Neuroimaging. 2011; 194:163–175.

Macrini D, Dickinson S, Fleet D, Siddiqi K. Bone graphs: medial shape parsing and abstraction. Computer Vision and Image Understanding. 2011; 115:1044–1061.

Madabhushi A, Agner S, Basavanhally A, Doyle S, Lee G. Computer-aided prognosis: predicting patient and disease outcome via quantitative fusion of multi-scale, multi-modal data. Computerized Medical Imaging and Graphics. 2011; 35:506–514. [PubMed: 21333490]

Mokhtarian F, Abbasi S. Shape similarity retrieval under affine transforms. Pattern Recognition. 2002; 35:31–41.

Monaco JP, Tomaszewski JE, Feldman MD, Hagemann I, Moradi M, Mousavi P, Boag A, Davidson C, Abolmaesumi P, Madabhushi A. High-throughput detection of prostate cancer in histological sections using probabilistic pairwise Markov models. Medial Image Analysis. 2010; 14:617–629.

Nasreddine K, Benzinou A, Fablet R. Variational shape matching for shape classification and retrieval. Pattern Recognition Letters. 2010; 31:1650–1657.

Pachauri D, Hinrichs C, Chung M, Johnson S, Singh V. Topology-based kernels with application to inference problems in alzheimer's disease. IEEE Transactions on Medical Imaging. 2011; 30:1760–1770. [PubMed: 21536520]

Persoon E, Fu KS. Shape discrimination using Fourier descriptors. IEEE Transactions on Systems, Man, and Cybernetics. 1977; 7:170–179.

Pizer S, Fritsch D, Yushkevich P, Johnson V, Chaney E. Segmentation, registration, and measurement of shape variation via image object shape. IEEE Transactions on Medical Imaging. 1999; 18:851–865. [PubMed: 10628945]

Qian X, Tagare HD, Fulbright RK, Long R, Antani S. Optimal embedding for shape indexing in medical image databases. Medical Image Analysis. 2010; 14:243–254. [PubMed: 20163981]

Rangayyan RM, Nguyen TM. Fractal analysis of contours of breast masses in mammograms. Journal of Digital Imaging. 2007; 20:223–237. [PubMed: 17021926]

Reuter M, Wolter FE, Peinecke N. Laplacebeltrami spectra as shape-dna of surfaces and solids. Computer-Aided Design. 2006; 38:342–366.

Rodgers JL, Nicewander WA. Thirteen ways to look at the correlation coefficient. The American Statistician. 1988; 42:59–66.

Rose K, Gurewitz E, Fox G. A deterministic annealing approach to clustering. Pattern Recognition Letters. 1990; 11:589–594.

Roweis S, Saul L. Nonlinear dimensionality reduction by locally linear embedding. Science. 2000; 290:2323–2326. [PubMed: 11125150]

Ruberto CD. Recognition of shapes by attributed skeletal graphs. Pattern Recognition. 2004; 37:21–31.

Rueda S, Udupa JK, Bai L. Shape modeling via local curvature scale. Pattern Recognition Letters. 2010; 31:324–336.

Sajjanhar A, Lu G, Zhang D, Hou J, Chen YP. Spherical harmonics and distance transform for image representation and retrieval. Proc. Intelligent Data Engineering and Automated Learning. 2009; vol. 5788:309–316.

Sebastian TB, Klein PN, Kimia BB. Recognition of shapes by editing their shock graphs. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2004; 26:550–571. [PubMed: 15460278]

Shi J, Malik J. Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Learning. 2000; 22:888–905.

Shinagawa Y, Kunii T, Kergosien Y. Surface coding based on morse theory. IEEE Computer Graphics and Applications. 1991; 11:66–78.

Siddiqi K, Shokoufandeh A, Dickinson SJ, Zucker SW. Shock graphs and shape matching. International Journal of Computer Vision. 1999; 35:13–32.

Sjöstrand K, Rostrup E, Ryberg C, Larsen R, Studholme C, Baezner H, Ferro J, Fazekas F, Pantoni L, Inzitari D, Aldemar GW, Group LS. Sparse decomposition and modeling of anatomical shape variation. IEEE Transactions on Medical Imaging. 2007; 26:1625–1635. [PubMed: 18092733]

Sparks R, Madabhushi A. Novel morphometric based classification via diffeomorphic based shape representation using manifold learning. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2010:658–665.

Street WN, Wolberg WH, Mangasarian OL. Nuclear feature extraction for breast tumor diagnosis. IS&T/SPIE 1999 International Symposium on Electronic Imaging:Science and Technology 1905. 1993:861–870.

Sun J, Ovsjanikov M, Guibas L. A concise and provably informative multi-scale signature based on heat diffusion. Computer Graphics Forum. 2009; 28:1383–1392.

Tabesh A, Teverovskiy M, Pang H, Kumar V, Verbel D, Kotsianti A, Saidi O. Multifeature prostate cancer diagnosis and Gleason grading of histological images. IEEE Transaction on Medical Imaging. 2007; 26:1366–1378.

Tenenbaum J, de Silvia V, Langford J. A global framework for nonlinear dimensionality reduction. Science. 2000; 290:2319–2323. [PubMed: 11125149]

Twining C, Marsland S, Taylor C. Measuring geodesic distances on the space of bounded diffeomorphisms. Proc. British Machine Vision Conference (BMVC). 2002:847–856. 2002.

Ward A, Hamarneh G. The groupwise medial axis transform for fuzzy skeletonization and pruning. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2010; 32:1084–1096. [PubMed: 20431133]

Xu Y, Liang G, Hu G, Yang Y, Geng J, Saha PK. Quantification of coronary arterial stenoses in cta using fuzzy distance transform. Computerized Medical Imaging and Graphics. 2012; 36:11–24. [PubMed: 21555207]

Yang W, Zhang S, Chen Y, Li W, Chen Y. Shape symmetry analysis of breast tumors on ultrasound images. Computers in Biology and Medicine. 2009; 39:231–238. [PubMed: 19178908]

Yushkevich PA, Zhang H, Gee JC. Continuous medial representation for anatomical structures. IEEE Transactions on Medical Imaging. 2006; 25:1547–1564. [PubMed: 17167991]

Zhang Z. Iterative point matching for registration of free-form curves and surfaces. International Journal of Computer Vision. 1994; 13:119–152.

Zhao H. Combining labeled and unlabeled data with graph embedding. Neurocomputing. 2006; 69:2385–2389.

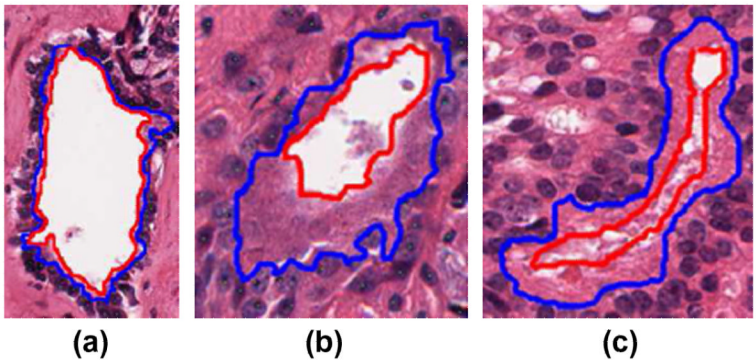

Fig. 1.

Three representative prostate glands on digitized needle biopsy histology specimens with lumen boundary (red) and nuclear boundary (blue) segmentations displayed. (a) A gland from benign prostate tissue, the gland has a regular oval structure. (b) A gland from a CaP region identified as Gleason grade 3, the gland is smaller with greater margin irregularity compared to the benign gland. (c) A gland from a CaP region identified as Gleason grade 4, the gland is highly irregular in shape with a shrunken lumen. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

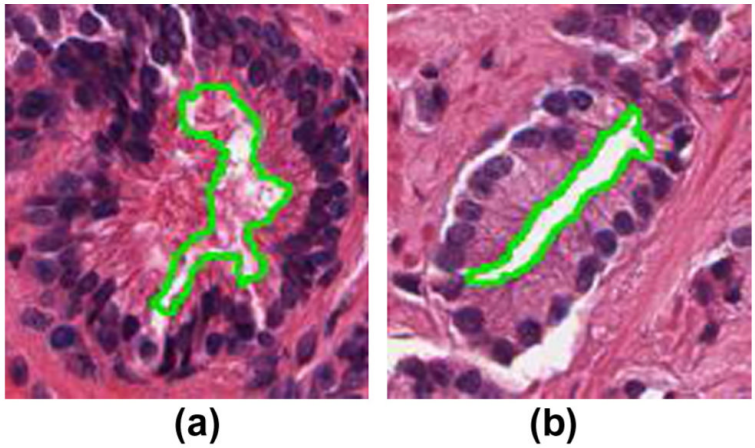

Fig. 2.

Example of two prostate glands with similar boundary-based descriptors despite obvious qualitative visual differences in morphology. (a) A benign prostate gland with a smoothness of 0.3 and a compactness of 0.57. (b) A Gleason grade 4 prostate gland with a smoothness of 0.3 and a compactness of 0.57. Our novel ESD methodology calculated values of 0.66 and 0.87 respectively for the glands, and hence is better able to quantify the differences between these two glands. All measures have a range of 0–1.

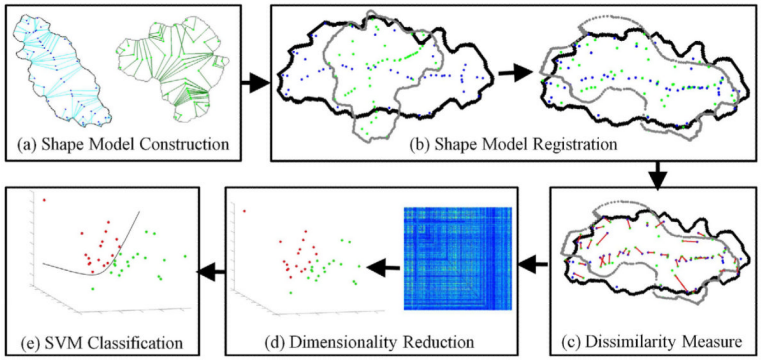

Fig. 3.

An illustration of the main modules for extracting explicit shape descriptors (ESDs). (a) A medial axis shape model (MASM) (blue, green) is fit to each object contour (black, gray). (b) Pairwise registration between MASMs is performed to align medial axes which then aids in (c) determining parameter correspondence between registered MASMs. Subsequently, pairwise differences between object shapes are computed which yields a $N \times N$ affinity matrix. (d) A NLDR scheme, GE, is then applied yielding a set of ESDs which quantify shape differences. Finally, (e) a Support Vector Machine (SVM) is trained to learn the optimal hyperplane which separates the ESD feature space into different object classes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

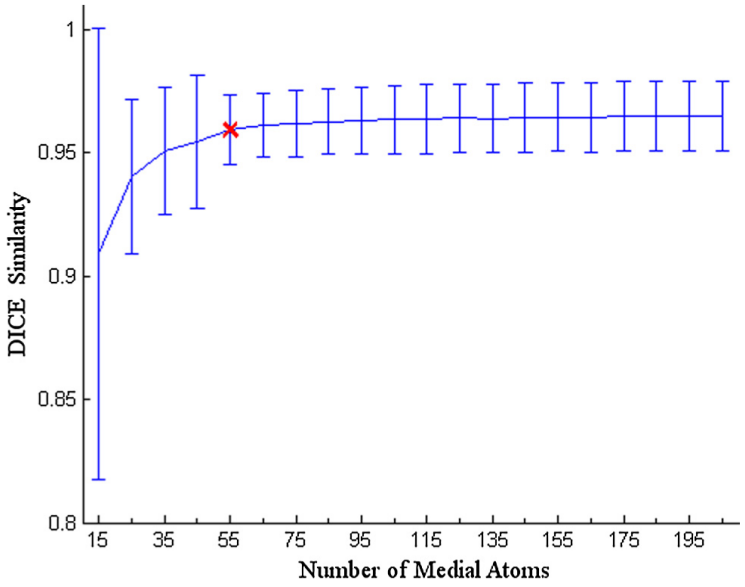

Fig. 4.

Prostate gland reconstruction accuracy as measured by DICE over number of medial atoms (blue). The number of medial atoms determined to give the highest reconstruction accuracy for the least computational cost is displayed (red cross). At first there is a large increase in DICE as more medial atoms are added to the MASM. After a certain point, adding more medial atoms does not significantly increase DICE. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

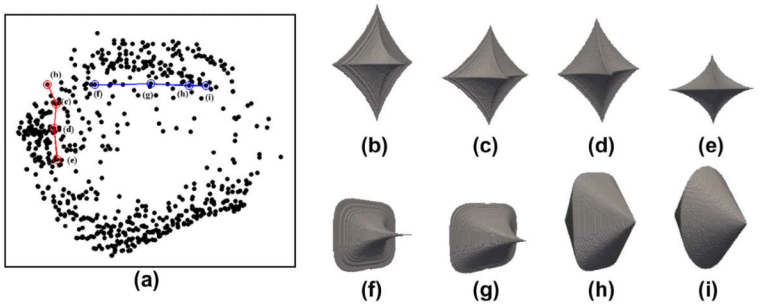

Fig. 5.

(a) The first and second ESDs are plotted on the X and Y axes respectively. Note that the manifold is curvilinear with the two axes corresponding roughly to the variation in ϵ_1 (red) and ϵ_2 (blue) respectively. (b)–(e) Ellipsoids with all parameters held equal except ϵ_2, resulting in subtle differences between object morphology. (f)–(i) Ellipsoids with all parameters held equal except ϵ_1, resulting in subtle differences between object shape. Finally note that the two ellipsoids farthest on the manifold, (e) and (i), are the most dissimilar. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

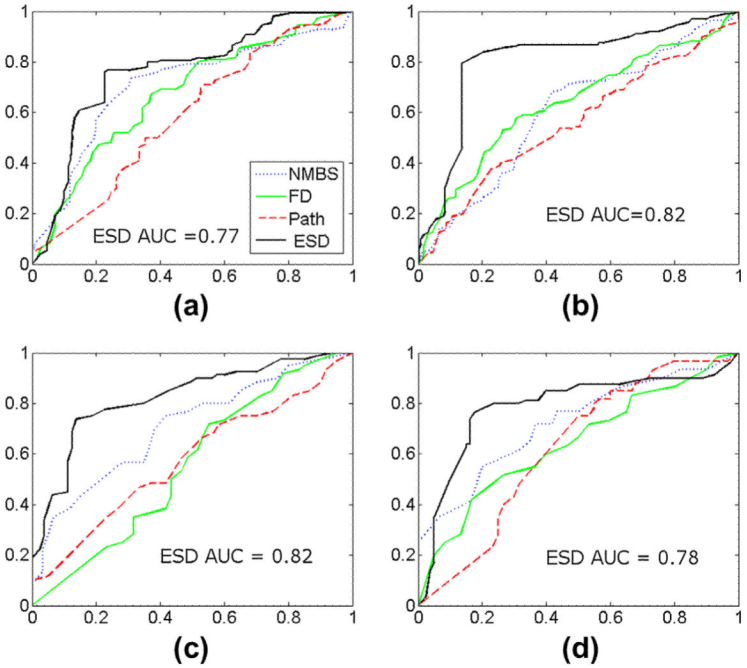

Fig. 6.

ROC curves for a SVM classifier trained using 3-fold cross-validation on 888 prostate glands for Gleason grade classification of prostate glands as seen on histopathology in four tasks: (a) BE versus Other (G3 and G4), (b) G3 versus other (BE and G4), (c) G4 versus other (BE and G3), and (d) G3 versus G4. Four feature sets were evaluated, Boundary (blue), FD (green), Path (red), and ESD (black). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

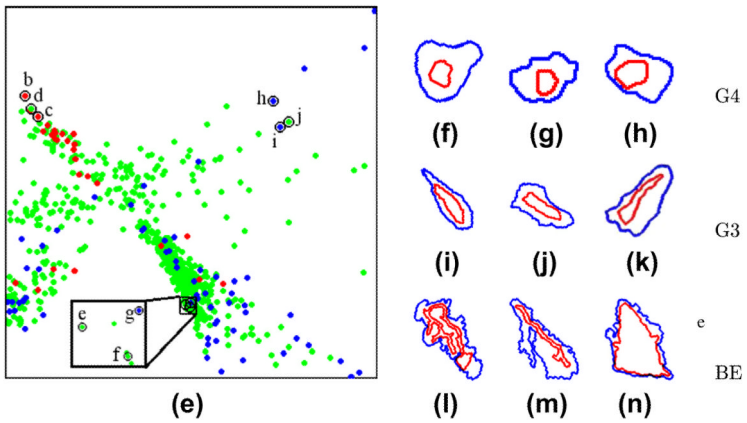

Fig. 7.

(e) ESD feature space for prostate digital histopathology with BE (blue), G3 (green), and G4 (red) glands. The first and second ESDs are plotted on the X and Y axes respectively. Lumen (red) and nuclear (blue) layers are shown, for glands labeled (f)–(h) G4, (i)–(k) G3, and (l)–(n) BE. Ground truth for mislabeled glands, displayed in the far right row, are (h) G3, (k) BE, (n) G3. Glands with similar shapes are embedded adjacent to each other on the manifold while glands with dissimilar shapes are embedded far apart. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1

Description of commonly employed notation and symbols in this paper

Symbol	Description	Symbol	Description
C	d-dimensional image scene.	T	Affine transformation function.
C	d-dimensional grid of voxels.	g^{kj}	kth cluster centroid at iteration j.
ξ	Contour for an object of interest.	$P(m\|g^{kj})$	Probability of $m \in M$ belonging to g^{kj}.
M	Medial axis shape model (MASM).	Q	Diffeomorphic transformation function.
M	Set of voxels on a medial axis.	G	Green's function.
$v_1(m), v_2(m)$	Surface vectors for $m \in M$.	$(\hat{u}, \hat{v})$	Correspondence between two medial axes.
$f(c)$	Signed distance function for $c \in C$.	A	$N \times N$ dissimilarity matrix.
X_i	ith direction.	**y'**	Set of n Explicit Shape Descriptors (ESDs).
N	Number of objects.	S	Super quadratic ellipsoid.

Table 2

A listing of the 6 boundary-based features utilized to evaluate object morphology and compared against our ESD feature set. Note that $|\xi|$ represents the cardinality of set ξ

Boundary feature	Description
Normalized average radial distance ratio	$\frac{\frac{1}{\lvert\xi\rvert}\sum_{p\in\xi}\lVert p-\bar{p}\rVert}{\max_{p\in\xi}\lVert p-\bar{p}\rVert}$ where $\bar{p}=\frac{1}{\lvert\xi\rvert}\sum_{p\in\xi}p$
Area overlap ratio	$\frac{\lvert\xi\rvert}{\pi r^2}$ where $r=\max_{p\in\xi}\lVert p-\bar{p}\rVert$
Standard deviation of distance ratio	$\sigma_I=\sqrt{I(p)-\mu_I^2}$ where $I(p)=\frac{\lVert p-\bar{p}\rVert}{\max_{p\in\xi}\lVert p-\bar{p}\rVert}$ and $\mu_I=\frac{1}{\xi}\sum_{p\in\xi}I(p)$
Variance of distance ratio	σ_I^2
Compactness	$\frac{F(\xi)^2}{\lvert\xi\rvert}$ where $F(\xi)=\sum_{p\in\xi,\,j\in\{1,\ldots,J\}}\lVert p^{j+1}-p^{j}\rVert$
Smoothness	$\sum_{p\in\xi,\,j\in\{1,\ldots,J\}}B(p^{(j)})$ where $B(p^{(j)})=\lVert p^{(j)}-\bar{p}\rVert-\frac{\lVert p^{(j-1)}-\bar{p}\rVert+\lVert p^{(j+1)}-\bar{p}\rVert}{2}$

Table 3

Object reconstruction accuracy for 2 datasets using the previously determined optimal number of medial atoms for each dataset. MAD and Hausdorff distances are shown in units of pixels. DICE and PPV are unitless ratios. Note that 3*D* objects require more medial atoms to accurately represent object morphology

Performance measure	Dataset (Dimensionality)	
	Ellipsoid (3*D*)	**Prostate gland (2*D*)**
Medial atoms	1000	55
MAD	3.40 ± 3.54	0.01 ± 0.01
Hausdorff	24.46 ±12.95	1.51 ±1.16
DICE	0.87 ±0.12	0.95 ± 0.03
PPV	0.94 ± 0.06	0.98 ± 0.02

Table 4

(a) CA and (b) AUC for a SVM classifier trained using 3-fold cross-validation on 888 prostate glands for distinguishing between Gleason grades on prostate histopathology. SVM classifiers were trained with 4 feature sets (Boundary, FD, Path, ESD). In total 16 classification studies (4 feature sets, 4 pairwise classification tasks) were performed. *p*-values comparing ESD to the comparison feature sets are reported, statistically significant *p*-values ($p < 0.01$) are bolded. The best CA and AUC across the feature sets is bolded

Classification task		**Feature**			
		Boundary	**FD**	**Path**	**ESD**
BE vs. other	CA	0.80 ± 0.05	0.70 ± 0.06	0.60 ± 0.05 $\gamma = 100$, $n = 9$	**0.83 ± 0.04** $\gamma = 1.2$, $n = 4$
	p-value	**0.005**	**0.001**	**0.005**	-
G3 vs. other	CA	0.70 ± 0.08	0.65 ± 0.04	0.61 ± 0.08 $\gamma = 1025$, $n = 14$	**0.87 ± 0.07** $\gamma = 2.1$, $n = 4$
	p-value	0.013	**0.006**	**0.003**	-
G4 vs. other	CA	0.75 ± 0.06	0.58 ± 0.14	0.65 ± 0.10 $\gamma = 1$, $n = 4$	**0.85 ± 0.03** $\gamma = 1.5$, $n = 4$
	p-value	**0.005**	0.024	0.016	-
G3 vs. G4	CA	0.68 ± 0.08	0.60 ±0.10	0.64 ± 0.06 $\gamma = 151$, $n = 5$	**0.89 ± 0.06** $\gamma = 1.2$, $n = 4$
	p-value	**0.004**	**0.001**	**0.002**	-
		NMBS			
BE vs. other	AUC	0.71 ± 0.07	0.67 ± 0.08	0.61 ± 0.09 $\gamma = 100$, $n = 9$	**0.77 ± 0.05** $\gamma = 1.2$, $n = 4$
	p-value	0.15	0.059	**0.007**	-
G3 vs. other	AUC	0.60 ± 0.09	0.64 ± 0.06	0.56 ± 0.08 $\gamma = 1025$, $n = 14$	**0.81 ± 0.05** $\gamma = 2.1$, $n = 4$
	p-value	**0.002**	**0.001**	**0.0004**	-
G4 vs. other	AUC	0.71 ± 0.08	0.57 ± 0.06	0.56 ± 0.20 $\gamma = 1$, $n = 4$	**0.82 ± 0.07** $\gamma = 1.5$, $n = 4$
	p-value	0.042	**0.0003**	0.023	-
G3 vs. G4	AUC	0.72 ± 0.04	0.58 ±0.17	0.64 ±0.17 $\gamma = 151$, $n = 5$	**0.78 ± 0.11** $\gamma = 1.2$, $n = 4$
	p-value	0.25	0.065	0.162	-