



3D object classification using salient point patterns with application to craniofacial research

Indriyati Atmosukarto^{a,*}, Katarzyna Wilamowska^a, Carrie Heike^b, Linda G. Shapiro^a

^a Department of Computer Science and Engineering, University of Washington, Box 352350, Seattle, WA, USA

^b Seattle Children's Hospital, Seattle, WA, USA

ARTICLE INFO

Article history:

Received 20 February 2009

Received in revised form

30 October 2009

Accepted 5 November 2009

Keywords:

3D object classification

Shape matching

Salient points

Craniofacial dysmorphism

ABSTRACT

This paper presents a new 3D shape representation and classification methodology developed for use in craniofacial dysmorphism studies. The methodology computes low-level features at each point of a 3D mesh representation, aggregates the features into histograms over mesh neighborhoods, learns the characteristics of salient point histograms for each particular application, and represents the points in a 2D spatial map based on a longitude–latitude transformation. Experimental results on the medical classification tasks show that our methodology achieves higher classification accuracy compared to medical experts and existing state-of-the-art 3D descriptors. Additional experimental results highlight the strength and advantage of the flexible framework that allows the methodology to generalize from specific medical classification tasks to general 3D object classification tasks.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Hospitals and clinics are now commonly scanning patients for diagnostic and clinical purposes. Patients typically encounter a variety of imaging modalities as an increasing amount of tests are accomplished through digital imaging. A patient may accumulate many different types of images such as 2D images, 3D volume CT scans, 3D surface meshes, 3D fMRI, and even 4D fMRI time series. Though different image modalities are typically acquired for clinical purposes, all resulting images are also commonly used for medical research studies.

Medical researchers at Seattle Children's Hospital Craniofacial Center (SCHCC) use CT scans and 3D surface meshes of children's heads to investigate head shape dysmorphism due to craniofacial disorders such as craniosynostosis, 22q11.2 deletion syndrome, deformational plagiocephaly, or cleft lip and palate. These researchers aspire to be able to store, quantify or classify, and retrieve 3D objects automatically. Current techniques for 2D image classification and video categorization cannot be directly translated and applied to 3D object classification as 3D objects have different data characteristics from the above modalities. Most existing 3D shape descriptors have only been developed and tested on general 3D object datasets, while those designed for

medical purposes must satisfy a specific medical application and dataset.

Classification of 3D objects requires that objects be represented in a way that captures the local and global shape characteristics of the object. This is achieved by creating a 3D object descriptor or signature, which summarizes the important shape properties of the object. Unfortunately, finding a good 3D object descriptor is not a trivial task. The global properties of an object capture the overall shape of an object, while the local properties capture the details of an object. The descriptor should be able to capture a good balance between the local and global shape properties of the object, so as to allow flexibility in performing different tasks.

Motivated by our existing collaborations with researchers at SCHCC, a new methodology for representing 3D objects was developed specifically for medical craniofacial applications. We are collaborating in two research studies at SCHCC for the study of craniofacial anatomy. The first study is of children with 22q11.2 deletion syndrome and the second study is of infants with deformational plagiocephaly.

22q11.2 deletion syndrome (22q11.2DS) is a genetic disease that is one of the most common multiple anomaly syndromes in humans [30]. This condition is associated with more than 180 clinical features, including over 25 dysmorphic craniofacial features. No single clinical feature occurs in every individual with 22q11.2DS, and there is no individual who has all the clinical features. Abnormal clinical features of individuals with 22q11.2DS include asymmetric face shape, hooded eyes, bulbous nasal tip, and retrusive chin, among others. The range of variation in

* Corresponding author. Tel.: +1 206 6856765; fax: +1 206 543 2969.

E-mail addresses: indria@cs.washington.edu (I. Atmosukarto), kasiaw@cs.washington.edu (K. Wilamowska), carrie.heike@seattlechildrens.org (C. Heike), shapiro@cs.washington.edu (L.G. Shapiro).

individual feature expression is very large. As a result, even experts have difficulty in diagnosing 22q11.2DS from frontal facial photographs alone [8]. Early detection of 22q11.2DS is important as many affected individuals are born with a conotruncal cardiac anomalies, mild-to-moderate immune deficiencies and learning disabilities, all of which can benefit from early intervention.

Deformational plagiocephaly (also known as positional plagiocephaly, or non-synostotic plagiocephaly) refers to the deformation of the head, characterized by a persistent flattening on the side resulting in an asymmetric head shape and misalignment of the ears. Deformational plagiocephaly is caused by persistent pressure on the skull of a baby before or after birth. Another possible factor, which can lead to deformational plagiocephaly, is torticollis, a muscle tightness in the neck resulting in a limited range of motion for the head that causes infants to look in one direction and to rest on the same spot of the back of the head. If left untreated, children with these abnormal head shape conditions may experience a number of medical issues in their lives, ranging from social problems due to abnormal appearance to delayed neurocognitive development.

Our methodology begins with a base framework that extracts the low-level and mid-level features of the 3D object. A learning approach is used to identify the interesting local features or salient points on a 3D object for a particular application. These feature points are represented in a 2D global spatial map that becomes a signature for the object. A classifier is trained on the signatures of the objects for a number of classes and can then be used to classify new unseen objects. We developed our methodology on a Heads database that was created especially for our work. This database is relevant for our craniofacial applications and yet more general, as the database includes head shapes of not only humans but also other species such as cats, dogs, horses, bears, and tigers. The methodology was then used in the two craniofacial applications: 22q11.2 deletion syndrome and deformational plagiocephaly. Finally, the methodology was tested for generality on the SHape REtrieval Contest (SHREC [49]) general object database, although this dataset did not satisfy all the assumptions on which our algorithm is based.

The remainder of this paper is organized as follows. First, existing shape descriptors for general 3D objects and specific medical studies of craniofacial features are discussed. Next, our base framework for feature extraction is described. The datasets used to develop and test the methodology are then described. The method for learning the salient points of a 3D object is explained and applied to the different applications. Then, a 2D longitude–latitude map signature that captures the pattern of salient points is defined. In the experimental results section, a set of classification experiments, using the 2D map signatures on the 22q11.2DS, Plagiocephaly, Heads and SHREC datasets, is described and analyzed. Finally, a summary and suggestions for future work are provided.

2. Related literature

Shape-based 3D object retrieval and classification has received increased attention in the past few years due to the increase in the number of 3D objects available. There have been several survey papers on the topic [51,52,21,59,17,27,13,14,12,34,41]. Most recently, an evaluation database was introduced to establish a benchmark in the field. The various databases in SHREC are now used to evaluate the effectiveness of various 3D object descriptors in various retrieval and classification tasks. Results so far show that no one descriptor performs the best for all kinds of task. Each

descriptor has its own strength and weakness for the different tasks.

There are three broad categories of 3D object representation: feature-based methods, graph-based methods, and view-based methods. Feature-based methods, which are the most popular, can be further categorized into (1) global features, (2) global feature distributions, (3) spatial maps, and (4) local features. Early work on 3D object representation and its application for retrieval and classification focused more on the global features and global feature distribution approaches. Global features computed to represent 3D objects include area, volume or moments [18]. Example of global feature descriptors include global shape distributions by Osada et al. [39] and modified shape distributions by Ohbuchi et al. [37]. The global feature methods are computationally efficient as they reduce the computation space of the 3D object by describing the object with fewer dimensions; however, these methods are not discriminative enough when the objects have small differences such as in intra-class retrieval cases or classification of very similar objects.

Spatial map representations describe the 3D object by capturing and preserving physical locations on the object. Saupe et al. [44] described a spherical extent function that calculates the maximal extent of a shape across all rays from the origin and compared two different kinds of representations of the function: using spherical harmonics and moments. The spherical harmonic coefficients reconstruct an approximation of the object at different resolutions. Kazhdan et al. [29] used this idea to show that spherical harmonics can be used to transform rotation dependent shape descriptors into rotation independent ones without the need to pose normalize the objects in advance. Laga et al. [32,31] and Zhenbao et al. [33] also used spherical wavelet descriptors to describe 3D objects. Assfalg et al. [2] captured the shape of a 3D object using the curvature map of the object's surface. Our method is quite related to this last approach, but it differs in that it does not use the curvature information directly.

Recent research is beginning to focus more on the local approach to representing 3D objects, as this approach has a stronger discriminative power when differentiating objects that are similar in overall shape [40]. Local features are often points that are considered to be interesting or salient on the 3D object. These points are computed in various ways. Some methods randomly select points on the surface of the object. Frome et al. [20], who developed a 3D shape context, and Johnson et al. [28], who designed spin image descriptors, both randomly selected points as their basis points. Shilane et al. [47,48] used random points with harmonic shape descriptors at four different scales. Castellani et al. [15] proposed a new methodology for detecting and matching salient points based on measuring how much a vertex is displaced after filtering. The salient points are characterized using a local description based on a hidden Markov model. Novatnack et al. [36,35] extracted corners and edges of a 3D model in a discrete scale-space. Akagunduz et al. [1] used a Gaussian pyramid at several scales to extract the surface extrema and represented the points and their relationships by a graphical model.

While feature-based methods use only the geometric properties of the 3D model to define the shape of the object, graph-based methods use the topological information of the 3D object to describe its shape. The graph that is constructed shows how the different shape components are linked together. The graph representations include model graphs, Reeb graphs [24], and skeleton graphs [50]. These methods are known to be computationally expensive and sensitive to small topological changes.

The most effective view-based shape descriptor currently is the light field descriptor (LFD) developed by Chen et al. [16]. A light field around a 3D object is a 4D function that represents the

radiance at a given 3D point in a given direction. Each 4D light field of a 3D object is represented as a collection of 2D images rendered from a 2D array of cameras distributed uniformly on a sphere. The light field descriptor was evaluated to be one of the best performing descriptors on the SHREC database. Ohbuchi et al. [38] used a similar view-based approach to the light field descriptor. However, their method extracted local features from each rendered image using the SIFT algorithm. Wang et al. [55] improved the space usage efficiency of the LFD descriptor by projecting a number of uniformly sampled random points along six directions to create six images that are then described using Zernike moments. Experimental results on the Princeton shape benchmark database [49] showed that their method's performance was comparable to the LFD descriptor for some categories.

Traditionally, medical studies of craniofacial features have been through physical examination combined with craniofacial anthropometric measurements [42]. Newer methods of craniofacial assessment use digital images, but hand measurements and hand labeling of landmarks of these digital images are still commonly used.

Automated methods for analyzing 22q11.2DS are limited, as craniofacial anthropometric measurements still prevail as the standard manual assessment method. Boehringer et al. [11] applied a Gabor wavelet transformation to 2D photographs of individuals with 10 different facial dysmorphic syndromes. The generated datasets were then transformed using principal component analysis (PCA) and classified using linear discriminant analysis, support vector machines and k-nearest neighbors. Hammond et al. [23] used the dense surface model approach to align training samples according to point correspondences. The approach then produces an "average" face for each population studied and represents each face by a vector of PCA coefficients. Neither method is fully automatic as both require manual landmark placement. Wilamowska, Wu, Heike and Shapiro [56,58] have tackled the quantification of 3D face shape for study of 22q11.2DS in ongoing work.

There have also been some semi-automated work on analysis of deformational plagiocephaly. Hutchison et al. [25,26] developed a technique called HeadsUp that involves taking a top view digital photograph of infant heads fitted with an elastic head circumference band. The elastic band is equipped with adjustable color markers to identify landmarks such as ear and nose position. The resulting photograph is then automatically analyzed to obtain quantitative measurements for the head shape, including cephalic index, head circumference, distance of ear to center of nose, oblique length, and oblique ratio. Their results showed that the cephalic index (CI) and oblique cranial length ratio (OCLR) can be used as a quantification measurement of shape severity, since the numbers differ significantly between cases and control. Zonenshayn et al. [60] used a headband with two adjustable points (nasion and inion of the head) and used photographs of the headband shape to calculate the cranial index of symmetry (CIS). These methods require consistency in setting up the band and placing the markers, which may lead to non-reproducible results. In addition, this is a 2D technique, but plagiocephaly and brachycephaly are three dimensional deformations. We propose a feature-based approach that uses a learning methodology to identify the interesting salient points on the object and creates a global spatial map of the salient point patterns. The approach described in this paper generalizes and improves our former work on classification of plagiocephaly [6,3].

3. Base framework

Our methodology was developed for single 3D object classification as required in craniofacial applications. It does not

handle objects in cluttered 3D scenes nor occlusion. A surface mesh, which represents a 3D object, consists of points p_i on the object's surface and information regarding the connectivity of the points. The base framework for our methodology starts by rescaling the objects to fit in a fixed size bounding box. The framework then consists of two phases: low-level feature extraction and mid-level feature aggregation. The low-level feature extraction starts by applying a low-level operator to every point on the surface mesh. After the first phase, every point p_i on the surface mesh will have either a single low-level feature value v_i or a small set of low-level feature values, depending on the operator used. The second phase performs mid-level feature aggregation and computes a vector of values for a given neighborhood of every point p_i on the surface mesh. The feature aggregation results of the base framework are then used to construct our 3D object representation [5,4].

3.1. Low-level feature extraction

The low-level operators extract local properties of the surface points by computing a feature value v_i for every point p_i on the mesh surface. All low-level feature values are convolved with a Gaussian filter to reduce noise effects. Three low-level operators were implemented to test the methodology's performance: absolute Gaussian curvature, Besl–Jain curvature categorization, and azimuth–elevation of surface normal vectors. Fig. 1(a) shows an example of the absolute Gaussian curvature values of a 3D model. Fig. 1(b) shows the results of applying a Gaussian filter over the low-level Gaussian curvature values, while Fig. 1(c) shows the results of applying the Gaussian filter over the low-level Besl–Jain curvature values.

3.1.1. Absolute Gaussian curvature

The absolute Gaussian curvature low-level operator computes the Gaussian curvature estimation K for every point p on the surface mesh:

$$K = 2\pi - \sum_{f \in F} interior_angle_f$$

where F is the list of all the neighboring facets of point p and the interior angle is the angle of the facets meeting at point p . This calculation is similar to calculating the angle deficiency at point p . The contribution of each facet is weighted by the area of the facet divided by the number of points that form the facet. The operator then takes the absolute value of the Gaussian curvature as the final low-level feature value for each point.

3.1.2. Besl–Jain curvature

Besl and Jain [9] suggested surface characterization of a point p using only the sign of the mean curvature H and Gaussian curvature K . These surface characterizations result in a scalar surface feature for each point that is invariant to rotation, translation and changes in parameterization. The eight different categories are: (1) peak surface, (2) ridge surface, (3) saddle ridge surface, (4) plane surface, (5) minimal surface, (6) saddle valley, (7) valley surface, and (8) cupped surface. Table 1 lists the different surface categories with their respective curvature signs.

3.1.3. Azimuth–elevation angles of surface normal vectors

Given the surface normal vector $n(n_x, n_y, n_z)$ of a 3D point, the azimuth angle θ of n is defined as the angle between the positive xz plane and the projection of n to the x plane. The elevation angle

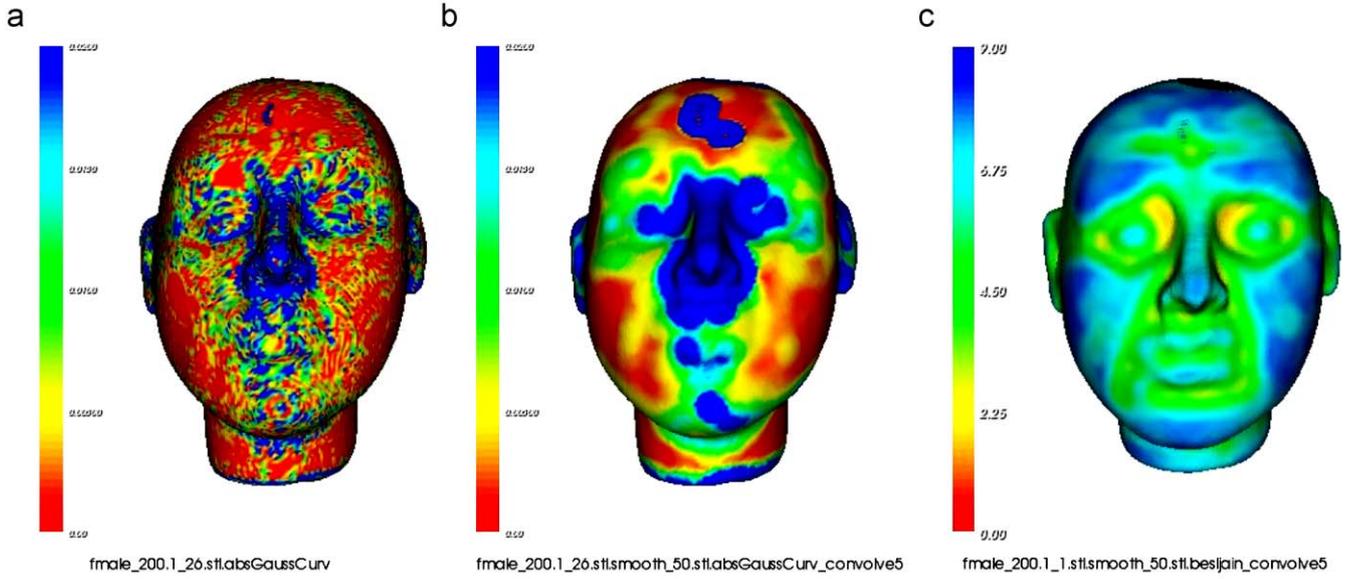


Fig. 1. (a) Absolute Gaussian curvature low-level feature value. (b) Smoothed absolute Gaussian curvature values after convolution with the Gaussian filter. (c) Smoothed Besl-Jain curvature values after convolution. Higher values are represented by cool (blue) colors, while lower values are represented by warm (red) colors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Besl-Jain surface characterization.

Label	Category	H	K
1	Peak surface	$H < 0$	$K > 0$
2	Ridge surface	$H < 0$	$K = 0$
3	Saddle ridge surface	$H < 0$	$K < 0$
4	Plane surface	$H = 0$	$K = 0$
5	Minimal surface	$H = 0$	$K < 0$
6	Saddle valley	$H > 0$	$K < 0$
7	Valley surface	$H > 0$	$K = 0$
8	Cupped surface	$H > 0$	$K > 0$

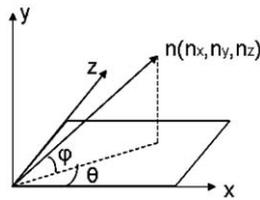


Fig. 2. Azimuth and elevation angle of a 3D surface normal vector.

ϕ of n is defined as the angle between the x plane and vector n (Fig. 2).

$$\theta = \arctan\left(\frac{n_z}{n_x}\right), \quad \phi = \arctan\left(\frac{n_y}{\sqrt{n_x^2 + n_z^2}}\right)$$

where $\theta = [-\pi, \pi]$ and $\phi = [-\pi/2, \pi/2]$. The azimuth–elevation low-level operator computes the azimuth and elevation values for each point on the 3D surface.

3.2. Mid-level feature aggregation

The second phase of our base framework performs mid-level feature aggregation and computes a number of values for a given neighborhood of each point p_i on the surface mesh. In this work, we use local histograms to aggregate the low-level feature values of each point. The histograms are computed by taking a

neighborhood around each point and accumulating the low-level features in that neighborhood. The size of the neighborhood is determined by multiplying a constant c , $0 < c < 1$, with the diagonal of the object’s bounding box. This ensures that the size of the neighborhood is scaled according to the object size, and that the results are comparable across different objects. The value of c was determined empirically, in our experiments we used $c = 0.05$. Aggregating the single-valued low-level feature values results in a 1D histogram with d histogram bins for every point on the surface mesh. Aggregating the pair-valued low-level feature values (such as the azimuth–elevation angle feature values) results in a 2D histogram constructed of $a \times b$ bins, where a and b are the two different dimension sizes. Fig. 3(a) shows an example of a 1D histogram aggregating the absolute Gaussian curvature low-level feature values from points on the nose of a 3D head object. Fig. 3(b) shows an example of the 2D histogram aggregating the azimuth–elevation low-level feature values on a head.

The mid-level features are used to determine salient points of an object, which are learned for each different application. Thus, before describing salient point learning and detection, we will briefly describe the four different datasets used.

4. Datasets

We obtained four different databases to develop and test our methodology. Each database has different characteristics that help explore the different properties of our methodology. The 22q11.2DS database contains 3D face models of individuals affected and unaffected by 22q11.2 deletion syndrome. The Plagiocephaly database contains 3D head models of individuals affected and unaffected by deformational plagiocephaly. The Heads database contains head shapes of different classes of animals, including humans. These three databases help explore the performance of the methodology on data of similar overall shape with subtle distinctions—the type of data for which our method was designed. The SHREC 2008 classification benchmark database was obtained to further test the performance of our

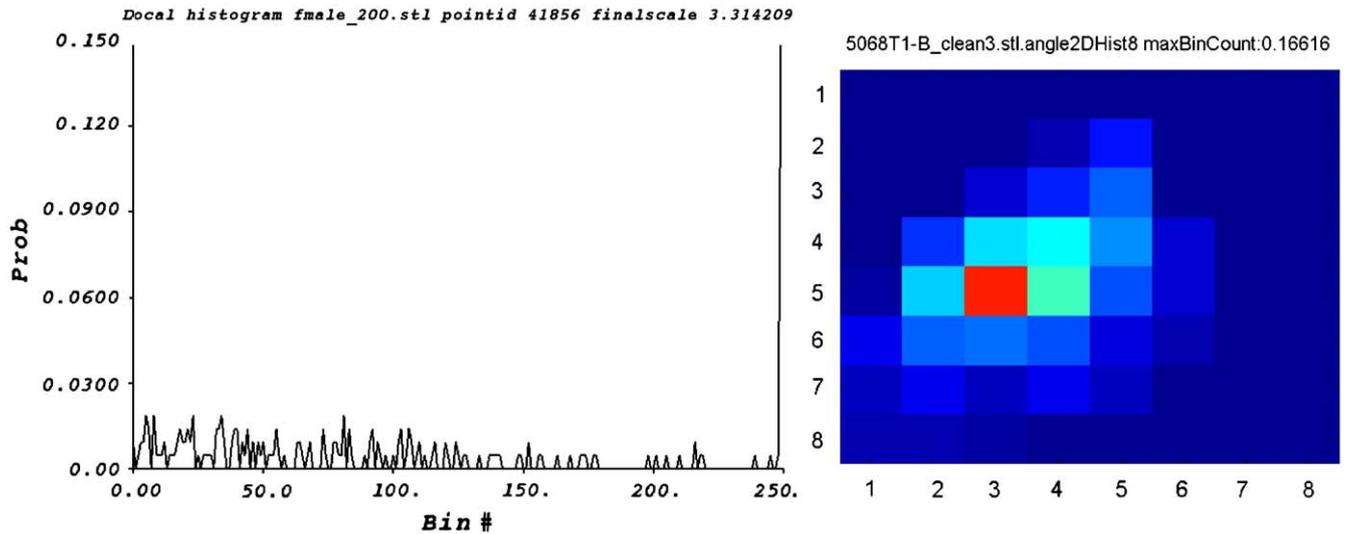


Fig. 3. (a) 1D histogram aggregating the absolute Gaussian curvature values from points on the nose of a human head. (b) 2D histogram aggregating the azimuth–elevation vector values at a point on the back of the head.

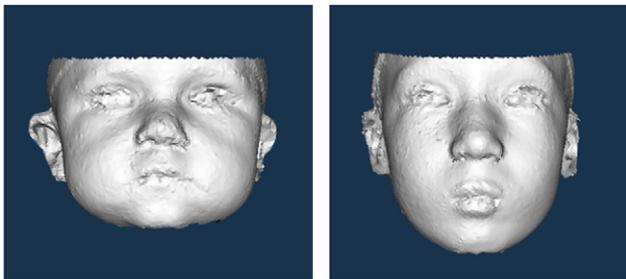


Fig. 4. Example of 3D face mesh data of children with 22q11.2 deletion syndrome.

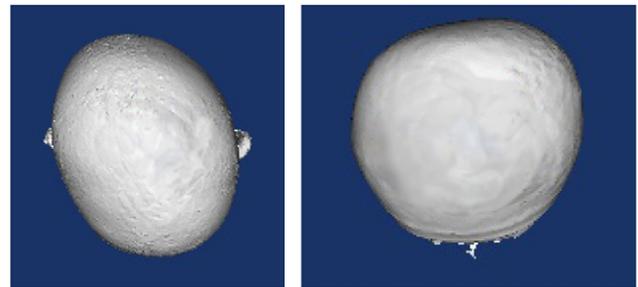


Fig. 5. Tops of heads of children with deformational plagiocephaly.

methodology on general 3D object classification, where objects in the database are not very similar.

4.1. 22q11.2DS dataset

The 3D face models used for these experiments were collected by the Craniofacial Center of Seattle Children’s Hospital using the 3dMD imaging system. The 3dMD imaging system uses four camera stands, each containing three cameras. Stereo analysis yields 12 range maps that are combined using 3dMD proprietary software to yield a 3D mesh of an individual’s head and a texture map of the face. Our system uses only the 3D meshes, due to human subject regulations. An automated system to align the pose of each mesh was developed, using symmetry to align the yaw and roll angles and a height differential to align the pitch angle. Although faces are not truly symmetrical, the pose alignment procedure can be cast as finding the angular rotations of yaw and roll that minimizes the difference between the left and right sides of the face. The pitch of the head was aligned by minimizing the difference between the height of the chin and the height of the forehead. In some cases, manual adjustments were necessary to pose normalize the faces [56,46]. Fig. 4 shows two examples of affected individuals in the dataset.

The dataset contains 3D meshes for 189 individuals. Metadata for each 3D mesh consisted of the age, gender, and self-described ethnicity of the individual plus a label of affected or unaffected. The dataset consisted of 53 affected individuals and 136 control individuals. The groundtruth for the individual’s label was determined through laboratory confirmation for 22q11.2 deletion syndrome.

4.2. Plagiocephaly dataset

The datasets for the plagiocephaly classification experiments were obtained through a similar data acquisition pipeline as the 22q11.2DS dataset. The resulting 3D meshes are also automatically pose normalized. Fig. 5 shows two examples of individuals diagnosed with deformational plagiocephaly.

The original dataset consisted of 254 3D head meshes consisting of 100 controls and 154 cases. Each mesh in the original dataset was assessed by two human experts who assigned discrete severity scores based on the degree of the deformation severity at the back of the head: category 0 (normal), category 1 (mild), category 2 (moderate), and category 3 (severe). To avoid inter-expert score variations, heads that were assigned different scores by the two human experts were removed from the dataset. The trimmed dataset consisted of 140 3D head meshes with 50 control individuals in category 0 and 90 case individuals: 46 in category 1, 35 in category 2, and 9 in category 3.

4.3. Heads dataset

For the Heads database, the digitized 3D objects were obtained by scanning hand-made clay toys using a Roldand-LPX250 laser scanner with a maximal scanning resolution of 0.008 in for plane scanning mode [43]. Raw data from the scanner consisted of 3D point clouds that were further processed to obtain smooth and uniformly sampled triangular meshes of 0.9–1.0 mm resolution. To increase the number of objects for

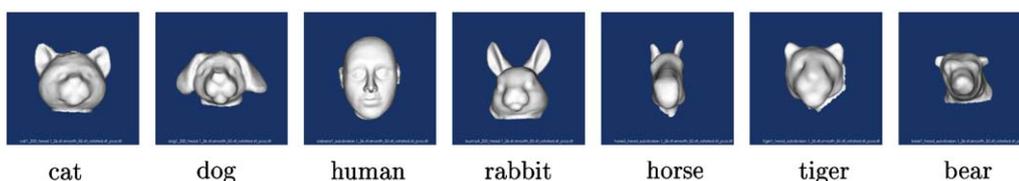


Fig. 6. Example of objects in the Heads database.

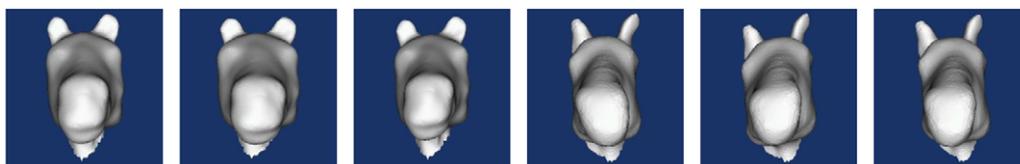


Fig. 7. Example morphs from the horse class. Morphs were generated by stretching, twisting, or squeezing the original object with different parameters.

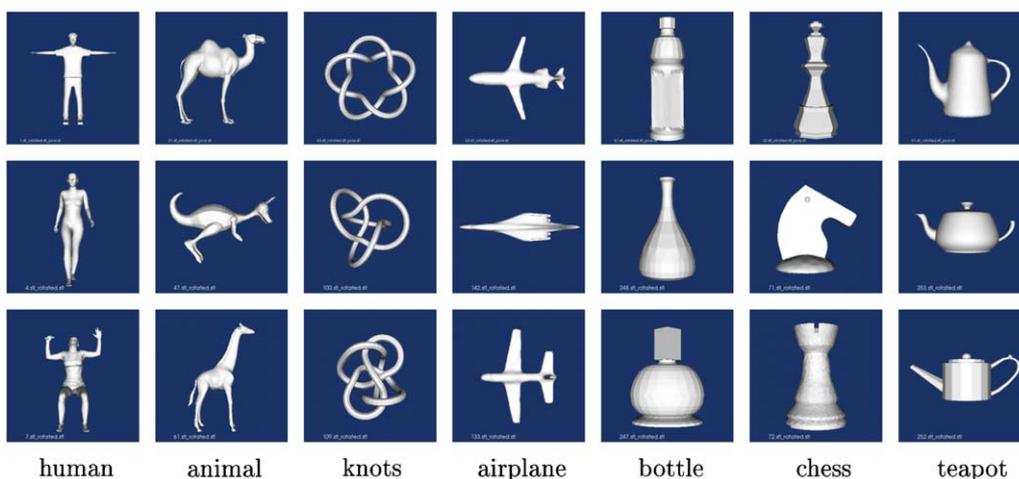


Fig. 8. Example of objects in the SHREC 2008 classification database. It can be seen that the intra-class variability in this dataset is quite high as objects in the same class have quite different shapes.

training and testing our methodology, we created new objects by deforming the original scanned 3D models in a controlled fashion using 3D Studio Max software [7]. The concept of morphable model was originally developed by Blantz and Vetter [10] in the context of computer graphics. Global deformations of the models were generated using morphing operators such as tapering, twisting, bending, stretching and squeezing. The parameters for each of the operators were randomly chosen from ranges that were determined empirically. Each deformed model was obtained by applying at least five different morphing operators in a random sequence.

Fifteen objects representing seven different classes were scanned. The seven classes are: cat head, dog head, human head, rabbit head, horse head, tiger head and bear head. Each of the 15 original objects were randomly morphed to increase the size of the database. A total of 250 morphed models per original object were generated. Points on the morphed model are in full correspondence with the original models from which they were constructed. Fig. 6 shows examples of objects from each of the seven classes, while Fig. 7 shows example of morphs from the horse class.

4.4. SHREC dataset

For the SHREC database, we selected the dataset from the track called “Classification of Watertight Models” [22]. The models in

the track were chosen by the organizer to ensure a high level of shape variability to make the track more challenging, so this dataset did not satisfy the assumptions on which our algorithm is based. The models in the database were manually classified using three different levels of categorization. At the *coarse* level of classification, the objects were classified according to both their shapes and semantic criteria. At the *intermediate* level, the classes were subdivided according to functionality and shape. At the *fine* level, the classes were further partitioned based on the object shape. For example, at the coarse level some objects were classified into the *furniture* class. At the intermediate level, these same objects were further divided into *tables*, *seats* and *beds*. At the fine level, the objects were classified into *chairs*, *armchairs*, *stools*, *sofa* and *benches*. We chose to use the intermediate level of classification as the fine level had too few objects per class, while the coarse level had too many objects that were dissimilar in shape grouped into the same class. The dataset consisted of 425 pre-classified objects. Fig. 8 shows examples of objects in the benchmark database.

5. Learning salient points

Given the base framework’s ability to compute low-level feature values at each point of a 3D mesh and to aggregate these features in neighborhoods about the point, we can now explore the use of this framework to create a representation for 3D

objects. Before constructing our 3D object signature, we want to identify interesting or salient points on the 3D object and use the characteristics of those points when constructing our signatures. The identified salient points are application dependent. We wanted our framework and methodology to be specifically applicable to classification of craniofacial disorders, such as 22q11.2 deletion syndrome and deformational plagiocephaly, but also be appropriate for general use in 3D shape classification.

To find salient points on a 3D object, a learning approach was selected. A salient point classifier is trained on a set of marked training points on the 3D objects provided by experts in the particular application. Histograms of low-level features of the training points obtained using the base framework are then used to train the classifier. The classifier will then learn the salient points on the 3D objects based on the application, which lead to salient regions in our signatures.

5.1. Learning salient points for 22q11.2 deletion syndrome

Traditionally, studies of individuals with craniofacial disorders such as 22q11.2 deletion syndrome have been performed through in-person clinical observation coupled with craniofacial anthropometric measurements derived from anatomic landmarks [19]. These landmarks are located either visually by clinicians or through palpation of the skull. Fig. 9 shows the landmark points that are commonly used for craniofacial measurements.

The salient point classifier was trained on a subset of the craniofacial anthropometric landmarks marked on 3D head objects. This was done so that these craniofacial landmarks would be included in the set of interesting or salient points for classification of the craniofacial disorders. The particular subset of landmarks was selected to be well-defined points that both experts and non-experts could easily identify. The training set consisted of human heads selected from the Heads database. Fig. 10 shows an example of manually marked salient points on the training data. Histograms of low-level features obtained using the base framework were used to train a support vector machine (SVM) [45,53] classifier to learn the salient points on the 3D surface mesh. We used the SVM implemented in WEKA for our experiments [57]. A training set, consisting of 75 morphs of 5 human heads, was used to train the classifier to learn the characteristics of the salient points for faces in terms of the histograms of their low-level features.

Although the salient training points were selected only to be commonly used craniofacial landmark points, our empirical studies determined that the classifier actually finds salient regions with a combination of high curvature and low entropy values. This result can be observed in the different histograms of salient and non-salient points in Fig. 11. In the figure, the salient point histograms have mainly low bin counts in the bins corresponding

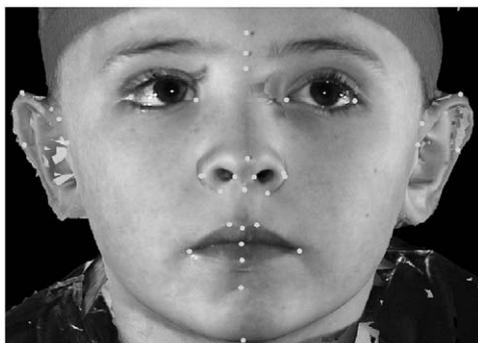


Fig. 9. Craniofacial anthropometric landmarks.

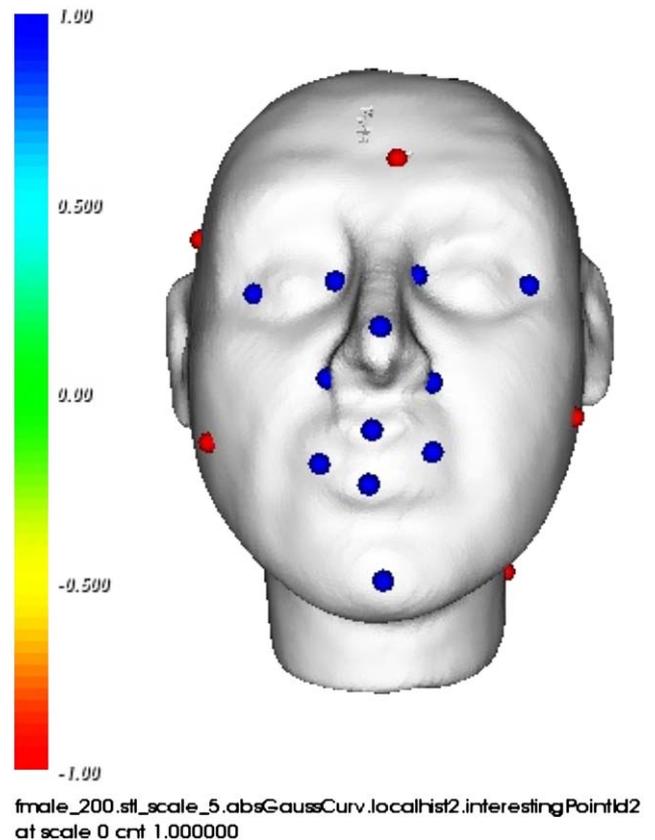


Fig. 10. Example of manually marked salient (blue color) and non-salient (red color) points on a human head model. The salient points include corners of the eyes, tip of the nose, corners of the nose, corners of the mouth, and chin. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to low curvature values and a high bin count in the last (highest) curvature bin. The non-salient point histograms have mainly medium to high bin counts in the low curvature bins and in some cases a high bin count in the last bin. The entropy of the salient point histograms also tends to be lower than the entropy of the non-salient point histograms. The classifier approach avoided the use of brittle thresholds.

Fig. 12 shows results of the salient points predicted on two faces in the 22q11.2DS database, which include not just the manually marked points but other points with the same characteristics. The salient points are colored according to the assigned classifier confidence score. Non-salient points are colored in red, while salient points are colored in different shades of blue with dark blue having the highest prediction score.

5.2. Learning salient points for deformational plagiocephaly

A similar learning-based approach was used to find salient points for 3D heads with deformational plagiocephaly. The salient point classifier for deformational plagiocephaly was trained on a set of points marked on the flat areas at the back of the head of individuals with deformational plagiocephaly. The training salient points consisted of 10 marked points on the flat areas of 10 heads with deformational plagiocephaly, while the non-salient training points were selected from 10 heads without deformational plagiocephaly. Histograms of the azimuth–elevation low-level features obtained using the base framework were used to train a support vector machine (SVM) classifier to learn the salient points on the 3D heads. After training was complete, the classifier was

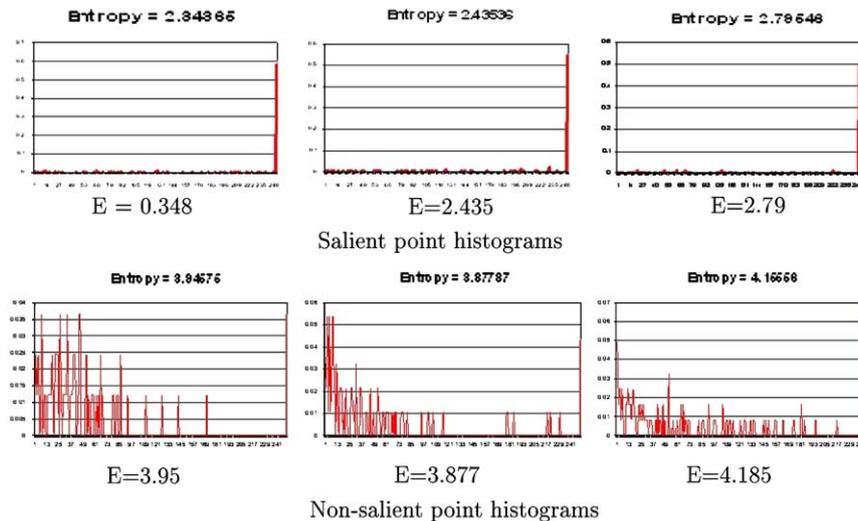


Fig. 11. Example histograms of salient and non-salient points. The salient point histograms have a high value in the last bin illustrating a high curvature in the region, while low values in the remaining bins in the histogram. The non-salient point histograms have more varied values in the curvature histogram. In addition, the entropy E of the salient point histogram is lower than the non-salient point histogram (listed under each histogram).

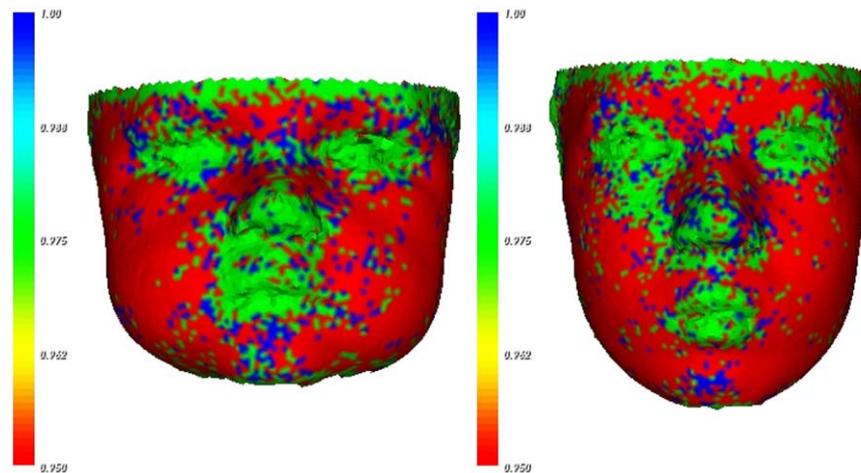


Fig. 12. Salient point prediction for two faces in the 22q11.2DS dataset. Non-salient points are colored in red, while salient points are colored in different shades ranging from green to blue, depending on the classifier confidence score assigned to the point. A threshold ($T = 0.95$) was applied to include only salient points with high confidence scores. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

able to label each point on a 3D head as either salient or non-salient and provide a confidence score for each decision. The same threshold, $T = 0.95$, was applied to the confidence scores for the salient points.

5.3. Learning salient points for general 3d object classification

The salient point classifier for general 3D object classification was trained on selected objects from the Heads database using the craniofacial landmark points that were used in the 22q11.2DS application. A small training set consisting of 25 morphs of the cat head model, 25 morphs of the dog head model, and 50 morphs of human head models was used to train the classifier to learn the characteristics of salient points for general 3D object classification. Histograms of low-level features obtained using the base framework were used to train a support vector machine (SVM) classifier to learn the salient points on general 3D objects. A threshold $T = 0.95$ was also applied to the confidence scores for the classifier salient points. Fig. 13 shows results of the salient

points predicted on instances of the cat, dog and human head class in the Heads, which include, as previously mentioned, not just the manually marked points, but other points with the same characteristics. The salient points are colored according to the assigned classifier confidence score. Non-salient points are colored in red, while salient points are colored in different shades of blue with dark blue having the highest prediction score. While the classifier was only trained on cat heads, dog heads, and human heads, it does a good job of finding salient points on the other classes of heads, and the 3D patterns produced are repeatable across objects of the same class. Fig. 14 shows the predicted salient points on new object classes that were not included in the training phase. We have also tested the trained classifier on our fourth dataset, the SHREC 2008 Classification database, and found that the labeled salient points were quite satisfactory, even though not intended for that purpose. Fig. 15 shows the salient points predicted on a number of objects from the SHREC 2008 database. Note that on this database, which has a lot of intra-class shape variance, the salient point patterns are not consistent across all members of each class.

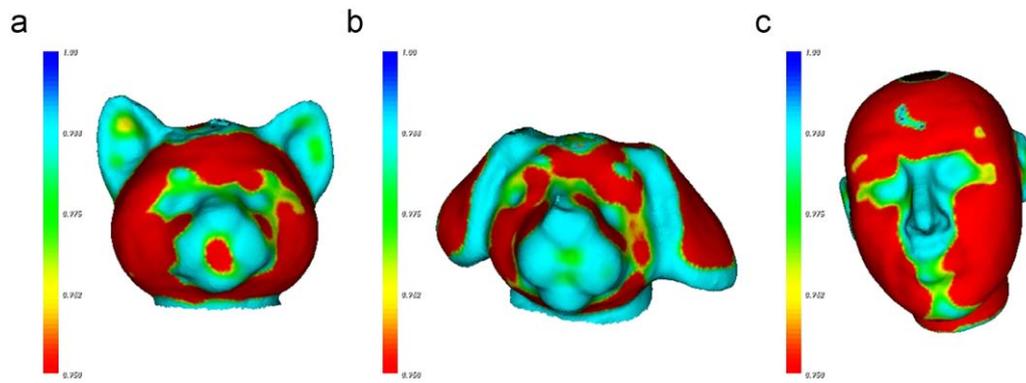


Fig. 13. Salient point prediction for (a) cat head class, (b) dog head class, and (c) human head class. Non-salient points are colored in red, while salient points are colored in different shades ranging from green to blue, depending on the classifier confidence score assigned to the point. A threshold ($T = 0.95$) was applied to include only salient points with high confidence scores. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

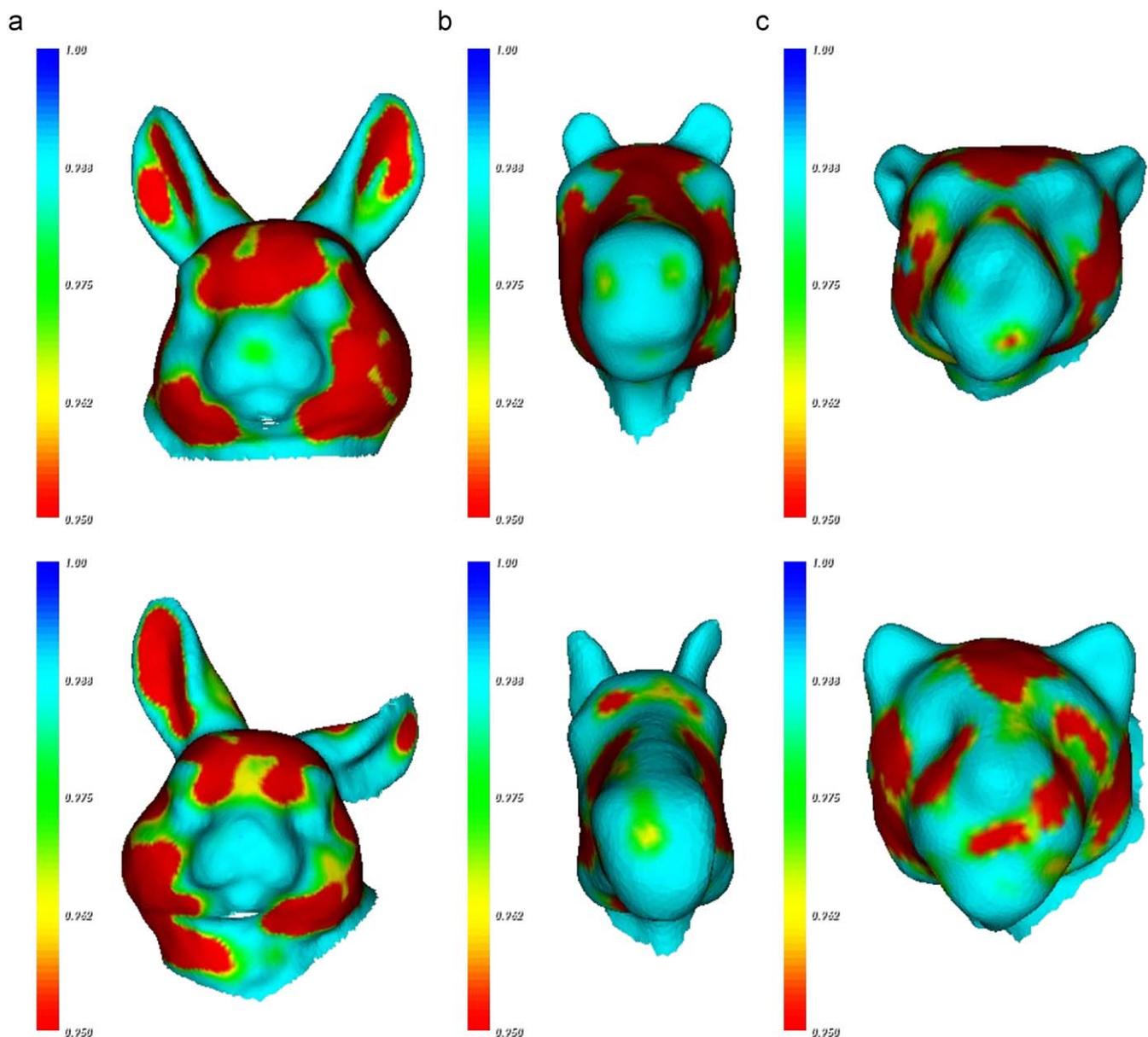


Fig. 14. Salient point prediction for (a) rabbit head class, (b) horse head class, and (c) leopard head class from the heads database. Even though all three classes were not included in the training, the training model was able to predict salient points across the classes.

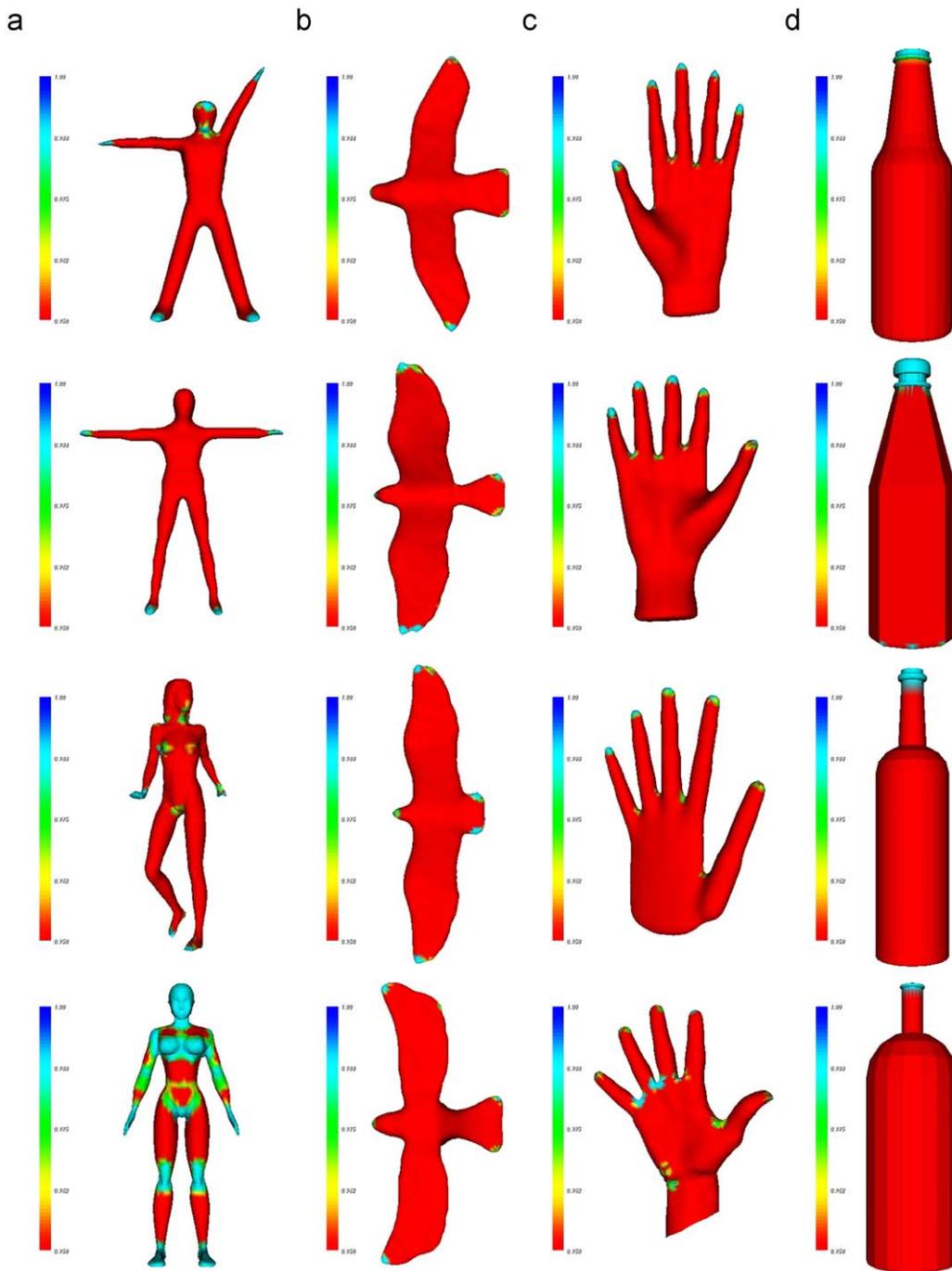


Fig. 15. Salient point prediction for (a) human class, (b) bird class, (c) human hand class, and (d) bottle class from the SHREC 2008 database. Note that for classes that have a lot of intra-class shape variance the salient point patterns are not consistent across all members of those classes as seen in column (a).

6. 2D longitude–latitude map signature

Most 3D object classification methods require the use of a 3D descriptor or signature to describe the shape and properties of the 3D objects. Our signature is based on the salient point patterns of the 3D object mapped onto a 2D plane via a longitude–latitude transformation. Classification of 3D objects is then performed by training a classifier using the 2D spatial maps of the objects.

Before mapping the salient point patterns onto the 2D plane, the salient points are assigned a label according to the classifier confidence score assigned to the point. The classifier confidence score range is discretized into a number of bins. For our experiments, at confidence level 0.95 and above, we chose to discretize the confidence

score range into 5 bins. Each salient point on the 3D mesh is assigned a label based on the bin into which its confidence score falls.

To obtain the 2D longitude–latitude map signature for an object, we calculate the longitude and latitude positions of all the 3D points on the object’s surface. Given any point $p_i (p_{ix}, p_{iy}, p_{iz})$, the longitude position θ_i and latitude position ϕ_i of point p_i are calculated as follows:

$$\theta_i = \arctan\left(\frac{p_{iz}}{p_{ix}}\right), \quad \phi_i = \arctan\left(\frac{p_{iy}}{\sqrt{(p_{ix}^2 + p_{iz}^2)}}\right)$$

where $\theta_i = [-\pi, \pi]$ and $\phi_i = [-\pi/2, \pi/2]$.

A 2D map of the longitude and latitude positions of all the points on the object's surface is created by binning the longitude and latitude values of the points into a fixed number of bins. A bin is labeled with the salient point label of the points that fall into that bin. If more than one label is mapped to a bin, the label with the highest count is used to label the bin. Fig. 16 shows salient point patterns for the cat head, dog head, and human head model in the Heads database and their corresponding 2D map signatures. Fig. 17 shows how different objects that belong to the same class will have similar 2D longitude–latitude signature maps.

To reduce noise in the 2D longitude–latitude map signatures, we applied a wavelet transformation to the 2D map signatures. In our experiments, we treated the 2D longitude–latitude map signatures as 2D images and decomposed the 2D images using image-based Haar wavelet function. The wavelet function decomposes the 2D image into approximation and detail coefficients. The approximation and detail coefficients at the second level were collected and concatenated into a new feature vector with dimension $D = 13 \times 13 \times 4$. This final feature vector became the descriptor for each object in the database and was used for classification. For most of the experiments, the noise

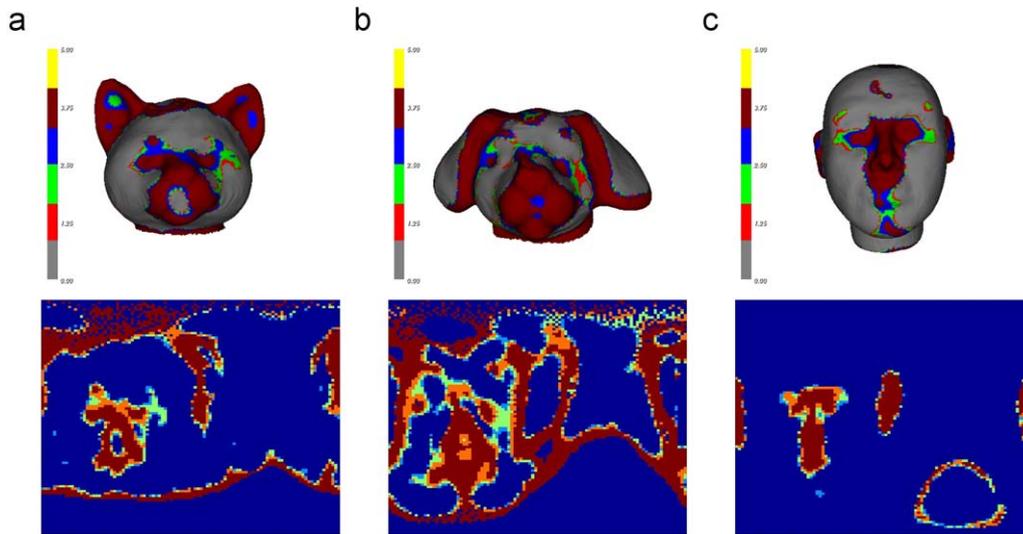


Fig. 16. Salient point patterns on 3D objects of Fig. 12 and their corresponding 2D longitude–latitude map signatures.

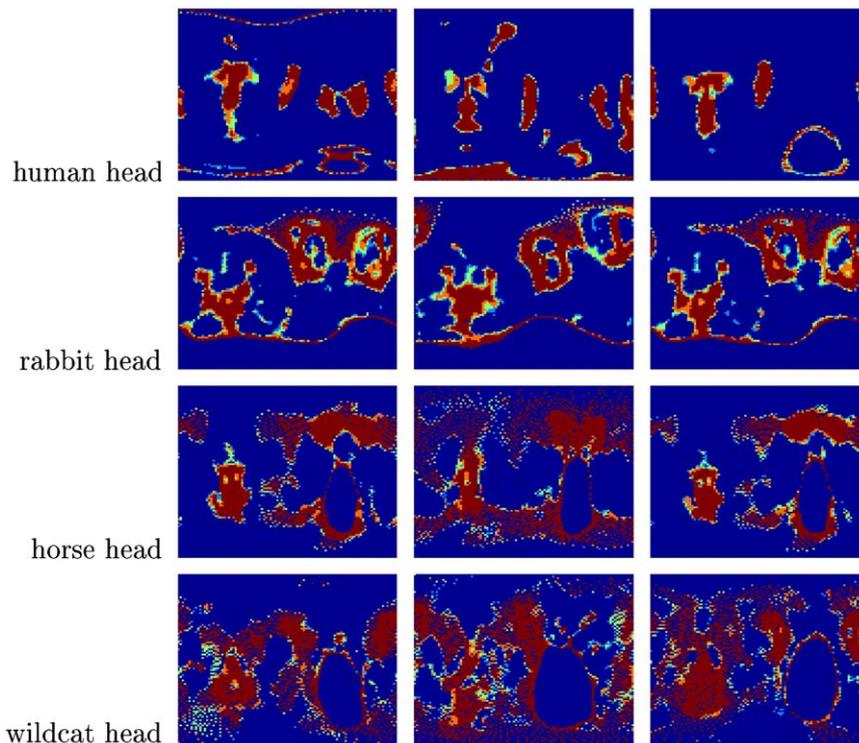


Fig. 17. Objects that are similar and belong to the same class will have similar 2D longitude–latitude signature maps.

reduction step was not found to improve the classification accuracy except for the SHREC dataset (Section 7.4).

7. Experimental results

By creating a signature for each of the 3D objects, we are able to perform classification of the 3D objects in the database. We performed several classification experiments on each of the acquired datasets.

7.1. Experiments on 22q11.2DS dataset

The goal of this experiment is to classify each individual in the dataset as either affected or unaffected and to measure the classification accuracy. The salient points classifier was trained on a subset of the craniofacial anthropometric landmarks marked on 3D human head models as explained in Section 5. Table 2 shows the classification performance with two different classifiers: Adaboost and SVM. The classification accuracy for the higher scoring SVM classifier is 86.7%, beating a study of three human experts whose mean accuracy was 78% [56].

The classification accuracy of our methodology was compared to some of the state-of-the-art and best performing 3D object descriptors in the literature. The following existing descriptors were used for comparison: light field descriptor (LFD) [16], ray-based spherical harmonics (SPH) [29], shape distribution of distance between random points (D2) [39], and absolute angle distance histogram (AAD) [37]. The light field descriptor (LFD) is a view-based descriptor that extracts features from 100 2D silhouette image views and measures the distance between two 3D objects by finding the best correspondence between the set of 3D views for the two objects. The Spherical Harmonics method calculates the maximal extent of a shape across all rays from the origin and uses spherical harmonics to represent the function. The shape function D2 represents 3D objects by calculating the global shape distribution of distances between two random points, while the AAD method enhances the D2 shape function by measuring not only the distance between two random points, but also the mutual orientation of the surfaces on which the pair of points is located. Table 3 provides an overall comparison of our method with the four existing shape descriptors. Results in Table 4 show that our method achieves higher classification

Table 2
Classification performance for 22q11.2DS.

Classifier	Accuracy	Prec.	Recall	F-measure	TP rate	FP rate
Adaboost	0.804	0.795	0.804	0.791	0.804	0.387
SVM	0.867	0.866	0.868	0.861	0.868	0.27

Table 3
Overall comparison of the various shape descriptors.

	LFD	SPH	D2	AAD	Our method
Type	Global view-based	Global spatial map	Global feat.dist	Global feat. dist.	Global local feat.
Efficiency	Medium	Fast	Fast	Fast	Medium
Pose-normalization	No	No	Yes	Yes	Yes
Discriminative power for large shape diff.	High	High	Medium	Medium	Medium
Discriminative power for subtle shape diff.	Medium	Medium	Low	Low	High
Applications	General 3D	General 3D	General 3D	General 3D	General 3D Medical 3D

accuracy of 22q11.2 deletion syndrome than any of these state-of-the-art methods.

Classification of 22q11.2DS will lead to better understanding of the connection between the 22q11.2 deletion syndrome genotype and the phenotype of this syndrome. Being able to connect facial features to the genetic code will allow for understanding the etiology of craniofacial malformation and pathogenesis of 22q11.2DS, which, in turn, will be informative of the genetic control needed for normal craniofacial development.

7.2. Experiments on plagiocephaly dataset

The goal of this experiment is to classify each individual as either control or case affected by the plagiocephaly condition and to measure the classification accuracy. The salient points for the map signature were obtained by using the salient flat point classifier as explained in Section 5.

Table 5 shows the classification accuracy of the method on the full 254 individual dataset. The groundtruth for the classification was the referral doctors' originally assigned patient status: case or control. Table 6 shows the classification accuracy of the method on the trimmed 140-individual dataset in which the experts agreed. The Adaboost classifier obtains a 80.3% classification accuracy on the full dataset and an improved 87.9% accuracy on the trimmed dataset.

Table 4
Comparison of classification accuracy for 22q11.2DS.

Dataset	Our method	LFD	SPH	D2	AAD
F189	0.867	0.741	0.746	0.619	0.73

Table 5
Classification performance for plagiocephaly using the full 254 individuals dataset.

Classifier	Accuracy	Prec.	Recall	F-measure	TP rate	FP rate
Adaboost	0.803	0.805	0.803	0.804	0.803	0.208
SVM	0.787	0.787	0.787	0.787	0.787	0.233

Table 6
Classification performance for plagiocephaly using the trimmed 140 individuals dataset.

Classifier	Accuracy	Prec.	Recall	F-measure	TP rate	FP rate
Adaboost	0.879	0.878	0.879	0.878	0.879	0.156
SVM	0.85	0.849	0.85	0.849	0.85	0.19

Table 7
Comparison of classification accuracy for plagiocephaly.

Dataset	Our method	LFD	SPH	D2	AAD
Full 254 dataset	0.803	0.72	0.673	0.650	0.685
Trimmed 140 dataset	0.879	0.714	0.743	0.779	0.721

The classification accuracy of the methodology for this application was also compared to existing state-of-the-art descriptors. Table 7 shows that our method achieves higher classification accuracy of deformational plagiocephaly compared to other existing methods, including the LFD descriptor and others discussed in Section 2.

Classification of this condition can be incorporated into epidemiologic research on the prevalence and long-term outcome of deformational plagiocephaly, which may eventually lead to improved clinical care for infants with deformational plagiocephaly.

7.3. Experiments on Heads dataset

The Heads database can be thought of as a first step toward testing our method on more general shapes still in the craniofacial category, but for multiple different animals where face shapes can be quite different.

In the first set of experiments, all objects in the Heads database were pose-normalized by rotating the heads to face the same orientation, as was the case for the medical craniofacial datasets. Classification of the 3D objects in the database was performed by training a SVM classifier on the salient point patterns of each class using the 2D longitude–latitude map signature of the objects in the class. The classifier was trained using the signatures of 25 objects from each class for all seven classes in the database and tested the classifier model with a new test set consisting of 50 objects per class for each of the seven classes. The classifier achieved 100% classification accuracy in classifying all the pose-normalized objects in the database.

Since 3D objects may be encountered in the world at any orientation, rotation-invariant classification is desirable. The second set of experiments explored rotation invariance. To achieve rotation invariance for classification, we trained the classifier with a number of rotated versions of the 2D map longitude–latitude signature for each training object. The first experiment in this set tested the classification accuracy by training a classifier with rotated versions of the training data signatures in 45° increments for all three axes. This resulted in $8 \times 8 \times 8$ rotated signatures for each object in the database. The classifier was tested on new objects in the same classes. Rotated versions of the testing data signatures were generated using the same rotation degree increments as in the training. The classifier again achieved 100% classification accuracy when classifying objects that were rotated in this way.

In the second experiment in this set, the classification method was tested using 15 new testing instances per class that were rotated randomly. For example, a rotation of (250, 4, 187) was one of the random rotations that did not match any of the training rotations. The classifier was still able to achieve 100% classification accuracy.

The third set of experiments was to explore the degradation in the classification accuracy by varying the training rotation angle increment when generating the signatures for the training data. Fig. 18 shows the degradation in the classification accuracy as the training angle increment increases and the number of rotated training signature instances decreases. The graph shows that the classification accuracy steadily decreases as the number of rotated

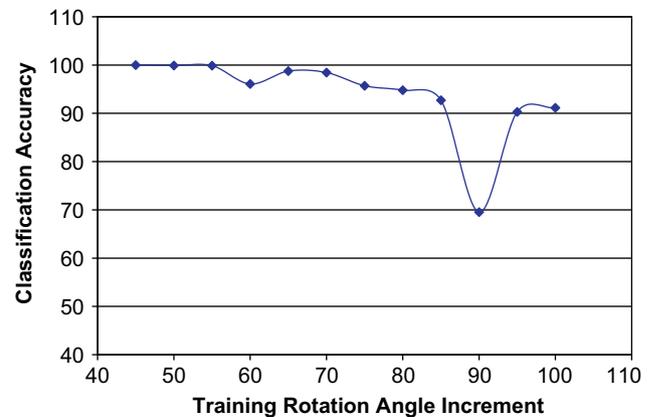


Fig. 18. Classification accuracy vs. training rotation angle increment.

training signatures decreases. In addition, there is a big dip in the classification performance when the training signatures are generated at 90° angle increments. This is because the signatures produced at 90° increments are not representative of angles in between the multiples of 90°. Note that the classifier is still able to achieve 91% classification accuracy with training signatures generated at 100° increments with only $3 \times 3 \times 3 = 27$ rotated training signatures per training object, which is much better than the $8 \times 8 \times 8 = 512$ signatures that were originally used.

7.4. Experiments on SHREC dataset

The SHREC dataset was used to challenge our method on data unlike those it was designed for and to compare it to other methodologies that were designed for more general object models and for many different classes. For this dataset, rotational invariance was a requirement. To achieve it, we tested two different pose-normalization methods. The first method, *4ContPCA*, is an extension to the commonly used principal component analysis (PCA) method that aligns 3D models to a canonical coordinate system. *4ContPCA* extends the PCA method by taking the mesh resolution and sizes of the triangles into consideration when aligning the models [54]. The second approach to achieve rotation-invariant classification, *IncRot*, was to rotate each 3D object at 100° increments for all three axes and generated the 2D longitude–latitude map signature of the object at each rotated pose. This resulted in a total of $3 \times 3 \times 3 = 27$ map signatures for each object in the database. In both the *4ContPCA* approach and the *IncRot* approach, the distance between two objects in the database was the minimum distance between the various rotated map signatures of the two objects. Since the *IncRot* method had better classification performance than the *4ContPCA*, we will only report the classification results using the *IncRot* pose-normalization method. The map signatures were further transformed using the wavelet computation. The wavelet coefficient feature vectors were used for classification.

As with the medical datasets, we compared our results against the light field descriptor (LFD) [16], the ray-based spherical harmonics (SPH) [29], the shape distribution of distance between random points (D2) [39], and the absolute angle distance histogram (AAD) [37]. Since the number of objects in each class in the dataset varied greatly, creating an unbalanced dataset, we were not able to use machine learning algorithms such as SVM and Adaboost to classify the objects in the dataset. As a result, we computed the pairwise distance matrix between every object in the dataset, and measured the classification performance using four different commonly used statistics: (1) nearest-neighbor

classification accuracy, (2) first-tier classification accuracy, (3) second-tier classification accuracy, and (4) F-measure. The first three statistics indicate the percentage of the top K nearest neighbor of a given object to be classified. The nearest neighbor statistics provide an indication of how well a nearest neighbor classifier performs where $K=1$. The first-tier and second-tier statistics indicate the percentage of top K matches that belong to the same class as a given object where $K=C-1$ and $K=2(C-1)$, respectively, where C is the class size of the classified object. The F-measure is a composite measure of precision (P) and recall (R) where $F=2 * P * R / (P + R)$. Table 8 shows the comparison results. For this dataset, the LFD method, which was developed to differentiate between very different shape classes, rather than subtle distinctions in the shape of a common object, was the best performer.

7.5. Timing studies

We have performed timing experiments to investigate the runtime performance of our methodology. In this experiment, we compare the runtime speed of our method to the existing light field descriptor (LFD) method on all four datasets. These experiments were performed on a PC running Windows XP. The runtime performance of the light field descriptor can be divided into two main phases: feature extraction and feature comparison and classification. The runtime performance of our methodology can be divided into five main phases: (1) low-level feature extraction, (2) mid-level feature aggregation, (3) salient point prediction, (4) signature generation, and (5) classification. Table 9 shows the runtime for each of the phases of the light field descriptor on all four datasets, while Table 10 lists the runtime for

Table 8
Comparison of classification accuracy for SHREC 2008 dataset.

Method	NN	1st tier	2nd tier	F-measure
AbsGaussCurv	0.569	0.285	0.375	0.246
Besljain	0.516	0.278	0.379	0.244
LFD	0.759	0.437	0.549	0.365
SPH	0.715	0.365	0.483	0.321
D2	0.502	0.278	0.382	0.238
AAD	0.549	0.266	0.388	0.252

Table 9
Timing of each phase of the light field descriptor on the four datasets.

Phase	22q11.2DS (189)	Plagiocephaly (254)	Heads (105)	SHREC (425)
Feat. extraction	23.6 m	47.1 m	21.4 m	63.3 m
Feat. compare and classify	14 m	17 m	13 m	140 m

m refers to minutes while h refers to hours. The number of objects in each dataset is listed in brackets.

Table 10
Timing of each phase of our methodology on the four datasets.

Phase	22q11.2DS (189)	Plagiocephaly (254)	Heads (105)	SHREC (425)
Low-level feat. extraction	1.5 m	3 m	1 m	4 m
Mid-level feat. aggregation	7 m	7 m	3 m	32 m
Salient point prediction	51 m	19 m	106 m	53 m
Signature generation	2 m	2 m	2 m	11 m
Signature classification	2 m	1 m	1 m	2 m

m refers to minutes. The number of objects in each dataset is listed in brackets.

each of the phases of our methodology on the same four datasets. The bottleneck of our method is in the salient point prediction phase where the classifier labels each point on the mesh as either salient or non-salient. Depending on the salient point learned model and the number of points on the objects in the dataset, this phase may take a longer time, however, accuracy of results, not speed is most important in the medical applications.

8. Conclusion

We propose a new base framework and methodology for 3D object representation and discuss its application in 3D object classification. The methodology starts by extracting and aggregating low-level features in the base framework. Motivated by existing collaborations on classification of craniofacial disorders, a learning approach to identify interesting or salient points was used. The classifier learns the characteristics of interesting points based on the extracted feature values in a neighborhood of each point. The salient point patterns of the 3D objects are then mapped onto a 2D plane via a longitude–latitude transformation. Classification of various 3D objects is then performed by using the map signatures to train a classifier.

We have tested our methodology on two different medical craniofacial applications: classification of 22q11.2 deletion syndrome and deformational plagiocephaly. Experimental results show that our methodology achieves higher classification accuracy compared to both medical experts and existing state-of-the-art descriptors. This methodology will be used in medical research studies whose purpose is to reveal genotype–phenotype disease associations. From a clinical standpoint, offering a standard automated filter may aid physicians in concentrating on the more difficult cases and provide insights into the shapes that are considered most telling for a specific dysmorphological syndrome. A further strength of our methodology is its flexibility to generalize from specific medical applications to general 3D object classification as shown in experimental results on two additional 3D object datasets.

In future work, we want to extend the same framework and methodology to produce local shape signatures instead of global map signatures. We hope that the local signatures will provide more discriminative power and allow us to explore other problems such as partial matching.

Acknowledgments

This research was supported by the National Science Foundation under Grant number DBI-0543631 (PI: L. Shapiro), Grant number 1 R01 HD046565 National Institute of Child Health and Human Development (NICHD) to Dr Speltz, and Grant number K23-DE017741 to Dr Heike. We would also like to thank our collaborators at the Seattle Children's Hospital for providing the medical datasets and for their discussion and collaboration, especially to Michael Cunningham, MD, PhD, Matthew Speltz, MD, Jaqueline Starr, PhD, Brentt Collett, PhD, and Anne Hing.

References

- [1] E. Akagunduz, I. Ulusoy, 3D object representation using transform and scale invariant 3D features, in: ICCV 3dRR Workshop, 2007, pp. 1–8.
- [2] J. Assfalg, A.D. Bimbo, P. Pala, Content-based retrieval of 3D models through curvature maps: a cbr approach exploiting media conversion, *Multimedia Tools Application* 31 (2006) 29–50.
- [3] I. Atmosukarto, L. Shapiro, J.R. Starr, C.L. Heike, B. Collett, M.L. Cunningham, M. Speltz, 3d head shape quantification for infants with and without deformational plagiocephaly, *The Cleft-Palate Craniofacial Journal* (2009).

- [4] I. Atmosukarto, L.G. Shapiro, A learning approach to 3D object representation for classification, in: International Workshop on Statistical+Structural and Syntactic Pattern Recognition, 2008.
- [5] I. Atmosukarto, L.G. Shapiro, A salient-point signature for 3D object retrieval, in: ACM Multimedia Information Retrieval, 2008.
- [6] I. Atmosukarto, L.G. Shapiro, M. Cunningham, M. Speltz, Automatic 3D shape severity quantification and localization for deformational plagiocephaly, in: SPIE Medical Imaging: Image Processing, vol. 7259, 2009, pp. 725852–725958.
- [7] Autodesk, 2009. 3dsmax <<http://www.autodesk.com>>.
- [8] D. Becker, T. Pilgram, L. Martyr-Grames, D. Govier, J. Marsh, A. Kane, Accuracy in identification of patients with 22q11.2 deletion by likely care providers using facial photographs, *Plastic Reconstructive Surgery* 114 (6) (2004) 1367–1372.
- [9] P.J. Besl, R.C. Jain, Three-dimensional object recognition, *Computing Surveys* 17 (1) (1985) 75–145.
- [10] B. Blantz, T. Vetter, A morphable model for the synthesis of 3D faces, *ACM Transactions on Graphics* (1999) 187–1994.
- [11] S. Boehringer, T. Vollmar, C. Tasse, R.P. Wurtz, G. Gillissen-Kaesbach, B. Horsthemke, D. Wiczorek, Syndrome identification based on 2d analysis software, *European Journal of Human Genetics* 14 (2006) 1082–1089.
- [12] B. Bustos, D. Keim, D. Saupe, T. Schreck, Content-based 3D object retrieval, *IEEE Computer Graphics and Applications* 27 (4) (2007) 22–27 (Special Issue on 3D Documents).
- [13] B. Bustos, D. Keim, D. Saupe, T. Schreck, D. Vranic, Feature-based similarity search in 3D object databases, *ACM Computing Surveys* 37 (4) (2005) 345–387.
- [14] B. Bustos, D. Keim, D. Saupe, T. Schreck, D. Vranic, An experimental effectiveness comparison of methods for 3D similarity search, *International Journal on Digital Libraries* 6 (1) (2006) 39–54.
- [15] U. Castellani, M. Cristani, S. Fantoni, V. Murino, Sparse point matching by combining 3D mesh saliency with statistical descriptors, *Computer Graphics Forum* 27 (2) (2008) 643–652.
- [16] D. Chen, X. Tian, Y. Shen, M. Ouhyoung, On visual similarity based 3D model retrieval, *Computer Graphics Forum* 22 (3) (2003) 223–232.
- [17] A. Del Bimbo, P. Pala, Content-based retrieval of 3D models, *ACM Transactions on Multimedia Computing, Communications, and Applications* 2 (1) (2006) 20–43.
- [18] M. Elad, A. Tal, S. Ar, Content based retrieval of vml objects: an iterative and interactive approach, in: Eurographics workshop on Multimedia 2001, pp. 107–118.
- [19] L. Farkas, M. Katic, C. Forrest, International anthropometric study of facial morphology in various ethnic groups/races, *The Journal of Craniofacial Surgery* 16 (4) (2005) 615–646.
- [20] A. Frome, D. Huber, R. Kolluri, T. Bulow, J. Malik, Recognizing objects in range data using regional point descriptors, in: Proceedings of the European Conference on Computer Vision (ECCV), 2004, pp. 224–237.
- [21] T. Funkhouser, M. Kazhdan, P. Min, P. Shilane, Shape-based retrieval and analysis of 3D models, *Communications of the ACM* 48 (6) (2005) 58–64.
- [22] D. Giorgi, S. Martini, Shape retrieval contest 2008: classification of water tight models, in: IEEE Shape Modeling and Applications, 2008, pp. 219–220.
- [23] P. Hammond, The use of 3D face shape modeling in dysmorphology, *Archives of Disease in Childhood* 92 (2007) 1120–1126.
- [24] M. Hilaga, Y. Shinagawa, T. Kohmura, Topology matching for fully automatic similarity estimation of 3D shapes, in: SIGGRAPH, 2001, pp. 203–212.
- [25] B. Hutchison, L. Hutchison, J. Thompson, E.A. Mitchell, Plagiocephaly and brachycephaly in the first two years of life: a prospective cohort study, *Pediatrics* 114 (2004) 970–980.
- [26] L. Hutchison, L. Hutchison, J. Thompson, E.A. Mitchell, Quantification of plagiocephaly and brachycephaly in infants using a digital photographic technique, *The Cleft Palate-Craniofacial Journal* 42 (5) (2005) 539–547.
- [27] N. Iyer, S. Jayanti, K. Lou, Y. Kalyanaraman, K. Ramani, Three-dimensional shape searching: state-of-the-art review and future trends, *Computer Aided Design* 37 (2005) 509–530.
- [28] A. Johnson, M. Hebert, Using spin images for efficient object recognition in cluttered 3D scenes, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21 (5) (1999) 433–449.
- [29] M. Kazhdan, T. Funkhouser, S. Rusinkiewicz, Rotation invariant spherical harmonic representation of 3D shape descriptors, in: Eurographics, 2003, pp. 156–164.
- [30] L. Kobrynski, K. Sullivan, Velocardiofacial syndrome, digeorge syndrome: the chromosome 22q11.2 deletion syndromes, *The Lancet* 370 (9596) (2007) 1443–1452.
- [31] H. Laga, M. Nakajima, A boosting approach to content-based 3D model retrieval, in: GRAPHITE, 2007, pp. 227–234.
- [32] H. Laga, H. Takahashi, M. Nakajima, Spherical wavelet descriptors for content-based 3D model retrieval, in: Shape Modeling and Applications, 2006, pp. 15–23.
- [33] Z. Liu, J. Mitani, Y. Fukui, S. Nishihara, Multiresolution wavelet analysis of shape orientation for 3D shape retrieval, in: ACM Multimedia Information Retrieval, 2008, pp. 403–410.
- [34] K. Lou, N. Iyer, S. Jayanti, Y. Kalyanaraman, S. Prabhakar, K. Ramani, Effectiveness and efficiency of three-dimensional shape retrieval, *Journal of Engineering Design* 16 (2) (2005) 175–194.
- [35] J. Novatnack, K. Nishino, Scale-dependent 3D geometric features, in: International Conference on Computer Vision, 2007, pp. 1–8.
- [36] J. Novatnack, K. Nishino, A. Shokoufandeh, Extracting 3D shape features in discrete scale space, in: 3D Data Processing Visualization, and Transmission, 2006, pp. 946–953.
- [37] R. Ohbuchi, T. Minamitani, T. Takei, Shape-similarity search of 3D models by using enhanced shape functions, *International Journal of Computer Applications in Technology* 23 (2) (2005) 70–85.
- [38] R. Ohbuchi, K. Osada, T. Furuya, T. Banno, Salient local visual features for shape-based 3D model retrieval, in: Shape Modeling International, 2008, pp. 93–102.
- [39] R. Osada, T. Funkhouser, B. Chazelle, D. Dobkin, Shape distributions, *ACM Transactions on Graphics* 21 (2002) 807–832.
- [40] G. Passalis, T. Theoharis, Intraclass retrieval of nonrigid 3D objects: Application to face recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29 (2) (2007) 218–229 member-loannis A. Kakadiaris.
- [41] Z. Qin, J. Jia, J. Qin, Content based 3D model retrieval: a survey, in: International Workshop CBMI, 2008, pp. 249–256.
- [42] J.T. Richtsmeier, V. DeLeon, S. Lele, The promise of geometric morphometrics, *Yearbook of Physical Anthropology* 45 (2002) 63–91.
- [43] S. Ruiz-Correa, L. Shapiro, M. Meila, G. Berson, M. Cunningham, R. Sze, Symbolic signatures for deformable shapes, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (1) (2004) 75–90.
- [44] D. Saupe, D. Vranic, 3D model retrieval with spherical harmonics and moments, in: DAGM Symposium on Pattern Recognition, 2001, pp. 392–397.
- [45] B. Scholkopf, A.J. Smola, *Learning with Kernels*, Cambridge University Press, Cambridge, 2002.
- [46] L.G. Shapiro, K. Wilamowska, I. Atmosukarto, J. Wu, C.L. Heike, M. Speltz, M. Cunningham, Shape-based classification of 3D head data, in: International Conference on Image Analysis and Processing, 2009.
- [47] P. Shilane, T. Funkhouser, Selecting distinctive 3D shape descriptors for similarity retrieval, in: Shape Modeling International, 2006, pp. 18–25.
- [48] P. Shilane, T. Funkhouser, Distinctive regions of 3D surface, *ACM Transactions on Graphics* 26 (2) (2007).
- [49] P. Shilane, P. Min, M. Kazhdan, T. Funkhouser, The Princeton shape benchmark, in: Shape Modeling International, 2004, pp. 167–178.
- [50] H. Sundar, D. Silver, N. Gagvani, S. Dickenson, Skeleton-based shape matching and retrieval, in: Shape Modeling International, 2004, pp. 130–138.
- [51] J. Tangelder, R. Veltkamp, A survey of content based 3D shape retrieval methods, in: Shape Modeling International, 2004, pp. 145–156.
- [52] J. Tangelder, R. Veltkamp, A survey of content based 3D shape retrieval methods, *Multimedia Tools Application* 39 (3) (2008) 441–471.
- [53] V.V. Vapnik, *Statistical Learning Theory*, Wiley, New York, 1998.
- [54] D. Vranic, D. Saupe, Tools for 3D object retrieval: Karhunen–Loeve transform and spherical harmonics, in: IEEE Workshop on Multimedia Signal Processing, 2001, pp. 293–298.
- [55] Y. Wang, R. Liu, T. Baba, Y. Uehara, D. Masumoto, S. Nagata, An images-based 3D model retrieval approach, in: Advances in Multimedia Modeling, 2008, pp. 90–100.
- [56] K. Wilamowska, L.G. Shapiro, C.L. Heike, Classification of 3D face shape in 22q11.2 deletion syndrome, in: IEEE International Symposium on Biomedical Imaging, 2009.
- [57] I.H. Witten, E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, second ed., Morgan Kaufmann, San Francisco, 2005.
- [58] J. Wu, K. Wilamowska, L. Shapiro, C. Heike, Automatic analysis of local nasal features in 22q11.2ds affected individuals, in: IEEE Engineering in Medicine and Biology, 2009.
- [59] Y. Yang, H. Lin, Y. Zhang, Content-based 3D model retrieval: a survey, *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews* 37 (6) (2007).
- [60] M. Zonenshayn, E. Kronberg, M. Souweidane, Cranial index of symmetry: an objective semiautomated measure of plagiocephaly, *Journal of Neurosurgery (Pediatrics)* 5 100 (2004) 537–540.

About the Author—INDRIYATI ATMOSUKARTO is currently a Ph.D. student in the Department of Computer Science and Engineering at the University of Washington in Seattle. She is advised by Prof. Linda Shapiro. She has obtained a Master of Science from the University of Washington in 2006 and a Master of Science from the National University of Singapore in 2002. Before that, she obtained her Bachelor in Computer Science from the National University of Singapore in 2000. She was also a research engineer at the National University of Singapore.

Her research interests are in computer vision, computer graphics, and machine learning. Her current research work is in applying computer vision and machine learning techniques to analyze and quantify 3D shapes for similarity-based retrieval and classification for general computer vision applications. She is also currently collaborating with medical doctors at the Seattle Children's Hospital Craniofacial Center to quantify the 3D shape variation of various medical condition.

About the Author—KATARZYNA WILAMOWSKA is currently a National Library of Medicine Informatics Post-doctoral Research Fellow in the Division of Biomedical and Health Informatics of the Department of Medical Education and Biomedical Informatics of the School of Medicine, University of Washington. She received a B.S. degree in Computer Science from the University of Wyoming in 2001 and M.S. and Ph.D. degrees in Computer Science and Engineering from the University of Washington in 2006 and 2009, respectively. Her research interests include translational medicine, bioinformatics, computer vision, and computational biology.

About the Author—CARRIE L. HEIKE, MD, MS is a pediatrician in the Children's Craniofacial Center at Seattle Children's Hospital, and an Assistant Professor in the Department of Pediatrics, University of Washington. She participates in clinical research to assess the etiology of craniofacial conditions and improve care for children with craniofacial disorders.

About the Author—LINDA G. SHAPIRO was born in Chicago, IL, 1949. She received the B.S. degree in mathematics from the University of Illinois, Urbana, in 1970, and the M.S. and Ph.D. degrees in computer science from the University of Iowa, Iowa City, in 1972 and 1974, respectively.

She was an Assistant Professor of Computer Science at Kansas State University, Manhattan, from 1974 to 1978 and was an Assistant Professor of Computer Science from 1979 to 1981 and Associate Professor of Computer Science from 1981 to 1984 at Virginia Polytechnic Institute and State University, Blacksburg. She was Director of Intelligent Systems at Machine Vision International in Ann Arbor from 1984 to 1986. She is currently Professor of Computer Science and Engineering and of Electrical Engineering at the University of Washington. Her research interests include computer vision, image database systems, artificial intelligence, pattern recognition, and robotics. She has co-authored three textbooks, one on data structures and two on computer vision.

Dr. Shapiro is a Fellow of the IEEE and a Fellow of the IAPR. She is a past Chair of the IEEE Computer Society Technical Committee on Pattern Analysis and Machine Intelligence and is currently an editorial board member of *Computer Vision and Image Understanding* and of *Pattern Recognition*. She has served as Editor-in-Chief of *CVGIP: Image Understanding*, Associate Editor of *IEEE Transactions on Pattern Analysis and Machine Intelligence*, co-Program Chair of the IEEE Conference on Computer Vision and Pattern Recognition in 1994, General Chair of the IEEE Workshop on Directions in Automated CAD-Based Vision in 1991, and General Chair of the IEEE Conference on Computer Vision and Pattern Recognition in 1986. She was the Co-Chair of the Medical and Multimedia Applications Track of the International Conference on Pattern Recognition for 2002 and Co-Chair of CVPR 2008. She has also served on the program committees of numerous vision and AI workshops and conferences.