3D Zernike Descriptors for Content Based Shape Retrieval

Marcin Novotni
University of Bonn
Institute of Computer Science II
Römerstr. 164
D-53117 Bonn, Germany
marcin@cs.uni-bonn.de

Reinhard Klein
University of Bonn
Institute of Computer Science II
Römerstr. 164
D-53117 Bonn, Germany
rk@cs.uni-bonn.de

ABSTRACT

Content based 3D shape retrieval for broad domains like the World Wide Web has recently gained considerable attention in Computer Graphics community. One of the main challenges in this context is the mapping of 3D objects into compact canonical representations referred to as descriptors, which serve as search keys during the retrieval process. The descriptors should have certain desirable properties like invariance under scaling, rotation and translation. Very importantly, they should possess descriptive power providing a basis for similarity measure between three-dimensional objects which is close to the human notion of resemblance.

In this paper we advocate the usage of so-called 3D Zernike invariants as descriptors for content based 3D shape retrieval. The basis polynomials of this representation facilitate computation of invariants under the above transformations. Some theoretical results have already been summarized in the past from the aspect of pattern recognition and shape analysis. We provide practical analysis of these invariants along with algorithms and computational details. Furthermore, we give a detailed discussion on influence of the algorithm parameters like type and resolution of the conversion into a volumetric function, number of utilized coefficients, etc. As is revealed by our study, the 3D Zernike descriptors are natural extensions of spherical harmonics based descriptors, which are reported to be among the most successful representations at present. We conduct a comparison of 3D Zernike descriptors against these regarding computational aspects and shape retrieval performance.

Categories and Subject Descriptors

I.4.10 [**Computing Methodologies**]: Image Representation —*Image Representation*; I.3.5 [**Computing Methodologies**]: Computer Graphics—*Computational Geometry and Object Modeling*; H.3.3 [**Information Systems**]: Information Storage and Retrieval—*Information Search and Retrieval*

General Terms

Algorithms, Theory

SM'03, June 16–20, 2003, Seattle, Washington, USA.

Keywords

shape retrieval, 3D Zernike moments, shape descriptor, invariants

1. INTRODUCTION

It can be observed that the proliferation of a specific digital multimedia data type (e.g. text, images, sounds, video) was followed by emergence of systems facilitating their content based retrieval. With the recent advances in 3D data acquisition techniques, graphics hardware and modeling methods, there is an increasing amount of 3D objects spread over various archives: general objects commonly used e.g. in games or VR environments, solid models of industrial parts, etc. On the other hand, modeling of high fidelity 3D objects is a very cost and time intensive process – a task which one can potentially get around by reusing already available models. Another important issue is the efficient exploration of scientific data represented as 3D entities. Such archives are becoming increasingly popular in the areas of Biology, Chemistry, Anthropology and Archeology to name a few. Therefore, since recently, concentrated research efforts are being spent on elaborating techniques for efficient content based retrieval of 3D objects.

One of major challenges in the context of data retrieval is to elaborate a suitable canonical characterization of the entities to be indexed. In the following, we will refer to this characterization as a descriptor. Since the descriptor serves as a key for the search process, it decisively influences the performance of the search engine in terms of computational efficiency and relevance of the results. A simple approach is to annotate the entities with keywords, however, due to the inherent complexity and multitude of possible interpretations this proved to be incomplete, insufficient and/or impractical for almost all data types, cf. [30, 13].

Guided by the fact that for a vast class of objects the shape constitutes a large portion of abstract object information, we focus in this paper on general shape based object descriptors. We now can state some requirements that a general shape based descriptor should obey:

1. **Descriptive power** - the similarity measure based on the descriptor should deliver a similarity ordering that is close to the application driven notion of resemblance.
2. **Conciseness and ease of indexing** - the descriptor should be compact in order to minimize the storage requirements and accelerate the search by reducing the dimensionality of the problem. Very importantly, it should provide some means of indexing and thereby structuring the database in order to further accelerate the search process.
3. **Invariance under transformations** - the computed descriptor values have to be invariant under an application depen-

dent set of transformations. Usually, these are the similarity transformations, however, some applications like e.g. retrieval of articulated objects may additionally demand invariance under certain deformations, etc.

A list of additional requirements may be given, which apply in case of search for general 3D objects that may be found in various archives on the World Wide Web: insensitiveness to noise and small extra features, independence of 3D object representation, tessellation, or genus, robustness against arbitrary topological degeneracies.

In this paper we present a 3D content based retrieval method relying on 3D Zernike moments. These moments are computed as a projection of the function defining the object onto a set of orthonormal functions within the unit ball – the 3D Zernike polynomials introduced by Canterakis [7]. From these Canterakis has derived affine invariant features of 3D objects represented by a volumetric function. To our knowledge, these results have not been applied to the content based retrieval of 3D objects so far. In this paper we perform a comparison with previous description methods. To this end we apply our method to a small database of general objects collected from the WWW and compare the results with those yielded by the recently introduced spherical harmonic descriptors (SH descriptors) [13], which are reported to be among the most powerful at present. As it turns out, the construction of SH and 3D Zernike descriptors is closely related, which enables a detailed comparison regarding the structure and performance.

The outline of the rest of the paper is as follows: in the next section we shortly review the relevant previous work. In Section 3 we present a general theoretical framework for the computation of rotationally invariant descriptors and delineate the 3D Zernike descriptors in this framework. We also examine the 3D Zernike descriptors for accordance with the above criteria and 3D shape retrieval performance. Section 4 gives a discussion on practical issues concerning the implementation of 3D Zernike descriptors paying special attention to numerical stability of computations. In Section 5 we present our results and conclude in Section 6.

2. PREVIOUS WORK

2.1 Systems

Up to date numerous systems for 2D image retrieval have been introduced. To gain a good overview over the state-of-the-art in this area we refer to the survey papers [30, 14, 27]. As for content based retrieval of general 3D objects, the first system was introduced in [25], which was followed by [31]. A very recent result is presented in [13, 22]. Considering systems covering narrower domains, [1] deals with anthropological data, [3, 9] facilitate the retrieval of industrial solid models, [4] explores protein databases.

2.2 Spatial domain

The spatial domain shape analysis methods yield non-numeric results, usually an attributed graph, which encodes the spatial and/or topological structure of an object. Notably, in his seminal work Blum introduced the Medial Axis Transform (MAT) [6], which was followed by a number of extensions like shock graphs, see e.g. [29], shock scaffold [20], etc. Forsyth et al. [12] represent 2D image objects by spatial relationships between stylized primitives, [26] uses a similar approach. A further technique having a long tradition is the geon based representation [5]. As for 3D industrial solid models, [8, 21] capture geometric and engineering features in a graph, which is subsequently used for similarity estimation. Tangelder and Veltkamp [33] describe an approach representing the polyhedral objects as weighted point sets. Hilaga et al. [15] presented a method for general 3D objects utilizing Reeb graphs based on geodesic distances between points on the mesh, which enabled a deformation invariant recognition. The methods in this class are attractive since they capture the high level structure of objects. Unfortunately though, they are computationally expensive, most of them suffer from noise sensitivity, and the underlying graph representation makes the indexing and comparison of objects very difficult.

2.3 Scalar transform

The scalar transform techniques capture global properties of the objects generating numbers (scalars or vectors) as shape descriptors.

2.3.1 Histograms

A number of methods in 2D rely on color histograms measuring the color distribution in an image [32]. In [25] this is generalized to constructing 3D histograms of normal vector, color, material, etc. distributions. Ankerst et al. [4] subdivide the space into shells and sectors around the center of gravity of an object, the resulting partitions correspond to the bins of the 3D shape histogram. Kazhdan et al. describe a reflective symmetry descriptor in [18], Osada et al. [24] compute histograms based on geometric statistics of the boundary of 3D objects. Unfortunately, these descriptors usually provide an insufficient discrimination between objects.

2.3.2 Projection based techniques

Some techniques both in 2D and in 3D are based on coefficients yielded by compression transforms like the cosine [28] or wavelet transform e.g. in [17]. Fourier descriptors [37] have been applied in 2D, however, these are hard to generalize to 3D due to the difficulties in parameterizing 3D object boundaries.

Moments can generally be defined as projections of the function defining the object onto a set of functions characteristic to the given moment. Since Hu [16] popularized the usage of image moments in 2D pattern recognition, they have found numerous applications. Teague [34] was first to suggest the usage of orthogonal functions to construct moments. Subsequently, several 2D moments have been elaborated and evaluated [35]: geometrical, Legendre, Fourier-Mellin, Zernike, pseudo-Zenike moments. 3D geometrical moments have been used in [11, 23], and a spherical harmonic decomposition was used by Vranic and Saupe [36]. The main drawback of these methods is that prior to computations a canonical pose of objects has to be determined, which often proves to be instable, as discussed in [13]. Funkhouser et al. [13] profit from the invariance properties of spherical harmonics and present an affine invariant descriptor. The main idea behind this is to decompose the 3D space into concentric shells and define rotationally invariant representations of these subspaces. In this way a descriptor was constructed which was proven to be superior over other 3D techniques with regard to shape retrieval performance. In [35] 2D Zernike moments were found to be superior over others in terms of noise sensitivity, information redundancy and discrimination power. Guided by this, Canterakis [7] generalized the classical 2D Zernike polynomials to 3D, however, in his work Canterakis considered exclusively theoretical aspects.

3. 3D ZERNIKE MOMENTS AND DESCRIPTORS

This section gives a systematic construction of 3D Zernike moments and descriptors. We attempt to describe a framework pro-

viding a general approach to handle this issue. We also recall the relevant results of Canterakis [7] and describe our improvements.

3.1 Moments

Moments in the context of shape analysis are defined as projections of the (square integrable) object function $f \in L^2$ onto a set of functions $\Psi = \{\psi_i\},\ i \in \mathbb{N}$ over the domain Ω:

$$\mu_i = \langle f, \psi_i \rangle = \int_\Omega f(\mathbf{x}) \cdot \overline{\psi_i(\mathbf{x})}\, d\mathbf{x}.$$

The behavior and properties of a particular moment based representation are therefore determined by the set of functions Ψ.

We now consider the desirable properties of a descriptor based on moments and subsequently give a general formula for computation of moments obeying these properties for the two and three dimensional case.

1. **Invariance.** Let $\mathscr{F}(f)$ be a set of descriptors computed on the function f defining the object, and let G be a group of transformations. The invariance of $\mathscr{F}$ under the action of G can be defined as follows:

   $$\mathscr{F}(gf) = \mathscr{F}(f),$$

   where $g \in G$. A typical requirement is the invariance under the action of similarity transformations, i.e. uniform scaling, reflection, translation and rotation.

2. **Orthonormality.** The collection of functions Ψ is orthonormal, if

   $$\langle \psi_i, \psi_j \rangle = \delta_{ij},$$

   where $\psi_i, \psi_j \in \Psi$ and δ_{ij} is the Kronecker delta.

3. **Completeness.** The set of functions Ψ forms a complete system if for any $f \in L^2$,

   $$\lim_{n\to\infty} ||f - \sum_{i=0}^{n} \langle f, \psi_i \rangle \psi_i||^2 = 0,$$

   where $||f||$ denotes the L_2-norm. Complete orthonormal function collections are said to form a basis of the function space on the domain Ω.

Most approaches transform the object into a canonical pose: translate the center of gravity of the object into the origin and normalize the area/volume or radius of the bounding circle/sphere. The rotation invariance may subsequently be achieved by aligning the principal axes of the object with the coordinate system axes. However, as has been investigated by [13], this last step is often unstable and leads to reduced retrieval performance. Based on these observations, in our choice of Ψ we will favor representations yielding a more stable rotation invariance.

The orthogonality of our function collection, i.e. the mutual independence of computed features is an important property, since it implies that a set of features will not contain redundant information. The non-orthogonality (as in the case of geometrical moments based on monomials) means that some characteristics of the objects will be over-represented during the comparison. The classical 2D Zernike polynomials are orthonormal within the unit circle. They therefore deliver independent features, and are shown to be superior over the geometrical moments in terms of retrieval performance. The additional normalization is essentially a convenience criterion, since this property allows for a canonical formulation of projections of functions.

The completeness property implies that we are able to reconstruct approximations of the original object from moments. The approximations are getting finer with increasing number n of moments and converge to the original object at infinity. This is of considerable practical importance, since the ability to reconstruct allows us to infer a higher bound on the amount of object information encoded by a given number of moments.

3.2 Selection of basis functions

To sum up, we are looking for sets of functions forming complete orthogonal systems and allowing for construction of moments that are invariant under rotation transformations. As it turns out, a straightforward solution is essentially a tensor product formulation and consists of two ingredients. First, one has to choose an angular function set $\{S_l(\varphi)\}$ or $\{S_l^m(\varphi,\vartheta)\}$ defined on the circle or sphere, respectively, that is orthogonal and has subspaces invariant under the action of the rotation group. Second, the circular or spherical function is modulated by a suitable radial function $R_{nl}^m(r)$ while maintaining the orthonormality. Note that in general a particular function R may depend on indices l and m, which implies a dependency on the angular function. Also note that for 3D Zernike moments this dependency is reduced to l, and no dependency at all for 2D Zernike moments.

The general formula for the generation of moments μ possessing the above properties is:

$$\mu_{ln} = \langle f, R_{nl} S_l \rangle = \int\int f(r,\varphi)\overline{R_{nl}(r) S_l(\varphi)} d\varphi dr$$

and

$$\mu_{ln}^m = \langle f, R_{nl}^m S_l^m \rangle = \int\int\int f(r,\varphi,\vartheta)\overline{R_{nl}^m(r) S_l^m(\varphi,\vartheta)} d\vartheta d\varphi dr$$

for the two and three dimensional case, respectively. The choice of an appropriate angular function seems to be crucial, therefore we first summarize some observations that have been made in the 2D case and then move on to 3D.

3.2.1 2D Zernike moments

In 2D, a suitable angular function has proven to be:

$$S_l(\varphi) = e^{il\varphi}, \tag{1}$$

which is essentially the familiar Fourier basis function. It has been shown e.g. by Khotanzad and Hong [19] that for such functions the following relation applies:

$$|\langle f(\varphi + \varphi_0), e^{il(\varphi+\varphi_0)} \rangle| = |\langle f(\varphi), e^{il(\varphi)} \rangle|.$$

This implies that by projecting a function f defined on the circle onto a basis of above functions (Eqn. 1), and computing the norms of these projections, we obtain descriptors of f that are invariant under the action of 2D rotations. The radial polynomial R_n for the 2D Zernike functions is defined so that the resulting basis $R_n S_l$ is orthonormal.

3.2.2 3D Zernike moments

Using the general construction rule derived above, we now derive the 3D Zernike moments (see [7] for details).

Spherical harmonics. Motivated by the facts summarized in the previous subsection and recalling that spherical harmonics on the sphere have properties similar to the functions of Eqn. 1, we continue with the description of spherical harmonics.

Spherical harmonics form a Fourier basis on a sphere much like the familiar sines and cosines do on a line or a circle. Spherical harmonics Y_l^m are given by:

$$Y_l^m(\vartheta,\varphi) = N_l^m P_l^m(\cos\vartheta) e^{im\varphi},$$

where N_l^m is a normalization factor

$$N_l^m = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}},$$

and P_l^m denotes the associated Legendre functions.

Invariance properties. The vector of spherical harmonics

$$\mathbf{Y}_l = (Y_l^l, Y_l^{l-1}, Y_l^{l-2}, \ldots, Y_l^{-l})^t \tag{2}$$

for a given l forms the basis for a $(2l+l)$-dimensional subspace which is invariant under the operations of the full rotation group[1]. This can be formulated as

$$\mathbf{Y}_l(\vartheta+\vartheta_0, \varphi+\varphi_0) = \mathbf{o}_l(\vartheta_0, \varphi_0)\mathbf{Y}_l(\vartheta, \varphi), \tag{3}$$

where $\mathbf{o}_l$ is a unitary matrix referred to as l-th representation of the three dimensional rotation group $SO(3)$. Furthermore, this subspace is irreducible that is, it cannot be split into smaller subspaces which are also invariant under the rotation group. Since rotations do not change the norm of functions, in consequence of Eqn. 3, after projecting a function f defined on the sphere onto the functions of the vector $\mathbf{Y}_l$, we obtain invariant features μ_l of f by computing the norms of the so computed vectors:

$$\mu_l = \left\| \begin{array}{c} \langle f, Y_l^l(\vartheta+\vartheta_0, \varphi+\varphi_0)\rangle \\ \langle f, Y_l^{l-1}(\vartheta+\vartheta_0, \varphi+\varphi_0)\rangle \\ \vdots \\ \langle f, Y_l^{l-1}(\vartheta+\vartheta_0, \varphi+\varphi_0)\rangle \end{array} \right\| = \left\| \begin{array}{c} \langle f, Y_l^l(\vartheta, \varphi)\rangle \\ \langle f, Y_l^{l-1}(\vartheta, \varphi)\rangle \\ \vdots \\ \langle f, Y_l^{l-1}(\vartheta, \varphi)\rangle \end{array} \right\| \tag{4}$$

As a next step, we have to augment this representation to cover the three dimensional space.

Harmonic polynomials. Canterakis based his derivations on harmonic polynomials which finally enabled him to formulate the 3D Zernike polynomials as homogenous polynomials in the Cartesian coordinates *x*, *y* and *z*.

Let us define the conversion between Cartesian and spherical coordinates by $\mathbf{x} = |\mathbf{x}|\xi = r\xi = r(\sin\vartheta\sin\varphi, \sin\vartheta\cos\varphi, \cos\varphi)^T$. The harmonic polynomials e_l^m are defined as

$$e_l^m(\mathbf{x}) = r^l Y_l^m(\vartheta, \varphi).$$

Using the integral formula for associated Legendre functions [10] and converting into Cartesian coordinates, we can express the harmonic polynomials as

$$\begin{aligned} e_l^m(\mathbf{x}) &= c_l^m r^l \left(\tfrac{ix-y}{2}\right)^m z^{l-m} \\ &\quad \cdot \sum_{\mu=0}^{\lfloor \frac{l-m}{2} \rfloor} \binom{l}{\mu}\binom{l-\mu}{m+\mu}\left(-\tfrac{x^2+y^2}{4z^2}\right)^\mu, \end{aligned} \tag{5}$$

where c_l^m are normalization factors:

$$c_l^m = c_l^{-m} = \frac{\sqrt{(2l+1)(l+m)!(l-m)!}}{l!}.$$

The above formula yields homogenous polynomials for $m > 0$. For $m < 0$ the following symmetry relation is used:

$$e_l^{-m}(\mathbf{x}) = (-1)^m \overline{e_l^m(\mathbf{x})}, \tag{6}$$

which yields homogenous polynomials in this case as well. It is easy to see that an invariance relation similar to that of Eqn. 4 applies for the harmonic polynomial.

[1] A set $\{\psi_i\}$ of vectors is said to span an invariant subspace V_s under a given set of group operations $\{g_j\}$ if $g_j\psi_i \in V_s\ \forall i, j$.

Derivation of 3D Zernike moments. The 3D Zernike functions Z_{nl}^m are defined as

$$Z_{nl}^m(\mathbf{x}) = R_{nl}(r)\cdot Y_l^m(\vartheta, \phi)$$

while restricting l so that $l \le n$ and $(n-l)$ be an even number. The above equation can be rewritten in Cartesian coordinates using the harmonic polynomials e_l^m:

$$Z_{nl}^m(\mathbf{x}) = \sum_{\nu=0}^{k} q_{kl}^\nu |\mathbf{x}|^{2\nu} e_l^m(\mathbf{x}), \tag{7}$$

where $2k = n-l$ and the coefficients q_{kl}^ν are determined to guarantee the orthonormality of the functions within the unti sphere:

$$\begin{aligned} q_{kl}^\nu &= \frac{(-1)^k}{2^{2k}}\sqrt{\frac{2l+4k+3}{3}}\binom{2k}{k}(-1)^\nu \\ &\quad \cdot \frac{\binom{k}{\nu}\binom{2(k+l+\nu)+1}{2k}}{\binom{k+l+\nu}{k}}. \end{aligned}$$

The orthonormality relation reads as follows:

$$\frac{3}{4\pi}\int_{|\mathbf{x}|\le 1} Z_{nl}^m(\mathbf{x})\cdot\overline{Z_{n'l'}^{m'}(\mathbf{x})}d\mathbf{x} = \delta_{nn'}\delta_{ll'}\delta^{mm'}$$

In case of the 3D Zernike functions the same invariance relation applies as in case of spherical harmonics. If we collect the functions into $(2l+1)$-dimensional vectors $\mathbf{Z}_{nl} = (Z_{nl}^l, Z_{nl}^{l-1}, Z_{nl}^{l-2}, \ldots, Z_{nl}^{-l})^t$ for each l, for an arbitrary rotation $\mathbf{P}$ we obtain the relation

$$\mathbf{Z}_{nl}(\mathbf{Px}) = \mathbf{o}_l(\mathbf{P})\mathbf{Z}_{nl}(\mathbf{x}). \tag{8}$$

We are now able to define the 3D Zernike moments Ω_{nl}^m of an object defined by f as

$$\Omega_{nl}^m := \frac{3}{4\pi}\int_{|\mathbf{x}|\le 1} f(\mathbf{x})\overline{Z_{nl}^m(\mathbf{x})}d\mathbf{x}.$$

It is worthwhile noting that due to the symmetry relation of Eqn. 6, a similar relation holds for the Zernike moments:

$$\Omega_{nl}^{-m}(\mathbf{x}) = (-1)^m\overline{\Omega_{nl}^m(\mathbf{x})}. \tag{9}$$

It is important to notice that the 3D Zernike moments Ω_{nl}^m are not invariant under rotations. In order to achieve invariance, we apply the approach followed in case of spherical harmonics (cf. Eqn. 4): we collect the moments into $(2l+1)$-dimensional vectors $\mathbf{\Omega}_{nl} = (\Omega_{nl}^l, \Omega_{nl}^{l-1}, \Omega_{nl}^{l-2}, \ldots, \Omega_{nl}^{-l})^t$ and define the rotationally invariant 3D Zernike descriptors F_{nl} as norms of vectors $\mathbf{\Omega}_{nl}$:

$$F_{nl} := ||\mathbf{\Omega}_{nl}||. \tag{10}$$

Reconstruction. Since the functions Z_{nl}^m form a complete orthonormal system, it is possible to approximate the original function f by a finite number of 3D Zernike moments Ω_{nl}^m:

$$\hat{f}(\mathbf{x}) = \sum_n\sum_l\sum_m \Omega_{nl}^m\cdot Z_{nl}^m(\mathbf{x}). \tag{11}$$

Here, we sum over $n \in [0, N]$, $l \in [0, n]$ such that $(n-l)$ be an even number and $m \in [-l, l]$. We use the reconstruction to verify how much of the original object information is included in a set of 3D Zernike moments up to a given order $n = N$.

4. COMPUTATION OF 3D ZERNIKE DESCRIPTORS

We now consider the computational details. First, let us expand Z_{nl}^m of Eqn. 7 using Eqn. 5:

$$\begin{aligned} Z_{nl}^m(\mathbf{x}) &= c_l^m 2^{-m} \Sigma_{v=0}^{k} q_{kl}^v \\ &\cdot \Sigma_{\alpha=0}^{v} \binom{v}{\alpha} \Sigma_{\beta=0}^{v-\alpha} \binom{v-\alpha}{\beta} \\ &\cdot \Sigma_{u=0}^{m} (-1)^{m-u} \binom{m}{u} i^u \\ &\cdot \Sigma_{\mu=0}^{\lfloor \frac{l-m}{2} \rfloor} (-1)^{\mu} 2^{-2\mu} \binom{l}{\mu} \binom{l-\mu}{m+\mu} \\ &\cdot \Sigma_{v=0}^{\mu} \binom{\mu}{v} \\ &\cdot x^{2(v+\alpha)+u} \\ &\cdot y^{2(\mu-v+\beta)+m-u} \\ &\cdot z^{2(v-\alpha-\beta-\mu)+l-m}. \end{aligned}$$

Substituting $r = 2(v+\alpha)+u$, $s = 2(\mu-v+\beta)+m-u$ and $t = v-\alpha-\beta-\mu)+l-m$ and setting

$$\begin{aligned} \chi_{nlm}^{rst} &= c_l^m 2^{-m} \Sigma_{v=0}^{k} q_{kl}^v \\ &\cdot \Sigma_{\alpha=0}^{v} \binom{v}{\alpha} \Sigma_{\beta=0}^{v-\alpha} \binom{v-\alpha}{\beta} \\ &\cdot \Sigma_{u=0}^{m} (-1)^{m-u} \binom{m}{u} i^u \\ &\cdot \Sigma_{\mu=0}^{\lfloor \frac{l-m}{2} \rfloor} (-1)^{\mu} 2^{-2\mu} \binom{l}{\mu} \binom{l-\mu}{m+\mu} \\ &\cdot \Sigma_{v=0}^{\mu} \binom{\mu}{v}, \end{aligned}$$

Z_{nl}^m can be written in a more compact form as a linear combination of monomials of order up to n

$$Z_{nl}^m(\mathbf{x}) = \sum_{r+s+t \le n} \chi_{nlm}^{rst} \cdot x^r y^s z^t \tag{12}$$

Finally, let us observe that using Eqn. 12, the 3D Zernike moments Ω_{nl}^m of an object can be written as a linear combination of geometrical moments of order up to n:

$$\Omega_{nl}^m = \frac{3}{4\pi} \sum_{r+s+t \le n} \overline{\chi_{nlm}^{rst}} M_{rst}, \tag{13}$$

where M_{rst} denotes the geometrical moment of the object scaled to fit in the unit ball:

$$M_{rst} := \int_{|\mathbf{x}| \le 1} f(\mathbf{x})\, x^r y^s z^t \, d\mathbf{x}, \tag{14}$$

where $\mathbf{x} \in \mathbb{R}^3$ denotes the vector $\mathbf{x} = (x,y,z)^t$. An important fact implied by Eqn. 13 is that in order to compute the 3D Zernike functions we only have to compute the geometrical moments instead of evaluating the complex exponential and associated Legendre function of spherical harmonics.

4.1 Algorithm

The above observations lead to the following algorithm to compute the 3D Zernike descriptors F_{nl}^m. The computations have to be conducted for all n,l,m index combinations for $n \in [0,N]$, $l \in [0,n]$ such that $(n-l)$ be an even number and $m \in [-l,l]$.

The values χ_{nlm}^{rst} for $r+s+t \le n$ have to be determined before starting the algorithm. Note that this step is independent of a particular object and may be done offline. Since for an (n,l,m) triple, there will typically be a lot of zero coefficients, we store the values of these together with the indices r,s,t indexing the corresponding geometrical moment in a list $List_{\chi}^{nlm}$.

We now give the steps needed to compute the 3D Zernike moments and descriptors:

1. **Normalization.** Compute the center of gravity of the object, transform it to the origin, and scale the object so that it will be mapped into the unit ball.

2. **Geometrical moments.** Compute all geometrical moments M_{rst} for each combination of indices, such that $r,s,t \ge 0$ and $r+s+t \le N$. Refer to the next subsection for details on this computation.

3. **3D Zernike moments.** Compute all Zernike moments Ω_{nl}^m according to Eqn. 13. Note that the summation has to be conducted only for the nonzero coefficients χ_{nlm}^{rst} stored in the list $List_{\chi}^{nlm}$. Also note that for $m \le 0$, Ω_{nl}^m may be computed using the symmetry relation of Eqn. 9.

4. **3D Zernike descriptors.** Compute all F_{nl} according to Eqn. 10.

4.2 Geometrical moments

The computation of the geometrical moments is of central importance with respect to the overall computational efficiency end numerical accuracy of our method.

A typical approach to computing the geometrical moments of an object represented by a 2D image or a 3D voxel grid is the following:

1. Fix a coordinate system with origin at a corner of the grid and axes aligned with the grid axes. Subsequently, sample all monomials of order up to N at the grid point positions.

2. Compute the geometrical moments according to Eqn. 14 but integrating over the whole voxel grid.

3. Transform the geometrical moments according to the normalization transformation of the object. This can easily be accomplished, since scaling can be achieved by scaling the moments, the moments of the translated object can be represented in terms of a linear combination of original moments of not greater order.

The first two steps introduce numerical problems. First, the sampling at grid points implies that we treat the monomial as a function having a constant value within a voxel, which is determined by the value of the monomial e.g. in the center of the voxel. For rapidly changing functions, like the monomials of high order, this results in inaccuracy. Second, for a 64^3 grid for instance, the precision of the double precision floating point number is exceeded already at the order of 9. According to our experience, moments up to order of 20 are needed to provide a good descriptor.

We treat the first issue by computing the geometrical moments in terms of monomials integrated over the voxels. Since for high orders the 3D Zernike descriptors seem to discard the values of voxels close to the origin, we normalize the object prior to computation of moments, thus obtaining considerably better numerical accuracy and providing a cure to the second problem. These procedures are described in the remainder of this section.

4.2.1 Integration.

Let us first consider the 1D case. The function f is sampled at the sample points $\{x_i\}$, $0 \le i \le N-1$. We treat f as having constant

values f_i within intervals $[x_i, x_{i+1})$:

$$\begin{aligned} M_p &= \int f(\alpha)\alpha^p \, d\alpha \\ &= \sum_{i=0}^{N-1} f_i \int_{x_i}^{x_{i+1}} \alpha^p d\alpha \\ &= \sum_{i=0}^{N-1} f_i \frac{x_{i+1}^{p+1} - x_i^{p+1}}{p+1}. \end{aligned}$$

The computation of geometrical moments of order p for $0 \le p \le P$ can be formulated in matrix form:

$$\underbrace{\begin{bmatrix} M_0 \\ M_1 \\ \vdots \\ M_P \end{bmatrix}}_{\mathbf{M}} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{3} \\ \vdots \\ \frac{1}{P+1} \end{bmatrix} \underbrace{\begin{bmatrix} x_0 & x_1 & \cdots & x_N \\ x_0^2 & x_1^2 & \cdots & x_N^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_0^P & x_1^P & \cdots & x_N^P \end{bmatrix}}_{\mathbf{X}} \cdot$$

$$\cdot \underbrace{\begin{bmatrix} -1 & & & \\ 1 & -1 & & \\ & 1 & \cdots & \\ & & \cdots & -1 \\ & & & 1 \end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{N-1} \end{bmatrix}}_{\mathbf{F}}$$

We note that $\mathbf{X}$ is a Van der Monde with dimensions $(P+1)\times(N+1)$, matrix $\mathbf{D}$ is of dimensions $(N+1)\times N$. During the computation, we first conduct the multiplication $\mathbf{DF}$ yielding a vector $\mathbf{F}'_0$ of differences:

$$\mathbf{F}'_0 = \mathbf{DF} = \begin{bmatrix} f'_{0,0} \\ \vdots \\ f'_{0,N-1} \\ f'_{0,N} \end{bmatrix} = \begin{bmatrix} -f_0 \\ f_0 - f_1 \\ \vdots \\ f_{N-2} - f_{N-1} \\ f_{N-1} \end{bmatrix}.$$

Subsequently, we generate the vectors $\mathbf{F}'_i$ by successively multiplying componentwise with the vector of samples $\mathbf{S} = [x_0, x_1, \ldots, x_N]^T$:

$$\mathbf{F}'_{n+1} = \begin{bmatrix} f'_{n+1,0} \\ \vdots \\ f'_{n+1,N-1} \\ f'_{n+1,N} \end{bmatrix} = \begin{bmatrix} x_0 f'_{n,0} \\ \vdots \\ x_{N-1} f'_{n,N-1} \\ x_N f'_{n,N} \end{bmatrix}$$

The 1D geometrical moments M_p can thus be computed by adding up the components of $\mathbf{F}'_p$ and multiplying by a factor:

$$M_p = \frac{1}{p+1} \sum_{l=0}^{N} f'_{p,l}.$$

Since the 3D geometrical moments M_{pqr} can be written as:

$$M_{pqr} = \sum_{i=0}^{N-1} \frac{x_{i+1}^{p+1} - x_i^{p+1}}{p+1} \cdot \sum_{j=0}^{N-1} \frac{y_{j+1}^{q+1} - y_i^{q+1}}{q+1} \cdot \sum_{k=0}^{N-1} \frac{z_{k+1}^{r+1} - z_i^{r+1}}{r+1} f_{ijk}.$$

As the formula already suggests, the three dimensional case can be split into 1D cases and may thus be computed in exactly same manner as above.

4.2.2 Pre-scaling

The radial polynomials of the 3D Zernike function tend to have small values near the origin. As a consequence, the projections of the object function f in the vicinity of the origin are suppressed – these values have very small impact on the final value of a 3D Zernike moment Ω_{nl}^m. On the other hand, scaling the object to fit into the unit ball means that we shift the most severe numerical inaccuracy caused by the floating point representation to the vicinity of the origin, since high order monomials have values with high negative exponent in this area.

Consequently, in order to obtain an improved numerical accuracy, as a first step we translate and scale the object according to the normalization transformation discussed above, and compute the geometrical moments afterwards. As is proven by our results in Section 6, this considerably improves the numerical accuracy of the final 3D Zernike descriptors.

5. RESULTS

In this section we describe some practical results of our approach. In our experiments, the volumetric object functions were generated by voxelizing the polygonal boundary representations of geometric models. To this end we used the vxt software library [2]. We investigated a number of voxelization methods provided by this package; in particular, voxelization with linear and gaussian decay function, i.e. the values corresponding to voxels are determined as a linear or gaussian function of the distance from the object surface. Furthermore, a binary function may be generated by simply thresholding the values of the voxel grid. As subject of experiments, we used a small set of 655 objects of general categories downloaded from www.3dcafe.com. During our investigations we used the Euclidean distance between vectors representing 3D Zernike descriptors as a measure of similarity between the objects.

In the remainder of this section we demonstrate the numerical accuracy of the algorithm and the reconstruction ability of 3D Zernike moments based representation. Furthermore, we discuss the dependency of the descriptor on the voxel grid resolution, voxelization method and number of used invariants.

5.1 Numerical accuracy

Figure 1: The objects used to generate the results of Table 1.

In order to verify the numerical accuracy of our method, we used the GNU MP arbitrary precision arithmetic library (http://www.swox.com/gmp/) to generate an accurate reference. As already mentioned in the Section 4.2, the accuracy of computations is decisively influenced by the numerical stability of geometrical moments. We have implemented three versions of our software: with pre-scaling, with integration and with both scaling and integration. The results for the set of objects depicted in the Fig. 1 are presented in Table 1. We obtained these results by computing the 3D Zernike descriptors from Zernike moments of order n up to 20 and calculating the L_2 norm of the difference between a vector of invariants yielded by the repective version and the precise values computed using the GNU MP. In order to be able to assess the values of Table 1 we note that we compute the similarity between two objects as the Euclidean distance of vectors containing their respective 3D Zernike descriptors. These distances are typically on the order of 0.1, which implies that the inaccuracy caused by incautious computation of geometrical moments render the approach unusable. Therefore, both the scaling and the integration are important components of the numerical calculations.

We note that the computations using the GNU MP library were about two orders of magnitude slower compared to those using the built-in double precision arithmetic.

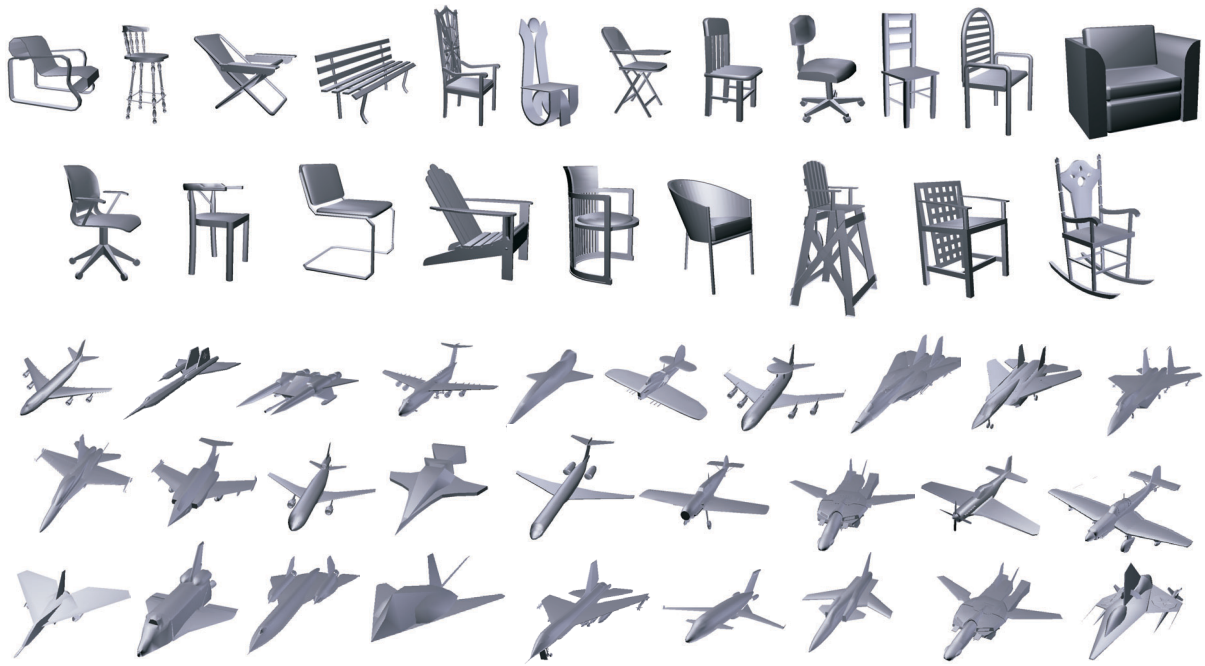

Figure 2: The classes of chairs and planes selected manually to inspect the retrieval performance of our algorithm.

	Integrated	Pre-scaled	Integrated, Pre-scaled
GUITAR	484.5	22.7	0
T_INVADER	85215.7	796.2	0
INVCHAIR	874.4	997.2	1.6e-17
747	2.6e5	56.7	1.3e-17
BALL	1.7e7	2289.7	3.7e-9

Table 1: The L_2 error in dependence of the geometrical moment calculation method.

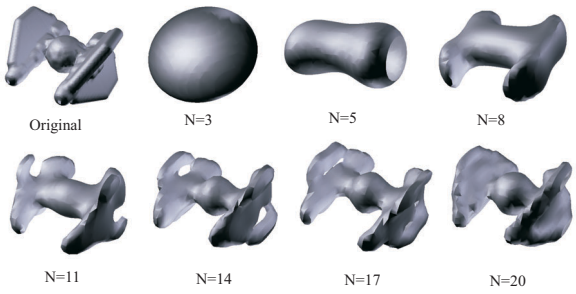

Figure 3: Reconstruction of a spaceship shown as isosurfaces of the reconstructed volumetric function. The upper left is the voxelized original object. The numbers below the images indicate the number of Zernike moments that have been used for the reconstruction.

5.2 Reconstruction

Figure 3 demonstrates the reconstruction property of the 3D Zernike moments. We reconstructed the object using Eqn. 11 on a 24^3 grid. As it can be seen, the moments of order up to 20 allow for reconstructing the main object characteristics while discarding small details.

5.3 Parameter dependency

In order to estimate the retrieval performance of the 3D Zernike descriptors, we experimented with two classes of objects manually selected from our small database: a class of 21 chairs and another one consisting of 28 planes, see Figure 2. To measure the quality of the query results we used the precision-recall diagrams commonly used in information retrieval. These diagrams may be interpreted as follows: Having a class C of objects and the top n matches delivered by the retrieval system, the recall R indicates the number of already found objects belonging to C, while the precision $P(R) = \frac{R}{n}$ denotes the ratio of this number to n. High precision values indicate good results according to this measurement method. Since we intend to investigate the retrieval quality with respect to several parameter values, a single scalar would be of more use. To achieve this, we average the precision values P_i of ith class members of C with $1 \leq i \leq |C|$ for a specific recall R_j

$$\overline{P(R_j)} = \frac{\sum_{i=1}^{|C|} P_i(R_j)}{|C|},$$

and simply sum up the precision over all recall values and normalize the sum by the count of object belonging to each class, thereby quantifying the *overall precision* P_o^C of class C

$$P_o^C = \frac{\sum_{j=1}^{|C|} \overline{P(R_j)}}{|C|}.$$

Hence, P_o^C is essentially the integral of normalized precision-recall diagrams averaged over the members of respective classes.

We analyze the influence of the following parameters:

- **Voxelization method**: we voxelized the polygonal boundary

Resolution	Voxelization	Kernel width	N	P_o^{Chairs}
32^3	Linear	2	10	0.31
		4	9	0.29
		5	8	0.02
48^3	Linear	2	14	0.41
		4	12	0.34
		6	12	0.3
64^3	Linear	2	18	0.44
		4	15	0.43
		6	15	0.38
64^3	Binary	1	20	0.45
		2	20	**0.48**
		3	21	0.4

Chairs

Resolution	Voxelization	Kernel width	N	$P_o^{Airplanes}$
32^3	Linear	2	10	0.28
		4	10	0.38
		5	10	0.39
48^3	Linear	2	14	0.31
		4	12	0.35
		6	12	0.36
64^3	Linear	2	18	0.31
		4	15	0.35
		6	15	0.37
64^3	Binary	1	23	0.32
		2	22	**0.42**
		3	23	**0.43**

Airplanes

Table 2: The average retrieval precision results for chairs and planes. Note that the highlighted best values are achieved for similar parameter values in case of both classes.

of the objects using a radial linear- and binary kernel with varying widths. We also experimented with gaussian kernels with compact support, however, in this case the results were very similar to those yielded by utilizing the linear kernel. The support width of the kernel indicates the distance in voxel units where the kernel values decrease to zero.

- **Number N of moments**: this indicates that Zernike moments Ω_{nl}^m with $0 \leq n \leq N$ were used to generate the invariants. Utilizing low values of N we discard the high frequencies of the objects, while for high values detail information is incorporated into the comparison as well.

- **Resolution of the voxel grid**: we used three resolutions: 32^3, 48^3 and 64^3. As is confirmed by our results, the coarser and finer grids correspond to less or more detailed representations of objects, respectively.

Some results concerning the retrieval accuracy are summarized in Table 2. Here, we did not include the results for *all* numbers N of highest order moments, but only for those yielding the highest overall precision. The results for each class of objects may be explained as follows:

- **Airplanes**: Kernels with large support (which essentially corresponds to thickening the boundary of the objects) yield a representation that is apparently most characteristic to this object class. The application of such kernels discards the "disturbing" details, while emphasizing the indeed characteristic structure of planes: a cylinder like bulk with wings emanating symmetrically.

- **Chairs**: Although there are some common characteristics, this class exhibits a greater manifoldness – we typically need more details, i.e. high resolution, to describe these objects. It is interesting to see in this example that in spite of the intuitional high-frequency nature of chairs (relatively thin legs, thin planes of the lean and seat), the optimal number of moments is lower for this class compared to the planes. This may be explained by the fact that addition of detail beyond a specific bound makes the objects too different. Thus, on the one hand we need a high grid resolution to preserve enough detail in the volumetric representation. On the other though, we need to discard some of these details in terms of high order coefficients to maintain similarity.

It can be seen from Table 2 that the best results are achieved in case of binary voxelization in a 64^3 grid. In what follows, we apply a kernel width of 2 and maximal moment order $N = 20$.

While creating both of the above classes, we classified the objects according to shape and function. As it can be seen in Figure 2, though, we generated classes containing objects of considerable variance in terms of shape. It should be emphasized that other classifications (e.g. including only dining room chairs, etc.) and/or other databases of objects may and typically will yield other retrieval accuracies and other optimal parameter values. Moreover, different descriptors may be more or less suitable for a particular classification. We conjecture that it is possible to classify in such a way that the precision values are considerably higher than in our experiment. Furthermore, a particular classification may be more or less appropriate for a given application. Our goal was to demonstrate the performance of the 3D Zernike descriptors as a general 3D shape characterization, which guided our choice of example object classes.

In order to test the influence of small geometric deformations on the performance of the 3D Zernike descriptor, we gradually deformed a chair using the Free Form Deformation, see Figure 4. The Euclidean distances from the original chair are indicated as ratio to the distance of the best match given by our algorithm. As it can be seen, the 3D Zernike descriptors are relatively insensitive to such small geometric deformations, since the deformed versions would still be the best matches.

Figure 4: Influence of small geometric deformation on the 3D Zernike descriptors. The leftmost object is the original chair, the other two were generated by deforming it. The values below the deformed objects indicate the ratio of their distances to the best match of the original object. Since the values are smaller then 1, these objects still lie closer to the original one then the best match yielded by the retrieval.

5.4 Comparison with spherical harmonic descriptors

The spherical harmonic descriptors were introduced very recently by Funkhouser et al. [13]. These descriptors essentially fit into

the model of rotationally invariant descriptor construction we presented in the Section 3, with the difference that the authors do not use radial polynomials to modulate the spherical harmonics, but sample the three dimensional space as concentric shells, where the shells are defined by equal radial intervals. Subsequently, they discretize the shells into equiangular bins, and define a binary spherical function defined as 1 if there is an object point in such a bin and 0 otherwise. Consequently, their object representation consists of a full spherical harmonic decomposition for each shell. The objects are voxelized into a 64^3 grid after a normalization transformation similar to that described in the previous section. The authors use 32 concentric shells to define the spherical functions and 16 rotationally invariant spherical harmonic descriptors for each shell, this gives a vector of 512 scalar values for a single object. During the search of a database, the similarity of objects is calculated as the Euclidean distance between these vectors.

Both 3D Zernike descriptors and spherical harmonic descriptors achieve rotation invariance by exploiting the invariance properties of the spherical harmonics. However, by merely sampling the space in radial direction, the latter descriptor does not capture object "frequencies" or coherence in this direction, thereby incorporating less object characteristic information. Finally, we note that using up to 20 moments, the frequency resolution of the Zernike descriptors will be similar to that of the spherical harmonic descriptors. However, using the former we have 122 scalar values for an object, which delivers a more compacte descriptor. In consequence, the dimensionality of the search problem will be reduced, which supports the effectiveness of the search process, and there is less storage overhead for an object.

We present some experimental search results in Fig. 5. The objects in the database have been voxelized and for each object a 3D Zernike descriptor was computed based on moments of order up to 20 which resulted in 122 scalar values for each object. The computation of such descriptors for a grid of 64^3 lasted 0.3 seconds each on a 1.8 GHz Pentium4. As for spherical harmonic descriptors, we used the setting suggested in [13], i.e. 32 concentric spheres and 16 coefficients for each sphere. The query object is the first one in each row, the left to right ordering reflects the similarity ordering yielded by the retrieval process. Beneath each result given by our approach, the result delivered using spherical harmonic descriptors are presented for reference. Although there are differences, the descriptors deliver similar retrieval performance. To compare the performance of both methods we present average precision-recall diagrams for pre-classified sets of chairs and planes, see Figure 6. The results are very similar for the class of planes, whereas for chairs the 3D Zernike descriptors apparently perform considerably better. We once again emphasize though, that we use optimized parameters as described in the previous subsection. We did not experiment with the spherical harmonic descriptors in such terms, we merely used the setting suggested by the authors.

Similarly to the spherical harmonic descriptors, the representation as 3D Zernike descriptors is insensitive to geometric and topological artifacts common to freely available objects. This implies that despite of its compactness, the 3D Zernike descriptors can successfully compete with the retrieval performance of spherical harmonic descriptors.

We note that in case of a large database of 3D objects, the underlying frequency metaphor may be used to accelerate the search process. We recall that the spherical harmonics form essentially a Fourier basis on the sphere and the radial polynomial may also be interpreted analogously in terms of their order. This allows us to generate a natural hierarchy of representations and enables the utilization of an efficient hierarchical search algorithm.

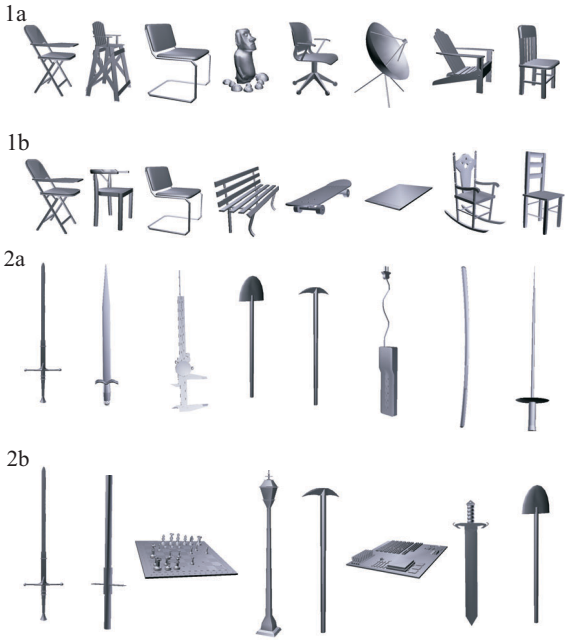

Figure 5: Excerpt of the search results. The leftmost object is the query in each row, the ordering from left to right reflects the similarity ordering. Beneath the results provided by our method (denoted by 'a') the results yielded by spherical harmonic descriptors (denoted by 'b') are shown for reference.

6. CONCLUSIONS AND FUTURE WORK

In this paper we utilized the 3D Zernike descriptor for the purpose of content based retrieval of 3D objects. We discussed some general rules for the construction of affine invariant object descriptors and derived the 3D Zernike descriptors within this framework. We furthermore considered the implementational issues: the severe instability of geometrical moments and hence the 3D Zernike descriptors in case of high orders. As a cure to this problem, we applied analytical integration within each voxel and scaled the object prior to computations, thereby achieving high accuracy even for high orders of Zernike moments. The quality of the descriptor regarding the retrieval performance was analyzed and verified also with respect to another related recent technique. As it turns out, the 3D Zernike descriptors compare favorably to the best descriptors for general 3D objects in terms of retrieval performance and robustness against topological and geometrical artifacts plaguing a most of freely available models.

As for short term future work we plan to investigate the usage of further radial functions: a wavelet based function seems to be promising, as such basis would allow for a multi-resolution radial localization of frequenciese. Moreover, we intend to elaborate a new distance function between the descriptors, as we suspect that different coefficients contribute to different a extent to the overall shape information.

7. ACKNOWLEDGMENTS

We thank Roland Wahl and Patrick Degener for the insightful discussions. This work was partially funded by the German Research Foundation (DFG) within the initiative V^3D^2 "Distributed Processing and Delivery of Digital Documents".

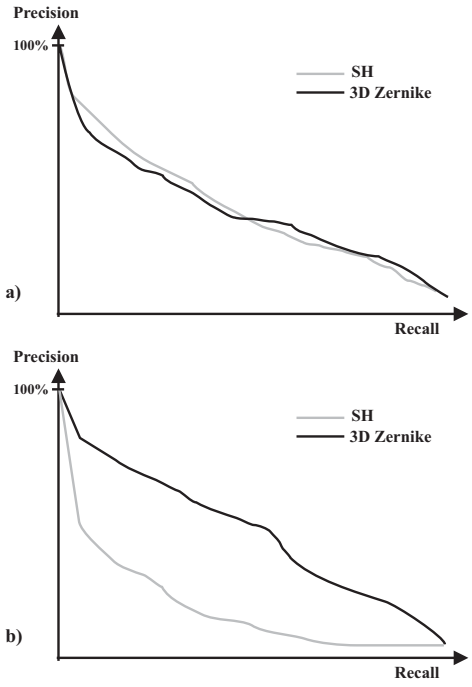

Figure 6: Precision-recall diagrams corresponding to the a) class of planes and b) class of chairs.

8. REFERENCES

[1] 3D Knowledge, `http://3dk.asu.edu`.

[2] Milos Sramek's volume visualization page, `http://www.viskom.oeaw.ac.at/~milos/`.

[3] National design repository, `http://edge.mcs.drexel.edu/repository/frameset.html`.

[4] M. Ankerst, G. Kastenmuller, H.-P. Kriegel, and T. Seidl. 3D shape histograms for similarity search and classification in spatial databases. In *Symposium on Large Spatial Databases*, pages 207–226, 1999.

[5] I. Biederman. Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94:115–147, 1987.

[6] H. Blum. Biological shape and visual science. *Journal of Theoretical Biology*, 38:205–287, 1973.

[7] N. Canterakis. 3D Zernike moments and zernike affine invariants for 3D image analysis and recognition. In *11th Scandinavian Conf. on Image Analysis*, 1999.

[8] V. Cicirello and W. C. Regli. Machining feature-based comparisons of mechanical parts. pages 176–185. Int'l Conf. on Shape Modeling and Applications, May 2001.

[9] J. Corney, H. Rea, D. Clark, J. Pritchard, M. Breaks, and R. MacLeod. Coarse filters for shape matching. *IEEE Computer Graphics and Applications*, 22(3):65–74, 2002.

[10] H. Dym and H. McKean. *Fourier series and integrals*. Academic Press, 1972.

[11] M. Elad, A. Tal, , and S. Ar. Content based retrieval of VRML objects - an iterative and interactive approach. In *Eurographics Multimedia Workshop*, pages 97–108, 2001.

[12] D. Forsyth, J. Malik, M. Fleck, and J. Ponce. Primitives, perceptual organization and object recognition. Technical report, Computer Science Division, University of California at Berkeley, Berkeley, CA 94720, 1997.

[13] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs. A search engine for 3D models. *ACM Transactions on Graphics*, 22(1), 2003.

[14] A. Gupta and R. Jain. Visual information retrieval. *Communications of the ACM*, 40(5):70–79, 1997.

[15] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L. Kunii. Topology matching for fully automatic similarity estimation of 3D shapes. In *Proceedings of ACM SIGGRAPH*, 2001.

[16] M. K. Hu. Visual pattern recognition by moment invariants. *IRE Trans. Information Theory*, 8(2):179–187, 1962.

[17] C. E. Jacobs, A. Finkelstein, and D. H. Salesin. Fast multiresolution image querying. In *Proceedings of SIGGRAPH '95*, pages 277–286, 1995.

[18] M. Kazhdan, B. Chazelle, D. Dobkin, T. Funkhouser, and S. Rusinkiewicz. A reflective symmetry descriptor for 3D models, 2003. To appear in Algorithmica.

[19] A. Khotanzad and Y. Hong. Invariant image recognition by Zernike moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5), 1990.

[20] F. Leymarie and B. Kimia. The shock scaffold for representing 3d shape. In *Proc. of 4th International Workshop on Visual Form (IWVF4)*, 2001.

[21] D. McWherter, M. Peabody, W. C. Regli, and A. Shokoufandeh. Transformation invariant shape similarity comparison of solid models. ASME Design Engineering Technical Confs., 6th Design for Manufacturing Conf. (DETC 2001/DFM-21191), Sep 2001.

[22] P. Min, J. A. Halderman, M. Kazhdan, and T. A. Funkhouser. Early experiences with a 3D model search engine. In *Proc. Web3D Symposium*, 2003.

[23] R. Ohbuchi, T. Otagiri, M. Ibato, and T. Takei. Shape-similarity search of three-dimensional models using parameterized statistics. In *Pacific Graphics*, 2002.

[24] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin. Matching 3D models with shape distributions. In *International Conference on Shape Modeling and Applications*, 2001.

[25] E. Paquet and M. Rioux. A content-based search engine for VRML databases. In *CVPR Proceedings*, pages 541–546, 1998.

[26] E. Petrakis. Design and evaluation of spatial similarity approaches for image retrieval. *Image and Vision Comp.*, 20(1):59–76, 2002.

[27] Y. Rui, T. S. Huang, and S.-F. Chang. Image retrieval: Past, present, and future. In *International Symposium on Multimedia Information Processing*, 1997.

[28] M. Schneier and M. Abdel-Mottaleb. Exploiting the jpeg compression scheme for image retrieval. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(8):849–853, 1996.

[29] K. Siddiqi, A. Shokoufandeh, S. J. Dickinson, and S. W. Zucker. Shock graphs and shape matching. In *ICCV*, pages 222–229, 1998.

[30] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.

[31] M. T. Suzuki, T. Kato, and N. Otsu. A similarity retrieval of 3D polygonal models using rotation invariant shape descriptors. In *IEEE International Conference on Systems, Man, and Cybernetics (SMC2000)*, pages 2946–2952, 2000.

[32] M. Swain and D. Ballard. Color indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991.

[33] J. W. Tangelder and R. C. Veltkamp. Polyhedral model retrieval using weighted point sets. *International Journal of Image and Graphics*, 3(1):1–21, 2003.

[34] M. Teague. Image analysis via the general theory of moments. *Journal Optical Society of America*, 70(8):920–930, 1980.

[35] C.-H. Teh and R. T. Chin. On image analysis by the methods of moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4):496–513, 1988.

[36] D. V. Vranic and D. Saupe. Description of 3D-shape using a complex function on the sphere. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME 2002)*, pages 177–180, 2002.

[37] C. T. Zahn and R. Z. Roskies. Fourier descriptors for plane closed curves. *IEEE Transactions on Computers*, 21:269–281, 1972.