Eurographics Workshop on 3D Object Retrieval (2017)
I. Pratikakis, F. Dupont, and M. Ovsjanikov (Editors)

Directed Curvature Histograms for Robotic Grasping

Rodrigo Schulz, Pablo Guerrero, and Benjamin Bustos

Department of Computer Science, University of Chile, Chile

Abstract

Three-dimensional descriptors are a common tool nowadays, used in a wide range of tasks. Most of the descriptors that have been proposed in the literature focus on tasks such as object recognition and identification. This paper proposes a novel three-dimensional local descriptor, structured as a set of histograms of the curvature observed on the surface of the object in different directions. This descriptor is designed with a focus on the resolution of the robotic grasping problem, especially on the determination of the orientation required to grasp an object. We validate our proposal following a data-driven approach using grasping information and examples generated using the Gazebo simulator and a simulated PR2 robot. Experimental results show that the proposed descriptor is well suited for the grasping problem, exceeding the performance observed with recent descriptors.

Categories and Subject Descriptors (according to ACM CCS): I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Curve, surface, solid, and object representations I.2.9 [Artificial Intelligence]: Robotics—Manipulators I.4.7 [Image Processing and Computer Vision]: Feature Measurement—Size and shapes

1. Introduction

The use of three-dimensional (3D) sensory data has become very popular and almost ubiquitous in recent years. This tendency can be partially explained by the availability of low-cost 3D sensors like *Microsoft Kinect* [Zha12], released in 2012, which made 3D information accessible and easy to generate. Also, the introduction of Point Cloud Library (PCL) [RC11], and similar libraries for 3D data manipulation and analysis have boosted the use of 3D data.

This tendency is reflected in the research world, where the development of 3D descriptors is now an active field. Also, in Robotics and Computer Vision many new algorithms and applications have been developed using 3D data. Many problems in these areas may find simpler or easier solutions using 3D sensory information, instead of plain 2D sensors such as cameras.

Robotic grasping is a discipline in the field of Robotics, which aims to develop algorithms for the generation of strategies for a robot to grasp an object. These strategies will depend on the features of the robot and the form of the object, and usually are strongly influenced by the visual features from which they are based on.

We believe that the local curvature of an object is one of the most significant features which allows us to determine feasible grasping strategies. This belief stems from the observation that humans usually grasp an object following its smaller dimension [FBD14], which can be also viewed as aligning the hand according to the direction of higher curvature in the surface of the object. This can be explained because such grasping strategy allows a person to close the hand as much as possible, which yields a strong and reliable grasp.

In this work, we propose a 3D local descriptor, designed to be used in robotic grasping tasks. The proposed descriptor is focused on encoding how the curvature changes around the target point, by analyzing the surface of the object in different directions.

We evaluate our proposal using the Gazebo simulator [Zha12] and a simulated PR2 robot, to test the applicability to the grasping problem, using a data-driven resolution approach. We show that our proposal is robust in relation to the orientation of the robot grip, allowing to define zones of feasible orientations for it.

This work is structured as follows: Section 2 defines the robotic grasping problem and makes a brief review of the current approaches to it. Section 3 reviews the related works, focusing on 3D local descriptors. Section 4 presents the details of the proposed descriptor. Section 5, discusses the the applicability of the proposed descriptor to real-life clouds and presents a method to deal with noisy sensory data. Section 6 describes the experimental setup used for evaluation. Section 7 presents and discusses the experimental results obtained. Section 8 presents the conclusions of this work.

2. The Robotic Grasping Problem

Robotic grasping refers to the problem commonly described as *determining the grasp required to carry out certain manipulation*

DOI: 10.2312/3dor.20171060

tasks on an object [LMSB14]. The grasp is typically performed using a robotic arm along with a robotic hand or gripper, addressed as *end-effector* or, simply, *effector*. In general, the main goal here is to find a parameterization of the effector trajectory or final 3D pose, satisfying the criteria defined for the current task, which together with a suitable point or zone located on the surface of the object, would allow for successfully grasp it.

2.1. Current Approaches

In the literature, there are different approaches to the grasping problem. Sahbani et al. [SEKB12] divided these methodologies into two groups: i) analytic methods; and ii) empirical methods.

Analytic methods are characterized by the use of multi-fingered effectors (robotic hands) and focus on modeling the physics and dynamics of the effector-object interaction. According to Shimoga [Shi96], this type of methods generate force-closure grasps, which are stable in equilibrium and exhibit a certain dynamic behavior. In a review, Bohg et al. [BMAK13] point out that analytic methods are commonly formulated as an optimization problem, constrained to the criteria of the grasping task and the dynamics of the effector. This was the most popular approach until the past decade.

In contrast to analytic methods, which construct a grasp according to the physics of the problem, empirical or data-driven methods rely in the use of sensory information for sampling grasp candidates from the solution space and ranking them according to a particular metric. This process is usually based on previous grasping experience, grasping prototypes or examples provided by human experience. For this reason, it has also been dubbed *knowledge based* [Shi96]. This approach typically uses simpler effectors, for example two-fingered grippers, and has been the most popular strategy in recent years.

Bogh et al. [BMAK13] divided data-driven methods into 3 groups according to the assumptions they make about the target object: i) grasp of a known object; ii) grasp of a familiar object; and iii) grasp of an unknown object.

Grasp-of-known-objects methods assume that the target object has been seen before, and in consequence grasps are already generated for the object. These methods focus on recognition and identification of the object and estimation of its pose, so the previously generated grasp can be reused. **Grasp-of-familiar-objects methods** assume that the target object share some similarities with some of the already seen objects. This approach assumes that similar objects (or zones) can be grasped in similar ways, taking advantage of previous grasping experience. **Grasp-of-unknown-objects methods** do not assume any kind of grasping experience. They analyze the geometry of the object to identify suitable zones to perform a grasp.

3. Related Work

3D sensors and sensory data are increasingly present nowadays, making the development of descriptors for 3D data an active research field. 3D descriptors can be classified in two types: i) global; and ii) local. Global descriptors [BKS*05] aim to encode the information present in the whole set of sensory data, which makes them better suited for tasks like object recognition. Local descriptors [BBGO11], on the other hand, encode the information present in contiguous portions of the sensory data, making them better for detection of relevant elements or recognition of zones of interest in the set of sensory data. 3D local descriptors are of particular interest for the grasping task, since they provide information to identify suitable grasping points or zones.

One of the older approaches is *Spin Images*, proposed by Johnson [Joh97]. This method first represents the data using cylindrical coordinates (ρ, θ, h) centered at the target point, which are then projected of the plane $\rho - h$ to generate a 2D version. The projected data is binned and transformed into a grayscale image, where the darker areas mark higher concentrations of points.

A common strategy for description is to use the spatial distribution of data, represented through different histograms. Using this methodology, Frome et al. [FHK*04] proposed *3D Shape Context* (3DSC), which defines a sphere around the target point, orienting its north pole according to the surface normal at the target point. The sphere is then divided along azimuth, elevation and radius, generating multiple spatial bins. Since the rotation of the sphere (along the azimuth) is not fixed, the spatial binning is performed using different rotations to provide some rotation invariance. Then, a vector with a weighted count of the points in each bin is generated for each rotation. Finally, the descriptor is formed by concatenating the weighted count vector of each rotation.

Following a similar strategy, Tombari et al. [TSDS10] proposed *Signature of Histograms of Orientations* (SHOT). This descriptor defines a sphere around the target point, and then a local reference frame which is used to fix the sphere rotation and provides it with rotation invariance. The sphere is split in 8 parts along the azimuth, 2 along the elevation and 2 along the radius, generating 32 regions. For each region, a histogram of the angle between the target point and each point in the region is calculated. The descriptor is formed by the concatenation of the histograms of each region.

A different approach was followed by Rusu et al. [RMBB08] when proposed *Point Feature Histograms* (PFH) in an attempt to encode information which describes the geometry in the vicinity of the target point. This descriptor computes a fixed frame for each pair of points in the vicinity. Then, the difference between the surface normal at each point is described using 3 angles plus the distance between the points, producing 4 description variables per pair of points. Finally, a histogram is generated for each variable and then concatenated to form a single vector. Since PFH is computationally expensive, Rusu et al. [RBB09] also proposed *Fast Point Feature Histograms* (FPFH), a lighter version of PFH which performs a reduced number of comparisons between points in the vicinity.

Guo et al. [GSB*13] proposed a different strategy in *Rotational Projection Statistics* (RoPS). This descriptor first generates a mesh from the 3D sensory data and extracts the surface in a neighborhood of the target point, over which a local reference frame is defined. The extracted surface is rotated around the x axis and then projected over the planes xy, xz and yz, computing different statistics for each projection. These statistics are concatenated to form the descriptor.

All the descriptors mentioned capture different features from the observed object. However, they all fail to properly retrieve enough information useful for the grasping problem in particular. *Spin Images* provides a rough representation of how an object changes around the target point, but the nature of this descriptor makes it difficult to identify and compare different zones according to their curvature.

The approach expressed by 3DSC and SHOT proves useful for object recognition, shape retrieval and 3D feature matching. Despite not having been developed for the grasping problem, spatial description of data is usually well suited for the identification of features related to the shape of the object. This approach typically allows for the identification of different curvature zones, based in the spatial distribution of the points in the surface of the object. An important drawback of 3DSC is the high dimensionality of the final descriptor, which strongly affects the performance of algorithms and tasks.

PFH and FPFH are specially designed to take into account the curvature of the object, and both descriptors attempt to encode as much information as possible about the surface. The main disadvantage of this descriptor is that the computation is complex and time expensive, which is especially problematic in the context of robotics.

Finally, the main drawback in RoPS is the inevitable construction of a mesh, which considerably affects the computation time of the descriptor and might be prohibitive in the context of robotic grasping.

4. Directed Curvature Histograms

Our hypothesis is that the local curvature of an object is a highly significative feature for the grasping task, and consequently it can be used for the generation of grasping strategies. We propose a novel 3D local descriptor, that focuses on encoding the curvature of the surface of the underlying object in different directions.

4.1. Proposed Approach

Our proposed descriptor aims to encode the local curvature, based on the evolution of the normal vector along N radial bands of data, arranged symmetrically around the target point (Figure 1).

For this work, we will assume that sensory data is stored as a cloud of 3D points or *point-cloud*. Nevertheless, all the definitions can be applied to different representations, such as meshes. Additionally, we will represent curvature change between two points as the angle formed by the surface normal on both points. When the descriptor is computed, this change will be the angle between the target point and the points in each band, and in consequence this will describe the curvature changes around the target point in different directions. This representation has three advantages over the direct use of the curvature of the surface:

i) The angle between two vectors is usually easier and more stable to calculate than the curvature of the surface.
ii) Angles are limited in range ($\alpha \in [-\pi, \pi]$).
iii) Special cases, such as a flat surface, are easy to represent and identify using angles.

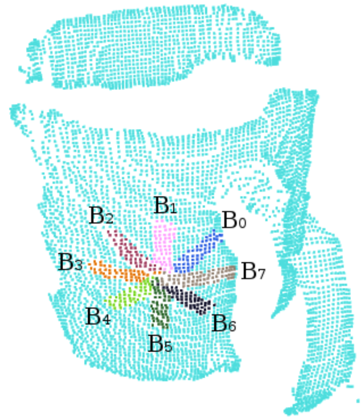


Figure 1: *3D data representing an object with a set of 8 bands highlighted on its surface. Each band provides information about the curvature along a different direction.*

4.2. Calculation Process

The calculation process is divided into two consecutive steps: i) bands extraction; and ii) descriptor calculation.

4.2.1. Bands Extraction

The goal of this step is to extract the set of 3D points that belong to each of the bands going along arbitrary directions around the target point $\vec{p}_*$ (Figure 1). For this purpose, we define a sphere of radius r centered at the target point, which generates the neighborhood set Φ. The radius of the sphere will determine the extent of the encoded curvature information.

To determine which points are part of each band, a plane, $f(\vec{x})$, tangent to the surface at the target point is defined using the Equation 1, where $\vec{p}_*$ is the target point and $\vec{n}_*$ its surface normal.

$$f(\vec{x}): \quad \vec{n}_* \cdot (\vec{x} - \vec{p}_*) = 0 \tag{1}$$

A local 2D coordinate system $\tilde{O}$ is defined inside the plane, with its origin at $\vec{p}_*$ and an arbitrary orientation (this will be later useful for the grasping task). Every point $\vec{p}_i \in \Phi$ is then projected on $f(\vec{x})$ and translated to the local coordinate system $\tilde{O}$. This yields an *image* set of points $\vec{q}_i \in \Phi_{IMG}$, completely contained in $f(\vec{x})$.

To define the orientation of each band, a set D of N unitary vectors is generated. Each $\vec{d}_k \in D$ is defined inside $f(\vec{x})$, starting at $\tilde{O}$. The set of N vectors is arranged symmetrically around $\tilde{O}$, sorted counter clockwise, starting with the first vector, $\vec{d}_1$, parallel to the axis $\tilde{O}\tilde{x}$. Consequently, every $\vec{d}_k$ will be separated from the following vector $\vec{d}_{k+1}$ by an angle $\gamma = 2\pi/N$ (Figure 2).

The orientation of each band is defined by the Equation 2.

$$l_k(t) = \vec{p}_* + t \cdot \vec{d}_k \tag{2}$$

The perpendicular distance between each point $\vec{q}_i \in \Phi_{IMG}$ to each line $l_k(t)$, $k \in \{1, \ldots, N\}$ is measured. The points with a distance smaller than a threshold w are associated to the respective band. It

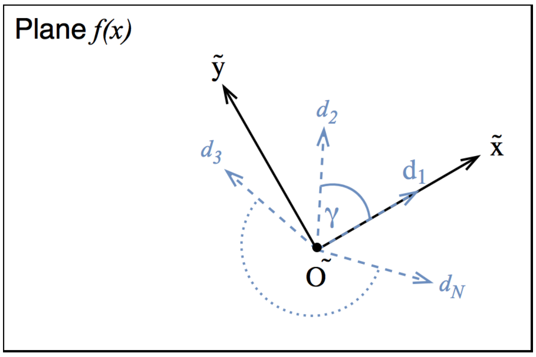


Figure 2: *The set of vectors $\vec{d}_k \in D$ are arranged symmetrically around $\tilde{O}$. Each $\vec{d}_k$ acts as the director vector for a band.*

is important to notice that this definition allows for the presence of a point in more than one band. This is usually the case of points that are close to the target point.

Finally, each band is defined as a subset of points $B_k \subseteq \Phi$ such that the image point $\vec{q}_i$ of each $\vec{p}_i$ is at a distance equal or less than w.

$$B_k = \{\vec{p}_i \in \Phi \mid \vec{q}_i \in \Phi_{IMG} \wedge dist(\vec{q}_i,\, l_k(t)) \leq w\} \tag{3}$$

Since the set of bands $\{B_1, B_2, \ldots, B_N\}$ was extracted from Φ, then the length of each band will be at most r and its width $2 \cdot w$.

4.2.2. Descriptor Calculation

The goal of this step is to encode how the surface normal changes along each band. For this purpose, we propose the generation of a histogram for each band to describe the change in orientation of the surface normal along it.

In detail, for each band B_k the angle between the surface normal at the target point, $\vec{n}_*$, and the surface normal at each point present in the band, $\vec{n}_i \mid \vec{p}_i \in B_k$, is computed as shown in Equation 4.

$$\alpha_i = arctan\left(\frac{\|\vec{n}_* \times \vec{n}_i\|}{\vec{n}_* \cdot \vec{n}_i}\right)^{-1} \tag{4}$$

After evaluating Equation 4 for each point, a histogram H_k is generated for each band B_k.

A histogram is used as a means to encode the evolution of the orientation in the normal vector, while avoiding the effect of noise and errors in the estimation of the normals. Using one histogram per band allows the descriptor to encode information individualizing the evolution of the surface normal in different directions. At the same time, the use of several bands, symmetrically arranged, provides an overview of the surrounding surface. For this reason, we called this approach *Directed Curvature Histograms* (DCH).

Finally, the descriptor is formed by the concatenation of the histograms associated to each band.

$$DCH = \begin{bmatrix} h_{k1} \\ h_{k2} \\ \vdots \\ \vdots \\ h_{N1} \\ h_{N2} \\ \vdots \end{bmatrix} = \begin{bmatrix} v_{11} \\ v_{12} \\ \vdots \\ \vdots \\ v_{N1} \\ v_{N2} \\ \vdots \end{bmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} band\ 1 \\ \vdots \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} band\ N \end{matrix} \tag{5}$$

Note that the resolution used for the histograms (size of the bins) will affect the size of the final descriptor. Smaller bins are more descriptive and provide a more fine-grained representation of the curvature of the object, but it also generates longer histograms, which will translate into a final descriptor of higher dimensionality.

4.3. Coordinate System Orientation

As mentioned in Section 4.2.1, the coordinate system defined for the computation of DCH is arbitrarily oriented. Such feature is an advantage that can be exploited according to the specific problem being approached. The formulation of the proposed descriptor allows users to define the most convenient reference frame for the application they are working on, ranging from an ad-hoc orientation to a general reference frame, thus making it applicable to diverse problems.

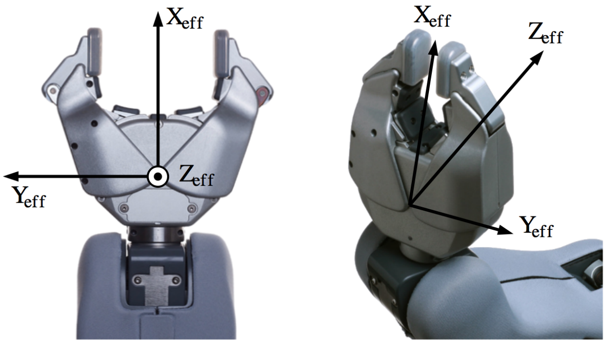
The only requirement for a general formulation, which can be used across different instances of the descriptor and different experiments, is that the reference frame has to be repeatable and consistent across instances of the descriptor. Such formulation can be achieved by simply using a fixed orientation, defined in a global coordinate system. Different approaches can be used to select orientation of the proposed descriptor. For example, this can also be done by using the definition of a repeatable local reference frame, like the one proposed by Tombari et al. [TSDS10]. In this work a repeatable local reference frame proposes is proposed, defined by the principal directions computed through a Eigen Value Decomposition (EVD) or a Singular Value Decomposition (SVD) of the covariance matrix of the point coordinates, plus an heuristic procedure for the disambiguation of the sign of the reference frame vectors.

4.3.1. Coordinate System for Robotic Grasping

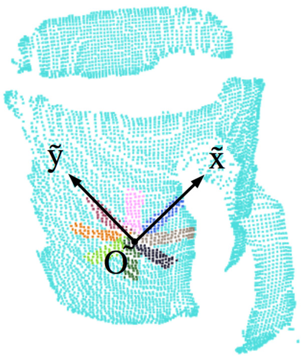
In the particular case of the robotic grasping problem, DCH can used to solve it by following the approach of data-driven methods (Section 2).

Since the orientation of the bands in the descriptor is defined arbitrarily, we can select it to match the orientation of the coordinate system of the effector, defined as shown in Figure 3a. According to this diagram, for matching the orientations of the effector and the descriptor, the angle of the axis $\tilde{x}$ has to be aligned with the vertical axis of the coordinate system of the effector, i.e., axis Z_{eff}. When the orientation of the coordinate systems is matched, the orientation of the band B_0 will be the same as the orientation of the effector.

This formulation allows the descriptor to encode changes in the



(a) *Coordinate system placed on the effector.*



(b) *Coordinate system of the descriptor, used for the definition of the orientation of the bands.*

Figure 3: *Coordinate system with origin at the effector and coordinate system used for the generation of the bands. The proposed descriptor can be used to solve the grasping problem by aligning axes Z_{eff} and $\tilde{x}$.*

curvature of the surface, as observed from the coordinate system of the effector, i.e., the local shape of the object from the perspective of the effector. Then, DCH can be used to perform a supervised learning process, that focuses on extracting a characterization of the local shape observed when a successful/unsuccessful grasp is performed. This process attempts to mock the way humans decide a grasp zone for an object, and how grasping experience of similar objects is applied to new objects. According to the classification presented in Section 2.1, this approach can be identified as *grasping of familiar objects*.

5. Dealing With Real Sensors

In this work, we used the simulated environment provided by Gazebo to perform the evaluation of the ability of the proposed descriptor to predict the outcome of a grasping candidate. Despite Gazebo is a realistic simulator, which runs the actual software of the simulated robots, necessarily the sensory data has to be synthetically produced to complete the simulation process. This might hide sensitivity problems of the proposed descriptor to noise or perturbations. This issue is specially delicate for the surface normals, whose computation is particularly sensitive to noise, and this could affect the computation of DCH when using data from real sensors.

In general, methods for the estimation of the normals of a surface involve fitting a plane tangent to the surface of the object, and then disambiguating the orientation using an adequate algorithm or criterion. For the development of this descriptor, we use the method proposed by Rusu [Rus10], based on fitting a tangent plane through a least-squares (LS) estimation and then disambiguating the orientation using Principal Component Analysis (PCA). To provide robustness against noise, we added a filtering step where a Gaussian smoothing was performed, to reduce the effect of noise in the position of the points from the point-cloud.

Figure 4 shows the effect of the Gaussian filtering step on the estimation of the normals in a point-cloud that was obtained using a real sensor. Figures 4a, 4c and 4e show the point-clouds, while Figures 4b, 4d and 4f show the computed normals for each point-cloud. These images show that the effect produced by the Gaussian filtering step has a positive effect on the computation of the normals, when dealing with real sensory data, and make the normals behave more similar to the normals computed on a synthetic point-cloud (Figure 4f).

6. Experimental Evaluation

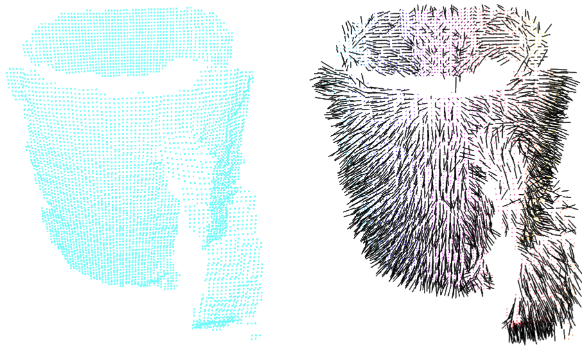
To evaluate the applicability of DCH to the grasping problem, we used the Gazebo simulator [KH04] and a simulated instance of a PR2 robot. Both were used in order to generate experimental data which could be used to perform the supervised learning process, as well as to evaluate the capabilities of the descriptor.

6.1. Experimental Data Generation

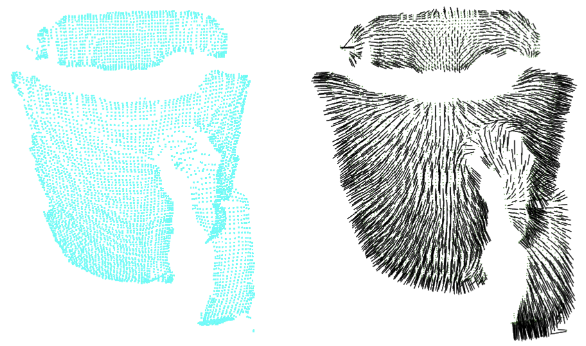
Using a simulated experimental environment, we performed several grasps attempts on three different objects, each one arranged in different positions (shown in Figure 5):

- A can, in 2 different positions.
- A cordless drill, in 3 different positions.
- A hammer, in 2 different positions.

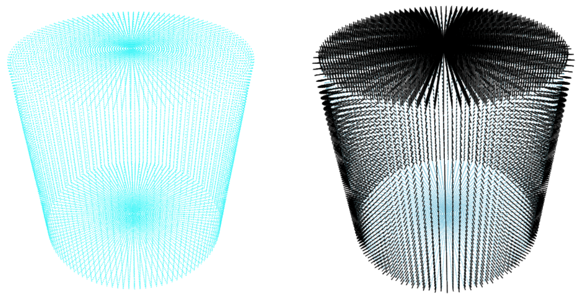
To determine the points where a grasp will be attempted, dubbed *grasping points*, we sampled the surface of each object, using a method based on unsupervised learning, assuring that each meaningful surface is represented by the set of produced samples. Such points, plus an orientation for the effector, produced a grasping candidate, which was used to perform a grasping attempt. Each experiment was performed using 4 and 5 different angles for the orientation of the effector (Figure 6a and Figure 6b). Finally, we recorded the result of each grasping attempt, identifying 3 outcomes: i) grasp successful, the object was grasped; ii) grasp failed, the object was not grasped; and iii) grasp unfeasible, the grasp could not be performed. Notice that the later outcome is produced because data-driven methods do not take into consideration physical of mechanical limitations of the robots. Therefore, it is possible to generate grasp candidates which are not feasible. It is also important to notice that this step is equivalent to sample the space of grasp candidates, to the goal of identifying which are good candidates and which are not. Therefore, since the space of possible grasps



(a) *Raw point-cloud.* **(b)** *Normals computed with the raw cloud.*



(c) *Filtered point-cloud* **(d)** *Normals computed with the filtered cloud.*



(e) *Synthetic point-cloud* **(f)** *Normals computed with the synthetic point-cloud.*

Figure 4: *Comparison of the computed normal vectors.*

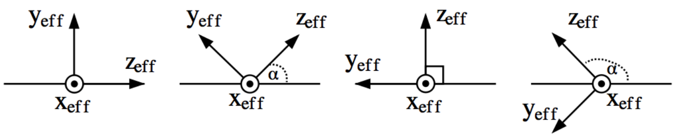
mostly contains configurations which are not capable of grasping the object, then many of the grasp attempts would produce failed attempts.

The results of the production of experimental data are presented in Table 1 and Table 2.

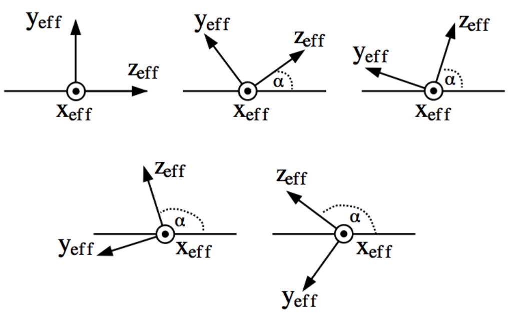
6.2. The Experiment

To evaluate the suitability of DCH for the grasping task, we performed a supervised learning process. For this process, we used only the results from the successful and failed grasping attempts.

Figure 5: *Set of objects used to produce the experimental grasping data. The information collected (grasp performed and its result) was used to evaluate the applicability of DCH to the grasping problem.*



(a) *Set of 4 orientations used for the generation of the data for the experiments (0°, 45°, 90°and 135°).*



(b) *Set of 5 orientations used for the generation of the data for the experiments (0°, 36°, 72°, 108°and 144°).*

Figure 6: *Orientations for the effector which were used to generate experimental data. Each one is represented using the reference frame of the effector, as shown in Figure 3a.*

Table 1: *Summary of the generated data.*

Angles	Success	Failed	Unfeasible	Total
4	33	29	34	96
5	31	45	59	135
Total	64	74	93	231

Table 2: *Details of the generated data.*

	Angles	Success	Failed	Unfeasible	Total
Can	4	5	14	1	20
	5	4	15	6	25
Drill	4	8	5	31	44
	5	9	13	48	70
Hammer	4	20	10	2	40
	5	18	17	5	35
Total	-	64	74	93	40

The results coming from unfeasible grasps attempts were discarded.

We trained a Support Vector Machines (SVM) [SC08] classifier using the descriptor as predictor variable and the result of the grasping attempt (successful or failed) as the predicted variable. The training was done using the data generated for two of the three objects available (a can and a cordless drill), and following a 10-fold approach for cross validation. Then, the performance of the classifier was evaluated using the data grasping data generated for the remaining third object (a hammer).

Note that since the descriptor was computed with an orientation matching the orientation of the effector, then the value of the descriptor itself will encode the difference between grasp attempts of the same point, but with different orientations for the effector.

Finally, since grasping data was generated using 4 and 5 different angles for the orientation of the effector, we performed 4 experiments combining this information:

1. Classifier trained and evaluated using data with 4 orientations.
2. Classifier trained and evaluated using data with 5 orientations.
3. Classifier trained using data with 4 orientations and evaluated using data with 5 orientations.
4. Classifier trained using data with 5 orientations and evaluated using data with 4 orientations.

The experiments were performed using DCH and three other descriptors, all of which have been tested and validated in tasks such as object recognition and/or shape retrieval: i) SHOT, ii) Spin Images, and iii) FPFH.

This evaluation allowed us to compare the performance of DCH against a set of existing descriptors, and also to evaluate its capability to characterize a surface suitable for grasping.

7. Results

The results of the performance evaluation done on the trained classifiers are presented in the Figure 7. These results show that the proposed descriptor is well suited to solve the grasping problem, using a data-driven resolution approach. In particular, for all the experiments carried out, the ROC curves presented in Figure 7 show that the classifiers trained using DCH exhibit a significantly better performance than those trained using the remaining descriptors. Since the training and evaluations procedures were carried out using a completely different set of objects, these the experiments reveal that the descriptor successfully extracts information which can be used to train a relation between the curvature observed from the reference frame of the effector, for successful grasping attempts.

The classifiers evaluated using datasets with the same number of orientations for the effector than the dataset used for training (experiments 1 and 2), performed better than the ones evaluated with a dataset with a different number of orientations (experiments 3 and 4). This was the expected result. However, the later experiments still displayed a good performance. This result shows that the trained classifier is capable of correctly predicting the outcome of a grasp candidate, even when a new angle for the effector is used.

The results obtained with the classifiers trained using SHOT, Spin Images and FPFH were notoriously worse than the ones obtained using DCH. This can be explained by the fact that all these descriptors are rotational invariant, thus they are unable to actually encode the differences between each orientation of the effector.

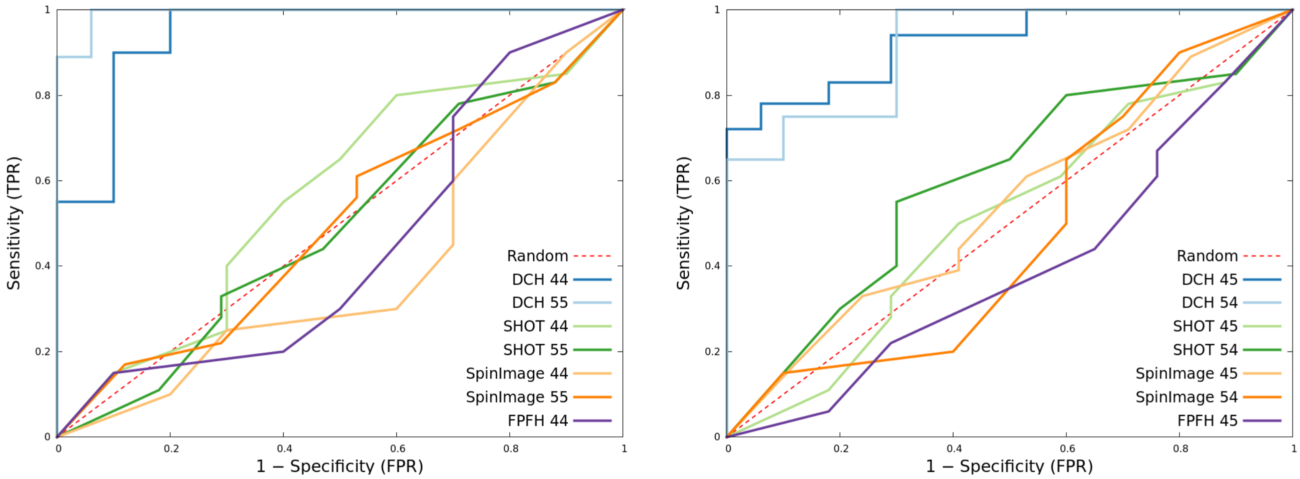
From our results, Spin Images is the descriptor with the worst performance, while SHOT showed the best results, leaving FPFH in an intermediate position. The particularly bad results obtained with Spin Images can be explained by taking into account the fact that this descriptor projects the neighborhood of the target point using a local coordinate system. However, if the sensory data do not cover a spatial region with enough extension, such projection might generate an image with too few features to provide a useful descriptor. One could tackle this problem by increasing the size of vicinity of the target point, but in turn this would increase the amount of data used to compute the descriptor. Similarly, for DCH if the vicinity is too small to provide enough information to encode the features of the underlying surface, or if it do not contains any useful information, the method may not be able to identify any re. This issue also explains most of the failed attempts during the evaluation with DCH.

8. Conclusions

The current work proposed a novel 3D local descriptor, designed for use in the problem of robotic grasping. This descriptor aims to encode information about how the curvature of an object changes around a point.

The experiments showed that the proposed descriptor, dubbed Directed Curvature Histogram, is well suited for solving the grasping problem following a data-driven approach. These tests showed that DCH resulted particularly helpful in the identification of the required orientation of the effector to successfully grasp an object.

Additionally, the results showed that the proposed descriptor is



(a) *Classification performance observed on the experiments 1 and 2.* **(b)** *Classification performance observed on the experiments 3 and 4.*

Figure 7: *Performance of the classifiers trained using DCH versus the classifiers trained using other descriptors.*

better suited for the grasping task than one recent and widely used descriptor. This highlights the advantage of using a task oriented descriptor to solve this problem.

As a future work, we plan to investigate the applicability of the presented descriptor to other tasks.

Acknowledgements

This research was partially founded by Conicyt-Fondecyt project 1140783.

References

[BBGO11] BRONSTEIN A. M., BRONSTEIN M. M., GUIBAS L. J., OVSJANIKOV M.: Shape google. *ACM Transactions on Graphics 30*, 1 (1 2011), 1–20. 2

[BKS*05] BUSTOS B., KEIM D., SAUPE D., SCHRECK T., VRANIĆ D. V.: Feature-based similarity search in 3D object databases. *ACM Computing Surveys 37*, 4 (12 2005), 345–387. 2

[BMAK13] BOHG J., MORALES A., ASFOUR T., KRAGIC D.: Data-Driven Grasp Synthesis: A Survey. *IEEE Transactions on Robotics 30*, 2 (4 2013), 1–21. 2

[FBD14] FEIX T., BULLOCK I. M., DOLLAR A. M.: Analysis of human grasping behavior: Object characteristics and grasp type. *IEEE Transactions on Haptics 7*, 3 (2014), 311–323. 1

[FHK*04] FROME A., HUBER D., KOLLURI R., BÜLOW T., MALIK J.: Recognizing Objects in Range Data Using Regional Point Descriptors. *Computer Vision - ECCV 2004: 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part III 3023* (2004), 224–237. 2

[GSB*13] GUO Y., SOHEL F. A., BENNAMOUN M., WAN J., LU M.: RoPS: A local feature descriptor for 3D rigid objects based on rotational projection statistics. In *2013 1st International Conference on Communications, Signal Processing and Their Applications, ICCSPA 2013* (2 2013), IEEE, pp. 1–6. 2

[Joh97] JOHNSON A.: Spin-images: a representation for 3-D surface matching. *Carnegie Mellon University*, CMU-RI-TR-97-47 (1997), 138. 2

[KH04] KOENIG N., HOWARD A.: Design and use paradigms for gazebo, an open-source multi-robot simulator. *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566) 3* (2004), 2149–2154. 5

[LMSB14] LEÓN B., MORALES A., SANCHO-BRU J.: Robot Grasping Foundations. In *Cognitive Systems Monographs*, vol. 19. 2014, pp. 15–31. 2

[RBB09] RUSU R. B., BLODOW N., BEETZ M.: Fast Point Feature Histograms (FPFH) for 3D registration. *IEEE International Conference on Robotics and Automation* (2009), 3212–3217. 2

[RC11] RUSU R. B., COUSINS S.: 3D is here: Point Cloud Library (PCL). In *Proceedings - IEEE International Conference on Robotics and Automation* (5 2011), IEEE, pp. 1–4. 1

[RMBB08] RUSU R. B., MARTON Z. C., BLODOW N., BEETZ M.: Persistent point feature histograms for 3D point clouds. *Intelligent Autonomous Systems 10, IAS 2008* (2008), 119–128. 2

[Rus10] RUSU R. B.: Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments. *KI - Künstliche Intelligenz 24*, 4 (11 2010), 345–348. 5

[SC08] STEINWART I., CHRISTMANN A.: *Support Vector Machines*, 1st ed. Springer Publishing Company, Incorporated, 2008. 7

[SEKB12] SAHBANI A., EL-KHOURY S., BIDAUD P.: An overview of 3D object grasp synthesis algorithms. *Robotics and Autonomous Systems 60*, 3 (2012), 326–336. 2

[Shi96] SHIMOGA K.: Robot Grasp Synthesis Algorithms: A Survey. *The International Journal of Robotics Research 15*, 3 (6 1996), 230–266. 2

[TSDS10] TOMBARI F., SALTI S., DI STEFANO L.: Unique signatures of histograms for local surface description. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 6313 LNCS*, PART 3 (2010), 356–369. 2, 4

[Zha12] ZHANG Z.: Microsoft kinect sensor and its effect. *IEEE Multimedia 19*, 2 (2012), 4–10. 1