Floating Visual Grasp of Unknown Objects Using an Elastic Reconstruction Surface

Vincenzo Lippiello, Fabio Ruggiero, and Bruno Siciliano

Abstract In this paper a new method for fast visual grasp of unknown objects is presented. The method is composed of an object surface reconstruction algorithm and of a local grasp planner, evolving in a parallel way. The reconstruction algorithm makes use of images taken by a camera carried by the robot, mounted in an eye-in-hand configuration. An elastic reconstruction sphere, composed by masses interconnected each other by springs, is virtually placed around the object. The sphere is let to evolve dynamically under the action of external forces, which push the masses towards the object centroid. To smoothen the surface evolution, spatial dampers are attached to each mass. The reconstruction surface shrinks toward its center of mass until some pieces of its surface intercept the object visual hull, and thus local rejection forces are generated to push out the reconstruction points until they stay into the visual hull. This process shapes the sphere around the unknown object. Running in parallel to the reconstruction algorithm, the grasp planner moves the fingertips, floating on the current available reconstructed surface, according to suitable quality measures. The fingers keep moving towards local minima depending on the evolution of the reconstruction surface deformation. The process stops when the object has been completely reconstructed and the planner reaches a local minimum. Quality measures considering both hand and grasp proprieties are adopted. Simulations are presented, showing the effectiveness of the proposed algorithm.

1 INTRODUCTION

Operating in unstructured environments is a challenging research field which has not been widely investigated as far as the problem of grasping unknown objects is concerned.

PRISMA Lab, Dipartimento di Informatica e Sistemistica, Università di Napoli Federico II, Italy; e-mail: {vincenzo.lippiello, fabio.ruggiero, bruno.siciliano}@unina.it.

Grasping and manipulation tasks, in general, require a priori knowledge about the object geometry. One of the first approaches to grasping in unknown environments can be found in [26], where visual control of grasping is performed employing visual information to track both object and fingers positions. A method to grasp an unknown object using information provided by a deformable contour model algorithm is proposed in [16]. In [25] an omnidirectional camera is used to recognize the shape of the unknown object while grasping is achieved on the basis of a grasping quality measure, using a soft-fingered hand.

From another point of view, it is straightforward to recognize that two main tasks have to be performed to achieve unknown objects grasping: object recognition/reconstruction and grasp planning.

In literature, different methods have been proposed to cope with 3D model reconstruction of unknown objects. The main differences lies in how the available object images are processed and on the algorithms used for object reconstruction. A number of algorithms can be classified under the so called *volumetric scene reconstruction* approach [6]. This category can be further divided into two main groups: the *shape from silhouettes* and the *shape from photo-consistency* algorithms. Another method, proposed in [24], considers a surface that moves towards the object under the influence of internal forces, produced by the surface itself, and external forces, given by the image data.

The finite-elements method is used in [5] to reconstruct both 2D and 3D object boundaries. Using an active contour model, data extracted from taken images are employed to generate a pressure force on the active contour that inflate or deflate the curve, making its behavior like a balloon.

A technique for computing a polyhedral representation of the *visual hull* [12] –the set of points in the space that are projected inside each image silhouette– is studied in [9]. Other approaches rely on the use of *apparent contours* [4, 19], where the reconstruction is based on the spatio-temporal analysis of deformable silhouettes.

On the other hand, grasp planning techniques rely upon the choice of grasp quality measures used to select suitable grasp points. Several quality measures proposed in the literature depend on the properties of the grasp matrix, that means on the grasp geometry; others are based on the area of the polygon created by the contact points, or on the external resistent wrench.

In [7] two generally optimal criteria are introduced, where the total finger force and the maximum finger force are considered, while in [15] simple geometric conditions to reach an optimal force closure grasp both in 2-D and in 3-D are found. The geometric properties of the grasp are also used in [13] to define quality measures, as well as suitable task ellipsoids in the wrench space of the object are proposed to evaluate grasp quality also with respect to the particular manipulation task.

Measures depending on the hand configurations [20] define a set of quality measures based on the evaluation of the capability of the hand to realize the optimal grasp. A rich survey of these grasp quality measures can be found in [22].

To plan a grasp for a particular robotic hand, quality measures depending both on grasp geometry and on hand configuration should be taken into account. Few papers

address the problem of grasping an unknown object using a given robotic hand, able to reach the desired contact points in a dexterous configuration [1, 3, 8, 10, 11].

A grasp control task is considered in [17], where several controllers are combined to reach different wrench closure configurations, while in [18] grasp prototypes – generalization of example grasps– are used as starting points in a search for good grasps.

In this paper, a new method for fast visual grasping of unknown objects using a camera mounted on a robot in an eye-in-hand configuration is presented. This method is composed of an *object surface reconstruction algorithm* and of a *local grasp planner*, which evolve in a synchronized parallel way. The reconstruction algorithm makes use of images taken with a camera carried by the robot. First, a rough estimation of the object position and dimensions is performed, and an *elastic reconstruction surface* with a spherical shape is virtually placed around the object. Then, the fingertips of the robotic hand are suitably attached on it, at a suitable *floating safety distance*. The reconstruction surface is sampled by points, which are endowed with a virtual mass, and are interconnected with each other by means of virtual springs resulting in a cross reticulum. A virtual spatial damper is also considered to smoothen the motion of the surface sample points. During the reconstruction process, the elastic surface evolves in a dynamic way leaning to shrink on itself under the action of reticular elastic forces and external forces, pointing to the contours of the visual hull defined by the object silhouettes. For each mass, the external *compression forces* becomes *repulsive forces* along the outgoing direction with respect to the visual hull if the reconstruction point comes into the visual hull, while they return to be attractive when the point comes out of the visual hull. The amplitude of the external forces is progressively reduced during the chattering around the contour of the visual hull to guarantee that a dynamic equilibrium between external and elastic forces is quickly reached. The current reconstruction surface is buffered and made available, with a fixed sample time, to the planner. This moves the fingers, floating on this current available reconstructed surface at an imposed security distance, according to suitable quality measures. Due to the consequent motion floating effect around the object, the proposed method has been called *Floating Visual Grasp*. The fingertips keep moving toward a local minimum on the current reconstructed surface as long as it continues to change due to the updates provided by the reconstruction process. Quality measures considering both hand and grasp proprieties are adopted for the local planner: the directions of the finger motion leading toward grasp configurations that are not physically reachable or causing collisions or loss of hand manipulability are discarded. Moreover, a discretized version of the quality measure proposed in [15] is applied to the survived possible motion directions to select those leading toward an optimal (in a local sense) grasp configuration. Notice that many other quality measures may be chosen in substitution of those proposed in [15], without affecting the the proposed framework.

Due to the intrinsic smoothness and safety with respect to collisions of the planned trajectories, the execution of the grasp can also be executed in real-time during the trajectory generation. This allows the proposed algorithm to be used for kinematic control purposes.

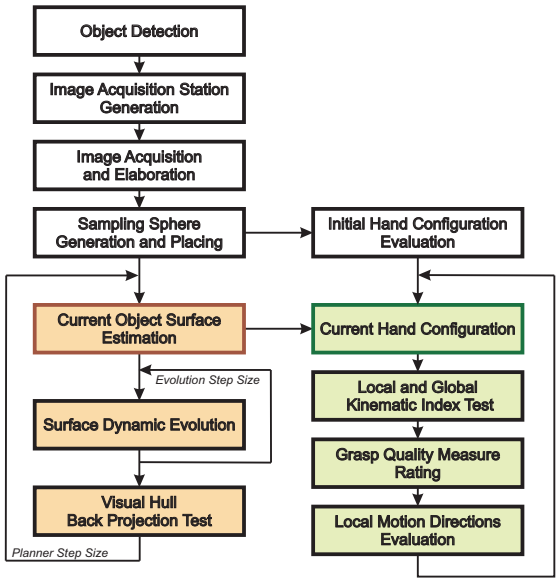


Fig. 1 Block diagram of the visual grasp algorithm.

Simulations results are presented to show the performance of the proposed algorithm.

2 FLOATING VISUAL GRASP ALGORITHM

The block diagram in Fig. 1 shows the data flow and the main elaboration steps of the proposed visual grasp algorithm. The elaboration processes may be arranged into three main groups: some *preparation steps*, the *object surface reconstruction algorithm*, and the *local grasp planner*.

The preparation steps of the algorithm start with a detection algorithm, that is based on a classical blob analysis, to evaluate the presence of an object in the field of view of the camera. Successively, by holding the optical axis perpendicular to the plane where the object has been detected, the camera is moved until the optical axis intercepts the centroid of the object. At the end of this step, the camera is exactly over the unknown object and ready to start the image acquisition process. Moreover, during this step, a rough estimation of the object center of mass is evaluated using the centroid of the object shape extracted from some images.

The image acquisition stations are chosen as illustrated in Fig. 2, while the concerning image acquisition step is carried out as follows: 1) an image is acquired from the top of the object; 2) a subset of n_1 images is taken from camera stations equally distributed over a circular path of radius r_1, with the optical axis of the camera pointing to the estimated center of the object and forming an angle α_1 with respect to the revolution axis $\mathbf{z}$; 3) a subset of n_2 images is acquired as in 2), but

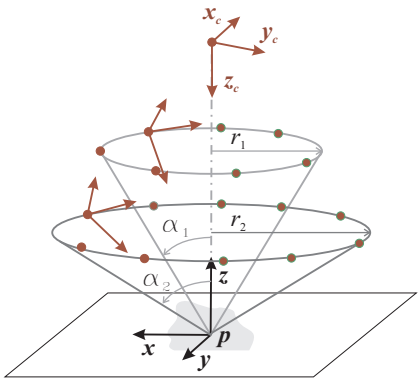


Fig. 2 Camera stations (bullets) and trajectories of the camera during image acquisition.

using a radius r_2 and an angle α_2. In the following, the total number of acquired images will be denoted as $n = n_1 + n_2 + 1$.

At this point, a blob analysis technique is employed to determine the silhouette of the object for each image. Each silhouette is also improved using suitable filtering techniques (e.g. dilatation and erosion iterative process) to reduce the effects of the image noise, and the centroid of the corresponding blob is evaluated. Then, the center of mass of the object, assuming homogeneous mass distribution, is estimated using a least-squares triangulation method.

On the basis of the dimension of each silhouette, the radius r_s of a 3D spherical surface that may contain the object is estimated, adding a suitable safety margin to the object estimated dimension. Finally, the *reconstruction sphere* with radius r_s, centered at the estimated center of mass of the object, and sampled with a number of n_s points is built. The position of each sample on the sphere is chosen in a way to achieve an initial uniform distribution of the reconstruction points, avoiding a thickening of points around the two poles (north and south) of the sphere.

A virtual mass is associated to each sample point of the reconstruction spherical surface, and four links are imposed with the closest cross points interposing springs. The resulting elastic reconstruction surface is shown in Fig. 3, where it is noticeable that the links between the masses dived the sphere into a certain number of parallels and meridians.

The initial grasp configuration of the hand can be set on the basis of the initial reconstruction sphere. In this paper, a three-fingered hand and point contact type at the fingertips are considered. Hence, a direct correspondence between the position of a point on the sphere and the position of each finger of the hand can be assumed. Due to the symmetry of the sphere, an infinite number of grasp configurations, ensuring force-closure grasp, could be initially selected. To this purpose, it is well known that a three contact grasp of a sphere is force-closure if the contact points are 120° apart on the same plane (neglecting moments and transversal forces). Therefore, the plane parallel to the floor halving the sphere, corresponding to the parallel with the maximal area that is graspable with respect to the hand kinematics, is chosen. Finally, three points 120° apart are selected on this parallel according to the orientation of

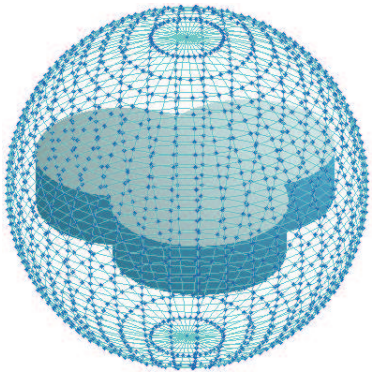


Fig. 3 Elastic reconstruction sphere surrounding the object.

the major and minor axis of the silhouette observed from the station just over the object –the observation station characterized by an optical axis perpendicular to the parallels of the sphere–.

At this point, both the object model reconstruction process and the local planner start in parallel and cooperate to the final goal. In particular, as shown in Fig. 1, the reconstruction algorithm updates in real-time the estimation of the current reconstructed object surface, while the local planner, on the basis of the current estimation, computes the fingers trajectories toward the current local optimal configuration for the grasp. Notice that a force optimization algorithm, e.g. [2], could be used for a proper distribution of grip forces.

These two parallel processes are independent and can be allocated under two different threads and, in a multi-processor system, also on different CPUs. In other words, the proposed method exhibits an intrinsic capability to be run in parallel. More details of these two crucial steps are provided in the next two sections.

3 OBJECT SURFACE RECONSTRUCTION

The object surface reconstruction algorithm employed in this paper is an evolution of the method proposed in [14].

As described in the previous sections, from the set of n silhouettes of the object a reconstruction spherical surface is created and sampled by points. A virtual mass is associated to each sample point, and four links are imposed with the closest cross points with springs, resulting in a cross reticular topology for the reconstruction surface (see Fig. 4). The two poles of the sphere are connected with all the masses of the nearest parallel of the resulting spherical reticulum.

Each parallel of the sphere should have the same n_m number of points, corresponding to the number of meridian, allowing the construction of a cross reticulum fully linked. In other words, for each point, the existence of a couple of corresponding points on the closest parallels of the spherical grid is guaranteed. Without loss

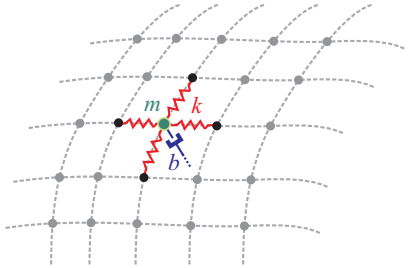


Fig. 4 Cross network topology of the reconstruction surface with the virtual mass, springs and spatial damper of the i-th sample points.

in generality, the number of parallels n_p is chosen equal to the number of meridians $n_p = n_m = \sqrt{n_s - 2}$. To avoid that the parallel nearest to the poles determines an unnecessary initial thickening of sample points around the poles, a suitable angular distribution of the parallels has been carried out, reducing (augmenting) the density of the parallels near the poles (equator).

The model of the system, defining the motion of each sample point of the reconstruction surface, is defined by the following dynamic equations:

$$m\ddot{\boldsymbol{x}}_{i,j} + b\dot{\boldsymbol{x}}_{i,j} + k(4\boldsymbol{x}_{i,j} - \boldsymbol{x}_{i-1,j} - \boldsymbol{x}_{i,j+1} - \boldsymbol{x}_{i+1,j} - \boldsymbol{x}_{i,j-1}) = \boldsymbol{f}_{i,j}, \tag{1}$$

for $i = 1,\ldots,n_m$ and $j = 1,\ldots,n_p$, denoting with $\boldsymbol{x}_{i,j}$ the position in the workspace of the sampling point at the intersection of the i-th meridian with the j-th parallel –reticular position (i,j)–, with m, k, and b the mass associated to the point, the constant spring linking the point with its nearest four points of the cross of the reticulum, and the constant spatial damper, respectively. Notice that subscript $i = j = 0$ ($i = n_m + 1$ and $j = n_p + 1$) for the representation of the four nearest points in the reticulum corresponds to $i = n_m$ and $j = n_p$ ($i = j = 1$), respectively.

The term $\boldsymbol{f}_{i,j}$ is the external force acting on the mass associated to the sample point (i,j), attractive with respect to the border of the visual hull, which is progressively reduced once the corresponding point comes in or out from the visual hull:

$$\boldsymbol{f}_{i,j} = \alpha_{i,j}(t_{i,j}) F_a \boldsymbol{n}_{i,j}, \tag{2}$$

where $\boldsymbol{n}_{i,j}$ is the unit vector pointing from the current point (i,j) to the estimated centroid of the object, that defines the direction of the force, and $\alpha_{i,j}(t_{i,j})F_a$ is the amplitude of the force. Also, F_a the maximum force module and $\alpha_{i,j}(t_{i,j}) \in (-1,1]$ a discrete numerical sequence of scale factors defined as follow:

$$\alpha_{i,j}(t_{i,j}) = -\varepsilon \alpha_{i,j}(t_{i,j} - 1) \tag{3}$$

where $\varepsilon \in (0,1)$ (e.g. $\varepsilon = 0.9$), $\alpha_{i,j}(0) = 1$, and $t_{i,j} = 0,\ldots,\infty$ is a discrete step time incremented every time the point (i,j) comes in or out of the visual hull.

The two poles have to be treated separately due to their topological peculiarity. The previous model becomes

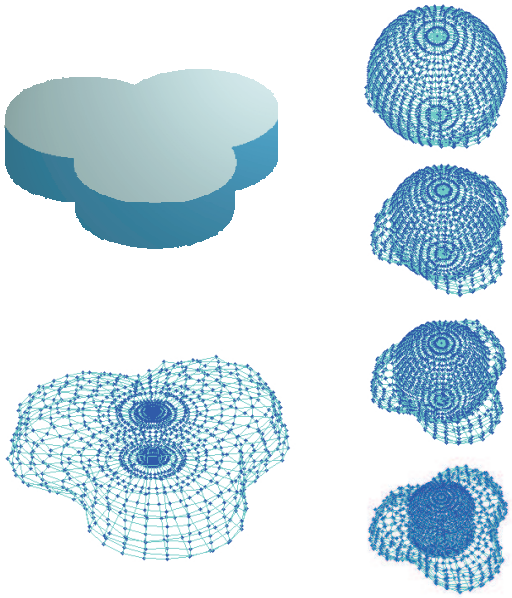


Fig. 5 Steps of the object surface reconstruction process.

$$m\ddot{\boldsymbol{x}}_n + b\dot{\boldsymbol{x}}_n + k(n_m\boldsymbol{x}_n - \sum_{j=1}^{n_m} \boldsymbol{x}_{1,j}) = \boldsymbol{f}_n \tag{4}$$

for the north pole, and

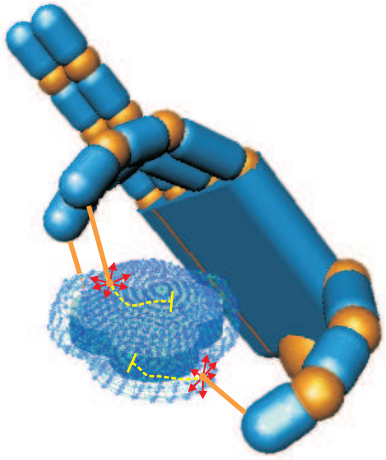
$$m\ddot{\boldsymbol{x}}_s + b\dot{\boldsymbol{x}}_s + k(n_m\boldsymbol{x}_s - \sum_{j=1}^{n_m} \boldsymbol{x}_{n_p,j}) = \boldsymbol{f}_s \tag{5}$$

for the south pole, where the subscripts n and s indicate quantities refereed to the north and south pole, respectively.

The stability of the system, for any non-trivial initial condition of the sphere, leads the reconstruction elastic surface to contract toward its center of mass until the visual hull is intersected. The dynamic evolution of the system reaches the equilibrium when the shape of the surface produces a dynamic equilibrium between the elastic forces generated by the grid and the repulsive forces depending on the contours of the visual hull. The result is that the initial elastic reconstruction sphere shapes itself on the unknown object.

The accuracy of the reconstruction process depends on the number of views, on the distribution of the observation stations, and on the density of points of the reconstruction sphere. On the other hand, the computational time of the algorithm increases if n and/or n_s are increased. However, considering that the final goal of the process is the object grasping and not the model reconstruction, which can be considered as a secondary outcome of the proposed method, the accuracy of the

Fig. 6 Floating visual grasp.



reconstruction process needs only to be adequate for the requirements of the grasp planner algorithm.

In Fig. 5 some images showing intermediate steps of the reconstruction algorithm of a synthesized object are shown, with parameters: $\alpha_1 = 45°$, $\alpha_2 = 80°$, $n_1 = 4$, $n_2 = 8$, and $n_s = 1602$.

4 LOCAL GRASP PLANNER

The current estimation of the object surface, which is stored in a buffer and is employed by the local grasp planner for updating the fingertips trajectories, is continuously updated during the dynamic evolution of the elastic surface. The local grasp planner, in accordance with the current reconstructed object surface, generates the fingertips trajectories in a floating manner, keeping a fixed *floating safety distance* δ_f between the fingers and the corresponding sample point along the outgoing normals, on the basis of suitable quality indices (see Fig. 6).

Namely, starting from the initial grasp configuration, chosen as described in the previous sections, the planner generates the motion of the fingers from the current position to a new point of the updated surface. In particular, the contact points of the grasp are first "virtually" moved to the updated surface, achieving an initial "target" grasp configuration. Then, for each contact point of the current target grasp configuration, the contour made by the eight neighboring points of the surface is selected. Considering the contours of all the contact points, the set of all the combinations of possible reachable grasp configurations is evaluated on the basis of suitable quality measures. If the current grasp configuration of the set has a value better than the value of the target configuration, this is chosen as the new target grasp configuration, establishing "de facto" the motion direction of each finger, as shown in Fig. 7.

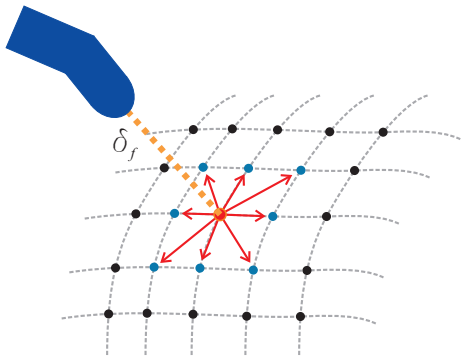


Fig. 7 Contour of neighbor points of the current target grasp point.

The process is repeated in a recursive manner until there are no more improvements of the quality measures, and it restarts at the next step time. The whole process ends when the object reconstruction algorithm reaches an equilibrium and the planner computes the final grasp configuration. Hence, the safety floating distance δ_f is progressively reduced to achieve the desired grasp action.

The floating distance is used to avoid collisions of the fingers with the object during the object reconstruction and approach, and before the final grasp configuration is reached. Moreover, the final progressive reduction of the floating distance implies that each fingertip moves perpendicularly to the surface.

Notice that only few points are considered as candidates for the next grasp configuration, so that the number of combinations to be inspected is limited, resulting in a computationally fast algorithm; moreover, a certain number of grasp configurations are discarded during the evaluation process. Namely, a *local kinematic index* is used to discard all the candidate grasp configurations that cannot be reached. Then, a *global kinematic index* is adopted to discard all the configurations causing finger collisions or lack of manipulability for the hand. Finally, a *grasp quality measure* is applied to the remaining configurations to evaluate possible improvements of the grasp quality.

The computational efficiency of the local planner jointed with the possibility to be executed in parallel with the reconstruction algorithm, eventually on two different elaboration units, resulting in a faster solution with respect to traditional methods, which first reconstruct the object and then plan the grasp on the whole knowledge of the environment.

In the next subsections, the quality indices and the finger trajectory planner are presented.

4.1 Local and global kinematic indices

On the basis of the finger kinematics, the local kinematic index allows discarding all the candidate contact points that cannot be reached. With reference to a single

contact point and finger, the kinematic test is carried out for all the contour points (see Fig.7). Namely, for each point, finger joints are computed using an inverse kinematics algorithm. Hence, those points for which joint limits are exceeded, or that are too close to a kinematic singularity, are discarded. This latter condition is evaluated on the basis of the condition number of the finger Jacobian. Avoiding singularities in the finger Jacobian allows being far from hand singularities.

A standard CLIK algorithm [21] is adopted to compute the inverse kinematics; in particular, the scheme based on the transpose of the Jacobian has been used to achieve a faster computation, together with a Singular Value Decomposition technique for the evaluation of the Jacobian condition number.

For the remaining points of the contour, the global kinematic index computes the distance between the fingers corresponding to all the possible grasp configurations. Hence, all the configurations for which the distances are under a given safety threshold, are discarded.

4.2 Grasp quality measure

The grasp quality is evaluated only for the configurations that are left after the kinematic tests. The method proposed in [15] is adopted, suitably modified to cope with the discretization of the grasp configurations, assuming neglectable moments and transversal forces.

Let us denote with $\boldsymbol{w} = \begin{bmatrix} \mathbf{f}^{\mathrm{T}} & \boldsymbol{\mu}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ the wrench vector collecting the force $\mathbf{f}$ and moment $\boldsymbol{\mu}$. Assuming that the finger forces are applied along the direction normal to the object surface, the force direction is specified only by the contact point.

Let $\mathscr{W}$ denote the space of wrenches, $\mathscr{W}_f \subset \mathscr{W}$ the space of unit forces acting in the grip plane, that is the plane containing the three contact points, through the center of grip, $\mathscr{W}_{\perp\mu} \subset \mathscr{W}$ the space of pure moments acting along the direction perpendicular to the grip plane. Moreover, let $\mathbf{g}^{-1}(-\mathbf{w})$ denote the set of finger forces which can resist the external wrench $\mathbf{w}$.

Finally, consider the quantity

$$Q_1 = \min_{\mathbf{w} \in \mathscr{W}_f} \left(\max_{\mathbf{f} \in \mathbf{g}^{-1}(-\mathbf{w})} \frac{1}{\| \mathbf{f} \|_f} \right), \tag{6}$$

which is a measure of the grasp ability to resist unit forces in the grip plane, and the quantity

$$Q_2 = \min_{\mathbf{w} \in \mathscr{W}_{\perp\mu}} \left(\max_{\mathbf{f} \in \mathbf{g}^{-1}(-\mathbf{w})} \frac{1}{\| \mathbf{f} \|_f} \right), \tag{7}$$

which is a measure of the grasp ability to resist unit moments normal to the grip plane.

The optimal grasp proposed in [15] is defined as the grasp that maximizes Q_2 among all grasps which maximize Q_1.

It can be proven (see [15]) that the optimum grasp with three fingers in a 2-D case under the above optimal criterion is reached when the normal forces are symmetric, with directions spaced 120° apart. Moreover, this grasp maximizes also the size of the outer triangle, defined as the triangle formed by the three lines perpendicular to the normal finger forces passing through the respective contact points. Under the same criterion, the optimum grasp with three fingers in a 3-D case is achieved when the maximum circumscribing prism-shaped grasp, that has the largest outer triangle, is selected among the grasps where the normal finger forces lie within the same grip plane and are in equilateral configuration.

Therefore, to reach the optimum in the 3-D case with three fingers, the planner has to seek three points in equilateral configuration on the object surface, so that the normal forces lie in the same grip plane, and for which the circumscribing prism grasp is maximum.

In the case presented in this paper, since the reconstructed object surface is sampled by points/masses, the above method cannot be directly applied. Differently from the continuous case, due to the presence of a finite set of sampled points, the existence of a "grip plane" containing all the normal forces is not guaranteed. This is mainly due to the fact that, because of the discretization, the surface normals are an approximation of the real ones. Considering that the optimal criterion requires that the desired normals have to be spaced 120° apart, a discretized implementation of the method of [15] is proposed here.

For each candidate configuration of three grasp points, the normal directions are estimated on the basis of the available point-wise approximation of the surface. Then, the unit vector normal to the grip plane containing the three points is evaluated. Denoting with ϑ_j the angle between the direction of the normal force applied to point j and the direction normal to the grip plane, a Coplanarity Error Index (CEI) can be defined as follow:

$$CEI = \frac{\sum_{j=1}^{3} \mid \vartheta_j - 90^\circ \mid}{3}. \tag{8}$$

Obviously, the closer CEI to zero, the more the normal forces lie in the same plane. The definition of a threshold Φ_{CEI} allows discarding all those configurations having a value of CEI higher than Φ_{CEI}; hence, all the remaining grasp configurations are assumed to have forces lying in the same grip plane and can be further processed.

The next step consists in looking for an equilateral grasp configuration. To this aim, for each grasp configuration, the unit vector normal to the object surface at each contact point is projected on the grip plane. Denoting with φ_j the angle between these projections for each of the 3 couple of points of the considered configuration, an Equilateral Error Grasp Index ($EEGI$) can be defined as:

$$EEGI = \frac{\sum_{j=1}^{3} \mid \varphi_j - 120^\circ \mid}{3}. \tag{9}$$

Clearly, the closer $EEGI$ to zero, the nearer the configuration to an equilateral grasp. The definition of a threshold Φ_{EEGI} allows discarding all those configurations with

a value of $EEGI$ higher than Φ_{EEGI}; hence, all the remaining grasp configurations are assumed to be equilateral.

Among all the equilateral configurations, the maximum circumscribing prism has to be found; if the grasp configuration associated with the largest prism is different from the current target configuration, this is taken as the new grasp configuration.

Notice that, in the case that the grasp configuration changes, the whole process starts again with the new contact points, by considering the new contours and applying the complete sequence of index-based tests starting from the kinematic ones. The algorithm stops if the best grasp configuration remains unchanged at the end of the optimization process, or in the case that all the candidate grasp configurations are discarded during the process.

4.3 Finger trajectory planner

The local grasp planner produces a sequence of intermediate target grasp configurations at each iteration of the object reconstruction algorithm. These end with the optimal grasp configuration (in local sense), as illustrated in Fig. 6. The intermediate configurations are used to generate the finger paths.

Namely, the sequence of intermediate configurations is suitably filtered by a spatial low-pass filter in order to achieve a smooth path for the fingers on the object surface. To this purpose, notice that only the final configuration needs to be reached exactly, while the intermediate configurations can be considered as via points.

With respect to the smooth paths through the points of the filtered configurations, the actual finger paths generated by the finger trajectory planner keep a floating distance δ_f along the normal to the surface (as explained in Section 4). This feature produces a floating effect of the fingers over the reconstructing object surface during the reconstruction process while they move according to the deformation of the reconstruction sphere. When the final configuration is reached, this safety distance is progressively reduced to zero, producing the desired grasp action.

5 SIMULATIONS

The proposed method has been tested with simulations using synthesized objects. The first used object is shown in Fig. 5, while for the trajectory planning only the first three fingers of the virtual hand shown in Fig. 6 have been considered.

The dynamic parameters of the reconstruction sphere have been chosen as follows: the total mass of the sphere $M = 10^{-3}\ kg$, $k = 0.3 \cdot 10^{-3}\ N/m$, $b = 0.09 \cdot 10^{-3}\ Ns/m$, and $F_a = 5\ N$.

By setting $\Phi_{CEI} = 15^\circ$ and $\Phi_{EEGI} = 10^\circ$, the grasp configuration and the finger trajectories are those of Fig. 8. The final grasp configuration (which can be proven to

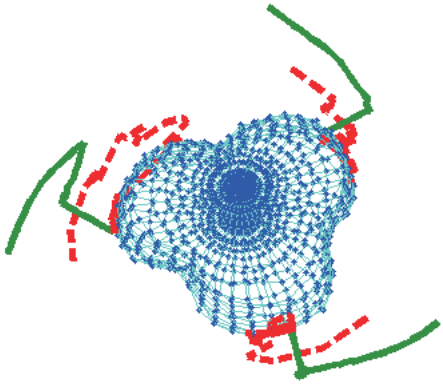


Fig. 8 Finger trajectories evaluated by the local grasp planner (continuous lines) and the corresponding sequence of grasp points on the reconstructed object surface (dotted lines).

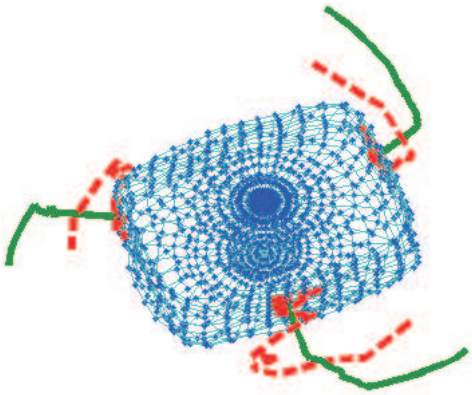


Fig. 9 Finger trajectories evaluated by the local grasp planner (continuous lines) and the corresponding sequence of grasp points on the reconstructed surface for the smooth prism (dotted lines).

be the global optimal grasp configuration) is characterized by the values $CEI = 13^\circ$ and $EEGI = 2.2^\circ$.

A prism with smooth lateral corners has also been considered. The grasp configuration and the corresponding trajectories are shown in Fig. 9. The values $CEI = 9.9^\circ$ and $EEGI = 1.04^\circ$ are obtained in the final configuration. Remarkably, an equilateral symmetry is achieved in the final grasp configuration: two fingers are placed on the smooth corners, and the other finger is placed in the middle of the opposite surface. This configuration corresponds to an opposite grasp ensuring force closure; as before, it can be proven that this grasp configuration is optimal also in a global sense, although the proposed approach can only guarantee that a local minimum is achieved.

6 CONCLUSION AND FUTURE WORK

6.1 Conclusion

A new method for fast visual grasp of unknown objects has been presented. The proposed method is composed of a fast iterative object surface reconstruction algorithm and of a local optimal grasp planner, which evolve in a synchronized parallel way. An eye-in-hand camera is adopted to acquire the images used by the reconstruction algorithm. A reconstruction elastic sphere sampled with points furnished of mass and linked with each other by springs, which is virtually placed around the object, dynamic evolves collapsing towards the visual hull of the object. An attractive force pushes each mass to the visual hull and is progressively reduced when the visual hull has been reached. During the reconstruction process, the planner moves the fingertips, floating on the current reconstructed surface at a safety distance, according to local and global kinematic indices and grasp quality measures. The effectiveness of the proposed method has been shown in simulation.

6.2 Future Work

The adoption of suitable pre-shaping techniques, that could be helpful also for the determination of the grasp as explained in [23], for the choice of the initial grasp configuration, which is based on visual information, may be a further improvement of the proposed algorithm. Moreover, quality indices connected to the tasks to be performed by the hand with the object and the adoption of other kinematic constraints, e.g. collision avoidance of the object with the hand's palm, could be considered. Finally, the extension of the proposed approach to the case of more than three fingers or to bi-manual manipulation, with the adoption of suitable quality measures, will be investigated.

Acknowledgements The research leading to these results has been supported by the DEXMART Large-scale integrating project, which has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement ICT-216239. The authors are solely responsible for its content. It does not represent the opinion of the European Community and the Community is not responsible for any use that might be made of the information contained therein.

References

1. C. Borst, M. Fischer and G. Hirzinger, "Calculating hand configurations for precision and pinch graps", *IEEE/RSJ Confonference on Intelligent Robots and Systems*, Lausanne, 2002.

2. M. Buss, H. Hashimoto and J. B. Moore, "Dexterous hand grasping force optimization", *IEEE Transaction on Robotics and Automation*, vol. 12, no.3, pp. 406-418, 1996.
3. E. Chinellato, R. B. Fisher, A. Morales and A. P. del Pobil, "Ranking planar grasp configurations for a three-fingered hand", *IEEE International Conference on Robotics and Automation*, Taipei, 2003.
4. R. Cipolla and A. Blake, "Surface shape from the deformation of apparent contours", *Internation Journal of Computer Vision*, vol. 9, no. 2, pp. 83-112, 1992.
5. L. D. Cohen, "On active contour models and balloons", *Computer Vision, Graphics, and Image Processing: Image Understanding*, vol. 53, no. 2, pp. 211-229, 1991.
6. C. R. Dyer, "Volumetric scene reconstruction from multiple views", *Foundations of Image Analysis*, L.S. Davis ed., Kluwer, Boston, 2001.
7. C. Ferrari and J. Canny, "Planning optimal grasps", *IEEE International Conference on Robotics and Automation*, Nice, 1992.
8. M. Fischer and G. Hirzinger, "Fast planning of precision grasps for 3D objects", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Grenoble, 1997.
9. J.-S. Franco and E. Boyer "Exact polyhedral visual hulls", *British Machine Vision Conference*, 2003.
10. Y. Guan and H. Zhang, "Kinematic feasibility analysis of 3D grasps", *IEEE International Conference on Robotics and Automation*, Seoul, 2001.
11. R. D. Hester, M. Cetin, C. Kapoor and D. Tesar, "A criteria-based approach to grasp synthesis", *IEEE International Conference on Robotics and Automation*, Detroit, 1999.
12. A. Laurentini, "How far 3D shapes can be understood from 2D silhouettes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 188–195, 1995.
13. Z. Li and S. S. Sastry, "Task-oriented optimal grasping by multifingered robot hands", *IEEE Journal of Robotics and Automation*, vol. 4, no. 1, pp. 32-44, 1988.
14. V. Lippiello and F. Ruggiero, "Surface model reconstruction of 3D objects from multiple views", *IEEE International Conference on Robotics and Automation*, Kobe, 2009.
15. B. Mirtich and J. Canny, "Easily computable optimum graps in 2-D and 3-D", *IEEE International Conference on Robotics and Automation*, San Diego, 1994.
16. D. Perrin, C. E. Smith, O. Masoud and N. P. Papanikolopoulos, "Unknown object grasping using statistical pressure models", *IEEE International Conference on Robotics and Automation*, San Francisco, 2000.
17. R. Platt, A. H. Fagg and R. A. Grupen, "Manipulation gaits: sequences of grasp control tasks", *IEEE International Conference on Robotics and Automation*, New Orleans, 2004.
18. N. S. Pollard, "Synthesizing grasps from generalized prototypes", *IEEE International Conference on Robotics and Automation*, Minneapolis, 1996.
19. S. Prakoonwit and R. Benjamin, "3D surface point and wireframe reconstruction from multiview photographic images", *Image and Vision Computing*, vol. 25, pp. 1509–1518, 2007.
20. K. B. Shimoga, "Robot grasp synthesis algorithms: A survey", *International Journal of Robotics Research*, vol. 15, n. 3, pp. 230-266, 1996.
21. B. Siciliano, L. Sciavicco, L.Villani and G. Oriolo, *Robotics: Modelling, Planning and Control*, Springer, London, 2009.
22. R. Suarez, M. Roa and J. Cornella, "Grasp quality measures", *Technical Report IOC-DT-P-2006-10*, Universitat Politecnica de Catalunya, Institut d'Organitzacio i Control de Sistemes Industrials, 2006.
23. D. Wren and R. B. Fisher, "Dextrous hand grasping strategies using preshapes and digit trajectories", *IEEE International Conference on Systems, Man and Cybernetics*, Vancouver, 1995.
24. C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow", *IEEE Transactions on Image Processing*, vol. 7, no. 3, 1998.
25. T. Yoshikawa, M. Koeda and H. Fujimoto, "Shape recognition and grasping by robotic hands with soft fingers and omnidirectional camera", *IEEE International Conference on Robotics and Automation*, Pasadena, 2008.
26. B. H. Yoshimi and P. K. Allen, "Visual control of grasping and manipulation tasks", *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Las Vegas, 1994.