

Ajuste probabilístico de múltiples primitivas cóncavas y convexas a nubes de puntos parciales

Menéndez, E.* , Martínez, S., Balaguer, C.

*RoboticsLab, Dpto. de Ingeniería de Sistemas y Automática, Universidad Carlos III de Madrid,
Av. De la Universidad 30, 28911 Leganés, España.*

Resumen

Este trabajo presenta un método probabilístico para ajustar múltiples primitivas geométricas (superelipsoides y superparaboloides con base desplazada) a nubes de puntos parciales obtenidas desde un único punto de vista mediante cámaras RGB-D. El ajuste se realiza mediante un proceso de estimación y maximización que refina iterativamente la asignación probabilística de los puntos y los parámetros que definen la primitiva geométrica. El modelo penaliza las formas que invaden el espacio vacío visible desde el sensor y emplea restricciones adicionales para favorecer reconstrucciones físicamente plausibles. Una contribución clave es la combinación secuencial de primitivas abiertas y cerradas, permitiendo representar con mayor fidelidad objetos de geometría compleja. El enfoque propuesto supera a métodos anteriores basados únicamente en formas cerradas, logrando reconstrucciones más precisas e interpretables. Esta capacidad resulta especialmente útil en escenarios de robótica asistencial, donde la reconstrucción precisa de objetos cotidianos es clave para una manipulación fiable y segura.

Palabras clave: Percepción y detección, Robots manipuladores, Reconstrucción de objetos, Ajuste probabilístico de formas, Nubes de puntos parciales, Primitivas geométricas

Probabilistic fitting of multiple convex and concave primitives to partial point clouds

Abstract

This work presents a probabilistic method for fitting multiple geometric primitives (superellipsoids and base-shifted superparaboloids) to partial point clouds captured from a single viewpoint using RGB-D cameras. The fitting process is formulated as an expectation-maximization procedure that iteratively refines both the probabilistic assignment of points and the parameters defining the geometric primitive. The model penalizes shapes that intrude into the visible free space of the sensor and incorporates additional constraints to promote physically plausible reconstructions. A key contribution is the sequential combination of open and closed primitives, which enables a more faithful representation of complex object geometries. The proposed approach outperforms previous methods that rely solely on closed shapes, achieving more accurate and interpretable reconstructions. This capability is particularly valuable in assistive robotics scenarios, where the precise reconstruction of everyday objects is essential for reliable and safe manipulation.

Keywords: Perception and detection, Robots manipulators, Object reconstruction, Probabilistic shape fitting, Partial point clouds, Geometric primitives

1. Introducción

La robótica asistencial plantea importantes desafíos, especialmente en lo que respecta a la percepción del entorno y la interacción física con objetos cotidianos. Para asistir eficazmente a las personas, un robot debe ser capaz de identificar, localizar

y manipular objetos de forma segura y precisa. Esta capacidad resulta especialmente compleja en escenarios reales, donde los objetos suelen estar ubicados sobre superficies compartidas y pueden estar parcialmente ocultos.

Tradicionalmente, el problema del agarre robótico se ha abordado mediante métodos analíticos que buscan configura-

*Autor para correspondencia: emenende@ing.uc3m.es
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

ciones de force-closure a partir de modelos físicos y geométricos precisos del objeto (Liu, 1999). Estos enfoques requieren información completa de la forma y propiedades del objeto, lo cual rara vez está disponible en entornos reales. Con el avance del aprendizaje automático, han surgido métodos basados en redes neuronales entrenadas con grandes bases de datos para predecir agarres directamente a partir de imágenes o nubes de puntos (Murali et al., 2020). Aunque han demostrado buenos resultados en tareas específicas, estos enfoques suelen tener dificultades para generalizar a objetos nuevos o escenas no vistas. En cambio, los humanos pueden decidir cómo agarrar un objeto desconocido casi instantáneamente. Esta capacidad se atribuye, en parte, a que el sistema visual humano abstrae los objetos en términos de primitivas geométricas simples en lugar de procesar directamente representaciones detalladas como nubes de puntos o mallas (Biederman, 1987).

En trabajos previos, abordamos el problema de la reconstrucción y manipulación de objetos en escenas reales mediante una representación basada en primitivas geométricas simples. En (Menendez et al., 2024a), propusimos un sistema que estima la forma de los objetos visibles a partir de una única vista del sensor RGB-D, ajustando supercuadráticas de nivel 1 (también conocidas como superelipsoides) a nubes de puntos segmentadas. Estas primitivas permiten describir una gran variedad de formas con un número reducido de parámetros, lo que facilita su interpretación y posterior uso en tareas de planificación. El procedimiento realiza una segmentación inicial y, para cada clúster, ajusta una única supercuadrática utilizando todos los puntos del mismo. En nuestro trabajo posterior (Menendez et al., 2024b), empleamos estas representaciones para calcular posiciones de agarre sobre los objetos detectados y planificar el movimiento del robot para ejecutar la manipulación correspondiente. No obstante, este enfoque presenta limitaciones importantes: cada objeto se representa mediante una única forma convexa, lo que impide modelar adecuadamente geometrías complejas o compuestas. Además depende en gran medida de que cada clúster esté correctamente segmentado y libre de ruido, ya que el ajuste se realiza utilizando la totalidad de sus puntos.

Estas limitaciones pueden abordarse utilizando el enfoque probabilístico propuesto por (Liu et al., 2022), que reformula el ajuste de supercuadráticas como un proceso de estimación y maximización (EM) basado en la verosimilitud de los puntos observados dada la primitiva definida por sus parámetros. Este método modela explícitamente la presencia de ruido y puntos atípicos mediante una mezcla Gaussiana-uniforme sobre la superficie paramétrica, permitiendo una asignación probabilística entre inliers y outliers y mejora la robustez frente a datos ruidosos o incompletos. Además, su algoritmo puede aplicarse secuencialmente para representar objetos complejos como combinaciones de supercuadráticas. No obstante, su propuesta se limita al uso de formas cerradas y no contempla restricciones físicas que eviten configuraciones inconsistentes con el entorno.

En este trabajo presentamos un método para el ajuste secuencial de primitivas geométricas que permite representar objetos complejos, incluso en presencia de oclusiones parciales, a partir de nubes de puntos obtenidas desde un único punto de vista. Frente a enfoques anteriores, proponemos una estrategia que combina superelipsoides (formas cerradas) y superparaboloides abiertos, permitiendo modelar tanto geometrías convexas como

cóncavas mediante la composición de múltiples formas por objeto. Este ajuste se formula dentro de un marco probabilístico, inspirado en el enfoque de (Liu et al., 2022), pero extendido con regularizaciones que tienen en cuenta las limitaciones de observación desde un único punto de vista, como la penalización por invasión del espacio vacío visible. Además, el sistema incorpora un mecanismo automático que selecciona, en cada etapa, si debe utilizar una forma abierta o cerrada según la evidencia observada en la nube de puntos.

2. Ajuste probabilístico de primitivas geométricas

Para hacer frente a las dificultades que implican las oclusiones y las observaciones parciales inherentes de trabajar con una nube de puntos capturada desde un único punto de vista, se formula el ajuste de primitivas como un problema de optimización probabilístico. En este trabajo inspirado por (Liu et al., 2022), se actualizan de forma iterativa los parámetros de forma, así como las probabilidades asociadas a cada punto de ser parte de la forma (inlier). En nuestro método se combina la optimización de una función de pérdida que combina la precisión geométrica con diferentes términos de regularización coherentes con su entorno físico.

2.1. Parametrización de Formas

Las primitivas geométricas permiten representar una amplia variedad de objetos utilizando pocos parámetros. En este trabajo empleamos dos tipos: superelipsoides para formas convexas y superparaboloides para cóncavas. Cada una se define en su sistema de coordenadas local mediante una función implícita con parámetros que determinan su geometría.

Superelipsoides. La ecuación (1) define su función implícita:

$$F(\mathbf{x}_i^{local}; \boldsymbol{\theta}) = \left(\left| \frac{x}{a_1} \right|^{\frac{2}{\epsilon_2}} + \left| \frac{y}{a_2} \right|^{\frac{2}{\epsilon_2}} \right)^{\frac{\epsilon_2}{\epsilon_1}} + \left| \frac{z}{a_3} \right|^{\frac{2}{\epsilon_1}} \quad (1)$$

donde $(x, y, z) = \mathbf{x}_i^{local}$ representa un punto en el sistema de coordenadas local. Los parámetros a_1, a_2, a_3 corresponden a los semiejes a lo largo de los ejes locales x, y y z , respectivamente. Los parámetros de forma ϵ_1 y ϵ_2 controlan la redondez de la superficie: ϵ_1 afecta la forma a lo largo del eje z , mientras que ϵ_2 lo hace en el plano ortogonal. Es importante destacar que ambos exponentes deben pertenecer al rango $(0, 2]$.

Superparaboloide con base desplazada. Este modelo permite representar formas abiertas que se ensanchan progresivamente con la altura. A diferencia del superparaboloide clásico, incluye un parámetro que permite definir un radio inicial distinto de cero en la base. La forma se describe mediante la función implícita dada en (2):

$$F(\mathbf{x}_i^{local}, \boldsymbol{\theta}) = \left(\left| \frac{x}{a_1 r(z)} \right|^{\frac{2}{\epsilon_2}} + \left| \frac{y}{a_2 r(z)} \right|^{\frac{2}{\epsilon_2}} \right)^{\frac{\epsilon_2}{\epsilon_1}}, \quad (2)$$

$$r(z) = r_0 + \left(\frac{z}{a_3} \right)^{1/\epsilon_1}$$

donde $(x, y, z) = \mathbf{x}_i^{local}$ representa un punto en el sistema de coordenadas local. Los parámetros a_1 y a_2 controlan el escalamiento en las direcciones x y y , respectivamente, mientras que

a_3 ajusta la escala en la dirección z . Los exponentes ϵ_1 y ϵ_2 determinan la curvatura: ϵ_1 afecta la forma en altura y también el crecimiento del radio con z , y ϵ_2 controla la forma en el plano xy . Finalmente, r_0 define el radio inicial en la base, es decir, el valor de $r(z)$ cuando $z = 0$. Todos los exponentes deben estar en el rango $(0, 2]$.

La función implícita $F(\mathbf{x}_i^{local}, \theta)$ permite determinar si un punto está sobre, dentro o fuera de la superficie de las primitivas. En concreto, se cumple que $F(\mathbf{x}_i^{local}, \theta)$ indica que el punto se encuentra exactamente sobre la superficie; si $F(\mathbf{x}_i^{local}, \theta) < 1$, el punto está dentro de la forma; y si $F(\mathbf{x}_i^{local}, \theta) > 1$, el punto está fuera. En el caso del superparaboloide, esta función solo está definida para valores de $z \geq 0$.

La parametrización permite generar una amplia variedad de geometrías ajustando los parámetros de tamaño (a_1, a_2, a_3) y de forma ϵ_1, ϵ_2 y r_0 , en el caso de los superparaboloides. La Figura 1 muestra ejemplos correspondientes a distintos parámetros para ambas familias: superelipsoides y superparaboloides con base desplazada.

Dado que la nube de puntos está definida en el sistema de referencia global y las funciones implícitas se evalúan en el sistema local de cada forma, transformamos los puntos al sistema de referencia local mediante una transformación rígida. Esta se define por una rotación (mediante los ángulos de Euler $\theta_z, \theta_y, \theta_x$) y una traslación (t_x, t_y, t_z) , y se aplica utilizando (3).

$$\mathbf{x}_i^{local} = \mathbf{R}^\top (\mathbf{x}_i^{global} - \mathbf{t}) \quad (3)$$

Cada primitiva se describe mediante sus parámetros geométricos y esta transformación. De esta forma, el vector de parámetros es $\theta = (\epsilon_1, \epsilon_2, a_1, a_2, a_3, \theta_z, \theta_y, \theta_x, t_x, t_y, t_z)$ para superelipsoides y añadiendo r_0 para los superparaboloides.

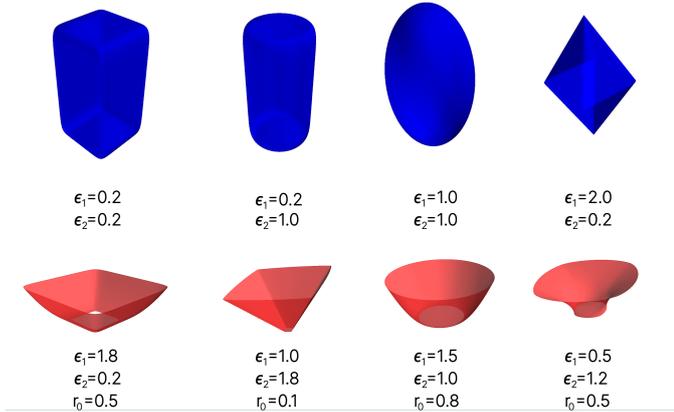


Figura 1: Superelipsoides (arriba) y superparaboloides con base desplazada (abajo) generados con distintos parámetros de forma, cuyos valores se indican bajo cada modelo.

2.2. Modelado probabilístico de la asignación de puntos

Para ajustar primitivas paramétricas, como superelipsoides y superparaboloides, a segmentos de nubes de puntos 3D, seguimos un enfoque probabilístico inspirado en el método presentado en (Liu et al., 2022). Este enfoque interpreta el ajuste de formas como un problema de clasificación probabilística, en el que cada punto de la nube está asociado a una variable latente

$z_i \in \{0, 1\}$ que indica si el punto es un inlier de la hipótesis de forma actual ($z_i = 1$) o un outlier ($z_i = 0$).

2.2.1. Modelado de la verosimilitud

Dada una parametrización de la forma θ , definimos una función de distancia o residual $d_i(\theta)$, que mide cuánto se desvía un punto $x_i \in \mathbb{R}^3$ de la superficie ideal de la forma. Este residual se calcula en el sistema local de la primitiva, tras aplicar la transformación rígida definida por θ .

Para tener en cuenta la varianza espacial y la incertidumbre del modelo, asumimos que las distancias de los puntos inlier siguen una distribución gaussiana isotrópica definida como (4).

$$p(x_i | z_i = 1, \theta) = \mathcal{N}(d_i(\theta); 0, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{3/2}} \exp\left(-\frac{d_i^2(\theta)}{2\sigma^2}\right) \quad (4)$$

Los outliers se modelan como distribuidos uniformemente dentro de un volumen de encuadre V siguiendo (5):

$$p(x_i | z_i = 0) = \frac{1}{V} \quad (5)$$

Dada una probabilidad a priori de pertenencia w , que refleja nuestra confianza inicial en que un punto x_i pertenezca a la forma dada por los parámetros θ . La probabilidad de observar x_i bajo el modelo, teniendo en cuenta ambas posibilidades (inlier u outlier), se define como (6).

$$p(x_i | \theta) = w \cdot \mathcal{N}(d_i(\theta); 0, \sigma^2) + (1 - w) \cdot \frac{1}{V} \quad (6)$$

2.2.2. Probabilidad posterior de pertenencia

Aplicando la regla de Bayes, la probabilidad posterior de que un punto x_i sea un inlier, dado su distancia a la forma, se calcula mediante (7).

$$p(z_i = 1 | x_i, \theta) = \frac{w \cdot \mathcal{N}(d_i(\theta); 0, \sigma^2)}{w \cdot \mathcal{N}(d_i(\theta); 0, \sigma^2) + (1 - w) \cdot \frac{1}{V}} \quad (7)$$

Esta probabilidad posterior regula la asignación probabilística de cada punto y permite que el modelo refine progresivamente qué puntos contribuyen al ajuste de la forma.

2.2.3. Función de distancia

El residual $d_i(\theta)$ mide la desviación del punto x_i con respecto a la superficie implícita definida por el modelo paramétrico, y se calcula mediante (8):

$$d_i(\theta) = \|x_i^{local}\| \cdot |F(x_i^{local}, \theta) - 1| \quad (8)$$

donde x_i^{local} es el punto expresado en el sistema de coordenadas local de la forma, y $F(\cdot)$ representa su función implícita definida en (1) y (2).

2.2.4. Estimación adaptativa de la varianza

La varianza σ^2 controla el grado de suavidad en la asignación de puntos como inliers. En lugar de fijar este valor, lo adaptamos durante la optimización en función de los residuales actuales y de las probabilidades posteriores de pertenencia usando (9).

$$\sigma^2 \leftarrow \frac{2 \sum_i p(z_i = 1 | x_i, \theta) \cdot d_i^2(\theta)}{3 \sum_i p(z_i = 1 | x_i, \theta)} \quad (9)$$

Esta actualización se obtiene al maximizar la función de verosimilitud logarítmica esperada de los datos completos bajo la suposición de distribución gaussiana, y refleja qué tan bien la forma actual explica los puntos considerados inliers.

2.3. Función objetivo y componentes de loss

La optimización de los parámetros de la forma θ se realiza minimizando una función de loss compuesta que combina precisión geométrica con restricciones físicas coherentes con el entorno. La función de loss global se define como (10).

$$\mathcal{L}_{\text{total}}(\theta) = \mathcal{L}_{\text{fit}}(\theta) + \lambda_{\text{free}}\mathcal{L}_{\text{free}}(\theta) + \lambda_{\text{drop}}\mathcal{L}_{\text{drop}}(\theta) + \lambda_{\text{extent}}\mathcal{L}_{\text{extent}}(\theta) \quad (10)$$

Los términos de loss aplicados dependen de la primitiva a optimizar:

- Para los superelipsoides, se utiliza la *data fitting loss* \mathcal{L}_{fit} y la *free-space loss* $\mathcal{L}_{\text{free}}$.
- Para los superparaboloides con base desplazada, se añaden la *drop loss* $\mathcal{L}_{\text{drop}}$ y la *extent loss* $\mathcal{L}_{\text{extent}}$.

Cada término cumple una función específica dentro del proceso de optimización:

Data Fitting Loss \mathcal{L}_{fit} . Este término favorece que la forma se ajuste con precisión a la nube de puntos observada. Se define según (11):

$$\mathcal{L}_{\text{fit}}(\theta) = \sum_i p_i \cdot d_i^2(\theta) \quad (11)$$

donde $p_i = p(z_i = 1 | x_i, \theta)$ es la probabilidad posterior de que el punto sea un *inlier*, y $d_i(\theta)$ es el residual entre el punto x_i y la superficie de la forma. Esta loss guía el modelo a ajustarse a los puntos relevantes observados, descartando gradualmente los outliers.

Free-Space Violation Loss $\mathcal{L}_{\text{free}}$. Para evitar que la forma invada regiones conocidas como vacías (el espacio entre la cámara y la superficie visible del objeto), se evalúa la función implícita $F(r_j; \theta)$ en los puntos de espacio libre r_j . Si estos puntos atraviesan la primitiva, es decir, si $F(r_j; \theta) < 1$, se penaliza el modelo actual utilizando la ecuación (12).

$$\mathcal{L}_{\text{free}}(\theta) = \frac{1}{M} \sum_{j=1}^M \sigma(-\beta \cdot (F(r_j; \theta) - 1)) \quad (12)$$

Donde σ es la función sigmoide y β es una constante que determina la pendiente de la transición. Los puntos r_j se obtienen muestreando a lo largo de M rayos que parten desde la posición de la cámara hacia puntos del entorno del clúster actual. Como estos puntos de la nube son visibles, se asume que no debe existir ninguna superficie entre la cámara y ellos.

Drop Loss $\mathcal{L}_{\text{drop}}$. En formas abiertas como los superparaboloides con base desplazada, que no tienen una parte inferior cerrada, se busca evitar que la forma se extienda innecesariamente hacia abajo. Para ello, la ecuación (13) penaliza la base de la forma si cae significativamente por debajo del punto más bajo del clúster observado.

$$\mathcal{L}_{\text{drop}}(\theta) = \left[z_{\text{mín}}^{\text{points}} - z_{\text{base}}(\theta) \right]_+^2 \quad (13)$$

Aquí, $z_{\text{mín}}^{\text{points}}$ es el valor mínimo en z de la nube de puntos observada, y $z_{\text{base}}(\theta)$ representa la altura de la base de la forma tras aplicar la transformación. Esta loss fomenta que la forma permanezca cerca de la parte inferior del objeto observado.

Extent Constraint $\mathcal{L}_{\text{extent}}$. Este término evita que la forma se extienda en exceso en dirección vertical, más allá de la variación de altura observada en la nube de puntos. La ecuación (14) penaliza el parámetro de escala vertical a_3 cuando supera el rango de datos observados en el eje z

$$\mathcal{L}_{\text{extent}}(\theta) = \left[a_3 - (z_{\text{máx}}^{\text{points}} - z_{\text{mín}}^{\text{points}}) \right]^2 \quad (14)$$

Aquí, $z_{\text{máx}}^{\text{points}}$ y $z_{\text{mín}}^{\text{points}}$ son los valores máximos y mínimos en z entre los puntos considerados inliers. Esta loss restringe la escala vertical de la forma para mantenerla coherente con los datos reales, evitando una elongación excesiva hacia arriba. Esto es especialmente relevante en los superparaboloides con base desplazada, donde el perfil se expande con la altura y una escala sin restricciones puede generar ajustes irreales o sobredimensionados.

2.4. Procedimiento de optimización

Se optimizan los parámetros de la forma θ minimizando la función de loss total, la cual equilibra la fidelidad a los datos observados con las restricciones físicas impuestas. En cada iteración, se alternan dos pasos:

- **(1) Actualización de inliers.** Se recalculan las probabilidades de pertenencia $p_i = p(z_i = 1 | x_i, \theta)$ en función de los residuales actuales $d_i(\theta)$, que se utilizan para ponderar el término de ajuste en la ecuación (7).
- **(2) Actualización de parámetros.** Se realiza un paso de descenso por gradiente para minimizar la función de loss total $\mathcal{L}_{\text{total}}(\theta)$ (10) con respecto a los parámetros θ .

Adicionalmente, cada N_σ iteraciones se actualiza la varianza del ruido σ^2 usando la ecuación (9), para reflejar la distribución actual de los residuales de los inliers. Esta estrategia evita inestabilidades por reescalados frecuentes y permite una convergencia más estable.

Para ilustrar cómo se desarrolla el procedimiento de optimización en la práctica, en la Figura 2 visualizamos el proceso de ajuste a lo largo de varias iteraciones. La figura muestra cómo la forma se alinea progresivamente con el objeto observado mientras se actualizan las probabilidades de pertenencia.

Al inicio, la forma está pobremente inicializada y muchos puntos reciben probabilidades intermedias debido a una varianza elevada. A medida que los parámetros se ajustan y la varianza se reduce de forma periódica, la forma se vuelve más compacta y se adapta mejor al objeto. Los puntos periféricos o ruidosos son atenuados mediante asignación probabilística, mientras que la estructura principal se refuerza. Esta dinámica ilustra las ventajas de combinar un modelado probabilístico con regularización geométrica.



Figura 2: Evolución del ajuste de forma durante la optimización. La estimación de la forma (en naranja) se ajusta progresivamente al objeto a medida que se refinan las probabilidades de pertenencia. Los puntos con alta probabilidad aparecen en amarillo; los outliers en púrpura oscuro.

3. Ajuste secuencial de múltiples formas

Tras introducir el método para ajustar una única forma a un clúster de puntos, describimos ahora nuestra estrategia para ajustar múltiples formas de manera secuencial dentro de cada región segmentada. El objetivo es descomponer cada clúster en subregiones que puedan representarse con precisión mediante un conjunto compacto de primitivas paramétricas. A partir de un clúster de nube de puntos obtenido mediante segmentación o agrupamiento, se ajustan formas de manera iterativa sobre el conjunto actual de puntos no asignados. En cada paso del proceso:

1. Se ajusta una forma (inicialmente un superelipsoide) al subconjunto actual de puntos utilizando el método probabilístico descrito en la sección anterior.
2. Se seleccionan como soporte de la forma aquellos puntos con alta probabilidad posterior de pertenencia, y el resto se pasa a la siguiente iteración.
3. Se calcula una penalización por espacio libre mediante rayos lanzados desde el centro de la cámara hacia los puntos inliers. Si esta penalización supera un umbral, se interpreta que el superelipsoide está invadiendo el espacio libre y que la forma podría estar mejor representada por una geometría abierta.
4. En ese caso, se vuelve a ajustar la misma región de puntos usando un superparaboloide con base desplazada, lo que proporciona un mejor soporte para geometrías abiertas o cóncavas.

Es importante destacar que, independientemente de si el superparaboloide sustituye el ajuste inicial, el proceso secuencial

continúa sobre los puntos que no fueron explicados por el superelipsoide, preservando la coherencia del procedimiento de ajuste en cascada. Este procedimiento continúa hasta que el número de puntos no explicados cae por debajo de un umbral o se alcanza un número máximo de ajustes. De este modo, los objetos complejos pueden descomponerse progresivamente en primitivas geométricas.

Un ejemplo de este proceso se muestra en la Figura 3, donde una copa de vino es reconstruida utilizando tres formas. El primer ajuste con un superelipsoide produce una penalización elevada por espacio libre, por lo que se sustituye por un superparaboloide con base desplazada. La base y el tallo restantes se modelan luego con dos superelipsoides adicionales.

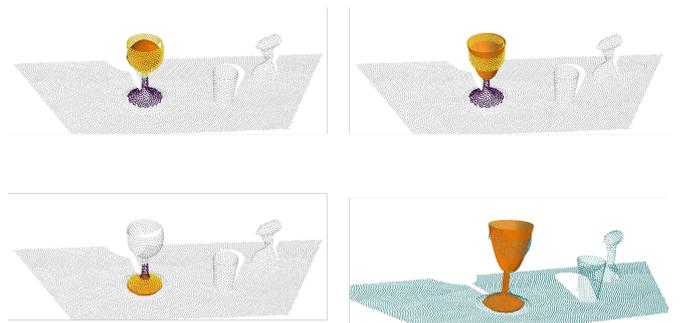


Figura 3: Ajuste secuencial de una copa de vino. La primera imagen muestra un ajuste inicial con un superelipsoide, que incurre en una alta penalización por espacio libre. La segunda imagen lo sustituye por un superparaboloide con base desplazada para modelar la parte superior abierta. En la tercera imagen, un superelipsoide ajusta los puntos restantes que forman la base. Finalmente, otro superelipsoide representa el tallo fino, dando lugar a una reconstrucción completa con tres primitivas.

4. Experimentos y Resultados

Evaluamos nuestro enfoque de forma cualitativa comparándolo con el método EMS propuesto en (Liu et al., 2022). La evaluación se centra en qué medida cada método recupera la geometría subyacente de objetos cotidianos colocados sobre una mesa, capturados desde una única vista RGB-D. El objetivo es analizar la robustez e interpretabilidad de las formas reconstruidas, especialmente en presencia de oclusiones parciales y estructuras complejas.

La escena evaluada incluye una selección de objetos representativos: recipientes abiertos (como un cuenco y un vaso), formas simples (una botella o una lata), y un objeto compuesto y estructuralmente más complejo como un taladro. Todos los objetos están dispuestos sobre una mesa y capturados desde una vista fija, simulando un escenario típico de percepción en robótica asistencial. La nube de puntos se procesa mediante un algoritmo de segmentación por proximidad, generando un clúster por objeto.

Nuestro método está implementado en PyTorch y utilizando el optimizador Adam, con una tasa de aprendizaje de $lr = 10^{-3}$ y un $weight\ decay = 0,001$. La probabilidad a priori de pertenencia se fija en $w = 0,1$, y la varianza σ^2 se actualiza cada $N_\sigma = 10$ iteraciones. Los distintos términos de la función de loss se ponderan de la siguiente manera: $\lambda_{fit} = 1$, $\lambda_{free} = 2$, $\lambda_{drop} = 10$, y $\lambda_{extent} = 20$.

Como baseline, utilizamos la implementación oficial de EMS publicada por los autores. Ajustamos el preprocesamiento

de entrada y la segmentación para que coincidieran con los mismos clústeres de nubes de puntos usados por nuestro método. El método EMS se limita a ajustar superelipsoides sin incorporar ninguna forma de regularización. Su configuración emplea una probabilidad $w = 0,1$, un máximo de 20 iteraciones EM, una tolerancia de convergencia de 10^{-3} , tolerancia relativa de 10^{-2} , y tres iteraciones del optimizador, según lo especificado en su repositorio.

La Figura 4 muestra una escena representativa que ilustra cómo nuestro método maneja de forma más precisa objetos abiertos y compuestos, evitando extensiones no plausibles que aparecen en la baseline por falta de regularización.



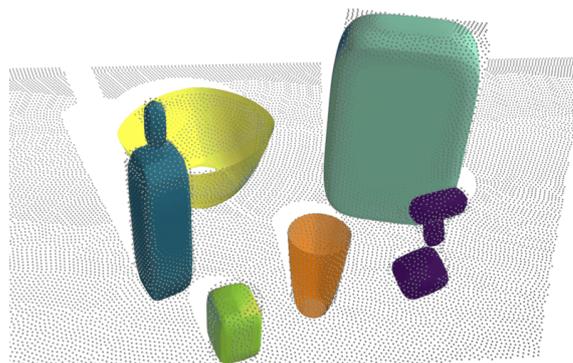
Figura 4: Nube de puntos de la escena utilizada para comparar el cualitativamente de los métodos de reconstrucción. La imagen muestra varios objetos cotidianos colocados sobre una mesa, incluyendo recipientes abiertos, objetos regulares y objetos de forma compleja.

La Figura 5 muestra la comparación visual con EMS. En ella se observa que nuestro enfoque genera reconstrucciones más fieles al combinar secuencialmente formas abiertas y cerradas, lo que permite representar una mayor variedad de geometrías.

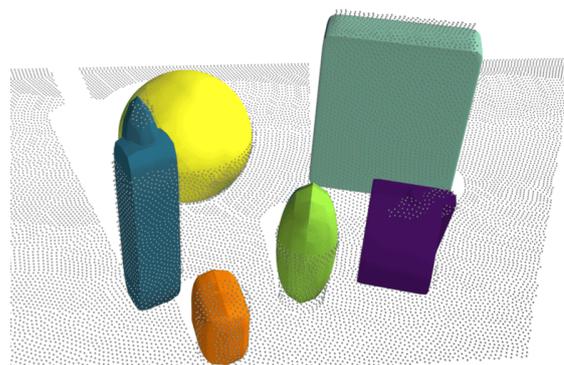
En el caso del cuenco y el vaso, EMS recurre a formas cerradas que invaden el espacio libre, mientras que nuestro método emplea superparaboloides abiertos para modelar la concavidad interna, logrando un ajuste más fidedigno. La botella, que puede aproximarse con formas cerradas, es bien representada por ambos métodos utilizando dos superelipsoides. En cambio, con la lata frontal, EMS produce una forma que se extiende hacia regiones vacías, mientras que nuestra propuesta penaliza estas invasiones y genera una reconstrucción más contenida.

Respecto al taladro, EMS ajusta dos elipsoides que no reflejan su estructura real y se extienden en el espacio vacío observado. Nuestro sistema, en cambio, lo descompone en tres primitivas bien alineadas que capturan de forma interpretable su empuñadura, cuerpo y base.

En conjunto, la capacidad de alternar entre primitivas abiertas y cerradas, junto con restricciones espaciales, permite generar reconstrucciones más realistas y ajustadas a la evidencia observada, especialmente en entornos con oclusiones y formas complejas.



(a) Reconstrucción de formas con nuestro método (superparaboloides + superelipsoides).



(b) Reconstrucción de formas con EMS (solo superelipsoides).

Figura 5: Comparación cualitativa entre nuestro método y EMS. Nuestro enfoque representa tanto formas abiertas como cerradas con mayor fidelidad, gracias al uso de restricciones específicas y múltiples primitivas paramétricas.

Agradecimientos

La investigación que ha conducido a estos resultados ha recibido financiación del proyecto iRoboCity2030-CM, Robótica inteligente para ciudades sostenibles (TEC-2024/TEC-62), financiado por Programas de Actividades I+D en tecnologías de la Comunidad de Madrid, y del proyecto SIROCO, con referencia PID2023-147343OB-I00, financiado por MICIU/AEI.

Referencias

- Biederman, I., 1987. Recognition-by-components: a theory of human image understanding. *Psychological review* 94 (2), 115.
- Liu, W., Wu, Y., Ruan, S., Chirikjian, G. S., 2022. Robust and accurate superquadratic recovery: A probabilistic approach. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 2676–2685.
- Liu, Y.-H., 1999. Qualitative test and force optimization of 3-d frictional form-closure grasps using linear programming. *IEEE Transactions on Robotics and Automation* 15 (1), 163–173.
- Menendez, E., Martínez, S., Díaz-de María, F., Balaguer, C., 2024a. Integrating egocentric and robotic vision for object identification using siamese networks and superquadratic estimations in partial occlusion scenarios. *Bio-mimetics* 9 (2), 100.
- Menendez, E., Martínez, S., Díaz-de María, F., Balaguer, C., 2024b. Selección y agarre robótico de objetos basada en el seguimiento de la mirada. In: *Actas del Simposio de Robótica, Bioingeniería y Visión por Computador: Badajoz*, 29 a 31 de mayo de 2024. Servicio de Publicaciones, pp. 127–132.
- Murali, A., Mousavian, A., Eppner, C., Paxton, C., Fox, D., 2020. 6-dof grasping for target-driven object manipulation in clutter. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 6232–6238.