1 Supplementary material

Explanation of Camera Sphere Parameters In order to facilitate the understanding of the mathematical concepts behind our sphere-based camera view angles generation method defined in Equations 1 and 2, we provide a detailed explanation of the underlaying parameters:

Equation 1: Number of Latitude Segments (N^{ϕ}) and Circles per Segment ($N_{\text{circ}}^{(i)}$)

- r_{circ}: The radius of the circles used for camera viewpoints, which influences the spacing and overlap of the viewpoints.
- r_{sph}: The radius of the sphere on which the camera viewpoints are distributed, i.e. the overall size of the area being considered for viewpoint placement.
- N^{ϕ}: The number of latitude floors on the sphere, which is calculated based on the ratio of the circle radius (r_{circ}) to the sphere radius (r_{sph}). N^{ϕ} is used to divide the sphere into horizontal segments, where each segment's height is determined in such a way that it roughly matches the diameters of the circles ($2r_{\text{circ}}$).
- $\lfloor \cdot \rfloor$: The floor function, which rounds down to the nearest integer, and ensures that the number of segments and circles per segment are whole numbers.
- $N_{\text{circ}}^{(i)}$: The number of circles in the i-th latitude segment, based on the circumference of the sphere at latitude θ_i divided by the diameter of the circles, ensuring even spacing of the viewpoints.
- θ_i: The colatitude angle for the i-th latitude segment, with $\theta_i \in \{0, \dots, \pi\}$

Equation 2: Spherical Camera Viewpoint Coordinates

- $(\theta_i, \phi_j^{(i)})$: The spherical coordinates that define the camera viewpoints. θ_i is defined above. $\phi_j^{(i)} = j \cdot \frac{2\pi}{N_{\text{circ}}^{(i)}}$ is the azimuthal angle for the j-th circle within the i-th latitude segment. The term $\frac{2\pi}{N_{\text{circ}}^{(i)}}$ ensures an even distribution around the latitude circle.

Grasp generation and rendering pipeline configuration Table 1 outlines the different parameters our frame generation pipeline utilizes, which influence the diversity and quantity of frames produced:

Using the variable notations of Table 1, the frame generation in our dataset can be expressed as in Algorithm 1 below:

Table 1: Frame generation configuration

Parameter	Description	Values
$(\Theta_{Vol})_j$	k-th ($k \in [1, .., N]$) probe rotation Euler angles, used within GrabNet	Manually selected set of three Euler angles $[\alpha, \beta, \gamma \vert \alpha, \beta, \gamma \in [0, ..., 360]]$
$(G_{[\Theta_{Vol}]})_k$	k-th manually selected GrabNet-based grasp out of all grasps generated by Θ_{Vol}	11 manually selected plausible grasps via visual inspection in MeshLab
r_{sph}	Sphere radius	0.8m
r_{circ}	Sphere surface circle radius	0.15m
Θ_k	k-th ($k \in [1, .., 92]$) camera view point Euler angles that depend on r_{sph} and r_{circ}	92 - 2 excluded = 90 values. Concrete Euler angles can be derived from applying the sphere concept described in section 2.2
z_{ego}	Camera distance (egocentric head to hand distance)	$[0.5, 0.8]$
$\mathcal{G}_{bkgr}$	Background image for rendered grasp frames	1 x plain white, 1 x consultation room, 3 x SPACE-FAN phantom, 3 x real pregnant mother belly (white / brown / black)
H_{rgb}	RGB values of glove texture	$\begin{bmatrix} 0.5647058824 \\ 0.5921568627 \\ 0.768627451 \end{bmatrix}, \begin{bmatrix} 0.38039215686 \\ 0.61960784314 \\ 0.8666666667 \end{bmatrix}$
A_{rgb}	RGB values of the arm	$[1.0, 0.6784313725, 0.3764705882]$

Algorithm 1 Combinatorics for frame generation based on a selected grasp

Input: Selected grasp $(G_{[\Theta_{Vol}]})_k$, egocentric viewpoint z_{ego}, background image $\mathcal{G}_{bkgr}$, hand texture H_{rgb}, camera view Euler angles Θ_k, sphere radius r_{sph}, circle radius r_{circ}
Output: Set of frames F

```
Initialize F ← ∅
for all (G_[Θ_Vol])_k, z_ego, G_bkgr, H_rgb, Θ_k do
    frame ← GenerateFrame((G_[Θ_Vol])_k, z_ego, G_bkgr, H_rgb, Θ_k, r_sph, r_circ)
    F ← F ∪ {frame}
end for
```