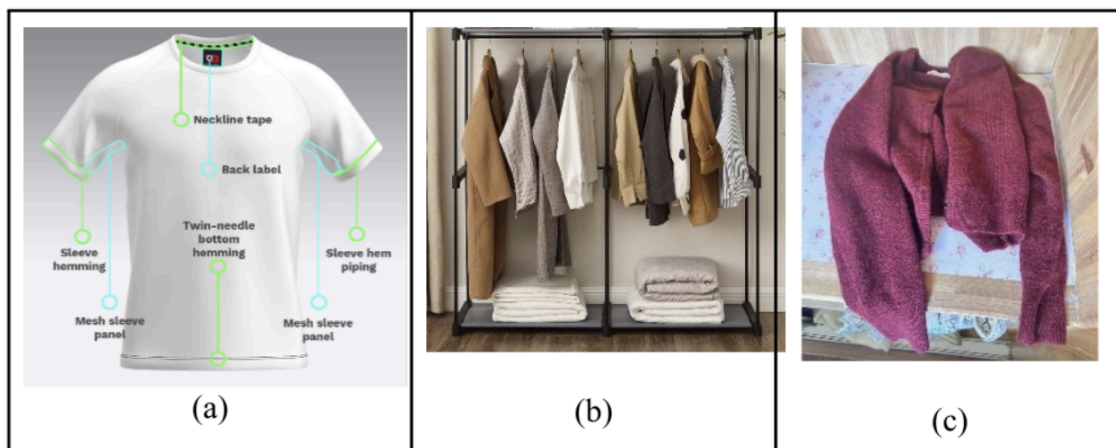


## Internship Positions on Garment Semantic Understanding

@Imagine Team of LIRIS lab, Ecole Centrale de Lyon

### Garment Semantic Understanding from Images

Recent statistics indicate an increasing demand for non-medical assistance for the elderly in the next years to come. Service robots that can help with the daily life chores of the elderly will prove to be quite useful. Additionally, such non-medical assistance could be beneficial for individuals recovering from a severe injury. In such a scenario, one can easily imagine a robo-assistant performing household chores dealing with garments such as fetching, folding, arranging clothes and even assisting with dressing or undressing the person in need of care. This makes the visual understanding and reasoning on the garment under manipulation an important research problem to solve.



The figure above gives an insight into the scenarios that we want to address. While fetching a garment from a well-arranged closet, one encounters topologically clean views, though the extent of visual obstruction may vary, primarily due to occlusion by other garments (Figure 1(b)). On the other hand, when dealing with crumpled clothes (Figure 1(c)), one encounters views that exhibit significant (self-)occlusions tightly coupled with the topological complexity of the garments, making it very challenging to interpret them. We will focus on these challenging scenarios of semantic understanding of garments from crumpled states or non-canonical forms. In the non-canonical form, the semantic segmentation is expected to be difficult. However, the topological comprehensible form allows the existence of some semantic labels that can be identified. To develop a complete comprehension of the garment semantics, following approaches would be studied.

Internship1: To recover semantic labels using self-supervised learning of garments in non-canonical states

Relying on the partial comprehension of semantic labels available in the non-canonical garment states from the images, we will build a self-supervised approach to deform the

garment from a canonical template to the configuration observed in the image. This representation will be improved by adapting our self-supervised approach of garment draping [1] with template-based reconstruction techniques [2] to deform the current representation such that it fits the garment seen in the images upon projection and restricting the identified garment semantic labels in T2 as almost fixed points.

Internship2: To recover semantic labels of garments in non-canonical states using generative AI

To identify the underlying canonical shape of the garment, it is important to estimate the deformations of the garment that has led to the current state, as done in [2]. This way, the occluded and invisible regions can also be substantially tracked. However this is a tedious process, requiring hours of computations on GPU. Contrary to this, GarSeM will seek a compact, and conditionable shape evolution for 4D garment models by combining self-supervised cloth simulation with generative models that have attempted to tackle similar yet different problems [3]. In addition, we aim to overcome the current limit of generative models that exhibit limited representation capabilities [4] and improve them so as to generate intricate details that are important for realistic garment shapes.

[1] R. Chen, L. Chen and S. Parashar. GAPS: Geometry-Aware, Physics-Based, Self-Supervised Neural Garment Draping. In 3DV, 2024.

[2] Kairanda N, Tretschk E, Elgharib M, Theobalt C, Golyanik V. f-sft: Shape-from-template with a physics-based deformation model. In IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022.

[3] K. Zou, S. Faisan, B. Yu, S. Valette, H. Seo, 4D Facial Expression Diffusion Model. To appear, ACM Trans. Multimedia Computing, Communications, and Applications 2025.

[4] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B., High-Resolution Image Synthesis with Latent Diffusion Models. In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021.

The project will be jointly supervised by Shaifali Parashar ([shaifali.parashar@cnrs.fr](mailto:shaifali.parashar@cnrs.fr)) and Prof. Liming Chen ([liming.chen@ec-lyon.fr](mailto:liming.chen@ec-lyon.fr)). Interested students should drop an email with CV and transcript.

Requirements:

1. Strong background in computer vision, machine learning and mathematics
2. Strong programming skills in C++ and python
3. Fluency in English

Project duration: 6 months

Location: Ecole Centrale de Lyon, France