BagIt! An Adaptive Dual-Arm Manipulation of Fabric Bags for Object Bagging

Peng Zhou[1], Jiaming Qi[2], Hongmin Wu[3], Chen Wang[4], Yizhou Chen[4], Zeqing Zhang[4]

***Abstract*— Bagging tasks, commonly found in industrial scenarios, are challenging considering deformable bags' complicated and unpredictable nature. This paper presents an automated bagging system from the proposed adaptive Structure-of-Interest (SOI) manipulation strategy for dual robot arms. The system dynamically adjusts its actions based on real-time visual feedback, removing the need for pre-existing knowledge of bag properties. Our framework incorporates Gaussian Mixture Models (GMM) for estimating SOI states, optimization techniques for SOI generation, motion planning via Constrained Bidirectional Rapidly-exploring Random Tree (CBiRRT), and dual-arm coordination using Model Predictive Control (MPC). Extensive experiments validate the capability of our system to perform precise and robust bagging across various objects, showcasing its adaptability. This work offers a new solution for robotic deformable object manipulation (DOM), particularly in automated bagging tasks. Video of this work is available at `https://youtu.be/6JWjCOeTGiQ`.**

I. Introduction

The primary challenges in DOM include the need for precise control, adaptability to varying material properties, and real-time responsiveness to complicated changes in the fabric state. In this paper, we employ a constraint-aware SOI planning framework to envelop 3D objects with a deformable fabric bag, as shown in Fig. 1. Our method draws on the concept of the Region of Interest (ROI) in image processing, indicating that fully estimating the state of a manipulated deformable object is not crucial in DOM. Specifically, in the bagging task, the edge of a fabric bag acts as the SOI. By focusing on the state estimation of this edge, the robotic system successfully carries out the bagging operation. Utilizing dual-arm operations, we represent the SOI as ellipses and introduce a two-stage manipulation strategy based on the object's bottom shape. An MPC-based controller ensures both arms accurately follow the planned trajectory. Our experiments demonstrate that in our framework, it is adequate to employ the SOI extracted merely from the bag opening rim to represent the whole fabric states during the bagging task. Furthermore, the presented system employs two robot arms working together, directed by an advanced planning framework that considers the object's structural limitations and the intended final configuration. Utilizing 3D-printed connectors, the robots achieve exceptional precision and stability in manipulating the bag.

Main contributions:

[1] The Great Bay University, Guangdong, China.
[2] Northeast Forestry University, Heilongjiang, China.
[3] Guangdong Academy of Sciences, Guangdong, China.
[4] The University of Hong Kong, Hong Kong, China.

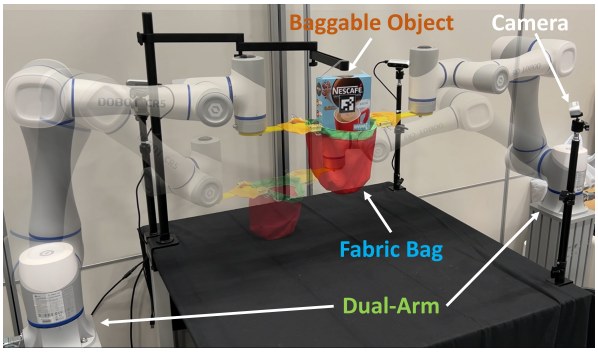

Fig. 1. The dual-arm system grasps the handles of a fabric bag to control the SOI for the bagging operation.

- We propose a robotic framework for automated bagging tasks, covering SOI estimation, bagging SOI generation, SOI planning, and a local planner for a dual-arm system.
- We introduce a constraint-aware planning strategy for SOI, enabling two robotic arms to successfully handle bags over various objects and achieve desired bag states.
- We integrate a vision-based control system that operates without prior knowledge of bag material properties, improving flexibility and adaptability in practice.

II. Problem Statement

The goal of the bagging task is to manipulate the handles of a fabric bag and wrap it around the bottom of the object $\mathcal{B}$, as shown in Fig. 2. The bottom of $\mathcal{B}$ is represented by a set of vertices $\mathcal{V} = \{\mathbf{v}_i\}_{i=1}^{n_v}$. We assume that all points in $\mathcal{V}$ lie on the same plane, e.g., the gray surface under the apple in Fig. 2. We consider that the bag's initial shape should remain slightly open rather than fully closed. Furthermore, we assume that all bags involved in this task are of a type that humans can manage in some way.

In this paper, we propose that a comprehensive estimation of the entire fabric bag is not requisite for the task at hand. Rather, our approach emphasizes the real-time identification and analysis of the SOI, which, in this context, is delineated as the opening rim of the bag. The set composed of points that make up SOI at time t is denoted by $\mathcal{Q}_t = \{\mathbf{q}_{i,t}\}_{i=1}^{n_x}$. Furthermore, the $\mathcal{Q}_t$ can be extracted from the raw point clouds $\mathcal{P}_t = \{\mathbf{p}_{i,t}\}_{i=1}^{n_p}$ via our proposed extraction method. For example, in Fig. 2, in the beginning, blue dots refer to point clouds $\mathcal{P}_0$, and the red curve is formed by points of SOI $\mathcal{Q}_0$. Here $\mathbf{v}_i, \mathbf{q}_{i,t}, \mathbf{p}_{i,t} \in \mathbb{R}^3$ are all 3D point expressed in the world frame $\mathcal{F}_w$, as shown in Fig. 2. And ${}^m\mathbf{q}_{i,t}$ indicates the point $\mathbf{q}_{i,t}$ expressed in a frame $\mathcal{F}_m$. Note that unless otherwise noted, all points are expressed in $\mathcal{F}_w$ and will be omitted superscript ${}^w(\cdot)$ for simplicity. In this way, the bag

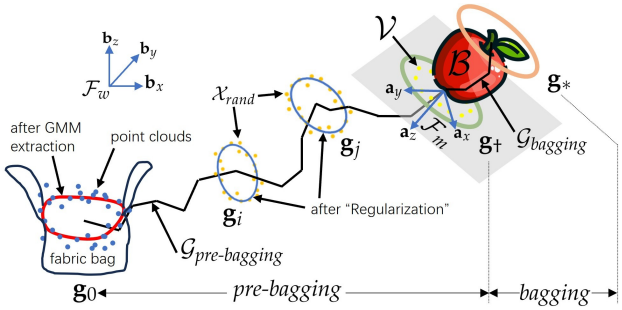

Fig. 2. Sketch of the proposed two-stage bagging strategy.

state $\mathbf{x}_t$ is defined by its SOI points, i.e.,:

$$\mathbf{x}_t = \left[\mathbf{q}_{1,t}^{\intercal}, \ldots, \mathbf{q}_{n_x,t}^{\intercal}\right]^{\intercal} \in \mathbb{R}^{3n_x}. \tag{1}$$

So, in the following, the bag state is equivalent to the SOI.

We partition the bagging task into two distinct phases based on whether the target object starts to enter the bag, as depicted in Fig. 2. This critical criterion is defined as the bagging SOI, denoted by $\mathbf{g}_\dagger$. During the packing process, $\mathbf{g}_\dagger$ naturally needs to satisfy two requirements. First, the size of the bag opening must be larger than the maximum cross-section of the object (i.e., size encircled by $\mathcal{V}$) to ensure that the object can smoothly pass through the opening and enter the bag. Second, the size of the bag opening cannot be arbitrarily large due to the physical constraints of the actual bag. Therefore, we assume that the size of the bag opening remains relatively unchanged during the bagging process, staying close to its initial size (i.e., encircled size of $\mathcal{Q}_0$) while satisfying the first condition.

From $\mathbf{g}_\dagger$, the goal SOI $\mathbf{g}_*$, where the bag fully warps the object, could be determined accordingly. Intuitively, the bag's initial SOI is given as $\mathbf{g}_0 := \mathbf{x}_0$, which can be specified manually before bagging. Based on $(\mathbf{g}_0, \mathbf{g}_\dagger, \mathbf{g}_*)$, the bagging process can be separated as *pre-bagging* and *bagging* stages, respectively. Then we employ a sampling-based motion planner to generate subgoals $\mathbf{g}_i$ in each stage, as shown in Fig. 2, and the collision-free bagging path can be formulated by:

$$\mathcal{G} := \{\mathcal{G}_{pre\text{-}bagging}, \mathcal{G}_{bagging}\} = \{\underbrace{\mathbf{g}_0, \mathbf{g}_1, \ldots,}_{\text{pre-bagging}} \underbrace{\mathbf{g}_\dagger, \ldots, \mathbf{g}_*}_{\text{bagging}}\}. \tag{2}$$

Given the planned path $\mathcal{G}$, an MPC-based shape servoing method is utilized to generate the velocity commands $\mathbf{u}_t$ from the robotic action space $\mathcal{A}$ based on the current SOI state $\mathbf{x}_t$ and the corresponding next subgoal $\mathbf{g}_{next} \in \mathcal{G}$, i.e.,

$$\mathbf{u}_t = \underset{\mathbf{u}_t \in \mathcal{A}}{\arg\min}\ \mathbb{E}(\mathbf{x}_t, \mathbf{g}_{next}), \tag{3}$$

where $\mathbb{E}$ is a measurable error function. Thus, the bagging task can be completed via a sequence of command velocities $\mathcal{U} = \{\mathbf{u}_t\}_{t=0}^{T}$ until $\mathbf{x}_t$ meets the goal SOI $\mathbf{g}_*$.

From a technical perspective, the bagging strategy presented in this paper is shown in Fig. 3. Firstly, the SOI extraction (Fig. 3-(b)) from visual perception ((Fig. 3-(a))) gives the states of the bag rim in a real-time manner. Then it will generate the bagging SOI according to the object bottom (Fig. 3-(c)), followed by the goal SOI generation (Fig. 3-(d)). The SOI planning algorithm (Fig. 3-(e)) will provide the available and collision-free path from the initial state to the bagging SOI and finally achieve the goal SOI. After that, the MPC controller (Fig. 3-(f)) is employed to control dual manipulators to execute the bagging task following the planned path (Fig. 3-(g)).

III. Experiments

Fig. 1 shows the experimental setup. Two CR5 robotic arms are equipped with custom 3D-printed mounts to grasp each handle of the bag, secured with zip ties to prevent slippage. A D455 depth camera (640x480) is fixed in an eye-to-hand configuration for top-down observation.

A. Dual-arm Bagging Manipulation

Dual-arm bagging experiments assess the effectiveness of the proposed manipulation method. Four distinct baggable objects are used in Exp. 1-4: a coffee box, canned pineapple, a grapefruit, and a 3D-printed triangular prism. The procedure involves the dual-CR5 robotic system guiding the bag from different initial positions along the pre-bagging trajectory $\mathcal{G}_{pre\text{-}bagging}$ to the intermediate configuration $\mathbf{g}_\dagger$, followed by bagging path along $\mathcal{G}_{bagging}$ to reach the final target shape $\mathbf{g}_*$ and complete the operation. To evaluate the performance, we compare two motion planning algorithms, FFG-RRT [1] and TS-RRT [2], alongside two control strategies, IBVS [3] and SSVS [4].

Fig. 5 shows the bagging process from Exp. 1-4. For a quantitative evaluation, we introduce three key metrics: planning success rate, planning duration, and manipulation success rate, reflecting the effectiveness of the planning and control algorithms. A detailed comparison is provided in Table I. The planning success rate indicates that CBiRRT outperforms other methods, achieving high reliability with reasonable computational cost. In contrast, FFG-RRT excels in speed, offering the shortest planning time due to its direct forward exploration strategy, while CBiRRT's bidirectional approach emphasizes stability, resulting in more search steps. Regarding manipulation success, the MPC clearly delivers the best results, surpassing the other control strategies. This is mainly because conventional shape servoing assumes a static desired shape, while the bagging task requires tracking evolving deformation trajectories. MPC's predictive capabilities align well with this dynamic need, enhancing tracking stability. These findings highlight MPC's suitability for complex robotic manipulation tasks. Additionally, $\mathbf{g}_\dagger$ serves as an intermediate buffer, segmenting the bagging process into pre-bagging and bagging phases, which enhances the robustness of the manipulation and contributes to a higher success rate.

B. Validations on Different Bags and Objects

To further validate the effectiveness of our method, we introduced additional bags and numerous objects with different poses. Here we used two bag types: one with color-contrasting red edges (Exp. 5-7) and another in solid color (Exp. 8-10). The results in Fig. 4 show that the proposed method remains effective in new task environments. Table I also provides a detailed comparison between our method and baselines. Exp. 5-7 shows a higher planning success

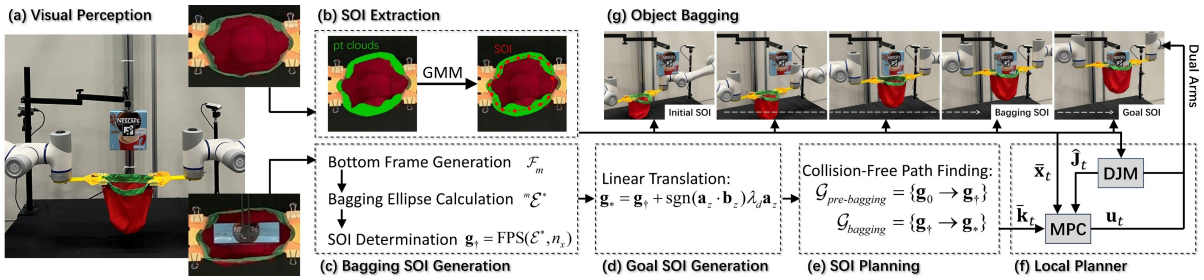

Fig. 3. Flowchart of our dual-arm manipulation strategy for the bagging task.

TABLE I

EXPERIMENTS RESULTS (S.R.: SUCCESS RATE. MANIP.: MANIPULATION)

Method	Coffee box (Exp 1)			Canned pineapple (Exp 2)			Grapefruit (Exp 3)			Triangular prism (Exp 4)			Tea caddy (Exp 5)		
	Planning S.R.	Planning time (s)	Manip. S.R.	Planning S.R.	Planning time (s)	Manip. S.R.	Planning S.R.	Planning time (s)	Manip. S.R.	Planning S.R.	Planning time (s)	Manip. S.R.	Planning S.R.	Planning time (s)	Manip. S.R.
FFG-RRT [1]	6/10	3.87 ± 1.97	8/8	8/10	2.37 ± 0.87	8/8	7/10	3.89 ± 1.18	8/8	6/10	3.58 ± 1.11	8/8	8/10	4.93 ± 1.23	7/8
TS-RRT [2]	7/10	6.32 ± 1.08	8/8	8/10	5.58 ± 1.13	8/8	9/10	6.85 ± 0.56	8/8	7/10	7.32 ± 1.34	8/8	8/10	7.46 ± 0.97	8/8
IBVS [3]	-	-	4/8	-	-	7/8	-	-	5/8	-	-	6/8	-	-	7/8
SSVS [4]	-	-	5/8	-	-	7/8	-	-	6/8	-	-	7/8	-	-	8/8
Ours	9/10	5.13 ± 1.26	8/8	10/10	4.21 ± 0.98	8/8	10/10	4.98 ± 1.93	8/8	9/10	5.32 ±1.56	8/8	10/10	6.16 ± 1.58	8/8

Method	Work bin (Exp 6)			Tea bucket (Exp 7)			Plastic packing (Exp 8)			Tea bucket (Exp 9)			3D-printed cuboid (Exp 10)		
	Planning S.R.	Planning time (s)	Manip. S.R.	Planning S.R.	Planning time (s)	Manip. S.R.	Planning S.R.	Planning time (s)	Manip. S.R.	Planning S.R.	Planning time (s)	Manip. S.R.	Planning S.R.	Planning time (s)	Manip. S.R.
FFG-RRT [1]	7/10	3.88 ± 0.98	8/8	8/10	3.93 ± 1.06	8/8	7/10	3.67 ± 1.66	6/8	8/10	2.26 ± 0.97	7/8	7/10	3.76 ± 1.03	6/8
TS-RRT [2]	8/10	6.13 ± 1.21	8/8	9/10	5.98 ± 0.72	8/8	7/10	5.96 ± 1.23	7/8	8/10	5.12 ± 0.98	7/8	9/10	5.15 ± 0.83	7/8
IBVS [3]	-	-	7/8	-	-	7/8	-	-	5/8	-	-	5/8	-	-	6/8
SSVS [4]	-	-	7/8	-	-	7/8	-	-	6/8	-	-	6/8	-	-	6/8
Ours	9/10	4.92 ± 0.96	8/8	10/10	5.06 ± 1.36	8/8	8/10	5.68 ± 1.29	7/8	9/10	4.06 ± 1.13	8/8	10/10	5.31 ± 1.48	7/8

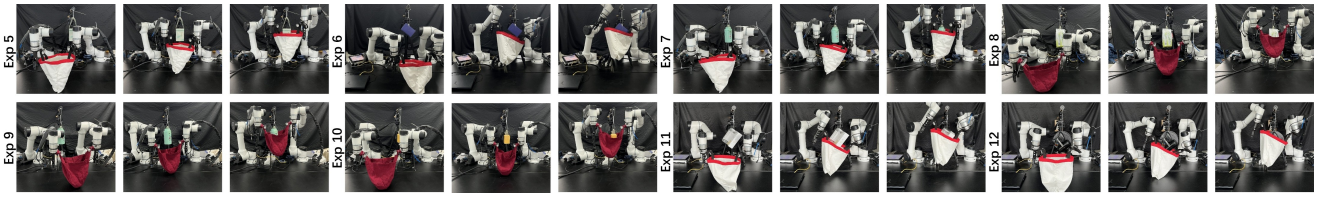

Fig. 4. Dual-arm manipulation using a red-edged bag and a solid-color bag for single (Exp. 5-10) and bound (Exp. 11-12) objects.

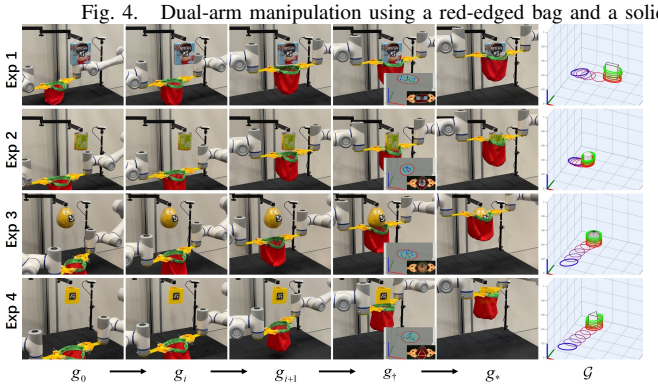

Fig. 5. Dual-arm bagging manipulation. Insets in the 4-th column show the determined bagging SOIs. The last column displays the planned SOI trajectories. The animation of the bagging process can be seen in the video.

rate than Exp. 8-10, likely due to larger bag dimensions. Notably, our method maintains the highest success rate among all approaches. In experiments with smaller objects, all three methods exhibited similar planning times, though our approach was slightly faster. The larger bags in Exp. 5-7 correspond to higher manipulation success rates, attributed to two factors: (1) the wider bag openings facilitate easier expansion, and (2) occlusion issues during manipulation for the solid-color bags stem from the neural network's detection accuracy limitations. In addition, Exp. 11 and 12 demonstrate the bagging operations for bound objects. From Fig. 4, the dual-arm system precisely handles the bag into the bottom of the bound objects (2 objects bundled in Exp. 11 and 3 objects bundled in Exp. 12) and completely wraps around them. Notably, in Exp. 6, 11, and 12, the bag successfully envelops the tilted objects along their bottom axis. In summary, our system has been validated to accurately and robustly complete various bagging tasks.

C. Comparison with Baselines

We replicate three dual-arm manipulation approaches, i.e., *ShakingBot* [5], *AutoBag* [6], and *BimaManip.* [7] for the task in Exp. 1. The first two use action primitives, while *BimaManip.* and this study employ SOI-based shape servoing. Since the initial state of bags in our paper is already open, we have ignored the shaking operations originally employed in [5] and [6]. Results are shown in Table II, where **Error** measures the distance between the SOI center and the object's center after task completion, indicating alignment accuracy. From Table II, *Ours* has the highest success rate and lowest error, indicating superior reliability and accuracy. Since *ShakingBot* and *AutoBag* use motion primitive-based action encoding for faster manipulation, *AutoBag* achieves the shortest time. In contrast, this study focuses on tracking transition shapes from CBiRRT, resulting in the slowest speed but the lowest error. *BimaManip.* offers balanced performance with moderate time and accuracy. Overall, *Ours* shows the most robust and precise bagging performance.

TABLE II

BAGGING COMPARISON ACROSS MULTIPLE APPROACHES.

	Manip. S.R.	Manip. Time (s)	Error (cm)
ShakingBot [5]	4/8	39.86 ± 3.74	6.47 ± 1.13
AutoBag [6]	6/8	32.54 ± 4.11	5.36 ± 0.87
BimaManip. [7]	7/8	44.21 ± 5.81	3.50 ± 0.93
Ours	8/8	53.09 ± 2.92	2.10 ± 0.77

REFERENCES

[1] O. Roussel, M. Taïx, and T. Bretl, "Motion planning for a deformable linear object," in *European workshop on deformable object manipulation*, 2014, pp. 153–158.

[2] C. Suh, T. T. Um, B. Kim, H. Noh, M. Kim, and F. C. Park, "Tangent space rrt: A randomized planning algorithm on constraint manifolds," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 4968–4973.

[3] X. Ren, H. Li, and Y. Li, "Image-based visual servoing control of robot manipulators using hybrid algorithm with feature constraints," *IEEE Access*, vol. 8, pp. 223 495–223 508, 2020.

[4] M. Hao and Z. Sun, "A universal state-space approach to uncalibrated model-free visual servoing," *IEEE/ASME Transactions on Mechatronics*, vol. 17, no. 5, pp. 833–846, 2011.

[5] N. Gu, Z. Zhang, R. He, and L. Yu, "Shakingbot: dynamic manipulation for bagging," *Robotica*, vol. 42, no. 3, pp. 775–791, 2024.

[6] L. Y. Chen, B. Shi, D. Seita, R. Cheng, T. Kollar, D. Held, and K. Goldberg, "Autobag: Learning to open plastic bags and insert objects," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 3918–3925.

[7] P. Zhou, P. Zheng, J. Qi, C. Li, H.-Y. Lee, Y. Pan, C. Yang, D. Navarro-Alarcon, and J. Pan, "Bimanual deformable bag manipulation using a structure-of-interest based neural dynamics model," *IEEE/ASME transactions on mechatronics*, 2024.