Intelligent Soft Surgical Robots for Next-Generation Minimally Invasive Surgery

Jiaqi Zhu, Liangxiong Lyu, Yi Xu, Huageng Liang, Xiaoping Zhang, Han Ding, and Zhigang Wu

Endowed with the expected visions for future surgery, minimally invasive surgery (MIS) has become one of the most rapid developing areas in modern surgery. Soft robotics, which originates from interdisciplinary advances in materials, fabrication, and electronics, featuring better adaptability and safer interaction, holds great promises in addressing current technical challenges in MIS, which are difficult to be solved with current rigid robotic technologies. For the first time, herein, the expected characteristics of next-generation MIS from the surgeons’ perspectives are analyzed and the recent progress of soft surgical instruments from three different aspects is comprehensively summarized: engineering design, fabrication techniques, and human–robot interaction. Perspectives of next-generation soft surgical robots are then discussed, where some exciting possibilities are emphasized. It is believed that further developments of intelligent soft robotics enable the next-generation MIS to agiley navigate to the target and conduct dexterous diagnostic or therapeutic procedures without any trade-offs in invasiveness and ultimately be a propitious solution for future surgery.

1. Introduction

Minimally invasive surgery (MIS) has received extensive attention in recent years, which represents a clear trend to further enhance the capability of surgical treatment, while considering the surgical experiences of patients and surgeons. To reach better surgical performance, MIS often pursues smaller, less incisions or even noninvasive, safer, more effective and faster surgical procedures, as well as more user-friendly manipulation. The huge advantages of MIS over open surgery are indisputable, but its further development still faces many challenges.

The main challenge lies in the surgical instruments (usually endoscopes) which are the core elements of MIS. Most surgical instruments currently used in clinics are rigid and have obvious application specificity. Although rigid structures mean higher precision, their limited flexibility and relatively large diameter are the main obstacles for MIS to further expand operation space and reduce trauma during the operations.

It is worth noting that some newly proposed MIS surgical instruments (most of which are still in the research phase) embody a gradual softening trend. This trend can bring better flexibility, biocompatibility, and operational possibilities, but it also means new technical problems and challenges. Considering the immature technical foundation and interdisciplinary characteristics of soft robots, the development of next-generation MIS technology is likely to require collaborative efforts of researchers from totally different fields, e.g., surgery, biomedical science, materials science, robotics, and other engineering sciences. This Review targets to provide a frame of soft surgical robots for MIS and encourages the researchers from diverse fields to discover new/better solutions.

This Review first gives a brief historical overview of MIS. Then, we identify important characteristics expected by the next generation of MIS through analysis from the perspective of surgeons. Then, we brief the characteristics of existing technologies and analyze the possible future trends of soft robotics technologies for MIS from three perspectives: engineering design, fabrication techniques, and human–robot interaction, as shown in Figure 1. Considering the new possibilities that soft materials bring to MIS, at the end of this Review, we try to envision new features of next-generation soft surgical robots from the following four aspects: new design concepts, new structural paradigms, new functions/characteristics, and new applications.

1.1. Brief Historical Overview

The history of MIS can be dated back to the pioneering work of the gynecologist Professor Kurt Semm in 1980s on conducting
therapeutic surgery using laparoscopes, which were traditionally only used for diagnostics before that.[8] Differing from the traditional open surgery, MIS usually uses elongated rigid or flexible surgical instruments to reach the target anatomy through single/multiple incisions or natural orifices within the body or on the skin and conduct surgical operations.[9,10] MIS can be divided into four categories based on the intervention approaches: extraluminal, intraluminal, transluminal, and hybrid,[11] as shown in Figure 2. Extraluminal procedures[12] intervene through one or more skin incisions, whereas intraluminal[13] and transluminal (NOTES)[14] procedures intervene into the body through natural orifices, e.g., esophagus, anus, vagina, and urethra. Transluminal procedures are an expansion of intraluminal ones, which can be further penetrated into the body cavities through a controllable incision on the luminal wall. Hybrid procedures[15] serve as a combination of the aforementioned intervention methods, mainly for some special operations that require multidevice collaboration.

For a clearer comparison, some key characteristics of the first three intervention methods are shown in Figure 2, respectively. As shown in Figure 2, the instrument shaft of extraluminal procedures is usually rigid, and one or more soft/rigid tools can be directly inserted to target points through the internal cannulas for surgical operations. Two typical examples are laparoscopy and thoracoscopy, through which various operations including genitourinary, hepatopancreaticobiliary, gynecologic procedures, and foreign bodies removal can be conducted. In contrast, intraluminal and transluminal procedures have to navigate in the tubular anatomical structures; thus, flexible and soft endoscopes are often involved. Intraluminal procedures can be used for the diagnosis and therapeutic procedures of hollow organs, e.g., tissue imaging, biopsy and excision, thrombus removal, and laser lithotripsy, whereas transluminal procedures are more inclined to partially replace the extraluminal procedures to further reduce the invasiveness.

Compared with open surgeries, MIS shows several significant advantages: smaller surgical trauma, higher safety, shorter postoperative recovery time, less sequelae and pain, and improved cosmesis.[16,17] The number of publications for MIS surgical instruments have shown an upward trend during the past 10 years, as shown in Figure 3. In 2020, more than 10,700 related achievements (papers and patents) have been published (Google Scholar search results, accessed 13.02.21; 9:30 a.m.).[18] After decades of development, MIS has become a widespread paradigm in the field of surgical medicine and has been widely used in many fields such as endoscopy, spine surgery, and percutaneous needle and neurosurgery.[8] Figure 4 shows an overview of the evolution of surgical instruments and platforms for MIS.

MIS workspaces are often dynamically changing, unstructured, narrow and fragile, and have poor visibility during surgery.[17] The rigid structures of existing MIS surgical instruments have poor compliance and flexibility. Some long and hard surgical instruments, e.g., laparoscopy, even have a fulcrum effect, which will magnify the surgeon’s hand shaking and induce impaired hand–eye coordination.[19,20] These factors all make surgeons face greater difficulty in operating the surgical

Figure 1. Technologies of soft surgical robots for MIS: engineering design, fabrication techniques, and human–robot interaction.

Figure 2. An illustration of three typical intervention methods: extraluminal, intraluminal, and transluminal based on the example of abdominal surgery. Their respective operative workspaces are shown in blue.[15]
instruments directly by their hands, and the corresponding problems include increased risk, low patient comfort, high surgeon labor intensity, long learning cycle, and low operating precision.[11]

The introduction of a robot-assisted platform has significantly alleviated these problems. The robot-assisted platform is an intermediary equipment between surgeons and MIS surgical instruments, through which surgeons can directly manipulate the remote operation console to control the movement of modified surgical instruments within the patient’s body. During this process, the robot system brings accuracy, stability, and flexibility (larger rotation angle than that of our wrists) to surgical operation, and the surgeons’ perception ability is restored to a certain extent by receiving information, e.g., force and position, from the sensing system.[21] The remote console provides an ergonomic interface, allowing surgeons to conduct surgical operations in a comfortable way with a shortened learning cycle and reduced labor intensity.[22]

Robot-assisted platforms fully release the capabilities of existing surgical instruments. However, MIS seems to encounter a bottleneck when trying to further expand its capabilities. The root cause is the superficial accessibility of surgical instruments. The intrinsic impedance mismatch between rigid surgical robots and soft human body makes it difficult to navigate in tortuous human tracts. The introduction of robot-assisted platform only partially alleviated this issue, and rigid surgical instruments still cannot provide access to all anatomy.[23]

In recent years, the emerging soft robot technology has gained great attention due to its intrinsic advantages in excellent compliance, infinite degrees of freedom (DOFs), adjustable stiffness, and low cost.[24–26] Soft robots are mainly composed of soft materials with high compliance and easy access, allowing multiple actuation mechanisms and having a rich selection of fabrication techniques. Considering the medical applicability of soft robots, introducing them into MIS is likely to generate many exciting opportunities, especially for intraluminal and transluminal procedures. Several recently published reviews have confirmed the importance of soft robot technology for MIS,[2,5] and Runciman et al. gave a short summary of recent research on MIS soft devices.[27]

Immature technical systems and many challenges are main obstacles restricting soft robots’ further popularization and massive application of MIS. Just to name a few here, the obstacles include low output force, poor controllability/predictability of
behavior, and limited adaptability of fabrication techniques.\textsuperscript{27,28} How to exploit their advantages, as well as break their limitations in medical usage, is on the central stage of soft robot research for next-generation MIS. Hence, this Review aims to give a comprehensive overview on soft surgical robots, discuss possible engineering design/implementation, and try to provide some new and meaningful perspectives/thoughts for the next-generation MIS soft robots.

1.2. Clinic Demands

This section will analyze the main expected characteristics of next-generation MIS from the perspective of surgeons.

As shown in Figure 5, hollow organs, e.g., urogenital system, gastrointestinal system, and pulmonary bronchial system, are directly connected to the ambient environment and usually suffer from high tumor incidence rate. For example, the global tumor incidence rates of lung cancer, gastrointestinal cancer, urinary tract tumor, and genital tract tumor rank first, second, fourth, and ninth in all malignant tumors.\textsuperscript{29} These organs generally are featured by the following. 1) They are often adjacent to important organs or large vessels, leading to a high risk of extraluminal procedures, e.g., percutaneous puncturing. 2) Their internal lumens can be utilized to conduct intraluminal or transluminal procedures. 3) However, all these lumens have a long, narrow, and tortuous hierarchy structure, which imposes demanding requirements on the accessibility of surgical robots. Consequently, in clinic practices, there are no standard effective diagnostic and therapeutic methods for the minimally invasive treatment of these targets. Besides the apparent flexibility and steerability, dimension also matters in developing surgical robots for these targets. To begin with, the length of the surgical robot should be long enough (usually 1.5–2 m), to make the reachable space fully cover the target.\textsuperscript{17} Moreover, the diameter of the surgical robot is also constrained. For example, minimally invasive procedures through oral and esophagus require a surgical robot’s feature size to be less than 30 mm in diameter,\textsuperscript{30} whereas intravascular procedures require it to be less than 3 mm.\textsuperscript{31} In a manner of speaking, further miniaturization often means "reach deeper and broader."

Apart from good accessibility, expected surgical instruments also have to work effectively after reaching the target, specifically including stability, precise movement or positioning performance, and sufficient force exertion.\textsuperscript{12} The lack of structural shaping capability makes it easy for current surgical instruments to deviate from the desired target under the influence of luminal boundaries floating or even the loss of luminal support.\textsuperscript{11} Such a limitation hinders effective implementation of transluminal procedures under current technical conditions. What’s more, a continuous and stable force exertion is also an important measure for smooth operation of surgery. All of these make it necessary for surgical instruments to own adjustable stiffness and achieve sufficient stiffness when needed. Introducing variable stiffness mechanism is one of the ideas to solve this dilemma, and related research progress will be discussed in the next section.

High cost, short durability, and cumbersome maintenance procedures of current surgical tools are also severe issues. Taking a soft ureteroscope (FURS) that is widely used in diagnosis and treatment of upper urinary tract as an example, although the risk of extraluminal procedure is avoided, reusable FURS is still fragile, with durability reports ranging from 3 to 159.\textsuperscript{33–36} According to the statistics, excluding the initial purchase cost, the amortized cost per use of FURS is $ 848.10.\textsuperscript{37} What’s more, almost all current surgical instruments (including FURS) are actuated by cables,\textsuperscript{3} and consequently they can only be replaced entirely when partly damaged. High costs and expenses have greatly hindered the popularity of MIS technology.\textsuperscript{38} New design paradigms and fabrication techniques are anticipated to replace the traditional counterparts.

![Figure 5. Complex anatomic structures in hollow organs, e.g., urogenital system, gastrointestinal system, and pulmonary bronchial system. Black dotted line represents common surgical paths to access the target anatomy.](image-url)
Some expectations for future reusable surgical instruments include lower costs, better durability, as well as an easier maintenance process, and the fabrication techniques are expected to be more efficient, more reliable, and have abundant optional materials. However, considering the further miniaturization of MIS surgical tools (slenderer and longer), improving the durability will be more difficult. Disposable devices or components may be an ideal solution to face this dilemma. The introduction of soft robot technology makes disposable surgical instruments possible, through which the durability and maintenance costs may no longer have to be considered. Moreover, disposable instruments also allow them to be tailored to the patient, which can further enhance the surgical effect. Despite these advantages, disposable surgical instruments for future MIS have to place more emphasis on the quality stability and efficiency of fabrication techniques compared with the reusable ones.

Excellent human–robot interaction capabilities are also very important for next-generation MIS, which can be further divided into patient–robot system and surgeon–robot system interaction. The robot system usually refers to a combination of robot-assisted platforms and surgical instruments, e.g., endoscopes. The interaction between patient and robot system is mainly reflected in the robot system collecting information, e.g., images, contact force, and temperature, from the patient with the help of a sensor system and then consciously making some adjustments, e.g., force control and posture change, and conducting surgical operations. Most of the current surgical instruments are constricted to image transmission and very limited tactile sensing, and thus the surgical operation heavily relies on the surgeon’s experience under the condition of perceptual separation from the operative site. This will undoubtedly increase the risk of surgery. Considering this issue, the next-generation MIS robot system is expected to offer comprehensive, high-quality, stable, and real-time information on the human body, in which advanced sensors play a key role, and can autonomously implement force/position control and compliance deformation to a certain extent. The interaction between surgeon and robot system is mainly reflected in the robot system processing information from patient and surgical instruments, e.g., bend angle and position, and transmitting them to surgeons; meanwhile, the surgeons then convey motion instructions to robot system through consoles. For surgeons, an ideal interaction should be featured by a humanized form of information transmission, ergonomic manipulation, and a certain degree of intelligence. In addition, as most of the existing robot-assisted platforms have a large footprint, high costs, and complex maintenance processes, correspondingly, miniaturized, lower-cost, and easier-to-maintain counterparts are preferred. The desired characteristics of next-generation MIS are shown in Figure 6.

2. Engineering Design

Engineering design directly determines the surgical capability of instruments. According to the earlier discussion, expected surgical instruments should provide enough accessibility (high flexibility, compliance, small size, and long enough length) and effectiveness (precise positioning, stability, and sufficient output force). Current rigid devices have good effectiveness, but the

Figure 6. Desired characteristics of next-generation MIS covering three aspects: operation capability, manufacturing and maintenance, as well as human–robot interaction.

| Operation capability | Manufacturing and maintenance | Human–Robot interaction |
|----------------------|-------------------------------|-------------------------|
| **Increased stiffness** | **Generality** | **Patient & Robot system** |
| Accessibility | Acceptable cost | High-quality, comprehensive, real-time, stable sensor information |
| Effectiveness | High automated degree | Autonomous force/position control |
| | Adaptability | Certain intelligence |
| | Various optional materials | Surgeon & Robot system |
| **Reduced stiffness** | | Humanized information transmission mode |
| | | Ergonomic manipulate mode |
| | | Information intelligent processing |
| **Reusable instruments** | | Robot-assisted platform characteristics |
| | | Small footprint |
| | | Low cost |
| | | Easier maintenance process |
accessibility is very limited.\cite{19} To obtain better accessibility, soft surgical instruments (SSIs) have been proposed, but their relative low effectiveness is becoming a severe issue.\cite{27} An inherent dilemma lies in that surgical instruments with good accessibility should have low stiffness, whereas good effectiveness requires high stiffness, as shown in Figure 7. Introducing variable stiffness mechanism is an excellent solution for such a contradiction. Notably, it’s unrealistic and unnecessary to completely replace the current rigid surgical instruments with their fully soft counterparts. Incorporating some soft robotic techniques into current instrument design for a balanced performance might be a pragmatic solution during instrumental evolution.

Surgical instruments for MIS represented by endoscopes usually have a slender body with a circular cross section, which sometimes contains working channels of different sizes. These working channels are used to deliver devices, e.g., forceps for physical intervention, catheter for perfusion or drainage, or a high-density energy-based scalpel such as laser to ablate tissues. Only a small portion of space is available for other components, e.g., soft actuators and variable stiffness units. Figure 7a shows a typical section view of state-of-the-art endoscopes with available space marked out. In this section, we will discuss the engineering design of SSIs based on the general configuration of endoscopes. It can be divided into the selection and design of both soft actuators (also the basis of manipulation) and variable stiffness components. Of course, engineering design not only affects the surgical capabilities, but also the corresponding fabrication techniques, which will be discussed later.

2.1. Soft Actuation

Actuation provides surgical instruments the capability of locomotion and manipulation. Traditional actuation methods (usually cable driven) rely on the relatively movement of rigid mechanical components and consequently lead to poor compliance.\cite{44,45} However, soft actuation mainly involves soft materials and achieves actuation by material/structural deformation. It opens up an opportunity for designers to achieve adequate soft and compact surgical instruments with enhanced accessibility.

Through reasonable selection and structural design, various instrumental motions, e.g., bending, twisting, elongation, and expanding, can be achieved. Of course, it is necessary to trade off among various characteristics of different actuation principles based on the concrete application and choose the most suitable one. A comparison of the existing and potential soft actuation methods for next-generation MIS is shown in Table 1.

2.1.1. Cable Driven

Cable driven is the most commonly used actuation method in current multijoint rigid surgical instruments,\cite{46–48} where cables are used as a force transmission media to transmit force and movement from an electrical motor or human hand to the places cable passes by. Due to its relatively mature technical foundation, cable driven has been widely used in many SSIs.\cite{49–51} Compared with the rigid ones, cable-driven soft instruments can usually return to their initial positions due to the natural elasticity. Figure 8a shows an example of such compliant instruments for cardiac ablation.\cite{52} The use of cable driven can achieve greater output force, and the material properties of cables can also allow quick response, light weight, and small footprint. However, uneven force loading caused by the friction between cable and soft material during actuating complicates the modeling and increases the control difficulty. Introducing a hollow hose (e.g., made from PVC) is a feasible solution; however, the instruments’ diameter will increase.\cite{52}

Future works on cable-driven SSIs may mainly focus on the miniaturization of the entire system and the precise control of force and position. An external motor is the main obstacle to the system miniaturization. A trade-off should be made between the output force and the size of auxiliary equipment. Replacing the motor with more compact soft actuators (e.g., shape memory alloy [SMA]) is an alternative solution.\cite{53} Furthermore, cable driven is expected to fully exert its unique advantages in force and accuracy in the future MIS, and feedback control with fused sensing is one of the most important trends. For instance, cable-driven soft manipulators using visual servo control can achieve precise tracking of the target point.\cite{54,55}
| Actuation mechanism          | Description                                                                 | Commonly used scale | Available strain | Common auxiliary equipment | Output stress | Response velocity | Available deformation | Other characteristics                                                                                       | References |
|-----------------------------|------------------------------------------------------------------------------|---------------------|------------------|---------------------------|--------------|-------------------|-----------------------|----------------------------------------------------------------------------------------------------------------|------------|
| Cable driven                | Pull and release cables to control deformation                               | mm-cm               | High             | Electromotor              | High (up to 150 Mpa) | Fast (< 0.1 s) | Bending, contraction | Pros: compact structure; lightweight Cons: easy to wear; difficulty in controlling force exertion                | [46-55,449] |
| Pneumatic driven            | Positive-pressure driven Adjust gas pressure to control chamber deformation    | μm-cm               | High (≈10–40%)   | Air pressure-control system | High          | Fast (≈0.05–1 s) | Bending, expansion, contraction | Pros: lightweight; high power density (>10–10^3 kW⋅m⁻³); Cons: diameter expansion of the device when actuated; nonlinear response; risk of leakage; difficulty in long-distance transmission; error caused by gas compressibility | [56-63]    |
| Vacuum driven               | Adjust vacuum degree to control chamber shrinkage                             | mm-cm               | Medium (≈10–40%) | Vacuum degree control system | High          | Fast (≈0.1–1 s) | Bending, contraction | Pros: no risk of leakage; high power density Cons: difficulty in long-distance transmission; relatively complex structure; error caused by gas compressibility | [66,67,71,304,415] |
| Hydraulic driven            | Adjust liquid pressure to control chamber deformation                         | μm-cm               | High (≈10–50%)   | Hydraulic control system  | High          | Medium (≈0.5–2 s) | Bending, expansion, contraction | Pros: high control accuracy; high power density Cons: diameter expansion of the device when actuated; risk of liquid leakage; nonlinear response | [64,65]    |
| Shape memory materials      | SMA Adjust the temperature to control phase change causing deformation         | μm-cm               | Wires: low (up to 5%) | Power supply (low voltage and high currents) | Wires: high (several hundreds of MPa) | Medium (≈0.2–2 s) | Bending, contraction | Pros: self-sensing; variable stiffness capacity; lightweight; high power density (>10^3–10^4 kW⋅m⁻³); low structural complexity Cons: low-energy conversion efficiency (>10%); hysteresis | [80,82,84,85,87–89] |
| SMP                         | Control the temperature-sensitive crystallographic change to realize deformation | μm-cm               | High (≈50–400%)  | External heaters | Low (recovery stress 1–3 Mpa) | Slow (≈5–60 s) | Bending | Pros: variable stiffness capacity; lightweight Cons: variable stiffness capacity; lightweight | [79–81,83] |
| Magnetic actuation          | External magnetic field controls the deformation of soft structures containing ferromagnetic particles | μm-cm               | Medium           | Magnetic field generator | Medium        | Fast (< 10 μs) | Bending | Pros: occupy less space; high control accuracy Cons: single form of motion | [94–104]    |
| Electrical actuation        | DEA Electrostatic force leads to compression deformation of elastomer materials | μm-cm               | High (≈1–1000%)  | Power supply (≤1–20 kV) | Medium (≤7 Mpa) | Fast (< 200 μs) | Bending, expansion | Pros: self-sensing, high electromechanical efficiency; high power density (>10^3–10^4 kW⋅m⁻³); lightweight Cons: risk of leakage currents and electrical breakdown | [111,112,115–117,119] |
| IPMC                        | Voltage promotes ion migration to achieve deformation                         | μm-cm               | Medium (≈1–10%)  | Power supply (≤1–5 V) | Low           | Slow (≤0.3 Mpa) | Bending | Pros: self-sensing; low-energy conversion efficiency (≈30%); low longevity; high costs | [113,114,118,120] |
| HASSEL                      | Electrostatic force leads to the transfer of intermediate dielectric fluid    | mm-cm               | High (≈1–60%)    | Power supply (≤5–20 kV) | Medium        | Fast (≈0.05–1 s) | Bending, expansion, contraction | Pros: self-sensing Cons: risk of high-voltage leakage; difficult to miniaturize | [121,122,124] |
Figure 8. Some existing/potential soft actuation methods for next-generation MIS. The depth of the background color of different methods represents their relevant technical foundations for their use in MIS as a soft continuum robot. a) Cable-driven soft manipulator for cardiac ablation. Reproduced with permission.\textsuperscript{[52]} Copyright 2013, IEEE. b) Pneumatic soft tentacles with 3D mobility, which can be used for liquid delivery, diagnosis, and organ manipulation in MIS. Reproduced with permission.\textsuperscript{[63]} Copyright 2013, Wiley-VCH. c) Vacuum-driven soft actuator module and soft continuum robot composed by it. Reproduced with permission.\textsuperscript{[67]} Copyright 2017, AAAS. d) Structure and mechanism of multidirectional-bending electric endoscope using SMA wires. Reproduced with permission.\textsuperscript{[84]} Copyright 2016, Springer Nature. e) Schematic illustration of the magnetically responsive tip of the ferromagnetic soft continuum robots. Reproduced with permission.\textsuperscript{[100]} Copyright 2019, AAAS. f) Actuation mechanism and demonstration of high-performance HASEL actuators for untethered soft robots. Reproduced with permission.\textsuperscript{[124]} Copyright 2019, Wiley-VCH. g) Tendril-like soft robot based on reversible osmotic actuation. Reproduced with permission.\textsuperscript{[118]} Copyright 2019, Nature Publishing Group.
2.1.2. Fluidic Actuation

Currently, fluidic actuation is one of the most widely used actuation methods in soft robotics, mainly because of easy implementation and low cost.\[96,37\] Based on the types of fluidic media, the fluidic actuation can be divided into two types: pneumatic driven and hydraulic driven. Both are actuated by regulating fluid pressure in soft chambers. To generate required movement or shape deformation in soft devices when pressurized or vacuumed, various methods, e.g., asymmetry in chamber distribution or material stiffness,\[58,59\] embedding of nonstretchable materials,\[60,61\] and actuating sequence regulation,\[49,62\] have been proposed. A typical example of pneumatic tentacles that can be used for liquid delivery, diagnosis, and organ manipulation in MIS is shown in Figure 8b.\[63\] Compared with the pneumatic ones, hydraulic surgical instruments generally exhibit greater output stress and better control accuracy,\[64,65\] but higher inertia and density of liquids may introduce extra errors and complicate control strategy. In addition, most of the existing fluidic SSIs are driven by positive pressure, and their common shortcomings are the radial expansions during actuating and the risk of fluid leakage. Vacuum driven (usually pneumatic) can avoid these problems\[66\] but is rarely used in MIS, mainly because the chambers of vacuum-based actuators usually have to be filled with support materials, e.g., sponge, to provide support and recovery force, which in turn increases structural complexity and further limits its miniaturization, as shown in Figure 8c.\[67\] Another issue may be its slower recovery speed, which severely hinders the dexterity of devices.

With relatively mature technology, fluid actuation has been becoming one of the mainstream choices for MIS SSIs. However, there are still many challenges for widespread applications, e.g., bulky and noisy pneumatic/hydraulic control systems, high control difficulty caused by nonlinear dynamics,\[57,59\] and development of fluidic actuators with higher energy density and lower radial expansion.\[17,58\] Recently, some attempts to miniaturize pneumatic/hydraulic control systems show promising results.\[68–70\] They are hopeful in increasing the integration and portability of fluidic-driven soft devices in the future MIS. Nonlinear responses and viscoelastic properties of soft materials are always the main obstacles for fluidic-driven soft robots to reach precise control, and hence the strategies using traditional model-based control only achieve very limited progress.\[71\] Feedback control using sensors is an ideal way to jump over these barriers, e.g., the works by Tapia\[72\] and Marchese.\[73\] Nevertheless, increasing the integration of soft devices without affecting the size requires more powerful fluidic actuators. Introducing nonstretchable materials, e.g., cloth\[74\] and bellows,\[75\] or changing the transmission medium (such as particle drive)\[76,77\] is a possible solution that has been proven effective in other applications. In addition, there are still some other technologies that can be used in future fluid-driven SSIs, e.g., self-healing materials and related actuators that are expected to avoid the risk of liquid leakage.\[78\]

2.1.3. Shape Memory Materials

Soft actuators based on shape memory alloy (SMA) and shape memory polymer (SMP) are special actuators that can return to their initial shape when subjected to certain stimulus (often temperature changes) and exhibit great stiffness variation during phase transition. The specific working principles and more technical details can be found in other reviews.\[79–83\] Although these kinds of actuators have been applied in numerous fields, e.g., soft grippers, biomedical devices, as well as aerospace, their applications in MIS are relatively limited until now.\[82,83\] The possible concerned issues may be the safety problems from heat dissipation and the relatively slow response of SMA/SMP. Compared with other actuation methods, SMA/SMP actuation can realize lighter, smaller, and simpler robotic systems similar to the cable-driven ones, as shown in Figure 8d.\[84,85\] SMA actuators usually exhibit greater recovery stress and can be directly actuated by electrical Joule heating, but SMP actuators have lower costs and biocompatibility, which usually need introducing external heaters.\[86\] It’s worthwhile to note that large stiffness variation and low recovery stress\[83\] of SMP make it more suitable as a variable stiffness component rather than an actuator. Therefore, most of the existing SSIs actuated by shape memory materials use SMA instead of SMP.\[84,85,87,88\]

SMA/SMP is hopeful to be further used in future MIS, so as to exploit its advantages in the miniaturization of robot systems. However, before that, intensive researches have to address some critical issues, e.g., high-temperature risk,\[89\] low response speed (especially cooling speed), and recovery error.\[90\] The first two can be alleviated by improving cooling efficiency. For instance, the introduction of cooling elements, e.g., circulating water, can be considered,\[90\] but it increases the structural complexity. The high-temperature risk can also be compromised by deploying an insulating layer that can protect the soft tissues from high-temperature damage. In addition, it has been proved that development of composite materials is helpful to improve the response speed and recovery error of SMA/SMP.\[91–93\]

2.1.4. Magnetic Actuation

As an emerging actuation method in soft robotics, magnetic actuation has gained extensive attention in recent years due to its unique capabilities of remote control and fast response (up to 100 Hz) in a confined space.\[94\] Compared with other mechanisms, magnetic actuation can easily build small-scale (down to submillimeter-scale) soft robots without considering the integration of actuation elements, and it holds great promise in diverse areas, especially in medical applications.\[95,96\] Typical magnetic soft actuators are made of soft matrix with magnetic fillers embedded. Under the stimulus of an external magnetic field, magnetic soft actuators can exhibit different shapes and deformations due to the reorientation of internal magnetic fillers.\[97–99\] A representative ferromagnetic soft device (outer diameter: 500 μm) proposed by Kim et al. can navigate freely in a complex cardiovascular network and conduct surgical treatment, as shown in Figure 8e.\[100\] They also introduced a hydrogel skin that can reduce the friction during navigation.\[101\] Another work presents a magnetically controlled soft microrobot attached to the top of a traditional guide wire.\[102\]

The major challenge faced by magnetically driven soft robots is the insufficient capability of magnetic field-generating equipment, which is mainly reflected in limited controllable space, large footprint, and huge power consumption.\[103,104\]
Future focus on performance improvement will heavily rely on the magnetic field-generating equipment, specifically including equipment miniaturization and increasing electromagnetic (EM) conversion efficiency. Tuning the properties, e.g., stiffness and magnetic permeability, of soft magnetic composites is another way to alleviate this problem, which in turn lowers their demands on magnetic field strength. In addition, it should be reminded that strong magnetic field may intervine or disturb the function of other electronic devices that already exist in the patient’s body, and it is preferable for surgeons to manipulate remotely due to the potential health issues in long-term exposure to strong magnetic fields.

2.1.5. Electrical Actuation

Another popular actuation method potentially used in future MIS is electrical actuation, which includes several different subtypes. All of them can be directly actuated electrically without any additional auxiliary equipment. Dielectric elastomer actuators (DEAs) and ionic polymer–metal composites (IPMCs) are two popular ones among them. Their actuation states or amplitudes can be flexibly controlled by modulating the waveform of a power supply signal, e.g., amplitude, frequency, and phase. DEAs are arranged in an electrode/dielectric sheet/electrode structure and actuated by the expansion of dielectric sheet, which is due to the electrostatic attracting force generated by a high-voltage direct-current (HVDC) signal applied to the pairing electrodes. IPMCs are driven by an ion-migration swelling mechanism and can exhibit a bending behavior under low-voltage stimulation (1–5 V). DEAs have been demonstrated in various fields, e.g., robotic gripping, crawling, and swimming, and IPMCs have also been used as soft actuators in various robotic applications, as shown in Figure 8g. However, their applications in MIS surgical instruments are rarely demonstrated. The possible reason is that DEAs require high actuating voltage (generally >1 kV) and often suffer from the risk of electrical breakdown, whereas IPMCs show small output stress and relatively slow response. Recent work on hydraulically amplified self-healing electrostatic (HASEL) actuators presents a new type of electrical soft actuator, which effectively addresses the issue of electrical breakdown and improves the output stress by replacing the dielectric sheet of DEAs with a dielectric fluid. HASEL actuators are like the combination of hydraulic actuators and DEAs, inheriting the large output stress and fluidity of the former and the direct electric drive of the latter, and are suitable for future MIS. Application of HASEL actuators has been demonstrated in a few fields, e.g., robotic gripping and artificial muscles, as shown in Figure 8f. We believe that HASEL may become an alternative actuation method for soft MIS surgical instruments in the future.

The future development of DEA-based soft devices for MIS has to eliminate the risk of leakage currents and electrical breakdown. Further research on IPMCs actuators might mainly focus on improving the response speed and output stress, but recent progress is very limited. Several researchers have tried to introduce closed-loop control for increasing the response speed, but the progress is very limited. The main obstacle hindering the application of HASEL actuators to surgical tools might be the manufacturing techniques for further miniaturization, specifically including heat sealing technology with higher resolution (will be discussed in Section 3), stable packaging technology of fluid medium at small scale, as well as integration technology of finer electrodes and tapes. Of course, possible liquid leakage is also one of the issues that has to be considered.

2.2. Variable Stiffness Mechanism

Variable stiffness is a powerful measure for the effective operation of SSIs. The close cooperation between variable stiffness elements and soft actuators is expected to well reproduce various surgical operations of the human hand in open surgery. Desired variable stiffness elements has to exert as little influence as possible when deactivated, while presenting sufficient stiffness variation and response speed when activated. It should be noted that the specific required value is difficult to give due to the diverse application scenarios and highly coupled stiffness values. Except this, there are many other factors that have to be traded off when selecting a suitable variable stiffness mechanism based on the specific demands, e.g., safety, biocompatibility, stiffening range, controllability, yield stress, auxiliary equipment, space occupation, and structural complexity. For convenience, Table 2 shows some characteristics of several existing and potential variable stiffness mechanisms for MIS. It should be emphasized that here we only introduce variable stiffness mechanisms, which are relatively common or, in our opinion, potential to be used in SSIs soon, and more variable stiffness mechanisms can be found in other studies.

2.2.1. Jamming

Jamming is a common variable stiffness mechanism widely used in soft robotics and has a relatively mature technical foundation. It mainly relies on the friction force between active units, e.g., rigid particles, elastic layer, and wire, generated by the external positive pressure to achieve relative position/structure locking and can obtain continuous stiffness changes within a certain range. Jamming based on rigid particles (granular jamming) usually exhibits better fluidity when deactivated and thus has less effect on the overall stiffness of soft devices, as shown in Figure 9a, whereas jamming based on elastic layers or wires is expected to obtain a smaller overall diameter of SSIs due to its low thickness structure, as shown in Figure 9b. The specific stiffness range of jamming is influenced by several different parameters, e.g., the membrane thickness, the size and shape of basic locking elements (particles, layers, and wires), as well as the structural configuration. Hence, it is important to design them properly according to the actual needs. Apart from the pressure difference activation method, active force application and the phase change of locking elements made by special materials can also be used as the stimuli of jamming. Compared with the pressure difference activation, the latter two can be convenient for the miniaturization of surgical instruments in certain applications but may lose some controllability.
Table 2. Comparison of various variable stiffness mechanisms.

| Variable stiffness mechanism | Description | Control method | Common auxiliary equipment | Response velocity | Stiffening range | Absolute stiffness | Number of stiffness states | Other characteristics | References |
|------------------------------|-------------|----------------|----------------------------|-------------------|------------------|---------------------|--------------------------|----------------------|-----------|
| Jamming-based stiffening     | Granular friction lock between rigid particles | Regulate the vacuum/apply force | Vacuum control system/soft actuator | Medium (≈0.1–1.1 s) | Low (≈1–20) | High | Continuous (external pressure) | Pros: increase the weight and size of instruments; prone to nonuniform stiffness distribution | [127,128,134,135] |
| Layer jamming               | Friction lock between planar material layers | Regulate the vacuum/apply force | Vacuum control system/soft actuator | Medium (≈0.1–1 s) | Low (≈1–10) | High | Continuous (external pressure) | Cons: influence the stiffness of instruments when deactivated; difficulty in miniaturization | [129,130,136] |
| Rheological property        | ERF Viscoelastic behavior under high electric fields | Regulate the electric field intensity (typically up to 5 kV·mm⁻¹) | Electric field generator | Fast (< 10 ms) | Medium (≈1–100) | Low (≈140 kPa) | Continuous (electric field intensity) | Pros: relatively low-energy consumption; less impact when deactivated | [141–143,146,148,149] |
| MRF Viscoelastic behavior in high magnetic fields | Regulate the magnetic field intensity (up to 500 mT) | Magnetic field generator | Fast (< 10 ms) | Medium (≈1–100) | Low (≈200 kPa) | Continuous (magnetic intensity) | Cons: relatively high-energy consumption | [144,145,147,149] |
| Phase transition            | SMA Temperature-controlled phase transition between martensite and austenite | Regulate the heating temperature | Power supply (low voltage and high currents) | Medium (≈0.2–2 s) | Low (≈1–10) | High (low-temperature phase: 10–83 GPa; high-temperature phase: 0.1–41 GPa) | 2 | Pros: relatively high absolute stiffness; lightweight | Cons: long cooling time | [80,82] |
|                            | SMP Phase transition under external stimulus (often temperature changes) | Regulate external stimulus (usually the heating temperature) | External heaters | Slow (≈5–60 s) | High (≈100–300) | High (low-temperature phase: 0.6–3 GPa; high-temperature phase: 0.1–10 GPa) | 2 | Pros: relatively high stiffness variation; lightweight | Cons: long cooling time | [79,81,150–154] |
|                            | LMPA Solid–liquid conversion under temperature changes | Regulate the heating temperature (typical melting temperature: 28–62 °C) | Power supply | Slow (melting time: 1–10 s; solidifying time: ≈30 s) | High (≈25–5000) | High (≈3–9 GPa) | 2 | Cons: slow cycle time | [157–161] |
|                            | LMPP Phase transition under temperature changes | Regulate the heating temperature | External heaters | Slow (usually takes a few seconds) | – | – | – | – | – | [162-165] |
|                            | Antagonistic arrangement | Multiple actuating units act in an antagonistic way | – | – | – | – | – | – | Pros: small occupied space | Cons: low controllability | [52,63,169–172,174] |
Jamming methods have relatively quick response (0.1–1 s) and dramatic stiffness changes. However, larger stiffness requires more active units, e.g., particles, layers, or wires, which mean increased volume. Also, more active units (especially layers and wires) will bring higher stiffness to surgical instruments when deactivated. Based on these dilemmas, future research on jamming-based variable stiffness mechanism may include promoting further structural miniaturization, eliminating the impact on instruments’ stiffness when deactivated, and achieving uniform distribution of stiffness when activated. Further subdividing the active units, e.g., finer particles and thinner layers, and increasing their mutual friction through increased surface roughness or introducing specific structures, e.g., barbs, wedges, might be a possible way to improve the performance of jamming mechanism.

2.2.2. Stimulus–Responsive Soft Materials

Besides jamming, phase transition or rheological property regulation achieved by external stimuli, e.g., magnetic field, electric field, and temperature change, can also be used for stiffness variation. According to the behind mechanisms, related materials can be divided into three categories: electrorheological fluid (ERF) and magnetorheological fluid (MRF), SMA and SMP, and low-melting-point alloys (LMPA) and low-melting-point polymers (LMPP). The stiffness modulation of ERF/MRF mainly depends on the transformation of its rheological characteristics under a strong magnetic/electric field, and its stiffness is proportional to the electric field/magnetic field strength within a certain range, as shown in Figure 9e. A detailed description
of the internal conversion mechanism can be found in various studies (ERF,[141–143] and MRF.[144,145]) ERF/MRF has been demonstrated in some robotic applications, e.g., rehabilitation robotics,[146] robotic joints,[147] and robotic fish.[148] as well as some medical applications, e.g., catheters and prosthesis penile.[44] MRF usually shows greater stiffness changes than ERF,[149] whereas ERF has lower energy consumption. As described in Section 2.1, SMA/SMP exhibits a stiffness change during its phase transition with shape memory effect and can also be used as a variable stiffness element, as shown in Figure 9c-d.[79–81,150] Compared with the large stiffness range of SMP (100–300),[150] the stiffness change of SMA is very low (1–10).[82] which makes it more suitable to be used as an actuator in MIS. The use of SMP as a variable stiffness component in MIS is rarely reported,[151] but related applications such as robotic fingers are relatively common.[152–154] LMPA/LMPP represents a class of materials that can be transformed between multiphase states with temperature changes.[155,156] LMPA has multiple types and is able to transit from solid to liquid at a relatively low temperature (usually 28–62 °C), as shown in Figure 9f.[157] Due to its considerable stiffness changes (25–9000),[158,159] LMPP has been used in various fields, e.g., SSIs[160] and robotic grippers,[161] and is expected to find broader applications in the future MIS. LMPP is a joint name of temperature-sensitive multiphase polymers, which has emerged as a variety of new materials with different characteristics in recent years.[162–164] For example, Zhou et al. proposed a complex multiphase organogelreadable with a precisely controllable thermo-induced step-wise switching, which can be used in adaptive grasping of soft grippers.[165] A common type of LMPP is thermoplastics, e.g., ABS, PLA, Nylon, PMMA, and PET, most of which are widely used in 3D printing. Minh et al. compared the properties of several thermoplastic materials, selected relatively superior PET (biocompatibility, high strength, high chemical resistance, and low cost) to construct variable stiffness components, and conducted ex vivo experiments in fresh pig tissues with a flexible manipulator.[166]

Slow response, especially cooling time, and high-temperature risk are two shared challenges of SMA/SMP and LMPA/LMPP.[161,167] Both of which rely on thermal excitation, and the possible solutions can refer to the relevant discussion in Section 2.1. Among them, the conductive SMA and LMPA can be directly heated by Joule heating, whereas SMP and LMPP usually need extra heaters. The high current (usually several amperes) required for electric heating may also cause safety hazards for use within the human body, so ensured insulation is essential. Although ERF/MRF has no problem in response speed (≤ 10 ms)[141] its dependence on external electric/magnetic field-generation equipment and relatively low absolute stiffness, e.g., hundreds of kPa,[141,168] is the obstacle that cannot be ignored in the further applications in SSIs.

2.2.3. Antagonistic Arrangement

Antagonistic arrangement is inspired by the bionic principle of trunk and octopus tentacles, which obtains overall or local stiffness enhancement by activating the muscle groups of different parts to fight against each other. Numerous combinations of soft actuators and variable stiffness elements (based on same or different principles) can be implemented in this mechanism,[169–172] and a more detailed summary can be found in other studies.[126,173] Among them, two common examples are fluidic actuators and cable-driven actuators. The former can achieve flexible regulation of local stiffness by actuating a set of mutually opposing chambers, but this process will affect the instruments’ diameters.[63,174] The latter can obtain relatively large stiffness changes by tensioning the cables in opposite positions but can only change the overall stiffness.[82]

Although an antagonist arrangement mechanism is possible to achieve integration of actuators and variable stiffness elements, low controllability and limited stiffness variation range still limit its further applications in MIS. What’s more, certain combinations are also accompanied by some negative effects. For example, antagonistic arrangements involving positive-pressure fluidic actuators often result in an expansion of instrument diameter. When multiple fluidic actuators are involved, this effect will even be magnified. Therefore, appropriate evaluations and trade-offs are necessary during the design process.

2.3. Summary

This section provides a brief summary of the reported potential soft actuation and variable stiffness principles for next-generation SSIs.

For soft actuation, the novel actuation mechanisms based on soft materials will greatly enrich the DOFs of surgical devices and are expected to subvert the original design paradigm. First, as mentioned earlier, the deformation of soft actuators will be extremely diverse through appropriate design, including but not limited to bending, twisting, elongation, expanding, shrinking, spiral winding, and folding, most of which are unimaginable for traditional rigid surgical instruments. These deformations can be coupled or independently controlled, resulting in some attractive features. Second, SSIs can completely get rid of the limitations of fixed joints and motors to achieve high-density, arbitrary-position soft actuator distribution. The size and deformation of different actuators can also be flexibly changed, and the upper limit of the DOFs will no longer be limited but entirely depends on the specific needs. Third, the excellent compliance of soft materials also brings countless passive DOFs, which are far more than the traditional multijoint rigid surgical devices, and thus can better fit the nonstructural complex environment within the human body to achieve high compliance and reliability.

For variable stiffness principles, variable stiffness elements can greatly expand the available stiffness range of soft surgical devices, and the upper limit of stiffness is even several times higher than that of current commercial rigid ones. An original function of the variable stiffness element is to assure that the SSI is stiff enough when needed to achieve precise positioning, stable support, and transfer force. What’s more, the variable stiffness elements can also be used to construct temporary bones, which can deactivate and activate at any time as needed, to improve the controllability and performance of the soft surgical devices. We can also imagine that the variable stiffness elements can cooperate with soft actuators to freely construct some handy
tools with specific shapes, e.g., hook shaped, to facilitate surgical operations.

For both soft actuation and variable stiffness principles, considering the fragile working environment of soft surgical devices in the human body, there are several issues that have to be seriously considered during design. The first is safety. Several most important safety concerns for soft surgical devices include fluid leakage, high-temperature risk, and electrical leakage. Measures to isolate heat and electricity should be done. Fluids, e.g., dielectric fluid for HASEL, fluidic actuation media, or potential fluids, e.g., LMPA or thermoplastic, should be carefully sealed and potential toxic substance leakage should be avoided. The second is biocompatibility. Biocompatible materials should be used or covered in the parts of SSIs that directly or possibly contact tissues to avoid any negative immune reaction and other side effects. The third is sterilization. SSIs should be designed to be easy to sterilize, or even self-sterilizing, which is especially compulsory for reusable ones.

Furthermore, system compactness is also important. Except magnetic actuation, all other principles have to occupy a certain degree of space in the instruments and hence require the actuator or variable stiffness elements to be as powerful or effective as possible in a relatively small volume. Optimization of single-element performance is one way to achieve such a goal. For example, the output force of the fluid actuators can be increased by improving the pressure-bearing capacity of chambers. Increasing the friction between particles can expand the variable stiffness range of particle jamming. Furthermore, the coordination between different components achieved by reasonable selection and design is another way to reduce the occupied space and obtain a compact system. For example, matching the fluid actuators with the jamming-based variable stiffness elements, through appropriate design, can make them share an air pressure control system, thereby reducing system complexity. Also, the cooperation of SMA as actuators and SMP as variable stiffness elements can achieve similar effects, as they have similar principles and stimulation methods. Notably, another point is that actuation and variable stiffness are often coupled for soft devices, because the actuation usually involves uneven distribution of material deformation, which will lead to increased stiffness. However, the stiffness variation range achieved by this is relatively limited, and it is difficult to control them independently because of the coupling relationship between stiffness and deformation. Antagonistic arrangement is a decoupling method. As discussed earlier, the antagonistic arrangement can realize the integration of actuators and variable stiffness elements, but its variable stiffness range is still limited. Nevertheless, it is still an attractive choice for some less-demanding occasions.

In the current stage, SSIs are more likely to be an integrated system of soft and rigid structures, as the existing soft robotic technologies are not necessary to replace all rigid components, e.g., cautery knife, clamp, and miniature camera. Therefore, seamless integration of rigid–soft structures is worth exploring. First, some methods for connecting rigid structures can be used, e.g., bolting, bonding, hinge, friction locking, buckle, and stuck connection. Among them, the compliance of soft materials allows more flexible stuck connection design. The soft and rigid structures can be embedded into each other through some protruding structures, e.g., bosses, which is impossible for two rigid structures. Second, soft materials also allow some completely new connection methods. For instance, the rigid components can be cast into soft structures during the fabrication process, e.g., SDM, which will be discussed in Section 3. Moreover, the soft structures can also be stitched together or even be interpenetrated/one-way permeable with the rigid structures. All of these connection methods can be used alone or simultaneously to enhance the rigid–soft coupling structures.

Different actuation or variable stiffness principles usually have their own characteristics. Good design has to comprehensively evaluate the actual requirements, make full use of the advantages of different principles, and overcome/minimize their disadvantages. To achieve a better overall effect, a combination of several different actuation or variable stiffness methods may be an option. As long as the size and structural complexity of the device are not significantly affected, an ingenious combination may be complementary to each other. For example, the soft manipulator proposed by Shiva et al. combines the pneumatic actuators with the cable actuators, which demonstrate the compliance of pneumatics and the control accuracy of cable drives. The posture and stiffness of the manipulator can also be controlled simultaneously through the cooperation of the two actuators. Based on the similar approach, the soft manipulator proposed by Stilli et al. has the capability to control stiffness and greatly shrink the volume. Some other studies, e.g., the combination of SMA and pneumatics also demonstrate the potential of the approach. To obtain better results, we encourage researchers to freely try any new potential solutions.

Figure 10 shows possible modules for a typical SSI system. Various fabrication techniques represent different ways to achieve predetermined structures or functions of such a system.
As stated in Section 1.2, an ideal manufacturing technology should trade off among various factors, e.g., cost, degree of automation, instrument durability, efficiency, adaptability, optional materials, and quality stability, based on the instrument attributes (reusable or disposable) and the actual demands. This section will discuss some existing and future possible fabrication techniques for MIS SSIs, and possible industrialization will be the focus during our analysis. It should be noted that we only list some common or promising fabrication techniques, and a more detailed summary can be found in other reviews.\textsuperscript{[180–183]}

Table 3 shows a brief comparison of various fabrication techniques.

3.1. Conformal Replication

A classic idea of shaping soft materials is to partially or entirely conformably deposit the liquid form precursor onto a specific shape before they are fully cured/solidified and then maintain such a state until the shape is finalized. During this process, the uncured soft materials are required to have a certain degree of fluidity, and an external force, e.g., gravity, centrifugal force, or wetting force, is often needed to drive the conformal progress. This idea is usually simple and effective, through which the morphological features, e.g., microstructure and texture, of the conformed surface can be accurately replicated by soft materials. In recent years, around this solution, researchers have explored various conformal replication techniques. We will mainly focus on the three of them: molding/casting, soft lithography, and coating methods.

3.1.1. Molding/Casting

Molding/casting is the most popular fabrication technology used in soft robot community due to its low equipment dependency and relatively mature technical foundation. It can be used for easy and fast iteration of prototypes with complex 3D shapes. Molding usually uses catalyst embedded prepolymer, e.g., mixed silicone precursors, which are poured into premade molds, heated or left for curing, and then taken out, as shown in Figure 11a. Several post-treatment processes, e.g., fiber reinforcement and bonding, are also required for the finished products. To fabricate more complex structures, various

| Fabrication process | Description | Efficiency | Resolution | Other characteristics | References |
|---------------------|-------------|------------|------------|-----------------------|------------|
| Conformal replication | Molding/casting | Pour liquid soft materials into the mold and cure by heating | Small batch: low | Medium (millimetre scale) | Low device dependency; limited available materials; easy to produce bubbles and other defects; capability to reproduce quickly | [59,60,73,184–191] |
| Soft lithography | A series of technologies share the use of soft polymer stamps replicating from an original hard master | Small batch: low | High (micrometre scale) | Low device dependency; suitable for small and delicate structure fabrication; capability to reproduce quickly | [192–208,214,215] |
| Coating methods | A series of coating-based thin-layer manufacturing technologies | High (usually one step) | High (micrometre scale) | Limited available structure (only thin layers); multiple alternative methods | [218–229] |
| 3D printing (additive manufacturing) | FDM | Layer-by-layer printing thermoplastic materials | Low (nozzle layer-by-layer printing) | Medium (millimetre scale) | Easily accessible device; good expansibility of device functions; difficult to print long-span suspended structures; limited available materials | [51,234-237] |
| DIW | Layer-by-layer printing thermosetting or photosensitive viscoelastic inks | Low (nozzle layer-by-layer printing) | Medium (millimetre scale) | Easily accessible device; prone to deformation under gravity before curing; good expansibility of device functions; difficult to print long-span suspended structures | [100,153,239–258] |
| Inkjet printing | Jet-based light curing layer-by-layer printing | Low (nozzle layer-by-layer printing) | Medium (micrometre scale) | Usually high cost; difficult to print long-span suspended structures | [153,176,231,259–266] |
| SLA | Selective layer-by-layer photopolymerization of liquid resin | Medium (layer-by-layer photopolymerization) | High (micrometre scale) | Device functions are difficult to expand; easy to print soft and overhanging structures; multiple alternative methods; limited multimaterial printing capabilities | [267–278] |
| SLS | Selectively fuse powder materials layer by layer | Medium (layer-by-layer sintering) | High (micrometre scale) | Device functions are difficult to expand; easy to print soft and overhanging structures; limited multimaterial printing capabilities; limited available materials | [279–288] |
| SDM | Combine material deposition with removal processes | Low (multistep) | Medium (millimetre scale) | Multistep process; allow convex structure; easy to embed components; easy to fabricate heterogeneous material structures | [289–297] |
| Heat sealing | Programmable heating bonding between low-melting materials | High (only one layer of printing) | Medium (millimetre scale) | Low cost; limited available materials; easily accessible device; limited available structure (only suitable for fabrication of 2D fluid actuators) | [122,124,298–305] |
Figure 11. Possible fabrication techniques for next-generation MIS soft devices. The depth of the background color of different methods represents their relevant technical foundations for their use in soft robot community. a) Schematic illustration of the molding process. Reproduced with permission. Copyright 2012, Wiley-VCH. b) Schematic illustration of a flexible fluidic micro-actuator fabricated by molding. Reproduced with permission. Copyright 2013, IOP Publishing Ltd. c,d) Schematic illustration of a typical soft lithography procedure (up) and a flexible pneumatic twisting actuator fabricated by it (down). Reproduced with permission. Copyright 2014, Elsevier. e,f,g) Schematic illustration of spin coating and a microscale trapezoidal soft pneumatic actuator fabricated by it. Reproduced with permission. Copyright 2014, IOP Publishing Ltd. h) Schematic illustration of FDM and a TPU-made cable-driven soft manipulator fabricated by it. Reproduced with permission. Copyright 2015, ASME. i,j) Schematic illustration of DIW and a four-channel pneumatic soft tentacle fabricated by it. Reproduced with permission. Copyright 2016, IEEE. k,l) Schematic illustration of SLA and two miniature soft pneumatic robots fabricated by it. Reproduced with permission. Copyright 2017, IEEE. m,n) Schematic illustration of SLS and a bellow soft actuator fabricated by it. Reproduced with permission. Copyright 2018, IEEE. o,p) Schematic illustration of SDM and a deployable manipulator prototype fabricated by it. Reproduced with permission. Copyright 2019, Wiley-VCH. q,r) Schematic illustration of heat sealing. Reproduced with permission. Copyright 2019, Association for Computing Machinery. s) Schematic illustration of a micro-thin-film fluid actuator fabricated by heat sealing. Reproduced with permission. Copyright 2019, Wiley-VCH.
methods, e.g., lost wax casting, retractable pin casting, as well as lamination casting, have been proposed, and flexible use of them can achieve most of the desired soft structures with relatively smooth surfaces from submillimeter scale to centimeter scale, as shown in Figure 11b. However, molding also faces some problems including possible nonevacuated bubbles and rupture-prone of seams. The former can be eliminated by bubble-removing technologies, e.g., centrifugal casting and vacuum casting, both of which can be used for small-scale perfusion, whereas the latter can be improved by modifying the actuator structure or using rotational casting. For a more detailed account of molding, other reviews are recommended. To obtain wide applications in future industrialization of soft devices, low degree of automation and poor adaptability (the molds are not universal) may be the two main issues that should be considered.

3.1.2 Soft Lithography

Soft lithography was first proposed in the 1990s to serve as an alternative, rapid, and low-cost fabrication technique, and has been widely applied in multiple areas such as microelectronics, biochemistry, microfluidics, as well as soft robotics. Soft lithography includes a family of techniques, which share the utilization of a soft polymeric stamp, e.g., polydimethylsiloxane (PDMS), replicated from an original hard master that is fabricated by photolithography, as shown in Figure 11c. Soft lithography can easily construct polymer microchannels (by replica molding) or microstructures with extremely high resolution (down to 30 nm) which are valuable for the precision manufacturing of complex soft devices (usually fluid driven) and have been demonstrated in various soft robot components, e.g., actuators (Figure 11d), and controllers. Related research in microfluidics can be found in various studies. Although soft lithography has many advantages, e.g., relatively low cost, simple procedure, and high resolution, it still has some limitations. One of which is that the use of soft lithography is often limited by the availability of photolithography-fabricated masters that are usually expensive and time-consuming. Rapid development of high-resolution 3D printing technologies provides new ways for this problem. It can be used to replace photolithography with direct printing masters. Other molds obtained from solidified ferrofluids or even original biological structures can also provide new ideas for obtaining soft lithography masters. Another limitation lies in the difficulty for soft lithography to form or replicate truly 3D hierarchical structures by one-step molding technique. Several recent progress, e.g., ingenious using elastic crack, may pave the way for breaking this limitation, through which some complex bionic structures with special characteristics, e.g., hydrophobic or super-lubricating, can be easily integrated onto the surface of soft devices (it is very attractive for MIS SSIs). In addition, grayscale lithography technology can also be introduced to give soft lithography the capability of replicating complex 3D micro–nano topography. Soft lithography may continue to play an important role in high-resolution manufacturing of microstructures in future soft robots.

3.1.3 Coating Methods

The coating methods can be used to fabricate 2D or 3D elastomeric thin layers with precise thickness, which can be freely cut or stacked to form complex structures with specific functions. Two common coating methods used in soft robots are blade coating and spin coating, both of which are used for the fabrication of flat films. The former mainly scrapes the uncured material manually through a precision blade to remove the excess materials, whereas the latter uses a high-speed rotating disk to throw off the excess material by centrifugal forces, as shown in Figure 11e. Blade coating only requires a set of lightweight custom blades and can easily and conveniently be completed by hand, whereas spin coating exhibits higher accuracy, a larger usable area, and can be easily automated. Other coating methods, e.g., dip coating, spray coating, and drop coating, are not listed here. However, they provide new approaches for the fabrication of complex 3D elastomer films, which may help to fabricate the complex soft structures in future. Coating methods share high efficiency, but their available structures are very limited (only thin film). Materials with relatively low viscosity are preferred for coating to achieve high resolution (down to 10 μm) and avoid adhesion (possible in blade coating), and high elastic modulus is also a must in fluidic-driven actuator applications, as shown in Figure 11f. Several soft materials that can be processed by coating include PDMS, silicone elastomers, and dielectric elastomers (DEs) together with some additives, e.g., magnetic particles, can further enrich the functions of processed thin layers. Although it is convenient to use coating methods to manufacture sensors and microactuators with multimaterial/multilayer structures, these methods can usually only serve as a certain step or several steps in the entire process. Some other processing equipment,
e.g., laser cutters and engraving machines, are often required to cooperate with the spin coater,[223] and integrating them together can be considered in future developments.

3.2. 3D Printing

3D printing, also called additive manufacturing, provides the capability of easily building complex geometric structures, and its core feature is the selective layer-by-layer solidification of inks, resins, or powders to create desired 3D shapes. Considering that most of the rigid molds used in molding are processed by 3D printing currently, it is naturally a valuable idea to directly 3D print soft materials, through which a high degree of automation in a small batch is possible. Extensive research has proved its feasibility in recent years.[176,230–232] Based on different mechanisms, various 3D printing technologies, e.g., fused deposition modeling (FDM), direct ink writing (DIW), inkjet printing, stereolithography (SLA), as well as selective laser sintering (SLS), have been proposed to print soft materials. For more available types and more details, other reviews are recommended.[182,183,213]

3.2.1. Fused Deposition Modeling (FDM)

As one of the most popular 3D printing methods, FDM is cost effective, highly reliable, and easy to use. During its printing process, a heated nozzle melts and extrudes filaments, and then the molten filaments are allowed to cool and solidify freely to be deposited on the stage, as shown in Figure 11g. Although using FDM technology can easily achieve 3D printing of multiple materials[234] and polymer composites,[235] its reliance on material thermoplasticity greatly limits the optional materials and in turn leads to its limited application in soft devices.[236] A relatively common soft material for FDM is thermoplastic urethane (TPU), which has been demonstrated in several applications, e.g., a cable-driven soft manipulator (Figure 11h)[51] and a pneumatic soft actuator.[237] SMP has also been reported to be 3D printable via FDM in a variable-stiffness soft gripper study.[238] To further expand the applications of FDM in soft robots, limited available material is the first issue to be solved.

3.2.2. Direct Ink Writing (DIW)

Another typical extrusion-based 3D printing technology is DIW, which is very suitable for printing soft polymer materials.[238] As shown in Figure 11i, a pressure source, e.g., air pressure, a piston, or a screw, is usually needed for DIW to force the viscoelastic ink to be extruded and deposited onto the stage, together with some additional processes, e.g., thermal curing or photopolymerization, to solidify the printed objects.[239] One of the most significant advantages of DIW is that it can provide a relatively broad range of printable materials,[238] including silicone elastomer,[239,240] polyurethane (PU),[241] and hydrogels.[242] The convenient introduction of additives or matrix can also endow the printing materials with new characteristics[243,244] or improve their mechanical properties,[245,246] thereby further enriching the freedom of DIW. For example, Kim et al. proposed DIW-printed soft robots with programmed ferromagnetic domains by embedding magnetic particles to silicone ink.[247] Furthermore, the structure of DIW printing devices can be flexibly modified and expanded to meet different requirements, e.g., UV-assisted DIW process,[248,249] coaxial printing,[250] and DIW printing with multicomponent mixing function (a printed soft actuator is shown in Figure 11j).[251] Abundant optional materials and extrusion-based printing endow DIW with powerful multimaterial 3D printing capability, and the latest research has achieved rapid fabrication of voxelated soft matter with multimaterial multinozzle 3D (MM3D) printing.[252] Naturally, as the deposited soft materials usually cannot solidify instantly, the deformation of printed objects under gravity is an important issue that should be concerned during printing process, which may greatly affect the processing accuracy. Possible solutions include introducing heating beds[251] or UV lamps[248] embedded printing in support fluid (such as Carbopol),[253] as well as printing easily removed support materials.[254] We are optimistic about the applications of DIW in future MIS, and it has been demonstrated in the fabrication of various soft devices’ components, e.g., magnetic soft actuators,[100] fluidic actuators,[255] DEAs,[256] sensors,[257] circuits,[133] and seamless transition between stiffness gradients.[258] However, the integrated manufacturing capacity of current DIW technology is still limited, and further research is expected to fill this gap.

3.2.3. Inkjet Printing

Different from FDM and DIW, inkjet printing is droplet based and usually contains multiple jetting heads. As shown in Figure 11k, during printing, the entire layer of low-viscosity liquid material is jetted onto the substrate and then cured (usually under UV light). The deposition process of a single layer of soft materials is similar to the printing process of 2D documents, where millions of different ink droplets are ejected simultaneously and form a pattern with a resolution of down to 50 nm.[259] Based on this principle, inkjet printing can easily achieve voxel-level multimaterial 3D printing, thereby allowing the high-precision programming of entities’ mechanical properties.[176,260–263] To further improve the resolution, electrohydrodynamic inkjet printing (e-jet) has been proposed, which uses the electric field force to form a Taylor cone and can flexibly construct micro/nanoscale 3D shapes down to 50 nm.[264,265] E-jet printing technology is hopeful to be widely used as a powerful method for the future construction of micro-/nano-objects. However, such high precision also makes the droplet formation condition of inkjet printing very strict, which requires fine regulation of both the fluid properties and the printing parameters.[259,266] and this in turn limits the diversity of printable materials. There are already some commercial multimaterial 3D printers based on inkjet printing available, e.g., Objet Connex 350 3D printer, and they can be used to fabricate soft devices with relatively complex 3D structures. Numerous soft robots based on these printers have been reported (Figure 11l),[151,231,262] but high cost and still limited integrated manufacturing capability are the main obstacles for their further massive adoption. Inkjet printing is expected to exploit its advantages in micro-/nanoscale high-resolution multimaterial 3D printing in future MIS as well as other related fields. Expanding the printable materials and reducing costs will help the early arrival of this moment.
3.2.4. Stereolithography (SLA)

SLA is the first proposed additive manufacturing technology based on vat polymerization, and various variants have been produced during its development process, including digital projection lithography (DLP),\textsuperscript{267} continuous liquid interface production (CLIP),\textsuperscript{268} two-photon polymerization (2PP),\textsuperscript{269} and computed axial lithography (CAL).\textsuperscript{270} The common characteristic of these methods is the selective photopolymerization of liquid resin using specific light (usually UV light) via layer by layer (Figure 11m) or one-shot (CAL only) exposure. Synthesis in liquid raw materials with the same density can provide self-supporting buoyancy, which is very helpful to print soft and overhanging structures. With the help of micromirror array devices\textsuperscript{271} or dynamic liquid-crystal masks,\textsuperscript{267} DLP and CLIP can solidify an entire layer once and thus obtain a higher printing efficiency than SLA, which relies on a point-source illumination to cure. CLIP can be regarded as a developed DLP. It can further reduce the printing time from hours to minutes with a relatively high resolution (below 100 μm).\textsuperscript{266} As another emerging 3D printing technique, CAL uses superposition of light energy from multiple angles to achieve concurrent printing of all points in printed parts and hopes to improve printing efficiency by several orders of magnitude compared with the layer-by-layer printing.\textsuperscript{270} Unlike other variants, 2PP pays more attention to the resolution and it can provide the highest lateral resolution (around 100 nm) at the cost of small printable volume (limited to 1 cm$^3$).\textsuperscript{272–274} It should be emphasized that all these variants are just trade-offs or selectivity highlighting among various performances of 3D printing, e.g., build speed, volume, as well as resolution,\textsuperscript{270} and none of them can completely replace other variants currently. Although 3D printing technologies based on vat photopolymerization can provide the highest printing precision and considerable printing efficiency, their multimaterial 3D printing capability is poor. More importantly, it is difficult to print composite materials as the additive particles will change the transparency, viscosity, and light scattering rate of raw materials and are easy to settle, all of which are not good for the usual printing of these methods. Although the two main drawbacks will greatly limit the capability of SLA and its variants to print soft devices with functional gradients or new properties in the future MIS, these methods can still serve important roles due to their benefits in resolution and efficiency. The former can be used to fabricate soft robots with complex microstructures and geometries, whereas the latter helps to implement possible commercial large-scale manufacturing. In addition, some attempts have been made to use vat photopolymerization to print multimaterials (Figure 11n)\textsuperscript{275–277} and composite materials,\textsuperscript{278} which will help to further liberate the application potential of SLA and its variants in the future.

3.2.5. Selective Laser Sintering (SLS)

In SLS, a laser raster scans across a powder bed to selectively fuse the powder materials together. Then, a new layer of powder is evenly deposited along the bed, and the printing process is repeated until the 3D part is complete (Figure 11o). During this process, the powders in nonfused regions can serve as supports to the printed structures and can be fully recycled after printing to achieve almost 100% utilization. To ensure fluidity and resolution, the granulated powder size should be small enough (usually between 10 and 100 μm), and the existing SLS resolution can reach 100 μm under some optimal conditions.\textsuperscript{280} In addition, the printing efficiency of SLS can be lifted considerably by keeping the entire powder bed at a temperature slightly below the melting point of powder, thereby reducing the time/energy required for laser to fuse the powders.\textsuperscript{281} The main obstacle of SLS lies in the limited printable materials, which have to meet relatively strict conditions, including sufficient fluidity.\textsuperscript{282} compactibility,\textsuperscript{283} aging stability,\textsuperscript{284} and good thermal properties.\textsuperscript{285} and it is also difficult to process soft materials into usable powders. Due to these obstacles, the multimaterial 3D printing capabilities of SLS based on soft materials are very limited. Expanding the printable soft materials will help SLS to be further used in future soft devices’ integrated manufacturing.\textsuperscript{286} Prior to this, SLS can still take advantage of its high resolution and high efficiency in printing soft parts, which has been demonstrated in some applications, as shown in Figure 11p.\textsuperscript{287,288}

3.3. Shape Deposition Modeling (SDM)

Shape deposition modeling (SDM) refers to a cyclic process for rapid prototyping applications with complex and possibly multimaterial structures, rather than a specific manufacturing method with dedicated equipment. It is based on an alternating use of additive and subtractive manufacturing and combines the former’s capability to rapidly construct arbitrarily complex shapes and the latter’s advantages in high accuracy and quality surface finishes, as shown in Figure 11q. During processing, individual segments of parts and support material structures are gradually deposited into near-net shapes, which are then machined (usually by CNC milling) to net shape before depositing and shaping new materials.\textsuperscript{289,290} SDM allows designers to easily create structures by combining heterogeneous materials and has been demonstrated in several applications, e.g., cockroach limbs with soft polyurethane materials,\textsuperscript{291} surgical graspers (Figure 11r)\textsuperscript{292,293} and hexapedal robots\textsuperscript{294,295} However, it is not easy to achieve a fully automated process, as multiple manufacturing tools must be utilized together and a high degree of control is required to ensure processing quality.\textsuperscript{296,297} Furthermore, fatigue failure due to imperfect processing and insufficient interfacial bonding between different materials also should be considered. Although SDM does not have distinct advantages in resolution and efficiency, some multimaterial complex structures processed by it\textsuperscript{291} and the allowable embedding of actuation or sensing elements\textsuperscript{295,297} are difficult to achieve by other processes. We look forward to the emergence of SDM-based integrated manufacturing platforms soon, and this may greatly enrich the optional structures of surgical instruments in next-generation MIS.

3.4. Heat Sealing

As a mature and inexpensive process, heat sealing has been widely used in plastic packaging for a long time, where the heat-sealing areas of multiple thermoplastic film layers are
heated to a viscous state and then pressurized to stick together, as shown in Figure 11s. Several variants of traditional heat-sealing process have been used for the fabrication of thin-film-based fluid-driven actuators in recent years. These variants can be divided into three categories: manual sealing, heat press sealing, and robotic sealing. Available materials for heat sealing include plastics, e.g., PET, PVC, PE, ABS, paper, and fabric. All are flexible nonstretchable materials, and Ou et al. give a summary of different materials’ fabrication parameters. For materials that are nonthermoplastic or nonsealed, e.g., paper and fabric, coating of thermoplastic polyurethane (TPU) is an ideal solution. Thin-film actuators obtained by heat sealing have attractive design flexibility. Designers can not only combine different materials at will to share their respective advantages, but also the actuation medium can be replaced freely to obtain various characteristics. For example, unthethered actuation can be achieved using a low-boiling-point liquid or dielectric liquid (HASEL) as the medium. The former is controlled by Joule heating, whereas the latter is controlled by the loaded high-voltage electricity. Using low-cost aluminum-coated PET, the circuit can also be directly etched on the actuators’ surfaces to realize the integration of sensors or heaters. Most existing heat sealings are contact heat sealing, and the available fluid chamber resolution is limited to 2–5 cm. Lu et al. recently proposed a so-called noncontact hot air sealing (NoHAS) through which a high resolution (down to 5 mm) of chambers can be achieved, as shown in Figure 11t. However, the use of heat sealing in soft robots is still limited by now, but its high degree of automation and programmability can provide a valuable reference for the development of new fabrication techniques. It is possible to stack or loop these 2D actuators to form 3D structures, which can provide rapid and large strain and has a small footprint (thickness down to 20 μm). These characteristics are very attractive for soft devices used in MIS.

3.5. Summary

All these fabrication techniques involve a process of selectively shaping and curing various soft materials to yield parts with special structures or functions. Among them, conformal replication emphasizes the direct control of shape; 3D printing shares the characteristic of materials’ layer-by-layer addition; SDM is an alternating process of adding and subtracting materials; and heat sealing directly programs existing materials. Different shaping ideas endow them with different pros and cons and thus determine their respective application areas. Both molding and soft lithography can be manually operated without the need for special equipment, although vacuum pumps (to discharge bubbles) or ovens (to speed up curing) are usually involved. Molding is good at rapid building of complex 3D geometric structures, but its poor repeatability and 3D shape error bringing by some manual steps are difficult to eliminate at the present stage. Soft lithography focuses on the high-precision replication of planar/3D microshapes. It is worth noting that both these techniques have the potential to achieve considerable economies of scale, as their molds can be reused, and their current low degree of automation is only a trade-off between the cost of automation and the existing demand in the market.

3D printing and heat sealing have achieved a high degree of automation for prototyping or small-batch manufacturing, in which the 3D/2D structures can be easily designed on the PC side and directly manufactured according to the CAD models. 3D printing offers a series of alternative methods with different characteristics. Among them, SLS and SLA have prominent advantages in printing resolution (especially SLA) and efficiency, but their multimaterial printing capabilities are very poor. On the contrary, FDM, DIW, and inkjet printing show inherent advantages in multimaterial printing, both in structural form and in optional materials (FDM excluded), whereas their efficiency and resolution are relatively low. Heat sealing is another highly automated fabrication technique, which has considerable efficiency and low cost and is suitable for rapid prototyping of millimeter-scale fluid-driven actuators, but its available structure and optional materials still have to be further expanded. The superiorities of the remaining two methods (SDM and coating methods) are both highlighted in the fabrication of some specific structures. SDM can be used to obtain multimaterial structures with smoother convex interfaces or surfaces, and it is convenient to embed components into materials during the fabricating process. However, its implementation often requires multidevice collaboration and is difficult to automate. Coating methods include several different thin-film processing methods, all of which share high efficiency, but the single available structures make them usually only serve as one or several steps in the entire fabricating process.

The further development of these methods, in our opinion, should be oriented to the industrialization demands of next-generation soft devices, in which the integrated manufacturing capabilities are at the core position. However, the integrated manufacturing abilities of current fabrication techniques are still very limited, and the main limitations focus on two parts: material and structure. As shown in Figure 12, the next-generation fabrication techniques are not only expected to have more optional materials and abundant material composites, but also the capabilities of fabricating higher resolution/complexity geometry structures, heterogeneous material structures, as well as multiscale structures. Furthermore, considering the possible commercial benefits in future, the next-generation fabrication
techniques also have to be more automated, cheaper, more reliable, and more efficient. To achieve this vision, the research on new techniques and the improvement and integration of existing methods are all feasible measures. For example, it is possible to integrate inkjet printing, FDM, and DIW in one printer for collaborative printing, which can share the printable materials of each of the three methods and obtain a broader resolution range. Flexible switching between molding, soft lithography, coating methods, and SDM can also obtain more complex structures. In addition, it should be emphasized that the aforementioned fabrication techniques are only some relatively common methods that can be used to fabricate soft actuators and 3D shapes, although most of them can also be used to fabricate other components, e.g., sensors, functional structures, and circuits. In addition to these approaches, there are also some other techniques specifically for integrating sensors on ready-made soft devices, e.g., laser-tuned selective transfer printing and wet transfer process. Among them, the laser-tuned selective transfer printing technology can also be used for the assembly and reconstruction of soft magnetic robots through which we can perceive its potential of being used to construct complex 3D soft structures in a layer-by-layer manner. Introducing them as a postprocessing step or directly integrating them with other methods can be considered in future.

4. Human–Robot Interaction

To improve surgical performance, the next generation of MIS surgical instruments will inevitably move toward being more precise, more flexible, and multifunctional. It will make it unrealistic for surgeons to direct holding instruments by hand for surgery, although hand-held operations will be still feasible in some occasions. Moreover, complex surgical environments and complicated sensing information will also increase the surgeons’ psychological pressure and workload. It will make it more difficult and exhausting for surgeons to only rely on their own experience to complete the entire surgical operation. Therefore, the introduction of intelligent robot-assisted platforms to achieve smooth human–robot collaboration will be an important trend for MIS. Currently, there are many types of robot-assisted platform used in different occasions that are commercially available. To better understand the realization of human–robot collaboration in MIS, we take the most typical da Vinci surgical robot auxiliary platform as an example to abstract the interactions among surgeon, robot-assisted platform, and patient into the circulation of energy and information flow, and thus a clear working principle diagram is given, as shown in Figure 13.

It should be emphasized that all directed closed loops consisting of energy or information flow with length greater than 2 in the figure mean a possible feedback control. As described in Section 1.2, human–robot interaction in MIS consists of two parts: the patient–robot system interaction and the surgeon–robot system interaction. The former emphasizes a comprehensive and high-quality collection of patient-side information (requires a high-performance end-sensing system), whereas the latter emphasizes an efficient and accurate transmission of information between surgeons and robot-assisted platforms (requires an intelligent and ergonomic robot-assisted platform). Moreover, a series of properties of soft surgical devices led by
nonlinear characteristics also bring huge challenges to their accurate modeling and force/position control compared with the rigid ones and related research is still in its infancy. Therefore, establishing reliable control strategies is also essential. Excellent human–robot interaction can exploit instruments’ surgical potential, reduce surgical risk, and improve surgical experiences of both patients and surgeons. To summarize the existing technologies, this section will introduce from the three key parts of such an interaction: sensors, control, and robot assistance.

4.1. Sensors

Sensors serve as robots’ perception core elements. For MIS, an ideal sensor system can provide comprehensive and accurate information for surgeons and robot-assisted platforms, so as to make up for the surgeons’ perception loss and achieve a certain degree of intelligent control. To achieve precise control of surgical process, sensing systems have to obtain comprehensively real-time information from devices, environments, and the interactions between them. Based on the source of information, sensors can be divided into three categories: proprioception, environmental perception, and interactive perception. Notably, the data fusion of different types of sensors is critical, as the information provided by a single type of sensor is usually less useful or even useless. For example, without knowing the real-time position of the surgical instruments, the environmental information returned by the front-end camera will become unrefereced. Different from their rigid counterparts, new characteristics of soft robots also raise new requirements for integrated sensors. As discussed by Polygerinos et al.\(^\text{[57]}\), the ideal integrated sensing elements have to meet the following requirements: 1) sufficient compatibility is essential to avoid them restricting or dramatically modifying the properties of soft devices; 2) elasticity and extensibility are also needed to avoid failure over many cycles of motion; and 3) possessing features that act as stress concentrators are prohibited to avoid damage. Besides these requirements, the interface matching between sensors and surgical instruments is also one of the most important factors to be considered.

4.1.1. Proprioception Sensors

Proprioception is the monitoring of surgical instruments on their own state or position and has the highest priority. The detailed monitoring parameters can be diverse according to the actual requirements. The common ones include strain and end position. Strain information is usually used for the closed-loop motion control of instruments, whereas the end-position information can assist the precise navigation of surgical instruments in complex and dynamic environments. Compared with the others, current sensing technologies for measuring strain are more abundant and there are three common types: optical waveguide, magnetic, and stretchable resistance or capacitance sensors.

Optical waveguide sensors are a kind of sensor that converts the measured physical quantities into various optical ones, e.g., intensity, phase, polarization state, frequency, and wavelength. These types of sensors share the advantages of strong resistance to EM interference, high sensitivity, and good electrical insulation. Among them, fiber Bragg grating (FBG) sensor is one of the most widely used ones in the strain measurement of soft devices in recent years.\(^\text{[100–113]}\) The FBG sensors are capable of reflecting a particular narrow-range wavelength of a broadband light source input with full spectrum, and the reflected wavelength can be affected by the ambient temperature and applied strain. Following this, a relationship between wavelength change and strain can be established. Using the FBG sensors, strain information of each segment of soft devices can be obtained, so as to reconstruct complex shapes. For example, Wang et al. embedded four optical fibers, each of which has four Bragg gratings, into a soft manipulator at 90° to each other to achieve high-precision 3D shape reconstruction with the help of a shape-sensing algorithm.\(^\text{[109]}\) Regardless of their miniaturized outer diameter (≈200 μm) and rich functions, FBG sensors still face many challenges, including error caused by temperature affected by temperature, high cost of customization, and strain transfer reduction.\(^\text{[313,314]}\) To reduce costs and improve availability, Zhao et al. proposed new optical waveguide sensors,\(^\text{[315,316]}\) which can be buried inside the soft materials and their optical loss is proportional to the deformation of the soft devices, as shown in Figure 14a. These stretchable sensors can provide an easy-to-manufacture, low-cost, and highly repeatable way to measure the soft instruments’ strain, through which excellent precision (a signal-to-noise ratio of > 50) is possible.\(^\text{[316]}\) However, the relatively large cross-sectional area (about 3 mm × 3 mm) of the reported fibers still has to be further miniaturized before their use in microscale soft devices, and it is difficult to use them to measure the local strain of soft devices, which has been resolved by a follow-up work recently proposed.\(^\text{[317]}\) Also, the fibers are generally embedded in soft materials in a ‘U’ shape, which may be unfriendly to some segmented and slender SSIs. The ‘I’-shaped optical waveguide sensors based on a similar principle used in SSIs were later proposed by Al Jaber et al.\(^\text{[314]}\) But these sensors often have to align to a specific light source or camera at the instrument’s end, which may affect the operation of possible end manipulators. In addition, there are some soft optical sensors based on other principles, e.g., optically diffuse\(^\text{[118,119]}\) and silicone diaphragm reflection.\(^\text{[120]}\) However, due to their technical difficulties in miniaturization or integration with SSIs, no further introduction is given here.

As another type of emerging strain sensor, magnetic sensors usually use Hall elements to monitor changes in local magnetic flux density generated by micromagnets (≈5 mm) embedded in soft materials to sense the deformation of soft devices.\(^\text{[321–323]}\) As shown in Figure 14b, during fabrication, the Hall element on a flexible circuit and the magnet are positioned in a specific way and are encapsulated in the substrate independently of each other to achieve contact-free sensing, which has a negligible effect on soft material stiffness. Magnetic sensors are proven to be highly accurate (up to 7.5 Hz) and repeatable and can be fabricated and integrated on assembly lines.\(^\text{[321]}\) However, it is difficult to manually calibrate each sensor when the required positioning accuracy is very high. The rigid elements (both Hall element and magnet), although very small, can easily peel off from the substrate when the soft device is subjected to large strains. To avoid delamination, a gradient distribution of material stiffness can be adopted.\(^\text{[177]}\)
Compared with the first two, stretchable resistance or capacitance sensors are the most widely used and own the richest research foundation. Amjadi et al. gave a systematic overview.\footnote{324} With different configurations and sensing mechanisms, resistive-type sensors and capacitive-type sensors transit mechanical strains to electrical signals. The former is usually a soft substrate doped or embedded with conductive materials, and the resistance will change as the applied strain changes. The latter usually shares a sandwich structure, where a pair of stretchable electrodes sandwich a dielectric layer and the applied strain brings two electrodes closer to increase the capacitance. Compared with the capacitive-type ones, the resistance-type sensors show higher strain sensitivity and stretchability but have nonlinear manners and hysteresis. On the contrary, although the linearity and hysteresis performance of the capacitive-type sensors are better, they have lower strain sensitivity. There seems to be a trade-off between sensitivity, stretchability, and linearity of these sensors.\footnote{325}

Up to now, electrically conductive materials used to fabricate stretchable strain sensors include carbon nanotubes (CNTs),\footnote{326} graphene,\footnote{327} carbon black (CB),\footnote{328,329} nanoparticles,\footnote{330} nanowires,\footnote{331} or even liquid alloy (usually eutectic gallium indium, as shown in Figure 14c),\footnote{332} and to facilitate integration and reduce their impact on soft devices, the substrate of the stretchable sensors is often composed by silicone-based elastomers that are similar to or same as the body materials of the soft devices, e.g., PDMS,\footnote{332} Ecoflex,\footnote{326,329} and Dragon Skin. One of the biggest challenges of stretchable sensors is the lack of conductive materials with low modulus of elasticity, and the doping of these conductive additives will increase the stiffness of the substrate. Although liquid metals have desirable low modulus and can be easily integrated by transfer printing,\footnote{336} they are relatively expensive. Recently, the introduction of some new principles such as kirigami\footnote{333} may bring new possibilities to solve this issue.

Figure 14. Several sensors, respectively, for proprioception, environmental perception, and interactive perception in MIS SSIs. a) Optical waveguide sensors for bending angle estimation of soft fingers. Reproduced with permission.\footnote{316} Copyright 2016, AAAS. b) Magnetic sensor (left) is integrated into a soft snake robot (right) for proprioceptive curvature measurement. Reproduced with permission.\footnote{321} Copyright 2015, Elsevier. c) Vacuum-powered soft actuator with embedded liquid-metal-based strain sensors. Reproduced with permission.\footnote{319} Copyright 2019, IEEE. d) Soft continuum robot integrated with EM sensors for effective endoscopic navigation. Reproduced with permission.\footnote{318} Copyright 2017, Mary Ann Liebert, Inc. e) Schematic (left) and photograph (right) of a soft-end effector integrated with optical waveguide for endoscope fluorescence imaging diagnosis. Reproduced with permission.\footnote{310} Copyright 2015, IEEE. f) Thin-film pressure sensor inserted into the constraint layer of soft finger for force perception. Reproduced with permission.\footnote{337} Copyright 2019, Springer Nature. g) Schematic illustrations of four typical transduction mechanisms for tactile sensing, from left to right are piezoresistive, piezocapacitive, piezoelectric, and triboelectric. Reproduced with permission.\footnote{330} Copyright 2017, Elsevier.
In addition to the aforementioned three methods, there are other methods that can also be used to measure the strain of soft devices, e.g., embedded single-electrode triboelectric curvature sensors (S-TECS)\textsuperscript{[334]} and some actuators with self-sensing abilities.\textsuperscript{[152]} Even by measuring fluid pressure (only for fluidic-driven actuators)\textsuperscript{[64,65]} or cable tension (only for cable-driven actuators),\textsuperscript{[172]} an approximate strain range can be inferred indirectly. Furthermore, the existing commercially available strain gauges, which are low cost and flexible but have a nonstretachable body, can also serve as sensing elements for soft devices. However, so far, they are rarely used, and the possible reason is the incompatibility between their substrate materials and the commonly used materials in soft robots, but we suppose that if necessary, they will be ideal sensors to combine with the actuators fabricated by heat sealing, as they share similar materials.\textsuperscript{[301,303]}

To monitor the body motion comprehensively and accurately, multiple strain sensors should be reasonably embedded in different positions of soft devices. However, most of the current research is based on experience to design the distribution of sensors, and this will be more difficult as the complexity of soft devices’ movements and structures increases. To solve this problem, Tapia et al. recently presented a computational method that can automatically compute a minimal strain sensor network for specific soft robotic structures, through which proprioception of soft robots with any shape and size is possible.\textsuperscript{[72]}

Although strain measurement can help reconstruct the instrument shape, the end-position error in the process is difficult to control. Considering such a defect, it is necessary to introduce end-tracking technology in some precision operations to obtain higher accuracy. EM tracking is a common end-tracking technology and has been demonstrated in many continuum robots,\textsuperscript{[119,315,336]} as well as in several commercial products.\textsuperscript{[337,338]}

EM sensors are based on mutual inductance mechanism and can directly provide position and direction information without line-of-sight problems.\textsuperscript{[139]} Due to their small size, the EM sensors can be easily integrated with little influence on the mechanical properties of instruments. By fusing the EM data with machine learning technology, Lee et al. demonstrated accurate trajectory tracking and closed-loop position control (mean error $\leq 2.49^\circ$) of a soft endoscope end even under dynamic external disturbance ($\leq 1$ N), which will be helpful for effective navigation of endoscopes in the complex and changeable environment of the human body, as shown in Figure 14.d\textsuperscript{[336]}. Despite the advantages, EM sensors still have some limitations. For example, they are prone to producing errors due to the magnetic field distortion caused by surrounding electronic or metal devices\textsuperscript{[338,339]} and their measurement accuracy is unevenly distributed in a limited workspace.\textsuperscript{[313,340]}

To solve the former limitation, Sadjadi et al. proposed a simultaneous localization and calibration method, through which the tracking error can be reduced to less than half of the original one.\textsuperscript{[341,342]} For the latter, Reichl et al. deployed an EM field generator at the end of a robotic arm to follow the movement of the EM sensor to keep the sensor in an optimal subspace of the tracking volume.\textsuperscript{[149]} In addition to EM sensor, Liu et al. demonstrated another mechanism to realize end-position control, in which four sets of optical fiber proximity sensors are mounted on a soft endoscope’s end to measure the real-time distance between the end and organ to achieve closed-loop control of the end position.\textsuperscript{[343]} Unlike the EM counterparts, optical fiber proximity sensors are not interfered by magnetic field distortion and have good biocompatibility, but they can only be used to measure the relative position between the instrument end and the organs, and its accuracy may also be affected by the uneven reflective properties of the organ surface.

4.1.2. Environmental Perception Sensors

Current environmental perception mainly focuses on the image perception of the surgical environment. Temperature, humidity, pH, and other information can also be included in the perception range in future MIS. Commonly used medical image-sensing technologies can be divided into two categories: structure-based imaging, e.g., fluoroscopy and ultrasound, and surface-based optical imaging, e.g., flexible optical fiber and miniature camera.\textsuperscript{[344]}

The former can provide surgeons with intuitive image information, e.g., the shape and relative position of tissues and instruments, but has lower image resolution and a certain time-lag.\textsuperscript{[345–348]}

The latter can provide high-resolution real-time color images, which are important for pathological diagnosis and instrument navigation, but the field of view (FOV) is very limited and hardware modification is often required.\textsuperscript{[344,349,350]}

For structure-based imaging technology, fluoroscopy can obtain a relatively clear vision, but corresponding equipment is expensive and bulky and requires the patient to be exposed to extensive radiation. Although ultrasound is relatively safe and can detect depth information of both instruments and tissues, its resolution is limited. Considering the complementary characteristics of the two methods, the fusion of ultrasound and fluoroscopy technology has been demonstrated in some MIS operations such as cardiac catheterization in recent years.\textsuperscript{[351,352]}

During surgery, fluoroscopy can provide high-quality 2D images, whereas ultrasound can supplement depth information in the FOV.

Surface-based optical imaging is also very important for some surgical operations, as the structure-based one can only achieve limited resolution and suffers from lacking information such as color and cavity wall structure. Two common methods for realizing surface-based optical imaging are flexible optical fibers and miniaturized camera, both of which can provide a clear color FOV. Among them, the miniature camera is relatively commonly used because of its mature technology and commercial availability.\textsuperscript{[55,349]}

However, for some surgical instruments that require higher flexibility and smaller size, flexible optical fiber imaging may be a better choice, as it can achieve a smaller size without significantly increasing the cost and has a lower impact on the flexibility of the device.\textsuperscript{[344]}

In addition, some biophotons probe technologies, e.g., fluorescence imaging, can also be introduced to integrate the diagnostic capabilities. For example, a soft end effector integrated with optical waveguide for endoscopic fluorescence imaging diagnosis is shown in Figure 14.e.\textsuperscript{[300]}

Two common challenges for monocular endoscopes are the limited FOV and lack of depth information. To solve these problems, binocular endoscopes can be considered to achieve 3D reconstruction and image stitching.\textsuperscript{[353,354]}

In addition to the earlier techniques, there are also some explorations on the bioinspired electronic eye cameras based on stretchable optoelectronics, through which a wide
FOV and low aberrations are feasible,[355,356] but these progresses still have a long way to go from maturity, standardization, and commercialization.

4.1.3. Interactive Perception Sensors

Interactive perception is used to collect interactive information (usually the magnitude and distribution of force) between surgical instruments and surrounding tissues. Good interactive perception capability can protect tissues from equipment damage and help surgeons for better manipulation. Commonly available technical solutions include thin-film pressure sensors as well as flexible and stretchable tactile sensors. Commercially available thin-film pressure sensors are flexible and have mature technology, but they are generally not stretchable and their substrate materials are not compatible with the commonly used materials for soft devices. Some soft devices integrate them in the strain-limiting layer or the parts with low elongation rate to realize force feedback, as shown in Figure 14f,[357,358] Also, similar to the strain gauge described earlier, integration of them with actuators made by heat sealing may also be an ideal choice.

Compared with thin-film force sensors, flexible tactile sensors based on flexible electronics technology are more commonly used and promising. They usually have better stretchability, better softness, better compatibility, and higher sensitivity and have been richly developed in recent years.[359–362] In addition to capacitive and resistive types that can also be used to measure strain, commonly used ones also include piezoresistive and triboelectric types.[360] The mechanisms are shown in order in Figure 14g. All types of sensors realize sensing by converting strain generated by normal/shear force into electrical signals, and a detailed summary of them can be found in the other reviews.[360] Among them, both piezoelectric and triboelectric sensors can be self-powered, which may be very attractive for slender soft surgical devices that require end sensing, but their flexibility is relatively low. In addition, force detection can also be achieved by means of air pressure sensing[364] or machine vision.[365] The specific implementation method can be flexibly chosen according to the structural characteristics of the soft devices.

Considering the deformable bodies of soft devices, a major challenge of force sensing lies in the decoupling of signals induced by strain and force.[364] Introducing a reference state of deformation might be a solution to achieve decoupling,[365] but it is impractical for soft surgical robots with time-varied deformation. A compromising solution might be miniaturizing force-sensing units and deploying them on the areas with small strain.

4.1.4. Summary

The next generation of SSIs will undoubtedly need to fuse the information comprehensively from multiple sensors (same or different types), while their bodies will become even slender and flexible. Among the sensors mentioned earlier, all sensors except structure-based imaging sensors (fluoroscopy and ultrasound) require different degrees of hardware modification. This makes high-quality and high-density sensing in a limited volume with little impact on the instruments a major challenge for the sensors used in next-generation MIS. The multiplexing of sensor information as well as the integration of sensors and other functional components may be two feasible methods. For the first method, for example, by deploying a set of trackers distributed along the instrument length and fusing the data with kinematics model, EM sensors can be used for shape reconstruction and end tracking at the same time.[313] In addition, the signal coupling problem of force and strain can also be used to reduce the number of sensors, as when one of the two is known, the other or both of them can be calculated. For the second method, some actuators with self-sensing characteristics, e.g., SMA, DEA, IPMC, and HASEL, can act as strain sensors during the actuation, as their resistance/capacitance will change with their body deformation. Based on the same principle, some heaters made of conductive soft materials, e.g., conductive TPU and liquid alloy, can also serve as strain sensors.[152,200] Further, the increase in sensing density and the possible introduction of other components, e.g., power, signal conditioning, and communication units, in next-generation MIS also bring great difficulties to the stable packaging and integration of the sensor systems, which have to make a trade-off between various factors, e.g., sensitivity, material matching, heat dissipation, tightness, etc.

Reliable integration between sensors and soft devices is another technical challenge. The main problem lies in the mismatch between the interface and stiffness of different materials. Magnetic sensors, EM sensors, and minicameras for image perception will inevitably introduce rigid components. Although thin-film pressure sensors and force gauges are flexible, they are almost nonstretchable, and their substrate materials are often incompatible with soft devices. Integrating these sensors with SSIs has to avoid delamination and signal instability. Reducing their size and placing them in parts with low strain are two possible mitigation measures. Stretchable sensors based on flexible electronic technology have a native advantage in integration with soft devices, as their substrate materials can be similar or the same as soft device materials, e.g., PDMS or Ecoflex. However, most of the existing stretchable sensors have limited stretchability, and their structure is also vulnerable to repeated contact or deformation.[166,167] This is mainly due to lack of ideal stretchable conductive materials, and there seems to be a balance between the conductivity and softness. Using conductive liquids, e.g., conductive silver ink,[368] liquid metal,[369] or even sodium chloride (NaCl) solution,[370] may be a promising way to break the balance. Due to their good fluidity and conductivity, these conductive liquids can be embedded manually or even directly integrated into the soft devices through embedded 3D printing.[257] However, the possible leakage is an issue to be considered.

Except the challenges mentioned earlier, there is still a long way to go to develop sensors with high sensitivity, good stretchability, small size, low cost, and suitability for integration with soft devices. In many cases, a trade-off seems necessary among these sensor factors, as two or more of them are in conflict. For example, for reusable slender SSIs, high sensitivity, good stretchability, and small size seem to be more important than cost and manufacturing difficulty, whereas for the less-demanding disposables SSIs, cost and manufacturing difficulty are obviously more important. In addition, up to now, soft surgical devices
with integrated soft sensors are still rare, and preclinical/clinical trials are even scarce.\cite{Kim2021} Kim et al. gave a good demonstration, in which a multifunctional balloon catheter with integrated sensing was developed. The modified balloon catheters provide the ability to sense temperature, flow, tactile, optical and electrophysiological data, as well as completing local ablation of tissue. A series of in vivo experiments in live animal models, e.g., pig and rabbit, were conducted to illustrate their operation and prove their effectiveness.\cite{Kim2021} With the further development of soft technologies, more similar demonstrations are highly desired to be conducted, using more powerful and complex soft surgical devices, which are significant to verify the unique advantages of soft robotic technologies in MIS. Moreover, we are looking forward to seeing sensors based on new paradigms or improved existing paradigms being proposed, and even brand new concepts, such as multimodal structures and multifunctional materials, can be involved to achieve such a goal.

4.2. Control

Precise force/position control is undoubtedly essential for surgical applications, especially for precision operations, e.g., lithotripsy, excision, and suture. However, the nontrivial properties of soft surgical robots, e.g., infinite DOFs, nonlinear characteristics, and diverse properties based on different actuation principles, have made their accurate modeling and control become one of the top challenges in this field. Considering this situation, in this section, we will briefly summarize and analyze the existing control approaches applicable to typical soft surgical devices (a continuum robot with a slender body) and more detailed related reviews can be found in other studies.\cite{Renda2021, Qin2021} Existing controllers for soft robots can be divided into three categories: model-based, model-free, and hybrid controllers.

The model-based controllers highly rely on the analytical models and are by far the most commonly used ones. Numerous efforts have been made to construct accurate and applicable static/kinematic and dynamic models for soft robots.\cite{Kim2021, Qin2021, Renda2021} However, it is known that an accurate and simple kinematic model has been very difficult to develop; hence, a dynamic model based on this is even more difficult. For a static/kinematic model, a steady-state assumption (static equilibrium at each step) is usually involved to facilitate modeling. Up to now, the most adopted modeling method is piecewise constant curvature (PCC),\cite{Kim2021, Qin2021} which is a simplified method with several assumptions, including torsionless body, negligible external load, and uniformly symmetrical configuration. To further improve the accuracy and adaptability, other more complex modeling methods, e.g., beam theory,\cite{Renda2021} cosserat-rod theory,\cite{Kim2021, Qin2021, Renda2021} and finite element method (FEM),\cite{Kim2021, Qin2021} have been developed. However, the improved accuracy also brings sharply increased computational demand and thus greatly limits their usage. A compromise method, called piecewise constant strain (PCS), was proposed by Renda et al., which discretizes the cosserat method and assumes piecewise constant deformation along the soft body (learn from the PCC idea) to achieve simplification.\cite{Renda2021} This method inherits the better adaptability of cosserat method (by considering shears and torsion) and provides a valuable reference for follow-up research. For the dynamic model, related research is still in its early stage. Due to the progressive relationship between them, the performance, e.g., accuracy and complexity, of kinematic models can directly affect that of the dynamic ones. Therefore, the cosserat-rod theory-based\cite{Kim2021} and FEM-based dynamic models will undoubtedly be more precise than the PCC-based ones,\cite{Kim2021, Qin2021} with higer computational demands. To promote the practical use of these dynamic control strategies, some model reduction techniques (especially for FEM) were proposed to reduce the dimension and complexity of the original complex models, with some promising results.\cite{Kim2021, Qin2021} The further development of dynamic controllers requires the coordinated improvement of several factors, e.g., computational power, sensing capabilities, algorithms and modeling methods, and the cooperation of researchers from multiple fields is eagerly anticipated.

The model-free controllers are controllers constructed using some data-based methods, e.g., machine learning and empirical methods, and a related progress report can be found in the study by Chin et al.\cite{Chin2021} One native advantage of model-free controllers is that they are independent of the system complexity and can perform well even in an extremely complex system, which is almost impossible for the model-based ones. However, the model-free approaches can only achieve black box models, and a large amount of data and training time are essential to obtain satisfactory performance. Two common strategies for using these methods to achieve soft manipulator control are learning the inverse statics/kinematics\cite{Chin2021} and learning the forward dynamics.\cite{Chin2021} The former is simpler and more suitable for some slow time-scale tasks with relatively low precision requirements as the dynamic coupling between steps is ignored, whereas the latter can obtain smoother, accurate, and efficient tracking, but also requires longer training time. With the rapid progress of computational power and training algorithms as well as the increased promising related research, we are optimistic about the future use of model-free dynamic controllers in continuum robots.

Compared with the model-free ones, the model-based controllers are more accurate and reliable for some soft manipulators with uniform, compact, and known configurations that operate in a relatively stable and controllable environment. Correspondingly, for manipulators with complex or specific structures or uncertain working environments, model-free controllers provide a better choice. For most surgical procedures, the working environments are known and relatively reliable, so the surgical instruments are to be reasonably designed to enable the model-based controllers. However, for some complicated surgical instruments that have to be specially designed, model-free approaches can also be considered.

Recently, a new trend of soft robot control is to combine model-based and model-free controllers to reinforce each other.\cite{Chin2021, Qin2021, Renda2021} Such hybrid controllers can outperform the two and obtain better performance. For example, Tang et al. combined the merits of model predictive control (MPC, model based) and iterative learning control (ILC, model free) to propose a hybrid method called iterative learning model predictive control (ILMPC), which uses the initial input provided by MPC to decrease the iteration number and refines the model accuracy iteratively.\cite{Chin2021} Also, Gillespie et al. proposed a deep neural network (DNN) model predictive controller, which uses neural
networks to linearize the model, so as to facilitate the model predictive controller to quickly construct complex dynamic models. Several other typical examples also demonstrated the great potential of this research line. In fact, the hybrid way can be diverse, but the hybrid idea may be valuable for further improving the control performance of the next-generation soft surgical robots.

4.3. Robotic Assistance

The surgeons’ advantages are subjective decision-making and judgment based on experience, whereas the robot systems’ advantages are high precision, stability, and powerful data-processing capabilities. A smooth cooperation between surgeons and the robot systems can well utilize their respective advantages and obtain a better overall surgical effect. Unlike the rigid ones, the actuation mechanism of actuators and variable stiffness components for SSIs are far from limited to cable driven. Therefore, the structure of the corresponding robot-assisted platforms will also be more diversified. The main differences will be concentrated on the SSI control end, whereas the other parts, such as consoles, can be universal. For example, the fluid-driven SSIs require the control end to provide accurate and real-time changeable multichannel air/hydraulic pressure, which needs the introduction of a pneumatic/hydraulic control system, whereas the magnetically driven ones have to introduce magnetic field-generating equipment to tune the deformation. Compared with others, the control end of electrically actuated SSIs can be relatively simple, as only the voltages for actuation are required. In addition, depending on the specific human body parts and manipulation needs of the surgical operation, the robot-assisted systems are also different, but a certain range of universality can still be achieved. As the SSIs are not yet popular, so far there are not many reports on related robot-assisted platforms. Considering that many technologies are universal, this section will introduce several relatively mature robot-assisted platforms for rigid instruments. For more examples, readers can refer to other studies.

A prime example of such a robotic platform is the da Vinci surgical system from Intuitive Surgical, Inc., Sunnyvale, CA. Targeting minimally invasive robotic-assisted surgery (RAS), the da Vinci System, is a teleoperated surgical robotic system with a master–slave architecture. The system is composed of three modular hardware components: a patient cart, a surgeon console, and a vision cart. The vision cart serves as the “brain” and provides communication across the entire system, whereas the patient cart and surgeon console are responsible for the interactions among system, patient, and surgeon. In the da Vinci system, haptic feedback is provided via visually displayed cues, whereas the surgeon console along with its proprietary instrument for vision yields an immersive 3D-HD view of the operative field. Under the feedback from the patient cart where surgical operations are conducted, the surgeon controls the whole surgical procedures on the surgeon console via user-friendly finger loops, which compensate fulcrum effect and offer hand-eye coordination recovery. In addition, the da Vinci System provides various proprietary multifunctional instruments to enable a wide spectrum of procedures, which have trusted dexterity by mimicking the movement of human hands and precision through tremor reduction and motion scaling. Due to these unique characteristics, the da Vinci surgical system has been reported to reduce peri- and postoperative complications and hospital stays compared with traditional laparoscopes, thus widely used in the extraluminal procedures on various organs, e.g., prostate, hepatobiliary, and uterus.

The Flex robotic system from Medrobotics Corp., Raynham, MA, is another example. After integrating robotic scope, this robot-assisted platform can define a nonlinear path to the surgical site. Once the robotic scope is in position, Medrobotics Flex proprietary instruments or other compatible instruments can be inserted into the guide tubes along the scope and reach the target. In addition, a familiar joystick-like controller is provided for surgeons, resulting in a short learning curve. Such a platform has been validated to be suited for transoral procedures and gained FDA approval in 2015. There are other robotic platforms designed for specific minimally invasive procedures, e.g., Sensei X robotic catheter system from Hansen Medical Inc., Mountain View, CA, used for cardiac catheter insertion and NeoGuide Endoscopy System from NeoGuide Endoscopy System Inc, Los Gatos, CA, for colonoscopy.

In recent years, some robot-assisted platforms have innovated in different ways by integrating other new technologies. For example, Senhance robotic system (TransEnterix, Morrisville, NC) utilizes eye-tracking technology to enhance the interaction experience of surgeons with the system, which always centers the display at the point the surgeon is looking at and allows surgeon to control the display in a more natural way during surgery. These platforms represent a typical and quite successful human–robot interaction solution for MIS, but they still have some common shortcomings. Apart from the large footprint and high cost, their intelligence and ergonomics also have to be further improved.

To achieve better human–computer interaction and collaboration, the current interaction methods between robot-assisted platforms and surgeons that are mostly limited to vision and sound (a few tones from the devices) are still not enough. The intervention of more sensory organs led by tactile is worthy of being explored, which is also a potential advantageous space for soft robotic technologies. The surface texture, topography, stiffness, deformation, applied force, and other desired information of the internal organs of the patient can be measured and fed back by the integrated sensing system of the surgical tools and then transmitted to the surgeon through tactile reproduction provided by the soft actuator array. Some researches using soft robotic technologies for haptic feedback have emerged in recent years, and a recent review of them can be found in the study by Yin et al. Common haptic feedback methods based on mechanical stimulation include vibrotactile stimulation, normal pressing, and skin stretch, and the introduction of soft technologies can achieve safe, compact, and lightweight system integration as well as good simulation of tactile sensation of soft organs.

5. Perspectives of Next-Generation SSIs

The previous sections give an overview of the recent progresses in SSIs, where the configuration of surgical instrument is
defined as a slender body with circular multichannel cross section. Such a configuration covers almost all existing rigid surgical instruments. However, with the introduction of soft materials, richer intelligence and functions are allowed to be embedded in the structures. Various actuation mechanisms, fabrication techniques, and sensing strategies also greatly enrich the design freedom. It may be improper to continue to use the previous configuration. Therefore, we intend to discuss new configurations/possibilities of SSIs, which may help us liberate our minds and sketch future research, as shown in Figure 15.

5.1. New Design Concepts

Compared with rigid ones, soft and smart materials can achieve larger and flexible changes, e.g., shape, volume, and stiffness. Such a feature allows novel design concepts that may be not applicable to rigid surgical instruments to be introduced to further empower SSIs with more capabilities. An obvious example is modular design. Consequently, the structure of next-generation soft surgical robots may be significantly changed from many aspects. Different functional components, e.g., actuators, sensors, variable stiffness components, etc., can be modularized. Surgeons may freely assemble them according to the practical requirements and easily replace the damaged parts or other functional modules. Further, the soft devices can be automatically split into multiple intelligent modules to realize on-demand reconfiguration, cooperative operation, or just pass through some narrow wiggly channels. However, real-time communication, collaborative movement, and self-reconfiguration between multiple modules in unstructured environments will be technical challenging. Propulsion method is another point with potential for breakthroughs. Avoiding abrasion between devices and tissues will be a core issue. Although lubricated surfaces or the cooperation between end-tip active turning and proximal insertion will be two measures, some ingenious design concepts inspired from nature may be more attractive. For instance, the next-generation MIS soft devices may be growable like a plant, self-propelled like a worm, or even propelled by controllable tissue peristalsis (through safe electrical stimulation, if possible). In addition, some design principles in life, e.g., spiral precession and crawler travel, can also serve as references. To gain high dexterity in limited spaces, origami or kirigami can be introduced in MIS soft devices, through which in situ stiffness manipulation and motion programming of spatial structures can be realized. Possible large fold-out ratio brought by origami is another attractive feature. To improve the output force and dynamic response speed of soft devices, the implantation of some reinforcement components, e.g., spring, can also be considered. Here, we remind the readers that we just mentioned a few typical examples. To solve some specific issues,
seeking inspiration from analogous ones or phenomena in nature or daily life will be very helpful. For example, to improve the adhesion of self-propelled soft devices, various prototypes, e.g., octopus, gecko, beetle, and snail, are potential references. We also encourage our readers to freely exploit or fuse these principles and propose some ingenious approaches according to the actual needs. The introduction of these concepts can further tap the potential of soft materials and demonstrate their unique advantages.

5.2. New Structural Paradigm

For soft materials, a certain degree of physical intelligence can be programmed into their bodies and hence their behaviors can be controlled/triggered through various external stimuli, e.g., electrical, magnetic, or thermal field. Therefore, functional soft robots/components can usually work across multiple scales and are no longer restricted by traditional individual component modes. These features allow SSIs to have extremely rich and free structural design paradigms, which will empower them with new possibilities and stronger capabilities in unstructured environments. In terms of scale, the next-generation MIS soft device family will likely be a multiscale hybrid. Soft manipulators or robots from nanometer to centimeter scale will find their unique applications in their own areas. Small-scale soft robots, e.g., micromagnetic robots, micromanipulators, and nanorobots will exert their unique advantages in dexterity, maneuverability, and accessibility, whereas macroscale soft devices will continue to focus on macroscale navigation, diagnosis, and operation.

Cross-scale collaboration is very attractive, through which their respective advantages may be combined. For example, small-scale soft robots can be attached to a larger-scale one to complete rapid migration and accurately drop near the target point. In some cases, the scale is also changeable. It can be imagined that by some special soft materials, small-scale soft robots can be used separately or as a multi-scale one, which can also be re-distributed at the right time/place. Except for the scale issue, next-generation MIS soft devices will also be freer in structure. The soft device can be not only a multicomponent-integrated system, but also a piece of programmed soft material, or even a controlled liquid or DNA strand. In addition, the shape of the soft devices can also be diverse, e.g., capsule or sphericity, to achieve better acceptance or accessibility to specific targets. Further, a possible parasitic/symbiotic relationship between the soft robots and the human body is also very attractive. Soft robots may have the capability of achieving growth or self-supply by collecting material or energy resources within the human body, e.g., body fluids, temperature, salinity, etc., and can be easily discharged after completing their work. To utilize natural abilities of some organisms or cells (such as sperm), semibiological or cell-based soft robots, although still somewhat distant, can also be considered in future MIS, where the behavior of organisms or cells is controlled by artificial stimuli. These new structural paradigms will greatly broaden the capabilities of MIS and provide surgeons with new opportunities for future precision operation or new therapy methods.

5.3. New Functions/Characteristics

The development of functional soft materials in recent years allows a wealth of soft materials with unique characteristics/functions, which may be very attractive for MIS, to be introduced into MIS soft devices to create new possibilities. For example, earthworm-inspired soft coatings can endow soft devices with superior low friction and excellent antifouling property, which may greatly improve the accessibility and avoid abrasions. Similar effects can also be achieved by hydrogel skins.

Another example is soft materials with self-healing properties, through which fast self-assembly and excellent durability of soft devices are possible. A review of self-healing materials in automated robots can be found in the study by Tan et al. Furthermore, some other characteristics of soft materials, e.g., bacterialid ability, edibility, degradability, etc., are also attractive. We can also imagine that the next-generation MIS soft devices may be auto-fluorescent, just like some deep-sea fish, so as to get rid of the dependence on light guide components, e.g., optical fibers. In addition, some characteristics of specific soft materials can change under external stimulus, e.g., light, thermal, magnetic field, etc. With these characteristics, it is expected to conduct complex control of the behavior of soft robots. The recently reported hydrogel–metal soft robot controlled by light and magnetic field is a good example. We can also combine multiple soft materials with tunable characteristics and fine tune their spatial distribution and structure to achieve complex programming or even logic control. These soft materials with new characteristics have proven their great potential in soft robots. In the future, we expect to see more novel soft materials boldly introduced into the design of robots.

5.4. New Applications

The new design concepts, structural paradigms, and functions/characteristics introduced earlier are expected to significantly expand the capabilities of MIS. Thus, it is necessary for us to think and redefine the application scope of next-generation MIS. First, we believe that the next generation of MIS soft devices will undoubtedly exploit their advantages of multiscale hybrids. Smaller scale means better accessibility. Soft robots of millimeter, micrometer, and nanometer scale will have the capability of directly reaching most of the lesions and conducting works, e.g., drug delivery, thrombolysis, or targeted therapy. During the surgery process, multiscale soft robot operations can be conducted simultaneously to achieve a certain degree of collaboration. Second, more optional surgical operations, e.g., internal wound bonding, can also be achieved in future MIS. Although extensive research has been conducted with soft materials, e.g., hydrogels, to achieve tissue adhesion, there is almost no research to apply them to MIS. We think it will be attractive to use them for wound treatment during surgery process. The adhesive materials may serve as a detachable part of the soft device and can be automatically peeled off when needed. Third, the application of future MIS will probably no longer be limited to disease treatment. Less trauma and discomfort mean that it can also be used for the implantation of some small-scale soft robots or electronic
components, which may perennially or temporarily stay in the human body to achieve functions, e.g., readable and writable information storage, weight management, or location marking. Except the applications listed earlier, there are many other attractive possibilities that can be achieved in future MIS. For example, the soft robot in human body can not only interact with the surgeons but also directly interact with the patient through in vivo stimulation methods, e.g., vibration, electrical stimulation, and heating. Then, the patient can use some specific tools, e.g., a special magnet, to conveniently conduct some simple instructions to the soft robots in any place without the need to seek help from a surgeon.

6. Conclusion and Outlook

The fundamental purpose of MIS is to improve the surgical capability with reduced risks and obtain better surgical experience of both surgeons and patients. Targeting at this purpose, researchers from various fields have achieved considerable progress in the performance, materials, fabrication techniques, sensors, and control strategies of MIS surgical instruments. However, the current instruments’ capabilities from many aspects are still far away from what we expected. Looking ahead to the technical outline of the next generation of MIS and objectively assessing existing technical solutions are of great importance for further development of MIS. For the first time, this Review systematically sketches some key expected characteristics of the next generation of MIS from various aspects. Based on these characteristics, the existing technologies are summarized and analyzed from three different parts: engineering design, fabrication techniques, and human–robot interaction. Finally, perspectives of next-generation SSIs are given. The next generation of MIS will be the product of multidisciplinary integration, and the booming soft robot technology will have the opportunity to burst out its unique advantages in MIS.

To achieve more powerful soft devices for next-generation MIS, the development of advanced soft materials is very crucial, which greatly determines the upper limit of the capabilities of soft robots. New active soft materials with high power density, fast response, and low cost can serve as key components, e.g., actuators or variable stiffness components, to improve the performance of SSIs, whereas soft materials with characteristics, e.g., superhydrophobic, self-lubricating, degradable, and biocompatibility, can bring more possibilities for future MIS. In addition, the integration of different parts of soft devices will inevitably involve the interface matching between various materials. Research on the stable and strong combination between different material interfaces and developing soft materials with better compatibility are very important for the durability of the entire system.

The main bodies composed of soft materials bring great challenges to modeling and control of soft surgical robots. Although acceptable force/position control can be achieved currently, it is still a long way to achieve precise, smooth, and fast navigation of SSIs, where reliable and efficient dynamic control is of great importance. To achieve such a goal, researchers from related research fields, e.g., control theory, artificial intelligence, mathematics, and intelligent algorithms, are highly demanded to be involved. Furthermore, automation is an inevitable trend for surgical robots, and advanced robotic systems will increasingly take over tasks that originally belonged to surgeons, e.g., navigation, stitching, diagnosis, and decision-making. Of course, the relationship between robotic system and surgeons will change from unilateral control to division of labor and finally to possible complete replacement. An in-depth discussion can be found in the study by Thai et al. [403] Although it seems still far away, we look forward to such a day.

The low cost and easy availability of soft robots make disposable SSIs possible, which will also bring a broader space for MIS-oriented soft devices. The partial or overall disposable of surgical instruments can not only save considerable maintenance costs, but also facilitate the flexible customization of some key parameters of the instruments according to the practical conditions of the patient, so as to obtain better surgical results. For some infectious surgical environments, disposable surgical instruments are also an ideal alternative to the traditional surgical instruments. The popularization and application of disposable surgical instruments will be closely related to the development of corresponding manufacturing technologies. Controlling costs while ensuring quality and efficiency will be the key to achieve such vision. Due to their high frequency of replacement, future disposable surgical instruments are more likely to be manufactured by a highly automated and integrated manufacturing platform/assembly line, whereas reusable instruments may introduce some manual process in the manufacturing process if necessary.

Finally, the next generation of SSIs for MIS will emphasize the seamless fusion of multiple technologies. For example, machine learning can not only be used for the control of soft robots, but also for the arrangement of sensors and the structural design of soft actuators. Image processing technology can be used for debouncing of visual field and dynamic calibration. In addition, VR technology and holographic projection can help surgeons understand the surgical environment more intuitively. The introduction of these technologies will definitely enhance the capabilities of MIS.

Acknowledgements

The materials based on the work were partly financially supported by National Key R&D Program of China (2017YFB1303100) and National Science Foundation of China (U1613204).

Conflict of Interest

The authors declare no conflict of interest.

Keywords

fabrication and integrations, human–robot interactions, intelligent soft surgical robots, minimally invasive surgery, robotic-assisted surgery

Received: January 15, 2021
Revised: February 15, 2021
Published online: May 5, 2021
**Jiaqi Zhu** is currently a postgraduate student in Huazhong University of Science and Technology (HUST), School of Mechanical Engineering. His research interests include intelligent soft machines, human–robot interaction, and stretchable electronics.

**Liangxiong Lyu** is currently a postgraduate student in HUST, School of Mechanical Engineering. His research interests include stretchable electronics, soft robotics and surgical instruments.

**Yi Xu** is currently working in the Department of Mechanical Engineering-Intelligent Soft Materials Lab (HUST) as a postdoctoral fellow. His research interests include lab-on-chip systems and soft robotics for bio/medical applications.

**Huageng Liang** is currently an associate chief physician of Department of Urology, Wuhan Union Hospital, Tongji Medical College, HUST, and an associate professor in School of Materials Science and Engineering, HUST. He has 12 years of clinical experience in urology, and skillful to various surgical instrument operations.

**Xiaoping Zhang** is a professor and the chairman of Department of Urology, Wuhan Union Hospital, Tongji Medical College, HUST. He is currently a National member of Chinese Urological Association Committee, and a member of American Urological Association and European Urological Association. Zhang has nearly 20 years of first-line clinical and scientific research experience, and worked in BIDMC and MGH Urology Department of Harvard Medical School for 5 years. He is skilled in various surgical treatments of urological organ tumors including bladder, kidney, prostate, adrenal gland, testis, and penis, and has rich experience and understanding of surgical instrument operation.

**Han Ding** is a professor in State Key Laboratory of Digital Equipment and Manufacturing, HUST, China, and an academician of Chinese Academy of Sciences. He has long been engaged in the research of robotics and digital manufacturing theory and technology, combining robotics and manufacturing technology, and has a deep understanding of rigid robotic systems. In recent years, he paid more attention to emerging robotic technologies/concepts, e.g., soft robotics, teleoperation, machine learning, and their applications in medical and aerospace fields.
Zhigang Wu is currently a professor in State Key Laboratory of Digital Equipment and Manufacturing, HUST, China. Prior to this, he was an associate professor in the division of Microsystems Technology, Uppsala University, Sweden, and held a research fellow at Swedish Research Council. His research interests are concentrated on processing and integration technologies of soft electronics and machines for various applications, e.g., in surgical medicine, industrial, life sciences, rehabilitation and home healthcare, and fundamental research and applications in microfluidics for point-of-care healthcare.