論文の内容の要旨
Thesis Summary

Development of a Self-Contained Image Mapping Framework for Ultrasound-Guided Fetoscopic Procedures via Three-Dimensional Dynamic View Expansion

(超音波画像誘導下胎児外科手術のための外部センサを用いない内視鏡三次元画像マッピングシステムの開発)

氏　名：　ヤン　リャンジン

Although surgical navigation technologies address many challenges faced in minimally invasive procedures of modern surgical treatments and medical interventions, applications of the state-of-the-art technology for case-specific minimally invasive procedures remain challenging due to their respective unique clinical requirements. One example is the minimally invasive fetoscopic procedure, which includes laser photocoagulation treatment on abnormal blood vessels in the case of Twin-to-Twin Transfusion Syndrome (TTTS). Without treatment, prenatal loss associated with this abnormality is 90% with high risk of neurological impairment in survival case.

Like many other forms of endoscopy, fetoscopic imaging is associated with a narrow field-of-view. Therefore, it is extremely challenging to navigate surgical instruments within the womb given limited visual information. The close-range imaging and poorly textured placental surface further worsens the problem. Without a navigational global perspective of the vast vasculature network, surgeons may have to revisit explored areas repetitively. While current clinical practices include intraoperative ultrasound image-guidance for navigation, surgeons have to rely on non intuitive mental registration between the two separate sources of visual information adding further complexity to the challenging hand-eye coordination required in minimally invasive procedures. These issues have inspired several researches to develop techniques that provide navigational visualization in fetoscopic procedures. However, existing related works have either not address the problem of non intuitive visual guidance in fetoscopic procedures or requires external navigation system that burdens the medical team including the surgeons. The literatures of three relevant subjects are reviewed. These subjects include ultrasound image-guided navigation, vision-based tracking, and dynamic view expansion. While technologies developed in these three areas have been promising for surgical navigation applications in various medical fields, there are clear research gaps in translating these techniques for fetoscopic procedures. The close-range imaging and poorly textured scene associated with placental imaging make pure vision-based techniques inadequate without integrating with other source of information. Innovations in integrated framework design and application-specific modular techniques are expected to close up these research gaps.

To address the clinical need for intuitive intraoperative guidance and inadequacy in existing researches on surgical navigational technology for minimally invasive fetoscopic procedures, an ultrasound-guided 3D dynamic view expansion method is proposed and evaluated as a self-contained image mapping framework for surgical

navigation and visualization. Unlike most surgical navigation technologies, the proposed framework requires no external tracking device as it obtains spatial information from ultrasonography and endoscopy, which are already in use during ultrasound image-guided fetoscopic procedures.

The development process adopted a design-centric approach that begins with the conceptualization of the framework. This comprises problem identification, design consideration and the proposal of an appropriate solution. The problem identification process aligns the research scope defined in this thesis with actual clinical needs and engineering applications while the design-centric development methodology ensures sound conceptualization with well-defined specifications derived from the basis of needs, functionalities, and usability. The outcome of the conceptualization process is a 3D dynamic view expansion method implemented through a tracker-less approach proposed as a self-contained image mapping framework for the integration of ultrasonography and endoscopy.

The framework consists of four core modules namely, ultrasound-based initialization (**U/S**), endoscopic image alignment (**IA**), vision-based endoscopic tracking (**V-B**), and texture mapping (**Tx**). These modules combine to achieve image mapping between the photorealistic endoscopic vision and the 3D navigational ultrasound model.

The developed **U/S** module is a suite of operations that comprises 3D reconstruction of the scene geometry, instrument localization, outlier removal, and an optimization process based on the minimally invasive kinematic constraint. Conceptually, positional information obtained from ultrasound image-based localization can be used to overlay endoscopic views directly to the scene. However in practice, timely updates are limited by the acquisition and processing speed, which is highly resource dependent. In view of this practical developmental constraint, ultrasound-based localization is proposed as a module for initialization and relocalization in this study.

The endoscopic **IA** module adopts a feature-based approach, which is also a suitable option for developing the **V-B** module as will be discussed later. The differences in these two modules lie in the type of transformation estimated. In **IA**, matching features are used to solve for homographies that transform 2D images onto a single mosaic plane. This is a 2D-2D transformation that projects a given view in alignment with a 2D mosaic or its previous view. In vision-based tracking however, the goal is to estimate the transformation between camera coordinate systems at different time frame. This tracking process solves the 2D-3D registration problem through pose estimations based on continuous corresponding of 2D features and 3D interest points. Endoscopic vision can consequently be registered to the 3D geometrical scene.

While **U/S**, **IA** and **V-B** provide positional information for initialization or registration, image mapping of the photorealistic endoscopic views to the 3D geometrical scene is eventually realized through texture mapping using the developed **Tx** module. This is done by updating the texture map of the 3D patched surface with the color scale-indexed images acquired from endoscopic camera. The registration information through **U/S**, **IA**, or **V-B** is used to map the relevant texture information from the indexed images of the endoscopic views to the rightful patches comprising the 3D scene.

In general, the image mapping framework begins with a one-time ultrasound initialization through ultrasound-based 3D scene construction and initial camera position localization (**U/S**). Subsequent updates of the moving endoscopic vision on the ultrasound-constructed 3D model of the placental surface can be achieved through the mapping of instantaneous 2D views to the 3D scene using spatial information obtained by vision-based tracking of the endoscopic motion (**V-B**). Alternatively, the 3D dynamic view expansion can be propagated through image alignment of adjacent views (**IA)** onto a mosaic plane that has been registered to the scene via the one-time ultrasound-based initialization. This alternative variation is a useful simplification of the framework to achieve 3D dynamic view expansion within a limited locality. Module **U/S** can also be used to reinitialize i.e. relocalize camera position in event when cumulative error becomes excessively large due to the sequential estimations in **IA** and **V-B**. By combining **U/S**, **IA** and **V-B** in a strategic manner discussed in the thesis, timely updating can be achieved with better accuracy. In this thesis, two variations of the framework for collaborative updating of image map are introduced. The coarse-to-fine scheduling scheme (**C2F**) updates texture mapping using the three different modules in a staggered coherent manner. Untracked image mapping using **IA** module deployed at the finest frequency bypasses the need for camera position estimation during image mapping updates. At mid-range frequency, the more accurate 2D-3D registration updates is realized by **V-B** to relief the homography-based **IA** updates from further estimation errors. This spreading out of **V-B** updates also minimizes overfitting. This is a problem caused by over reliance on vision sensing. Since **V-B** is not constrained by the 2D homography transformation the higher degree 3D perspective transformation may lead to larger error if the innate sensor uncertainty of the camera is large. Image mapping update using **U/S** being the most time consuming approach is scheduled at the coarsest interval. To further incorporate adaptability of the collaborative scheme, a heuristic approach termed hybrid mapping (**Hyb**), which determines the appropriate updating modules based on conditional evaluation of observations, is also proposed. This scheme directs the process flow to the necessary updating modules based on a set of heuristics that aims to avoid module-related failures.

Simulation analysis, phantom model experiments, and ex vivo validations are performed in the evaluation study to ensure soundness of the proposed framework. In the simulation analysis, simulated camera trajectories constrained by a pivotal point mimicking the incision port during a fetoscopic surgery are synthesized to study the influence of homography-based mapping accuracy under different perspectives of the camera. The simulated conditions are synthesized from actual endoscopic camera parameters and phantom placenta geometry acquired from 3D ultrasound scanning. According to the simulation analysis, the valid endoscopic camera workspace for homography-based image alignment appears to be limited to a conical workspace that spans 120° around a pivotal apex with its camera-scene distance approximately 20-40 mm in order to maintain RMS reprojection error below 3 mm, which is a dimension comparable to a photocoagulation lesion sizes. The observation represents an upper bound limitation of the homography-based image alignment under ideal simulated conditions.

In the phantom model experiments, imaging is carried out on a phantom placenta model in a water tank environment. The experiments encompass analysis and tests under controlled conditions. This includes evaluation of ultrasound-based localization, vision-based static localization and motion tracking, and the consistency of the final image mapping results. Image alignment measurement error in RMS is maintained at 3.0 mm with standard deviation of 1.9 mm within 80 processed frames of mapping, which is equivalent to 112 s. This is conservatively sufficient to cover an overlay of 30 cm^2 before requiring positional update from either **V-B** or **U/S**. The absolute mean errors based on 4 physical landmark measurements of 20-frame image mapping results using direct ultrasound-based localization (**U/S**) and sequential vision-based tracking (1x **U/S** + **V-B**) are 1.3 mm and 1.6 mm, respectively.

In the ex vivo validation studies, the framework static localization accuracy is evaluated and visual inspection of the image mapping using the image alignment approach is performed on a monkey placenta. The overall consistency of static localization is within the order of a millimeter and a degree despite six times increase in error magnitude along the maximum error axis compared with the results of the phantom model experiment.

Finally, profiling of the computation performance is carried out to provide a conservative gauge of the practicality of the implementation. While the current prototype is not designed for speed performance it is able to ensure processing rate of at least 2 fps with (1x **U/S** + **V-B**). In practice, higher rate can be achieved with more lenient parameters but at the cost of lower accuracy. As the system specifications have yet to be finalized depending on actual accuracy requirements and users' priorities, it is not in the interest of this current study to design for speed performance.

While the feasibility of the proposed framework has been demonstrated through the evaluation method discussed, there are limitations and certainly room for improvement. One such area is the limitation in the realism of the largely phantom-based experiment method used in this study. While the phantom model experiment is important in performing repeatable analytical study with controlled conditions to evaluate motion consistency, there are many other factors that affect the performance of image mapping in actual clinical condition. Current phantom and ex vivo studies do not adequately represent clinical conditions in fetoscopic procedures. With the technical feasibility demonstrated in this thesis, the next step is to validate the method on in vivo data and conditions.

It should be noted that the proposed framework does not directly enhance the contrast of the vessels despite mapping photorealistic endoscopic vision to the 3D scene model. Nevertheless, it contributes as a platform for the incorporation of contemporary development in vessel enhancement techniques during placental imaging including, but not limited to, the use of contrast agents and hyperspectral imaging. While effective visualization of placental vasculature cannot be achieved perfectly in this thesis alone, current work focuses on the improvement of navigational perspective to enhance surgical visualization for treating communicating vessels.

In conclusion, the contribution of this thesis is the introduction of a self-contained image mapping framework between ultrasonography and endoscopy through 3D dynamic view expansion done via a methodical research and development approach that includes literature survey, conceptual designing, formulating and engineering of modular techniques, and integration of modules for collaborative workflows. Scientific investigation for evaluation of the framework is carried out with simulation analysis, experiment studies, and ex vivo validation. While the method is demonstrated on fetoscopic placental imaging, it can also be extended to general tissue imaging with potential applications in diagnostic visualization, and surgical augmented reality and robotics. It is hoped that this study will open up many new possibilities for advancement in the application of computer-aided-surgery technologies for rare disease treatments including those requiring fetoscopic procedures.