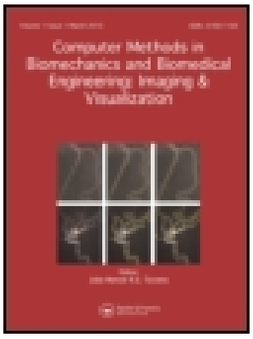

Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tciv20

KST-Mixer: kinematic spatio-temporal data mixer for colon shape estimation

Masahiro Oda, Kazuhiro Furukawa, Nassir Navab & Kensaku Mori

To cite this article: Masahiro Oda, Kazuhiro Furukawa, Nassir Navab & Kensaku Mori (2023): KST-Mixer: kinematic spatio-temporal data mixer for colon shape estimation, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, DOI: 10.1080/21681163.2022.2151938

To link to this article: https://doi.org/10.1080/21681163.2022.2151938

Published online: 30 Jan 2023.

COMPUTER METHODS IN BIOMECHANICS AND BIOMEDICAL ENGINEERING: IMAGING & VISUALIZATION
https://doi.org/10.1080/21681163.2022.2151938

KST-Mixer: kinematic spatio-temporal data mixer for colon shape estimation

Masahiro Oda [a,b], Kazuhiro Furukawa [c], Nassir Navab [d] and Kensaku Mori [b,e]

[a]Information and Communications, Nagoya University, Nagoya, Japan; [b]Graduate School of Informatics, Nagoya University, Nagoya, Japan; [c]Department of Gastroenterology and Hepatology, Nagoya University Graduate School of Medicine, Nagoya, Japan; [d]Chair for Computer Aided Medical Procedures, Technical University of Munich, Munich, Germany; [e]Research Center for Medical Bigdata, National Institute of Informatics, Tokyo, Japan

ABSTRACT

We propose a spatio-temporal mixing kinematic data estimation method to estimate the shape of the colon with deformations caused by colonoscope insertion. Endoscope tracking or a navigation system that navigates physicians to target positions is needed to reduce such complications as organ perforations. Although many previous methods focused to track bronchoscopes and surgical endoscopes, few number of colonoscope tracking methods were proposed because the colon largely deforms during colonoscope insertion. The deformation causes significant tracking errors. Colon deformation should be considered in the tracking process. We propose a colon shape estimation method using a Kinematic Spatio-Temporal data Mixer (KST-Mixer) that can be used during colonoscope insertions to the colon. Kinematic data of a colonoscope and the colon, including positions and directions of their centerlines, are obtained using electromagnetic and depth sensors. The proposed method separates the data into subgroups along the spatial and temporal axes. The KST-Mixer extracts kinematic features and mix them along the axes multiple times. We evaluated colon shape estimation accuracies in phantom studies. The proposed method achieved 11.92 mm mean Euclidean distance error, the smallest of the previous methods. Statistical analysis indicated that the proposed method significantly reduced the error compared to the previous methods.

ARTICLE HISTORY
Received 21 October 2022
Accepted 19 November 2022

KEYWORDS
Colon; colonoscope tracking; kinematic data estimation; multi layer perceptron

1. Introduction

[1]CT colonography (CTC) is performed to find colonic polyps from CT images. If colonic polyps or early-stage cancers are found in a CTC, a colonoscopic examination is performed to endoscopically remove them. A physician controls the colonoscope based on its camera view during a colonoscope examination. However, its viewing field is limited and unclear because the camera is often covered by fluid or the colonic wall. Furthermore, the colon changes shape significantly during colonoscope insertion. Physicians require much skill and experience to estimate how the colonoscope travels inside the colon. Inexperienced physicians overlook polyps or cause such complications as colon perforation. A colonoscope navigation system is needed that guides the physician to the polyp position. A colonoscope tracking method is necessary as its core.

Many tracking methods of endoscopes are proposed (Deligianni et al. 2005; Mori et al. 2005; Gildea et al. 2006; Schwarz et al. 2006; Peters and Cleary 2008; Rai et al. 2008; Deguchi et al. 2009; Ching et al. 2010; Liu et al. 2013; Luo et al. 2015; Fukuzawa et al. 2015; Visentini-Scarzanella et al. 2017; Oda et al. 2017; Wang et al. 2021; Banach et al. 2021; Yao et al. 2021). They can be classified into the image-based, the sensor-based, and the hybrid methods. The bronchoscope is the main application of bronchoscope tracking. Many researchers proposed image- and sensor-based methods in recent several decades. Image-based tracking methods estimate the camera movements based on image registrations. Registrations between temporally continuous bronchoscopic images (Peters and Cleary 2008) or between real and virtualised bronchoscopic images (Deligianni et al. 2005; Rai et al. 2008; Deguchi et al. 2009) are used for tracking. Recent approaches use deep learning-based depth estimation results to improve image-based tracking accuracy (Visentini-Scarzanella et al. 2017; Wang et al. 2021; Banach et al. 2021). Sensor-based tracking methods use small sensors to obtain bronchoscope position (Gildea et al. 2006; Schwarz et al. 2006). Hybrid methods of them use both image and sensor information to accurately estimate bronchoscope position (Mori et al. 2005; Luo et al. 2015). Some research groups propose tracking methods for colonoscope. In colonoscope tracking, the image-based method (Liu et al. 2013) is difficult to apply because unclear colonoscopic images appear frequently. Unclear image removal and disparity map estimation are utilised to improve tracking accuracy (Yao et al. 2021). Electromagnetic (EM) sensors are used to obtain colonoscope shapes (Ching et al. 2010; Fukuzawa et al. 2015). Unfortunately, they cannot guide physicians to polyp positions because they cannot map the colonoscope shape to a colon in a CT volume, which may contain polyp detection results. Combining the colonoscope position or shape information with polyp position information, which can be detected in a CT volume taken prior to colonoscope insertion, is essential in colonoscope navigation. Such navigation is called colonoscope-CT-based navigation.

CONTACT Masahiro Oda moda@mori.m.is.nagoya-u.ac.jp Information and Communications, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, Nagoya 464-8601, Japan

A few colonoscope tracking methods are applicable to perform colonoscope-CT-based navigation. Reference (Oda et al. 2017) used colonoscope shape measured by an EM sensor and CT volume for colonoscope tracking. The method obtains two curved lines representing the colon and colonoscope shapes to estimate the colonoscope position on a CT volume coordinate system. The method enables real-time tracking regardless of the colonoscopic image quality. However, the method does not consider colon deformation caused by colonoscope insertion. Such deformation caused significant tracking error. Large tracking errors were observed at the transverse and sigmoid colons, which are significantly deformed by a colonoscope insertion. To reduce tracking errors, estimation methods of the shape of the colon with deformations caused by colonoscope insertion were proposed (Oda 2018a,b). One of them (Oda 2018a) uses the shape estimation network that has a long short-term memory (LSTM) layer (Hochreiter and Schmidhuber 1997) to estimate the colon shape and the other (Oda 2018b) uses regression forests. Estimation accuracies of them need to be improved to perform accurate colonoscope tracking.

We propose a novel shape estimation method of the colon for colonoscope tracking using a Kinematic Spatio-Temporal data Mixer (KST-Mixer). The proposed method estimates the colon from time-series shape data of the colonoscope, which is inserted into the colon. The KST-Mixer extracts kinematic features from the data of the colonoscope using simple multi-layer perceptrons (MLPs). The extracted features are mixed along the spatial and temporal axes in spatio-temporal mixing blocks to estimate the colon shape dynamically. Because the KST-Mixer has a simple processing flow, it provides estimation results in a short processing time. It is suitable to be used in real-time colonoscope navigation systems.

Contributions of this paper are summarised as: **(1)** to propose a novel colon shape estimation method that utilises spatial and temporal kinematic features extraction and mixing processes, **(2)** to enable short computation time in estimation used in real-time colonoscope tracking, and **(3)** to achieve the smallest shape estimation error among the previous methods.

2. Method

2.1. Overview

The proposed method estimates the colon shape from the colonoscope shape. They are time-series data measured at a specific time interval. The KST-Mixer is trained to estimate a colon shape from time-series colonoscope shapes. After the training, a trained model estimates a colon shape during a colonoscope insertion.

2.2. Colon and colonoscope shape representation

We represent the colon and colonoscope shapes as point sets. The colonoscope shape (Figure 1(a)) is a set of 3D positions $\mathbf{p}_n^{(t)}$ and 3D directions $\mathbf{d}_n^{(t)}$, that is represented as

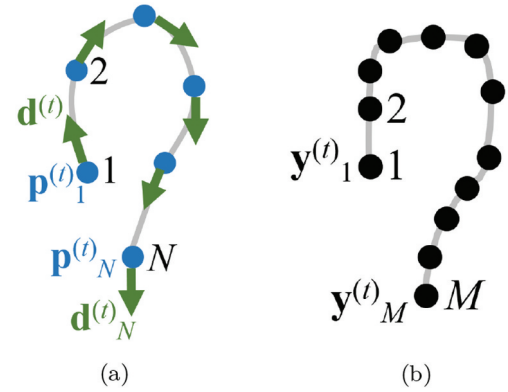

Figure 1. Examples of (a) colonoscope and (b) colon shapes.

$$\mathbf{X}^{(t)} = \{\mathbf{p}_n^{(t)}, \mathbf{d}_n^{(t)}; n = 1, \ldots, N\}, \tag{1}$$

where t $(t = 1, \ldots, T)$ is the index of time, T is the total number of time frames, and N is the number of points in the colonoscope shape. $\mathbf{p}_n^{(t)}$ is a point aligned along the colonoscope centerline. $\mathbf{p}_1^{(t)}$ corresponds to the position of the colonoscope tip. $\mathbf{d}_n^{(t)}$ is a tangent direction of the colonoscope tube at $\mathbf{p}_n^{(t)}$.

The colon shape (Figure 1(b)) is a set of 3D points $\mathbf{y}_m^{(t)}$ that is represented as

$$\mathbf{Y}^{(t)} = \{\mathbf{y}_m^{(t)}; m = 1, \ldots, M\}, \tag{2}$$

where M is the number of points in the colon shape. $\mathbf{y}_m^{(t)}$ is a point aligned along the colon centerline. $\mathbf{y}_1^{(t)}$ and $\mathbf{y}_m^{(t)}$ correspond to the caecum and the anus positions, respectively.

2.3. Kinematic spatio-temporal data mixer (KST-Mixer)

2.3.1. Overview of KST-Mixer

The KST-Mixer estimates a colon shape from time-series colonoscope shapes. Its architecture is based on MLPs that are repeatedly applied across the spatial or temporal axes. This architecture is inspired by the MLP-Mixer (Tolstikhin et al. 2021), which classifies images utilising spatial locations and image features. The MLP-Mixer has competitive image classification performance to current methods such as Vision Transformers (ViT) (Dosovitskiy et al. 2021) and provides a short processing time. We utilise the MLP-based architecture to process kinematic data in shape estimation tasks.

2.3.2. Data preparation

To generate input data of the KST-Mixer, we rearrange a time-series colonoscope shape data as a matrix with spatial and temporal axes. Components of the 3D point and direction are represented as $\mathbf{p}_n^{(t)} = (p_{n,x}^{(t)}, p_{n,y}^{(t)}, p_{n,z}^{(t)})$ and $\mathbf{d}_n^{(t)} = (d_{n,x}^{(t)}, d_{n,y}^{(t)}, d_{n,z}^{(t)})$. From $\mathbf{p}_n^{(t)}$, we define the positional matrix $(3N \times \tau)$ of time period $t = t_c, \ldots, t_c - \tau + 1$ as

$$\mathbf{P}^{(t_c)} = \begin{pmatrix} p_{1,x}^{(t_c)} & \cdots & p_{1,x}^{(t_c-\tau+1)} \\ p_{1,y}^{(t_c)} & \cdots & p_{1,y}^{(t_c-\tau+1)} \\ p_{1,z}^{(t_c)} & \cdots & p_{1,z}^{(t_c-\tau+1)} \\ \vdots & \ddots & \vdots \\ p_{N,x}^{(t_c)} & \cdots & p_{N,x}^{(t_c-\tau+1)} \\ p_{N,y}^{(t_c)} & \cdots & p_{N,y}^{(t_c-\tau+1)} \\ p_{N,z}^{(t_c)} & \cdots & p_{N,z}^{(t_c-\tau+1)} \end{pmatrix}, \tag{3}$$

where τ is time length. We also define the directional matrix as $\mathbf{D}^{(t_c)}$ similarly.

Values in $\mathbf{P}^{(t_c)}$ are normalised to take values in the range $[0, 1]$. We regard the normalised matrix as a 2D image to generate S non-overlapping and homogeneous sized image patches. The image size is $(3N, \tau)$ and the size of each patch is (s_1, s_2). From them, the number of patches is calculated as $S = \frac{3N\tau}{s_1 s_2}$. Each patch contains spatially and temporally local data. Each patch is projected to a feature vector of hidden dimension C. As the result, we obtain a input matrix of positional data $\Pi^{(t_c)} \in \mathbb{R}^{\mathbb{S}\times\mathbb{C}}$. The order of feature values in the matrix is sensitive to both spatial and temporal axes. Therefore, such a position embedding technique as ViT employs is not necessary. We also obtain a input matrix of directional data $\Delta^{(t_c)}$ from $\mathbf{D}^{(t_c)}$. We make a matrix of colonoscope shape data $\Xi^{(t_c)} \in \mathbb{R}^{2S\times C}$ consisting of $\Pi^{(t_c)}$ and $\Delta^{(t_c)}$ elements. This process is illustrated in Figure 2.

We generate additional input data of the KST-Mixer from the insertion length of the colonoscope. The insertion length of colonoscope at time t is $l^{(t)}$. A set of insertion lengths in the period of time $t = t_c, \ldots, t_c - \tau + 1$ is represented as

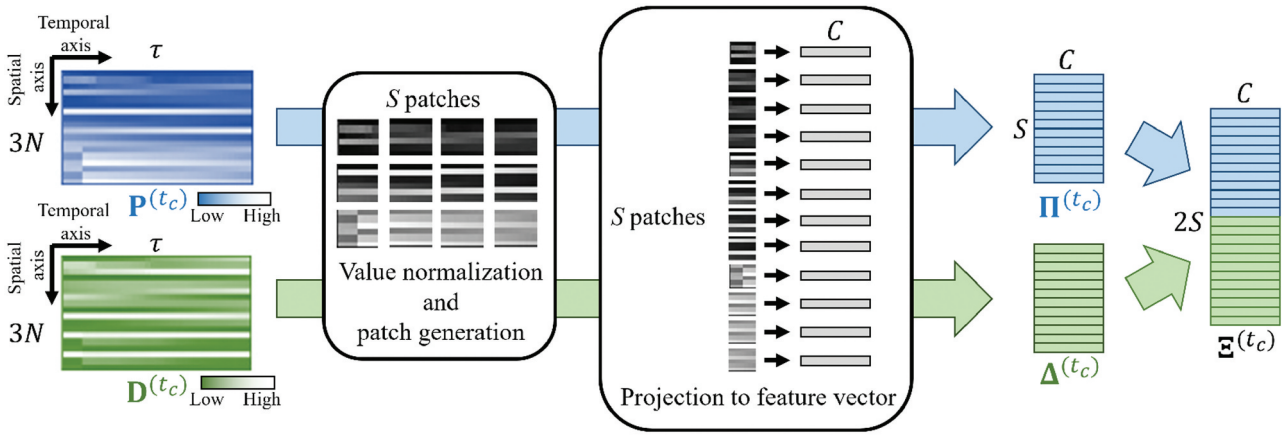

Figure 2. Data preparation process from positional and directional matrixes $\mathbf{P}^{(t_c)}$, $\mathbf{D}^{(t_c)}$ to $\Xi^{(t_c)}$.

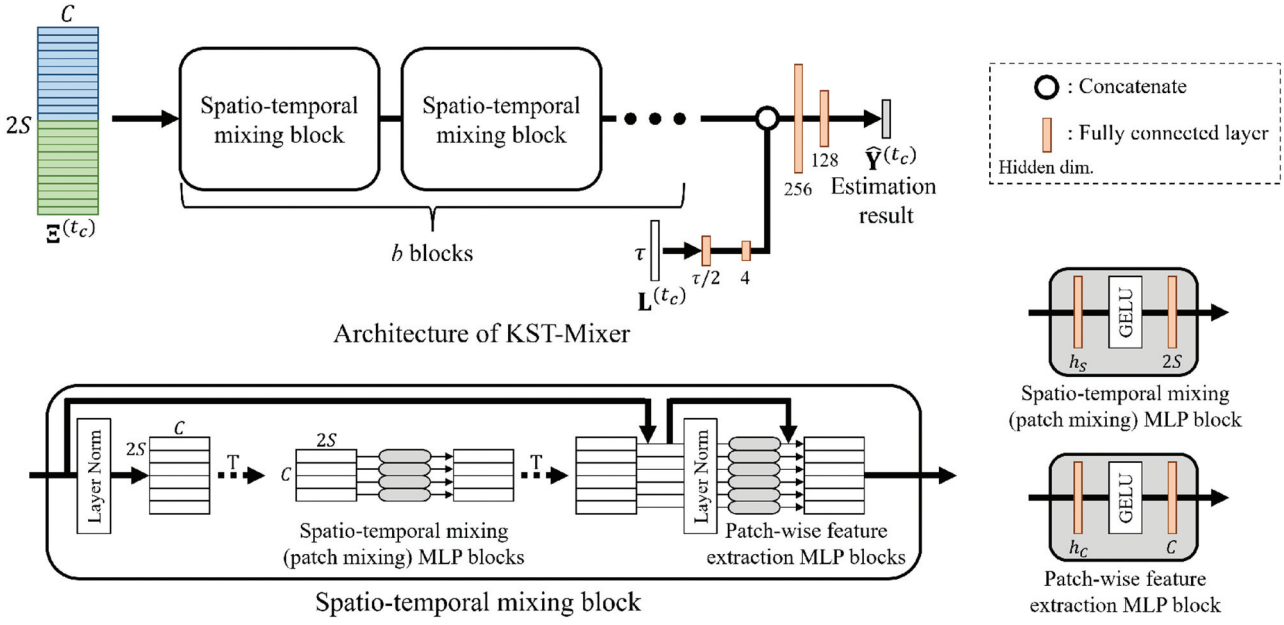

Figure 3. Architecture of KST-Mixer. It estimates colon shape $\hat{\mathbf{Y}}^{t_c}$ from colonoscope shape data Ξ^{t_c} and insertion length data $\mathbf{L}^{t_c}$. It extracts colonoscope shape features in multiple spatio-temporal mixing blocks. Features are then combined with insertion length feature. Fully connected layers process combined features to generate estimation results.

$\mathbf{L}^{(t_c)} = (l^{(t_c)} \dots l^{(t_c-\tau+1)})^T$, which is a column vector of size $\tau \times 1$. The insertion length data $\mathbf{L}^{t_c}$ is used in the process of the KST-Mixer.

2.3.3. Architecture of KST-Mixer

The input of the KST-Mixer are $\Xi^{(t_c)}$ and $\mathbf{L}^{(t_c)}$. The KST-Mixer outputs an estimated colon shape $\hat{\mathbf{Y}}^{(t_c)}$. The architecture of the KST-Mixer is shown in Figure 3. It has b spatio-temporal mixing blocks. Each of them consists of two MLP blocks similarly to the MLP-Mixer (Tolstikhin et al. 2021). The first is the spatio-temporal feature mixing MLP block (patch mixing MLP block). In the block, an input patch-wise feature vector is transposed and processed by the MLP block. The second MLP block is the patch-wise feature extraction MLP block.

Each MLP block has two fully connected layers and an activation function. Dropout with a probability of d_1 is performed. Operation in the spatio-temporal mixing MLP block can be represented as

$$\mathbf{U}_{*,i} = \mathbf{I}_{*,i} + \mathbf{w}_2\sigma(\mathbf{w}_1 N(\mathbf{I})_{*,i}), \tag{4}$$

$$\mathbf{O}_{j,*} = \mathbf{U}_{j,*} + \mathbf{w}_4\sigma(\mathbf{w}_3 N(\mathbf{U})_{j,*}), \tag{5}$$

where $\mathbf{I}, \mathbf{O}$ are input and output feature vectors, $\mathbf{w}_{1,\dots,4}$ are weight parameters of fully connected layers, and $i = 1, \dots, C$, $j = 1, \dots, 2S$. N is a layer normalisation function (Ba et al. 2016). σ is an GELU activation function (Hendrycks and Gimpel, 2016). $*$ indicates the row or the column vectors where operations are applied. Equation (4) is the calculation in the patch mixing MLP block. The calculation is performed for each column of $\mathbf{I}$. The number of hidden units of the first fully connected layer in this block h_S is used to control patch mixing. Equation (5) is the calculation in the patch-wise feature extraction MLP block. The calculation is performed for each rows of $\mathbf{U}$. The number of the hidden units of the first fully connected layer in this block h_C is used to control feature extraction from the patch.

After the processes of the spatio-temporal mixing blocks, feature values are mapped to a vector. It is combined with feature values calculated from the insertion length data $\mathbf{L}^{(t_c)}$ and then processed by some fully connected layers. Dropout with a probability of d_2 is performed here. The last layer outputs an estimated colon shape $\hat{\mathbf{Y}}^{(t_c)}$.

3. Experimental setup

We evaluated the colon shape estimation accuracy of the proposed method in a phantom study. We used a colon phantom (colonoscopy training model type I-B, Koken, Tokyo, Japan), a CT volume of the phantom, a colonoscope (CF-Q260AI, Olympus, Tokyo, Japan), an EM sensor (Aurora 5/6 DOF Shape Tool Type 1, NDI, Ontario, Canada), and a depth image sensor (Kinect v2, Microsoft, WA, USA). We measured colonoscope and colon shapes in our measurement environment shown in Figure 4. We assume the colonoscope tip is inserted up to the caecum when colonoscope tracking starts because physicians observe and treat the colon while retracting the colonoscope after its insertion up to the caecum. The colonoscope was moved from the caecum to the anus.

3.1. Colonoscope shape measurement

We measured colonoscope shapes using the EM sensor. The EM sensor is strap-shaped with six sensors at its tip and points along its strap-shaped body (one sensor is 6 DOF and remaining are 5 DOF). Each sensor provides a 3D position and a 3D/2D direction along the colonoscope by inserting the sensor into the colonoscope working channel. The measured colonoscope shape is $\mathbf{X}^{(t)} = \{\mathbf{p}_n^{(t)}, \mathbf{d}_n^{(t)}; n = 1, \dots, 6\}$ at every time t.

3.2. Colon shape measurement

We measured colon shapes from the colon phantom using the depth image sensor. Twelve position markers were attached to the surface of the colon phantom to scan its shape. The depth image sensor scanned colon shapes during the colonoscope insertions to the colon phantom. We automatically detect marker positions using YOLOv5 (Ultralytics 2022) from the scanned colour and depth images. Then, the detection results were manually corrected. The detected markers are described as $\mathbf{Y}^{(t)} = \{\mathbf{y}_m^{(t)}; m = 1, \dots, 12\}$ at every time t, used as colon

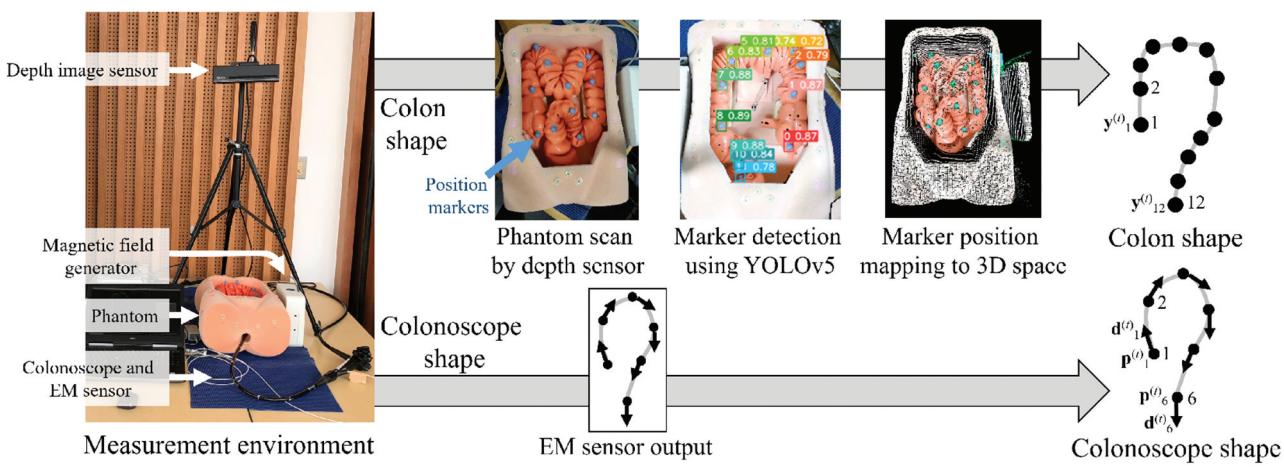

Figure 4. Colonoscope and colon shapes measurement environment. Processing flow of measured data is also illustrated.

Table 1. Mean and standard deviation of MED calculated in cross-validations performed using proposed and previous methods.

Method	MED (Mean ± S.D.) (mm)
KST-Mixer (Proposed)	**11.92 ± 1.75**
SEN (LSTM-based method) (Oda 2018a)	12.58 ± 2.08
Regression forests-based method (Oda 2018b)	13.08 ± 1.55

Table 2. Computation times of proposed and previous methods in estimation of one colon shape.

Method	Computation time (msec.)
KST-Mixer (Proposed)	7.3
SEN (LSTM-based method) (Oda 2018a)	2.9
Regression forests-based method (Oda 2018b)	8.9

shape. $\mathbf{y}_1^{(t)}$ and $\mathbf{y}_{12}^{(t)}$ correspond to the caecum and the anus positions, respectively.

3.3. Training and testing of KST-Mixer

We measured both $\mathbf{X}^{(t)}$ and $\mathbf{Y}^{(t)}$ during colonoscope insertions to the colon phantom. The shape recording frequency was six times per second. $\mathbf{X}^{(t)}$ and $\mathbf{Y}^{(t)}$ belong to the EM and depth image sensor coordinate systems. We registered them in the CT coordinate system using the iterative closest point (ICP) algorithm (Besl and McKay 1992) and manual registrations. Registered shape data was used to train and test the KST-Mixer. Parameters used in the trainings were: $\tau = 18$, $(s_1, s_2) = (6, 3)$, $b = 7$, $d_1 = 0.1$, $d_2 = 0.3$, $h_S = 64$, $h_C = 128$, 50 minibatch size, and 200 training epochs. Mean squared error was used as the loss function in training. We implemented the KST-Mixer using the Keras build in TensorFlow 2.4.0 running on a Windows PC equipped with a NVIDIA RTX A6000 GPU. The KST-Mixer used 2.5 GBytes of GPU memory in trainings.

In the test step, we provide colonoscope shapes for testing to the trained KST-Mixer. We obtain estimated colon shape $\hat{\mathbf{Y}}^{(t_c)}$ of current time t_c from it.

4. Experimental results

We measured colonoscope and colon shapes during eight colonoscope insertions and recorded 1,388 shape pairs. An engineering researcher operated the colonoscope. Leave-one-colonoscope-insertion-out cross validation was performed for evaluation. We used mean Euclidean distance (MED) (mm) between $\mathbf{Y}^{(t)}$ and $\hat{\mathbf{Y}}^{(t)}$ as an evaluation metric. It is defined as

$$\mathrm{MED} = \frac{1}{12T}\sum_{t=1}^{T}\sum_{m=1}^{12} ||\hat{\mathbf{Y}}_m^{(t)} - \mathbf{y}_m^{(t)}||. \tag{6}$$

We compared the MED of the proposed method with previous colon shape estimation methods, including SEN (LSTM-based method) (Oda 2018a) and regression forests-based method (Oda 2018b). Table 1 shows results of the comparison. The proposed method achieved the smallest MED among the methods. Statistical analysis of the results indicated that the proposed method significantly reduced MED compared to the SEN (Oda 2018a) ($p < 0.05$ with paired t-test of MED values). We compared computation times in estimations of one colon shape among these methods. The results are shown in Table 2. From the results, both the proposed and previous methods have real-time performances. Examples of colon shape estimation results are in Figure 5. The figure shows that

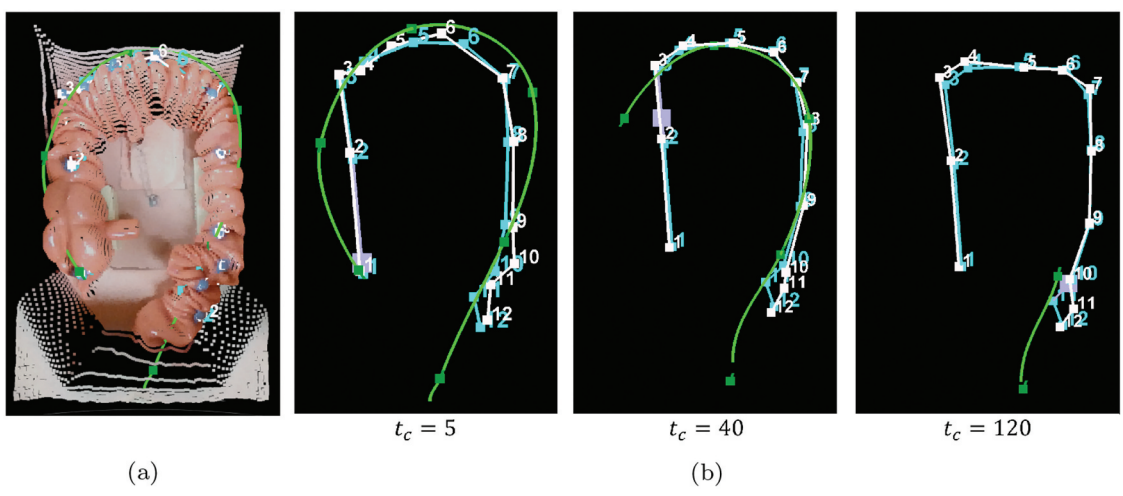

Figure 5. (A) Measurement results of depth image and EM sensors. (b) Colon shape estimation results of three frames. Colonoscope shapes are points on curved green lines. Colon shapes (ground truth) are white numbered points. Estimated colon shapes by proposed method are blue numbered points.

the differences between the ground truths and estimated colon shapes were tiny.

5. Discussion

The proposed KST-Mixer achieved the smallest error in the colon shape estimation among the previous methods. We designed the KST-Mixer to extract features from kinematic data using simple MLP blocks. Extracted features are then mixed in the spatio-temporal feature mixing MLP blocks to generate spatially and temporally global features. This structure is quite effective in processing time-series kinematic data because we achieved better estimation results than the LSTM-based method (Oda 2018a). The proposed method improves colonoscope tracking accuracy by accurately estimating deformed colon shape during colonoscope insertions. Furthermore, the computation time of the proposed method was short enough to be used in real-time applications.

The application of the proposed method is not limited to colon shape estimation alone. It can be applied to estimations of elastic organs in diagnosis and treatment. Organ shape estimation is essential in surgical assistances by computers. Accurate organ shape estimation contributes to the generation of real-time surgical navigation information and the automation of surgical assistance robots.

Although we have obtained promising results in colon shape estimation, many challenges are still remain for application of the proposed method to colonoscope tracking. Such challenges include (1) *collecting data containing variations of operators and colon shapes*, (2) *collecting in-vivo data*, (3) *development of intuitive visualisation method of tracking result*, and (4) *development of a colonoscope that have embedded EM sensors*. (1) *collecting data containing variations of operators and colon shapes* is necessary to improve robustness of the method to real situations. Colonoscope movements have variations among physicians depending on their years of experience. Furthermore, colon shapes also have variations among patients. Colon and colonoscope shapes data containing such variations is necessary to achieve better estimation model. We will measure shape data under operations of colonoscope by physicians of various years of experience. We also measure shape data using many colon phantoms and 3D printed phantoms that have variation of the shapes. (2) *collecting in-vivo data* is necessary to improve the proposed method from phantom level to clinically applicable level. (3) *development of intuitive visualisation method of tracking result* is required that presents deformed colon shapes in real-time during colonoscope insertions. Such visualisation helps physicians to understand how the colonoscope travelling in the colon. (4) *development of a colonoscope that have embedded EM sensors* is required to perform tracking in clinical situations.

6. Conclusions

This paper proposed a colon shape estimation method using the KST-Mixer from kinematic data. The KST-Mixer extracts kinematic features and mixes them along the spatial and temporal axes in multiple MLP blocks. We evaluated the method's estimation accuracy in the colon shape estimation from colonoscope shapes in the phantom study. The proposed KST-Mixer achieved the smallest estimation error in the comparative experiments. Future work includes improvement of the data number using other phantoms, evaluating the method using shape data measured during colonoscope operations by physicians, application to colonoscope tracking, and application to the human colon.

Note

1. Code and data of the proposed method are available at: https://github.com/modafone/kst-mixer.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Parts of this research were supported by the MEXT/JSPS KAKENHI Grant Numbers 21K12723, 17H00867 and the JSPS Bilateral International Collaboration Grants.

Notes on contributors

Masahiro Oda (Ph.D. Nagoya University '09) is an associate professor at the Nagoya University, Japan. His research is in medical image processing and mainly concerns computer-aided diagnosis and computer-assisted surgery in many application areas. He has (co-)authored over 200 peer-reviewed full papers in international conferences and journals and is the recipient of the Joint MICCAI 2022 Workshop on AE-CAI, CARE, and OR 2.0 Outstanding Paper Award and the RSNA Certificate of Merit (2009, 2014, 2019) Awards.

Kazuhiro Furukawa (M.D., Ph.D. Nagoya University '11) is an associate professor/lecturer at the Nagoya University Graduate School of Medicine, Japan. His research is in endoscopic diagnosis and treatment and computer-aided diagnosis.

Nassir Navab is a Full Professor and Director of the Laboratory for Computer-Aided Medical Procedures (CAMP) at the Technical University of Munich (TUM) and Adjunct Professor at Laboratory for Computational Sensing and Robotics (LCSR) at Johns Hopkins University (JHU). He is a fellow of MICCAI society and received its prestigious Enduring Impact Award in 2021. He was also the recepient of 10 years lasting impact award of IEEE ISMAR in 2015, SMIT Society Technology award in 2010 and Siemens Inventor of the year award in 2001. He completed his PhD at INRIA and University of Paris XI, France, in January 1993, before enjoying two years of post-doctoral fellowship at MIT Media Laboratory and nine years of expeirence at Siemens Corporate Research in Princeton, USA.He has acted as a member of the board of directors of the MICCAI Society, 2007-2012 and 2014-2017, and serves on the Steering committee of the IEEE Symposium on Mixed and Augmented Reality (ISMAR) and Information Processing in Computer-Assisted Interventions (IPCAI). He is the author of hundreds of peer-reviewed scientific papers, with more than 52,000 citations and an h-index of 101 as of April 2022. He is the author of more than thirty awarded papers including 11 at MICCAI, 5 at IPCAI, 2 at IPMI and 3 at IEEE ISMAR. He is the inventor of 51 granted US patents and more than 60 International ones. His current research interests include medical augmented reality, computer-assisted surgery, medical robotics, computer vision and machine learning.

Kensaku Mori has received the B.S degree in electronics engineering, the M. S degree in Information Engineering, and the Ph.D in Information Engineering from Nagoya University, Japan, under the supervision of Prof.

Jun-ichiro Toriwaki, in 1992, 1994 and 1996 respectively. He was a research fellow of the Japanese Society for the Promotion of Science (JSPS) from 1994 to 1997, Research Associate at Department of Computational Science and Engineering, Nagoya University from 1997 to 2000, an assistant professor in 2000. He was Associate Professor at the Research Center for Advance Waste and Emission Management of Nagoya University from 2001 to 2003. He was also Visiting Associate Professor at Department of Neurosurgery, Stanford University from 2001 to 2002. He was Associate Professor of the Graduate School of Information Science, Nagoya University from 2003 to 2009. He has been Professor of Information and Communications Headquarters, Nagoya University since 2009. In 2016, he started his services as the Directory of Information Technology Center, Nagoya University. In 2017, he was appointed as the Vice Chair of Information and Information and Communication, Nagoya University. He has been a Professor of Graduate School of Informatics, Nagoya University since 2017. His current research interests include three-dimensional image processing, computer graphics, virtual reality, and their applications to medical image. Computer aided diagnosis and surgery are his major active research fields. Recently AI-based colonoscope diagnosis system developed in his laboratory has obtained an official certificate as a medical device. He has received many awards from several institutes includes the Awards for Science and Technology, the Ministry of Education, Culture, Sports, the best paper awards from the Japanese Society of Medical Imaging Technology. He served as the general chair of the 16th international conference of Medical Image Computing and Computer Assisted Intervention (MICCAI 2013). He was elected as the President of the International Society of Computer Aided Surgery (ISCAS) in 2018. He was the Chair of SPIE Medical Imaging CAD conference in 2018 and 2019. He is a member of the Institute of Electronics, Information and Communication Engineering of Japan, Japanese Society of Biomedical Engineering, and Japanese Society of Medical Imaging Technology, MICCAI, IEEE, and SPIE. He is an Affiliate member of the Science Council of Japan.

ORCID

Masahiro Oda http://orcid.org/0000-0001-7714-422X
Kazuhiro Furukawa http://orcid.org/0000-0003-0980-9095
Nassir Navab http://orcid.org/0000-0002-6032-5611
Kensaku Mori http://orcid.org/0000-0002-0100-4797

References

Ba JL, Kiros JR, Hinton GE 2016. Layer normalization. arXiv:1607.06450.

Banach A, King F, Masaki F, Tsukada H, Hata N. 2021. Visually navigated bronchoscopy using three cycle-consistent generative adversarial network for depth estimation. Med Image Anal. 73:102164. doi:10.1016/j.media.2021.102164.

Besl PJ, McKay ND. 1992. A method for registration of 3-D shapes. IEEE Trans Pattern Anal Mach Intell. 14(2):239–256. doi:10.1109/34.121791.

Ching LY, Moller K, Suthakorn J. 2010. Non-radiological colonoscope tracking image guided colonoscopy using commercially available electromagnetic tracking system. In 2010 IEEE Conference on Robotics, Automation and Mechatronics; June 28–30. Singapore.

Deguchi D, Mori K, Feuerstein M, Kitasaka T, CR M Jr, Suenaga Y, Takabatake H, Mori M, Natori H. 2009. Selective image similarity measure for bronchoscope tracking based on image registration. Med Image Anal. 3(14):621–633. doi:10.1016/j.media.2009.06.001.

Deligianni F, Chung A, Yang GZ. 2005. Predictive camera tracking for bronchoscope simulation with CONDensation. In: Duncan J Gerig G, editors. MICCAI 2005. Medical image computing and computer-assisted intervention. Palm Springs (CA): LNCS 3749; pp. 910–916, Oct 26–30.

Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, et al. 2021. An image is worth 16x16 words: transformers for image recognition at scale. In International Conference on Learning Representations; May 3–7. Vienna, Austria.

Fukuzawa M, Uematsu J, Kono S, Suzuki S, Sato T, Yagi N, Tsuji Y, Yagi K, Kusano C, Gotoda T, et al. 2015. Clinical impact of endoscopy position detection unit (UPD-3) for a non-sedated colonoscopy. World J Gastroenterol. 21(16):4903–4910. doi:10.3748/wjg.v21.i16.4903.

Gildea TR, Mazzone PJ, Karnak D, Meziane M, Mehta A. 2006. Electromagnetic navigation diagnostic bronchoscopy: a prospective study. Am J Respir Crit Care Med. 174(9):982–989. doi:10.1164/rccm.200603-344OC.

Hendrycks D, Gimpel K. 2016. Gaussian error linear units (GELUs). arXiv:1606.08415.

Hochreiter S, Schmidhuber J. 1997. Long short-term memory. Neural Comput. 9(8):1735–1780. doi:10.1162/neco.1997.9.8.1735.

Liu J, Subramanian KR, Yoo TS. 2013. An optical flow approach to tracking colonoscopy video. Comput Med Imaging Graph. 37(3):207–223. doi:10.1016/j.compmedimag.2013.01.010.

Luo X, Wan Y, He X, Mori K. 2015. Observation-driven adaptive differential evolution and its application to accurate and smooth bronchoscope three-dimensional motion tracking. Med Image Anal. 24(1):282–296. doi:10.1016/j.media.2015.01.002.

Mori K, Deguchi D, Akiyama K, Kitasaka T, CR M Jr, Suenaga Y, Takabatake H, Mori M, Natori H. 2005. Hybrid bronchoscope tracking using a magnetic tracking sensor and image registration. In: Duncan J Gerig G, editors. MICCAI 2005. Medical image computing and computer-assisted intervention. Palm Springs (CA): LNCS 3750; pp. 543–550, Oct 26–30.

Oda M, Kitasaka T, Furukawa K, Miyahara R, Hirooka Y, Goto H, Navab N, Mori K. 2018b. Machine learning-based colon deformation estimation method for colonscope tracking. In: Baowei F, Robert JW III, editors. Proceeding of SPIE medical imaging; Feb 10–15; Houston (TX); p. 10576:1057619.

Oda M, Kondo H, Kitasaka T, Furukawa K, Miyahara R, Hirooka Y, Goto H, Navab N, Mori K. 2017. Robust colonoscope tracking method for colon deformations utilizing coarse-to-fine correspondence findings. Int J Comput Assist Radiol Surg. 12(1):39–50. doi:10.1007/s11548-016-1456-6.

Oda M, Roth H, Kitasaka T, Furukawa K, Miyahara R, Hirooka Y, Goto H, Navab N, Mori K. 2018a Sep 16–20. Colon shape estimation method for colonoscope tracking using recurrent neural networks. In: Frangi A, Schnabel J, Davatzikos C, Alberola-López C, and Fichtinger G, editors. MICCAI 2018. Medical image computing and computer assisted intervention. Granada, Spain, LNCS 11073: Springer Cham; pp. 176–184.

Peters T, Cleary K. 2008. Image-guided interventions: technology and applications. New York (NY): Springer.

Rai L, Helferty JP, Higgins WE. 2008. Combined video tracking and image-video registration for continuous bronchoscopic guidance. Int J Comput Assist Radiol Surg. 3:315–329. doi:10.1007/s11548-008-0241-6.

Schwarz Y, Greif J, Becker HD, Ernst A, Mehta A. 2006. Real-time electromagnetic navigation bronchoscopy to peripheral lung lesions using overlaid CT images: the first human study. Chest. 129(4):988–994. doi:10.1378/chest.129.4.988.

Tolstikhin I, Houlsby N, Kolesnikov A, Beyer L, Zhai X, Unterthiner T, Yung J, Steiner A, Keysers D, Uszkoreit J, et al. 2021. MLP-Mixer: An all-MLP architecture for vision. arXiv:2105.01601.

Ultralytics YOLOv5. 2022. [accessed 2022 Feb 27]. https://github.com/ultralytics/yolov5.

Visentini-Scarzanella M, Sugiura T, Kaneko T, Koto S. 2017. Deep monocular 3D reconstruction for assisted navigation in bronchoscopy. Int J Comput Assist Radiol Surg. 12:1089–1099. doi:10.1007/s11548-017-1609-2.

Wang C, Hayashi Y, Oda M, Kitasaka T, Takabatake H, Mori M, Honma H, Natori H, Mori K. 2021. Depth-based branching level estimation for bronchoscopic navigation. Int J Comput Assist Radiol Surg. 16:1795–1804. doi:10.1007/s11548-021-02460-8.

Yao H, Stidham RW, Gao Z, Gryak J, Najarian K. 2021. Motion-based camera localization system in colonoscopy videos. Med Image Anal. 73:102180. doi:10.1016/j.media.2021.102180.