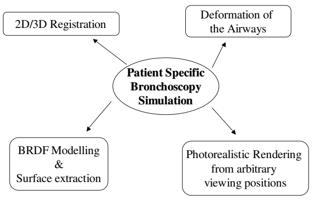
Predictive Camera Tracking for Bronchoscope Simulation with CONDensation

Fani Deligianni, Adrian Chung and Guang-Zhong Yang

Royal Society/Wolfson Foundation MIC Laboratory

Introduction



Flexible fiber-optic bronchoscopy is performed on patients who are fully awake or with light conscious sedation. The procedure can entail considerable discomfort if it is not handled properly. Training according to the classical apprenticeship scheme can result in prolonged surgical procedures with increased patient discomfort and a potential risk for further complications. Computer simulation, particularly the reliance on patient specific data for building anatomical models both in terms of biomechanical fidelity and photorealism has attracted extensive interests in recent years. However, accurate 2D/3D registration is important for recovering the pose of the camera in video bronchoscope sequences. Since bronchoscope images only provide localized views of the inner lumen, image-based technique cannot guarantee the convergence of the registration algorithm.

This work exploits the use of temporal information to minimize the ambiguity of camera motion tracking in bronchoscope simulation. The condensation algorithm has been used to track the tip of the bronchoscope in presence of semi-occlusion and image artefacts. Modular training provides a systematic learning procedure in order data from different subjects to be integrated and create a dynamic model that accommodates the learnt behavior. Experimental in-vivo results demonstrate a significant improvement in tracking accuracy especially in cases where there is airway semi-occlusion or airway deformation.

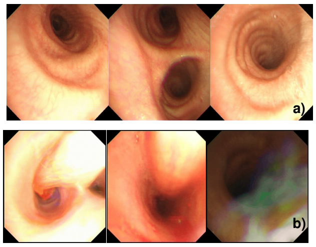


Fig. 1: a) Bronchoscope views provide only localised view of the airways, b) Blood, deformation and mucosa are common artefacts in bronchoscope video

Temporal Tracking

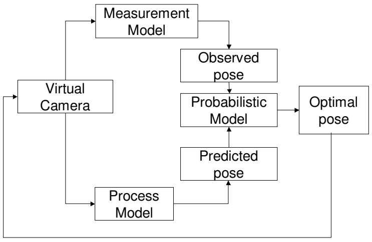


Fig. 2: The probabilistic framework that combines the prediction of a state-space representation and the observations of a measurement model.

To incorporate temporal constraints in the 2D/3D registration scheme, the probabilistic framework in fig2. was used. According to this the optimal statistical pose given the imprecise measurements of the 2D/3D registration and the process model can be estimated in an optimal statistical fashion.

For video bronchoscope navigation, Naga *et al* used Kalman filtering to increase the speed and accuracy of the registration algorithm [1]. Kalman filter, however, is generally restricted to situations where the probability distribution of the state variables is unimodal. In bronchoscopy, tissue deformation, inter-reflection and view dependent specularity due to mucosa can limit the accuracy of image-based algorithms. The resultant probability density function of the state vector is typically multi-modal. Therefore, the observation probabilistic model cannot be approximated as a Gaussian distribution.

Alternatively, particle filtering uses a stochastic approach that has no restrictions on the measurement model and the distribution of error sources, [2].

In order to construct a motion model for the endoscope camera that moves freely in the 3D tracheo-bronchial tree, a second order auto-regressive (AR) model is used. For a Kth-order auto-regressive model, the following equation applies:

$$x_t = \sum_{k=1}^{K} A_k x_{t-k} + d + Bwt$$

Where A_k represents the collection of damped harmonic oscillators associated with vibrational modes, d is a drift per unit time, and w the white noise with covariance coefficient B.

This is in contrast to the 'constant acceleration' model, [1], which implies that the camera acceleration is expected to be constant during bronchoscope tracking. In this study, the AR model takes into account that during bronchoscope navigation, motion occurs within a bounded area, and a rapidly moving camera is expected to slow down or change in direction rather than accelerate further.

Modular Training

In practice, it is possible to build a tracking model by approximating its general behavior to intuitive expectations of the observed motion. However a hand-built model is not appropriate in this study due to the high-dimensionality and complex motion involved. For the purpose of bronchoscope simulation, it is more meaningful to collect several training sets from the same as well as different operators and construct a representative dynamic model. This is accomplished by modular training [3].

To this end, the auto-correlation coefficients of each training set have been calculated individually and then combined in a linear fashion.

Since each of these dynamic systems may have a different mean value, the use of pre-estimated mean value of the system can result in a prediction strongly biased, fig3-(a). In this study, the mean value of the system is estimated on-line as part of the state vector.

Fig3-(b) demonstrates the effectiveness of the training process involved in this study. The ground truth data of the camera pose from four different patients have been used to train the AR model. The performance of the trained model was evaluated on the fifth patient data. The Euclidean distance between the first and subsequent camera positions predicted from the condensation algorithm was used for error analysis.

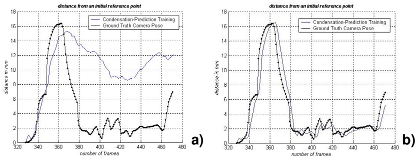


Fig. 3: Assessment of the accuracy of the training model and the effect of excluding a), and including b), the mean value estimation of the system as part of the state vector.

Results

For *in vivo* validation, bronchoscopy examination was performed in five patients. Effective image area used is 454×487 pixels. The CT images were acquired from the Siemens Somaton Volume Zoom four-channel multi-detector CT scanner with a slice width of 3mm and collimation of 1mm. The video frames were preprocessed in order to alleviate the effects of interlacing, lens distortion and unnecessary texture information.

It is evident that the 2D/3D registration accuracy has been increased significantly by the use of the proposed predictive tracking algorithm. More notably, the method permits more stable tracking results in patients where image artifact (*e.g.* partial occlusion of the images due to mucosa or bleeding) and sudden airway deformation due to coughing can introduce large propagational errors to the original registration technique. In Fig. 4 we demonstrate the extent of this effect on the actual bronchoscope views. The left column shows the original frames from the bronchoscope video, whereas the middle and right columns are the virtual views of the 3D model by using registration without and with predictive camera pose tracking.

In Fig. 5(a-b), the Euclidean distance between the first and subsequent camera positions predicted from the condensation algorithm was used for error analysis. Similar analysis for the error in orientation was also performed.

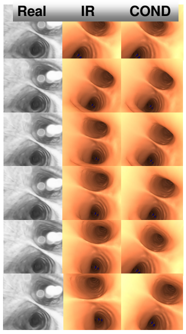


Fig. 4: Qualitative Results

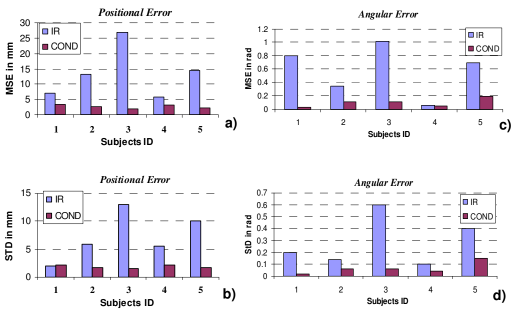
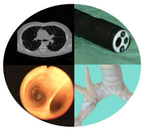


Fig. 5: Numerical Assessment of the accuracy of the registration process

Discussion



Our results demonstrate a significant improvement in tracking accuracy especially in cases where there is airway deformation and image artifacts. The use of AR model based on the principles of the maximum likelihood learning and modular learning has facilitated the incorporation of multiple sequences from different patients. The proposed method can be further extended to multi-class motion description such that the dynamic behavior of camera navigation in different parts of the tracheo-bronchial tree can be incorporated.

[1] Naga, J., Mori, K., et al., *MICCAI*, 2004, vol.2, 551-558.
[2] Isard, M., Blake, A.,, *International Journal of Computer Vision*, 1998, vol.29(1), 5-28.
[3] North, B., Blake, A. et al., *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, vol.22(9), 1016-1034.

For more information visit our webpage: http://vip.doc.ic.ac.uk/VISaVE/