

Efficient Seeding and Defragmentation of Curvature Streamlines for Colonic Polyp Detection

Lingxiao Zhao^a, Charl P. Botha^a, Roel Truyen^b, Frans M. Vos^a and Frits H. Post^a

^aDelft University of Technology, Mekelweg 4, Delft, the Netherlands;

^bPhilips Health Care, Veenpluis 6, Best, the Netherlands

ABSTRACT

Many computer aided diagnosis (CAD) schemes have been developed for colon cancer detection using Virtual Colonoscopy (VC). In earlier work, we developed an automatic polyp detection method integrating flow visualization techniques, that forms part of the CAD functionality of an existing Virtual Colonoscopy pipeline. Curvature streamlines were used to characterize polyp surface shape. Features derived from curvature streamlines correlated highly with true polyp detections. During testing with a large number of patient data sets, we found that the correlation between streamline features and true polyps could be affected by noise and our streamline generation technique. The seeding and spacing constraints and CT noise could lead to streamline fragmentation, which reduced the discriminating power of our streamline features.

In this paper, we present two major improvements of our curvature streamline generation. First, we adapted our streamline seeding strategy to the local surface properties and made the streamline generation faster. It generates a significantly smaller number of seeds but still results in a comparable and suitable streamline distribution. Second, based on our observation that longer streamlines are better surface shape descriptors, we improved our streamline tracing algorithm to produce longer streamlines. Our improved techniques are more efficient and also guide the streamline geometry to correspond better to colonic surface shape. These two adaptations support a robust and high correlation between our streamline features and true positive detections and lead to better polyp detection results.

Keywords: Visualization, Virtual Colonoscopy, Automatic Polyp Detection, Curvature Streamlines

1. INTRODUCTION

In recent years, significant progress has been made for using computer technology in clinical examinations. Computer imaging and visualization techniques are now widely used in visual diagnosis. Virtual colonoscopy is a patient-friendly technique that enables effective screening of the large intestines for colorectal polyps, which are well-known precursors of colon cancer.^{1,2} The interior colon wall is extracted from the patient data set, usually CT scans, and then rendered to allow for visual inspection by a radiologist. Due to the complexity of the colon wall, a thorough examination of the entire colon is very time consuming. In order to achieve high efficiency and effectiveness in virtual colonoscopy, computer-aided techniques are developed to pre-detect and visualize polyp candidates.

A large number of automatic polyp detection schemes have been developed as helpful additions to the virtual colonoscopy pipeline.³⁻¹⁰ Many of them are based on 3D surface curvature measures, i.e. scalar principal curvature values. An early example of the use of curvature concepts from differential geometry is due to Summers et al.¹¹ Yoshida et al.⁴ derived two geometric features, namely the shape index and volumetric curvedness, to characterize polyps, folds and colonic walls at voxels and indicate polyp candidates by a hysteresis thresholding.

Further author information: (Send correspondence to L. Zhao)

L. Zhao: E-mail: zlx@graphics.tudelft.nl, Telephone: +31 (0)1 52 78 46 98

C.P. Botha: E-mail: c.p.botha@tudelft.nl, Telephone: +31 (0)1 52 78 71 41

R. Truyen: E-mail: roel.truyen@philips.com, Telephone: +31 (0)4 02 76 26 74

F.M. Vos: E-mail: f.m.vos@tudelft.nl, Telephone: +31 (0)1 52 78 71 33

F.H. Post: E-mail: f.h.post@ewi.tudelft.nl, Telephone: +31 (0)1 52 78 25 28

Huang et al.⁶ developed a two-stage curvature estimation method and derived the sphericity index as a shape-based discriminator to find polyp candidates. Accurate and noise-insensitive curvature estimation is essential for any such scheme. Only using scalar curvature measures can result in many false-positive detections.

Other work has been done trying to find complementary solutions instead of only relying on curvature and corresponding derivatives. Hong et al.⁹ presented a CAD polyp detection pipeline based on an integration of texture and shape analysis with volume rendering and conformal 3D colonic surface flattening. Kiss et al.⁷ fitted a spherical model to the colonic surface normals and considered convex regions as suspected polyps. Van Wijk et al.⁸ considered that polyp shape and size vary greatly and are often far from the widely-used ellipsoidal polyp model, showing many irregularities. They proposed a shape and size invariant approach, in which the local colonic surface is deformed until the minimum curvature is close to zero. The amount of deformation is computed as a “protrusion” measure and used for detecting polyp candidates.

Our previous work¹⁰ explored the potential of surface principal curvature directions, which are represented as two orthogonal vector fields on the surface. We integrated flow visualization techniques in a complete polyp detection protocol. Useful features were calculated on curvature streamlines and used to reduce the number of false-positive detections, which are result of a pre-detection stage. However, during further testing with a larger number of patient data sets, we found that the correlation between our curvature streamline features and true polyp detections could be affected by CT noise and the characteristics of our streamline generation scheme. In this paper, we present two major improvements of our curvature streamline generation to strengthen the discriminating power of curvature streamline features. We adapt our streamline seeding strategy to the local surface geometry so that the integration of curvature streamlines becomes more efficient. Our experiments also suggested that longer curvature streamlines are more suitable for feature calculation. We improved our streamline tracing algorithm to produce longer streamlines that better correspond to colonic surface shape. These two adaptations support a better polyp detection result using curvature streamline features.

The rest of the paper is organized as follows: a more detailed discussion of our previous work is given in section 2. We present our surface-adaptive streamline seeding strategy and improved longer streamline tracing algorithm in section 3. In section 4, we give an analysis for using improved streamline features in automatic polyp detection. Finally, we summarize our work with some conclusions and propose for future work in section 5.

2. BACKGROUND

The proposed CAD process in our previous work¹⁰ consists of three steps: pre-detection of polyp candidates, candidate selection and finally enhanced visualization. Polyp candidates were pre-detected using the approach developed by Yoshida et al.⁴ based on the shape index and volumetric curvedness. These pre-detections included all true polyps as documented with the data sets used and a large number of false-positive detections. We made use of curvature streamlines as additional information to avoid false-positive detections without losing sensitivity.

The idea was based on a generalization of the polyp surface. We modeled a polyp as composed of the top part and the bottom part, namely the cap and the neck. Visualization of principal curvature direction fields around the polyp showed a circular pattern in the maximum curvature direction and a focus pattern in the minimum curvature direction, with consistently outward-pointing surface normals. To exploit such polyp characteristics, we distributed curvature streamlines in the neighborhood per polyp candidate. We developed techniques that trace surface-constrained streamlines for both explicit and implicit surface representations. In order to cover necessary surface shape information, curvature streamlines were distributed using a curvature-controlled spacing approach. Distance between neighboring streamlines is dependent on local surface curvatures. Geometric properties of curvature streamlines are good descriptors of local surface shape. Circular or closed streamlines that follow maximum curvature direction were inspected around the polyp neck area. Therefore such patterns could be used as indicators of colonic polyps. The winding angle and mean radius were calculated on curvature streamlines as features used in our polyp candidate selection process. The winding angle is defined as the cumulative change of direction along a streamline (fig. 1). Mean radius is a coarse measure of the polyp size. As the indicator of a polyp instance, a closed streamline has a winding angle of at least 2π . We performed an experimental study of these features on 5 patient data sets, including 50 true polyps and 281 false-positive

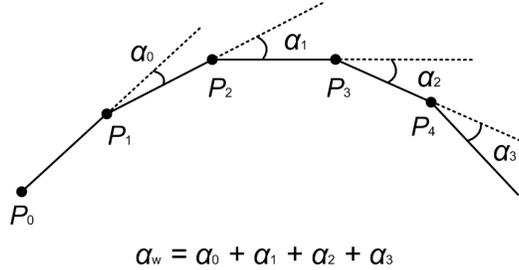


Figure 1. Winding angle is the cumulative change of direction of a streamline.

detections. The result showed a high correlation between curvature streamline features and true polyp detections. We used this correlation to separate true and false positive detections.

We performed further test with a larger number of data sets and found that this correlation sometimes could be significantly affected by noise and the characteristics of our curvature streamline generation scheme. Due to the thresholding for streamline spacing and inevitable data noise, many short streamline fragments were generated. Importantly, expected circular or closed streamlines could be fragmented. This streamline fragmentation made the winding angle feature less robust particularly for larger polyps. Often short streamline segments were distributed around the polyp neck area, and none of them had a winding angle of at least 2π . In this case, the enclosed polyp could not be characterized by the winding angle feature because of these short streamline fragments. Long streamlines are preferable for obtaining high values of winding angle. Therefore the streamline spacing technique should promote the generation of long streamlines.

In addition, our streamline seeding strategy was inefficient due to its greedy nature, in which a pair of seed points was placed on both sides of the streamline for each integration step. This fact increased the computational complexity of our method and slowed down the system.

An improvement of our curvature streamline based CAD system is required to achieve a low number of false positives, but without losing sensitivity. In the next section, we present two major improvements to increase the efficiency of our streamline computation and make our streamline features more robust.

3. METHODS

3.1 Adaptive curvature streamline seeding

The greedy streamline seeding strategy¹² used in our previous work generated a large number of seed points (Fig. 2(a)). For each integration step of streamline tracing, a pair of potential new seed points was placed perpendicular to the local streamline segment on both sides. All these seed points were put into a seed queue for starting new streamlines. After a new streamline was generated, seed points too close to existing streamlines were removed from the queue. Thus many potential seeds were calculated but not used for streamline tracing and it was always needed to validate these redundant seeds and remove them if necessary. These seeds were also very close to each other. Only a small number of them were actually used as starting points for tracing new streamlines. This significantly increased the computational complexity of our streamline generation.

In order to increase the efficiency of our streamline seeding strategy, the number of redundant seed points calculated on the fly can be reduced. In our streamline spacing, local surface curvature magnitudes are used to estimate the optimal distance between two neighboring streamlines. We adopt this idea to improve the streamline seeding. The density of potential seed points is adapted to the local surface shape, i.e. controlled by the local curvature values. Seed points are only calculated if they have a high probability to be actually used for tracing new streamlines.

This surface-adaptive streamline seeding is explained in Fig. 2(b). When tracing a streamline, integration steps are selected to place seed pairs. After a pair of seeds is placed at an integration step A , we calculate a

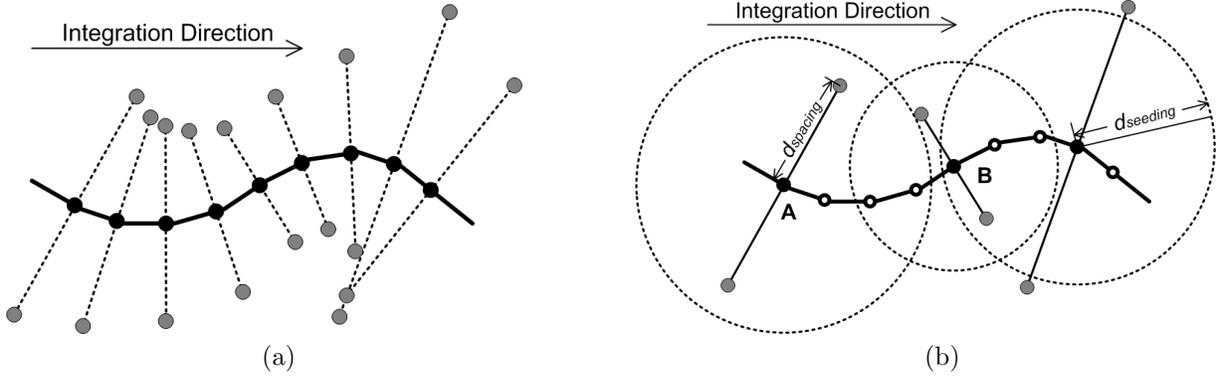


Figure 2. Curvature streamline seeding: (a) Greedy streamline seeding generated potential seed points for each streamline integration step, many redundant seed points were not actually used for streamline generation; (b) Adaptive streamline seeding improves the efficiency of streamline computation, the density of seed points is controlled by local surface geometry. Gray dots indicate seed points, black dots are integration steps where seeds are placed, and hollow white dots are integration steps where seeds are not placed.

seeding distance using:

$$d_{seeding} = \sqrt{\varepsilon_s \left(\frac{2}{|k|} + \varepsilon_s \right)} \quad (1)$$

where ε_s is a user-specified error boundary to constrain the local surface approximation and $|k|$ is the controlling curvature value that corresponds to the curvature direction that the streamline follows. This equation is similar to the one that we used in our previous work¹⁰ for the calculation of optimal streamline spacing distance. The only difference is that we are now using the same principal curvature that the streamline follows. With this seeding distance, we continue to integrate the streamline until we arrive at integration step B that is far enough, i.e. the Euclidean distance between integration A and B is no less than $d_{seeding}$. Then we place another pair of seeds and continue. We set a maximum seeding distance to ensure that necessary seed points can always be calculated. Although not all adaptively calculated seed points are actually used for streamline generation, the number of redundant seeds is significantly reduced.

3.2 Curvature streamline defragmentation

Our work employs flow visualization techniques, more specifically streamlines in vector fields, for polyp surface characterization. The streamline generation in our work is based on the algorithm for generating evenly-spaced streamlines proposed by Jobard and Lefer.¹² A problem introduced is the stringent thresholding for streamline spacing distance. This can result in many short streamlines, which are less meaningful for characteristic features. To promote the computation of many long streamlines, Jobard and Lefer suggested to reduce their fixed threshold by half for spacing.

We adapted the streamline spacing threshold to local surface curvatures and followed Jobard and Lefer’s suggestion. However our observation showed that curvature streamlines were often terminated when they could have continued. This fact made our curvature streamline features less robust for polyp candidate selection. To avoid this, we develop a streamline defragmentation approach and obtain sufficiently long streamlines.

Our streamline defragmentation is based on the correlation between two streamlines. Streamline correlation is a similarity measure that quantifies the probability that two streamline segments belong to the same streamline. An application of streamline similarity is streamline clustering. Many streamline clustering methods have been presented in the literature, particularly for fiber tracks in diffusion tensor imaging (DTI).¹³ Brun et al.¹⁴ defined a DTI fiber similarity based on the Euclidean distance between end points of fibers. Corouge et al.¹⁵ computed point pairs of two streamlines and used them to define fiber distance.

Our streamline defragmentation approach consists of two stages. In the first stage, we trace streamlines that are as long as possible so that they capture more surface information. In the second stage, neighboring fragments

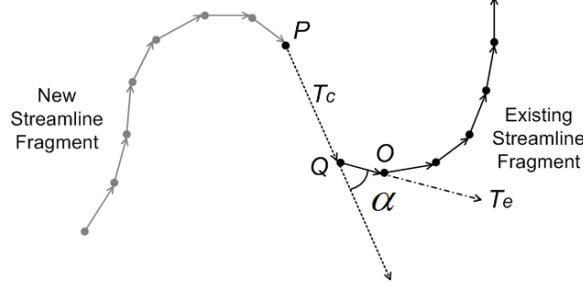


Figure 3. Connection of a newly integrated streamline with an existing streamline at the end points. The new streamline is shown in gray, the existing streamline is shown in black. Dotted line indicates a tentative connection, dash-dotted line indicates the direction at the end of the existing streamline.

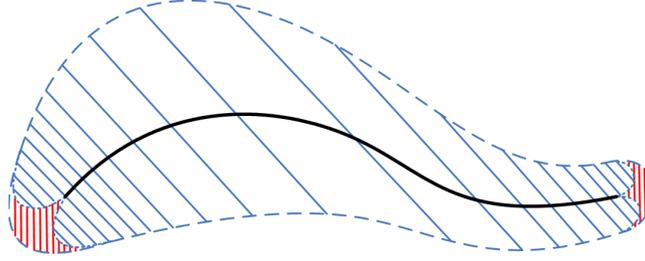


Figure 4. Correlation-based streamline spacing: For a streamline indicated by the central black line, the area filled with diagonal lines is the correlation-based spacing region. The area filled with vertical lines demonstrates the difference from purely curvature-controlled spacing.

with endpoints that are very close are connected, only if the direction of the connecting line segment is similar to the respective directions of the two streamline fragments at their connecting endpoints. Each connection combines two streamline fragments into a longer streamline fragment.

In our previous work, streamline spacing was dependent on local surface curvature. In order to generate longer streamlines, spacing is now also dependent on the correlation between streamline directions at their endpoints. We use a correlation-based streamline spacing distance to determine when and where a streamline stops. We compute the correlation between a newly integrated streamline point P and an existing streamline end point Q by (Fig. 3):

$$\begin{cases} R = T_c \cdot T_e = \cos \alpha, \\ \text{if } R < 0 \text{ then } R = 0. \end{cases} \quad (2)$$

with:

$$\begin{cases} T_c = \frac{\overrightarrow{PQ}}{\|PQ\|}, \\ T_e = \frac{\overrightarrow{QO}}{\|QO\|}. \end{cases} \quad (3)$$

where O is the previous streamline point of Q . For the correlation between P and an existing streamline point that is not an end point, R is always equal to 0. Then we compute the local spacing distance between two streamlines by:

$$d_{spacing} = l_s \sqrt{1 - R^2} \quad (4)$$

where l_s is the curvature-controlled streamline spacing distance defined in our previous work.¹⁰

Fig. 4 shows the spacing region of a streamline. New streamline integration stops at the region border. If a new streamline integration step is approaching the end of an existing streamline and has a high correlation with it locally, they can be very close at their end points. For streamline fragments that are arbitrarily close at

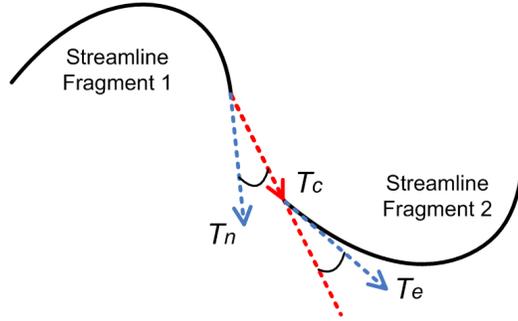


Figure 5. When two streamline fragments are very close and locally propagate almost in the same direction, we assume that they belong to one single streamline and connect them.

their end points, we need to examine both the tentative connection direction and the two streamline propagating directions at corresponding end points. We define a measure:

$$\eta = \sqrt{\frac{(T_n \cdot T_c)^2 + (T_c \cdot T_e)^2}{2}} \quad (5)$$

to quantify the probability that two streamline fragments belong to one single streamline (Fig. 5). T_n and T_e are local directions of the two streamline fragments and T_c is the tentative connection direction. All vectors are of unit length. We define a tolerance for this quantity by a suitable threshold and connect streamline fragments that fulfil the criteria.

4. EXPERIMENTAL RESULTS

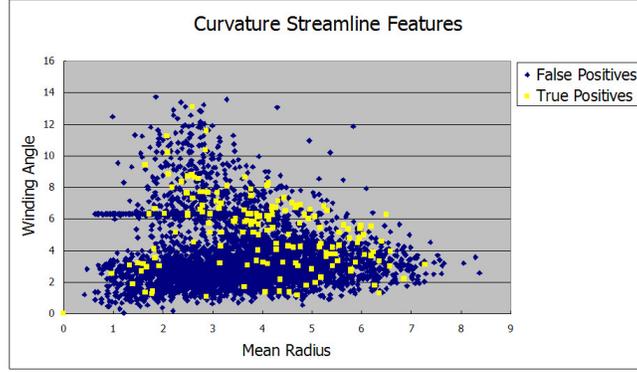
We have evaluated the effects of streamline defragmentation on our polyp detection strategy. We first compare the new surface-adaptive streamline seeding with the previous greedy strategy and show how the efficiency of curvature streamline computation is improved. We then demonstrate that the discriminating power of our curvature streamline features for characterizing true polyp detections is significantly strengthened by the streamline defragmentation. We also illustrate how our streamline defragmentation enhances the visualization of polyp surface shape.

4.1 Streamline computation efficiency

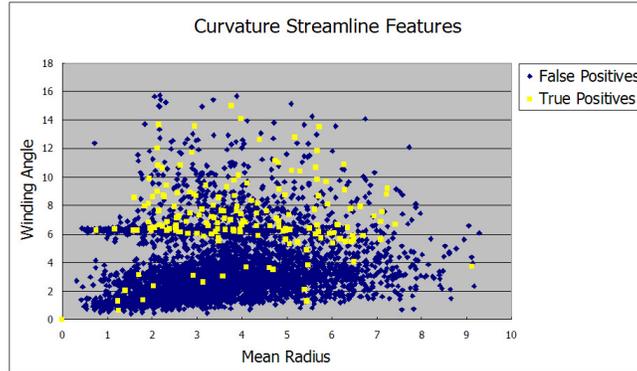
As in earlier work, streamline generation and feature calculation are integrated in a complete polyp detection protocol. Polyp candidates are pre-detected using an existing detection scheme.⁴ Many false-positive detections are included and all true-positive detections are annotated by radiologists. Colonic surfaces are represented as implicit iso-surfaces embedded in 3D volume data and rendered using volume ray casting technique. Surface-constrained curvature streamlines are directly traced in the 3D volume and distributed over a small neighborhood surrounding the polyp candidate. Linear interpolation and first-order integration were used in streamline tracing.

Table 1. Comparison of streamline seeding strategies: using surface-adaptive seeding, curvature streamline computation is accelerated to be twice as fast.

Per candidate area	Greedy Seeding	Surface-adaptive Seeding
Average number of streamlines	66.2	63.1
Average number of seeds	2056.8	631.6
Average number of streamline points	2352.3	2216.6
Average computation time (sec.)	15.233	5.786



(a)



(b)

Figure 6. Scatter plot of curvature streamline features for polyp candidates: (a) features calculated on curvature streamlines without using streamline defragmentation; (b) features calculated on defragmented curvature streamlines. By streamline defragmentation, the discriminating power of curvature streamline features for classifying true polyp detections is increased.

We compared our improved streamline computation with our previous work. The streamline computation also included implicit iso-surface curvature estimation. Our workstation had dual 3.20GHz Intel Xeon CPUs with 3GB memory. The average computation times for both seeding strategies are shown in table. 1. The result indicates that our surface-adaptive streamline seeding is more than twice as fast as the greedy seeding approach. A far smaller number of redundant seed points were generated while maintaining a good adaptation to the local surface shape.

4.2 Discriminating power of curvature streamline features

In previous work¹⁰ we presented an assessment of our curvature streamline features to discriminate between true- and false-positive polyp detections, based on 5 patient data sets. 50 sites were true polyps and 281 candidates were false-positive detections. These candidates were pre-detected based on the volumetric shape index. Curvature streamlines were applied to all 331 detections. A scatter plot of the streamline winding angle and mean radius features showed a high correlation with true polyp detections.

In further testing, we included another 13 patients. Each patient had 2 CT scans, performed using a Philips Mx8000 multislice scanner. The average voxel size of the 3D volume image is $0.75mm \times 0.75mm \times 1.60mm$. Some of these data sets have a higher noise level. 4598 polyp candidates were pre-detected in the preprocessing stage. An experienced radiologist classified all candidate areas as either true- or false-positive detections. 168 sites were annotated as true polyp detections, found in all 13 patients. Some of them were large (10 – 19mm) polyps and even masses (20mm and above). We calculated curvature streamline features, i.e. the winding angle and mean radius, on all these candidates. A scatter plot (Fig. 6(a)) based on the previous work shows that the

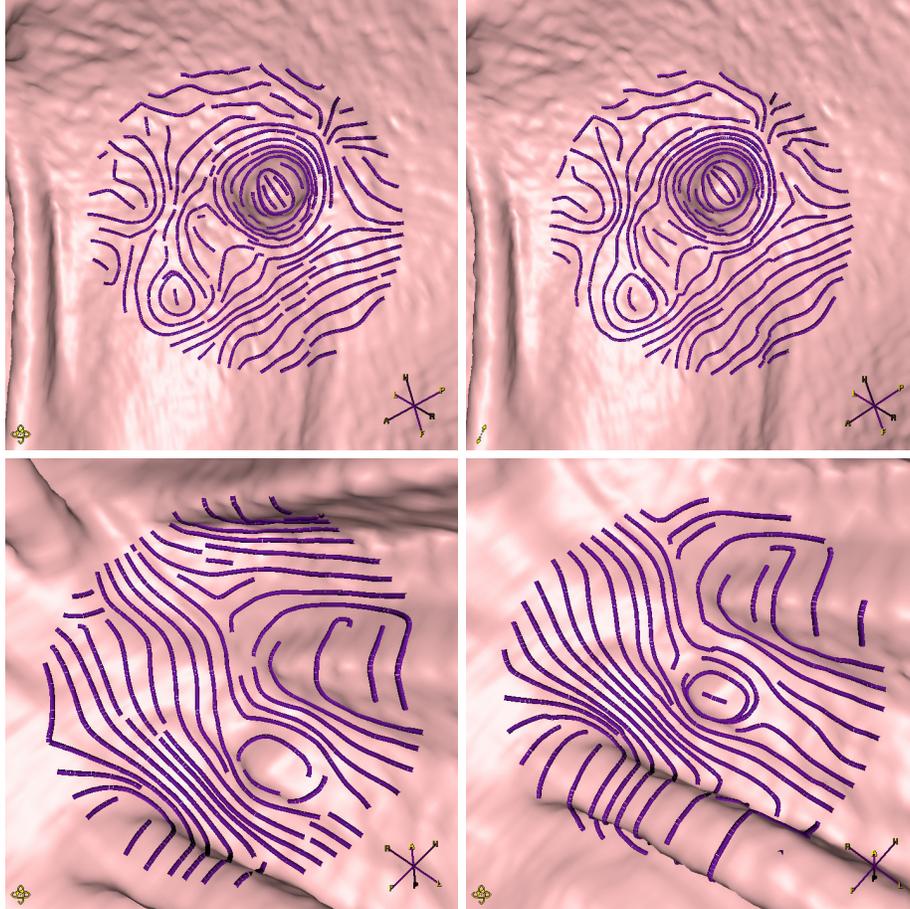


Figure 7. Comparison of previous streamline generation (left images) and improved approach using streamline defragmentation (right images): we often inspect streamline fragments using our previous tracing algorithm. If this case happens around a polyp surface, then we will not detect closed streamlines using winding angle feature. Our streamline defragmentation is a helpful addition to solve this problem.

discriminating power of curvature streamline features can be significantly reduced due to data noise and the size of the polyps. For large polyps, our previous winding angle feature became much less meaningful. An important reason for this was the streamline fragmentation.

We applied our curvature streamline defragmentation approach to improve the discriminatory power of curvature streamline features. The average streamline length was increased by 25%. Fig. 6(b) shows the scatter plot of our improved winding angle and mean radius features on all polyp candidates in the 26 patient data sets. The result indicates that the discriminating power of the curvature streamline features is significantly strengthened by streamline defragmentation, especially for large true polyp detections.

The visualization of the shaded colon wall is also improved by longer curvature streamlines. We compared our correlation-based long streamline approach with our previous work in Fig. 7. In these images, it appears that longer streamlines better emphasize polyp surface features. We expect that this will facilitate visual inspection and expert detection. In future work, this hypothesis will be experimentally evaluated.

5. CONCLUSIONS

Our curvature streamline based polyp detection scheme explored the characteristics of colonic polyps embedded in principal surface curvature direction fields. Streamlines of curvature were used to visualize these two vector

fields and extract specific features. A crucial problem was that streamline generation significantly influences the discriminating power of our curvature streamline features to classify true polyp detections. In this paper, we presented two major improvements to increase the efficiency and robustness of our previous work. We adapted our streamline seeding to local surface shape and reduced the computational complexity of streamline generation. The process was accelerated generating a significantly smaller number of seed points. Our streamline defragmentation approach yielded longer curvature streamlines that better characterize the colonic surface shape. We applied our curvature streamline technique on more data sets, of which some have a higher noise level. The result shows that our features calculated on longer curvature streamlines correlate more strongly with true polyp detections. In addition, the robustness of our improved streamline features will be further validated in later research. We will also work on deriving other useful features from curvature streamlines as helpful additions for our polyp candidate classification.

ACKNOWLEDGMENTS

This work was supported in part by Philips Health Care, Best, the Netherlands. The authors thank all colleagues in Philips for their advices and helpful suggestions, as well as providing the authors with patient CT scans. The authors also thank Ir. C. van Wijk for enlightening discussions.

REFERENCES

1. D. Bartz, "Virtual endoscopy in research and clinical practice," *Computer Graphics Forum* **24**, pp. 111–126, Mar. 2005.
2. P. J. Pickhardt, "Ct colonography (virtual colonoscopy) for primary colorectal screening: Challenges facing clinical implementation," *Abdominal Imaging* **30**, pp. 1–4, Jan. 2004.
3. J. Li, M. Franaszek, N. Petrick, J. Yao, A. Huang, and R. M. Summers, "Wavelet method for ct colonography computer-aided polyp detection," in *Proceedings of IEEE International Symposium on Biomedical Imaging*, pp. 1316–1319, 2006.
4. H. Yoshida and J. Näppi, "Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps," *IEEE Transactions on Medical Imaging* **20**, pp. 1261–1274, Dec. 2001.
5. R. M. Summers, J. Yao, P. J. Pickhardt, M. Franaszek, I. Bitter, D. Brickman, V. Krishna, and J. R. Choi, "Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population," *Gastroenterology* **129**, pp. 1832–1844, Dec. 2005.
6. A. Huang, R. M. Summers, and A. K. Hara, "Surface curvature estimation for automatic colonic polyp detection," in *Medical Image 2005: Physiology, Function and Structure from Medical Images, Proceedings of SPIE*, **5746**, pp. 393–402, 2005.
7. G. Kiss, J. van Cleynenbreugel, S. Drisis, D. Bielen, G. Marchal, and P. Suetens, "Computer aided detection for low-dose ct colonography," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005*, pp. 859–867, 2005.
8. C. van Wijk, V. F. van Ravesteijn, F. M. Vos, R. Truyen, A. H. de Vries, J. Stoker, and L. J. van Vliet, "Detection of protrusions in curved folded surfaces applied to automated polyp detection in ct colonoscopy," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2006*, pp. 471–478, 2005.
9. W. Hong, F. Qiu, and A. Kaufman, "A pipeline for computer aided polyp detection," *IEEE Transactions on Visualization and Computer Graphics* **12**(5), pp. 861–868, 2006.
10. L. Zhao, C. P. Botha, J. O. Bescos, R. Truyen, F. M. Vos, and F. H. Post, "Lines of curvature for polyp detection in virtual colonoscopy," *IEEE Transactions on Visualization and Computer Graphics* **12**(5), pp. 885–892, 2006.
11. R. M. Summers, W. S. Selbie, J. D. Malley, L. M. Pusanik, A. J. Dwyer, N. A. Courcoutsakis, D. J. Shaw, D. E. Kleiner, M. C. Sneller, C. A. Langford, S. M. Holland, and J. H. Shelhamer, "Polypoid lesions of airways: Early experience with computer-assisted detection by using virtual bronchoscopy and surface curvature," *Radiology* **208**, pp. 331–337, 1998.
12. B. Jobard and W. Lefer, "Creating evenly-spaced streamlines of arbitrary density," in *Proceedings of the Eurographics Workshop on Visualization in Scientific Computing*, pp. 45–55, 1997.

13. B. Moberts, A. Vilanova, and J. J. van Wijk, "Evaluation of fiber clustering methods for diffusion tensor imaging," in *Proceedings of IEEE Visualization 2005 (VIS'05)*, pp. 65–72, Oct. 2005.
14. A. Brun, H.-J. Park, H. Knutsson, and C.-F. Westin, "Coloring of dt-mri fiber traces using laplacian eigenmaps," in *Proceedings of EUROCAST 2003, Lecture Notes in Computer Science 2809*, pp. 564–572, Feb. 2003.
15. I. Corouge, S. Gouttard, and G. Gerig, "Towards a shape model of white matter fiber bundles using diffusion tensor mri," in *Proceedings of IEEE International Symposium on Biomedical Imaging*, pp. 344–347, 2004.