

Recognition of Protruding Objects in Highly Structured Surroundings by Structural Inference

Vincent F. van Ravesteijn¹, Frans M. Vos^{1,2}, and Lucas J. van Vliet¹

¹ Quantitative Imaging Group, Faculty of Applied Sciences,
Delft University of Technology, The Netherlands
V.F.vanRavesteijn@tudelft.nl

² Department of Radiology, Academic Medical Center,
Amsterdam, The Netherlands

Abstract. Recognition of objects in highly structured surroundings is a challenging task, because the appearance of target objects changes due to fluctuations in their surroundings. This makes the problem highly context dependent. Due to the lack of knowledge about the target class, we also encounter a difficulty delimiting the non-target class. Hence, objects can neither be recognized by their similarity to prototypes of the target class, nor by their similarity to the non-target class. We solve this problem by introducing a transformation that will eliminate the objects from the structured surroundings. Now, the dissimilarity between an object and its surrounding (non-target class) is inferred from the difference between the local image before and after transformation. This forms the basis of the detection and classification of polyps in computed tomography colonography. 95% of the polyps are detected at the expense of four false positives per scan.

1 Introduction

For classification tasks that can be solved by an expert, there exists a set of features for which the classes are separable. If we encounter class overlap, not enough features are obtained or the features are not chosen well enough. This conveys the viewpoint that a feature vector representation directly reduces the object representation [1]. In the field of imaging, the objects are represented by their grey (or color) values in the image. This sampling is already a reduced representation of the real world object and one has to ascertain that the acquired digital image still holds sufficient information to complete the classification task successfully. If so, all information is still retained and the problem reduces to a search for an object representation that will reveal the class separability.

Using all pixels (or voxels) as features would give a feature set for which there is no class overlap. However, this feature set usually forms a very high dimensional feature space and the problem would be sensitive to the curse of dimensionality. Considering a classification problem in which the objects are regions of interest \mathcal{V} with size N from an image with dimensionality D , the dimensionality of the feature space Ω would then be N^D , i.e. the number of pixels

in \mathcal{V} . This high dimensionality poses problems for statistical pattern recognition approaches. To avoid these problems, principal component analysis (PCA) could for example be used to reduce the dimensionality of the data without having the user to design a feature vector representation of the object. Although PCA is designed to reduce the dimensionality while keeping as most information as possible, the mapping unavoidably reduces the object representation.

The use of statistical approaches completely neglects that images often contain structured data. One can think of images that are very similar (images that are close in the feature space spanned by all pixel values), but might contain significantly different structures. Classification of such structured data receives a lot of attention and is motivated by the idea that humans interpret images by perception of structure rather than by perception of all individual pixel values. An approach for the representation of structure of objects is to represent the objects by their dissimilarities to other objects [2]. When a dissimilarity measure is defined (for example the 'cost' of deforming an object into another object), the object can be classified based on the dissimilarities of the object to a set (or sets) of prototypes representing the classes.

Classification based on dissimilarities demands prototypes of both classes, but this demand can not always be fulfilled. For example, the detection of target objects in highly structured surroundings poses two problems. First, there is a fundamental problem describing the class of non-targets. Even if there is detailed knowledge about the target objects, the class of non-targets (or outliers) is merely defined as all other objects. Second, if the surroundings of the target objects is highly structured, the number of non-target prototypes is very large and they all differ each in their own way, i.e. they are scattered all over the feature space. The selection of a finite set of prototypes that sufficiently represents the non-target class is almost impossible and one might have to rely on one-class classification.

The objective of this paper is to establish a link between image processing and dissimilarity based pattern recognition. On the one hand, we show that the previous work [3] can be seen as an application of structural inference which is used in featureless pattern recognition [1]. On the other hand, we extend the featureless pattern recognition to pattern recognition in the absence of prototypes. The role of prototypes is replaced by a single context-dependent prototype that is derived from the image itself by a specific transformation for the application at hand. The approach will be applied in the context of automated polyp detection.

2 Automated Polyp Detection

The application that we present in this paper is automated polyp detection in computed tomography (CT) colonography (CTC). Adenomatous polyps are important precursors to cancer and early removal of such polyps can reduce the incidence of colorectal cancer significantly [4,5]. Polyps manifest themselves as protrusions from the colon wall and are therefore visible in CT. CTC is a minimal-invasive technique for the detection of polyps and, therefore, CTC is considered a promising candidate for large-scale screening for adenomatous

polyps. Computer aided detection (CAD) of polyps is being investigated to assist the radiologists. A typical CAD system consists of two consecutive steps: candidate detection to detect suspicious locations on the colon wall, and classification to classify the candidates as either a polyp or a false detection.

By nature the colon is highly structured; it is curved, bended and folded. This makes that the appearance of a polyp is highly dependent on its surrounding. Moreover, a polyp can even be (partly) occluded by fecal remains in the colon.

2.1 Candidate Detection

Candidate detection is based on a curvature-driven surface evolution [3,6]. Due to the tube-like shape of the colon, the second principal curvature κ_2 of the colon surface is smaller than or close to zero everywhere (the normal vector points into the colon), except on protruding locations. Polyps can thus be characterized by a positive second principal curvature. The surface evolution reduces the protrusion iteratively by solving a non-linear partial differential equation (PDE):

$$\frac{\partial I}{\partial t} = \begin{cases} -\kappa_2 |\nabla I| & (\kappa_2 > 0) \\ 0 & (\kappa_2 \leq 0) \end{cases} \quad (1)$$

where I is the three-dimensional image and $|\nabla I|$ the gradient magnitude of the image.

Iterative application of (1) will remove all protruding elements (i.e. locations where $\kappa_2 > 0$) from the image and estimates the appearance of the colon surface as if the protrusion (polyp) was never there. This is visualized in Fig. 1 and Fig. 2. Fig. 1(a) shows the original image with a polyp situated on a fold. The grey values are iteratively adjusted by (1). The deformed image (or the solution of the PDE) is shown in Fig. 1(b). The surrounding is almost unchanged, whereas the polyp has completely disappeared. The change in intensity between the two images is shown in Fig. 1(c). Locations where the intensity change is larger than 100 HU (Hounsfield units) yield the polyp candidates and their segmentation (Fig. 1(d)). Fig. 2 also shows isosurface renderings at different time-steps.

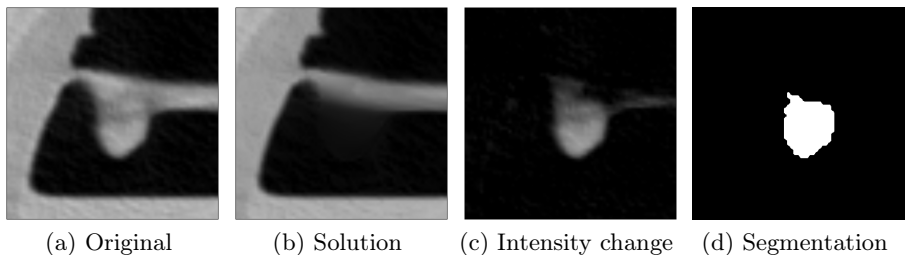


Fig. 1. (a) The original CT image (grey is tissue, black is air inside the colon). (b) The result after deformation. The polyp is smoothed away and only the surrounding is retained. (c) The difference image between (a) and (b). (d) The segmentation of the polyp obtained by thresholding the intensity change image.

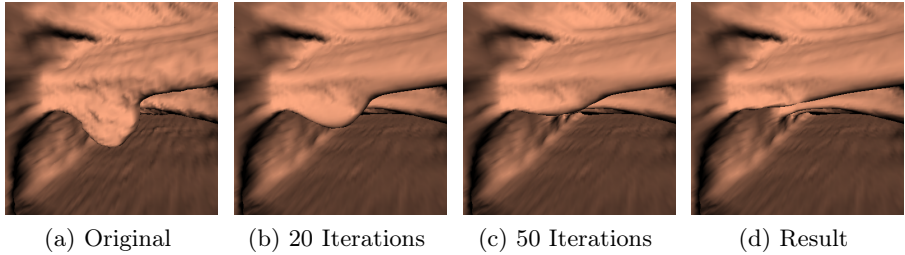


Fig. 2. Isosurface renderings (-750 HU) of a polyp and its surrounding. (a) Before deformation. (b–c) After 20 and 50 iterations. (d) The estimated colon surface without the polyp.

2.2 Related Work

Konukoglu et al. [7] have proposed a related, but different approach. Their method is also based on a curvature-based surface evolution, but instead of removing protruding structures, they proposed to enhance polyp-like structures and to deform them into spherical objects. The deformation is guided by

$$\frac{\partial I}{\partial t} = \left(1 - \frac{H}{H_0}\right) |\nabla I| \quad (2)$$

with H the mean curvature and H_0 the curvature of the sphere towards the candidate is deformed.

3 Structural Inference for Object Recognition

The candidate detection step, described in the previous section, divides the feature space Ω of all possible images into two parts. The first part consists of all images that are not affected by the PDE. It is assumed that these images do not show any polyps and these are said to form the surrounding class Ω_o . The other part consists of all images that are deformed by iteratively solving the PDE. These images thus contain a certain protruding element. However, not all images with a protruding element do contain a polyp as there are other possible causes of protrusions like fecal remains, the ileocecal valve (between the large and small intestine) and natural fluctuations of the colon wall.

To summarize, three classes are now defined:

1. a class $\Omega_o \subset \Omega$; all images without a polyp: the surrounding class,
2. a class $\Omega_f \subset \Omega \setminus \Omega_o$; all images showing a protrusion that is not a polyp: the false detection class, and
3. a class $\Omega_t \subset \Omega \setminus \Omega_o$; all images showing a polyp: the true detection class.

Successful classification of new images now requires a meaningful representation of the classes and a measure to quantify the dissimilarity between an image and a certain class. Therefore, Section 3.1 will describe how the dissimilarities can be defined for objects of which the appearance is highly context-dependent, and Section 3.2 will discuss how the classes can be represented.

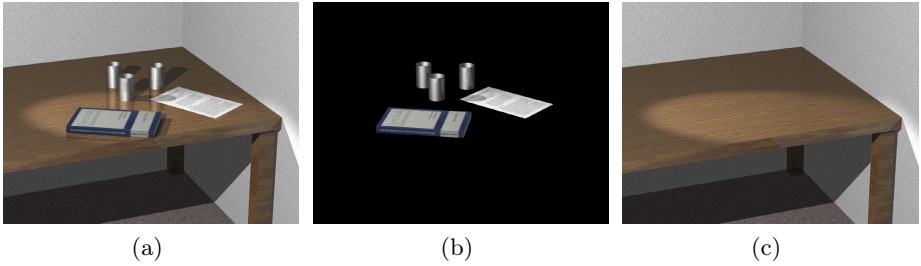


Fig. 3. (a) Objects in their surroundings. (b) Objects without their surroundings. All information about the objects is retained, so the objects can still be classified correctly. (c) The estimated surrounding without the objects.

3.1 Dissimilarity Measure

To introduce the terminology and notation, let us start with a simple example of dissimilarities between objects. Fig. 3(a) shows various objects on a table. Two images, say \mathbf{x}_i and \mathbf{x}_j , represent for instance an image of the table with a cup and an image of the table with the book. The dissimilarity between these images is hard to define, but the dissimilarity between either one of these images and the image of an empty table is much easier. This dissimilarity may be derived from the image of the specific object itself (Fig. 3(b)).

When we denote the image of an empty table as \mathbf{p}_o , this first example can be schematically illustrated as in Fig. 4(a). The dissimilarities of the two images to the prototype \mathbf{p}_o are called d_{i_o} and d_{j_o} . If these dissimilarities are simply defined as the Euclidean distance between the circles in the image, the triangle-inequality holds.

However, if the dissimilarities are defined as the spatial distance between the objects (in 3D-space), all objects in Fig. 3(a) have zero distance to the table, but the distance between any two objects (other than the table) is larger than zero. This shows a situation in which the dissimilarity measure violates the triangle-inequality and the measure becomes non-metric [8]. This is schematically illustrated in Fig. 4(b). The prototype \mathbf{p}_o is no longer a single point, but is transformed into a blob Ω_o representing all objects with zero distance to the table. Note that all circles have zero Euclidean distance to Ω_o .

The image of the empty table can also be seen as the background or surrounding of all the individual objects, which shows that all objects have exactly the same surrounding. When considering the problem of object detection in highly structured surroundings this obviously no longer holds. We first state that, as in the first example given above, the dissimilarity of an object to its surrounding can be defined by the object itself. Secondly, although the surroundings may differ significantly from each other, it is known that none of the surroundings contain an object of interest (a polyp). Thus, as in the second example, the distances between all surroundings can be made zero and we obtain the same blob representation for Ω_o , i.e. the surrounding class. The distance of an object

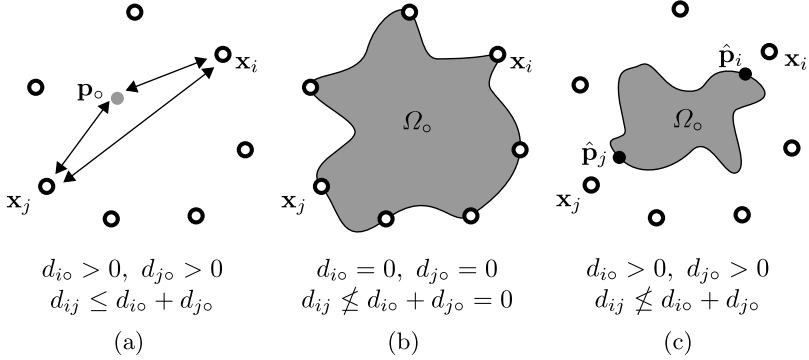


Fig. 4. (a) Feature space of two images of objects having the same surrounding, which means that the image of the surrounding (the table in Fig. 3(a)) reduces to a single point \mathbf{p}_o . (b) When considering spatial distances between the objects, the surrounding image \mathbf{p}_o transforms into a blob Ω_o and all distances between objects within Ω_o are zero. (c) When the surroundings of each object are different but have zero distance to each other, the feature space is a combination of (a) and (b).

to the surrounding class can now be defined as a minimization of the distance between the image of the object over all images \mathbf{p}_k from the set of surroundings Ω_o

$$d_{i_o} \triangleq d(\mathbf{x}_i, \Omega_o) = \min_k d(\mathbf{x}_i, \mathbf{p}_k) \text{ with } \mathbf{p}_k \in \Omega_o.$$

In short, this problem is a combination of the two examples and this leads to the feature space shown in Fig. 4(c). Both images \mathbf{x}_i and \mathbf{x}_j have a related image (prototype), respectively $\hat{\mathbf{p}}_i$ and $\hat{\mathbf{p}}_j$, to which the dissimilarity is the smallest. Again, the triangle inequality does no longer hold: two images that look very different may both be very close to the surrounding class. On the other hand, two objects that are very similar do have similar dissimilarity to the surrounding class. This means that the compactness hypothesis still holds in the space spanned by the dissimilarities. Moreover, the dissimilarity of an object to its surrounding still contains all information for successful classification of the object, which may easily be seen by looking at Fig. 3(b).

3.2 Class Representation

The prototypes $\hat{\mathbf{p}}_i$ and $\hat{\mathbf{p}}_j$ thus represent the surrounding class, but are not available a priori. We know that they must be part of the boundary of Ω_o and that the boundary of Ω_o is the set of objects that divides the feature space of images with protrusions and those without protrusions. Consequently, for each object we can derive its related prototype of the surrounding class by iteratively solving the PDE in (1). That is, $\Omega_s \triangleq \delta\Omega_o \cap (\delta\Omega_i \cup \delta\Omega_f)$ are all solutions of (1) and the dissimilarity of an object to its surroundings is the 'cost' of the deformation

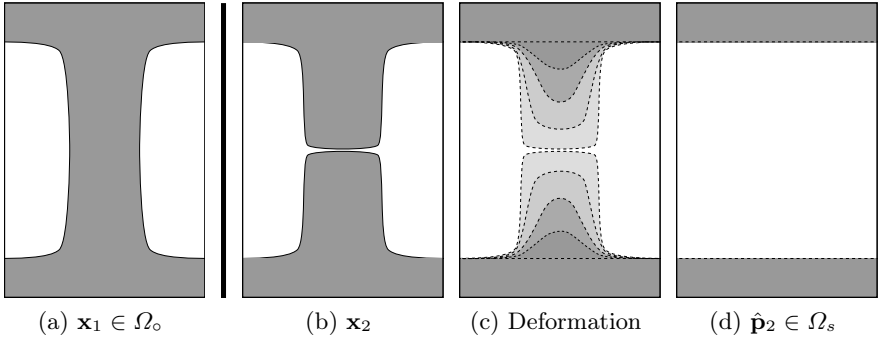


Fig. 5. (a–b) Two similar images having different structure lead to different responses to deformation by the PDE in (1). The object \mathbf{x}_1 is a solution itself, whereas \mathbf{x}_2 will be deformed into $\hat{\mathbf{p}}_2$. A number of structures that might occur during the deformation process are shown in (c).

guided by (1). Furthermore, the prototypes of the surroundings class can now be sampled almost infinitely, i.e. a prototype can be derived when it is needed.

A few characteristics of our approach to object detection are illustrated in Fig. 5. At first glance, objects \mathbf{x}_1 and \mathbf{x}_2 , respectively shown in Figs. 5(a) and (b), seem to be similar (i.e. close together in the feature space spanned by all pixel values), but the structures present in these images differ significantly. This difference in structure is revealed when the images are being transformed by the PDE (1). Object \mathbf{x}_1 does not have any protruding elements and can thus be considered as an element of Ω_o , whereas object \mathbf{x}_2 exhibits two large protrusions: one pointing down from the top, the other pointing up from the bottom. Fig. 5(c) shows several intermediate steps in the deformation of this object and Fig. 5(d) shows the final solution. This illustrates that by defining a suitable deformation, a specific structure can be measured in an image. Using the deformation defined by the PDE in (1), all intermediate images are also valid images with protrusions with decreasing protrudedness. Furthermore, all intermediate objects shown in Fig. 5(c) have the same solution. Thus, different objects can have the same solution and relate to the same prototype.

Let us propose to use a morphological closing operation as the deformation, then one might conclude that images \mathbf{x}_1 and \mathbf{x}_2 are very similar. In that case we might conclude that image \mathbf{x}_2 does not really have the structure of two large polyps, as we concluded before, but might have the same structure as in \mathbf{x}_1 altered by an imaging artifact. Using different deformations can thus lead to a better understanding of the local structure. In that case, one could represent each class by a deformation instead of a set of prototypes [1]. Especially for problems involving objects in highly structured surroundings, it might be advantageous to define different deformations in order to infer from structure.

An example of an alternative deformation was already given by the PDE in (2). This deformation creates a new prototype of the polyp class given an image and the 'cost' of deformation could thus be used in classification. Combining

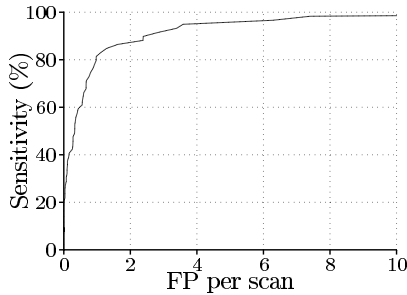


Fig. 6. FROC curve for the detection of polyps ≥ 6 mm

both methods thus gives for each object a dissimilarity to both classes. However, this deformation was proposed as a preprocessing step for current CAD systems. By doing so, the dissimilarity was not explicitly used in the candidate detection or classification step.

4 Classification

We now have a very well sampled class of the healthy (normal) images, which do not contain any protrusions. Any deviation from this class indicates unhealthy protrusions. This can be considered as a typical one-class classification problem in which the dissimilarity between the object \mathbf{x} and the prototype \mathbf{p} indicates the probability of belonging to the polyp class. The last step in the design of the polyp detection system is to define a dissimilarity measure that quantifies the introduced deformation, such that it can be used to successfully distinguish the non-polyps from the polyps. As said before, the difference image still contains all information, and thus there is still no class overlap.

Until now, features are computed from this difference image to quantify the 'cost' of deformation. Three features are used for classification: the length of the two principal axes (perpendicular to the polyp axis) of the segmentation of the candidate, and the maximum intensity change. A linear logistic classifier is used for classification. Classification based on the three features obtained from the difference image leads to results comparable to other studies [9,10,11]. Fig. 6 shows a free-response receiver operating characteristics (FROC) curve of the CAD system for 59 polyps larger than 6 mm (smaller polyps are clinically irrelevant) annotated in 86 patients (172 scans). Results of the current polyp detection systems are also presented elsewhere [3,6,12].

5 Conclusion

We have presented an automated polyp detection system based on structural inference. By transforming the image using a structure-driven partial differential

equation, knowledge is inferred from the structure in the data. Although no prototypes are available a priori, a prototype of the 'healthy' surrounding class can be obtained for each candidate object. The dissimilarity with the healthy class is obtained by means of a difference image between the image before and after the transformation. This dissimilarity is used for classification of the object as either a polyp or as healthy tissue. Subsequent classification is based on three features derived from the difference image. The current implementation basically acts like a one-class classification system: the system measures the dissimilarity to a well sampled class of volumes showing only normal (healthy) tissue. The class is well sampled in the sense that for each candidate object we can derive a healthy counterpart, which acts as a prototype.

Images that are very similar might not always have the same structure. In the case of structured data, it is this structure that is most important. It was shown that the transformation guided by the PDE in (1) is capable of retrieving structure from data. Furthermore, if two objects are very similar, but situated in a different surrounding, the images might look very different. However, after iteratively solving the PDE, the resulting difference images of the two objects are also similar. The feature space spanned by the dissimilarities thus complies with the compactness hypothesis. However, when a polyp is situated, for example, between two folds, the real structure might not always be retrieved. In such situations no distinction between Figs. 5(a) and (b) can be made due to e.g. the partial volume effect or Gaussian filtering prior to curvature and derivative computations. Prior knowledge about the structure of the colon and the folds in the colon might help in these cases.

Until now, only information is used about the dissimilarity to the 'healthy' class. The work of Konukoglu et al. [7] offers the possibility of deriving a prototype for the polyp class given a candidate object just as we derived prototypes for the non-polyp class. A promising solution might be a combination of both techniques; each candidate object is then characterized by its dissimilarity to a non-polyp prototype and by its dissimilarity to a polyp prototype. Both prototypes are created on-the-fly and are situated in the same surrounding as the candidate. In fact, two classes have been defined and each class is characterized by its own deformation.

In the future, the patient preparation is further reduced to improve patient compliance. This will lead to data with increased amount of fecal remains in the colon and this will complicate both the task of automated polyp detection as well as electronic cleansing of the colon [13,14]. The presented approach to infer from structure can also contribute to the image processing of such data, especially if the structure within the colon becomes increasingly complicated.

References

1. Duin, R.P.W., Pekalska, E.: Structural inference of sensor-based measurements. In: Yeung, D.-Y., Kwok, J.T., Fred, A., Roli, F., de Ridder, D. (eds.) SSPR 2006 and SPR 2006. LNCS, vol. 4109, pp. 41–55. Springer, Heidelberg (2006)

2. Pekalska, E., Duin, R.P.W.: *The Dissimilarity Representation for Pattern Recognition, Foundations and Applications*. World Scientific, Singapore (2005)
3. van Wijk, C., van Ravesteijn, V.F., Vos, F.M., Truyen, R., de Vries, A.H., Stoker, J., van Vliet, L.J.: Detection of protrusions in curved folded surfaces applied to automated polyp detection in CT colonography. In: Larsen, R., Nielsen, M., Sporning, J. (eds.) *MICCAI 2006*. LNCS, vol. 4191, pp. 471–478. Springer, Heidelberg (2006)
4. Ferrucci, J.T.: Colon cancer screening with virtual colonoscopy: Promise, polyps, politics. *American Journal of Roentgenology* 177, 975–988 (2001)
5. Winawer, S., Fletcher, R., Rex, D., Bond, J., Burt, R., Ferrucci, J., Ganiats, T., Levin, T., Woolf, S., Johnson, D., Kirk, L., Litin, S., Simmang, C.: *Colorectal cancer screening and surveillance: Clinical guidelines and rationale – update based on new evidence*. *Gastroenterology* 124, 544–560 (2003)
6. van Wijk, C., van Ravesteijn, V.F., Vos, F.M., van Vliet, L.J.: Detection and segmentation of protruding regions on folded iso-surfaces for the detection of colonic polyps (submitted)
7. Konukoglu, E., Acar, B., Paik, D.S., Beaulieu, C.F., Rosenberg, J., Napel, S.: Polyp enhancing level set evolution of colon wall: Method and pilot study. *IEEE Trans. Med. Imag.* 26(12), 1649–1656 (2007)
8. Pekalska, E., Duin, R.P.W.: Learning with general proximity measures. In: *Proc. PRIS 2006*, pp. IS15–IS24 (2006)
9. Summers, R.M., Yao, J., Pickhardt, P.J., Franaszek, M., Bitter, I., Brickman, D., Krishna, V., Choi, J.R.: Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population. *Gastroenterology* 129, 1832–1844 (2005)
10. Summers, R.M., Handwerker, L.R., Pickhardt, P.J., van Uitert, R.L., Deshpande, K.K., Yeshwant, S., Yao, J., Franaszek, M.: Performance of a previously validated CT colonography computer-aided detection system in a new patient population. *AJR* 191, 169–174 (2008)
11. Näppi, J., Yoshida, H.: Fully automated three-dimensional detection of polyps in fecal-tagging CT colonography. *Acad. Radiol.* 14, 287–300 (2007)
12. van Ravesteijn, V.F., van Wijk, C., Truyen, R., Peters, J.F., Vos, F.M., van Vliet, L.J.: Computer aided detection of polyps in CT colonography: An application of logistic regression in medical imaging (submitted)
13. Serlie, I.W.O., Vos, F.M., Truyen, R., Post, F.H., van Vliet, L.J.: Classifying CT image data into material fractions by a scale and rotation invariant edge model. *IEEE Trans. Image Process.* 16(12), 2891–2904 (2007)
14. Serlie, I.W.O., de Vries, A.H., Vos, F.M., Nio, Y., Truyen, R., Stoker, J., van Vliet, L.J.: Lesion conspicuity and efficiency of CT colonography with electronic cleansing based on a three-material transition model. *AJR* 191(5), 1493–1502 (2008)