Endoluminal surface registration for CT colonography using haustral fold matching ☆

Thomas Hampshire [a,*], Holger R. Roth [a], Emma Helbren [b], Andrew Plumb [b], Darren Boone [b], Greg Slabaugh [c], Steve Halligan [b], David J. Hawkes [a]

[a] Centre for Medical Image Computing, University College London, Gower Street, London WC1E 6BT, UK
[b] Centre for Medical Imaging, University College London, 255 Euston Road, London WC1 2BU, UK
[c] Department of Computer Science, City University, London EC1V OHB, UK

ARTICLE INFO

ABSTRACT

Computed Tomographic (CT) colonography is a technique used for the detection of bowel cancer or potentially precancerous polyps. The procedure is performed routinely with the patient both prone and supine to differentiate fixed colonic pathology from mobile faecal residue. Matching corresponding locations is difficult and time consuming for radiologists due to colonic deformations that occur during patient repositioning.

We propose a novel method to establish correspondence between the two acquisitions automatically. The problem is first simplified by detecting haustral folds using a graph cut method applied to a curvature-based metric applied to a surface mesh generated from segmentation of the colonic lumen. A virtual camera is used to create a set of images that provide a metric for matching pairs of folds between the prone and supine acquisitions. Image patches are generated at the fold positions using depth map renderings of the endoluminal surface and optimised by performing a virtual camera registration over a restricted set of degrees of freedom. The intensity difference between image pairs, along with additional neighbourhood information to enforce geometric constraints over a 2D parameterisation of the 3D space, are used as unary and pair-wise costs respectively, and included in a Markov Random Field (MRF) model to estimate the maximum a posteriori fold labelling assignment.

The method achieved fold matching accuracy of 96.0% and 96.1% in patient cases with and without local colonic collapse. Moreover, it improved upon an existing surface-based registration algorithm by providing an initialisation. The set of landmark correspondences is used to non-rigidly transform a 2D source image derived from a conformal mapping process on the 3D endoluminal surface mesh. This achieves full surface correspondence between prone and supine views and can be further refined with an intensity based registration showing a statistically significant improvement ($p < 0.001$), and decreasing mean error from 11.9 mm to 6.0 mm measured at 1743 reference points from 17 CTC datasets.

1. Introduction

1.1. Motivation

Colorectal cancer is a leading cause of cancer mortality with 1.23 million individuals developing the disease and 608,000 deaths annually (World Health Organisation, 2009). Population screening offers the best prospects to reduce in mortality and aims to prevent the development of advanced cancers by early detection and removal of both localised cancers or premalignant adenomas, from which more than 80% of cancers are thought to arise (Cunningham et al., 2010). Optical colonoscopy is the current gold standard method to inspect the whole-colon; however colonoscopy is time consuming and uncomfortable for the patient, and is occasionally associated with serious complications (Pignone and Sox, 2008) such as colonic perforation. Computed tomographic colonography (CTC) is now widely considered the preferred radiological technique for detecting cancer and polyps, and has comparable sensitivity to optical colonoscopy (Johnson et al., 2008), while being more acceptable to patients (von Wagner et al., 2012) and relatively safe (Burling et al., 2006). Patients undergo a full bowel preparation to cleanse the colon, which is then insufflated with gas immediately before helical CT imaging of the abdomen and pelvis

* Corresponding author. Tel.: +44 0 20 7679 6713; fax: +44 0 20 7679 0221.
E-mail addresses: thomas.hampshire.09@ucl.ac.uk (T. Hampshire), h.roth@ucl.ac.uk (H.R. Roth), Emma.Helbren@uclh.nhs.uk (E. Helbren), aaoplumb@doctors.org.uk (A. Plumb), drdarrenboone@gmail.com (D. Boone), gregory.slabaugh.1@city.ac.uk (G. Slabaugh), s.halligan@ucl.ac.uk (S. Halligan), d.hawkes@ucl.ac.uk (D.J. Hawkes).

(gas insufflation maximises attenuation contrast between endoluminal surface and intraluminal space). Graphics rendering software is used to generate high resolution 'virtual colonoscopy' three-dimensional images of the endoluminal surface, simulating those obtained using conventional colonoscopy. The procedure is performed routinely with the patient both prone and supine to redistribute gas and residue within the colon (Taylor et al., 2007). This helps differentiate fixed colonic pathology from mobile faecal residue because abnormalities whose position remains fixed in both acquisitions are likely to be true polyps. Also, using two data acquisitions increase the chance of discovering pathology occluded by retained fluid or hidden by luminal collapse. Matching corresponding locations between prone and supine endoluminal colonic surfaces is therefore an essential aspect of interpretation by radiologists. However, interpretation can be difficult and time-consuming due to the considerable colonic deformations that often occur during repositioning of the patient (Punwani et al., 2009). These deformations can induce diagnostic error and increase interpretation time. Hence, a method for automatic registration of prone and supine datasets has the potential to improve efficiency and diagnostic accuracy.

1.2. Related work

A number of methods have been proposed to find correspondence between prone and supine positions. For example, centreline-based methods extract and align colonic centrelines by stretching and shrinking based on path geometries (Li et al., 2004; Wang et al., 2009). However, these methods are inherently restricted to achieving a registration along a single dimension and do not give any information about the degree of torsion of the colon wall between views. Anatomical landmarks can be used to help align the two datasets by first identifying a stable set of anatomical features, such as the caecum, rectum and flexures (Wang et al., 2009; Näppi et al., 2005). However, these do not account for the varying compression and expansion along the length of the colon nor the degree of colonic torsion between acquisitions. Voxel-based methods provide a further possibility for registration (Suh and Wyatt, 2009). However, these methods rely to varying extents upon continuous prone and supine colonic segmentations, free from luminal occlusion by fluid or collapse; a situation achieved infrequently in clinical practice, despite optimal bowel preparation (Taylor et al., 2003).

Fukano et al. (2010) proposed a registration method based on haustral fold matching. A second-order derivative difference filter was used to extract folds; their volume and relative positions along the centreline with respect to a set of locations of high centreline curvature were then used to establish correspondence. The method relied on prior automatic identification of a set of landmark locations for registration. They reported correct registration of 65.1% of large folds (9.3% could not be judged), and 13.3% of small folds (32.7% could not be judged).

Methods that involve conformal mapping of the endoluminal colonic surface have been proposed in order to reduce the complexity of the three-dimensional task. For example, Zeng et al. (2010a) combined conformal mapping with feature matching between the prone and supine endoluminal surfaces. The prone and supine colonic segmentations were mapped onto five rectangle pairs. Correspondences were established using a feature matching method based upon mean curvature. The method relied on accurately determining five matching segments in the prone and supine datasets, which is difficult to achieve and may be impossible with local endoluminal collapse. The method proposed by Roth et al. (2011), aims to overcome these limitations by mapping the entire endoluminal surface to a cylinder. Dense surface correspondence was then achieved by a conformal mapping of the prone and su-

pine endoluminal surfaces to 2D cylindrical domains using Ricci flow (Jin et al., 2008; Zeng et al., 2010b), followed by a non-rigid cylindrical intensity based registration using a B-spline method (Rueckert et al., 1999) with a sum-of-squared-differences similarity metric based on shape index (SI) (Koenderink, 1990). However, this method can be susceptible to mis-registration of contiguous sections by one or two haustral folds due to the repetitive similarity of neighbouring features and may not achieve sufficient accuracy if the prone and supine acquisition differ considerably in terms of distension or endoluminal collapse.

Most recently, Wang et al. (2012) used a graph matching algorithm to register a set of features extracted from colon segmentations. A set of key points are detected by the n-SIFT algorithm and matched using an algorithm based on mean field theory. The main advantage of the method is that definition of a colon centreline is unnecessary; however reported mean registration error was 37.6 mm.

We propose a novel method to establish correspondence between the two acquisitions automatically. Previous methods attempt to match corresponding haustral folds based on spatial location and size alone, e.g. (Fukano et al., 2010; Zeng et al., 2010a). We use a novel fold-matching metric, based on depth map images of the endoluminal surface at fold positions as well as local geometric information. The problem is first simplified by detecting haustral folds using a graph cut method applied to a surface mesh generated from the segmentation of the colonic lumen. A virtual camera is used to create a set of images which provide a metric for matching pairs of folds between the prone and supine acquisitions. Image patches are generated at the fold positions using an endoluminal surface mesh depth map rendering and optimised by performing a virtual camera registration over a restricted set of degrees of freedom. The intensity difference between image pairs, along with additional neighbourhood information to enforce geometric constraints over a 2D parameterisation of the 3D space, are used as unary and pair-wise costs respectively, and included in a Markov Random Field (MRF) model to estimate the maximum a posteriori fold labelling assignment. This new haustral fold matching method is the principal contribution of the paper and is described in detail in Section 2. A new initialisation method is also introduced. First the sparse positions and displacements of the corresponding fold matches are mapped onto a 2D domain created by performing a conformal mapping using the Ricci flow algorithm (Jin et al., 2008; Zeng et al., 2010b), to construct an underlying function based on multilevel B-splines that can be evaluated at any point to give a transformation from the prone to the supine data. This transformation is refined further by the intensity based registration in Roth et al. (2011). This method explicitly addresses the problem of colonic collapse. Full details of the full surface-based registration and initialisation methods are provided in Section 3.

2. Haustral fold matching

2.1. Haustral fold segmentation

Haustral folds are elongated, ridgelike endoluminal structures and can be identified by extracting curvature measurements from a triangular mesh representation of the colonic wall. The maximum and minimum values of the normal curvature at a point are called the principal curvatures, k_1 and k_2 respectively. A metric based on the principal curvatures is used to perform a binary classification of each vertex as fold, or non-fold:

$$M = k_1 - \gamma \|k_2\|. \tag{1}$$

This recognises that at a fold, one expects $k_1 \gg 0$ and $k_2 \approx 0$. The γ parameter penalises the metric against curvature in any direction

other than in the maximum, helping to separate the folds at the te-naie coli. The surface mesh is first simplified using an edge collapse transformation process (Hoppe, 1999) to a resolution of $\sim$0.2 poly-gons/mm^2, at which level the haustal folds are still clearly visible. Following this, the mesh is treated as a graph, with graph nodes defined by the mesh vertices and graph edges defined by the mesh edges. Using a virtual sink and source with the given weighting, a graph cut segmentation (Boykov and Kolmogorov, 2004) is performed which minimises the energy function:

$$\sum_p E_p(f_p) + \sum_p \sum_{q \in N_p \setminus q} \delta_{pq}(f_p, f_q), \tag{2}$$

where f_p and f_q are the binary labels of neighbouring nodes p and q, corresponding to fold or non-fold, and N_p is the neighbour set of p defined by the directly connected vertices. The function E_p is defined as:

$$E_p(f_p) = \begin{cases} M & \text{if } f_p \text{ is fold} \\ -M & \text{if } f_p \text{ is non-fold} \end{cases} \tag{3}$$

and δ_{pq} is a Potts energy function smoothing term:

$$\delta_{pq}(f_p, f_q) = \begin{cases} 0 & \text{if } f_p = f_q \\ \zeta & \text{if } f_p \neq f_q \end{cases}. \tag{4}$$

This model considers only the equality or inequality of labels and captures the assumption that labelling should be piecewise constant. This results in a label assignment of fold or non-fold over the entire surface mesh. The centre of each fold is taken as the vertex with the shortest maximum distance to any vertex lying on the border of the segmented region. Qualitative visual results are promising (Fig. 1).

2.2. Markov random field modelling

The matching of prone and supine haustral folds is formulated as a labelling problem. First, m haustral folds are detected in the supine data and these are uniquely labelled; the objective is to then assign labels to the detected prone folds, achieved by solving an MRF defined as such:

- A set of sites $\mathbf{S} = \{1, \ldots, n\}$ corresponding to the haustral folds in the prone data set.
- A neighbourhood system $\mathbf{N} = \{N_i | \forall i \in \mathbf{S}\}$, defining the extent of local connections between sites.
- A pair-wise clique defined on $\mathbf{N}$ and $\mathbf{S}$, $\mathbf{C}_2 = \{\{i, i'\} | i \in \mathbf{S}, i' \in N_i\}$, allowing the inclusion of a priori knowledge of geometric dependencies between label configurations.
- A set of random variables $\mathbf{F} = \{F_1, \ldots, F_n\}$ taking on a discrete label $\mathbf{f} = \{f_1, \ldots, f_n\}$ taken from the set of haustral folds identified in the supine data set.

The maximum a posteriori (MAP) estimate of the optimum labelling is computed, which is equivalent to minimising the energy function:

$$\mathbf{f}^{(MAP)} = \arg \min_{\mathbf{f}} \left[\sum_{i \in \mathbf{S}} \Theta_i(f_i) + \sum_{i \in \mathbf{S}} \sum_{j \in N_i} \Theta_{ij}(f_i, f_j) \right], \tag{5}$$

where $\Theta_i(f_i)$ is the unary term, a cost function for assigning the label f_i to site i. Similarly, the pair-wise term $\Theta_{ij}(f_i, f_j)$ is the cost for assigning neighbouring sites i and j labels f_i and f_j.

2.2.1. Unary cost function

The aim is to calculate an $n \times m$ unary cost matrix, where $n = \|\mathbf{S}\|$ is the number of sites or prone folds, and m is the number of labels or supine folds. To calculate the cost $\Theta_i(f_i)$, depth map images are rendered at the fold positions, visualising the internal colon wall (Fig. 2). The depth map images provide a description of the colon surface at the position of haustral folds, as well as the configuration of colonic pathology in the local region. The perspective projections provide a degree of invariance to colon wall deformation due to patient repositioning, whilst also allowing for a direct comparison between images. The images are generated by exploiting the Z-buffer from the graphics rendering model. Normally this buffer contains values that relate to the distance of an object from the camera position and is used to specify the order that polygons should be rendered in order to generate an image of the scene. Using a perspective camera model, the z buffer value z' is a value in the range $[-1, 1]$ specified in terms of the distances of *near* and *far* clipping planes of the viewing frustum with respect to the real object distance z (Lengyel, 2004):
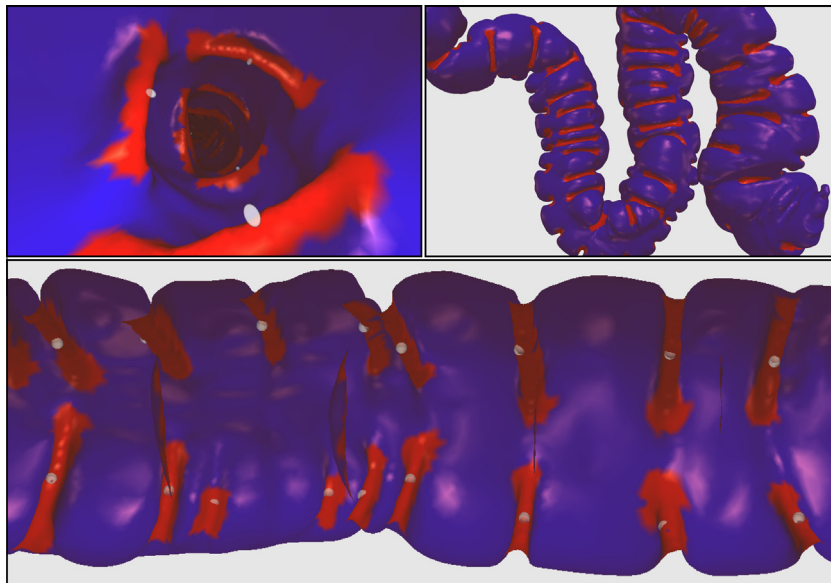
Fig. 1. Virtual colonoscopy (left), external (right) and internal (bottom) views of segmented haustral folds with marked centres. Red and blue sections represent fold and non-fold labelled vertices respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
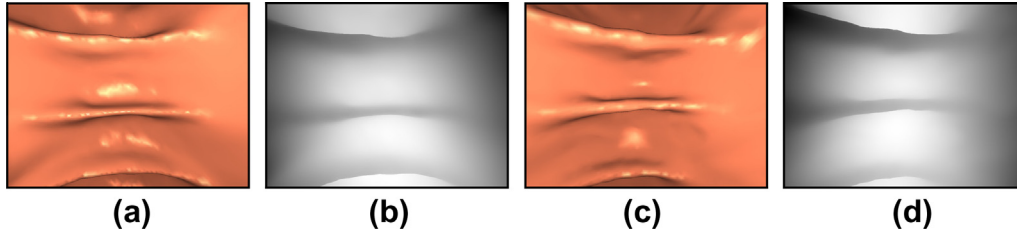
Fig. 2. Internal views of rendered haustral fold in the prone (a) and supine (c) view, with their corresponding depth map images (b) and (d).

$$z' = \frac{far + near}{far - near} + \frac{1}{z}\left(\frac{-2 \cdot far \cdot near}{far - near}\right). \qquad (6)$$

This can be rearranged in terms of z to retrieve true depth values:

$$z = \frac{2 \cdot far \cdot near}{far + near - z'(far - near)}. \qquad (7)$$

The resulting images are then compared using a mean-squared-error (MSE) metric. An optimisation over the external parameters of the virtual camera used to visualise the supine dataset accounts for any inaccuracies in the fold point identification. Restricting the number of degrees of freedom of camera search ensures that the camera focus remains on the correct fold. The degrees of freedom (see Fig. 3) are as follows:

- Elevation (θ) – the fold centre and camera right vector give a position and axis about which the camera is rotated
- Roll (ϕ) – rotation around the camera view direction
- Dollying (τ) – translation along the camera right vector
- Zoom (η) – the distance from camera position to fold centre

Given the four parameters θ, ϕ, τ, η; the optimisation finds the local minimum in a MSE between the depth map images I_1 and I_2 using Powell's gradient descent method (Fletcher and Powell, 1963). The camera position is initialised to the closest intra-luminal centreline point to the fold with the camera 'up' vector set to the tangent to the centreline curve and viewpoint centred on the fold. A multi-resolution approach is used to increase the chance of the optimisation converging on the correct minima. Adding a scaling parameter W allows the weighting of unary to pair-wise costs:

$$R(I_1, I_2) = W\sqrt{\text{MSE}(I_1, I_2 | \theta, \phi, \tau, \eta)}. \qquad (8)$$

Additionally, a constraint is added so that the matching folds must lie in a similar region. The normalised fold centreline positions in the prone and supine $\hat{v}_c = [0, 1]$, are used to limit corresponding fold matches to a fraction ω of the total colon centreline length. Finally, a constant unary cost α is associated with the assignment of the null label f_0 to any given node, allowing for missing labels. The unary costs are then defined:
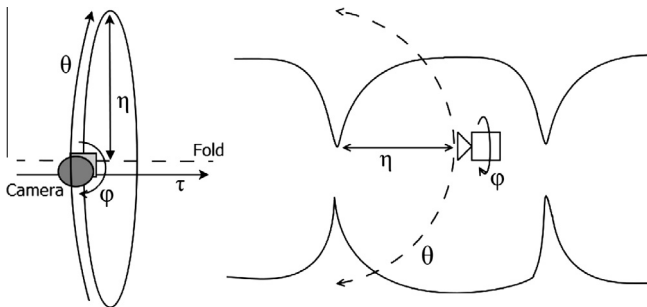

Fig. 3. Virtual camera with restricted set of degrees of freedom for optimisation. Image shows view along (left) and perpendicular to (right) the axis of the colon.

$$\Theta_i(f_i) = \begin{cases} \alpha & \text{if } f_i = f_0 \\ R(I_1, I_2) - \min(R(I_1, \cdot)) & \text{if } \hat{v}_c^{prone} = \hat{v}_c^{supine} \pm \omega, \\ \infty & \text{otherwise} \end{cases} \qquad (9)$$

where $\min(R(I_1, \cdot))$ is the minimum cost for depth map I_1 over all depth maps.

2.2.2. Pair-wise cost function

To improve labelling performance, geometric information about neighbouring fold positions can be used. In this work an adapted frame ($\mathbf{t}, \mathbf{u}, \mathbf{v}$) on the centreline curve $r(\xi)$ is used. An adapted frame is a set of orthonormal vectors, where $\mathbf{t}$ is the unit tangent and $\mathbf{u}$ and $\mathbf{v}$ span the curve normal plane. The most familiar case of an adapted frame is the Frenet frame which defines $\mathbf{u}$ and $\mathbf{v}$ as the curve normal $\mathbf{n}$ and binormal $\mathbf{b}$ respectively:

$$\mathbf{t} = \frac{\frac{d\mathbf{r}}{d\xi}}{\left|\frac{d\mathbf{r}}{d\xi}\right|}, \quad \mathbf{n} = \frac{\frac{d\mathbf{t}}{d\xi}}{\left|\frac{d\mathbf{t}}{d\xi}\right|}, \quad \mathbf{b} = \mathbf{t} \times \mathbf{n}. \qquad (10)$$

Unfortunately, the Frenet frame suffers from indeterminacies at inflexions where r'' is parallel to r', or vanishes, and therefore $\mathbf{n}$ and $\mathbf{b}$ are undefined. As such, the rotation of such a frame about the tangent of a general curve often leads to undesirable twists in frame orientation (Fig. 4). In this work a Rotation Minimising Frame (RMF), specifically the double inflexion method in Wang and Joe (1997) is employed to define a moving frame that does not rotate about the instantaneous tangent of the centreline curve $r(\xi)$ (Fig. 4) by minimising the global error E_g of the magnitude of the angle between the reference vectors of frames U_i and U_{i+1}:

$$E_g = \sum_{i=0}^{n-1} |\angle(U_i, U_{i+1})|. \qquad (11)$$

The resulting set of reference frames can then be use to describe the relative position of each fold to its neighbours: $\mathbf{v} = [v_c, v_\theta]^T$; where v_c is the difference in fold position along the centreline and $v_\theta = [\pi/2, -\pi/2]$ is the difference in angle of rotation around the centreline with respect to the RMF. For each fold position p, the corresponding centreline position c can be found with a surface to centreline correspondence method. We use the conformal mapping technique described in (Roth et al., 2010). A degree of rotation θ can be defined by interpolating a RMF corresponding to centreline position c, and projecting the vector from centreline position to fold $\mathbf{p}$ onto the curve normal plane defined by $\mathbf{u}$ and $\mathbf{v}$ (Fig. 5). The angle θ between this vector $\mathbf{p}'$ and the RMF normal $\mathbf{u}$ then gives a relative degree of rotation with respect to the RMF and can be used to compare neighbouring folds.

This 2D parameterisation simplifies the description of the translation between corresponding pairs of folds between the prone and supine as the centreline v_c and rotational v_θ displacement should be similar ($\mathbf{v}^{prone} \approx \mathbf{v}^{supine}$). Alternatively we can state $\mathbf{v}^{prone} = \mathbf{v}^{supine} + \epsilon$, where ϵ represents some uncertainty, and can be modelled with a zero mean bivariate normal distribution $\epsilon \sim \mathcal{N}(0, \Sigma)$, with $\Sigma = diag(\sigma_c^2, \sigma_\theta^2)$. Finally we recognise that the position of a neighbouring site becomes more uncertain as the displacement
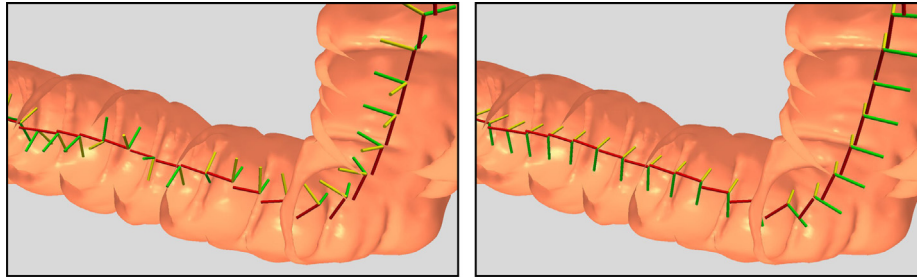
Fig. 4. Parameterisation based on Frenet-Serret (left) and Rotation Minimising (right) frames.
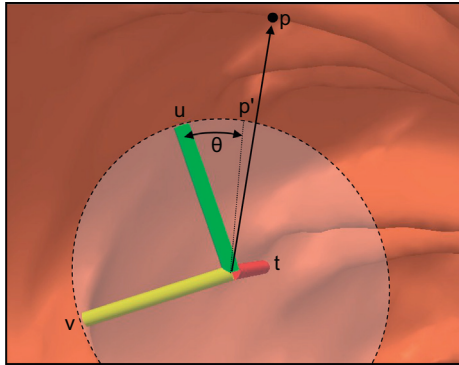

Fig. 5. Calculation of relative angle of rotation of haustral fold position with respect to an adapted frame.

along the centreline increases. In order to model this uncertainty we look at the distribution of stretching and rotation of corresponding fold pairs between the prone and supine views, with respect to the centreline distance between the two folds.

By observing the variance in centreline and rotational (Fig. 6) displacement over a set of training data (described in detail in Section 4.2), a second order polynomial function can be fitted to model the error:

$$\sigma_c^2 = a_2 v_c^2 + a_1 v_c + a_0, \tag{12}$$

$$\sigma_\theta^2 = b_2 v_c^2 + b_1 v_c + b_0. \tag{13}$$

With this information, a pair-wise cost for assigning neighbouring sites i and j label configurations f_i and f_j, is defined by the negative log-normal distribution:

$$\varphi_{ij}(f_i, f_j) = \frac{(\boldsymbol{v}^{prone} - \boldsymbol{v}^{supine})^T \boldsymbol{\Sigma}_{ij}^{-1} (\boldsymbol{v}^{prone} - \boldsymbol{v}^{supine})}{2}. \tag{14}$$

As the angle element of vector $(\boldsymbol{v}^{prone} - \boldsymbol{v}^{supine})$ can only be in the range $[\pi/2, -\pi/2]$, $\pm\pi$ is added until this requirement is met. A local neighbourhood system is defined in order to enforce local geometric constraints on neighbouring fold positions. The local neighbourhood of a site is set to be:

$$N_i^{local} = \{\{j\} | j \neq i, \|(v_c)_i - (v_c)_j\| < k\}, \tag{15}$$

where k is a threshold distance. Pairs of sites that are separated by a local colonic collapse are removed from the neighbourhood set. A uniqueness constraint is also enforced so any two sites may not be assigned the same label. This is included in the pair-wise cost function by connecting each site with every other site in a global neighbourhood system:

$$N_i^{global} = \left\{ \{j\} | j \neq i, j \notin N_i^{local} \right\} \tag{16}$$

and defining the pair-wise cost of assigning the same label to two different nodes to be infinity, except in the case of a null label assignment. The full pair-wise cost function is:

$$\Theta_{ij}(f_i, f_j) = \begin{cases} \infty & \text{if } f_i = f_j \neq f_0 \\ \beta & \text{if } S_j \in N_i^{local} \wedge \\ & (f_i = f_0 \vee f_j = f_0) \\ \varphi(f_i, f_j) & \text{if } S_j \in N_i^{local} \wedge \\ & f_i \neq f_0 \wedge f_j \neq f_0 \\ 0 & \text{otherwise} \end{cases}. \tag{17}$$
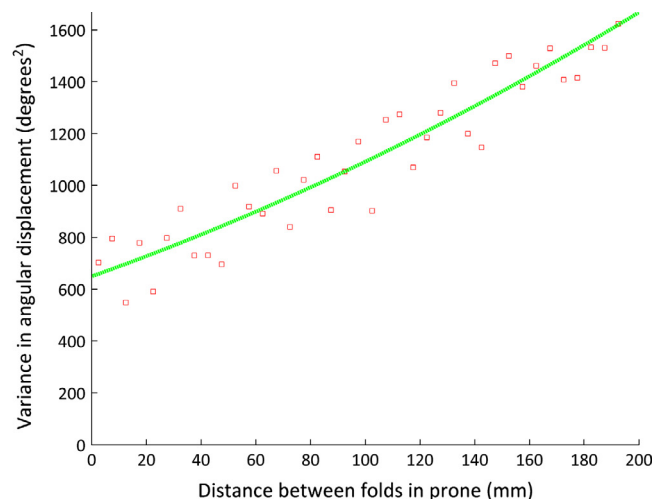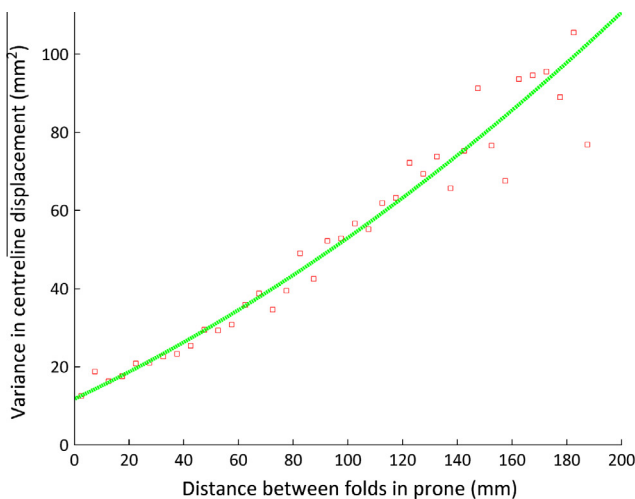

Fig. 6. Distance between folds against variance in centreline (left) and rotational displacement (right) between corresponding fold pairs.

2.2.3. MRF inference

Submodularity influences the choice of inference algorithm that may be applied to the problem of generating the optimal labelling. A submodular energy function is one which, for all labels, satisfies:

$$\Theta(f_d, f_e) + \Theta(f_g, f_g) \leqslant \Theta(f_d, f_g) + \Theta(f_g, f_e), \tag{18}$$

which must hold for all labels $f_d, f_e, f_g \in \mathbf{f}$. The uniqueness constraint on the pair-wise costs means the problem of solving the MRF is non-submodular as $\Theta(f_g, f_g) = \infty$, which restricts the possible algorithm choice for MAP inference. The min-sum Belief Propagation (BP) algorithm is suitable for this purpose (Weiss and Freeman, 2002). It is known that BP is exact on acyclic tree-like graphical models, but has been shown to give a good MAP estimate in graphs with loops. The BP algorithm works by passing messages between nodes of a graph defined by the set of sites $\mathbf{S}$, with edges defined by the site neighbourhoods $\mathbf{N}$. Each message M is an i dimensional vector, with i equal to the number of possible labels. At each iteration at time t, every node sends messages to each of its neighbours in parallel, whilst also receiving messages itself. Let $m_{p \to q}^t$ be the message that node p passes to node q at iteration t. All entries in $m_p^0 \to q$ are initialised to zero. At each iteration, new messages are computed as follows:

$$m_{i \to j}^t(f_j) = \min_{f_i}\left(\Theta_i(f_i) + \Theta_{ij}(f_i, f_j) + \sum_{s \in N_i \setminus j} m_{s \to i}^{t-1}(f_i)\right), \tag{19}$$

where $N_i \setminus j$ denotes all neighbours of i other than j. After T iterations, the belief vector for each node may be computed:

$$b_j(f_j) = -\left[\Theta_j(f_j) + \sum_{s \in N_j} m_{s \to j}^T(f_j)\right]. \tag{20}$$

The belief vector $b_j(f_j)$ expresses the negative cost of assigning each label f_j to site j. The algorithm will converge on a solution if the system reaches a state where the messages stabilise; however, this is not guaranteed in a graph with loops as the system may flip between a pair of states in each iteration. Convergence can be aided with the use of *message damping*, where the actual messages at time t are computed:

$$m_{p \to q}^{t'}(f_q) = (1 - \tau)m_{p \to q}^t(f_q) + \tau m_{p \to q}^{t-1}(f_q). \tag{21}$$

Once the algorithm has terminated, each node is assigned the label having the maximum belief:

$$f_q^* = \arg\max_{f_q \in \mathbf{f}} b_q(f_q). \tag{22}$$

2.2.4. Inverse consistency

In the methods presented, the algorithm is configured with the haustral folds in the prone view as the set of nodes in the MRF, and the haustral folds in the supine view as the labels (or vice versa). The algorithm is run in both directions and the intersection of the labelling results is used. For each node where null labels are present in both directions or the labelling in one direction differs from the labelling in the other direction, the null label is assigned. This results in an increase in the accuracy of the fold labelling at the expense of an increase in null label assignments.

2.2.5. Parameter training

The parameters to model the error in the pair-wise cost function are derived from the statistical analysis of reference standard fold matches in the training data described in Section 4.2. W, k, α, β and τ; the weighting of unary to pair-wise cost functions, node neighbourhood threshold, costs for assigning the null label in the unary and pair-wise cost function, and the message damping parameter respectively, are found using a gradient ascent optimisation. Parameters are trained on separate datasets to those used for validation and are listed below.

$a_2 - 8.19 \times 10^{-4}$
$a_1 - 0.331$
$a_0 - 11.8$
$b_2 - 6.67 \times 10^{-3}$
$b_1 - 3.76$
$b_0 - 650$
$W - 8.87 \times 10^{-3}$
$k - 200$
$\alpha - 6.23 \times 10^{-2}$
$\beta - 0.839$
$\tau - 0.813$

3. Surface based registration

There are scenarios where obtaining a one-to-one surface correspondence is required, such as to locate a possible polyp position in both the prone and supine views. In this case, the results of this fold matching method can be used to provide automated initialization for a surface-based registration technique (Roth et al., 2011).

The intensity-based registration (IBR) algorithm proposed by Roth et al. (2011) recognises that the colon is topologically cylindrical and reduces the complexity of the registration by mapping each point on the endoluminal surface onto a cylindrical representation with the use of a conformal mapping technique (Hamilton, 1982). This allows the registration to account for the large 3D deformations between the prone and supine views as a more simple 2D cylindrical deformation. The registration is then represented as a transformation between the two cylinders and includes non-linear stretching along the cylinder, and local torsion and rotation. A shape index metric is calculated at each point of the cylindrical image and used to drive a B-spline intensity based registration (Rueckert et al., 1999) in a cylindrical domain to achieve correspondence between the two views.

The following different methods are compared:

- *Linear Scaling Initialisation with Intensity Based Registration (LSI w/ IBR)*: We use the locations of the hepatic and splenic flexures as in Roth et al. (2011). These are automatically detected based on local maxima of the z-coordinate of the centreline. The flexure positions may be discarded if their centreline distance vary by more than a certain threshold. The positions of found flexure are mapped onto the cylindrical images and used to provide a linear scaling of prone image in the direction of the centreline. This is used as an initialisation to the intensity-based registration in Roth et al. (2011)
- *B-Spline Initialisation without Intensity Based Registration (BSI w/o IBR)*: The positions of the detected haustral folds and the corresponding locations in the prone and supine views are used to perform a multilevel B-spline point based deformation of the prone image (see Section 3.1).
- *B-Spline Initialisation with Intensity Based Registration (BSI w/ IBR)*: The same as above, but used as an initialisation to the intensity-based registration.

To determine the registration error, each reference standard point is transformed from one dataset to the other using the registration result, and the 3D Euclidean distance between this and the corresponding reference standard point is measured.

3.1. B-Spline Initialisation (BSI)

We wish to approximate a smooth function f which relates the (x, y) points in the prone unfolded image, to their (x', y') positions in the supine image over domain $\Omega = (x, y)|0 \leqslant x < m, 0 \leqslant y < n$. To

do this we use the set of folds $P = \{(x_c, y_c, v_c)\}$ where (x_c, y_c) is a point in Ω and v_c is the value at (x_c, y_c). We define the mapping in terms of two functions: $x' = X(x, y)$ and $y' = Y(x, y)$. As these functions can be derived simultaneously, we use the notation $v = (x', y') = f(x, y)$. To approximate the data P, we use function f as a uniform bicubic B-spline, defined by control lattice Φ overlaid on domain Ω using the method in Lee et al. (1997). We also assume Φ is an $(m + 3) \times (n + 3)$ lattice, where m and n are the image dimensions defined in lattice control points. We define ϕ_{ij} as the value of ij-th control point on lattice Ω for $i = -1, 0, \ldots, m + 1$ and $j = -1, 0, \ldots, n + 1$. We can then define the approximation function f:

$$f(x, y) = \sum_{k=0}^{3} \sum_{l=0}^{3} B_k(s) B_l(t) \phi_{(i+k)(j+l)}, \tag{23}$$

where $i = \lfloor x \rfloor - 1$, $j = \lfloor y \rfloor - 1$, $s = x - \lfloor x \rfloor$, and $t = y - \lfloor y \rfloor$. B_k and B_l are basis functions:

$$\begin{aligned} B_0(t) &= (1 - t)^3 / 6 \\ B_1(t) &= (3t^3 - 6t^2 + 4)/6 \\ B_2(t) &= (-3t^3 + 3t^2 + 3t + 1)/6 \\ B_3(t) &= (t^3)/6, \end{aligned} \tag{24}$$

where $0 \leqslant t < 1$. For every point in $P = \{(x_c, y_c, v_c)\}$ a different value ϕ_c of each of the control points ϕ_{ij} is defined:

$$\phi_c = \frac{w_c v_c}{\sum_{a=0}^{3} \sum_{b=0}^{3} w_{ab}^2}, \tag{25}$$

where $w_c = w_{kl} = B_k(s) B_l(t)$, $k = (i + 1) - \lfloor x_c \rfloor$, $l = (j + 1) - \lfloor y_c \rfloor$, $s = x_c - \lfloor x_c \rfloor$, $t = y_c - \lfloor y_c \rfloor$. Only data points in the 4×4 neighbourhood of each control point are taken into consideration. To choose a value for each ϕ_{ij} from the contributions from each point ϕ_c the error $e(\phi_{ij}) = \sum_c (w_c \phi_{ij} - w_c \phi_c)^2$ is minimised by differentiating $e(\phi_{ij})$ with respect to ϕ_{ij} giving:

$$\phi_{ij} = \frac{\sum_c w_c^2 \phi_c}{\sum_c w_c^2}. \tag{26}$$

To allow for a smooth function over the entire domain and more accurate local deformations, a multilevel B-spline approximation is used to generate a hierarchy of control lattices from coarse to fine. A refinement process is used to reduce the sum of these functions into one B-spline function. For each level of control lattice Ψ_k we can derive a finer control lattice Ψ'_k such that $F(\Psi'_{k-1}) = F(\Psi_{k-1})$ and $F(\Psi'_0) = f_0$ where $f_0(x_c, y_c) = (x_c, y_c)$. We then derive control lattice Φ_k to approximate data $P_k = \{(x_c, y_c, \Delta_k v_c)\}$, where $\Delta_k v_c = v_c - \sum_{i=0}^{k-1} f_i(x_c, y_c) = \Delta_{k-1} v_c - f_{k-1}(x_c, y_c)$, and $\Delta_0 v_c = v_c$. Each function defined by control lattice Φ_k serves to remove the residual error from the refined coarser lattice Ψ'_{k-1} at each level. We can now define a progressive control lattice $\Psi_k = \Psi'_{k-1} + \Phi_k$ from the coarsest to finest levels. As the end points of the colon are fixed, the image should not be transformed in the direction of the x-axis at the x-axis extrema, so at each refinement we calculate the residual error Φ_k and set the x-component of the data $\Delta x' = 0$ at points $\{\phi_{ij} | i = -1, 0, m, m + 1; j = -1, 0, \ldots, n + 1\}$.

We apply this technique to the images created by a conformal mapping (Roth et al., 2011) of the endoluminal surfaces of prone and supine images onto a rectangular domain. The sparse set of data points $P = \{(x_c, y_c, v_c)\}$ have their positional information $\{(x_c, y_c)\}$ taken from the positions of haustral folds mapped onto the 2D domain, and the vertical and horizontal displaced positions $\{v_c = (x'_c, y'_c)\}$ of the corresponding positions in the supine image. To allow for a pseudo-continuous function over the y-axis, the image is repeated and stacked over the y-axis. Due to the true cylin-

drical nature of the registration problem, there is an ambiguity over the direction of vertical displacement in the 2D images, resulting from the periodicity of the boundary in the vertical dimension. To create a smooth displacement, the B-spline fitting is repeated and at each iteration the datum P_c with the maximum error between the y component of the estimated and true displacement $e_y = |(F(x_c, y_c) - v_c)_y|$ is adjusted such that $v'_c = v_c + \text{sign}((F(x_c, y_c) - v_c)_y) \cdot y_{max}$ where y_{max} is the size of the image in the y-direction. The image is then shifted in the y-direction so as to minimise v_c and the full multi-level B-spline fitting is repeated to give the final function $F(\Psi)$ with control lattice Ψ.

Now for every position in the prone image $P_{prone} = \{(x, y)\} \in \Omega$ we can use the function F to find the corresponding position in the supine image $P_{supine} = \{(x, y)\} \in \Omega$. We can use this transformation as an initialisation to the intensity-based B-spline registration function presented in Roth et al. (2011) to create a finer composite registration.

4. Experimental results

4.1. Clinical validation

Ethical approval and informed consent were obtained to use anonymised CT colonography data. Colonic cleansing and insufflations had been performed in accordance with current recommendations (Taylor et al., 2007) (see Fig. 7).

4.2. Haustral fold matching

For the purpose of establishing a fold labelling between the prone and supine acquisitions, we selected the same 34 patient cases used in a previous publication (Roth et al., 2011). In 24 of the cases, the colon was optimally distended in both views, and where fluid tagging (allowing for digital cleaning of residual fluid) was used or little fluid remained. This allowed a continuous segmentation over the full length of the colon using the methods described in Roth et al. (2011). The other 10 cases exhibit local colonic collapse and are used to further validate the method. The datasets were randomly allocated into training and validation sets using a random permutation, resulting in training and validation sets both with 17 prone and supine cases of which 5 exhibited one or more areas of local colonic collapse (see Table 1 for details), and 4 cases which had been excluded from the previous study due to marked differences in local distension and therefore different surface features (cases 9–12). A subset of the cases used in this study are shown in Fig. 8.

Two radiologists (EH, AP) and a computer scientist (TH) with experience of reading CTC images, independently established a reference standard by matching haustral folds using virtual colonoscopic reconstructions, external renderings of the endoluminal surface, and unfolded images achieved by performing a conformal mapping of the endoluminal surface mesh onto a plane. Any folds where confident manual correspondences between the two views could not be established were excluded from the derived reference standard. All readers were unaware of the algorithm results. The reference standards were compared for consistency, and any discrepancies were resolved by the three readers in consensus. This resulted in a total of 1743 corresponding fold pairs over the 17 validation datasets. To assess algorithm performance, for each case the maximum a posteriori labelling solution is compared against the reference standard described above. The number of reference standard points that were correctly matched by the algorithm is given by the number of 'hits'. Similarly, the number of reference standard points that were incorrectly matched by the algorithm is given by the number of 'misses'. The 'misses' consist of the union
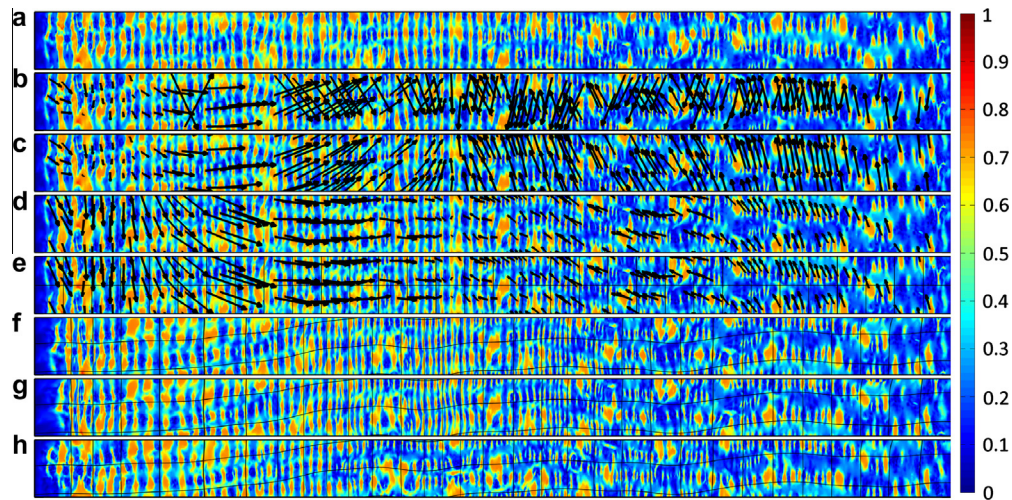
Fig. 7. Images of the endoluminal surface produced from the conformal mapping technique (case 11) showing the full registration with B-spline initialisation. The colour scheme shows the Shape Index (SI) and the vectors show the displacements generated from the landmark registration. Images show: (a) the source (prone) image; (b) the ambiguous vector direction on the source image; (c) the sorted displacements; (d) the source image vertically aligned to reduce displacements; (e) the source image with displacement vectors and regular grid; (f) the result of the landmark B-spline initialisation with transformed image and grid; (g) the refinement with the intensity based registration (with same grid); (h) the target (supine) image.

Table 1
Information of cases exhibiting local luminal collapse. For each case, the number of collapsed regions in the prone and supine images are displayed, along with the Euclidean distance across each region from centreline end-point to centreline start-point. Locations of collapse are given (DC: descending colon; SC: sigmoid colon).

Case	Prone			Supine		
	No. collapses	Location	Distance (mm)	No. collapses	Location	Distance (mm)
13	1	DC	65.0	0	–	–
14	1	DC	245.1	1	DC	272.4
15	0	–	–	1	SC	26.0
16	3	DC	6.5	0	–	–
		DC	34.4		–	–
		SC	8.0		–	–
17	0	–	–	1	DC	18.3

of the set of reference standard points that have been assigned the incorrect label, and the set of labels that have been assigned to the incorrect reference standard point. The resulting accuracy for each case is calculated as *hits/(hits + misses)*.

Table 2 shows the performance of the algorithm using the unary cost function alone. Table 3 shows the performance using the pairwise cost function only, with and without inverse consistency. In all cases the unary prior based on centreline distance and the pairwise uniqueness constraint are used to allow for a fair comparison. Table 4 shows the results using the full MRF cost function, with and without the inverse consistency constraint. Table 7 and summarise the performance of the individual cost functions with respect to the case features (fully distended or collapsed), in terms of accuracy and total number of output labels. It is apparent that across all datasets, using only the unary cost function results in a

Table 2
Fold labelling performance using the unary costs only, with and without inverse consistency. The number of folds identified in the Reference Standard (RS) are shown along with the number of algorithm output labels. From the intersection of the reference standard folds and the algorithm labelled folds, the number of correct hits and incorrect misses are displayed along with the accuracy calculated from these.

Case	RS points	Unary costs only				Unary costs with inverse consistency			
		Labels	Hits	Misses	Accuracy (%)	Labels	Hits	Misses	Accuracy (%)
Fully distended									
1	76	77	24	19	55.8	28	16	7	69.6
2	134	106	37	42	46.8	24	14	8	63.6
3	127	103	44	39	53.0	39	31	4	88.6
4	64	101	33	23	58.9	20	13	0	100.0
5	85	95	37	33	52.9	22	12	3	80.0
6	137	121	35	46	43.2	33	16	8	66.7
7	166	137	42	64	39.6	28	19	4	82.6
8	76	79	23	22	51.1	15	12	1	92.3
9	92	114	16	46	25.8	20	7	7	50.0
10	70	89	13	27	32.5	13	7	1	87.5
11	177	164	34	80	29.8	36	17	11	60.7
12	72	80	21	16	56.8	18	6	4	60.0
Subset mean	106.3	105.5	29.9	38.1	45.5	24.7	14.2	4.8	75.1
Collapsed									
13	96	88	31	26	54.4	22	19	3	86.4
14	94	94	17	44	27.9	14	5	8	38.5
15	61	87	20	27	42.6	19	15	2	88.2
16	107	122	22	50	30.6	14	7	1	87.5
17	109	148	36	35	50.7	34	21	7	75.0
Subset mean	93.4	107.8	25.2	36.4	41.2	20.6	13.4	4.2	75.1
Total mean	102.5	106.2	28.5	37.6	44.3	23.5	13.9	4.6	75.1
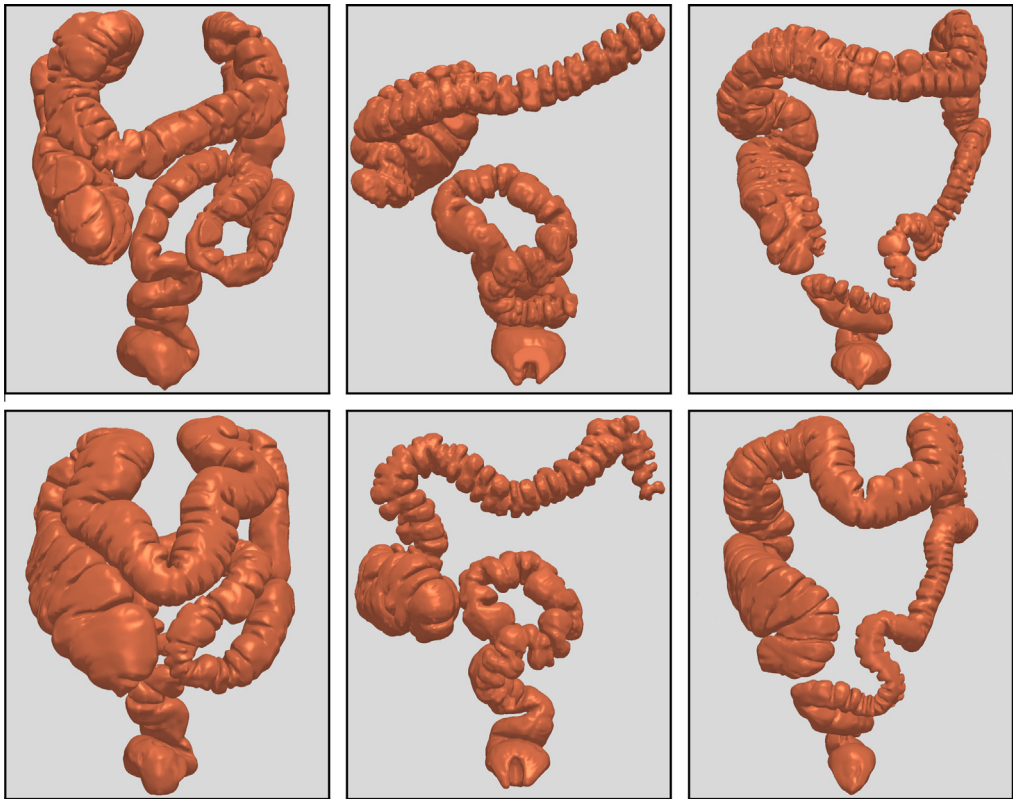
Fig. 8. Surface rendered examples of a subset of the cases used for validation. Top row shows prone view. Bottom row shows supine view. Cases shown from left to right are: 9, 14 and 16.

Table 3
Fold labelling performance using the pair-wise costs only, with and without inverse consistency. The number of folds identified in the Reference Standard (RS) are shown along with the number of algorithm output labels. From the intersection of the reference standard folds and the algorithm labelled folds, the number of correct hits and incorrect misses are displayed along with the accuracy calculated from these.

Case	RS points	Pair-wise costs only				Pair-wise costs with inverse consistency			
		Labels	Hits	Misses	Accuracy (%)	Labels	Hits	Misses	Accuracy (%)
Fully distended									
1	76	71	39	9	81.3	59	34	6	85.0
2	134	104	87	6	93.5	103	87	5	94.6
3	127	132	101	7	93.5	131	101	7	93.5
4	64	72	30	12	71.4	0	0	0	N/A
5	85	79	73	0	100.0	75	69	0	100.0
6	137	145	98	8	92.5	123	81	7	92.0
7	166	150	136	2	98.6	149	135	2	98.5
8	76	78	59	1	98.3	77	58	1	98.3
9	92	85	47	7	87.0	65	44	0	100.0
10	70	74	35	10	77.8	40	34	1	97.1
11	177	156	79	39	66.9	105	77	6	92.8
12	72	76	38	7	84.4	46	33	4	89.2
Subset mean	106.3	101.8	68.5	9.0	87.1	81.1	62.8	3.3	94.6
Collapsed									
13	96	88	59	11	84.3	79	59	6	90.8
14	94	78	22	32	40.7	23	22	0	100.0
15	61	69	39	8	83.0	58	39	3	92.9
16	107	100	24	40	37.5	69	23	25	47.9
17	109	136	85	4	95.5	110	83	4	95.4
Subset mean	93.4	94.2	45.8	19.0	68.2	67.8	45.2	7.6	85.4
Total mean	102.5	99.6	61.8	11.9	81.5	77.2	57.6	4.5	91.8

poor labelling accuracy with a mean of 44.2%, which is unsurprising since neighbouring haustral folds appear similar. Performance by using the pair-wise cost function alone is considerably better with mean accuracy of 81.6%. It is interesting to compare the pair-wise cost function results with the full model: using the sub-set of cases with a fully distended colon, mean accuracy only increases from 87.1% to 90.1%, compared to the much larger difference in the subset of cases exhibiting local colonic collapse where mean accuracy increases from 68.2% to 79.4%. There are two possible explanations: Firstly, the pair-wise cost function

Table 4
Fold labelling performance using the full MRF model, with and without inverse consistency. The number of folds identified in the Reference Standard (RS) are shown along with the number of algorithm output labels. From the intersection of the reference standard folds and the algorithm labelled folds, the number of correct hits and incorrect misses are displayed along with the accuracy calculated from these.

Case	RS points	Full model				Full model with inverse consistency			
		Labels	Hits	Misses	Accuracy (%)	Labels	Hits	Misses	Accuracy (%)
Fully distended									
1	76	95	57	9	86.4	93	57	8	87.7
2	134	137	115	5	95.8	136	115	5	95.8
3	127	143	118	1	99.2	140	116	1	99.1
4	64	85	50	4	92.6	78	49	3	94.2
5	85	97	82	1	98.8	94	82	0	100.0
6	137	168	116	7	94.3	152	115	4	96.6
7	166	176	154	3	98.1	175	153	3	98.1
8	76	95	69	1	98.6	94	69	1	98.6
9	92	106	43	18	70.5	61	39	0	100.0
10	70	101	58	3	95.1	69	56	0	100.0
11	177	176	124	18	87.3	151	123	2	98.4
12	72	83	27	15	64.3	45	27	5	84.4
Subset mean	106.3	121.8	84.4	7.1	90.1	107.3	83.4	2.7	96.1
Collapsed									
13	96	110	65	19	77.4	82	63	5	92.6
14	94	76	26	28	48.1	18	18	0	100.0
15	61	87	49	5	90.7	68	45	1	97.8
16	107	148	84	11	88.4	104	81	3	96.4
17	109	147	86	7	92.5	123	84	6	93.3
Subset mean	93.4	113.6	62.0	14.0	79.4	79.0	58.2	3.0	96.0
Total mean	102.5	119.4	77.8	9.1	86.9	99.0	76.0	2.8	96.1

Table 5
Comparison of the individual readers with the consensus reference standard. The number of folds identified in the consensus Reference Standard (RS) are shown along with the number of reader output labels. From the intersection of the reference standard folds and the reader labelled folds, the number of correct hits and incorrect misses are displayed along with the accuracy calculated from these.

| Case	RS points	Reader 1 performance				Reader 2 performance			
		Labels	Hits	Misses	Accuracy (%)	Labels	Hits	Misses	Accuracy (%)
Fully distended									
1	76	68	62	2	96.9	76	61	1	98.4
2	134	133	126	0	100.0	103	94	2	97.9
3	127	142	124	0	100.0	109	104	0	100.0
4	64	58	45	0	100.0	61	45	2	95.7
5	85	79	74	1	98.7	84	80	0	100.0
6	137	136	120	0	100.0	120	100	8	92.6
7	166	135	134	0	100.0	157	148	1	99.3
8	76	86	74	0	100.0	74	59	4	93.7
9	92	85	75	0	100.0	92	66	5	93.0
10	70	77	66	0	100.0	81	45	5	90.0
11	177	188	168	0	100.0	138	111	10	91.7
12	72	64	55	0	100.0	89	59	2	96.7
Subset mean	106.3	104.3	93.6	0.3	99.6	98.7	81.0	3.3	95.8
Collapsed									
13	96	75	73	0	100.0	98	80	6	93.0
14	94	52	52	0	100.0	97	85	0	100.0
15	61	61	56	0	100.0	76	53	4	93.0
16	107	96	85	0	100.0	113	83	1	98.8
17	109	133	102	0	100.0	103	87	5	94.6
Subset mean	93.4	83.4	73.6	0.0	100.0	97.4	77.6	3.2	95.9
Total mean	102.5	98.1	87.7	0.2	99.7	98.3	80.0	3.3	95.8

model will more likely converge to the correct solution when there is similarity in the pattern of segmented folds in the 2D parameter-ised space. In cases where local collapse is exhibited, the folds are less likely to be segmented correctly and therefore the likelihood of pattern similarity between the two acquisitions decreases. Secondly, the pair-wise cost function is more likely to converge on the correct solution when there is a fully connected graph along the entire colonic length. The local neighbourhood network of edges that make up the graph structure in the MRF are disconnected where there is local collapse as the geometric relationships

between pairs of folds across this space cannot be as easily determined. This creates two subgraphs which may converge on the incorrect labelling solution without the unary costs in place. The inverse consistency constraint gives a better performance in terms of accuracy for the unary cost only, pair-wise cost only, and full model increasing accuracy from 44.3% to 75.1%, 81.5% to 91.8%, and 86.9% to 96.1% respectively. This is however, achieved by reducing the mean number of labels: from 106.2 to 23.5, 99.6 to 77.2, and in the case of the full model from 119.4 to 99.0. It is of interest to note that applying the pair-wise only model with

Table 6
Comparison of the third reader and algorithm output with the consensus reference standard. The number of folds identified in the consensus Reference Standard (RS) are shown along with the number of reader/algorithm output labels. From the intersection of the reference standard folds and the reader/algorithm labelled folds, the number of correct hits and incorrect misses are displayed along with the accuracy calculated from these.

Case	RS points	Reader 3 performance				Algorithm performance			
		Labels	Hits	Misses	Accuracy (%)	Labels	Hits	Misses	Accuracy (%)
Fully distended									
1	76	70	60	2	96.8	93	57	8	87.7
2	134	113	108	0	100.0	136	115	5	95.8
3	127	102	94	0	100.0	140	116	1	99.1
4	64	55	52	0	100.0	78	49	3	94.2
5	85	94	82	1	98.8	94	82	0	100.0
6	137	97	89	1	98.9	152	115	4	96.6
7	166	136	125	1	99.2	175	153	3	98.1
8	76	62	60	2	96.8	94	69	1	98.6
9	92	60	55	0	100.0	61	39	0	100.0
10	70	72	63	1	98.4	69	56	0	100.0
11	177	137	121	6	95.3	151	123	2	98.4
12	72	76	63	1	98.4	45	27	5	84.4
Subset mean	106.3	89.5	81.0	1.3	98.5	107.3	83.4	2.7	96.1
Collapsed									
13	96	75	72	0	100.0	82	63	5	92.6
14	94	67	44	15	74.6	18	18	0	100.0
15	61	63	52	3	94.5	68	45	1	97.8
16	107	84	79	0	100.0	104	81	3	96.4
17	109	86	77	1	98.7	123	84	6	93.3
Subset mean	93.4	75.0	64.8	3.8	93.6	79.0	58.2	3.0	96.0
Total mean	102.5	85.2	76.2	2.0	97.1	99.0	76.0	2.8	96.1
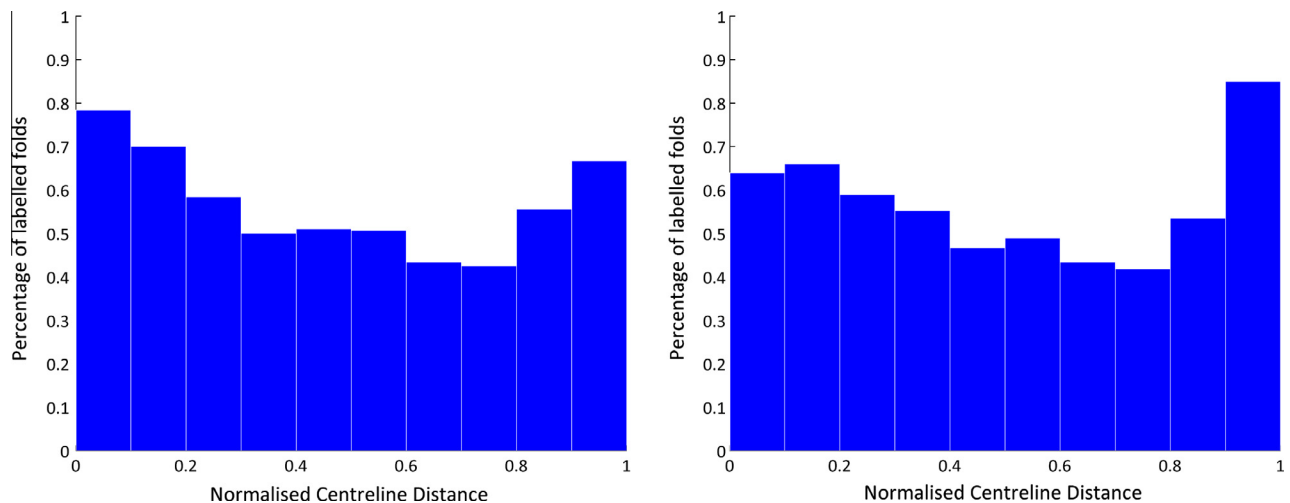
Fig. 9. Normalised distributions of labelled folds with respect to normalised centreline distance (caecum at 0, rectum at 1) in the reference standard (left) and the MAP labelling (right).

Table 7
Summary of mean values from Tables 2–4 with respect to case feature and cost function used.

Case feature	Cost function	RS points	Labels	Hits	Misses	Accuracy (%)
Fully distended	Unary only	106.3	106.3	29.9	38.1	45.5
	Unary with IC	106.3	24.7	14.2	4.8	75.1
	Pair-wise only	106.3	101.8	68.5	9.0	87.1
	Pair-wise with IC	106.3	81.1	62.8	3.3	94.6
	Full model	106.3	121.8	84.4	7.1	90.1
	Full model with IC	106.3	107.3	83.4	2.7	96.1
Collapsed	Unary only	93.4	107.8	25.2	36.4	41.2
	Unary with IC	93.4	20.6	13.4	4.2	75.1
	Pair-wise only	93.4	94.2	45.8	19.0	68.2
	Pair-wise with IC	93.4	67.8	45.2	7.6	85.4
	Full model	93.4	113.6	62.0	14.0	79.4
	Full model with IC	93.4	79.0	58.2	3.0	96.0
Total	Unary only	102.5	106.2	28.5	37.6	44.3
	Unary with IC	102.5	23.5	13.9	4.6	75.1
	Pair-wise only	102.5	99.6	61.8	11.9	81.5
	Pair-wise with IC	102.5	77.2	57.6	4.5	91.8
	Full model	102.5	119.4	77.8	9.1	86.9
	Full model with IC	102.5	99.0	76.0	2.8	96.1

Table 8
Mean fold registration error (mm) for each of the validation cases. Results are shown individually for the full registration with the Linear Scaling Initialisation with Intensity Based Registration (LSI w/ IBR) and the new B-Spline Initialisation (BSI) with and without subsequent IBR.

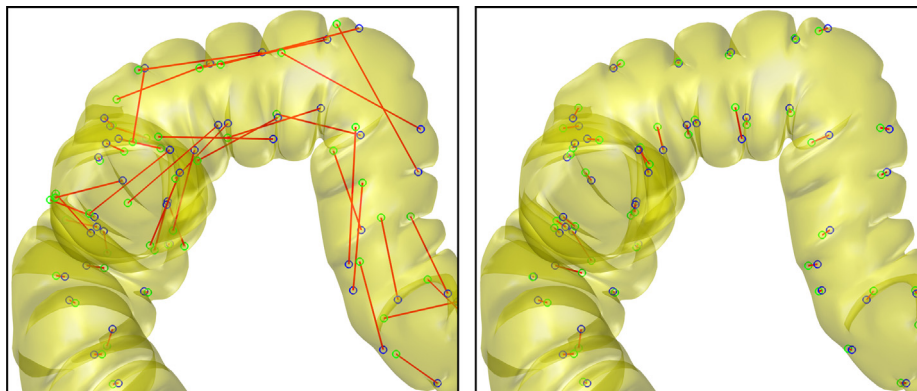| Case	LSI w/ IBS	BSI w/o IBS	BSI w/ IBS
Fully distended			
1	6.9	7.1	6.4
2	6.9	7.4	5.4
3	4.6	5.4	4.5
4	9.2	4.3	4.3
5	5.0	4.4	5.1
6	3.5	4.6	3.6
7	5.0	5.5	5.0
8	5.2	6.0	4.9
Subset mean	5.8	5.6	4.9
Subset Std	1.7	1.1	0.8
Previously excluded cases			
9	48.1	18.8	9.1
10	12.9	9.0	7.8
11	19.7	9.9	5.8
12	14.2	9.9	6.8
Subset mean	23.7	11.9	7.4
Subset Std	14.3	4.0	1.2
Collapsed			
13	27.5	14.9	11.7
14	4.8	12.2	5.7
15	5.1	9.5	5.0
16	17.4	8.1	5.9
17	6.2	6.7	5.4
Subset mean	12.2	10.3	6.7
Subset Std	8.9	3.0	2.5
Total mean	11.9	8.5	6.0
Total Std	11.1	3.8	1.9

and the distribution of labelled folds expressed as a percentage of total folds in that region. Fig. 9 shows that the percentage of labelled folds in the reference standard (total 58.5%) and in the algorithm output (total 56.5%) follow similar distributions, with a higher confidence of fold labelling in the caecum (at 0.0) and rectum (at 1.0), and lower in the transverse and descending colon. This is due to less ambiguity in the positions of corresponding haustral folds at the terminations of the colon as a result of a reduced level of deformation and more straightforward identification of stable anatomical landmarks for which the relative positions of folds can be located.

4.3. Full surface registration

Registration error was assessed by performing the full surface registration using the Linear Scaling Initialisation with Intensity Based Registration (LSI w/ IBR) and the new B-Spline Initialisation (BSI) with and without subsequent IBR using the same cases and reference standard as in Section 4.2. Results are shown in Table 8. It is clear that the BSI w/ IBR method outperforms both the BSI w/o IBR and LSI w/ IBR registration methods with a mean Euclidean error of 6.0 mm (±1.9 mm), compared to 8.5 mm (±3.8 mm) and 11.9 mm (±11.1 mm) respectively. Using a Related Samples Wilcoxon Signed Rank Test on the Euclidean error of all 1743 fold pairs in the final reference standard, the differences in error between the BSI w/o IBR and LSI w/ IBR, BSI w/ IBR and BSI w/o IBR, and BSI w/ IBR and LSI w/ IBR methods are statistically significant with $p = 0.043$, $p < 0.001$ and $p < 0.001$ respectively.

5. Conclusions

We present a novel method for establishing correspondence between two CT colonography acquisitions with the patient in prone and supine positions. First, haustral folds are segmented with a graph-cut method applied to a triangular mesh representation of the colonic lumen segmentation. The method uses depth map images to drive a virtual camera optimisation to provide a unary cost value for the matching of folds between the two views. An additional pair-wise cost function compares the geometric relationship between pairs of haustral folds in the prone and supine CTC images. A parametrisation of the image space exploits the quasi-cylindrical form of the colon and simplifies the description of this geometric relationship by reducing the dimensionality from 3D to 2D. The problem is modelled using a Markov Random Field, and a Belief Propagation algorithm used to estimate the optimal labelling. This process can establish an accurate correspondence

inverse consistency to case 4 gives no labels, as when running algorithm (prone to supine, supine to prone) it converges on entirely different solutions. Also, using the full model with inverse consistency gives a 29% increase in the mean number of output labels compared to using pair-wise cost only.

In Tables 5 and 6 the performance of the algorithm is compared to the performance of the individual readers. It is clear that the algorithm achieves a very similar level of mean accuracy at 96.1% compared to the mean reader accuracy of 97.5%. It also labels a similar mean number of folds at 99.0 per case, compared to the mean reader number of folds at 93.9 per case.

To analyse the distributions of fold labelling in the reference standard and in the output of the algorithm, areas of the colon have been discretised with respect to the normalised centreline distance

Fig. 10. External surface renderings of the transverse colon in the supine image of case 16. The set of reference standard points in the supine view (blue) and the corresponding points transformed from the prone view (green) and shown using the results from the LSI w/ IBS (left) and BSI w/ IBS (right) registration methods. The red lines show the Euclidean distance error. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

between the a set of positions in the two views even in cases where endoluminal collapse occurs, which is very common in clinical practice. We have also given an example of how this method can be applied to initialise a dense intensity-based registration technique, in this case a surface-based method, and show that it significantly reduces mean registration error. While the intensity-based registration method alone can be susceptible to misregistration by one or more haustral folds (Fig. 10), or by a degree of rotation around the tenaie coli, the new composite method defends against this problem by using a landmark-based initialisation. Moreover, the initialisation could be generalisable to other registration methods.

The work flow presented is fully automated, taking as input a prone and supine colon lumen segmentation, and in disconnected cases, the ordering of those segments. The consistency of results across cases exhibiting varying characteristics indicates that the composite method is more robust than those previously reported, especially in more 'difficult' cases showing marked differences in distension, or exhibiting areas of endoluminal collapse. This situation is very common in routine practice and algorithms must be able to cope in order to have clinical utility.

Although algorithm performance is similar for both well-distended cases and those cases exhibiting one or more areas of endoluminal collapse, the number of identified corresponding folds can decrease. In future work we will investigate inclusion of an additional unary prior based on the relative location of stable anatomical landmarks to aid MRF convergence and increase the number of output correspondences in difficult cases. The current method relies on manual ordering of collapsed segments, so a proposed extension will automatically find their order using the same MRF model. It would also be interesting to investigate an interactive system such that landmark points could be selected manually on the virtual colonoscopic views, and used in the MRF optimisation. We will also research MRF optimisation schemes that include the inverse consistency constraint directly in the optimisation, rather than after algorithm convergence.

To conclude, we have presented a novel method for matching haustral folds between prone and supine CT colonography acquisition. We give an example of how this may be used to improve the results of a full surface-based registration by validation using a set of 1743 reference standard points over 17 patient cases exhibiting a variety of characteristics.

Acknowledgements

References

Boykov, Y., Kolmogorov, V., 2004. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1124–1137.

Burling, D., Halligan, S., Slater, A., Noakes, M., Taylor, S., 2006. Potentially serious adverse events at ct colonography in symptomatic patients: National Survey of the United Kingdom. Radiology 239, 464–471.

Cunningham, D., Atkin, W., Lenz, H., Lynch, H., Minsky, B., Nordlinger, B., Starling, N., 2010. Colorectal cancer. The Lancet 375, 1030–1047.

Fletcher, R., Powell, M., 1963. A rapidly convergent descent method for minimization. The Computer Journal 6, 163–168.

Fukano, E., Oda, M., Kitasaka, T., Suenaga, Y., Takayama, T., Takabatake, H., Mori, M., Natori, H., Nawano, S., Mori, K., 2010. Haustral fold registration in CT colonography and its application to registration of virtual stretched view of the colon. Proceedings of SPIE 7624, 762420.

Hamilton, R., 1982. Three-manifolds with positive ricci curvature. Journal of Differential Geometry 17, 255–306.

Hoppe, H., 1999. New quadric metric for simplifiying meshes with appearance attributes. In: Proceedings of the Conference on Visualization'99: Celebrating Ten Years. IEEE Computer Society Press, pp. 59–66.

Jin, M., Kim, J., Luo, F., Gu, X., 2008. Discrete surface ricci flow. IEEE Transactions on Visualization and Computer Graphics 14, 1030–1043.

Johnson, C., Chen, M., Toledano, A., Heiken, J., Dachman, A., Kuo, M., Menias, C., Siewert, B., Cheema, J., Obregon, R., et al., 2008. Accuracy of ct colonography for detection of large adenomas and cancers. New England Journal of Medicine 359, 1207–1217.

Koenderink, J., 1990. Solid shape. MIT Press Cambridge, MA, USA.

Lee, S., Wolberg, G., Shin, S., 1997. Scattered data interpolation with multilevel B-splines. IEEE Transactions on Visualization and Computer Graphics 3, 228–244.

Lengyel, E., 2004. Mathematics for 3D game programming and computer graphics. Cengage Learning.

Li, P., Napel, S., Acar, B., Paik, D., Jeffrey Jr., R., Beaulieu, C., 2004. Registration of central paths and colonic polyps between supine and prone scans in computed tomography colonography: pilot study. Medical Physics 31, 2912.

Näppi, J., Okamura, A., Frimmel, H., Dachman, A., Yoshida, H., 2005. Region-based supine-prone correspondence for the reduction of false-positive CAD polyp candidates in CT colonography. Academic Radiology 12, 695–707.

Pignone, M., Sox, H., 2008. Screening for colorectal cancer: US preventive services task force recommendation statement. Annals of Internal Medicine 149.

Punwani, S., Halligan, S., Tolan, D., Taylor, S., Hawkes, D., 2009. Quantitative assessment of colonic movement between prone and supine patient positions during ct colonography. British Journal of Radiology 82, 475–481.

Roth, H., McClelland, J., Boone, B., Modat, M., Cardoso, M., Hampshire, T., Hu, M., Punwani, S., Ourselin, S., Slabaugh, G., Halligan, S., Hawkes, D., 2011. Registration of the endoluminal surfaces of the colon derived from prone and supine CT colonography. Medical Physics 38, 3077–3089.

Roth, H., McClelland, J., Boone, D., Hu, M., Ourselin, S., Slabaugh, G., Halligan, S., Hawkes, D., Medicsight, P., 2010. Conformal Mapping of the Inner Colon Surface to a Cylinder for the Application of Prone to Supine Registration.

Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M., Hawkes, D., 1999. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Transactions on Medical Imaging 18, 712–721.

Suh, J., Wyatt, C., 2009. Deformable registration of supine and prone colons for computed tomographic colonography. Journal of Computer Assisted Tomography 33, 902.

Taylor, S., Halligan, S., Goh, V., Morley, S., Bassett, P., Atkin, W., Bartram, C., 2003. Optimizing colonic distention for multi-detector row CT colonography: effect of hyoscine butylbromide and rectal balloon catheter. Radiology 229, 99.

Taylor, S., Laghi, A., Lefere, P., Halligan, S., Stoker, J., 2007. European society of gastrointestinal and abdominal radiology (esgar): consensus statement on CT colonography. European Radiology 17, 575–579.

von Wagner, C., Ghanouni, A., Halligan, S., Smith, S., Dadswell, E., Lilford, R., Morton, D., Atkin, W., Wardle, J., 2012. Patient acceptability and psychologic consequences of CT colonography compared with those of colonoscopy: results from a multicenter randomized controlled trial of symptomatic patients. Radiology 263, 723–731.

Wang, S., Petrick, N., Van Uitert, R., Periaswamy, S., Wei, Z., Summers, R., 2012. Matching 3-d prone and supine ct colonography scans using graphs. IEEE Transactions on Information Technology in Biomedicine 16, 676–682.

Wang, S., Yao, J., Liu, J., Petrick, N., Van Uitert, R., Periaswamy, S., Summers, R., 2009. Registration of prone and supine CT colonography scans using correlation optimized warping and canonical correlation analysis. Medical Physics 36, 5595.

Wang, W., Joe, B., 1997. Robust computation of the rotation minimizing frame for sweep surface modeling. Computer-Aided Design 29, 379–391.

Weiss, Y., Freeman, W., 2002. On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs. IEEE Transactions on Information Theory 47, 736–744.

World Health Organisation, 2009. Cancer.

Zeng, W., Marino, J., Gurijala, K., Gu, X., Kaufman, A., 2010a. Supine and prone colon registration using quasi-conformal mapping. IEEE Transactions on Visualization and Computer Graphics 16, 1348–1357.

Zeng, W., Samaras, D., Gu, D., 2010b. Ricci flow for 3D shape analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 32, 662–677.