2007

Using Pareto Fronts to Evaluate Polyp Detection Algorithms for CT Colonography

Adam Huang
National Institutes of Health

Jiang Li
National Institutes of Health, jli@odu.edu

Ronald M. Summers
National Institutes of Health

Nicholas Petrick
U.S. Food and Drug Administration

Amy K. Hara
Mayo Clinic

Original Publication Citation

Huang, A., Li, J., Summers, R. M., Petrick, N., & Hara, A. K. (2007) Using Pareto fronts to evaluate polyp detection algorithms for CT colonography. In M.L. Giger, N. Karssemeijer (Eds.), *Medical Imaging 2007: Computer-Aided Diagnosis, Proceedings of SPIE Vol. 6514* (651407). SPIE of Bellingham, WA. https://doi.org/10.1117/12.709426

Using Pareto Fronts to Evaluate Polyp Detection Algorithms for CT Colonography

Adam Huang[a], Jiang Li[a], Ronald M. Summers[a], Nicholas Petrick[b], and Amy K. Hara[c]*

[a]Diagnostic Radiology Department, Clinical Center
National Institutes of Health, Bethesda, MD 20892-1182
[b]LAMIS Image Analysis Laboratory, Food and Drug Administration
12720 Twinbrook Parkway, Rockville, MD 20852
[c]Diagnostic Radiology, Mayo Clinic in Scottsdale
13400 East Shea Boulevard, Scottsdale, AZ 85259

ABSTRACT

We evaluate and improve an existing curvature-based region growing algorithm for colonic polyp detection for our CT colonography (CTC) computer-aided detection (CAD) system by using Pareto fronts. The performance of a polyp detection algorithm involves two conflicting objectives, minimizing both false negative (FN) and false positive (FP) detection rates. This problem does not produce a single optimal solution but a set of solutions known as a Pareto front. Any solution in a Pareto front can only outperform other solutions in one of the two competing objectives. Using evolutionary algorithms to find the Pareto fronts for multi-objective optimization problems has been common practice for years. However, they are rarely investigated in any CTC CAD system because the computation cost is inherently expensive. To circumvent this problem, we have developed a parallel program implemented on a Linux cluster environment. A data set of 56 CTC colon surfaces with 87 proven positive detections of polyps sized 4 to 60 mm is used to evaluate an existing one-step, and derive a new two-step region growing algorithm. We use a popular algorithm, the Strength Pareto Evolutionary Algorithm (SPEA2), to find the Pareto fronts. The performance differences are evaluated using a statistical approach. The new algorithm outperforms the old one in 81.6% of the sampled Pareto fronts from 20 simulations. When operated at a suitable sensitivity level such as 90.8% (79/87) or 88.5% (77/87), the FP rate is decreased by 24.4% or 45.8% respectively.

Keywords: CT colonography, Virtual colonoscopy, computer-aided diagnosis, polyp detection, Pareto front

1. INTRODUCTION

Colonic polyps are abnormal growths originated from cells of the colonic mucosa, the inner wall of the colon. The majority are elliptical protrusions of the mucosa which can be detected by curvature-based region growing algorithms[1]. Early reports on various CT colonography (CTC) computer-aided detection (CAD) systems suggest high sensitivity is possible[2]. However, evaluation and optimization of these region growing algorithms remain manual and empirical.

The performance of a polyp detection algorithm involves two conflicting objectives, minimizing both false negative (FN) and false positive (FP) detection rates. This problem does not produce a single optimal solution but a set of solutions known as a Pareto front[3-5]. Any solution in a Pareto front can only outperform other solutions in one of the two competing objectives. Using evolutionary algorithms to find the Pareto fronts for multi-objective optimization problems has been common practice for years. However, they are rarely investigated in any CTC CAD system because the computation cost is inherently expensive. The cost of evaluating a CTC CAD algorithm comes mainly from two sources. First, a typical algorithm has several adjustable operating parameters. The dimensions of the searching space for optimal operations are equal to the number of these parameters. Second, colonic polyps vary in shape and size. Consequently, an adequate evaluating data set has to be large enough to cover the broad spectrum of polyp variations. Combining high parametric dimensions with large evaluating data makes algorithm evaluation an extremely expensive computation task. To circumvent this problem, we have developed a parallel program that is implemented on a Linux

* adam_huang2003@yahoo.com, lij3@cc.nih.gov, rms@nih.gov, nicholas.petrick@fda.hhs.gov, hara.amy@mayo.edu

Medical Imaging 2007: Computer-Aided Diagnosis, edited by Maryellen L. Giger, Nico Karssemeijer,
Proc. of SPIE Vol. 6514, 651407, (2007) · 1605-7422/07/$18 · doi: 10.1117/12.709426

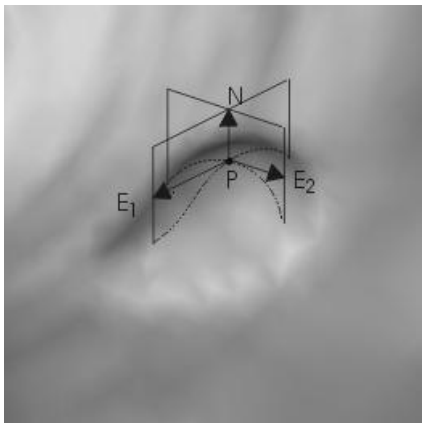

Figure 1. A 6mm adenomatous polyp appears like an elliptical protrusion. **N** is the surface normal and $\mathbf{E}_1$ and $\mathbf{E}_2$ are the principal tangent directions at point ***p***.

cluster environment to make the computation time acceptable. This parallel program is tested on an existing one-step, curvature-based region growing algorithm and helps to derive a new two-step, curvature-based region growing algorithm for colonic polyp detection. We demonstrate that Pareto fronts are effective visual tools to evaluate and compare algorithms for CTC CAD systems.

2. METHODS

2.1. Polyp Candidate Segmentation

Surface curvatures are local geometric properties which quantitatively describe how the surface curves or bends locally. Given air-distended colon surfaces, polyps appear like elliptical protrusions oriented inward toward the colon lumen (Fig 1). This elliptical feature can be characterized by two principal curvatures which are the maximum and minimum normal curvatures along the principal tangent directions[6]. Let k_1 and k_2 denote the maximum and minimum principal curvatures respectively and $\mathbf{E}_1$, $\mathbf{E}_2$ their corresponding principal directions which are perpendicular to each other. By definition, curves that bend toward and away from the normal direction **N** have positive and negative curvature values respectively. Consequently polyps can be identified as regions with negative k_1 and k_2[1].

Before we elaborate on polyp candidate filtering criteria, we give definitions of three additional curvature-related properties. Gaussian curvature K, mean curvature H, and sphericity index SI are derived from principal curvatures k_1 and k_2 and are defined as:

$$K = k_1 k_2 \tag{1}$$

$$H = (k_1 + k_2)/2 \tag{2}$$

$$SI = 2\left|\frac{k_2 - k_1}{k_2 + k_1}\right|, \text{ where } k_2 \le k_1 < 0. \tag{3}$$

K and H are simply product and mean of two principal curvatures. SI describes how round an elliptical surface is and ranges from 0 (sphere) to 2 (ridge). Any value in between represents an ellipsoid.

In this paper two curvature-based region growing algorithms for polyp detection are implemented and compared. The first algorithm, denoted as RA, clusters the vertices satisfying one set of curvature-based criteria on a triangular

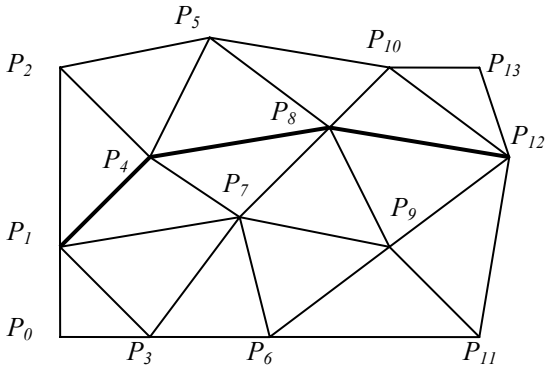

Figure 2. A graph path (thick line) from p_1 to p_{12} is noted as $P_{1,12} = \{p_1, p_4, p_8, p_{12}\}$.

mesh surface; the second algorithm, denoted as RA+RB, clusters the vertices based on two separate sets of curvature-based criteria. The algorithms are described in more detail in the following two sections.

2.1.1. One-step region growing algorithm RA

After the colonic surface is segmented and principal curvatures are estimated by using a kernel method[1], a one-step, curvature-based region growing algorithm, denoted as RA, is applied to screen through all surface vertices to identify polyp candidates. A region growing algorithm typically consists of three procedural components: 1) identifying seed points based on desired characteristics, 2) sequentially adding qualified immediate neighbor points to grow clusters, and 3) screening the clusters based on certain properties such as the member population and shape characteristics. However, in order to facilitate our discussion on the algorithm properties, we outline RA in the form of sets instead of procedural descriptions.

First, vertices satisfying 1) elliptical, 2) mean, and 3) Gaussian curvatures criteria are put in a set S_A:

$$\begin{aligned} S_A = \{p_i \mid\ & 1)\ k_1(p_i) < 0 \text{ and } k_2(p_i) < 0; \text{ and} \\ & 2)\ H_{\min} < H(p_i) < H_{\max}; \text{ and} \\ & 3)\ K_{\min} < K(p_i) < K_{\max}\} \end{aligned} \tag{4}$$

where $k_1(p_i)$, $k_2(p_i)$, $H(p_i)$, and $K(p_i)$ are the maximum, minimum, mean, and Gaussian curvatures at vertex p_i; $H_{\min}$, $H_{\max}$, $K_{\min}$, and $K_{\min}$ are adjustable operating parameters for selectivity.

Second, neighboring vertices in S_A are then clustered to form n disjoint regions RA_k, k = 1 to n, based on connectivity using regular region growing methods. Let $P_{i,j}$ represent a graph path between two vertices p_i and p_j ,

$$P_{i,j} = \{p_i, p_x, \ldots, p_j\} \tag{5}$$

where its elements are connected in the order of p_i, p_x, …, and p_j and $P_{i,i} = \{p_i\}$. Fig 2 illustrates a graph path on a triangular mesh. Regions RA_k are defined as:

$$\begin{aligned} RA_k = \{\, p_i \mid\ & \text{Let } p_s \text{ be a seed point, that is, } p_s \in S_A \text{ and } p_s \in RA_k; \\ & p_i \in RA_k \text{ if there exists a graph path } P_{i,s} \text{ such that } P_{i,s} \subset S_A\} \end{aligned} \tag{6}$$

In addition, RA_k satisfy the following two properties:

$$S_A = \bigcup_{k=1}^{n} RA_k , \tag{7}$$

$$RA_i \bigcap RA_j = \phi , \text{ if } i \neq j. \tag{8}$$

Third, the initial clusters RA_k are screened based on their vertex population and average SI to remove noise-induced small bumps and ridge-like folds respectively, that is,

$$\| RA_k \| \geq N_A \tag{9}$$

$$\text{Average } SI\,(RA_k) = 2\left| \frac{\frac{1}{\| RA_k \|} \sum_{p_i \in RA_k} k_2(p_i) - \frac{1}{\| RA_k \|} \sum_{p_i \in RA_k} k_1(p_i)}{\frac{1}{\| RA_k \|} \sum_{p_i \in RA_k} k_2(p_i) + \frac{1}{\| RA_k \|} \sum_{p_i \in RA_k} k_1(p_i)} \right| < SI_A. \tag{10}$$

$\| \, . \, \|$ represents the cardinality of a set, that is, the number of elements in a set. N_A and SI_A are two additional operating parameters.

In summary, surface point clusters, grown from vertices in set S_A, satisfying shape and size criteria SI_A and N_A are valid polyp candidates by algorithm RA. The resultant set is represented as

$$\boldsymbol{S_{RA}} = \{\, RA_k \mid \| RA_k \| \geq N_A \text{ and Average } SI\,(RA_k) < SI_A \,\}. \tag{11}$$

2.1.2 Two-step region growing algorithm RA+RB

We introduce another algorithm by adding a second region growing step, denoted as RB, to the first algorithm RA. The new two-step algorithm is denoted as RA+RB for brevity. Procedurally speaking, the first step of algorithm RA+RB is identical to algorithm RA while the second step, algorithm RB, uses valid polyp candidates in $\boldsymbol{S_{RA}}$ as bases to form new regions by a different set of curvature-based criteria. Algorithm RB first selects new seed points from a valid candidate RA_k and its immediate non-member neighboring vertices using a different set of curvature-based criteria. Second, vertices that are not members of existing, valid polyp candidates RA_i, $i \neq k$, and satisfy the new set of criteria are added to form regions RB_k. Third, newly found RB_k are screened based on their member populations. In order to facilitate our discussion on the properties of algorithm RB, we rephrase the procedures in more formal set descriptions.

First, any vertex satisfying a new, different set of curvature criteria is put in a new set S_B:

$$\begin{aligned} S_B = \{p_i \mid \quad & 1)\ k_1(p_i) < 0,\ k_2(p_i) < 0; \text{ and} \\ & 2)\ H(p_i) < H_B; \text{ and} \\ & 3)\ SI(p_i) < SI_B\} \end{aligned} \tag{12}$$

where $k_1(p_i)$, $k_2(p_i)$, $H(p_i)$, and $SI(p_i)$ are the maximum, minimum, mean curvatures, and sphericity index at vertex p_i; H_B and SI_B are adjustable operating parameters for selecting vertices which satisfy a different set of shape characteristics. For a valid polyp candidate RA_k in (11), a set of new seed points, SE_k, are extracted from members of RA_k and their non-member, immediate neighbors.

$$\begin{aligned} SE_k = \{p_i \mid \ & 1)\ p_i \in S_B \text{ and} \\ & 2)\ p_i \in RA_k \text{ or} \\ & \quad p_i \text{ is an immediate neighbor of the boundary vertices of } RA_k \text{ where } p_i \notin RA_j,\ RA_j \in \boldsymbol{S_{RA}} \,\}. \end{aligned} \tag{13}$$

Second, region RB_k is formed from SE_k by using a regular region growing algorithm. Let $P_{i,j}$, defined in (5), represent a graph path between two vertices p_i and p_j. Regions RB_k are defined as:

$$RB_k = \{\, p_i \mid \text{Let } p_s \text{ be any seed point, that is, } p_s \in SE_k\,, \; p_i \in RB_k \text{ if there exists a graph path } P_{i,s} \text{ such that } P_{i,s} \subset S_B \text{ and } P_{i,s} \cap (RA_j \cup RB_j) = \phi \text{ if } j \neq k\}. \tag{14}$$

Note that RB_k can consist of multiple disconnected vertex clusters since the curvature-based criteria in (4) and (12) are neither inclusive nor exclusive.

Third, the clusters RB_k are screened based on their vertex population, denoted as $\| RB_k \|$, that is,

$$\| RB_k \| \geq N_B\,. \tag{15}$$

In summary, surface point clusters in (11) form the initial regions of interest to apply the second region growing algorithm RB. The final set of valid polyp candidates formed by the 2-step algorithm RA+RB is represented as:

$$\begin{aligned} \boldsymbol{S}_{\boldsymbol{RA+RB}}(H_{\min}, H_{\max}, K_{\min}, K_{\max}, SI_{\mathrm{A}}, N_{\mathrm{A}}, H_{\mathrm{B}}, SI_{\mathrm{B}}, N_{\mathrm{B}}) \\ = \{\, RA_k \cup RB_k \mid \; & 1)\; RA_k \in \boldsymbol{S}_{\boldsymbol{RA}} \text{ and} \\ & 2)\; RB_k \text{ is the additional region initiated from } RA_k \text{ and} \\ & 3)\; \| RB_k \| \geq N_B \,\}. \end{aligned} \tag{16}$$

2.2 Pareto Fronts

The performance evaluation and optimization of a region growing algorithm in polyp detection involves two conflicting objectives, minimizing both false negative (FN) and false positive (FP) detection rates. True polyps that are not detected are FN's; detections localized on normal colon surfaces are FP's. This problem does not produce a single optimal solution but a set of possible solutions known as a Pareto optimal set, which results in a Pareto front in the objective space. A Pareto front is essentially an objective boundary such that any solution on the front can only be outperformed by another solution in at most one of the two competing objectives. Therefore, a Pareto optimal set is also called a Pareto non-dominated set.

Formally the optimization problem can be stated as minimizing a two-objective vector

$$\mathbf{F}(\mathbf{x}) = (\mathrm{FN}(\mathbf{x}), \mathrm{FP}(\mathbf{x})) \tag{17}$$

where $\mathbf{x}$ is the vector of algorithm operating parameters. A solution $\mathbf{x}_1$ is said to dominate $\mathbf{x}_2$ if and only if

$$\begin{aligned} &\mathrm{FN}(\mathbf{x}_1) \leq \mathrm{FN}(\mathbf{x}_2) \text{ and} \\ &\mathrm{FP}(\mathbf{x}_1) \leq \mathrm{FP}(\mathbf{x}_2) \text{ and} \\ &\{\mathrm{FN}(\mathbf{x}_1) < \mathrm{FN}(\mathbf{x}_2) \text{ or } \mathrm{FP}(\mathbf{x}_1) < \mathrm{FP}(\mathbf{x}_2)\} \end{aligned} \tag{18}$$

where $\mathbf{x}_2 \neq \mathbf{x}_1$.

The vector of operating parameters for algorithms RA and RA+RB can be explicitly expressed as

$$\mathbf{x}_{\mathrm{A}} = (H_{\min}, H_{\max}, K_{\min}, K_{\max}, SI_{\mathrm{A}}, N_{\mathrm{A}}) \tag{19}$$

and

$$\mathbf{x}_{\mathrm{A+B}} = (H_{\min}, H_{\max}, K_{\min}, K_{\max}, SI_{\mathrm{A}}, N_{\mathrm{A}}, H_{\mathrm{B}}, SI_{\mathrm{B}}, N_{\mathrm{B}}). \tag{20}$$

Let the set of all true polyps be represented as

$$\boldsymbol{S}_{\boldsymbol{polyp}} = \{\, S_j \mid S_j = \{\, p_{\mathrm{i}} \mid p_{\mathrm{i}} \text{ is a vertex marked as on polyp identify number } j\}\}, \tag{21}$$

the FN and FP can be expressed as

$$\mathrm{FN}(\mathbf{x}) = \|\{\, S_j \mid \ 1)\ S_j \in \boldsymbol{S}_{polyp} \text{ and } 2)\ S_j \cap R_k = \phi,\ \forall\, R_k \in \boldsymbol{S}\,\}\| \tag{22}$$

and

$$\mathrm{FP}(\mathbf{x}) = \|\{\, R_k \mid \ 1)\ R_k \in \boldsymbol{S} \text{ and } 2)\ S_j \cap R_k = \phi,\ \forall\, S_j \in \boldsymbol{S}_{polyp}\}\| \tag{23}$$

where $\mathbf{x}$ is $\mathbf{x}_{\mathrm{A}}$ or $\mathbf{x}_{\mathrm{A+B}}$, $\boldsymbol{S}$ is $\boldsymbol{S}_{RA}$ or $\boldsymbol{S}_{RA+RB}$ for algorithms RA and RA+RB respectively. $\|\,.\,\|$ represents the cardinality of a set, that is, the number of elements in a set.

As FN($\mathbf{x}$) and FP($\mathbf{x}$) are two conflicting objective functions with no closed form solutions, therefore, it is natural to look at this parameter setting problem as multiobjective evolutionary problem.

2.3. Multiobjective Evolutionary Algorithms

Since the first studies on evolutionary multiobjective optimization in the mid-1980s, many algorithms have been proposed in the literature[3, 5, 7] for finding or approximating the Pareto optimal set for multiobjective optimization problems. In a standard genetic algorithm, there are usually four steps in the evolutionary procedure: 1) randomly initialize the solution population, 2) evaluate and assign a fitness value for each individual in the population according to its performance, 3) select individuals based on their fitness values such that better individuals are more likely to be selected for producing the next generation, and 4) use crossover and mutation to produce next generation from the selected individuals. In this study, we use the SPEA2[5] algorithm to find the Pareto set for its fast convergence rate. The SPEA2 algorithm differs from the standard genetic algorithm in the following three aspects.

Environmental Selection: Aside from the regular population in a genetic algorithm, there is an archive containing all the nondominated solutions from the previous generation. An individual in the archive is removed only if 1) a solution has been found in the current generation that dominates it or 2) the maximum archive size is exceeded and its "fitness" is worse than that of other solutions in the archive.

Fitness Evaluation: The fitness evaluation for an individual is based on both the population and the archive. A good individual is assigned a smaller fitness value. Let $\mathbf{P}_t$ and $\overline{\mathbf{P}}_t$ denote the population and archive respectively, each individual operating parameter set $\mathbf{x}_i$ in $\mathbf{P}_t$ and $\overline{\mathbf{P}}_t$ is assigned a strength value $V(\mathbf{x}_i)$, the number of solutions it dominates,

$$V(\mathbf{x}_i) = \|\{\, \mathbf{x}_j \mid \mathbf{x}_j \in \mathbf{P}_t \cup \overline{\mathbf{P}}_t \text{ and } \mathbf{x}_i \succ \mathbf{x}_j\}\| \tag{24}$$

where $\|\,.\,\|$ represents the cardinality of a set and $\mathbf{x}_i \succ \mathbf{x}_j$ corresponds to the Pareto dominance relation, in which $\mathbf{x}_i$ dominates $\mathbf{x}_j$. Based on the value of $V(\mathbf{x}_i)$, a raw fitness value $R(\mathbf{x}_i)$ is given to the individual $\mathbf{x}_i$,

$$R(\mathbf{x}_i) = \sum_{\mathbf{x}_j \in \mathbf{P}_t \cup \overline{\mathbf{P}}_t,\, \mathbf{x}_j \succ \mathbf{x}_i} V(\mathbf{x}_j)\,. \tag{25}$$

The raw fitness assignment by (25) will fail if most individuals do not dominate each other. Therefore, additional density information is incorporated to discriminate individuals with identical raw fitness values. The density estimation technique used in SPEA2 in an adaptation of the k-th nearest neighbor method[8]. Let σ_i^k denote the distance (in the objective space) from the k-th nearest neighbor to an individual $\mathbf{x}_i$; the density estimation corresponding to $\mathbf{x}_i$ is defined by

$$D(\mathbf{x}_i) = \frac{1}{\sigma_i^k + 2}\,. \tag{26}$$

As a common setting, k is set as $\sqrt{N+\overline{N}}$, where N and $\overline{N}$ are the sizes of the population $\mathbf{P}_t$ and archive $\overline{\mathbf{P}}_t$ respectively. Last, adding $D(\mathbf{x}_i)$ to the raw fitness value $R(\mathbf{x}_i)$ of an individual $\mathbf{x}_i$ yields the final fitness value $F(\mathbf{x}_i)$:

$$F(\mathbf{x}_i) = R(\mathbf{x}_i) + D(\mathbf{x}_i). \tag{27}$$

Readers can refer to reference 5 for more details of the SPEA2 algorithm and references 4 and 9 for examples of implementations in colonic polyp detection.

2.4. Algorithm Implementation and Evaluation

The SPEA2 algorithm is implemented in Matlab; region growing algorithms RA and RA+RB are implemented in C++ under Visual Studio for Windows. To speed up computation, the region growing programs are recompiled in a Linux environment and multiple copies of the programs are run parallel in a Linux Beowulf computing cluster[10]. In the Linux cluster, the SPEA2 program runs as the main process, dispatching jobs and collecting results generation after generation until the Pareto front converges or the generation number reach a preset upper limit. The parallel computation is outlined as followed.

1. *Initialization.* The main program randomly generates a population $\mathbf{P}_{t=1}$ of $N = 100$ operating vectors $\mathbf{x}$ as the first generation and an empty archive $\overline{\mathbf{P}}_{t=1}$, $\overline{N} = 0$.
2. *Region Growing*. Apply $\mathbf{P}_t$ to a given evaluating data set of 56 CTC colon surfaces with 87 proven polyp detections sized 4 to 60mm. The data set is divided into n subsets and the region growing program is applied to each subset simultaneously on n machines. (The number of available machines is administrated by the system management depending on the system workload.)
3. *Environmental Selection*. The main program waits until all the data is processed for the current generation t and the resultant FN and FP rates are available. The SPEA2 algorithm updates the archive $\overline{\mathbf{P}}_{t=t+1}$ from $\mathbf{P}_t$ and $\overline{\mathbf{P}}_t$ based on Pareto Non-dominance. If the size $\overline{N}$ of the updated archive is large than 100, truncate the archive based on "fitness", defined by (27), until $\overline{N} = 100$.
4. *Mating Selection*. Select 100 individuals in the current archive with replacement using a binary tournament procedure.
5. *Reproduce*. Reproduce the next generation ($N = 100$) using standard crossover and mutation procedures. The crossover and mutation probability are set as 0.9 and 0.01 respectively in our experiments. Terminate if the generation number reaches the upper limit, otherwise go to step 2.

The upper generation limits of RA and RA+RB are 200 and 300 which are proportional to the number of operating parameters used as in $\mathbf{x}_A$ and $\mathbf{x}_{A+B}$ respectively.

We used a statistical comparison program developed by Knowles and Corne[3] to evaluate the performance difference between algorithms RA and RA+RB in the objective space (FN, FP). The basic idea of this metric is as follows: Suppose that two algorithms result in two non-dominated sets P1 and P2 respectively. The lines that join the solutions in P1 and P2 are called attainment surfaces (which are curves in two-dimensional cases). A number of lines are then drawn from the origin such that they intersect with the surfaces. The comparison is then individually done for each sampling line to determine which one outperforms the other. The method is illustrated in Fig 3. In this study, each algorithm was run 20 times and 500 intersection comparisons were sampled on each possible pair of P1 and P2 surfaces. Comparing classes of algorithms in CAD is essential for efficient algorithm development. Multiobjective evolutionary methods provide a tool for comparing classes of algorithms. However, these methods do have potential limitations in the selection of an appropriate performance metric and in the blurring between training and test data. While we are aware of these potential limitations, they are outside the scope of this paper.

3. RESULTS

In the 20 runs, the new algorithm RA+RB outperforms RA in 81.6% of sampled space; RA outperformed RA+RB in 0% of the same space and in 18.4% the two methods were equivalent. The average computation time is about 4 and 6 hours for each algorithm, respectively. This is a significant improvement compared to the time of more than 30 hours needed to run on a single machine.

The average results of the 20 runs of both methods are shown in Fig 4. By our problem definition, a better algorithm should have a lower Pareto front. The mean Pareto front of the new algorithm RA+RB is clearly on the lower left side of the existing algorithm RA. The average FP reduction in percentage by RA+RB at every sensitivity level is illustrated in Fig 5. It shows that RA+RB outperforms RA at all sensitivity levels except 18.4% (FN=71) where the FP rate increases by 4%. When operated at a suitable sensitivity level such as 90.8% (FN=8) or 88.5% (FN=10), the FP rate can be decreased by 24.4% or 45.8% respectively. Operating parameter vectors $\mathbf{x}_A$ and $\mathbf{x}_{A+B}$ at FN=10 are:

$$\mathbf{x}_A = (-10, -1.05, 0.59, 10.94, 0.99, 15) \tag{28}$$

and

$$\mathbf{x}_{A+B} = (-10, -0.16, -50, 6.25, 1.21, 19, -1.05, 0.67, 18). \tag{29}$$

The new method can distinguish a normal bump (Fig 6a) from a 6mm adenomatous polyp (Fig 6b) because the bump has fewer RB_k members and is now found to be less round than actual polyp candidates. Fig 7 illustrates another observation to help explain why the new algorithm performs better. As algorithm RA uses only one set of curvature criteria, the level of selectivity has to be compromising. This results in multiple, small detections (Fig 7a) for a large, irregular shaped polyp. The new algorithm consists of two sets of growing criteria such that one less selective set allows the detection to grow a larger cluster and one more selective set to guarantee that the detection has enough surface vertices which resemble the shape of an ellipsoid (Fig 7b).

4. CONCLUSION

We have developed a new systematic scheme to design new algorithms for the optimization of polyp candidate segmentation for CTC CAD systems using curvature-based criteria. We investigated the feasibility of applied multiobjective evolutionary algorithms to find the optimal operating parameters for an existing and a new two-stage region growing algorithm. Evolutionary algorithms are inherently expensive and are rarely introduced into the CTC CAD research field. To circumvent the problem, we have implemented the multiobjective evolutionary algorithm on a PC/Linux cluster environment. The average time to run a training session for algorithms RA and RA+RB is roughly 4 and 6 hours respectively.

In conclusion, the solutions in a Pareto optimal set can give more information and confidence to evaluate the performance differences between new and old algorithms. The results showed that the new algorithm can reduce the FP rate by 24.4 to 45.8% while operating at a high sensitivity level between 90.8 to 88.5%.

ACKNOWLEDGEMENT

This research was supported by the intramural programs of the NIH, the Diagnostic Radiologist Department, Clinical Center. This study utilized the high-performance computational capabilities of the NIH Biowulf PC/Linux cluster.

REFERENCES

1. Huang, A., Summers, R.M., and Hara, A.K.. *Surface curvature estimation for automatic colonic polyp detection*. Proc. Medical Imaging 2005, SPIE vol. 5746, p. 393-402.
2. Summers, R.M., Yao, J., Pickhardt, P.J., Franaszek, M., Bitter, Brickman, D., Krishna, V., Choi, J.R. *Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population.* Gastroenterology, 2005. **129**(6): p. 1832-44.
3. Knowles, J. and Corne, D., *Approximating the nondominated front using the pareto archived evolutionary strategy*. Evolutionary Computation, 2000. **8**(2): p. 149-172.
4. Li, J., Huang, A., Yao, J., Bitter, I., Petrick, N., Summers, R.M., Pickhardt, P.J., Choi, J.R. *Automatic colonic polyp detection using multiobjective evolutionary techniques*. Proc. Medical Imaging 2006, SPIE vol. 6144.
5. Zitzler, E., M. Laumanns, and L. Thiele, *Spea2: Improving the strength pareto evolutionary algorithm*. Technical Report, 2001, Swiss Federal Institute of Technology. Available at: http://e-collection.ethbib.ethz.ch/ecol-pool/incoll/incoll_324.pdf. Accessed June 12, 2006.
6. do Carmo, M.P., *Differential geometry of curves and surfaces*. 1976, Englewood Cliffs, NJ: Prentice Hall.
7. Schaffer, J.D. *Multiple objective optimization with vector evaluated genetic algorithms*. Proc. The First International Conference on Genetic Algorithms. 1985.
8. Silverman, B.W., *Density estimation for statistics and data analysis*. 1986, London: Chapman and Hall.
9. Li, J., Huang, A., Petrick, N., Yao, J., Summers, R.M., *Validating pareto optimal operation parameters of polyp detection algorithms for CT colonography*. To appear in Proc. Medical Imaging 2007, SPIE vol. 6514.
10. Bitter, I., Brown, J.E., Brickman, D., Summers, R.M., *Large-scale validation of a computer-aided polyp detection algorithm for CT colonography using cluster computing*. Proc. Medical Imaging 2004, SPIE vol. 5369, p. 290-294.

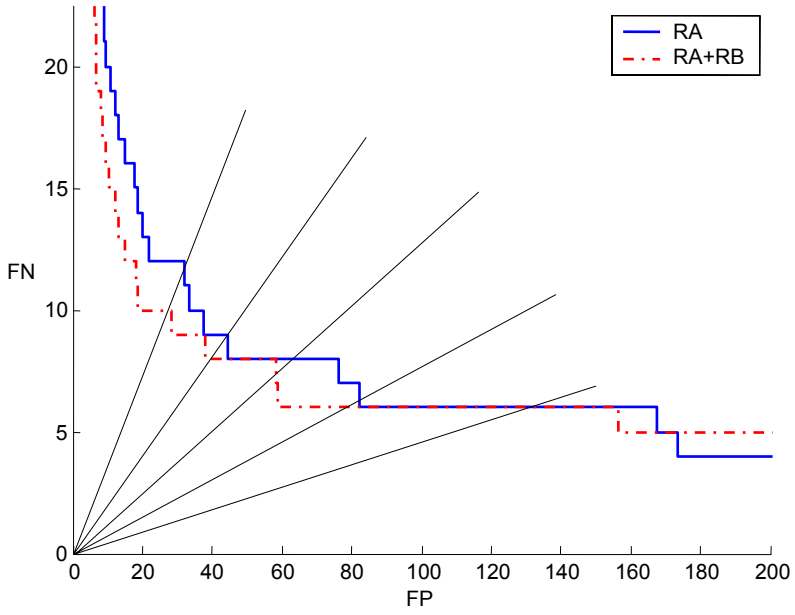

Figure 3. Compare RA and RA+RB performances by sampling a pair of non-dominated sets (Pareto fronts) using lines of intersections. Equally spaced radial lines originated at the origin are sent out to intersect both curves. A sampling line has three possible results: intersecting RA curve first (RA outperforms RA+RB), intersecting RA+RB curve first (RA+RB outperforms RA), intersecting both curves simultaneously (equivalent).

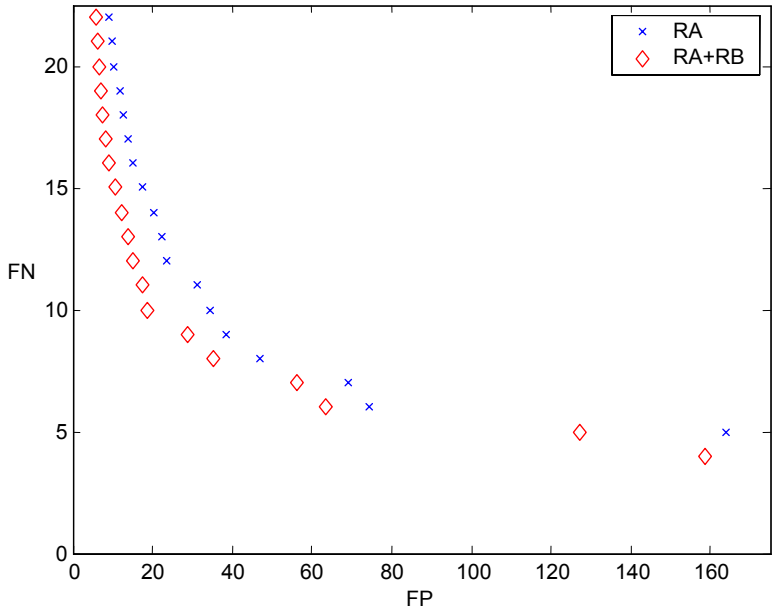

Figure 4. The mean Pareto fronts of algorithms RA and RA+RB from 20 runs. The RA+RB curve shows superior performance (fewer FN and FP).

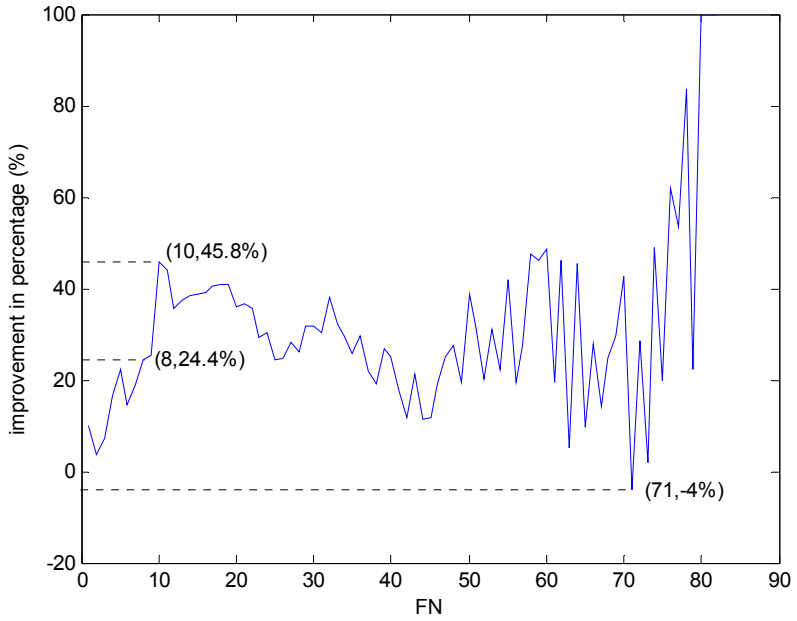

Figure 5. The improvement in reducing FP rates by using algorithm RA+RB when operated at the same FN rates using RA.

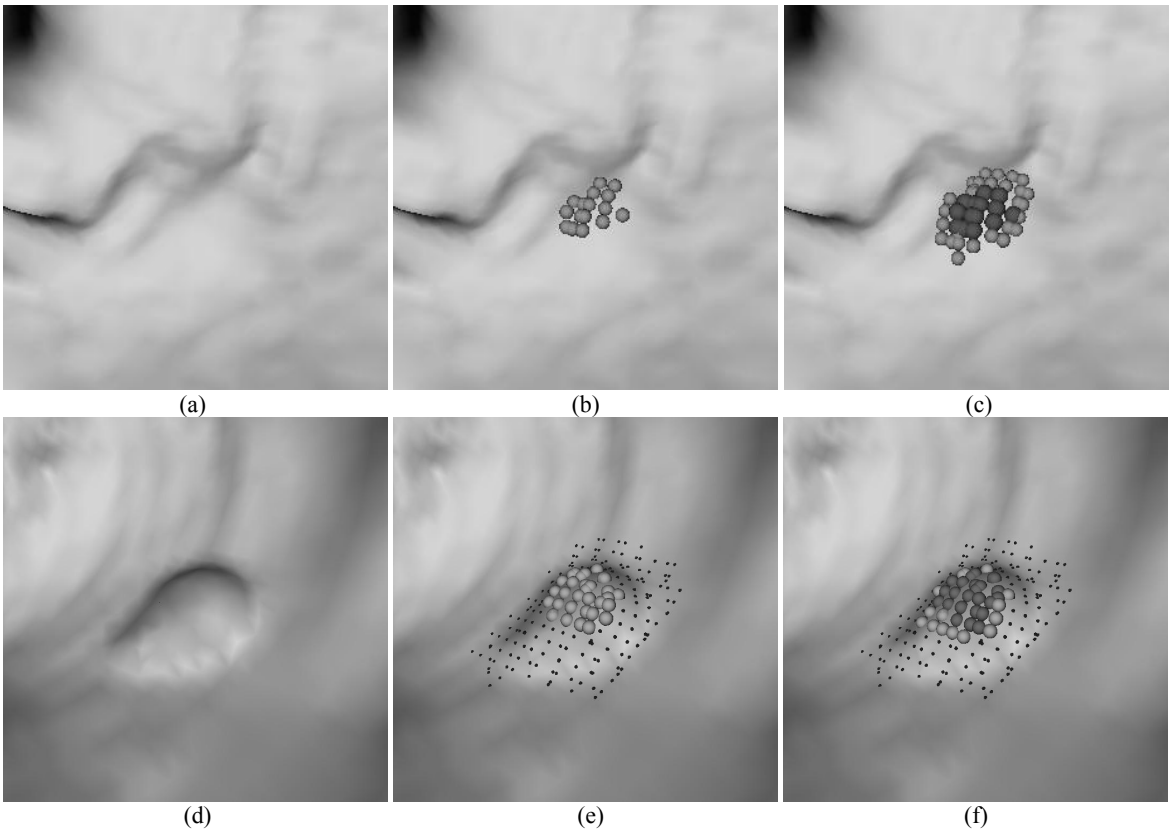

Figure 6. (a) A polyp-like bump. (d) A 6mm adenomatous polyp. (b, e) Resultant clusters of vertices using algorithm RA. (c, f) Resultant clusters using algorithm RA+RB. The clusters in (b, e) are valid candidates by a criterion set $\mathbf{x}_A$, The cluster in (f) is a valid candidate by a different criterion set $(\mathbf{x}_A)'+ \mathbf{x}_B$ but (c) is not because of a small population, $||RB_k||$. In (c, f) light gray spheres represent vertices that satisfy a less selective criterion set $(\mathbf{x}_A)'$; dark gray spheres satisfy both selective criterion sets $(\mathbf{x}_A)'$ and $\mathbf{x}_B$. In (e, f) small dots are manually marked vertices of the polyp.

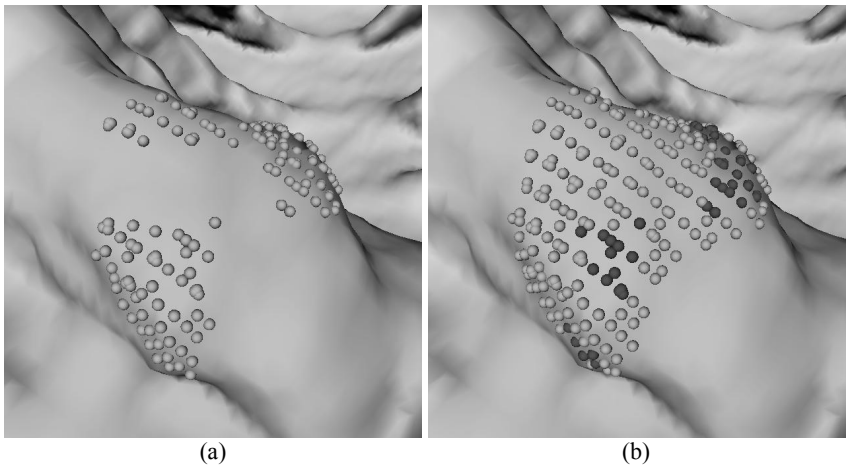

Figure 7. A partial view of a 4 cm irregular shaped polyp. (a) Multiple detections (clusters of light gray spheres) using RA. (b) A large, single detection is formed by using RA+RB. In (b) light gray spheres satisfy a less selective criterion set $\mathbf{x}_A$; dark gray spheres satisfy both selective criterion set $\mathbf{x}_A$ and $\mathbf{x}_B$.