

# Non-invasive differential diagnosis of dental periapical lesions in cone-beam CT scans

Kazunori Okada\*

*Department of Computer Science, San Francisco State University, San Francisco, CA*

Steven Rysavy

*Biomedical and Health Informatics Program,  
University of Washington, Seattle, WA*

Arturo Flores

*Computer Science and Engineering,  
University of California, San Diego, San Diego, CA*

Marius George Linguraru

*Sheikh Zayed Institute for Pediatric Surgical Innovation,  
Children's National Medical Center, Washington, DC and  
Departments of Radiology and Pediatrics,  
George Washington University, Washington, DC*

(Dated: January 31, 2015)

# Abstract

**Purpose:** This paper proposes a novel application of computer-aided diagnosis (CAD) to an every-day clinical dental challenge: the non-invasive differential diagnosis of periapical lesions between periapical cysts or granulomas. A histological biopsy is the most reliable method currently available for this differential diagnosis, however this invasive procedure prevents the lesions from healing non-invasively despite a report that they may heal without surgical treatment. A CAD using cone-beam computed tomography (CBCT) offers an alternative non-invasive diagnostic tool which helps to avoid potentially unnecessary surgery and to investigate the unknown healing process and rate for the lesions.

**Methods:** The proposed semi-automatic solution combines graph-based random walks segmentation with machine learning-based boosted classifiers and offers a robust clinical tool with minimal user interaction. As part of this CAD framework, we provide two novel technical contributions: 1) probabilistic extension of the random walks segmentation with likelihood ratio test and 2) LDA-AdaBoost: a new integration of weighted linear discriminant analysis to AdaBoost.

**Results:** A dataset of 28 CBCT scans is used to validate the approach and compare it with other popular segmentation and classification methods. The results show the effectiveness of the proposed method with 94.1% correct classification rate and an improvement of the performance by comparison with the Simon's state-of-the-art method by 17.6%. We also compare classification performances with two independent ground-truth sets from the histopathology and CBCT diagnoses provided by endodontic experts.

**Conclusion:** Our experimental results show that the proposed CAD system behaves in clearer agreement with the CBCT ground-truth than with histopathology, supporting the Simon's conjecture that CBCT diagnosis can be as accurate as histopathology for differentiating the periapical lesions.

Keywords: Computer-aided diagnosis, Dental imaging, Periapical lesion, Non-invasive diagnosis, Cone-beam computed tomography

## I. INTRODUCTION AND PURPOSE

Periapical lesions<sup>1</sup> are some of the important and common oral pathologies in endodontics - a field of dentistry that deals with tooth's pulp, root and its surrounding tissues by employing treatment techniques such as dental root canal therapy and maxillofacial surgery<sup>2</sup>.

5 Periapical lesions come in two main types: periapical cysts or granulomas<sup>1</sup>. Periapical cysts are inflammatory debris-filled closed cavities at the apices of teeth, lined by epithelia that contain liquid or semisolid. On the other hand, periapical granulomas are masses of chronically inflamed granulation tissue resulting from irritation following pulp disease or endodontic treatment. The incidence of these two types is comparable but varies largely in  
10 reports, ranging between 6 and 55% and between 9.3 and 87.1% for cysts and granuloma, respectively<sup>3</sup>.

Diagnostic differentiation of periapical lesions into these two types plays a crucial role in endodontic practices. Standard treatment for periapical lesions consists of the elimination of the infectious agents by root canal treatment, allowing healing of the lesion<sup>2</sup>. When  
15 treatment fails with a persisting lesion, even when asymptomatic, the endodontist should consider either retreatment of the canal, periapical surgery or extraction of the affected tooth<sup>2</sup>. A histological biopsy study of the periapical lesion can be used to confirm the diagnosis of periapical cyst and distinguish it from granuloma: a non-cystic inflammatory lesion<sup>1</sup>. This histological diagnosis is the most reliable method of differential diagnosis  
20 currently available<sup>3</sup>. Since this procedure is invasive, the chance for the granuloma to heal non-invasively is lost despite a report that a granuloma may heal without surgical treatment if given the opportunity<sup>4</sup>. As a result the patient is subjected to potentially unnecessary surgery and associated complications, including infection and discomfort. Furthermore, the healing process and rate for granulomas remain largely unknown due to the lack of non-  
25 invasive diagnostic tools. It is known that the chronic periapical granulomas can precede the cysts<sup>1</sup>. Although the statistical probability of cyst occurrence may be higher among larger lesions<sup>5</sup>, a definite relationship between lesion size and cystic condition has not been supported by histology.

Computed tomography (CT) has been suggested as a non-invasive method to make a dif-  
30 ferential diagnosis between these lesion types<sup>3,6,7</sup>, while other common diagnostic methods in dentistry, such as conventional radiographs<sup>8</sup>, periapical radiographs<sup>9</sup> and Papanicolaou

smears<sup>10</sup> have been shown ineffective for the differential diagnosis. Routine use of CT is, however, associated with high radiation risks<sup>11</sup>. Cone-beam computed tomography (CBCT)<sup>12</sup> is a recent 3D medical imaging technology that offers CT scans with much lower radiation dosage than the conventional CT scanners<sup>13</sup>. In their pioneering work, Simon et al.<sup>14</sup> demonstrated high correlation between CBCT-based predictions and histological ground-truths, suggesting the CBCT’s potential to be an effective and safe non-invasive differential diagnostic tool for periapical lesions. However, the proposed diagnostic procedure is technically unreliable and unrepeatable due to its simplistic and heavily interactive design. The method requires an endodontic expert to manually search the minimum intensity voxel corresponding to the cystic cavity among the whole lesion<sup>14</sup>, which is time consuming and prone to human error. Furthermore, the simple thresholding technique used in their method may be unreliable because the CBCT image’s grayscale values can be spatially inhomogeneous and do not correspond 1:1 to Hounsfield units<sup>15</sup>.

## A. Proposed Approach

Addressing these shortcomings, this paper proposes a framework for computer-aided diagnosis (CAD) of the periapical lesions, which improves the accuracy and repeatability of the above-described state-of-the-art method by Simon et al.<sup>14</sup>. The CAD framework provides a semi-automatic procedure that takes as inputs a CBCT scan and three user-specified click points, indicating a volume of interest (VOI), and outputs the differential diagnosis between cyst and granuloma. This design greatly improves the usability of the Simon’s method by reducing the amount of user interaction required per lesion. The proposed framework consists of three successive steps 1) graph-theoretic 3D lesion segmentation, 2) statistical intensity feature extraction, and 3) machine learning-based lesion classification. The segmentation of the CBCT periapical lesions is necessary for the diagnosis and also technically challenging because the low-dose CBCT images tend to be noisy and the interface between the lesion and the other soft tissues are often extremely vague. Our segmentation method follows the popular graph theoretic approach that has recently been applied to various applications successfully.<sup>16</sup> Then the automatic extraction of statistical intensity features from the segmented area helps to increase the diagnostic accuracy by incorporating more information than what is used in the Simon’s method<sup>14</sup> and by avoiding mis-detection of the minimum

intensity voxel due to human error. The machine-learning approach adapted in the lesion classification step plays another crucial role in improving the overall accuracy. Our approach learns a generic pattern classifier from examples to infer more reliable and flexible decision rules. For this classification step, we introduce the LDA-AdaBoost method that combines the AdaBoost classifier<sup>17</sup> with linear discriminant analysis (LDA)<sup>18</sup> used as weak learner. The proposed system is evaluated on the clinical CBCT dataset that includes the cases used in Simon et al.<sup>14</sup>, allowing the direct comparison of results.

## B. Contributions

To the best of our knowledge, this is the first semi-automatic CAD approach applied to the differential diagnosis of periapical lesions in CBCT data. Similar CAD framework has been widely and successfully applied in many clinical areas such as lung disease<sup>19</sup>, breast cancer<sup>20</sup>, colon polyp<sup>21</sup>, liver lesion<sup>22</sup> to name a few, providing physicians with a valuable tool. However, the field of dentistry has not yet fully benefited from the advancements of medical image analysis, especially with the new 3D imaging technologies, despite the commonality of dental procedures. The probabilistic extension of the random walks algorithm<sup>23</sup> with likelihood ratio test is another novel technical contribution of this work. We adapt a probabilistic data-driven decision rule with the likelihood ratio test<sup>24</sup> to increase the flexibility of the original random walks method, which was based on hard decision, and propose two ways to model the likelihood function that maximizes the lesion classification performance. In general, segmentation is an ill-posed problem so that an optimal algorithm should be designed under some domain-specific constraints. In our case, we design our solution to maximize the final diagnostic accuracy of the overall CAD system, creating a feedback loop between segmentation and classification. We also propose an LDA-AdaBoost algorithm for the classification step as an additional contribution in our paper. We first introduce sample weights in the standard linear discriminant analysis (LDA)<sup>18</sup> then adapt this weighted LDA as AdaBoost’s weak learner. Although AdaBoost<sup>17</sup> has been combined with various algorithms as weak learners, the combination with LDA<sup>18</sup> has scarcely been studied in literature. Exploiting LDA as AdaBoost’s weak learner helps to reduce the number of learners so that the final strong classifier has a better generalization performance than the standard AdaBoost. The proposed LDA-AdaBoost algorithm is applied to our dental CAD problem,

resulting in an accuracy of 94.1% or 17.6% improvement over Simon et al.<sup>14</sup>. Finally, we analyze the proposed CAD system with two independent ground-truth sets from biopsy and CBCT diagnoses. The results show that our system improves accuracy for both cases and  
95 behaves more in agreement with the CBCT diagnosis. This result supports a conjecture proposed in Simon et al.<sup>14</sup> that CBCT diagnosis can be as accurate as histopathology for differentiating the periapical lesions.

### C. Related Work

Early adaptations of image analysis to dentistry have mainly focused on 3D geometric  
100 model reconstruction of teeth and jaws<sup>25–28</sup> following the CAD/CAM approach<sup>29</sup> toward the applications of preoperative planning<sup>30</sup> and image-guided surgery<sup>31</sup>. More relevant to our study are early CAD systems on 2D radiographs for dental caries<sup>32,33</sup> and bone mineral density to detect osteoporosis<sup>34,35</sup>. On periapical lesions, Mol and van der Stelt reported one of the early studies on periapical bone defects, detecting the lesions semi-automatically and  
105 characterizing their size via texture analysis<sup>36–38</sup>. More recently, Li et al.<sup>39</sup> and Lee et al.<sup>40</sup> introduced semi-automatic CAD methods for segmentation of periapical/bifurcation lesions by level-set method and for detection of bone remodeling at treated sites by logistic regression, respectively. However, these studies neither address the differential diagnosis between granulomas and cysts nor deal with 3D CT data. Our target problem, differential diagnosis  
110 of periapical lesions in 3D CBCT scans, has attracted increasing interest in the endodontic community<sup>41</sup>, since the report by Cotti et al. demonstrated the potential of CT technology for this task<sup>42</sup>. The effectiveness of CBCT diagnosis for differentiating granulomas and cysts was first proposed by Simon et al.<sup>14,43</sup>. Aggarwal and Singla<sup>44</sup> also reported the differential diagnosis of periapical lesions using CT and showed its effectiveness toward nonsurgical le-  
115 sion management. These clinical studies, however, only addressed manual procedures with expert’s visual inspection, which may not be reproducible and are labor-intensive. Moreover, these studies did not aim to automate the procedures or to improve their accuracy by incorporating advanced image processing and machine learning methodologies. As a result, surgical biopsy and histopathological evaluation remain the standard procedures of  
120 the analysis of periapical lesions<sup>45</sup>, despite the advent of CBCT as a common option for diagnosis<sup>46,47</sup>, preoperative surgery/implant planning and postoperative assessment<sup>48–50</sup>.

There are a number of previous studies that are related to the technical methodologies proposed in this paper. Ding et al.<sup>51</sup> proposed a technique related to our likelihood ratio test-based extension of the random walks segmentation in their recursive segmentation and classification method. However, their work cannot be directly applied to our framework, since their classification step is designed to support segmentation, instead of the diagnostic classification task targeted in our CAD design. This design allows to tune our segmentation to improve overall performance of our CAD system.

Several previous studies are also related to the LDA-AdaBoost work. An outlier-class resistant approach to estimate the within-class covariance matrix in LDA for linear dimensionality reduction was presented by Tang et al.<sup>52</sup>. Although this work explored sample weights in their LDA formulation, they did not investigate its application to the boosting context of our interest. The adaptation of other classification models to AdaBoost has recently been explored<sup>53,54</sup>, however there are only very few reports that investigated LDA with AdaBoost despite the LDA's popularity<sup>55,56</sup>. Skurichina and Duin<sup>56</sup> also investigated boosting LDA classifiers, however their scheme did not aim to combine the two algorithms by introducing weights in LDA formulation. An interesting approach was proposed by Liu et al.<sup>55</sup> who adapted LDA to incorporate sample weights in an application related to indoor/outdoor digital image classification. Their approach, combining a weighted LDA and AdaBoost, differs from the proposed method in that 1) the Liu's weighted LDA<sup>55</sup> does not reduce to a classic LDA in the case of uniform weight distributions and 2) only pairs of features were used to train a LDA weak learner. The proposed method offers a form of the weighted LDA which reduces to the classic form when used with uniform weights and also includes feature triples to model more complex decision boundaries efficiently.

This paper extends our pilot studies<sup>57-60</sup> by updating the list of literature we surveyed, employing a larger size of dataset in our experiments, and expanding our discussion on the results.

## II. MATERIAL AND METHODS

### A. Lesion Segmentation

150 As the first step of our CAD framework, the segmentation of periapical lesions is formulated as a binary segmentation problem. We adapt a graph representation of 3D discrete data. We represent each 3D CT scan with  $N$  voxels as an undirected connected weighted graph  $G = (V, E, W)$  where  $V$  is a set of vertices/nodes  $v \in V, |V| = N$ ,  $E$  is a set of edges  $e \in E \subseteq V \times V$ , and  $W$  is a set of weights  $w(e) \forall e \in E$ . Each voxel  $i$  in the 3D image forms  
155 a node  $v_i$ . An edge  $e_{ij}$  between nodes  $v_i$  and  $v_j$  is included in  $E$  when the two voxels  $i$  and  $j$  are neighbors.  $w_{ij}$  denotes the weight on an edge  $e_{ij}$ , representing the contrast between intensities of the two nodes  $i$  and  $j$ .

A binary segmentation algorithm partitions the graph  $G$  into two disjoint subgraphs  $A$  and  $B$  by cutting edges connecting the two non-overlapping sets. As a result, all nodes  
160 within  $A$  and  $B$  are assigned with binary values 1 and 0 to annotate the target lesion (1) and the other (0) regions, respectively. Our semi-automatic segmentation algorithm allows the users to select the target lesion and have some control of the segmentation quality with minimum user-interaction, as explained below.

#### 1. Random Walks Segmentation: Review

165 The semi-automatic multi-label random walks algorithm proposed by Grady<sup>23</sup> is adapted and briefly reviewed bellow. The random walks algorithm requires a number of seed points to be specified by users for both foreground and background regions (or for each region corresponding to multi-labels) in an input image. The algorithm then efficiently computes the probability that a random walker starting from each unseeded pixel reaches one of the  
170 fore/back-ground seeds first. Then the final segmentation is given by assigning to each unseeded pixel the label with maximal probability. Grady showed that this random walker probability calculation is equivalent to the solution of a combinatorial Dirichlet problem of a sparse, symmetric, positive-definite system of equations

$$L_U x^s = -B^T m^s, \tag{1}$$



where, without the loss of generality, nodes in  $V$  are ordered as unseeded  $V_U$  and seeded  $V_M^s$  nodes with labels  $s = (1, \dots, K)$  in successive blocks. Then, for each label  $s$ ,  $x^s$  denotes the unknown of this system: a vector of probabilities  $x_i^s$  that a random walker starting from an unseeded  $v_i \in V_U$  first reaches one of the seeds with label  $s$ .  $m^s$  denotes the probability vector for the seeded nodes. In our binary segmentation,  $m^s$  is a vector of ones for foreground and zeros for background seeds with length  $|V_M|$ .  $L$  denotes the combinatorial Laplacian matrix defined as

$$L_{ij} = \begin{cases} d_i & \text{if } i = j \\ -w_{ij} & \text{if } v_i \text{ and } v_j \text{ are neighbors} \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

where  $d_i$  denotes the degree of  $v_i$  defined by  $d_i = \sum w_{ij}$  for all edges  $e_{ij}$  incident on  $v_i$ . The weights  $w_{ij}$  are computed from image by  $w_{ij} = e^{-\beta(g_i - g_j)^2}$ , where  $g_i$  represents the intensity of pixel  $i$  and  $\beta$  is the only free parameter in the algorithm.  $L_U$  and  $B$  in (1) represent submatrices/blocks of  $L$  due to the induced ordering of  $x$  according to the seeding labels.  $L_U$  corresponds to a block among all unseeded nodes  $v_i, v_i \in V_U$ , while  $B$  represents the cross-connection between seeded and unseeded nodes where  $v_i \in V_U$  and  $v_j \in V_M$ . Having computed the random walker probability map  $x_i^{(1)}$ , the following pixel-wise decision rule yields the final segmentation mask  $p_i$ .

$$p_i = \begin{cases} 1 & \text{if } x_i^{(1)} \geq 0.5 \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

## 2. Initialization of Random Walks Segmentation for 3D Data

A 3D adaptation of the Grady's original 2D implementation is straightforward. The 6-connected neighborhood in the 3D domain is used to construct the combinatorial Laplacian matrix  $L$ . The linear system in (1) can be solved by  $LU$  decomposition in polynomial time for small images. However, the size of the typical 3D CT images precludes this approach due to large memory requirements. Therefore, we apply the iterative biconjugate gradient stabilized method<sup>61</sup> to solve the combinatorial Dirichlet problem in (1).

A challenge for the 3D adaptation is the increasing difficulty in seeding the regions of interest for 3D images. In order to maintain the overall usability of the semi-automatic CAD system, it is imperative that this segmentation initialization requires only minimal and

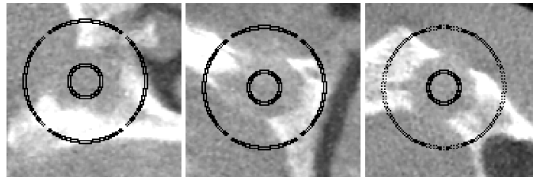


FIG. 1: Example seed points for 3D random walks segmentation. The inner sphere of seed points designates a region inside the lesion while the outer sphere of seed points designates a region outside the lesion.

intuitive user-interaction. We designed an initialization tool that specifies two concentric  
 200 spheres lying inside and outside of the target lesion’s boundary. Three parameters must  
 be specified to place these spheres: the center of the spheres, an inner radius and an outer  
 radius. The inner radius specifies a sphere small enough to be completely contained within  
 the lesion. Likewise, the outer radius specifies a sphere large enough to entirely enclose the  
 lesion. The tool offers a simple GUI to choose these parameters via three mouse clicks. Once  
 205 the two spheres are placed, voxels that lie on the inner and outer sphere are labeled with 1  
 and 0, respectively. An example is provided in Fig.1.

### 3. *Random Walks Segmentation by Likelihood Ratio Test*

Our overall dental CAD problem treats the segmentation step as a pre-process followed by  
 the classification step as a post-process. In this context, a small inclusion of the non-target  
 210 structures at the segmentation step can cause an overall mis-diagnosis while a conservative  
 segmentation missing some peripheral area may not influence the classification outcomes.  
 Hence more conservative segmentation may be preferred. A logical approach is to estimate  
 this problem-specific bias from data in order to achieve the best post-processing result by  
 incorporating classification criteria as a part of the segmentation process.

215 To realize this, we first extend the maximal-probability decision rule of the original  
 random walks segmentation in (3) by using likelihood ratio test (LRT) formalism<sup>24</sup> with  
 a threshold parameter  $k$  introduced to control the trade-off between under- and over-  
 segmentations,

$$p'_i = \begin{cases} 1 (in) & \text{if } p(i|in)/p(i|out) > k \\ 0 (out) & \text{otherwise,} \end{cases} \quad (4)$$

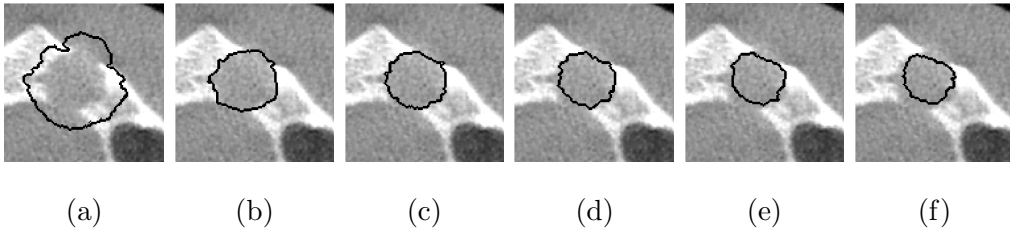


FIG. 2: LRT-extended random walks segmentation with varying threshold values. (a)  $k = 0.05$ , (b)  $k = 0.50$ , (c)  $k = 1.00$ , (d)  $k = 1.50$ , (e)  $k = 2.00$ , (f)  $k = 2.50$ . The original random walks segmentation by Grady<sup>23</sup> is given by (c).

where  $p'_i$  is a resulting binary-valued segmentation mask and  $p(i|in)$  and  $p(i|out)$  denote the spatial likelihood functions, indicating the probability that the voxel is inside or outside the lesion, respectively. These probabilities are computed directly by the random walks algorithm:  $x_i^{(1)} = p(i|in)$  and  $x_i^{(0)} = p(i|out) = 1 - p(i|in)$ . The threshold parameter  $k$  flexibly controls the segmentation results in the standard LRT sense. The 3D segmentation becomes more conservative as  $k$  increases, as exemplified in Fig.2. The original random walks segmentation is equivalent to the case with  $k = 1.0$ .

We propose two data-driven algorithms for optimizing the  $k$  value via the results of trained LDA classifiers<sup>18</sup>: the maximization of total accuracy of multiple LDA classifiers (maximum total accuracy) and the maximization of Bayesian posterior distribution using the classifier's cross validation results (Bayesian cross validation).

#### 4. Maximum Total Accuracy Criterion (MTA)

The first approach, MTA, derives the optimal threshold by maximizing the total accuracy of multiple LDA classifiers trained for varying feature combinations. Suppose a set of supervised training dataset with  $M$  cases, each consisting of a list of  $N$  features and a corresponding ground-truth label. We then consider  $K$  arbitrary subsets of the  $N$  features. For each of the  $K$  feature combinations, the leave-one-out cross-validation (LOOCV) of LDA classifier is performed for  $M' < M$  test cases chosen randomly from the training set.

Total accuracy is defined as the total number of accurately classified test cases among  $K \times M'$  cases. The best decision threshold is then given by the  $k$  value that maximizes the total accuracy measure. This approach seeks the segmentation threshold that maximizes

### 5. Bayesian Cross Validation Criterion (BCV)

The second approach, BCV, seeks to maximize the posterior probability of correct classification by a single LDA classifier with the best performing feature combination determined by a pilot study. LOOCV with  $M'$  randomly chosen test cases is performed to compute  
 245 three performance statistics for each  $k$  value: 1) Accuracy  $f_1(k) = TP + TN / (TP + FP + TN + FN)$ , 2) Sensitivity  $f_2(k) = TP / (TP + FN)$ , and 3) 1-Specificity  $f_3(k) = TN / (TN + FP)$ .  $TP$ ,  $FP$ ,  $TN$ , and  $FN$  denote the number of true positive, false positive, true negative, and false negative cases, respectively. A conditional distribution  $p(f_i|k)$  of the CAD performance given the decision parameter  $k$  is estimated by normalizing the per-  
 250 formance statistics  $f_i(k)$ . We model the prior of the parameter  $k$  by a Gaussian distribution  $p(k) = \frac{1}{\sigma(\sqrt{2\pi})} e^{-\frac{(k-\mu)^2}{2\sigma^2}}$ . The prior is centered at  $\mu = 1.0$  as this is the baseline random walks segmentation. The width  $\sigma$  is manually set broadly to  $\sigma = 1$  so that we focus more on the data term. The BCV criterion is defined as a maximum a posteriori (MAP) estimation with the posterior  $p(k|f_i)$  with three choices of the performance statistics  $\{f_1, f_2, f_3\}$ :  
 255  $k^* = \operatorname{argmax}_k p(k|f_i) = \operatorname{argmax}_k p(f_i|k)p(k)$ . This approach seeks the decision threshold that yields the best possible classification performance among various system settings.

### 6. Feature Extraction

For each segmented lesion, a set of intensity statistics are extracted as a feature vector for the next classification step. In Simon et al.<sup>14</sup>, only the minimum intensity value at the  
 260 center of the lesion was used for classification. In this study, we consider a set of eight features computed from the lesion’s intensity distribution: maximum, minimum, mean, median, standard deviation, skewness, kurtosis, and entropy. Skewness and kurtosis are the third and fourth-order standard central moments, respectively. Entropy follows the standard Shannon’s information entropy formulation by treating the normalized intensity distribution  
 265 as a probability mass function. Our features are then organized as a vector of 8 coefficients without attribute-wise normalization.

## B. Lesion Classification

The final step of our dental CAD framework is to classify lesions using the features extracted in the previous step. We train and compare a number of binary supervised classifiers that map an input feature set to either cyst or granuloma. To adapt the standard terminology for binary classifiers, we treat the cysts as positives and the granulomas as negatives, without the loss of generality.

### 1. Linear Discriminant Analysis and AdaBoost: Review

Linear discriminant analysis (LDA) is a standard supervised machine learning algorithm for a linear classifier with dimensionality reduction. In the binary classification case, the basis vector  $\mathbf{w}$  is used to project the  $N$ -dimensional feature vector  $\mathbf{x}$  into a 1D feature point  $y$  by a linear transformation  $y = \mathbf{w}^T \mathbf{x}$ . The classification of a novel case is given by the label  $q$  of the class mean  $\mu_q$  closest to the input's projection, which amounts to the following decision rule with threshold parameter  $w_0$ ,

$$q = \begin{cases} 1 \text{ (positive)} & \text{if } \mathbf{w}^T \mathbf{x}_i > w_0 = 0.5 * (\mathbf{w}^T \mu_0 + \mathbf{w}^T \mu_1) \\ 0 \text{ (negative)} & \text{otherwise.} \end{cases} \quad (5)$$

LDA estimates the basis vector  $\mathbf{w}$  that separates feature sets of different classes as far from each other as possible by maximizing the cost  $J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}}$ , where  $S_B$  and  $S_W$  denote between-class and within-class scatter matrices, respectively<sup>18</sup>. This study exploits LDA for learning the segmentation decision threshold as described in Sec. II A 3, as well as a part of the LDA-AdaBoost method as will be described in Sec. II B 3.

AdaBoost<sup>17</sup> is an ensemble learner which combines a set of classifiers of low accuracy (weak) in order to derive a high-accuracy (strong) classifier. The resulting classifier of the AdaBoost algorithm is given by

$$H(x) = \text{sign} \left( \sum_{t=1}^T \alpha_t h_t(\mathbf{x}) \right), \quad (6)$$

where  $h_t(\mathbf{x}) : \mathbf{x} \rightarrow \{-1, 1\}$  is the  $t$ -th weak classifier,  $\alpha_t$  is the weighted error rate of  $h_t$ , and  $H(\mathbf{x})$  is the strong classifier. AdaBoost assigns weights  $W(i)$  to the training samples  $\{(\mathbf{x}_i, l_i) | i = 1, \dots, M\}$ , which are initially set uniformly. At each boosting iteration  $t = 1, \dots, T$  for learning  $\{(h_t, \alpha_t, W_t(i))\}$ , the best performing weak classifier  $h_t$  that minimizes the

aggregate error  $\epsilon_j$  with respect to the sample weights is detected:  $h_t(\mathbf{x}) = \operatorname{argmin}_{h_j \in \mathcal{H}} \epsilon_j = \sum_{i=1}^M W_t(i) |_{l_i \neq h_j(\mathbf{x}_i)}$ . The resulting  $h_t$  and corresponding error  $\epsilon_t$  are then used to estimate  $\alpha_t$  and update  $W_t \rightarrow W_{t+1}$ :  $\alpha_t = \frac{1}{2} \ln \frac{1-\epsilon_t}{\epsilon_t}$  and  $W_{t+1}(i) = \frac{W_t(i) \exp(-\alpha_t l_i h_t(\mathbf{x}_i))}{Z_t}$ , where  $Z_t$  is a normalization factor to assure that  $\sum_i W_{t+1}(i) = 1$ . This procedure is repeated and more weak learners are added until either the specified maximum number of rules are added or low training error is reached.

## 2. Weighted LDA

In the standard LDA formulation<sup>18</sup>, there is no notion of sample weights. To this end, we first extend the between- and within-class scatter matrices with sample weights as follows,

$$S'_B = \sum_c \left( \sum_{i \in C} W(i) \right) (\mu'_c - \bar{\mathbf{x}}') (\mu'_c - \bar{\mathbf{x}}')^T \quad (7)$$

$$S'_W = \sum_c \sum_{i \in C} W(i) (\mathbf{x}_i - \mu'_c) (\mathbf{x}_i - \mu'_c)^T \quad (8)$$

$$\mu'_c = \frac{1}{\sum_{i \in C} W(i)} \sum_{i \in C} W(i) \mathbf{x}_i \quad (9)$$

$$\bar{\mathbf{x}}' = \sum_i W(i) \mathbf{x}_i, \quad (10)$$

where  $W(i)$  is the weight assigned to each  $i$ -th training sample with  $\sum_i W(i) = 1$  as was introduced for the AdaBoost algorithm in the previous section. An LDA algorithm defined over the maximization of the new cost  $J'(\mathbf{w}) = \frac{\mathbf{w}^T S'_B \mathbf{w}}{\mathbf{w}^T S'_W \mathbf{w}}$  is referred to as weighted LDA (wLDA). When the weights are uniformly distributed, wLDA becomes equivalent to the standard LDA such that  $J'(\mathbf{w}) = J(\mathbf{w})$ ,  $S'_B = \frac{1}{M} S_B$ , and  $S'_W = \frac{1}{M} S_W$ . For wLDA, we estimate the decision threshold  $w_0$  by minimizing the weighted Bayes error  $\sum_i W(i) |_{l_i \neq q}$  via an exhaustive search, yielding the Bayes optimal classifier.

## 3. LDA-AdaBoost

LDA is a relatively strong and stable learner, and it is generally believed that boosting is not well suited to use such a learner<sup>62</sup>. However, LDA can factor in between-feature correlations, while AdaBoost has shown effectiveness to accurately classify non-linearly separable data sets. We hypothesize that by combining the strengths of AdaBoost and LDA, it may

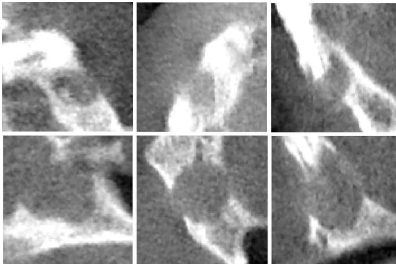


FIG. 3: Example of dental periapical lesions. Sagittal, coronal, and axial view of a granuloma (top) and a cyst (bottom) are displayed. Visually, they are hard to differentiate and include extensive weak boundaries.

be possible to produce an efficient and robust classifier that is superior to AdaBoost paired with the standard linear weak classifier.

315 wLDA is combined with the AdaBoost algorithm as its weak classifier  $h_j$  in the place of the standard linear classifier trained on each single feature by minimizing the Bayes error. We refer this single feature linear classifier as simple threshold classifier or ST. A set of wLDA classifiers with various feature combinations are used as the hypothesis space of AdaBoost. In this study, we consider all two-, three- and eight-feature combinations, while only two-  
 320 feature combinations were previously considered in Liu et al.<sup>55</sup>. At each boosting iteration  $t$ , a set of wLDAs  $\{h_j\}$  are trained with the current weights  $W_t(i)$  by using the methods described in Sec II B 2. The sample weights  $W_t(i)$  are then updated by using the procedure described in Sec II B 1. This process is iterated until the pre-determined minimum error is reached. The resulting strong classifier  $H(\mathbf{x})$  will take the same form as in (6). Note that,  
 325 as the sample weights are adapted, the resulting wLDA classifiers will also adapt over the iterations, yielding different basis and threshold values.

### III. RESULTS

#### A. CBCT Data

330 A dataset of 28 anonymous 3D dental CBCT scans are available for our study in the DICOM format. This study was approved by the institutional review board at University of Southern California where the data were collected. The CBCT images were captured using

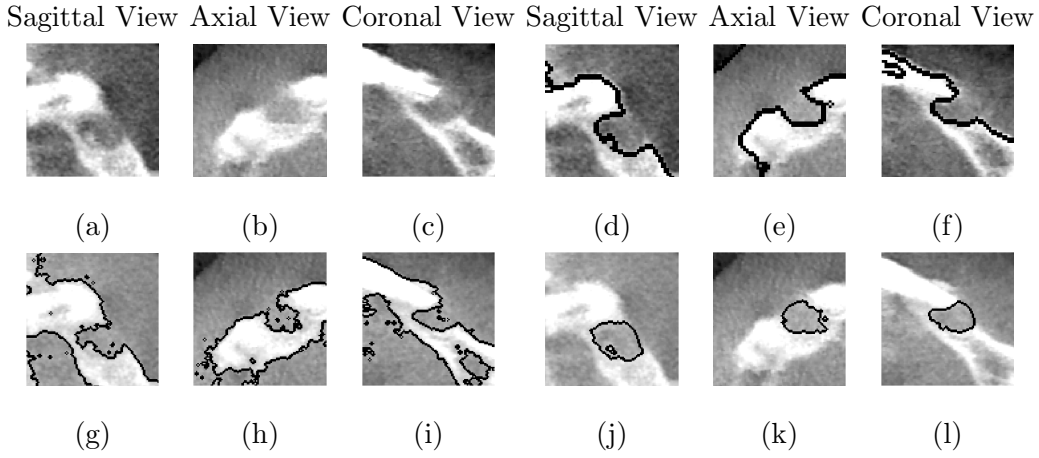


FIG. 4: 3D segmentation results of typical example for granuloma. Original image (a-c), normalized cut (d-f), graph cut (g-i), and random walks (j-l). The second through the third eigen vectors are used for the normalized cut. The thick boundary lines on (d-f) are artifacts from the Gaussian Sub-sampling method.

the NewTom 3G scanner, QR Srl, Verona, Italy. The NewTom 3G acquires 360 images at 1-degree intervals in 36 seconds, with reconstructed image resolution of 512\*512 pixels and 12 bits per pixel (4,096 grayscale). The pixel size was 0.25mm x 0.25mm. The axial slice thickness was 0.2mm. A VOI of 100 cubic voxels is cropped for each lesion to run our experiments efficiently. Fig.3 shows an example of each type of lesion. Each lesion was diagnosed as either a cyst or granuloma by endodontic experts using the CBCT data as recommended in Simon et al.<sup>14</sup>. There were 14 cyst and 14 granuloma cases in the dataset based on CBCT evaluation. The size of lesions was equal to or greater than 1 *cm* in diameter. Our dataset includes the 17 cases (11 cyst and 6 granuloma cases) used in the study of Simon et al.<sup>14</sup>. These cases also underwent a histological biopsy, providing an alternative set of ground-truth labels for these 17 cases.

Between the two ground-truth sets for the 17 cases used in Simon et al.<sup>14</sup>, there were four discordant, or split, cases which were categorized as cysts by the endodontists and as granulomas (apical periodontitis) by oral pathologists. Simon et al. argued the correctness of their endodontic labels based on careful interpretation of the histopathological notes which pointed to misdiagnoses due to their small size and fluid cavities which have lost their epithelium<sup>14</sup>.



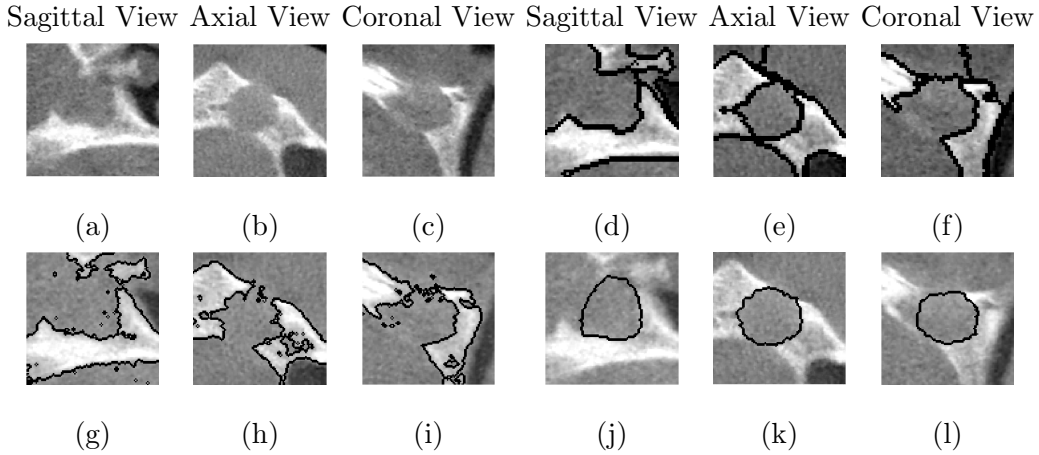


FIG. 5: 3D segmentation results of typical example for cyst. Original image (a-c), normalized cut (d-f), graph cut (g-i), and random walks (j-l). The second through the sixth eigen vectors are used for the normalized cut.

## B. Lesion Segmentation

350 First, we assess the effectiveness of the classic Random Walks segmentation<sup>23</sup> for our data, in comparison with two other well-known graph-based algorithms: normalized cut<sup>63</sup> and graph cut<sup>64</sup>. The default value of  $\beta = 90$  is used for the random walks. For the normalized cut, a lower resolution version of images using the Gaussian pyramid is used to handle its high memory requirement, and we manually chose eigen modes to be included in  
 355 order to achieve the best segmentation for each case. For the graph cut, we employ the  $\alpha$  expansion moves and initialize it by the results of K-means clustering of the voxel intensity. The same human inputs, indicating the inside and outside of a lesion used to initialize the Random Walkers algorithm, are used to initialize the two cluster centers for this K-mean clustering. We observed that the variation in the initial cluster has minimal impact  
 360 on the final segmentation. For the proposed Random Walks algorithm, our pilot study showed its insensitivity against varying initialization of the lesion center and that novice operators could quickly become adequate in performing this semi-automatic segmentation without errors after minimal training. Grady et al.<sup>65</sup> also reported a systematic sensitivity validation study with 5 cases over 1000 random trials, claiming that seed perturbations  
 365 within the target object produced changes in the segmentation under 5%, which indicates the insensitivity to initialization in more general settings.

Figs. 4 and 5 show illustrative examples of segmentation results by the three methods for granuloma and cyst cases, respectively. They demonstrate that the random walks algorithm clearly outperforms other segmentation methods on our data. The random walks algorithm (Figs. 4-5(j-l)) succeeds in delineating the weak lesion boundary and extracting a complete and closed lesion without much over-segmentation. We also observed that variance in the spherical radii did not have a large impact on the segmentation result. Normalize cut (Figs. 4-5(d-f)) failed to delineate weak boundaries. In some cases, the cuts occur outside the lesion boundary (e.g., Fig. 5). Without a cut along the boundary of the lesion, a complete and closed lesion segmentation is not possible by the normalized cut without additional user input. Graph cut (Figs. 4-5(g-i)) tends to result in over-segmentation. Like the normalized cut, weak lesion boundaries are not always segmented by this method. These examples shown here were representative for all available cases.

In general, however, the normalized and graph cut algorithms tend to delineate intensity edge better than the random walks algorithm. Segmented boundary by random walks sometimes cut through the high-intensity bone area near the boundary despite its superior performance in completing the weak boundaries. This poses a problem in terms of an overall CAD system's performance as discussed in Sec.II A 3.

Table I compares the time-complexity of the three segmentation algorithms compared above. The execution times presented here were obtained on a Pentium Core Duo 1.83 GHz. The algorithms were implemented in C/C++ (MEX) in MATLAB (Mathworks Inc., Natick, MA). Although the normalized cut algorithm uses the Lanczos method<sup>66</sup> for solving the eigen-value system, the running time of creating the sparse affinity matrix is quadratic. Both the graph cut and random walks algorithms have a near linear increase in the running time with larger image sizes. Random walks has higher complexity than graph cut especially for larger inputs, however the difference in accuracy outweighed our choice toward random walks.

Next, our proposed segmentation method with the LRT extension described in Sec. II A 3 is evaluated. Fig. 6(a) illustrates the MTA criterion for our dataset. The total accuracy over LDA classifiers with every possible two, three, and eight feature combinations ( $K = 85$ ) is computed for discrete  $k$  values of the LRT-threshold in the range  $[0.05, 2.50]$  in increments of 0.05 via LOOCV with all available test cases ( $M' = M = 28$ ). The maximum is observed at  $k = 1.5$  by an exhaustive search. Fig. 6(b,c) shows the BCV criterion for our dataset.

TABLE I: Computational complexity of Random walks, Graph cut, and Normalized cut segmentation.

| # of Voxels | Normalized Cut       | Graph Cut | Random Walks |
|-------------|----------------------|-----------|--------------|
| $25^3$      | 38 sec               | < 1 sec   | 2 sec        |
| $50^3$      | 2494 sec             | 1 sec     | 30 sec       |
| $75^3$      | Memory limit reached | 6 sec     | 80 sec       |
| $100^3$     | Memory limit reached | 12 sec    | 277 sec      |

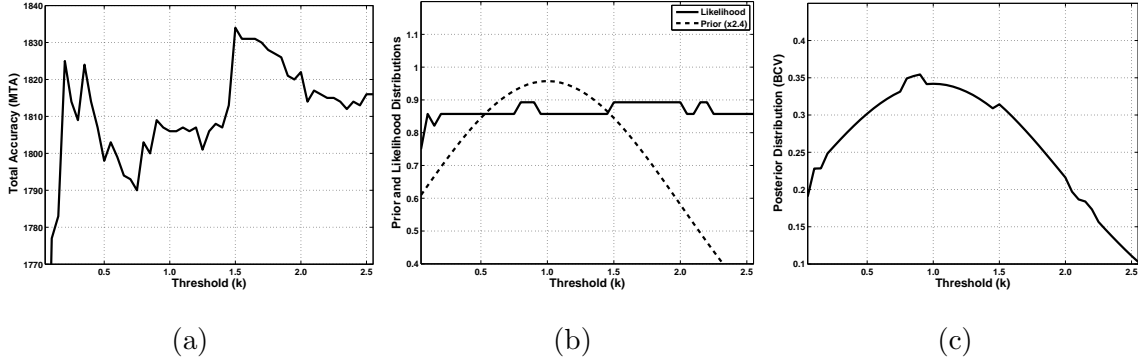


FIG. 6: Estimation of LRT threshold  $k$ . (a) MTA criterion: the total accuracy over varying threshold  $k \in [0.05, 2.50]$  in increments of 0.05. (b) Prior and likelihood distributions over varying threshold  $k$ . Gaussian prior centered at  $k = 1$  is scaled and translated for visual enhancement. (c) BCV criterion: posterior distributions over varying threshold  $k$ .

Fig. 6(b) demonstrates the likelihood and prior distributions and Fig. 6(c) demonstrates the  
 400 posterior distribution over varying  $k$ . We model the likelihood distribution by the LOOCV  
 accuracy measure  $f_1(k)$  with the best performing LDA classifier with the minimum and  
 entropy intensity features (See Sec. III C for experimental results validating this choice).  
 The maximum is observed at  $k = 0.9$ .

Overall, for all of the 28 cases, both MTA and BCV approaches result in an appropriate  
 405 segmentation, confirmed by visual inspection. Fig.7 shows two illustrative cases with three  $k$   
 values: 0.9 by BCV (a,d), 1.0 by the original random walk (b,e), and 1.5 by MTA (c,f). We  
 observe that the threshold identified by MTA resulted in conservative segmentations that  
 are preferred in our CAD application context. For example, the BCV’s result in Fig.7(a)

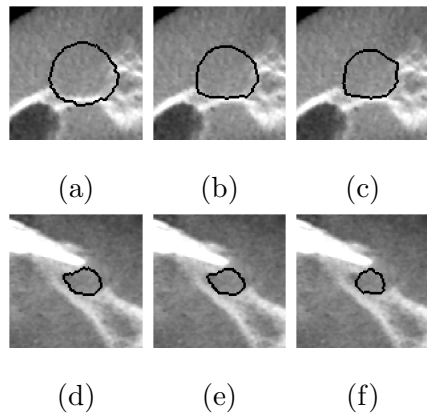


FIG. 7: Two illustrative cases for comparing MTA and BCV criteria. (a,d):  $k = 0.9$  estimated by BCV, (b,e):  $k = 1.00$  with the original random walks algorithm, (c,f):  $k = 1.5$  estimated by MTA.

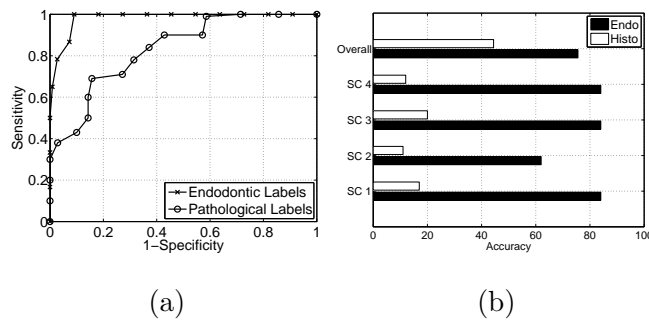


FIG. 8: Analysis of ground-truth data. (a) ROC curves generated with the two GT sets. (b) Comparison of LOOCV accuracy for each of the four split cases and the overall average over the split cases. Note the classification accuracy is significantly higher when using the endodontic diagnosis as ground-truth, as suggested in Simon et al.<sup>14</sup>.

incorrectly includes a large amount of dense bone tissue at the bottom of the lesion, which  
 410 is correctly removed by MTA with a higher threshold value as shown in Fig.7(c). Note  
 however that a lower threshold can sometimes yield more accurate segmentation as shown  
 in Fig.7(d). MTA yielded a less accurate result in this case shown in Fig.7(f), however  
 it still identified the majority of the lesion while simultaneously abstaining from including  
 non-target dense tissues.

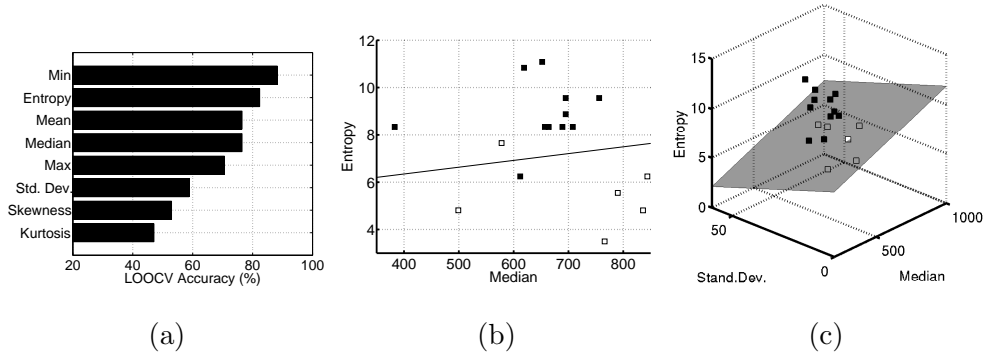


FIG. 9: Analysis of feature space. (a) LOOCV accuracy of Bayes optimal classifiers trained for each single feature. (b) 2D scatter plot with the LDA decision line for the best feature pair. (c) 3D scatter plot with the LDA decision plane for the best feature triple. Black boxes: cysts, white boxes: granulomas.

415 **C. Lesion Classification**

First, we analyze the classification ground-truth (GT) of our data by comparing the CBCT endodontic and histopathological labels. In the 17-case subset, there were four split cases which were categorized as cysts by the endodontists and as granulomas (apical periodontitis) by oral pathologists. In order to investigate these split cases, we trained our CAD system for the 17-case subset with the two GT sets and compared their performance. Fig. 8(a) compares ROC curves for the two GT sets, averaging the LOOCV statistics of the ten best performing two- and three- feature combinations among the eight features at varying threshold values. And Fig. 8(b) shows the LOOCV accuracy measures for the four split cases averaged over all two- and three-feature combinations. LOOCV is performed with the standard LDA classifier for its simplicity. The system with the endodontic GTs clearly performed better than that with the histopathological GT. Area-under-the-ROC-curves (AUCs) are 0.98 and 0.83, and sensitivities of 100.0% and 69.0% with false positive rates of 9.1% and 15.7% at the maximum accuracies of 94.1% and 78.9% are observed for the endodontic set and the histopathological set, respectively. For the four split cases, the endodontic GT also yielded largely higher accuracy than the histopathological GTs as shown in Fig. 8(b), lending further support to the claim that the endodontic diagnosis is more consistent, as argued in Simon et al.<sup>14</sup>. We use the endodontic GT set for the rest of experiments.

TABLE II: The proposed LDA-AdaBoost method compared with LDA, standard AdaBoost, and LDA-AdaBoost with wLDA proposed by Liu et al.<sup>55</sup>. Left: accuracy averaged over 7-fold cross validation with 28 cases. Right: p-value of Wilcoxon signed ranks test against LDA.

| Classifiers                              | Average Accuracy (%) $\pm$ (%) | p-value (vs LDA) |
|--|--------------------------------|------------------|
| LDA                                      | 75.0 $\pm$ 25.0                | N/A              |
| AdaBoost                                 | 75.0 $\pm$ 38.2                | 1.000            |
| LDA-AdaBoost (Liu et al. <sup>55</sup> ) | 78.6 $\pm$ 17.3                | 0.564            |
| LDA-AdaBoost (Proposed)                  | 85.7 $\pm$ 13.4                | 0.083            |

Next, we evaluate the set of eight intensity features. Fig. 9(a) compares the LOOCV accuracy of the Bayes optimal linear classifiers trained for each single feature. The results revealed that the minimum intensity and entropy features performed best at 88.2% and 82.4%, respectively. This agrees with the choice of intensity features used in the Simon’s method<sup>14</sup>. The 8D feature space is then explored by training LDA classifiers with all permutations of two (28 pairs), three (56 triples), and eight (1 8-tuple) features and comparing their LOOCV accuracy. The best accuracy of 88.2% was recorded by a number of feature pairs and triples. Figs. 9(b-c) shows the scatter plots with LDA decision boundary for the best performing pair (entropy and median) and triple (entropy, median, and standard deviation), respectively. Note that the feature pairs and triples that do not match with the greedy-selected top two or three features in Fig. 9(a) may still represent the best performing feature subsets. However the single feature classifier with the minimum intensity feature performed as good as any of the best performing LDA classifiers with feature combinations, which suggests that the LDA classifier itself does not improve the minimum intensity strategy of Simon et al.<sup>14</sup>.

Lastly, we evaluate the performance of the proposed LDA-AdaBoost with the complete data set of 28 cases. A 7-fold cross validation is performed using all the feature combinations considered in this study for the proposed system, as well as three other baseline methods: 1) LDA, 2) standard AdaBoost, and 3) LDA-AdaBoost with the wLDA formulation proposed in Liu et al.<sup>55</sup>. Table II summarizes the results. On average, the proposed method performed substantially better, however no statistically significant differences were observed among any

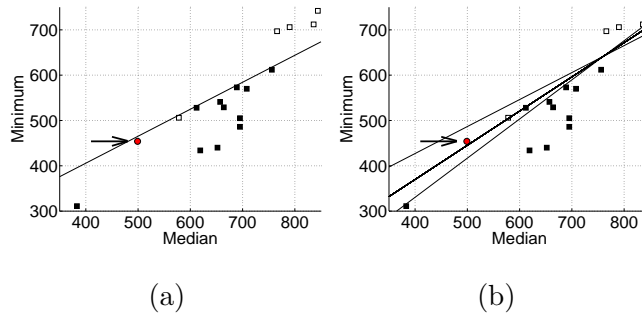


FIG. 10: Illustrative example of LDA-AdaBoost vs LDA. Circle with an arrow: a test granuloma case; Black boxes: cysts; White boxes: granulomas. (a) LDA, (b) LDA-AdaBoost.

classifier pairs by non-parametric Wilcoxon signed rank test:  $p = 0.083$  and  $p = 0.157$  for  
 455 our LDA-AdaBoost vs LDA and for our LDA-AdaBoost vs Liu et al.'s approach<sup>55</sup>, respec-  
 tively. Additionally, we performed feature selection by choosing the best performing feature  
 combination among the 85 tested combinations. The highest cross validation accuracy of  
 94.1% was recorded by the proposed LDA-AdaBoost using the feature pair of median and  
 minimum intensity. Fig. 10 illustrates how the proposed LDA-AdaBoost improves classifi-  
 460 cation accuracy over LDA. By adding a few more wLDA weak classifiers, LDA-AdaBoost is  
 able to correctly classify the test case which was missed by LDA.

#### IV. CONCLUSIONS

This paper presents a dental CAD system for non-invasive differential diagnosis of periapi-  
 cal lesion, which integrates graph-based random walks segmentation and machine learning-  
 465 based boosted classification algorithms. We proposed two novel approaches: LRT-based  
 extension of the random walks method and LDA-AdaBoost that adapts AdaBoost with  
 weighted LDA classifier used as weak learner.

The results of our experimental studies demonstrate the effectiveness of the proposed  
 method with 94.1% when we use the endodontic diagnosis as the ground-truth. Due to the  
 470 per-scan adaptive shifting of radiation to minimize radiation dosage to patients, the intensity  
 mapping is inconsistent across CBCT scans. Moreover, the morphology of the periapical  
 lesions can vary. Simple threshold segmentation and classification is therefore not a viable  
 option. As can be seen in Fig.3, it is non-trivial to segment or classify these lesions even to  
 the trained eyes. The interface between these periapical lesions and adjacent soft tissues is

475 often vague, making the parts of lesion contours to be ill-defined. Random walks has been shown to be suitable for delineating this type of objects<sup>23</sup>, supporting our methodological choice. Overall, the semi-automatic approach presented in this paper may prove to be more accurate, reproducible, and less prone to human error. The best performance at 94.1% by the proposed LDA-AdaBoost classifier improves those at 88.2% by the Bayes optimal classifier  
480 on single features shown in Fig.9(a) by 5.9%. This baseline Bayes classifier represents the best-case extension of the Simon’s method that relies on manual sampling of voxels within a lesion. Therefore the 5.9% performance increase by LDA-AdaBoost demonstrates its advantage beyond the virtue of our segmentation step.

Furthermore, the results of our ground-truth analysis reveals our dental CAD system  
485 agrees more with the CBCT diagnosis than the histopathological diagnosis. In 13 out of the 17 cases used in the Simon’s study<sup>14</sup>, the CBCT and histopathological diagnoses (ground-truth) coincide, but in the remaining four cases their diagnoses differ. Suppose the endodontic diagnosis is our gold standard, then the histopathological diagnosis is correct on 13 out of 17 cases with 76.5% accuracy, and vice versa. In this case, our method  
490 trained with the CBCT ground-truth yielding 94.1% best accuracy can be interpreted as an 17.6% improvement from the histopathology expert’s diagnosis. On the other hand, taking the histopathology as our gold standard, our method trained with the biopsy ground-truth yielding 78.9% accuracy is still slightly better than that of the endodontic experts, demonstrating the advantage of our approach for improving accuracy while not requiring intensive  
495 manual labor by experts.

Simon et al.<sup>14</sup> argued that the CBCT diagnosis can be more accurate than the histopathology for differentiating periapical lesions. Some cystic lesions may consist of small and/or fragmented epithelial areas that can also be destructed naturally<sup>14</sup>. Therefore, a small number of sections made commonly for the histological slides may miss epithelium or other  
500 evidence of cystic organization, causing the cysts to be misdiagnosed as granulomas<sup>14</sup>. A recent clinical study with 36 cases by Guo et al.<sup>43</sup> has also reported that the CBCT images can be used to provide moderately accurate differential diagnosis. The results of our study offer another support for their argument in that the CBCT-based diagnosis for periapical lesions is more accurate, safer and less invasive than histopathology.

505 Nair<sup>67</sup> argued that a type of radicular cysts containing epithelium-lined cavities that are open to the root canals may also heal after conventional root canal therapy, while those



containing cavities completely enclosed in epithelial lining may not. This argument, if true, implies that a non-invasive diagnosis provides clinical impact higher than those argued in this paper only involving the granulomas' potential to heal without surgery. An interesting  
510 problem will then be to differentiate the two cyst types non-invasively. Investigating the effectiveness of the CAD approach to this problem remains our future work.

Usability of the proposed CAD framework depends largely on its computational-complexity. The proposed framework would not be effective within dentists' clinical workflow if it required prohibitively long execution time even when it is highly accurate. In our CAD  
515 design, most of the run-time computation is accounted for the segmentation step, thus increasing its efficiency is crucial. Our choice of random walks segmentation is motivated by making the best trade-off between the segmentation step's efficiency and accuracy. The proposed LRT-based extension of the random walks incurs additional training-time computation which could take longer than the run-time of the segmentation, however they can be  
520 executed off-line and much less frequently, providing less impact to the overall workflow.

Beyond the specific focus of the CAD problem, the proposed LRT-extension of random walks and LDA-AdaBoost algorithms can be applicable to a wide range of clinical tasks related to general CAD. The proposed CAD design with the LRT-extension of random walks segmentation introduces a learning process to the segmentation step through a feedback from  
525 the classification step. Due to its generic design, the proposed algorithm is also applicable to any soft segmentation method that employs a probability/confidence mask, followed by a post classification process.

In order to transfer the proposed technology to clinical practice and to avoid overfitting in the resulting classifier, we will collect more data and consider other popular machine  
530 learning methods, such as naive Bayesian classifier, support vector machine, and random forest classifier, for further testing in our future work. Designing a problem-specific features could also boost our classifier's performance. To this end, our future work will incorporate distance-based features to better analyze the voxels surrounding the tip of the periapical root and also some geometric features, such as locations relative to tooth's apex, shape circularity, and well-definedness of boundary, explored in a recent clinical study by Guo  
535 et al.<sup>43</sup>.

## Acknowledgments

The authors would like to thank Reyes Enciso for providing the data and valuable feedback on interpreting our results. This work was partially supported by the California State  
540 University faculty-student collaborative research seed grant, by the San Francisco State University Center for Computing for Life Sciences, and by a philanthropic gift from the Government of Abu Dhabi to Children's National Medical Center.

---

\* Email: [kazokada@sfsu.edu](mailto:kazokada@sfsu.edu); Author to whom correspondence should be addressed.; Website:  
<http://online.sfsu.edu/kazokada/>

- 545 <sup>1</sup> S. Bhaskar, *Periapical lesion - types, incidence and clinical features*, Oral Surg Oral Med Oral Pathol **21** (1966), pp. 657–671
- <sup>2</sup> J. Ingle, L. Bakland and J. Baumgartner, *Ingle's Endodontics 6* (BC Decker Inc, 2007)
- <sup>3</sup> M. Fernandes and I. de Ataide, *Nonsurgical management of periapical lesions*, J Conserv Dent. **13** (2010) (4), pp. 240–245
- 550 <sup>4</sup> J. Simon, *Incidence of periapical cysts in relation to the root canal*, Journal of Endodontics **6** (1980) (11), pp. 845–848
- <sup>5</sup> S. White, J. Sapp, B. Seto and N. Mankovich, *Absence of radiometric differentiation between periapical cyst and granulomas*, Oral Surg Oral Med Oral Pathol Oral Radiol Endod **78** (1994), pp. 650–654
- 555 <sup>6</sup> M. Trope, J. Pettigrew, J. Petras, F. Barnett and L. Tronstad, *Differentiation of radicular cysts and granulomas using computerized tomography*, Endodontics and Dental Traumatology **5** (1989), pp. 69–72
- <sup>7</sup> E. Jorge, M. Tanomaru-Filho, M. Goncalves and J. Tanomaru, *Detection of periapical lesion development by conventional radiography or computed tomography*, Oral Surg Oral Med Oral  
560 Pathol Oral Radiol Endod **106** (2008), pp. 56–61
- <sup>8</sup> W. Priebe, J. Lazansky and A. Wuehrmann, *The Value of the Roentgenographic Film in the Differential Diagnosis of Periapical Lesions*, Oral Surg Oral Med Oral Pathol **7** (1954), pp. 979–983
- <sup>9</sup> J. McCall and W. S, *Clinical Dental Radiology, 4th Ed.* (Philadelphia: Saunders, 1954)

- 565 <sup>10</sup> F. Howell and V. De la Rosa, *Cytologic evaluation of cystic lesions of the jaws: a new diagnostic technique*, Calif Dent Assoc **36** (1968), pp. 161–166
- <sup>11</sup> K. Dula, R. Mini, P. van der Stelt, J. Lambrecht, P. Schneeberger and D. Buser, *Hypothetical mortality risk associated with spiral computed tomography of the maxilla and mandible*, European Journal of Oral Sciences **104** (1996), pp. 503–510
- 570 <sup>12</sup> C. Kau, M. Bozic, J. English, R. Lee, H. Bussa and R. Ellis, *Cone-beam computed tomography of the maxillofacial region - an update*, Int J Med Robotics and Comp Assist Surg **5** (2009), pp. 366–380
- <sup>13</sup> D. Schulze, M. Heland, H. Thurmman and G. Adam, *Radiation exposure during midfacial imaging using 4- and 16-slice computed tomography, cone beam computed tomography systems and conventional radiography*, Dentomaxillofacial Radiology **33** (2004), pp. 83–86
- 575 <sup>14</sup> J. Simon, R. Enciso, J.-M. Malfaz, M. Bailey-Perry and A. Patel, *Differential Diagnosis of Large Periapical Lesions Using Cone-Beam Computed Tomography Measurements and Biopsy*, Journal of Endodontics **32** (2006) (9), pp. 833–837
- <sup>15</sup> W. De Vos, J. Casselman and G. Swennen, *Cone-beam computerized tomography (CBCT) imaging of the oral and maxillofacial region: a systematic review of the literature*, Int J Oral Maxillofac Surg **38** (2009) (6), pp. 6009–25
- 580 <sup>16</sup> B. Peng, L. Zhang and D. Zhang, *A survey of graph theoretical approaches to image segmentation*, Pattern Recognition **46** (2013) (3), pp. 1020–1038
- <sup>17</sup> Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, in *European Conference on Computational Learning Theory* (1995), pp. 23–37
- 585 <sup>18</sup> R. O. Duda, H. P. E. and D. G. Stork, *Pattern Classification* (Wiley-Interscience Publication, 2000)
- <sup>19</sup> H. Chan, L. Hadjiiski, C. Zhou and B. Sahiner, *Computer-Aided Diagnosis of Lung Cancer and Pulmonary Embolism in Computed Tomography - A Review*, Acad Radiol **15** (2008), pp. 535–555
- 590 <sup>20</sup> J. Tang, R. Rangayyan, J. Xu, I. El Naqa and Y. Yang, *Computer-Aided Detection and Diagnosis of Breast Cancer with Mammography: Recent Advances*, IEEE trans. Info. Tech. in Biomed **13** (2009), pp. 236–251
- 595 <sup>21</sup> R. Summers, *Challenges for Computer-Aided Diagnosis for CT Colonography*, Abdom Imaging

**27** (2002), pp. 268–274

- <sup>22</sup> M. Linguraru, W. Richbourg, J. Liu, J. Watt, V. Pamulapati, S. Wang and R. Summers, *Tumor Burden Analysis from CT Data of Diseased Patients via Automated Liver and Tumor Segmentation*, IEEE Transactions on Medical Imaging **31** (2012) (10), pp. 1965–76
- <sup>23</sup> L. Grady, *Random Walks for Image Segmentation*, IEEE Transaction on Pattern Analysis and Machine Intelligence **28** (2006) (11), pp. 1768–1783
- <sup>24</sup> W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2 (Wiley, New York, 1971)
- <sup>25</sup> D. Paulus, M. Wolf, S. Meller and H. Niemann, *Three-dimensional computer vision for tooth restoration*, Med Img Anal **3** (1999), pp. 1–19
- <sup>26</sup> Y. Hirogaki, T. Sohmura, H. Satoh, J. Takahashi and K. Takada, *Complete 3-D Reconstruction of Dental Cast Shape Using Perceptual Grouping*, IEEE trans Med Imaging **20** (2001), pp. 1093–1101
- <sup>27</sup> S. Yamany, A. Farag, D. Tasman and A. Farman, *A 3-D Reconstruction System for the Human Jaw Using a Sequence of Optical Images*, IEEE trans Med Imaging **19** (2000), pp. 538–547
- <sup>28</sup> T. Kondo, S. Ong and K. Foong, *Tooth Segmentation of Dental Study Models Using Range Images*, IEEE trans Med Imaging **23** (2004), pp. 350–362
- <sup>29</sup> T. Miyazaki, Y. Hotta, J. Kunii, S. Kuriyama and Y. Tamaki, *A review of dental CAD/CAM: current status and future perspectives from 20 years of experience*, Dental Materials Journal **28** (2009), pp. 44–56
- <sup>30</sup> J. Chapuis, A. Schramm, I. Pappas, W. Hallermann, K. Schwenzer-Zimmerer, F. Langlotz and M. Caversaccio, *A New System for Computer-Aided Preoperative Planning and Intraoperative Navigation During Corrective Jaw Surgery*, IEEE trans Med Imaging **11** (2007), pp. 274–287
- <sup>31</sup> K. Verstreken, J. Van Cleynenbreugel, K. Martens, G. Marchal, D. van Steenberghe and P. Suetens, *An Image-Guided Planning System for Endosseous Oral Implants*, IEEE trans Med Imaging **17** (1998), pp. 842–852
- <sup>32</sup> N. Pitts and C. Renson, *Reproducibility of Computer-aided Image-analysis-derived Estimates of the Depth and Area of Radiolucencies in Approximal Enamel*, J Dent Res **64** (1985), pp. 1221–1224
- <sup>33</sup> A. Wenzel, *Bitewing and digital bitewing radiography for detection of caries lesions*, J Dent Res **83** (2004), pp. C72–C75

- <sup>34</sup> T. Southard and K. Southard, *Detection of Simulated Osteoporosis in Maxillae using Radiographic Texture Analysis*, IEEE trans Biomed Engineering **43** (1996), pp. 123–132
- <sup>35</sup> P. Allen, J. Graham, D. Farnell, E. Harrison, R. Jacobs, K. Nicopolou-Karayianni, C. Lindh,  
<sup>630</sup> P. van der Stelt, K. Horner and H. Devlin, *Detecting Reduced Bone Mineral Density From Dental Radiographs Using Statistical Shape Models*, IEEE trans Information Technology in Biomedicine **11** (2007), pp. 601–610
- <sup>36</sup> P. van der Stelt and W. Geraets, *Computer-Aided Interpretation and Quantification of Angular Periodontal Bone Defects on Dental Radiographs*, IEEE trans Biomed Engineering **38** (1991),  
<sup>635</sup> pp. 334–338
- <sup>37</sup> A. Mol and P. van der Stelt, *Application of Digital Image Analysis in Dental Radiography for the Description of Periapical Bone Lesions: A Preliminary Study*, IEEE trans Biomed Engineering **38** (1991), pp. 357–359
- <sup>38</sup> A. Mol and P. van der Stelt, *Application of computer-aided image interpretation to the diagnosis  
<sup>640</sup> of periapical bone lesions*, Dentomaxillofac Radiol **21** (1992) (4), pp. 190–4
- <sup>39</sup> S. Li, T. Fevens, A. Krzyzak, C. Jin and S. Li, *Semi-automatic computer aided lesion detection in dental X-rays using variational level set*, Pattern Recognition **40** (2007), pp. 2861–2873
- <sup>40</sup> T. Lee, H. Jeong, D. Kim and B. Lee, *Quantitative Analysis of Endodontic Treatment for Periapical lesions in Intraoral Radiographs*, in *IEEE Int Conf Bioinformatics and Bioengineering*  
<sup>645</sup> (2009), pp. 352–355
- <sup>41</sup> A. Petersson, S. Axelsson, T. Davidson, F. Frisk, M. Hakeberg, T. Kvist, A. Norlund, I. Mejare, I. Prtenier, H. Sandberg, S. Tranaeus and G. Bergenholtz, *Radiological diagnosis of periapical bone tissue lesions in endodontics: a systematic review*, Int Endod J **45** (2012) (9), pp. 783–801
- <sup>42</sup> E. Cotti, P. Vargiu, C. Dettori and G. Mallarini, *Computerized tomography in the management  
<sup>650</sup> and follow-up of extensive periapical lesion*, Endodontics and Dental Traumatology **15** (1999), pp. 186–189
- <sup>43</sup> J. Guo, J. Simon, P. Sedghizadeh, O. Soliman, T. Chapman and R. Enciso, *Evaluation of the Reliability and Accuracy of Using Cone-beam Computed Tomography for Diagnosing Periapical Cysts from Granulomas*, Journal of Endodontics **39** (2013), pp. 1485–90
- <sup>655</sup> <sup>44</sup> V. Aggarwal, A. Logani and N. Shah, *The evaluation of computed tomography scans and ultra-sounds in the differential diagnosis of periapical lesions*, Journal of Endodontics **34** (2008), pp. 1312–5

- 45 P. Rosenberg, J. Frisbie, J. Lee, K. Lee, H. Frommer, S. Kottal, J. Phelan, L. Lin and G. Fisch,  
*Evaluation of pathologists (histopathology) and radiologists (cone beam computed tomography)*  
660 *differentiating radicular cysts from granulomas*, Journal of Endodontics **36** (2010) (3), pp. 423–  
8
- 46 R. Walton, *Diagnostic imaging A. endodontic radiography*, in *Ingle's Endodontics 6* (BC Decker  
Inc, 2007), pp. 554–572
- 47 W. Scarfe, M. Levin, D. Gane and A. Farman, *Use of Cone Beam Computed Tomography in*  
665 *Endodontics*, Int J Dent **2009** (2009), pp. 1–20
- 48 Y. Nakagawa, K. Kobayashi, H. Ishii, A. Mishima, H. Ishii, K. Asada and K. Ishibashi, *Preop-*  
*erative application of limited cone beam computerized tomography as an assessment tool before*  
*minor oral surgery*, Int J Maxillofac Surg **31** (2002), pp. 322–327
- 49 M. Guerrero, R. Jacobs, M. Loubele, F. Schutyser, P. Suetens and D. van Steenberghe, *State-*  
670 *of-the-art on cone beam CT imaging for preoperative planning of implant placement*, Clin Oral  
Invest **10** (2006), pp. 1–7
- 50 L. Cevidanes, L. Bailey, S. Tucker, M. Styner, A. Mol, C. Phillips, W. Proffit and T. Turvey,  
*Three-dimensional cone-beam computed tomography for assessment of mandibular changes after*  
*orthognathic surgery*, Dentomaxillofacial Radiology **33** (2004), pp. 83–86
- 675 51 Y. Ding, G. Vachtsevanos, A. Yezzi Jr, Y. Zhang and Y. Wardi, *A Recursive Segmentation*  
*and Classification Scheme for Improving Segmentation Accuracy and Detection Rate in Real-*  
*Time Machine Vision Applications*, in *Int Conf Digital Signal Processing* (2002), vol. 2, pp.  
1009–1013
- 52 E. Tang, P. Suganthan, X. Yao and A. Qin, *Linear dimensionality reduction using relevance*  
680 *weighted LDA*, Pattern Recognition **38** (2005), pp. 485–493
- 53 X. Li, L. Wang and E. Sung, *Adaboost with SVM-based component classifiers*, Engineering  
Applications of Artificial Intelligence **21** (2008), pp. 785–795
- 54 T. Truyen, D. Phung, S. Venkatesh and H. Bui, *AdaBoost.MRF: Boosted Markov Random*  
*Forests and Application to Multilevel Activity Recognition*, in *Int Conf on CVPR* (2006), pp.  
685 1686–1693
- 55 X. Liu, L. Zhang, M. Li, H. Zhang and D. Wang, *Boosting image classification with LDA-based*  
*feature combination for digital photograph management*, Pattern Recognition **38** (2005) (6), pp.  
887–901

- <sup>56</sup> M. Skurichina and R. Duin, *Boosting in Linear Discriminant Analyses*, in *Int Workshop on Multiple Classifier Systems* (2000), pp. 190–199
- <sup>57</sup> S. Rysavy, A. Flores, R. Enciso and K. Okada, *Segmentation of Large Periapical Lesions toward Dental Computer-Aided Diagnosis in Cone-Beam CT Scans*, in *SPIE Medical Imaging* (San Diego, 2008), pp. 6914–153
- <sup>58</sup> S. Rysavy, A. Flores, R. Enciso and K. Okada, *Classifiability Criteria for Refining of Random Walks Segmentation*, in *Int Conf Pattern Recognition* (2008), pp. 1–4
- <sup>59</sup> A. Flores, S. Rysavy, R. Enciso and K. Okada, *Non-Invasive Differential Diagnosis of Dental Periapical Lesions in Cone-Beam CT Scan*, in *IEEE Int Symp Biomed Img* (2009), pp. 566–569
- <sup>60</sup> A. Flores, M. G. Lingurarur and K. Okada, *Boosted-LDA for Biomedical Data Analysis*, in *International Workshop on Machine Learning in Medical Imaging* (2010)
- <sup>61</sup> R. Barrett, *Templates for the solution of linear systems: building blocks for iterative methods* (Society for Industrial Mathematics, 1994)
- <sup>62</sup> J. Lu, K. Plataniotis and A. Venetsanopoulos, *Boosting linear discriminant analysis for face recognition*, in *International Conference on Image Processing* (2003), vol. 1, pp. I–657
- <sup>63</sup> J. Shi and J. Malik, *Normalized Cuts and Image Segmentation*, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22** (2000) (8), pp. 888–905
- <sup>64</sup> Y. Boykov, O. Veksler and R. Zabih, *Fast Approximate Energy Minimization via Graph Cuts*, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23** (2001) (11), pp. 1222–1239
- <sup>65</sup> L. Grady, T. Schiwietz, S. Aharon and R. Westermann, *Random Walks for Interactive Organ Segmentation in Two and Three Dimensions: Implementation and Validation*, in *Med Img Comp Comp Assit Interv* (2005), pp. 773–780
- <sup>66</sup> G. Golub and C. Van Loan, *Matrix computations* (Johns Hopkins University Press, 1996)
- <sup>67</sup> P. Nair, *New perspectives on radicular cysts: do they heal?*, *Int Endod J* **31** (1998), pp. 155–60