# RETRIEVING DENTAL RADIOGRAPHS FOR POST-MORTEM IDENTIFICATION

*Ayman Abaza, Arun Ross, Hany Ammar*

West Virginia University
Lane Department of Computer Science and Electrical Engineering
Morgantown, USA, 26506
(ayman.abaza, arun.ross, hany.ammar)@mail.wvu.edu

## ABSTRACT

Automating the process of postmortem identification of deceased individuals based on dental characteristics is receiving increased attention. With the large number of victims encountered in mass disasters (e.g., Asian tsunami), automating the identification process would enhance the scalability of this biometric. However, archiving and retrieving dental records from large databases is a challenging task and has received inadequate attention in the literature. This paper concerns itself with the task of efficient fast retrieving dental records from a database in order to assist the forensic expert in identifying deceased individuals in a rapid manner. The proposed method is an appearance-based technique that consolidates the evidence presented by individual teeth in a dental record, i.e., it 'moves' from tooth-to-tooth in order to render a record-to-record matching score. The proposed method is shown to reduce the searching time of record-to-record matching by a factor of hundred. Experimental results indicate that the proposed approach requires significantly less time compared to the other approaches suggested in the literature thereby underscoring its relevance in real-time applications.

**keywords:** Automated dental identification system (ADIS), Postmortem Identification, Record Retrieving, Eigen-teeth.

## 1. INTRODUCTION

Recent disasters have highlighted the significance of automated dental identification systems. For example, many of the victims of 9/11 were identified based on information contained in pieces of their jaw bones. Forensic experts using dental records identified around 20% of the 973 victims whose identification was determined in the first year after the attack. Identifying all 2,749 victims of the 9/11 disaster [1] took around 40 months [2]. The number of Asian tsunami victims identified during the first 9 months was 2,200 (out of an estimated total of 190,000). Most of these 2,200 victims were identified based on dental information (75% of the victims in Thailand were identified using dental records, compared to just 0.5% using DNA) [3]. Automating the process of dental identification will enhance human identification in such catastrophic events where the use of biometric identifiers such as face and fingerprints may not be possible [4, 5]. Given a post mortem (PM) dental record, the main goal of this paper is to design a search engine that can retrieve potential matches from a large repository of ante mortem (AM) dental records thereby assisting the forensic expert in her task of establishing identity.

Architecturally, ADIS is composed of three main components [6]:

- Dental Record Preprocessing Component, which involves several segmentation and classification steps in order to isolate each tooth and determine their class.

- The Search and Retrieval Component manages the archiving, searching and retrieval of dental records in order to produce a candidate list of matches. The desirable features of this stage are speed and accuracy.

- The Image Comparison Component registers and compares two sets of dental records and is used during the search process. This process can be computationally intensive.
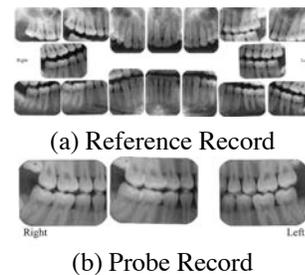

(a) Reference Record

(b) Probe Record

**Fig. 1**. Examples of dental records

In [5] and [7], the search and retrieval component is based on teeth contour alignment and matching which is used to determine the similarity between two dental images. These methods are typically very slow although they can be very useful in the image registration stage (Figure **??**). Both techniques use teeth contour features (shape based); the automatic

extraction of teeth contour is computationally intensive and the output itself may not be precise [8]. Nomir and Abdel-Mottaleb [9] describe a fast retrieval procedure based on teeth contour as one of their basic features. However, their technique can handle only bitewing films, which represents only $\sim 56\%$ of the available database [10].

The main goal of the proposed technique is to reduce the dental record retrieval time while maintaining reasonable accuracy. The proposed approach depends on dimensionality reduction using principle component analysis (PCA). The accuracy of class-based retrieval (where the tooth class is known: incisor, canine, premolar, or molar) versus class-independent retrieval (where the tooth class is unknown) is also studied. In section 2, we briefly summarize the teeth segmentation and normalization steps. In section 3, we describe the proposed technique that is based on Principal Component Analysis (PCA). In section 4, we discuss the experimental results. Finally in section 5, we summarize the contributions of this paper and suggest possible extensions.

## 2. TEETH PRE-PROCESSING

In designing a dental image retrieval system, there are many factors that dictate the choice of features. Computational cost and feature reliability are two key factors that determine our choice of practically feasible features. The automatic extraction of teeth contours is a challenging problem and current solutions can still result in inaccurate outputs [8]. This motivates us to consider alternative reliable appearance-based, low computational-cost features. To achieve this, we adopt the following preprocessing steps:

- **Teeth Segmentation:** We tackle the problem of teeth segmentation using a mathematical morphology approach. This segmentation algorithm has three stages: (i) noise filtering; (ii) thresholding to isolate the teeth from the background; and (iii) analyzing the resulting connected components labeling to determine the ROIs based on geometrical properties [11].

- **View Normalization:** An input tooth $t_q$ is geometrically-normalized. In practice, dental radiographic films show teeth that need not comply with the standard views often assumed in the literature. To perform geometric normalization of the segmented tooth, the most important step is to ensure that its lateral surface appears predominantly vertical [12].

- **Teeth Labeling:** We tackled this problem using a two-stage approach. The first stage utilizes low computational-cost, appearance-based features for assigning an initial class. The second stage applies a string matching technique, based on teeth neighborhood rules, to validate that initial class and hence assign a number corresponding to the tooth location in the dental chart [12].

## 3. EIGEN TEETH

We then proceed to generate the subspace for each of the four classes of teeth: incisors, canines, premolars, and molars. To construct the four subspaces, we resort to the principal component analysis (PCA). Because of some resemblance between the proposed method for teeth reconstruction and the eigenfaces [13] (at least at the abstract level), we refer to the resulting basis vectors defining the image subspaces as eigenteeth. First, representative tooth samples of each of the four classes are selected. Figure 2 depicts a sample of exemplars from each class, where different teeth segments, sizes and intensity levels are noted.
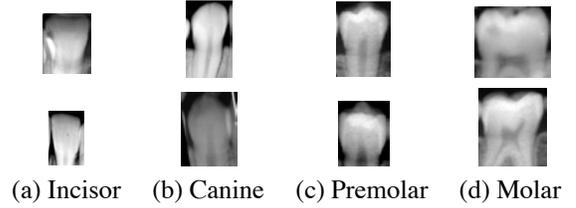


(a) Incisor    (b) Canine    (c) Premolar    (d) Molar

**Fig. 2**. Examples of teeth used in constructing the image subspaces of the four classes of teeth

One approach to deal with the problem of excessive dimensionality is to reduce dimensionality. The PCA approach, as a linear transformation method, seeks linear projection to a lower dimensional subspace that best represents the data in a least-square error sense.

With the sets of exemplars prepared, we proceed with the construction of the subspaces as follows:

- Compute the average of the four sample sets.

- Let $A_k = [(x_{1k} - \mu_k)....(x_{Nk} - \mu_k)]$ denote the matrix of the average-adjusted sample of class k, where $N_k$ is the number of exemplars of class k, and $x_{ik}$ is the vector representation of the $i^{th}$ exemplar in the $k^{th}$ class. Then the eigenvectors $\underline{u_{ik}}$ of the scatter matrices are computed as:

$$A_k^T * A_k * \underline{\nu_{ik}} = \lambda_{ik} * \underline{\nu_{ik}} \tag{1}$$

where $\lambda_{ik}$ is the $i^{th}$ eigenvalue of $A_k^T * A_k$. Finally, $\underline{\nu_{ik}}$ is linearly mapped to $\underline{u_{ik}}$, using $\underline{u_{ik}} = A_k * \underline{\nu_{ik}}$

- Select $r$ eigenvectors for each set whose corresponding eigenvalues are the largest. The eigenvectors are then arranged in a descending order according to their corresponding eigenvalues. Figure 4 shows image representations of the first 15 eigenvectors of each class.

In some cases, the class of the input tooth may not be known. To accommodate this condition, we create a general set of eigenvectors based on exemplars taken from multiple
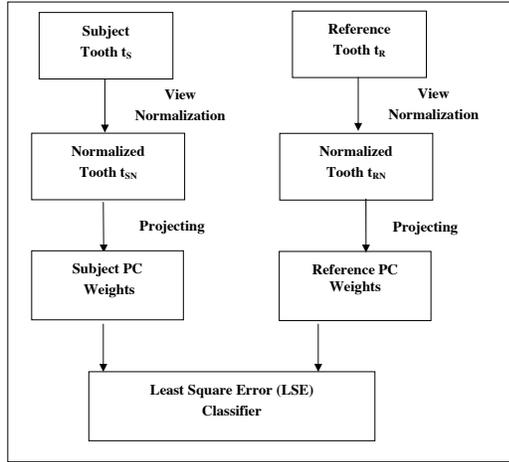
**Fig. 3**. Block diagram of proposed method

classes. The proposed approach for potential match search is illustrated in Figure 3, where the input tooth image and the reference tooth image are first normalized [12]. Here, the input image is projected onto the reduced subspace using the matrix of eigenvectors. The resulting vector is compared against a target vector using the Euclidean distance in order to generate a match score.

In some cases, a tooth may occur in multiple films resulting in multiple distance values corresponding to different representations of the same tooth. In such cases, we choose the minimum of all the distances as the match score. To move from tooth-to-tooth comparison toward record-to-record comparison, we combine the match scores generated by all component teeth via the min-rule or the mean-rule.
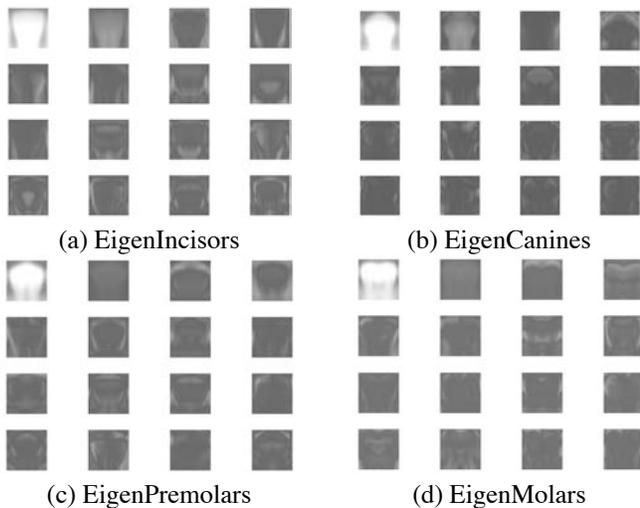


(a) EigenIncisors            (b) EigenCanines

(c) EigenPremolars          (d) EigenMolars

**Fig. 4**. The EigenTeeth representation for the four classes of teeth

## 4. EXPERIMENTAL RESULTS

To setup the image subspaces for the four teeth classes, a data set of exemplars was prepared using the dental image database described in [10]. 100 Molars, 100 Premolars, 30 Canines, and 60 Incisors from records were segmented and preprocessed. To ensure diversity, we used teeth from the upper and lower jaws as well as from the right and left sides of the mouth.

To test the proposed approach for teeth classification and numbering, a test set was constructed; this set was mutually exclusive with respect to the data set of exemplars [10]. The test data-set consisted of 104 records (about 500 bitewing and periapical films) involving more than 2000 teeth. There were 47 Antemortem (AM) records and 57 Postmortem (PM) records with 20 matched records.
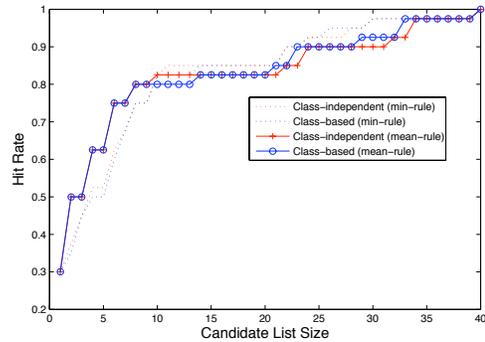


**Fig. 5**. CMC curves corresponding to the four best performing scenarios. See text for explanation

To select the optimal normalized teeth image size, we conducted a preliminary experiment based on the selected training exemplars. We used 220 teeth for training and 130 teeth for testing. We found that an image size of $32x32$ achieves optimal performance compared to, say, $16x16$ or $64x64$.

We conducted a second experiment to select the $r$ significant eigenvectors. We tested the performance using different number of eigenvectors ($15, 20, ... 55$ or $100$ eigenvectors) for both the class-specific and class-independent methods, and using the minimum and mean fusion rules to generate record-to-record matching scores. The resulting Cumulative Match Characteristic (CMC) curves were visually assessed and the top 4 results are shown in Figure 5. For the class-based retrieval method using the minimum fusion rule the value of $r$ was $45$, while for the class-independent method it was 35. For the class-based method using the mean-fusion rule the value of $r$ was 100, while for the class-independent method it was 50.

As can be seen from the curves in figure 5, the mean fusion rule and the minimum fusion rule have comparable performance. Second, the class-specific method needs to maintain more significant number of eigenvectors to achieve the

same performance as the class-independent one. Third, for $50\%$ of the cases the correct match (record) is within the top 3 ranks, and for $75\%$ of the cases it is within the top 6 ranks. On an average, the record to record comparison takes less than 0.1 seconds [1] and the average search time is $\sim 0.1n$ seconds, where $n$ is the number of records in the database. Compared to the image comparison technique [14] which takes about 20 minutes to do a single record-to-record matching, the proposed method significantly reduces the search time while still maintaining a comparable accuracy.

## 5. CONCLUSION AND FUTURE

The presented technique is a fast searching mechanism for an Automated Dental Identification System (ADIS). Initially each segmented tooth is compared to the corresponding reference tooth based on a lower dimensional image subspace established using Principal Component Analysis (PCA). By fusing multiple tooth matching scores, we calculate the similarity between the subject record and all the reference records in the database. Then we retrieve records based on the minimum Euclidean distance.

Testing results suggest that appearance-based features achieve fast and accurate retrieving with an average record-to-record matching of $\sim 0.1$ seconds. The efficiency of this method is based on two primary observations: (i) it does not require any complex feature extraction processes; and (ii) it does not require alignment of teeth images.

We tested and reported the performance of both class-specific and class-independed methods after using the minimum and mean fusion rules for integrating the match scores of multiple teeth available in a dental record. Ideally the performance of the class-specific method should have been better, but it was observed that that both achieve comparable performance because of the nature of the test database and the limited number of training samples. The fact that the class-specific method needs to retain more eigenvectors was expected as it has fewer number of training samples per class.

Our future work to enhance the performance of initial teeth classification includes: (i) adding weights to the score fusion scheme, for example based on teeth class; and (ii) fusion between feature based approach and the proposed appearance based approach.

## 6. REFERENCES

[1] "Forensic identification of 9/11 victims ends, New York, february 23, 2005," .

[2] Anil Jain, Patrick Flynn, and Arun Ross, *Handbook of Biometrics*, Springer, 2007.

[3] "Dental Records Beat DNA in Tsunami IDs, NewScientist.com News Service, 8 September 2005," .

[4] Anil Jain and Hong Chen, "Registration of dental atlas to radiographs for human identification," in *Proc. of SPIE Conference on Biometric Technology for Human Identification II*, Orlando, FL, 2005, pp. 292–298.

[5] Hong Chen and Anil Jain, "Dental biometrics: Alignment and matching of dental radiographs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1319–1326, 2005.

[6] Gamal Fahmy et al., "Toward an automated dental identification system," *Journal of Electronic Imaging*, vol. 14, no. 4, pp. 1 – 13, 2005.

[7] Jindan Zhou and Mohamed Abdel-Mottaleb, "A content-based system for human identification based on bitewing dental X-ray images," *Pattern Recognition*, vol. 38, no. 11, pp. 2132–2142, 2005.

[8] Samir Shah, Ayman Abaza, Arun Ross, and Hany Ammar, "Automatic tooth segmentation using active contour without edges," in *Proc. of Biometric Consortium Conference*, USA, 2006.

[9] Omaima Nomir and Mohamed Abdel-Mottaleb, "Human identification from dental X-ray images based on the shape and appearance of the teeth," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 2, pp. 188–197, 2007.

[10] "CJIS division - ADIS, digitized radiographic images (database)," August 2002.

[11] Eyad HajSaid, DiaaEldin Nassar, Gamal Fahmy, and Hany Ammar, "Teeth segmentation in digitized dental x-ray films using mathematical morphology," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 178– 189, 2006.

[12] DiaaEldin Nassar, Ayman Abaza, Xin Li, and Hany Ammar, "Automatic construction of dental charts for postmortem identification," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 2, pp. 234–246, 2008.

[13] Matthew Turk and Alex Pentland, "Face recognition using eigenfaces," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Maui, Hawaii, 1991, pp. 586–591.

[14] Diaa Eldin M. Nassar, *Automated Dental Identification: A Micro-Macro Decision-Making Approach*, Ph.D. thesis, West Virginia University, 2005.

---

[1] We carried out these tests in MATLAB running on a 2 GB RAM, 2.4 GHz Intel P4 computer