

Automated tracking of the mandibular canal in CBCT images using matching and multiple hypotheses methods

Brecht Moris
and Luc Claesen
EDM - tUL - iBBT
Hasselt University
Wetenschapspark 2,
3590 Diepenbeek Belgium

Yi Sun
and Constantinus Politis
Hospital East-Limburg ZOL
Campus Sint-Jan
Schiepse Bos 6,
3600 Genk Belgium

Abstract—A method for the automated tracking of the mandibular canal in Cone Beam Computed Tomography images is presented. In contrast to most proposed methods a fully automated method is presented including the detection of the mental and mandibular foramen. Extraction of these points is done by searching for high-contrast gaps in the mandible. The path of the mandibular canal can then be tracked in between these two points using a matching technique consisting of a template function which maps spatial coordinates to a cost. The robustness of the tracking is increased by recursively tracking multiple hypotheses.

I. INTRODUCTION

In the field of orthodontics and orthognathic surgery there are a lot of applications for which there are risks of damage and/or increased postoperative sensitivity or other postoperative sensory impairments to the inferior alveolar nerve (IAN). These applications cover a broad range of operations, from the removal of cysts up to osteotomies and dental implantology. To reduce these risks during an intervention, the surgeon should have knowledge of the exact course of the IAN. The IAN is situated in the mandibular canal, a tubular structure in the mandible which is detectable through computed tomography techniques. In figure 1 the canal is indicated by the red 'wire' on both sides of the mandible. The endpoints of this wire are the mental and the mandibular foramen which are located respectively on the outside of the mandible close to the chin and on the inside of the mandible behind the molars.

For health reasons Cone-Beam Computed Tomography (CBCT) is used instead of traditional CT in order to reduce the amount of radiation the patient is exposed to. However, the main disadvantage of CBCT is the low image quality as a direct result of the reduced amount of radiation that is used [1], [2]. The images that are produced contain a lot of noise and have a low contrast, especially in the canal region. This imposes an additional degree of difficulty in segmenting the mandibular canal in CBCT image compared to CT images. Because of the unpredictable nature of the mandibular canal it is often difficult to automatically extract its correct position in CT images. This is even more difficult to do in CBCT

images due to their low contrast. The data sets that are used are produced by a Galileos CBCT scanner from Sirona Dental Systems and consist of 512 voxels of $0.289 \times 0.289 \times 0.289$ mm in all three dimensions. The scans are made with a setting of 85kV and 42mAs. The scan time is fourteen seconds and the radiation dose is about $124.5 \mu\text{Sv}$

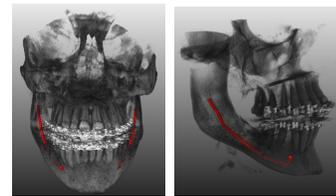


Figure 1: A 3D visualisation of a CBCT scan containing the mandibular canal (Indicated in red). Left image: data set viewed from a frontal perspective; Right Image: viewed from a the left side.

In most existing methods the extraction of the mental and mandibular foramen is done manually. In [3] a method is described to locate the mandibular foramen based on its position relative to landmarks. Kim et al. [7] use a texture analysis technique in 3D panoramic volume rendered images to extract the mental and the mandibular foramen. Soththivirat et al. [4] discusses a method to extract the mandibular canal in panoramic CT images based on a combination of contrast enhancement, thresholding, followed by morphological open and close operations. However, the low contrast in CBCT images makes it difficult to do this without excluding the canal or including too much noise in the process. A more robust method is presented by Friman et al. in [5] that uses a vessel template function in combination with a multiple hypotheses tracking framework for the segmentation of small 3D vessels. Kainmueller et al. [6] proposes a method to extract the mandibular canal from CBCT images that is based on a combined statistical shape model of the nerve and the bone and a Dijkstra-based optimization procedure. Kroon [6] uses an active appearance model in combination with weighted

intensities in order to increase the influence of the voxels representing the canal and decreasing the influence of other voxels. The results of these methods are shown in table I

Method	Dm (mm)	Dstd (mm)
Left nerve of Kainmueller et al. [6]	1,00	0,60
Right nerve of Kainmueller et al. [6]	1,20	0,90
Kim et al. (with speedterm) [7]	0,73	0,69
Kim et al. (without speedterm) [7]	2,41	2,82
Kroon et al. [10]	1,88	1,05
Own method, 1st quarter from MeF	10,54	24,84
Own method, 2th quarter from MeF	3,23	5,98
Own method, 3th quarter from MeF	1,64	1,65
Own method, 4th quarter from MeF	0,99	1,13

Table I: Comparison of the accuracy of the extracted nerves of different methods. Note: The own method depicts the accuracy of four equal parts of the nerve starting from the mental foramen.

In this paper, an automatic method is proposed to extract the mandibular canal in CBCT images including the detection of the mental and mandibular foramen, through which the canal enters the mandible. A method is discussed to extract the foramina based on the occurrence of visual characteristics in several subsequent 2D slices. The three dimensional dataset is considered from three different perspectives to produce a collection of two dimensional images called slices. These perspectives include a frontal, a sideways and a downward view. The extraction of the canal is done by searching several appropriate candidate positions in subsequent 2D slices between the mental and mandibular foramen. To determine appropriate candidate positions, a matching method is used to evaluate several positions within a search region. A cost is produced based on the similarities between the image at a specific position and an ideal model of the canal. To improve the robustness, multiple candidates are tracked recursively using a search tree. Additional restrictions, based on anatomical properties of the canal and its detected movement, are used to create a movement model that will help limit the number of considered positions.

First, an overview of the different steps of the processing pipeline is given in section I-A. In section I-B “Initialization”, the method that is used to extract the mental and the mandibular foramen is discussed. Next, section I-C “Tracking of the canal” explains the actual tracking of the canal between these two points in two directions. Both sections provide a detailed description of the preprocessing operations that are performed on the image. When both paths are obtained, they are combined into a single path representing the mandibular canal. This is explained in section I-D “Postprocessing”. Finally, the results of 25 nerves of 15 patients are compared to a gold standard path, which is manually determined. 13 female patients and 2 male patients are considered. The age ranges from 13 years to 48 years with an average age of 22 years. The average weight and height of the patients is 56kg and 173cm respectively. Some nerves have been excluded from the results

because they could not be located in a reliable way due to bad visibility.

A. Methods

The workflow is divided into three main operations. First, an initialization is performed by determining the position of the mental and mandibular foramen. Once the bounds of the canal are known, the images in between are processed in order to extract the canal. To improve the correctness, this is done in two different directions starting from both the mental and the mandibular foramen. Two partial paths of the canal are extracted this way, which are ultimately combined in a postprocessing step.

B. Initialization

In order to make the method fully automatic, the endpoints of the canal, mental and mandibular foramen, are searched before extracting the canal. This is done by observing a specific region of interest and check for certain image characteristics which indicate the presence of an endpoint.

1) Finding the Mental Foramen:

a) *Preprocessing:* The goal of the preprocessing is twofold. First, the size of the dataset is reduced to a minimum in all three dimensions in order to reduce the amount of redundant data that needs to be processed. When only the essential data remains, the images are enhanced to improve the visibility of the canal and reduce the amount of noise and unimportant features.

Data reduction: The region of interest containing the mental foramen is defined by a bounding box with a fixed size for every patient. Because no further information about the position of the mental foramen is known, a fixed region is chosen wherein it will always be included.

Image enhancement : Before the mental foramen is searched, the image is filtered to enhance certain image characteristics to reduce the amount of noise and make the mental foramen more visible. First, interval thresholding is used to reduce noise and other unimportant elements in the image. To do this, the intensities in the image are categorized into three categories. The first category consists of the intensities in which the mental foramen and the mandibular canal is perceived, the intensities of interest.

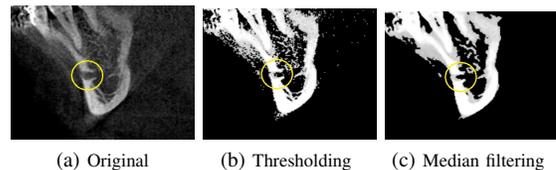


Figure 2: Effect of image enhancements on an image of the mental foramen viewed from the side. This point is characterized by a discontinuity of the bone on the left side.

In this method, the intensities of interest range from 1450 to 2000 gray values (GV). The mandibular canal is represented by intensities between 1450 GV and 1900 GV. Because the canal is often situated against the inner border of the

mandibular canal, intensities up to 2000 GV are considered. All intensities below this range do not represent the canal and are considered noise and other unimportant elements and are made zero (black). Finally, intensities that are above this range are clipped to the highest intensity in this range. Figure 4b shows the result of the interval thresholding on an image showing the mental foramen. Remaining noise is then removed by using a median filter, resulting in figure 4d.

b) Processing: When the images are preprocessed, the position of the mental foramen is searched. To do this, we consider the data from the side, in the sagittal view. In this view, the mental foramen is perceived as a discontinuity of the left border of the mandible, as shown in figure 3. In order to find the position of the mental foramen, the left border is evaluated from the bottom up until a certain height to check for such a discontinuity.

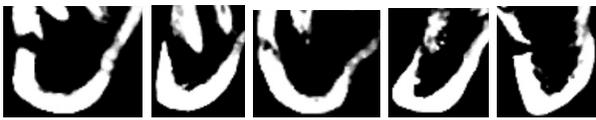


Figure 3: Possible appearances of the mental foramen viewed from the right side. A discontinuity of the bone in the left border of the mandible in the image is visible in all cases.

This gap in the border will be present in several subsequent slices. When such a gap is found in several subsequent slices at more or less the same position on the mandible, it is considered the mental foramen.

2) *Finding the Mandibular Foramen:* Next, the mandibular foramen is searched, which is situated on the inside of the mandible. The orientation of the mandibular canal is about vertical in the mandibular foramen and can, for that reason, be seen more easily from above, in the transversal view.

a) Preprocessing:

Data reduction: In contrast to the previous method, the position of the mental foramen can be used to help make a more informed decision in defining the three dimensional region of interest containing the mandibular foramen. By assuming a minimum and maximum distance between the mental and mandibular foramen in several patients, a range of slices of interest can be composed by selecting all slices between this minimum and maximum distance from the mental foramen.

Image enhancement: Again, the images are enhanced before they are processed. The method that is used to find the mandibular foramen does not utilize the exact value of the intensities in the voxels but merely the presence of the canal or bone. For this reason it suffices to have a binary image with only black and white, which is generated by a thresholding function that will set all intensities below a threshold to zero (black) and above this threshold to the maximum intensity (white). In this method a gray value of 1500 is used as threshold, in order to include all intensities that represent the mandibular canal. Because the mandibular is more prominent in the transversal view, the minimum intensity representing

the canal is increased 50GV in comparison to the range of intensities of interest in the extraction of the MeF. Doing this, enables us to remove more noise from the image. The resulting image is illustrated by figure 4b.

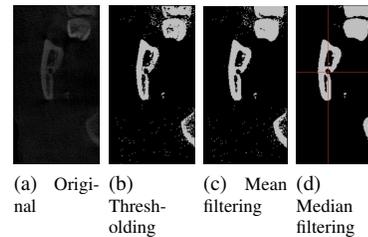


Figure 4: Effect of image enhancing on an image of the mandibular foramen viewed from above. This point is characterized by an inward notch in the right side of the mandible as seen in the images.

In order to smooth borders and fill small gaps in the mandible, a three dimensional mean filter is applied to the black voxels in each image (Figure 4c). Finally, remaining noise is reduced with a median filtering (Figure 4d).

b) Processing: Because the IAN runs through the bone of the mandible this is, again, characterized by a discontinuity of the border of the mandible. In the transversal view this is perceived as an inward notch (Figure 4d). Because this is always the case in the mandibular foramen, it is easy to distinguish from the rest of the mandible. In order to find this notch an approach similar to finding the mental foramen is taken. In every subsequent slice the position of the right border of the mandible is tracked from the lowest point up until a certain height. After this is done in a slice, the positions are evaluated to check for a notch.

A notch is characterized by an inward movement of the border followed by an outward movement of the border. An additional condition is imposed to increase robustness. A minimum and maximum height difference between the inward and the outward movement is used to eliminate false positives. In a typical patient, the inward notch resulting from the mandibular foramen is the only notch that will satisfy the imposed condition. However when the density of the bone is too low, additional gaps are introduced by the thresholding function. To prevent these gaps from being selected, a notch has to be present in several subsequent slices on a similar position in the mandible.

C. Tracking of the mandibular canal

To ensure correctness, the canal is tracked from both the foramina. The reason for this approach is due to the difference in orientation of the canal at the foramina. A static view is used in each approach as opposed to dynamically generating images from the dataset along the current path. Because a search tree is built to support multiple hypotheses, a static view is used to limit the computational complexity.

Often, the resulting paths will merely be partially correct because of the reduced visibility of the canal due to the changing orientation. When both paths are generated, they are

combined based on an overlap region. It suffices if only one of both paths in the overlapping region are correct or if both are merely partially correct.

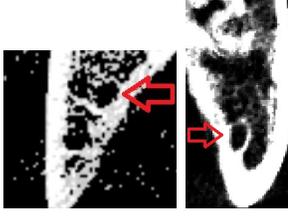


Figure 5: Different views used to track the canal. Left image: example of a the perception of the canal close to the mandibular foramen viewed from above. Right image: example of an image containing the canal close to the mental foramen seen from a frontal view. In both cases the canal is perceived as a dark circular structure surrounded by a band of light voxels.

1) Tracking a downward path:

a) Preprocessing:

Data reduction: It is clear that we only need data in between the foramina, hence a 3D bounding box can be constructed from these two points to serve as region of interest. Because the movement of the canal between two subsequent slices is limited to a few voxels, it suffices to search only inside a circular search region around the previously found position. However because these position are not yet known this can only be done at runtime.

Image enhancement: To prepare the images for the tracking of the mandibular canal an interval thresholding function is used first. This is done similar to the interval thresholding that was discussed earlier in section I-B1a. Intensities that are considered noise and/or background elements are made black. Intensities that are too bright, such as metal artifacts and bone, are clipped to a maximum intensity. The remaining intensities are left as is. Finally, a morphological close operation is used to bridge small gaps and fill small holes. The result of this operation can be seen in figure 6c.

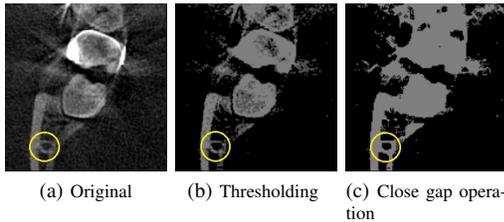


Figure 6: Effect of image enhancing on an image of the canal close to the mandibular foramen viewed from above. The circular structure is easy to perceive.

b) Processing:

Inner and outer region matching: After the preprocessing, the data set is suited for the actual processing. The basic principle for tracking the mandibular canal consists of a twofold matching function that compares the image at a given point in the image with an idealized model of the mandibular canal. The idealized model is defined as a dark, preferably black, circular structure surrounded by a brighter band. To

account for canals of different sizes and changing sizes, the idealized model is used with several different diameters ranging from seven to thirteen voxels, including the surrounding band.

In order to select a position a cost is calculated by the matching function for each position to express its appropriateness:

$$C(x, y, r) : \mathbb{R} \mapsto [0, 6000]$$

With r being the radius of the idealized model that is used. Mainly this is done by verifying the presence of the darker circular structure and the surrounding bright band respectively with the *inner region* and *outer region* matching methods. Basically these methods calculate the average intensity of the voxels inside the considered region and convert it to a cost as follows:

$$\begin{aligned} inner &= \text{Average intensity Inner region} \\ outer &= (\text{Max intensity} - \text{Average intensity Outer region}) \end{aligned}$$

Because a higher average intensity in the outer region should produce a lower cost, as it is more desired, it is inverted by the maximum intensity. Often, the outer region is only partially present as opposed to the inner region it surrounds. In other words, the average intensity of the outer region is allowed to be lower than assumed by the idealized model, but the inner region must always be relatively dark. More weight is therefore given to the inner region matching.

$$C(x, y, r) = 2 * inner(x, y, r) + outer(x, y, r)$$

This function enables us to find the correct position in most cases, however many situations exist in which this method is misled. Often the canal is situated against the bone of the mandible or its shape is far from circular. Figure 7 illustrates several such appearances of the canal. When the canal is partially embedded in the bone of the mandible, the calculation can be misled by the relatively high intensities of the bone, which will result in false positives. Additional factors are used in order to account for the shape of the canal.



Figure 7: Several possible perceptions of the mandibular canal. It is clear that the canal is not always visible as described by the idealized model.

Additional factors: The first, and most important, factor is a measure of how much voxels within the outer region are non-zero. In other words, this factor expresses the percentage of the outer region that is present without taking the actual intensities into account. This gives:

$$c(x, y, r) : \mathbb{R} \mapsto [0, 1]$$

However, this factor only expresses the amount of the border that is present in the image, but not how this is divided among the entire outer region. For example, a completeness factor of 0.5 can signify an outer region where the right half is

completely filled and where the left side is empty. However when the same amount of non-zero voxels is divided among the entire outer region the same factor is obtained. A second factor is therefore used to express the division of these voxels within the outer region.

1	1	1	2	2	3	3	3
1	1	1			3	3	3
1	1					3	3
8							4
8							4
8							4
7	7					5	5
7	7	7			5	5	5
7	7	7	6	6	5	5	5

Figure 8: Division kernel: The non-zero elements of the outer region kernel are divided into eight regions.

To obtain this factor, the outer region is divided in eight regions as shown in figure 8. When at least one non-zero voxel is present in a region it is considered to be 'filled'. For every opposing pair of filled regions an additional region is assumed to be filled. The number of filled regions is then divided by the maximum number of filled regions to obtain the factor.

$$d(x, y, r) : \mathbb{R} \mapsto [0.8, 1.0]$$

The impact of this factor is reduced to prevent it from introducing false results. Because both factors are inversely proportional to the cost, they are first inverted by subtracting them from one, after which they are integrated in the cost function as follows:

$$C(x, y, r) = 2 * inner(x, y, r) + outer(x, y, r) * (1 - c(x, y, r)) * (1 - d(x, y, r))$$

Multiple hypotheses: This matching method will be able to find the correct position even if the visibility of the canal is very low, however sometimes it is difficult or impossible to find the canal, even with the naked eye. When the canal 'disappears', a wrong position can be selected. When this happens, the canal can be outside the search region and the algorithm might not be able to recover. To prevent this, multiple positions are recursively tracked to increase robustness. A search tree is composed in which each position is stored along with its cost. The search tree is pruned regularly based on the total cost of each path to limit the complexity.

Search region: The risk of selecting an incorrect position can be reduced further by imposing additional restrictions on the search region based on anatomical properties of the mandibular canal, which runs down from the mandibular foramen towards the mental foramen at the front of the face. Because the canal will never move towards the back of the face, the vertical movement is limited to either no movement or an upward movement when viewed from above. This is illustrated in figure 9.



Figure 9: The course of the canal close to the mandibular foramen can only move up when viewed from above.

Also, the orientation of the canal will change gradually. This allows us to predict its movement between two slices based on the average movement of several previous slices. We calculated the average movement based on ten previous slices. Using less or more slices will make the average movement respectively too sensitive or not sensitive enough for a sudden change in movement. The search region can then be centered around this prediction.

2) *Tracking an upward path:* The same method is used to determine an upward path from the mental foramen. Because the canal runs more or less horizontally at this point, the data set is considered from a frontal view.

a) *Preprocessing:* Because, again, the slices of interest are contained within the mental and the mandibular foramen, data reduction is done similar to the previously described method. Also the region of interest in each slice is based on the search region and depends on the positions that are found in the previous slice. Enhancement of the image uses an interval thresholding function and a mean filter which is applied on black voxels.

b) *Processing:* To track an upward path, a similar method is used to the tracking of a downward path. The canal is often embedded against the inside of the border of the mandible, which consists of high intensity voxels. These bright voxels can corrupt the calculations of the outer region matching, causing an incorrect position to be selected against the border. To counteract this, an additional border threshold is introduced to the calculations of the outer region matching to exclude these high border intensities. We use a gray value of 1900 for this value.

D. Postprocessing

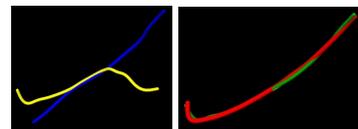


Figure 10: Left image: both upward path (yellow) and downward path (blue) are shown separately before the merging; Right image: the manually set nerve (red) and the merged path (green)

Finally, both the upward and the downward path are combined together to produce a single correct path. The leftmost image in figure 10 depicts a typical situation. Both the upward (yellow) and the downward (blue) path are correct and close to their starting points. However at a certain distance from that point the orientation of the canal changes. This causes the algorithm to select incorrect positions. When combining both

parts, we can assume that the first part of each path is correct. The remaining parts of the paths can be divided into chunks of a certain size, we consider five slices at the time, which can be compared to each other based on the average cost of the positions in that chunk. Also, the average movement in each chunk is considered to eliminate incorrect chunks, such as the part of the upward path that moves down. The result is shown in the rightmost image of figure 10 in green. The red nerve is the gold standard.

E. Results

Path	Dm (mm)	Dstd (mm)
1st quarter DP	1,02630	1,01706
2th quarter DP	1,95094	2,12987
3th quarter DP	6,06833	7,40743
4th quarter DP	14,84893	12,78999
Full DP path	6,10986	5,30818
Full merged path	4,25545	4,02101

Table II: Results from a comparison of the computed path with the manually selected gold standard (GST). Dm indicates the mean distance between the outer surface of the GST and the computed positions. DP indicates the downward path.

The methods described in this paper have been implemented in the MeVisLab software package. To test the method, 25 nerves were used from 15 different patients. Some nerves were excluded from the results because it was impossible to manually determine a gold standard due to a lack of visibility. The mental foramen is successfully obtained in 96% of all cases and the mandibular foramen was correct in all cases. The results of the calculated path are depicted in table II. Because no volume is calculated around each point, an average radius is obtained based on the size of outer region kernels that are selected for each position. The average radius of all tested nerves is 1,01568 mm. *Dm* resembles the mean distance between the gold standard and the outer surface of the calculated surface. *Dstd* is the standard deviation. The upper four rows represent the four quarters of the downward path. It is obvious that the downward path is very reliable close to the mandibular foramen. It is clear in figure 11 that the correctness of the downward path depends on the orientation of the canal.

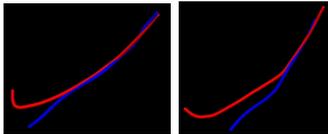


Figure 11: When the orientation of the nerve (in red) starts to go more horizontally, the canal, viewed from above, is not perceived as a circular structure anymore. As a result, the downward path (in blue) will not be able to locate the canal from a certain distance from the mandibular foramen.

II. CONCLUSION

This paper presented a fully automatic method to extract the mandibular canal in CBCT images, unlike most of the

existing methods no user input is required. Because the both the mandibular foramen and the mental foramen are a discontinuity of the relatively dense mandible, it is very easily extracted. However, when one of these reference points are extracted incorrect, the subsequent steps will have a probability to fail, depending of the degree of the error. Segmentation of the mandible is not necessary because the course of the mandibular is contained between the mental and mandibular foramen and its movement throughout the mandible is restricted within a dynamic search region. This search region is composed based on anatomical properties of the nerve and on a movement model that is composed from its previous movement. Because the contrast in CBCT images is very low, it is often difficult to distinguish the canal from noise or artifacts. By using multiple hypotheses robustness is increased significantly, allowing the visibility to have local minima without generating incorrect results.

However, the low contrast makes it difficult to distinguish the canal from other elements in the image. As a result, thresholding can exclude part of the canal or include too much noise which often makes it difficult to determine the correct position. Future work includes the integration of a dynamic thresholding function that computes a suitable threshold for each image. Also, because the canal is often not perceived as a circle and changes throughout its course, a dynamically adjustable idealized model is needed that adjusts itself based on its current shape to further increase the robustness.

ACKNOWLEDGMENT

The authors would like to thank the department of dentistry of the St. John's Hospital in Genk for providing the CBCT data. Also, I would like to thank MeVis Medical Solutions for providing me with a free license of MeVisLab.

REFERENCES

- [1] W. De Vos, J. Casselman and G. R. J. 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: 609-625, 2009.
- [2] W. C. Scarfe, A. G. Farman, and P. Sukovic. Clinical Applications of Cone-Beam Computed Tomography in Dental Practice. *Journal of the Canadian Dental Association*, 72(1):75-80, 2006.
- [3] O. Trost, V. Salignon, N. Cheynel, G. Malka and P. Trouilloud, A simple method to locate mandibular foramen: preliminary radiological study. *Surg Radiol Anat*, 32:927-931, 2010.
- [4] S. Sotthivirat and W. Narkbuakaew, Automatic Detection of Inferior Alveolar Nerve Canals on CT Images. *Biomedical Circuits and Systems Conference*, Nov 2006 pp. 142-145.
- [5] O. Friman, M. Hindennach and H.-O. Peitgen, Template-based Multiple Hypotheses Tracking of Small Vessels. In *Proceedings of the 5th IEEE International Symposium on Biomedical Imaging (ISBI'08)*, 2008.
- [6] D. Kainmueller, H. Lamecker, H. Seim, M. Zinser and S. Zachow, Automatic Extraction of Mandibular Nerve and Bone from Cone-Beam CT Data. In *Medical Image Computing and Computer-assisted Intervention (MICCAI'09)*, 2009.
- [7] G. Kim, J. Lee, H. Lee, J. Seo, Y.-M. Koo, Y.-G. Shin and B. Kim, Automatic Extraction of Inferior Alveolar Nerve Canal Using Feature-Enhancing Panoramic Volume Rendering. *IEEE Transactions on Biomedical Engineering*, Vol. 58, No. 2, February 2011.
- [8] D. Kroon, Segmentation of the mandibular canal in cone-beam CT data. *Thesis*, 2011.