Automatic Coronary Calcium Scoring Using Native and Contrasted CT Acquisitions

B. Michael Kelm¹ and Yefeng Zheng²

¹ Imaging and Computer Vision, Corporate Technology, Siemens AG, Erlangen, Germany
michael.kelm@siemens.com

² Imaging and Computer Vision, Corporate Technology, Siemens Corporation, Princeton, USA
yefeng.zheng@siemens.com

Abstract. The evaluation of coronary calcium is a standard procedure in clinical workflows for the diagnosis of patients presenting with suspected coronary artery disease. Based on a native cardiac CT scan, the coronary calcium present in each of the three main coronary arteries is determined. With most clinically available tools, the physician has to manually assign calcium clusters (voxels above 130 HU) to the coronary arteries in order to distinguish coronary calcium from bone, non-coronary calcium and noise. Here we propose an automatic approach for the assignment of coronary calcium to the coronary arteries which exploits the fact that often, native cardiac CT scans are acquired along with (contrasted) CT angiography scans in which the coronary vessels are clearly visible. Using previously developed segmentation approaches, automatically traced centerlines of the coronary vessel tree are mapped from the contrasted to the native scan. In combination with a machine learning classifier, this patient specific modeling allows for an accurate estimation of the patient’s coronary calcium burden. The evaluation on the multi-vendor CT data provided by the Coronary Calcium Scoring Challenge shows promising results.

1 Introduction

Patients with coronary artery disease (CAD) suffer from luminal narrowings of the coronary arteries which may lead to ischemic conditions in the heart muscle and cause myocardial infarction. CAD is indicated by the build-up of coronary plaques at the vessel wall that often accumulate calcium in particular as the disease progresses.

Using a native Computed Tomography (CT) scan of the heart, i.e. a scan without the administration of a contrast medium, calcified plaques in the coronary arteries can be detected with high sensitivity, segmented and quantified. From these, coronary calcium burden is estimated, e.g. in terms of the Agatston score [2], and used for patient risk stratification. Hence, current clinical CT protocols for CAD require the acquisition of a native CT scan prior to the acquisition of a contrasted scan which is used for a detailed analysis of the coronaries.

Potential coronary calcifications are easily detected in native CT scans as connected components of voxels above 130 HU. But often, calcium is also found at the heart valves and in the aorta; if close to the ostia it remains unclear whether to assign it to a coronary artery or discard it as non-coronary calcium. Adjacent bones such as sternum, ribs and
vertebrae also exhibit radiodensities above 130 HU and need to be excluded. Also, clusters of voxels that randomly exceed the calcium threshold due to image noise should not be counted towards the calcium score.

Most clinical tools for coronary calcium scoring are semi-automatic. The physician is left with distinguishing coronary from non-coronary calcium, e.g. by assigning the found candidate calcifications to one of three coronary arteries: right coronary artery (RCA), left circumflex (LCX) and left anterior descending (LAD) – coronary calcium found in the left main (LM) artery, i.e. the common part of LCX and LAD, is commonly assigned to the LAD. In light of the increasing clinical workload due to the availability of novel diagnostic tools (e.g. CT angiography and CT-based fractional flow reserve), it is desirable to automate the workflow for coronary calcium scoring as much as possible.

The automatic identification of coronary calcifications in native CT scans has also been subject of previous work. Wu et al. [7] presented an opt-out segmentation based method to exclude non-coronary high-density structures. The ascending and descending aorta are detected on each axial slice by exploiting the circular shape of the aorta. Afterwards, the calcifications inside the aorta can be excluded. The sternum and vertebrae are segmented to further reduce false positives. However, this method tends to over-estimate the calcium burden of a patient since the calcifications on the cardiac valves and pericardium are still included. Discriminative learning is widely used in the previous work to distinguish coronary calcium from others [5, 2, 3]. A classifier can be trained with various features, including size, shape, appearance, and spatial features. Although simple size or shape features may be sufficient to exclude bony structures, spatial features are important to distinguish calcifications of coronary arteries from those originated from the aorta, cardiac valves, or pericardium. To extract spatial features, the heart (sometimes, the aorta as well) needs to be segmented [5, 3]. In [2], explicit heart segmentation is avoided by exploiting a probabilistic coronary calcium atlas generated from a training set. During detection, the calcium atlas is registered to the input volume and the spatial features are then extracted based on the distance to the calcium probability map.

Another line of research has been on calcium scoring in contrasted CT. If reliable enough, this would allow to skip the native scan in cardiac CT protocols and thus save a significant amount of radiation dose. Due to the variations in the use of contrast agent and partial volume effect, the major difficulty in contrasted CT is to determine an optimal threshold for calcification detection. In [6], the coronary artery tree is automatically extracted and an optimal threshold is determined based on intensity histogram analysis. Since the calcium detection is constrained to the extracted coronary tree, this approach tends to under-estimate the calcium burden if an artery with calcification is missed by the automatic tree tracing algorithm.

In this work, we leverage information from both contrasted and native CT scans. The coronary tree is automatically extracted using a robust centerline tracing algorithm [9]. The extracted coronary tree is then mapped to the native scan using cardiac structures (i.e., the pericardium and aortic root) segmented from both contrasted and native scans. Exploiting the coronary tree extracted from a contrasted scan has two major advantages. 1) It helps to prune false positives (i.e., based on the distance to the coronary tree). 2) It provides a coronary branch label to the detected calcification. Besides the
overall calcium burden, vessel-specified calcium burden is often measured in clinical practice. Since the extracted coronary tree is properly labeled using our model-driven approach [9], it is straightforward to assign a branch label to the detected calcification based on nearest neighbor search. Such anatomical information is also beneficial to generate a more specific diagnosis report (e.g., heavy calcium on the proximal LAD). Unfortunately, the anatomical labeling of the detected calcification is often ignored in previous work.

2 Methods

Our automatic approach can be subdivided into three phases. In the first phase anatomical information is extracted from both, the contrasted and the native scans. While segmentations of the pericardium and the aortic root are obtained for the contrasted and native scans independently, centerlines of the coronary vessels are traced in the contrasted scan only and subsequently mapped to the native scan. In the second phase candidate calcifications are extracted and selected by simple rules that preserve sensitivity on the training set. In the final phase, a random forest classifier is used to distinguish between non-coronary structures and true coronary calcifications among the candidates. Details are provided in the following.

2.1 Cardiac Anatomy Segmentation

![Fig. 1. Native (left) and contrasted (right) CT scans of a patient from the training set (TRV1P7) with calcifications in the left coronary arteries; here the left anterior descending (green), the first diagonal (cyan) and the left circumflex (yellow) arteries are shown. The coronaries are automatically traced in the contrasted scan (right) and, by means of the segmented pericardium and aortic root, mapped to the native scan (left).](image)

In the first step, the pericardium and the aortic root are segmented in both the contrasted and the native scans using marginal space learning (MSL) as introduced in [11]. MSL is used to estimate the position, orientation, and size of the heart. A mean shape
(which is trained on a set of example shapes) is then aligned with the estimated pose as an initial estimate of the pericardium boundary. Learning-based boundary detectors are used to guide the boundary evolution under the active shape model (ASM) framework. However, the sternum and ribs are often included in the automatically segmented pericardium mesh if we apply MSL directly. As a post-processing step, we explicitly segment these bones to avoid their interference with calcium scoring [10]. Similarly, the aortic root is segmented with MSL and learning based boundary detectors using the segmented pericardium for a prior on the search location.

Centerlines for the coronary artery tree are automatically extracted from the contrasted scans using the approach proposed in [9]. A learning based verification step as described in [4] corrects for vessel tree parts erroneously traced into non-vessel structures. The coronary centerline tree is then mapped from the contrasted scan to the native scan based on the segmented pericardium meshes and the overlapping parts of the aortic root (starting from the hinges). Point correspondence between the segmentation meshes of the native and contrasted scan can be assumed since a model-based approach (ASM) has been used. To this end we used a thin-plate-spline (TPS) model like in [8] to interpolate the deformation field within the pericardium. While other point-based registration methods (rigid as well as deformable) could be applied as well, TPS has the advantages that 1) the interpolation is smooth with derivatives of any order, that 2) the model has no free parameters that need manual tuning and that 3) it has closed-form solutions for both warping and parameter estimation.

2.2 Candidate Extraction

In the native scan, candidate calcified plaques (CCPs) are extracted as connected groups of voxels (3D 6-connectivity) with intensities above $t_{\text{cal}} = 130$ HU. Additional constraints are imposed to reduce the number of (false positive) candidates which, given the small amount of training data, would harm classifier training. Thus, a CCP is discarded if

1. all voxels are outside the pericardium
2. the number of voxels is less than $n = 3$
3. the mean distance of its voxels to the coronaries is more than $d_v = 15$ mm
4. all voxels are within the aortic root
5. the mean intensity difference to the surrounding tissue is less than $D_{\text{max}} = 110$ HU

where the intensity of the surrounding tissue is estimated as the mean intensity from a box around the CCP. Finally, CCPs are labeled as belonging to the type of vessel (LAD, LCX, RCA – including side branches) to which most of its voxels are closest.

2.3 Candidate Classification

For the final phase, a random forest classifier [1] is trained to distinguish between true positive and false positive CCPs. To this end, a total of 36 appearance and geometric features are extracted for each CCP which are listed in Table 1. The random forest has been trained using 100 trees and 6 randomly selected features in each split. For evaluation, random forests have been trained in a leave-one-patient-out (LOPO) fashion.
Table 1. Features extracted for each candidate calcified plaque (CCP).

<table>
<thead>
<tr>
<th>name</th>
<th>parameter range</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>size</td>
<td>–</td>
<td>number of voxels</td>
</tr>
<tr>
<td>meangray</td>
<td>–</td>
<td>mean intensity</td>
</tr>
<tr>
<td>maxgray</td>
<td>–</td>
<td>maximum intensity</td>
</tr>
<tr>
<td>ccnt(&lt;n&gt;)</td>
<td>(&lt;n&gt; = 134 : 4 : 166)</td>
<td>number of voxels above (&lt;n&gt;) Hounsfield units</td>
</tr>
<tr>
<td>chist(&lt;n&gt;)</td>
<td>(&lt;n&gt; = 134 : 4 : 166)</td>
<td>(ccnt(&lt;n&gt;)) normalized by the number of voxels</td>
</tr>
<tr>
<td>lmeangray</td>
<td>–</td>
<td>mean intensity of voxels in local neighborhood</td>
</tr>
<tr>
<td>lmaxgray</td>
<td>–</td>
<td>maximum intensity of voxels in local neighborhood</td>
</tr>
<tr>
<td>lstdgray</td>
<td>–</td>
<td>intensity standard deviation in local neighborhood</td>
</tr>
<tr>
<td>ldgray</td>
<td>–</td>
<td>(meangray - lmeangray)</td>
</tr>
<tr>
<td>lddgray</td>
<td>–</td>
<td>(ldgray - lstdgray)</td>
</tr>
<tr>
<td>vdist(&lt;t&gt;)</td>
<td>(&lt;t&gt; \in {\ \text{mean, std, max, median, mad}})</td>
<td>statistics of the voxel distances to the closest centerline point</td>
</tr>
<tr>
<td>cog(&lt;t&gt;)</td>
<td>(&lt;t&gt; \in {x, y, z})</td>
<td>center of gravity relative to the pericardium</td>
</tr>
<tr>
<td>cog(&lt;t&gt;)</td>
<td>(&lt;t&gt; \in {\sin, \cos})</td>
<td>center of gravity in cylindrical coordinates</td>
</tr>
</tbody>
</table>

3 Results and Discussion

Native as well as contrasted scans from 64 patients have been provided by the organizers of the challenge (see http://orcascore.isi.uu.nl/). These were equally divided into a training (32) and a testing (32) set, each of which contained respectively 8 patients acquired on CT scanners from 4 different vendors. For the training set, ground truth annotations were also provided.

For evaluation, sensitivity and positive predictive value (PPV) have been computed on a voxel-wise basis (“Volume”) as well as on a lesion-wise basis (“Lesions”). Furthermore, the intraclass correlation coefficient (ICC) between true and estimated Agatston scores (by patient) has been computed. Results are listed in Table 2, where the “candidates” row shows the performance of the proposed approach without the final candidate classifier (third phase) and the “final” rows show the performance of the complete approach with a probability threshold of 0.5. While we have computed the cross-validated performance on the training data (LOPO-CV) ourselves, the results on the testing data have been reported by the challenge organizers. In addition to the values reported in Table 2, the challenge organizers reported ICCs for patient volume (1.00), LADLM volume (0.95), LCX volume (0.82) and RCA volume (0.97). The risk categorization according to the computed Agatston score was correct in 94% of the patients and a linearly weighted kappa of 0.95 was achieved.

Experiments were run single-threaded on a 64-bit desktop PC with Intel Xeon CPU operated at 2.6GHz (6 cores, 12 threads) and with 32GB of main memory. On average, pericardium segmentation took 0.49s per volume, aorta segmentation took 0.11s. Centerline tracing took 16.25s, including four-chamber heart segmentation in the contrasted scan and TPS-based mapping to the non-contrast scan. Thus, the extraction of cardiac anatomy (the first phase) required about 17.5s on average. CCP extraction and classification took 123.3s on average, where it should be considered that candidate extraction and feature computation were implemented in Matlab.
Figure 2 shows the Gini importance assigned to each feature by the random forest classifier [1]. Clearly, intensity-based features dominate the classifier decision. Both features based on local neighborhood intensities achieve top ranking – an indication that the classifier mainly distinguishes noise from true coronary calcification. Some information is still contained in the relative location within the heart and distance to the vessel tree, but their contribution is limited. Also, the size of the CCP does not seem to matter; presumably its information is already depleted by the candidate selection step (phase 2). Finally, in the current setting it becomes evident that (normalized) cumulative intensity histograms (\( \text{cHist}_{n} \)) are less useful than their un-normalized counterparts (\( \text{cCnt}_{n} \)) and could safely be removed from the feature pool. However, these results could easily change if the algorithm is trained on a larger set of data. Note, that in the current version data from only 32 patients was used for training (some of which without coronary calcifications) which resulted in an overall training set of 173 positive and 282 negative examples for the random forest classifier (cf. Table 2). With a larger training base we expect that the random classifier can handle more features and further improve accuracy.

Figure 3 visualizes the performance for different probability thresholds. The chosen operating point of our method (threshold of 0.5) achieves an optimal trade-off between lesion-wise sensitivity and positive predictive value, for both of which approximately 90% is attained (cf. Table 2). This choice might be favorable if the method is evaluated with respect to average ranking. As apparent from the right graph in Fig. 3, the best ICC and a more favorable operating point regarding FPR is obtained with a threshold of 0.7. Then, of course, sensitivity would suffer and decreases by about 10%.

4 Conclusion

Coronary calcium scoring is a routine step in CT-based diagnosis of coronary artery disease. Most clinically available tools still require the physician to manually distinguish coronary calcium from other structures with increased radiodensity. For a fully automatic calcium scoring we propose to extract the coronary artery tree from an accompanying contrasted CT angiography scan and map it to the native CT scan for calcium identification. A random forest classifier is used to finally identify true coronary calcium. The approach shows very promising results and, being trained and evaluated on a larger set of data, can be a reliable tool for a fully automatic calcium scoring.

Table 2. Performance numbers for extracted calcium candidates (phase 2), cross-validated classifier (phase 3) and testing results.

<table>
<thead>
<tr>
<th>Volume candidates (training data)</th>
<th>99.40% 76.50%</th>
<th>Lesions sens.</th>
<th>94.02% 38.02%</th>
<th>0.930</th>
</tr>
</thead>
<tbody>
<tr>
<td>final (LOPO-CV)</td>
<td>98.56% 88.83%</td>
<td>89.13% 90.11%</td>
<td>0.962</td>
<td></td>
</tr>
<tr>
<td>final (testing data)</td>
<td>97% 95%</td>
<td>93% 95%</td>
<td></td>
<td>0.997</td>
</tr>
</tbody>
</table>
Fig. 2. Gini feature importance computed with the random forest classifier. Most important are intensity-based features which indicates that the classifier mainly distinguishes noise from coronary calcium.

Fig. 3. Lesion-wise performance curves for different probability thresholds. Bold points indicate the operating points for the chosen threshold of 0.5. **Right:** positive predictive value (PPV) over sensitivity (SENS). Equal values for PPV and SENS (dashed line) are achieved the operating point. **Left:** false positive detections per volume (FPR) and intraclass correlation (ICC) over sensitivity. An alternative choice for the threshold of 0.7 (dashed line) seems to be advantageous when considering ICC and FPR.
References