

Region of Interest Queries in CT Scans

Alexander Cavallaro², Franz Graf¹, Hans-Peter Kriegel¹, Matthias Schubert^{1*},
and Marisa Thoma^{1*}

¹ Institute for Informatics, Ludwig-Maximilians-Universität München,
Oettingenstr. 67, D-80538 München, Germany,
{graf,kriegel,schubert,thoma}@dbs.ifi.lmu.de,

² Imaging Science Institute, University Hospital Erlangen, Maximiliansplatz 1,
D-91054 Erlangen, Germany,
Alexander.Cavallaro@uk-erlangen.de

Abstract. Medical image repositories contain very large amounts of computer tomography (CT) scans. When querying a particular CT scan, the user is often not interested in the complete scan but in a certain region of interest (ROI). Unfortunately, specifying the ROI in terms of scan coordinates is usually not an option because an ROI is usually specified w.r.t. the scan content, e.g. an example region in another scan. Thus, the system usually retrieves the complete scan and the user has to navigate to the ROI manually. In addition to the time to navigate, there is a large overhead for loading and transferring the irrelevant parts of the scan.

In this paper, we propose a method for answering ROI queries which are specified by an example ROI in another scan. An important feature of our new approach is that it is not necessary to annotate the query or the result scan before query processing. Since our method is based on image similarity, it is very flexible w.r.t. the size and the position of the scanned region. To answer ROI queries, our new method employs instance-based regression in combination with interpolation techniques for mapping the slices of a scan to a height model of the human body. Furthermore, we propose an efficient search algorithm on the result scan for retrieving the ROI with high accuracy. In the experimental evaluation, we examine the prediction accuracy and the saved I/O costs of our new method on a repository of 2 526 CT scans.

1 Introduction

Radiology centers all over the world currently collect large amounts of 3D body images being generated by various scanners like PET-CT, MRT, x-ray or sonography. Each of these methods generates a three dimensional image of the human body by transforming the echo of a different type of signal allowing a radiologist to examine the inner parts of a human body. In the following, we will particularly focus on CT body scans. However, the methods proposed in this paper are generally applicable to other types of scans as well.

* corresponding authors

Technically, the result of a CT scan is stored as a stack of 2D images representing 3D slices of the human body, i.e. each slice is considered to have a certain thickness. The scans in a radiology center are stored in a centralized picture archiving and communication system (PACS) and they are transferred via LAN to the workstation of a physician. In commercial PACS, querying CT scans is currently restricted to retrieving complete scans being annotated with certain meta information like patient name, date and type of the examination. Therefore, each time a CT scan is queried, the complete scan, potentially comprising several thousand high-resolution images, has to be loaded from the image repository. For example, the data volume of a thorax scan being generated by a modern scanner comprises around 1 GB of data. Considering that several physicians will simultaneously query a PACS, the loading time of a single CT scan is up to several minutes depending on network and server traffic.

However, in many cases it is not necessary to display the complete scan. For example, if a physician wants to see whether a certain liver lesion has improved between two scans, the user primarily requires the portion of both scans containing the liver. Therefore, the physician loses up to several minutes by loading unnecessary information and searching for the liver within both scans. Thus, a system retrieving the parts of both scans containing the liver, would save valuable time and network bandwidth.

Parts of a CT scan can be efficiently loaded by raster databases [2] as long as the coordinates of the ROI are specified. However, in the given context, the ROI is rather defined by the image content. In other words, the coordinates of organs and other anatomical regions may strongly vary because of differences in the patients' heights or in the scanned body region. Thus, raster coordinates cannot be used to align to CT scans w.r.t. the image content.

In this paper, we focus on a query-by-example setting. Therefore, the query is posed by selecting a certain body region in a scan. The result contains the part of the scan showing the corresponding body region in one or multiple result scans. For example, a radiologist could select a certain area in the scan being currently under examination. He or she might want to see the corresponding regions in scans of patients having the same disease or earlier examinations of the same patient.

The most established approach to answer this type of queries is based on landmark detection. [14] A landmark is an anatomically unique location in the human body which is well-detectable by pattern recognition methods. To use landmarks for query processing, it is first of all necessary to detect as many landmarks as possible in the example scan and all result scans. Let us note that landmark detection employs pattern recognition methods and thus, there is a classification error, i.e. some of the predicted landmark positions are error prone. Furthermore, it can happen that some of the landmarks are not detectable due to disturbances while recording the scan. However, having detected a sufficiently large number of landmarks, it is possible to align both scans and afterwards select the area from the target scan corresponding to the query.

An important aspect of this approach is that landmark detection should be done as a preprocessing step. Thus, the example scan and the target scans need to be annotated with the landmark position to allow efficient query processing. However, this causes a problem when allowing example scans not being stored in the same PACS. In this case, the query might not have any landmarks or it is not labelled with the same set of landmarks. If the example scan and the result scan are taken by CT scanners from different companies, the positioning systems might not be compatible. Another drawback of the landmark approach is the size of the scan. CT scans are often recorded for only a small part of the body. Thus, it cannot be guaranteed that the scanned body region contains a sufficiently large set of alignable landmarks. To conclude, a fixed and comparably small set of landmarks is often not flexible enough to align arbitrary scans.

In this paper, we propose a more flexible approach being based on similarity search on the particular slices of a CT scan. Our new method does not rely on any time-consuming preprocessing step, but it can be directly applied on any query and result scan. Whereas landmark-based approaches can only align scans with respect to a limited amount of fixed points to be matched, our new approach can generate the positions in the scan to be matched on the fly. Thus, we can even align scans being labelled with different types of landmarks or scans not having any detectable landmarks at all.

The key idea behind our method is to map single slices of a CT scan to a generalized height model describing the relative distances between concepts w.r.t. the height axis of the human body. The height model is independent of the individual size and proportions of a particular patient. Let us note that it is possible to use width and depth axes as well. However, the height axis is the predominantly used navigation axis for CT scans.

By mapping single slices to the model, we can better adjust to limited information about the scan and we are independent from the distribution of predefined landmark positions. Our prediction algorithm employs instance-based regression based on Relevant Component Analysis [1] and the X-Tree [3] for efficiently answering kNN queries.

ROI queries are answered as follows: In the first step, the user selects a certain region of interest in the example scan. Afterwards, we employ instance-based regression to determine the query position in the generalized height model. In the next step, we need to determine the part in each target scan corresponding to the query interval in the height model. Let us note that this second step is more complicated, since we cannot directly determine the slice belonging to a particular height value. One solution to this problem would be to label all available slices with the height value in the model. However, labelling all DICOM images in an average PACS would cause an enormous overhead in preprocessing. Since the majority of images will never be involved in answering an ROI query, we follow a different strategy. Instead of preprocessing each image in the PACS, determining height values for a given slice is done on the fly. To make this type of processing efficient, we propose a query algorithm that alternates regression and interpolation steps until the queried ROI is found in the result scan.

Let us note that although the solutions proposed in this paper are very problem-oriented, the solution principle can be extended to other data as well. For example, a similar processing scheme can be applied to video streams (e.g. procedure timing in surveillance videos) or text mining (e.g. news tickers, twitter streams, age classification in Internet forums).

The rest of the paper is organized as follows. Sect. 2 surveys methods that are related to our approach or parts of it. In Sect. 3, we formalize ROI queries and give an overview of our system. Afterwards, Sect. 4 introduces our method for predicting height values for particular CT slices. Sect. 5 first describes interpolation methods for aligning CT scans to a generalized height model and then presents our new query algorithm. The results of our experimental evaluation are shown in Sect. 6. The paper concludes with a brief summary and ideas for future work in Sect. 7.

2 Related Work

In medical imaging, there are various localization or registration approaches. Most of them are very domain specific, like the Talairach space brain atlas [15] or the MNI space [5]. Nevertheless, as these atlases are very specific to their domain, they were not designed to cover the entire body and they can thus hardly be used for general ROI queries.

Position mapping via landmark-detector-based approaches like the Theseus Medico system presented in [14] are more appropriate for our purpose. This prototype provides an image parsing system which automatically detects 22 anatomically relevant landmarks, i.e. invariant points, and 9 organs. [13] It is thus possible to query the database directly for ROIs which are equivalent to these automatically-annotated image regions. However, general queries for arbitrarily defined ROIs are not yet supported.

A more general, landmark-based interpolation approach for mapping a volume into a standardized height space has been proposed by [7]. However, it is very patient-specific and dependent on the used landmarks. Another approach that uses partial volumes as query is described in [6]. It localizes query volumes with sizes ranging from 4 cm to more than 20 cm by comparing the partial volume with an implicit height atlas based on Haar-like features. In [4], we presented an alternative method such that only a single query slice is needed in order to achieve comparable results.

In Sect. 5.2, we introduce an iterative interpolation and regression approach. In contrast to established regression methods, [9, 11] we enhance our model with newly generated information after each iteration in order to refine the final model until convergence is reached.

We experimented with several regression methods from the Weka machine learning package [8]. However, simple approaches like linear regression did not yield a sufficient prediction accuracy and more complicated approaches like support vector regression using non-linear kernel functions could not cope with the enormous amount of training data. Therefore, we decided to employ instance-

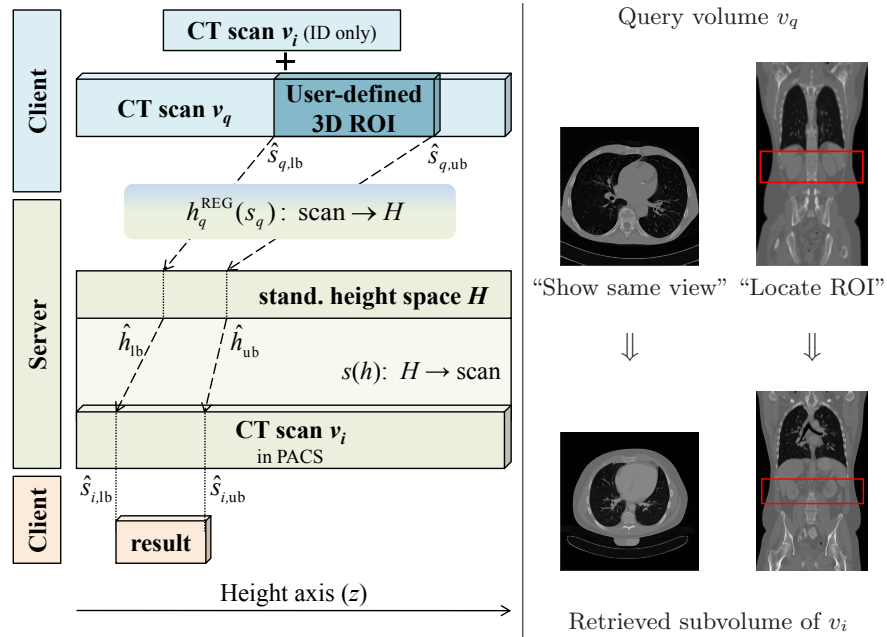


Fig. 1. Workflow of ROI retrieval and two example queries. The first query is specified by a ROI of only one slice, the second is given by a 3D ROI.

based regression which is robust and sufficiently fast when employing techniques of efficiently computing the k -nearest neighbors (k -NN). In particular, we employ k -NN queries being based on the X-Tree [3]. Let us note that there are multiple other index structures [12] for speeding up the same type of query. We decided to employ the X-Tree because it represents an extension of the standard R*-Tree [10] which is better suited for higher dimensionalities.

Current database systems like RasDaMan [2] already support conventional region of interest queries in raster data like CT scans. Nevertheless, the system needs to know the coordinate system in which the query is applied in order to navigate to the requested region. As we do not know the complete coordinate systems of the patients' CT scans in advance and since patients differ in height and body proportions, and thus, locations along the z -axis are not standardized, a globally fixed coordinate system will not be available in our setting. Therefore, our new approach represents a way to bridge the gap between the coordinates in the example scan and the coordinate system of the result scan.

3 Example-Based ROI Queries

In this section, we specify the proposed query process and give an overview of the proposed system. Formally, a dataset consists of n volumes $v_i \in \mathbb{N}^{x(i) \times y(i) \times z(i)}$ with $i \in \{1, \dots, n\}$ and varying voxel dimensions x, y, z . The height model H is

an interval $[h_{\min}, h_{\max}] \in \mathbb{R}_o^+$ representing the extension of the human body in the z -axis. A mapping function $h_i : \mathbb{N} \rightarrow H$ maps slices of volume v_i to a height value $h \in H$. Correspondingly, the reverse mapping function $s_i : H \rightarrow \mathbb{N}$ maps a position h in the height model to a slice number s in v_i . A matching point $p = (\mathbf{s}_p, \mathbf{h}_p, \mathbf{w}_p) \in \mathbb{N} \times H \times \mathbb{R}$ is a triple of a slice number, its corresponding height value in H and a reliability weight w . We use $p_{i,j}$ for naming the j^{th} matching point in scan v_i .

In our system, a region of interest (ROI) query is specified by a set of consecutive slices $(\hat{\mathbf{s}}_{e,\text{lb}}, \dots, \hat{\mathbf{s}}_{e,\text{ub}}) \subseteq \{0, \dots, z(e) - 1\}$ from an example scan v_e and it retrieves a consecutive sequence of CT slices $(\hat{\mathbf{s}}_{i,\text{lb}}, \dots, \hat{\mathbf{s}}_{i,\text{ub}}) \subseteq \{0, \dots, z(i) - 1\}$ from the result scan v_i .

Fig. 1 illustrates the complete workflow of query processing for example-based ROI queries. A user specifies the ROI query on the client computer by marking a region in an example scan v_e . Additionally, the queried scan v_i has to be identified for the server. Let us note that it is not necessary to transfer the complete marked subset of the example scan. Instead it is sufficient to transfer a scale-reduced version of the first and the last slice of the subset. After receiving the slices, the server performs a feature extraction step generating image descriptors for both slices. As an alternative, the client computer might directly compute the required image descriptors and only transfer the descriptors.

In the next step, the server employs a mapping function to predict height values \hat{h}_{lb} and \hat{h}_{ub} to describe the borders of the query interval in the height model H . In our system, $h_e^{\text{REG}}(s)$ is implemented by instance-based regression (cf. Sect. 4). Afterwards, our algorithm starts with aligning the result scan v_i to the height model H by employing the algorithm described in Sect. 5.2. In particular, the algorithm employs $h_i^{\text{REG}}(s)$ for generating matching points P_i which are required for the reverse mapping function $s_i(h)_{P_i}$, an interpolation function described in Sect. 5.1. Once the quality of $s_i(h)_{P_i}$ is satisfying, the server selects the sequence of slices $(\hat{\mathbf{s}}_{i,\text{lb}}, \dots, \hat{\mathbf{s}}_{i,\text{ub}}) \subseteq \{0, \dots, z(v_i) - 1\}$ from v_i corresponding to the height interval $[\hat{h}_{\text{lb}}, \hat{h}_{\text{ub}}]$ and returns them to the client as the query result. Let us note that $(\hat{\mathbf{s}}_{i,\text{lb}}, \dots, \hat{\mathbf{s}}_{i,\text{ub}})$ is extended by the amount of slices corresponding to 90% of the expected prediction error in order to compensate for the inaccuracy of $h_i^{\text{REG}}(s)$.

Table 1 displays an overview of the defined parameters including some additional annotations that will be introduced in the next sections.

4 Efficient Instance-Based Regression

In this section, we introduce our method for mapping a single slice into the standardized height scale H . We already mentioned that there exist methods for landmark and organ detection which mark slices in the scan with the detected landmarks or organs. [13] Using multiple landmarks detected at slices $\mathbf{s}_{i,j}$, which can be mapped to anatomical concepts with known standardized height positions \mathbf{h}_j and reliabilities \mathbf{w}_j , we can infer the standardized height of our queried slice. The landmark detector of [13] being used in our experiments was not

