A novel 3D shape descriptor for automatic retrieval of anatomical structures from medical images

Leila C. C. Bergamasco[a], Miguel A. G. Valverde[b], Marcel P. Jackowski[b], Fátima L. S. Nunes[a]

[a] *Laboratory of Computer Applications for HealthCare, School of Arts, Sciences and Humanities, University of São Paulo*
Graduate Program in Electrical Engineering, Polytechnic School, University of São Paulo
[b] *Department of Computer Science, Statistics and Mathematics Institute, University of São Paulo*

Abstract

Three-dimensional medical models are increasingly used for the detection of abnormalities in anatomical structures imaged by Magnetic Resonance Imaging (MRI) or Computer Tomography (CT) techniques. In this context, the development of descriptors that can capture information specific to organs or structures is desirable. In this work, we focus on the retrieval of two anatomical structures commonly imaged by MR techniques, namely, the left ventricle of the heart and blood vessels. Towards this aim, we developed a novel 3D local shape descriptor which uses mesh geometry information such as facet area and distance from centroid to surface to identify shape alterations. This novel descriptor showed good performance, anda 90%precision rate for retrieving both convex (left ventricle) and non-convex structures (vessels), allowing for detection of abnormalities associated with changes in shape. As a result, it has the potential to aid medical experts in the diagnosis of a wide range of vascular and cardiac diseases.

Keywords:
CBIR 3D, medical 3D models, local descriptor, vessels, CHF.

Introduction

Content-based Image Retrieval (CBIR) uses information about the content of an image to query for the most similar images from a database. CBIR can help to compose diagnoses by retrieving exams similar to a specific case. This technique has been broadly applied in Computer Aided Diagnosis (CAD) systems, in the two-dimensional (2D) context [1] but its use in three-dimensional (3D) scenario is still underexplored. Some complex medical exams use 3D models for diagnosis, like Magnetic Resonance Imaging (MRI) and Computadorized Tomography (CT). These modalities allow experts identify anomalies, such as aneurysm and artery coronary disease without invasive methods [2]. However, a single exam can generate hundreds of slices, which demands time to be analyzed. On the other hand, these modalities allow the generation of a 3D model from those slices, which can in turn be used to compose a diagnosis. In this context, techniques to quickly retrieve these models can be an important contribution to aid in medical diagnosis.

Global shape descriptors are frequently used in CBIR for 3D models. Features usually consider the geometric aspects, like vertices and facets. Direct derivation from 3D data structures, their satisfactory potential to represent the structurres and easiness in the implementation are some of their advantages. These descriptors are can be invariant to rotation, translation and scale. However, the computation processes are usually too time-consuming and they are insensitive when a dissimilarity occurs in a specific location of an object [3]. The latter problem can be minimized by using local shape descriptors.

This paper presents a novel local descriptor for 3D medical model retrieval, an "Area-Distance Local Descriptor" (ADLD). This approach is suitable for retrieving cases where deformations are located in specific regions of the model. To measure the efficiency of our approach, we applied the descriptor in convex models reconstructed from cardiac MRI images of the left ventricle (LV). These images are normally used to aid the detection of Congestive Heart Failure (CHF), a disease that causes deformation of the left ventricule shape. We also analyzed the performance of our descriptor using synthetic models.

In the medical imaging context, a few descriptors have been proposed. Wu *et al*. [4]'s and Glatard *et al.* [5] descriptors showed a good performance for 3D medical models of brain and heart, respectively. Our descriptor was developed with focus in convex structures, however the evaluation with non-convex models also showed a good performance.

This article is organized as follows: Section "Content-Based Image Retrieval" describes CBIR concepts in the 2D and 3D domains. In Section "Methods", ADLD and 3D medical models considered in the evaluation and the test methodology are described. Sections "Results" and "Discussion" present results obtained with convex and non-convex models and a discussion on the performance of the descriptor. In the Section "Conclusion", we present the final remarks about the local shape descriptor.

Content-Based Image Retrieval

CBIR systems retrieves from a database images that are more similar to an image provided as a query. The same concept is applied in the 3D context. Some steps are common in such systems, such as pre-processing, feature extraction and comparison between feature vectors using similarity functions. In the pre-processing stage, algorithms are applied to enhance regionsof interest or suppress noise [6]. In the feature extraction step, descriptors are defined to represent image features such as color, texture and shape. In general, several descriptors are developed for the same image. This set of features is organized in a feature vector. Authors have studied faster and more robust extraction that improves the accuracy of CBIR systems [7]. Similarity functions are used in the following phase to calculate the difference in content between two images based on their features. In addiiton, different metrics can be used to calculate distances between feature vectors, the most common beingthe Euclidean and Manhattan distances.

CBIR in 3D domain is an innovative approach. Some studies have tested techniques by using generic 3D objects found in classical 3D models databases such as Princeton

Benchmark [8]. However, it is important to analyze the effectiveness of the descriptors in specific situations to propose improvements and solve real problems. A challenge to develop local descriptors is how to correctly identify local differences related to specific deformations which often go unnoticed by the global descriptors. In spite of the additional cost of processing due to the more detailed analysis, the accuracy of the retrieval can be improved when local differences are common, frequently observed in the medical context.

Methods

The Area-Distance Local Descriptor - ADLD

The Area-Distance Local Descriptor (ADLD) is based on the Global Geometric Feature Map (GGFM), proposed by Wang *et al* .[9], which extracts information related to the facet area and the geometric distance between the centroid of the 3D model and each facet. Data are mapped in a two-dimensional matrix and indexed by spherical coordinates. The comparison between two models uses the Euclidean distance function.

The ADLD comprises two features vectors, as shown in Equation 1, where F_k is the k^{th} facet of the 3D model and it is composed by the spherical coordinates (θ and φ) of the k^{th} facet normal orientation. Those coordinates are defined on Equation 2. A_k is the k^{th} facet area and D_k is the distance between 3D model centroid and its facet.

$$F_k = [A_k;\ \theta;\ \varphi], [D_k;\ \theta; \varphi] \quad (1)$$

$$\theta = arccos\frac{z_k}{D_k}\ ;\quad \varphi = arctan\frac{y_k}{x_k} \quad (2)$$

The normal orientation of the k^{th} facet is defined by cartesian coordinates x_k, y_k and z_k (Equations 3, 4 e 5). The size of the normal vector D_k is computed by Euclidean distance between facet and centroid.

$$x_k = (y_2 - y_1) * (z_3 - z_1) - (z_2 - z_1) * (y_3 - y_1) \quad (3)$$

$$y_k = (z_2 - z_1) * (x_3 - x_1) - (x_2 - x_1) * (z_3 - z_1) \quad (4)$$

$$z_k = (x_2 - x_1) * (y_3 - y_1) - (y_2 - y_1) * (x_3 - x_1) \quad (5)$$

Equation 6 computes the facet area where AB, BC and AC are lengths of the edges that compose a triangular facet of the 3D model mesh. ~~Equation 7 shows how it is calculate the cosα~~

$$A_k = \frac{AB * BC * \sqrt{1 - \cos \propto}}{2} \quad (6)$$

$$\cos \propto = \frac{AB^2 + BC^2 - AC^2}{2 * AB + BC} \quad (7)$$

The last step is to compute the distance between the facets and the centroid, given by Equation 8, where x_c, y_c and z_c are the barycenter coordinates of each facet (~~Equations 8, 9 and 10~~).

$$D_k = \sqrt{x_c^2 + y_c^2 + z_c^2} \quad (8)$$

$$x_c = \frac{x_1 + x_2 + x_3}{3} \quad (9)$$

$$y_c = \frac{y_1 + y_2 + y_3}{3} \quad (10)$$

$$z_c = \frac{z_1 + z_2 + z_3}{3} \quad (11)$$

The data ~~above mentioned~~ compose two matrices: one for computed distances and other for the computed areas. To compare two different models, the Euclidean distance is calculated for each pair of matrices (MD for distance matrix and MA for area matrix). Lastly, the sum of these results is computed, described by Equation 12.

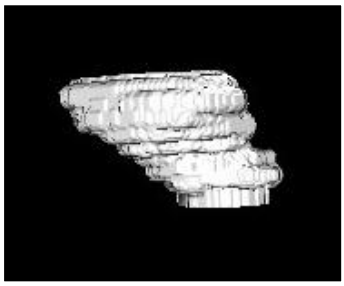
$$MRes = \sum_{i=1}^{n}\sum_{j=1}^{m}\sqrt{(MD1_{ij} - MD2_{ij})^2 + (MA1_{ij} - MA2_{ij})^2} \quad (12)$$

ADLD Validation

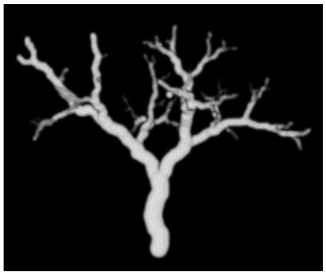
The ADLD performance was evaluated by using: (1) convex models of left ventricle from cardiac MRI images, and (2) non-convex models of synthetic blood vessels.

The cardiac models were provided by Instituto do Coração (InCor), one of the largest centers for treating heart diseases in Brazil. Experts from InCor classified the LV models used in this experiment into CHF and healthy (without CHF) groups. CHF affects the ventricle contractility and, consequently, blood pumping [10], causing deformation of this structure. 30 cases were selected for this evaluation (16 with CHF disease and 14 without CHF disease). A result is considered correct when a similar case is retrieved, that is, models with CHF disease should be retrieved when a case of this group (cases 1 to 16) is given and vice-versa.

Figure 1(a) presents an example of a left ventricle surface reconstructed.



(a)



(b)

Figure 1 – 3D medical models: (a) Surface of left ventricle reconstructed and (b) synthetic blood vessel image.

The second set is composed by synthetic blood vessel models defined by a grammar composed by parametric and stochastic rules[11]. The resulting expressions were used to create synthetic vessel models that mimic MRI and CT images. This methodology allows for vessel growth to be limited by arbitrary 3D surfaces, and the vessel intensity profile can be tailored to match real angiographic intensities. Isosurfaces were extracted from these synthetic angiographic images to be used in the ADLD.

A projection of one synthetic blood vessel generated using this methodology is shown in Figure 1(b). A total of 8 synthetic models, from 1 to 4 ramifications were used. Each model was modified in order to simulate deformations.

A result is considered correct when the same model with deformation is the first to be retrieved. After, the other models with similar amount of ramifications are expected. For example: when a model with two ramifications is given, the system should return first this model with deformation and after the model with three ramifications, four ramifications and so on.The descriptor performance was evaluated by using Precision versus Recall metric, which is broadly used to evaluate CBIR systems.

Results

Figure 2 presents the graph with Precision vs. Recall metric applied on the results from the first set of models. Precision reached 85% for all values of Recall, for for both situations: 3D medical models with CHF and without the disease. The line that represents non-CHF has a lower value of Recall (0.94) because has fewer cases. Figure 3 shows an example of left ventricle given as query and the three results most similar retrieved by prototype, in this case all of them presented CHF.

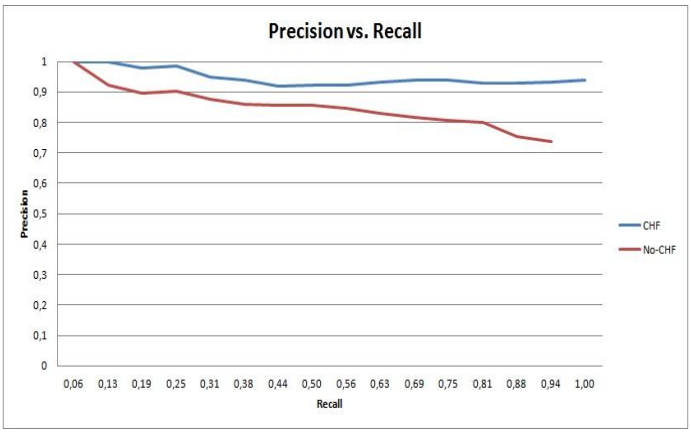


Figure 2– Precision vs. Recall metric of ADLD for 3D models of left ventricle with CHF and non-CHF.

Figure 3– Example of left ventricle given as query and the 3D models retrieved.

Figure 4 shows the plot of the Precision vs. Recall metric applied on the results from the second set of models. ADLD reached 88% of Precision for all values of Recall. This means that it is possible use the ADLD for different structures, when retrieval based on their shapes is desirable.

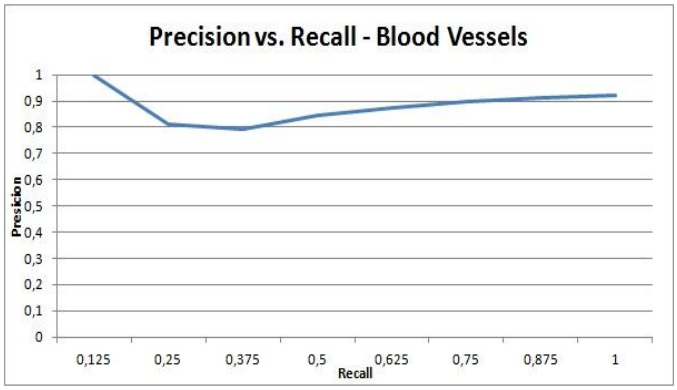


Figure 4– Precision vs. Recall metric of ADLD for 3D models of synthetic blood vessels.

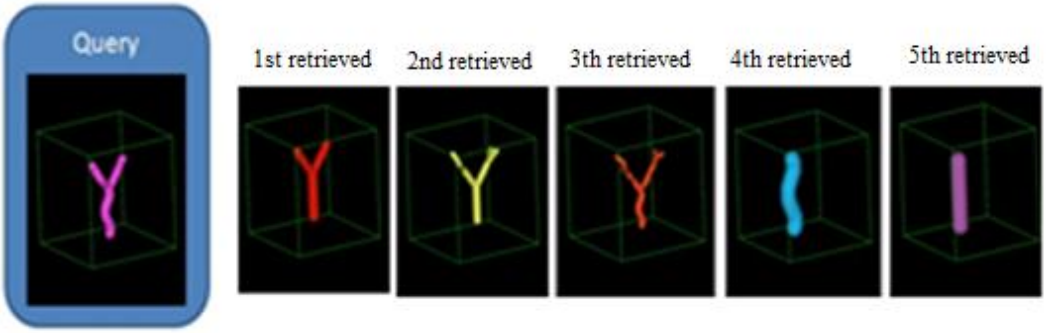


Figure 5– Example of synthetic blood vessel given as query and the retrieval results in order of similarity.

Figure 5 shows an example of the result of a query. As we can see, the models were retrieved in the correct order: the same 3D model without deformation, and after, those structures that have one ramification of difference (one more or one less ramification).

Figure 6 shows the Distance Matrix for the synthetic 3D blood vessel models. The "N1…N4" fields relate to normal synthetic models, without deformation. The "D1…D4" fields represent 3D blood vessel models with deformations. The number in each cell it is de Euclidean Distance between the 3D models.

This matrix shows the the design and implementation appropriateness of ADLD, since we obtained the expected results, as it was known in advance all synthetic 3D models and their ramifications. The distance from a normal 3D model to its respective deformed 3D model is smaller when compared to models with more ramifications, corroborating the fact that the deformation of the synthetic models proportionately increase by the amount of ramifications.

	N1	D1	N2	D2	N3	D3	N4	D4
N1	0	0.38	0.63	1.21	0.8	1.4	1.59	1.77
D1	0.38	0	0.58	0.98	0.71	1.3	1.49	1.67
N2	0.63	0.58	0	0.97	0.75	1.35	1.54	1.72
D2	1.21	0.98	0.97	0	0.54	0.64	0.84	1.02
N3	0.8	0.71	0.75	0.54	0	0.6	0.79	0.98
D3	1.4	1.3	1.35	0.64	0.6	0	0.28	0.43
N4	1.59	1.49	1.54	0.84	0.79	0.28	0	0.26
D4	1.77	1.67	1.72	1.02	0.98	0.43	0.26	0

Figure 6 – Distance Matrix of blood vessels synthetic 3D models.

Discussion

In both scenarios ─ convex and non-convex 3D medical models of anatomical structures ─ the ADLD showed good performance, and more than a 85% precision rate.

In regard to the modeling of the left ventricles, the descriptor enabled the identification of small changes in the structures, which includes the local deformation of the apex in patients with CHF disease. Spherical coordinates enables the local comparison between 3D models and the deformation degree can be identified by distance between centroid and 3D model facets.

The CHF disease was previously investigated using other descritors that considered global and local characteristics [12],[13]. In [12], a global descriptor which just analysed the distance between centroid and surface was used and yielded about 70% of precision. By using a similar method in which the 3D model was split into octants and a local reference was used, it was possible reach a mean rate precision rate of 80% [13]. Hence, the ADLD method described here yielded a superior precision score, when compared to those methods.

Regarding the blood vessel results, the ADLD also showed good performance, yielding a 88% precision rate. In this retrieval, the information related to the facet area was the factor which most contributed to obtain this performance. Increasing ramifications, result in increasing the sum of the facet area. As ramifications are located in specific regions of the model, ADLD can detect subtle differences or similarities in specific regions of two models.

Conclusion

Here we have introduced the ADLD ─ a novel 3D shape local descriptor for the retrieval of anatomical structures which take into consideration local deformations.

The descriptor presented a precision rate of 85% and 88%, for left ventricles and blood vessels models, respectively. These results indicated that the ADLD enables the detection of subtle local deformations in medical models. It is important to emphasize that both convex and non-convex models were considered. Since most of descriptors described in the literature shows good performance only for a restricted subset of models, this is another strength of the ADLD.

The innovation of this descriptor relies on considering both area and distance to centroid from facets of a mesh model. Thisinformation allows that the ADLD to be applied in the retrieval of a diverse set of anatomical structures.

Acknowledgement

Authors are grateful to the Brazilian National Council of Scientific and Technological Development (Process \#401745/2013-9), and the National Institute of Science and Technology Medicine Assisted by Scientific Computing for the financial support. They also thank to Heart Institute (InCor) and Professor Carlos Eduardo Rochitte for providing the images.

References

1. K. Doi. *Current status and future potential of computer-aided diagnosis in medical imaging*. In TheBritish Journal of Radiology, number 78, pages 3–19.2005.

2. V. M. Pereira, P. Bijlenga, A. Marcos, K. Schaller, and K. Lovblad. Di*agnostic approach to cerebralabeurysms*. European Journal of Radiology, 2012.Article in Press.

3. Y. Yubin, L. Hui, and Z. Yao. C*ontent-based 3-Dmodel retrieval: A survey*. In Proceedings of 7th IEEETransactions on Systems, Man, and Cybernetics.,volume 36, pages 1081–1098. IEEE Computer Society,2007. 07, IEEE Computer Society. p. 1081-1098.

4. H. Wu, J. Kim, W. Cai, and D. Feng. *Volume ofinterest (VOI) feature representation and retrieval ofmulti-di*mensional dynamic positron emissiontomography images. In Intelligent Multimedia, Videoand Speech Processing, 2004. Proceedings of 2004International Symposium on, pages 639 – 642, oct.2004.

5. T.Glatard, J. Montagnat, and I.E. Magnin, *Texture based medical image indexing and retrieval: application to cardiac imaging*, in *Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval*. 2004, ACM: New York, NY, USA. p. 135-142.

6. A. Smeulders, M. Worring, S. Santini, A. Gupta, andR. Jain. *Content-based image retrieval at the end ofthe early years*. Pattern Analysis and MachineIntelligence, IEEE Transactions on, 22(12):1349–1380, dec 2000. 7.

8. P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser.*The princeton shape benchmark*. In In Shape ModelingInternational, pages 167–178, 2004.
9. D. Wang and C. Cui, *3D Model Similarity Measurement with Geometric Feature Map Based on Phase-Encoded Range Image.* Optics Express, 2004: p. 103-110.
10. F. H. Martini. *Fundamentals of Anatomy andPhysiology*. Pearson, 17a edition, 2006.
11. M. A. Galarreta-Valverde, M. M. G. Macedo, C. Mekkaoui and M. P. Jackowski. *Three-dimensional synthetic blood vessel generation using stochastic L-system.* Proc. SPIE 8669, Medical Imaging 2013: Image Processing, 86691I (March 13, 2013); doi:10.1117/12.2007532.
12. L.C.C. Bergamasco and F.L.S. Nunes, *Applying Distance Histogram to retrieve 3D cardiac medical models*, in *AMIA Annual Symposium Proceedings*. 2013. p. 112-121.
13. L.C.C. Bergamasco andF.L.S.Nunes, *A new local feature extraction approach for content-based 3D medical model retrieval using shape descriptor,* in *Proceedings of the 29th Annual ACM Symposium on Applied Computing* . 2014, ACM: Gyeongju, Republic of Korea. p. 902-907.