34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

A Graph-Based Approach to the Retrieval of Volumetric PET-CT Lung Images

Ashnil Kumar, *Student Member, IEEE,* Jinman Kim, *Member, IEEE,* Lingfeng Wen, *Member, IEEE,* and Dagan Feng, *Fellow, IEEE*

***Abstract*—Combined positron emission tomography and computed tomography (PET-CT) scans have become a critical tool for the diagnosis, localisation, and staging of most cancers. This has led to a rapid expansion in the volume of PET-CT data that is archived in clinical environments. The ability to search these vast imaging collections has potential clinical applications in evidence-based diagnosis, physician training, and biomedical research that may lead to the discovery of new knowledge. Content-based image retrieval (CBIR) is an image search technique that complements conventional text-based retrieval by the use of image features as search criteria. Graph-based CBIR approaches have been found to be exemplary methods for medical CBIR as they provide the ability to consider disease localisation during the similarity measurement. However, the majority of graph-based CBIR studies have been based on 2D key slice approaches and did not exploit the rich volumetric data that is inherent to modern medical images, such as multi-modal PET-CT. In this paper, we present a graph-based CBIR method that exploits 3D spatial features extracted from volumetric regions of interest (ROIs). We index these features as attributes of a graph representation and use a graph-edit distance to measure the similarity of PET-CT images based on the spatial arrangement of tumours and organs in a 3D space. Our study aims to explore the capability of these graphs in 3D PET-CT CBIR. We show that our method achieves promising precision when retrieving clinical PET-CT images of patients with lung tumours.**

I. Introduction

Medical imaging plays a fundamental role in modern healthcare. This has led to a rapid expansion in the volume of medical images stored in databases and picture archiving and communication systems (PACS). In addition to retaining a record of patients' imaging results, these massive image collections also offer the opportunity for other applications such as evidence-based diagnosis, physician training and biomedical research [1], [2], [3]. However, these applications all rely upon the extraction and recognition of relevant image features.

Content-based image retrieval (CBIR) is an image search technique that complements the conventional text-based retrieval of images by using visual features as search criteria [4]. These features typically include shape, colour, texture and the spatial arrangement of objects within an image. In recent years, CBIR research has evolved from retrieving images of one modality to multi-modality images, providing the ability to form search queries using complementary image data. Multi-modality images, such as combined positron emission tomography and computed tomography (PET-CT), have become highly utilised in clinical environments due to their ability to present complementary data that improves clinical tasks. For example, PET-CT allows improved cancer diagnosis, localisation, and staging compared to single-modality PET or CT [5].

In previous work we presented the development of PET-CT CBIR using overlapping features [6]. We have also used feature vectors containing information extracted from both modalities: texture from CT, and the standard uptake value (SUV) of the PET radiotracer ^{18}F-Fluorodeoxyglucose (FDG) [7], [8]. However, this approach did not consider the relationships between functional and anatomical structures, which are powerful features for discriminating images.

We also proposed a graph-based approach that retained the relationship features between PET and CT regions [9]. Our method extended the single-modality graph-based CBIR method proposed by Petrakis et al [10], [11] in which regions of interest (ROIs) were represented as attributed graph vertices while attributed graph edges represented relationships between pairs of ROIs. Our choice of a graph-based approach was inspired by the fact that graphs are a powerful tool in pattern recognition applications, including image matching [12]. Furthermore, graphs have also been shown to have the highest retrieval accuracy when searching for images based upon the similarity of the spatial arrangement of their ROIs [10].

The staging system for lung cancer [13] is an inherently three-dimensional scheme as it classifies the disease according to the proximity, size, and relationships between tumours to surrounding anatomical structures. However, our previous work [9] investigated PET-CT CBIR retrieval with a 2D key slice based approach, where a single slice was chosen to represent the entire PET-CT volume. This approach did not take advantage of the full benefits of 3D features provided by the volumetric images. As tumours can occur in multiple locations it is unlikely that a single transaxial slice contained all the information necessary to completely represent the current condition of a patient.

In this paper, we investigate how indexing 3D spatial features introduces new capabilities by enabling the retrieval of PET-CT images based on the similarity of 3D tumour localisation. We propose a graph-based approach that indexes

This work was supported in part by ARC and PolyU grants.

A. Kumar, J. Kim, L. Wen, and D. Feng are with the Biomedical and Multimedia Information Technology (BMIT) Research Group, School of Information Technologies, University of Sydney, Australia.

L. Wen is also with the Department of PET and Nuclear Medicine, Royal Prince Alfred Hospital, Australia.

D. Feng is also with the Department of EIE, Hong Kong Polytechnic University, Hong Kong.

3D features extracted from volumetric ROI in both PET and CT images. Our similarity measurement algorithm uses a graph-edit distance calculation to retrieve images based on the spatial arrangement of tumours and organs in a 3D volume. We demonstrate the capabilities of our 3D graph-based CBIR framework in the retrieval of clinical PET-CT lung tumour studies.

II. Method

A. Data Set

We obtained 50 PET-CT studies of patients with lung tumours. The images were all acquired on a Siemens Biograph mCT scanner with a CT resolution of 512 x 512 pixels at $0.98mm$ x $0.98mm$, a PET resolution of 200 x 200 pixels at $4.07mm$ x $4.07mm$, and a slice thickness of $3mm$. We also obtained the clinical reports for each study, containing a description and locations of the tumours and nodal involvement in each case. Each study contained an average of 2.1 tumours (minimum 1, maximum 7).

We used a well-established adaptive thresholding algorithm [14] with refinements to segment the lung ROI from the CT. We applied connected thresholding to coarsely segment the brain and mediastinal tissue (including the heart). Tumours from the PET images were segmented with a 40% peak SUV (SUV_{40}) connected thresholding to detect hot spots indicated in the diagnosed reports [15]. We used the clinical reports to manually make minor corrections to the segmented ROI to ensure that the segments were well-defined.

We created a reference index of image similarity that specified for every PET-CT image in the data set the list of images that were similar to it. The reference index was created on the basis of the tumour locations (e.g. a particular lung lobe) specified within the clinical reports.

B. Graph Representation

We defined our multi-modality graph as follows: $G_C = (V_C, V_P, E, I)$, where V_C is the set of vertices representing CT ROI, V_P is the set of vertices representing PET ROI, $E = \{v_i v_j\}\ \forall\ v_i, v_j \in V_C \cup V_P$ where $v_i \neq v_j$ is the set of all edges, and I is the multi-modality image represented by G_C. By definition, all graph vertices were pairwise connected by an edge.

Every vertex $v_i \in V_C \cup V_P$ was defined to be a group of features: $v_i = (c, v, s, l)$. The vertex features were:

1) centroid (c) - the 3D centre of mass of the ROI. This feature was used for the calculation of other features and *not* for similarity measurement.
2) volume (v) - the number of voxels (individual 3D picture elements) within the ROI multiplied by the voxel size (mm^3).
3) surface area (s) - the number of pixels on the boundary of the ROI multiplied by the appropriate pixel size (mm^2).
4) length (l) - the maximum distance between two voxels on the boundary of the ROI (mm).

Every edge $e \in E$ was also defined to be a group of features: $e = (ev, d, ro, rv, md)$. The edge features were:

1) endvertices (ev) - the pair of vertices v_i, v_j with $v_i \neq v_j$ connected by this edge. This feature was used to calculate other features.
2) distance (d) - the distance between the centroids of the ROIs represented by the endvertices (mm).
3) relative orientation (ro) - the pitch and yaw angles between the centroids of the two ROI.
4) relative volume (rv) - the ratio of the volumes of the two ROI.
5) minimum distance (md) - the minimum distance between any two voxels within either ROI (mm).

We ensured that features that had high numerical values did not contribute more to the eventual similarity calculation by normalising features to the range $[0, 1]$. This was done by linearly scaling non-angle features to a random variable with zero mean and unit variance, and then shifting the value so that it was within the desired range [16]. The ro feature was normalised to the same range by taking the sine and cosine of the angles as suggested in [10].

Fig. 1 shows an example of an image and its corresponding graph representation. The CT image in Fig. 1(a) contains 3 ROI: the left lung, the right lung, and the mediastinal tissue. The corresponding PET image in Fig. 1(b) contains 4 tumour ROIs. Fig. 1(c) is the graph constructed from these seven ROI. While the figure depicts only a single 2D slice, our graph construction creates a single vertex for every 3D ROI.

C. Similarity Measurement

Our similarity measurement was based upon the concept of graph-edit distance, which can simply be explained as the cost to transform one graph into another. The brute force technique we used previously [9] had an exponential computational complexity and was not feasible for the large graphs that were used in this study. We therefore adapted the beam-search A* algorithm [17], which calculates graph edit distances by applying a beam to the popular A* algorithm [18]. This improved the speed of the graph edit distance calculation by limiting the number of combinations of vertex-to-vertex mappings that are compared.

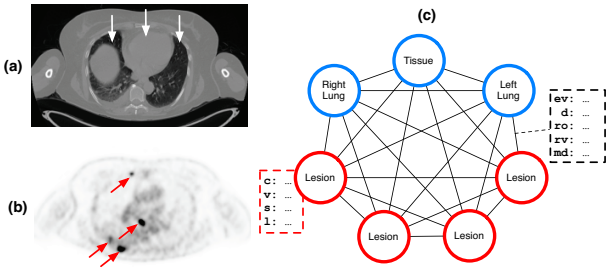

Fig. 1. Constructing a graph representation from a PET-CT image. (a) The CT image with 3 ROI. (b) The PET image with 4 ROI (tumours). (c) The graph representation.

We used our separation of the vertices into two sets (V_C and V_P) to ensure that CT vertices would be mapped only to other CT vertices, and PET vertices to other PET vertices. This meant that a PET vertex would never be mapped to a CT vertex when following the beam-search A* algorithm as described in [17]. The cost to match a vertex of edge of a query graph to an appropriate element of a graph representing an image in the data set was given by the Euclidean distance:

$$d(Q,S) = \sqrt{\sum_i (q_i - s_i)^2} \quad (1)$$

where Q is a vertex from the query graph, S is a vertex from a graph in the data set, and q_i and s_i are the i-th features of Q and S, respectively.

Extra vertices and edges on either the query or retrieved graph were accounted for with the following insertion and deletion cost, as obtained from [11]:

$$del\_ins(R) = \sqrt{\sum_i r_i^2} \quad (2)$$

where R is a vertex or edge being deleted or inserted and r_i is the the i-th features of R. Vertex deletion has a cascading effect because it also removes incident edges, thereby increasing the edit distance.

D. Experimental Procedure

We carried leave-one-out experiments using the entire data set and calculated the precision and recall of every retrieval. In addition, we also devised two specific retrieval experiments to demonstrate the ability of our CBIR method to find images with tumours with specific 3D localisations. The first was aimed at finding images with a single tumour within the upper lobe of the right lung, while the second attempted to find images with two tumours in different locations: the lower lobe of the left lung and the mediastinum.

Precision and recall are defined as:

$$precision = \frac{tp}{tp + fp} \quad (3)$$

and

$$recall = \frac{tp}{tp + fn} \quad (4)$$

where tp is a true positive (or similar) retrieved image, fp is a false positive (or dissimilar) retrieved result, and fn is a false negative or a similar image that was not retrieved.

III. Results and Discussion

We calculated the mean precision and recall across all the leave-one-out retrievals. Fig. 2 is an 11-point precision-recall plot of the averaged data. On average, our method was able to achieve $\geq 40\%$ precision at all levels of recall. In retrieval literature [19], the most similar images are expected within the first few retrieved images i.e. at low recall values (usually 4 or 5 images). The user then selects the correct images from the retrieved subset. Fig. 2 shows that our method is capable of achieving a mean precision of between 60-70% at low levels

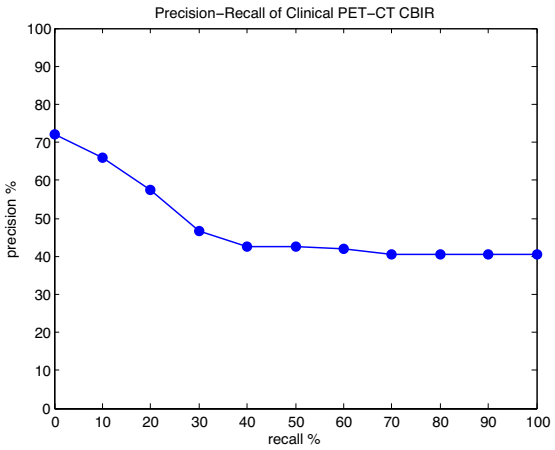

Fig. 2. Mean precision and recall when retrieving volumetric PET-CT images.

of recall (about $\leq 20\%$). In our experiments, this result means that on average more than half of the first 4 retrieved images are relevant to the query, which is in-line with other retrieval approaches [20].

Fig. 3 depicts the results of the experiments aimed at finding images that contain tumours with specific 3D localisations. For the purpose of clarity, the figure only shows coronal maximum intensity projects (MIPs) of the PET images of every case. The tumours within the images have been marked in red. The number beneath a retrieved image is the cost to transform the graph representing that image to the graph of the query image. Note that due to the nature of PET-CT acquisition, the left side of the body is depicted on the right side of these images.

The query in the first column contains a single tumour within the upper lobe of the right lung. It is easy to visually verify that each of the three retrieved images also has a single tumour within the upper lobe of the right lung. The tumour within the second retrieved image, however, extends into the mediastinum, and is much larger than the tumour within the third retrieved image. While all three retrieved images are similar, ideally the third image should have been ranked higher than the second image. The results obtained are expected as our method currently has graphs with the same structure, and does not weight any features or relationships. Weighting volume could have resulted in the ideal retrieval.

The query in the second column contains two tumours, one within the lower lobe of the left lung and the other within the mediastinum. The first retrieved image matches these localisations exactly. The second retrieved image has much smaller tumours; according to the clinical reports, one of the tumours is within the left lower lobe while the other is near the left hilum. The tumours in the third image are in the left upper lobe and the left hilum.

These results demonstrate that our 3D graph-based CBIR method is able to find similar 3D PET-CT data from the data set. In addition, the results have also identified ways in which our scheme could be improved. Firstly, every graph could index modality-specific features, such as SUV for PET and

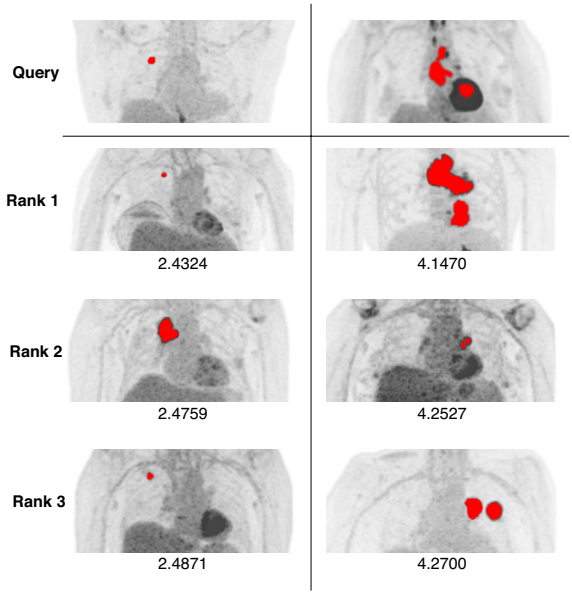

Fig. 3. Example retrievals using our proposed method. Tumours in the PET MIPs have been marked in red. The retrieved images are ranked in order from most similar (1) to less similar (3). The number below a retrieved image is the graph-edit distance (dissimilarity measure).

texture for CT, thereby enabling similarity measurement based on complementary features extracted from each modality.

In addition, the results in Fig. 3 suggest that there are elements of our graph that have a negative influence on the similarity measure. Currently, images with the same number of ROIs have graphs with same structure; all vertices are pairwise connected. This caused the result in the first column of Fig. 3 where the third retrieved image is more similar to the query than the second. In this example, removing edges between unrelated vertices would cause the second retrieved image to have a different structure (the tumour vertex would have edges with the left lung and the mediastinal tissue vertices) compared to the query and other retrieved graphs (the tumour vertex would have only one edge with the left lung vertex). A higher graph edit distance would then be calculated due to edge deletion operations. The final outcome would be that the image would be ranked lower than the third retrieved image (the ideal case, as described earlier).

IV. Conclusion and Future Work

In this study, we presented a graph-based CBIR algorithm that enabled the retrieval of volumetric PET-CT images through the analysis of volumetric ROI and the extraction of 3D spatial features. We demonstrated that our method was able to achieve high levels of retrieval precision on a data set of 50 clinical PET-CT images. In the future we will investigate the inclusion of additional graph features that are specific to a particular modality, such as texture for CT and SUV for PET. Furthermore, our graph currently contains edges that may represent redundant information; we will consider the removal of these edges by pruning the graphs.

Acknowledgment

The authors would like to thank our collaborators at the Royal Prince Alfred Hospital for their direct and indirect contributions to this research.

References

[1] H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content-based image retrieval systems in medical applications–clinical benefits and future directions," *Int. J. of Med. Inform.*, vol. 73, no. 1, pp. 1–23, 2004.

[2] H. Müller, Z. Xin, A. Depeursinge, M. Pitkanen, J. Iavindrasana, and A. Geissbuhler, "Medical visual information retrieval: State of the art and challenges ahead," in *IEEE ICME*, 2007, pp. 683–686.

[3] H. Müller, J. Kalpathy-Cramer, B. Caputo, T. Syeda-Mahmood, and F. Wang, *Overview of the First Workshop on Medical ContentBased Retrieval for Clinical Decision Support at MICCAI 2009*, 2009, vol. 5853, pp. 1–17.

[4] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. 22, no. 12, pp. 1349–1380, 2000.

[5] T. M. Blodgett, C. C. Meltzer, and D. W. Townsend, "PET/CT: Form and function," *Radiol.*, vol. 242, no. 2, pp. 360–385, 2007.

[6] J. Kim, L. Constantinescu, W. Cai, and D. D. Feng, "Content-based dual-modality biomedical data retrieval using co-aligned functional and anatomical features," in *Workshop on Content-based Image Retrieval for Biomedical Image Archives: achievements, problems and prospects, MICCAI*, 2007, pp. 45–52.

[7] Y. Song, C. Cai, Weidong, S. Eberl, M. J. Fulham, and F. Dagan, "A content-based image retrieval framework for multi-modality lung images," in *IEEE Symp. CBMS*, 2010, pp. 285–290.

[8] Y. Song, W. Cai, S. Eberl, M. J. Fulham, and F. Dagan, "Thoracic image case retrieval with spatial and contextual information," in *IEEE ISBI*, 2011, pp. 1885–1888.

[9] A. Kumar, J. Kim, W. Cai, S. Eberl, and D. Feng, "A graph-based approach to the retrieval of dual-modality biomedical images using spatial relationships," in *IEEE EMBC*, 2008, pp. 390–393.

[10] E. G. M. Petrakis, "Design and evaluation of spatial similarity approaches for image retrieval," *Imag. Vis. Comp.*, vol. 20, no. 1, pp. 59–76, 2002.

[11] E. G. M. Petrakis, C. Faloutsos, and K. I. Lin, "ImageMap: an image indexing method based on spatial similarity," *IEEE Trans. Knowl. Data Eng.*, vol. 14, no. 5, pp. 979–987, 2002.

[12] D. Conte, P. Foggia, C. Sansone, and M. Vento, "Thirty years of graph matching in pattern recognition," *Int. J. of Patt. Rec. and Art. Intel.*, vol. 18, no. 3, pp. 265–298, 2004.

[13] F. C. Detterbeck, "The new cancer staging system," *Chest*, vol. 136, no. 1, pp. 260–71, 2009.

[14] S. Hu, E. A. Hoffman, and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images," *IEEE Trans. Med. Imag.*, vol. 20, no. 6, pp. 490–498, 2001.

[15] J. Bradley, W. L. Thorstad, S. Mutic, T. R. Miller, F. Dehdashti, B. A. Siegel, W. Bosch, and R. J. Bertrand, "Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer," *Int. J. of Rad. Onc.*Biol.*Phys.*, vol. 59, no. 1, pp. 78–86, 2004.

[16] S. Aksoy and R. M. Haralick, "Feature normalization and likelihood-based similarity measures for image retrieval," *Patt. Rec. Let.*, vol. 22, no. 5, pp. 563–582, 2001.

[17] M. Neuhaus, K. Riesen, and H. Bunke, "Fast suboptimal algorithms for the computation of graph edit distance," *Struct. Synt. Stat. Patt. Rec.*, vol. 4109, pp. 163–172, 2006.

[18] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Trans. Sys. Sci. Cyber.*, vol. 4, no. 2, pp. 100–107, 1968.

[19] C.-R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. M. Aisen, and L. S. Broderick, "ASSERT: a physician-in-the-loop content-based retrieval system for HRCT image databases," *Comput. Vis. Image Underst.*, vol. 75, no. 1-2, pp. 111–132, 1999.

[20] G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, and C. Roux, "Wavelet optimization for content-based image retrieval in medical databases," *Med. Imag. Anal.*, vol. 14, no. 2, pp. 227–241, 2010.