Multi-object Segmentation for Abdominal CT Image Based on Visual Patch Classification

Pan Li, Jun Feng(✉), QiRong Bu, Feihong Liu, and HongYu Wang

School of Information & Technology, Northwest University Xi'an, Xi'an, China
fengjun@nwu.edu.cn

Abstract. We introduce a robust multi-object segmentation algorithm based on visual patch classification for abdominal CT segmentation. Firstly, the proximity of the pixels is expressed by both intensity and spatial distance. And then clustering framework is employed to form various visual patches. In this way, the noise and embedded small tissues such as blood vessels and tracheas which often make other segmentation algorithms failed are filtered out during the cluster iteration. Afterwards, the visual patches are further grouped by the way of classification in the criteria of spatial relationship of visual pitches. Specially, the algorithm can be viewed as effectively tradeoff of bottom-up methods and top-down methods. The approach has been applied to the multi-object segmentation of abdominal CT images, such as the liver, kidney, spleen and gallbladder. We have test the method in American published TCIA database, whose efficiency and robustness is evaluated through quantification results on both sectional level and volumetric level, which exhibit the optimistic application and prospect in the field of medical image processing.

Keywords: Abdominal CT image · Segmentation · Visual pitch · Classification · Spatial relationship

1 Introduction

Medical image segmentation is the first phase of medical imaging data analysis and visualization, which are also the precondition and crucial to computer-aided diagnosis, image-guided surgery, virtual endoscopy and many medical image applications [1]. Compared with osseous organs, human abdominal soft-tissues are more complex and deformable, including the liver, the kidney, the gallbladder, the spleen as well as vascular such as the veins and arteries. However, due to the limitation of imaging device and the peristalsis of tissues, they often exhibit intensity inhomogeneity and overlapping in abdominal CT series. Besides, the blurred organs and the ambiguous of the edge of lesions also bring some considerable difficulties for segmentation. And for abdominal multiple organs segmentation, there will be much more influential factors, including high similarity of adjacent organs, partial volume effects and the relatively high variations of organ position and shape. Therefore, multi-object segmentation in abdominal CT image is still a challenge task [2].

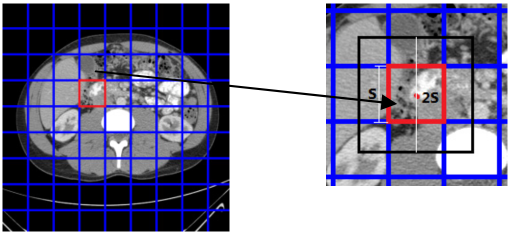
H. Zha et al. (Eds.): CCCV 2015, Part II, CCIS 547, pp. 130–138, 2015.
DOI: 10.1007/978-3-662-48570-5_13

Multi-object segmentation for abdominal CT image has become a research tendency, which has already reached some achievements. Daniel Freedman's algorithm, which is highly depend on learned shape and appearance models, compares the probability distributions instead of computing a pixelwise correspondence between the model and the image [3]. Robin Wolz et al. presented a multi-organ abdominal segmentation method based on a hierarchical atlas registration and weighting scheme that generates target specific priors from an atlas database, which finally obtain the segment result by applying an automatically learned intensity model [4]. Toshiyuki Okada et al proposed a method for finding and representing the interrelations based on canonical correlation analysis, which is developed for constructing and utilizing the statistical atlas [5]. Although the above two approaches are able to capture the organ location and appearance, it both relies on a subject-specific atlas model which impacted by inter-subject variability. Yinxiao Liu has developed an automatic threshold and gradient strength selection algorithm for unknown number of abdominal object regions by combining class uncertainty and spatial image gradient features, which is only in the view of mathematic and does not incorporate the spatial knowledge into segmentation[6]. To address the problem of spatial relationship among multiple organs of abdomen, Xiaofeng Liu et al extended the MAP framework by modeling the inter-organ spatial relations using a minimum volume overlap constraint, which focused on the posteriori probability and volume overlap without of characteristics of the organ themselves[7]. Besides, most of the traditional classification methods are put forward to single object with the ignorance of spatial information [8,9].

In this paper, we introduce a classification method for abdominal multi-object segmentation based on visual patch, which explicitly incorporate the spatial relationship of abdominal organs into classification. We introduce the definition of visual pitch, which is proved effective in color image pre-segmentation called superpixel [10], into the medical image segmentation. This pre-segmentation technology divides the image into several visual pitches by clustering the pixels with intensity similarity and spatial proximity, which are regarded as the unit of classification. Besides, through a large number of experiments, we discover an implicit spatial regularity among the visual patches of abdominal organs, which is implemented by establishing undirected adjacency graph of visual patches and finally integrated into classifier. In order to reducing the impact of similar intensity of different organs, we also consider the texture of visual patch into classification. Thus, the algorithm we proposed is the combination of low-level visual features and spatial physical characteristics, which also can be viewed as effectively tradeoff of bottom-up methods-clustering and top-down methods-classification. Experiments demonstrate that our method achieves the comparable performance with the state-of-art algorithms.

2 Visual Patches Generation Based on Superpixel

Here, we propose an unsupervised method for medical image, which generates the visual patch by clustering pixels based on both intensity similarity and spatial proximity. According to the complexity of image, there should be a K as the initial number of cluster center, and approximately equally divided the image into rectangle pitches shown in Figure 1(a). If the image is $N*M$, each initial visual patch is assigned about $N*M/K$ pixels. S is defined as the side-length of visual pitch, which is calculate by $S=\sqrt{N*M/K}$. The cluster region ranges $2S*2S$ around the center, which is shown in Figure 1(b).



(a) Initial visual patches (b) Cluster region

Fig. 1. Initial visual pitches

Intensity distance in grayscale space is perceptually effective for small distances, but it is no longer working when the space perception of the pixel exceeds the limit of intensity distance. Thus, we use the following measure instead of Euclidean distance:

$$d_{xy}=\sqrt{(x_i-x_j)^2+(y_i-y_j)^2} \tag{1}$$

$$d_g=\left|g_i-g_j\right| \tag{2}$$

$$D_s=\frac{\mu}{S}d_{xy}+d_{Gray} \tag{3}$$

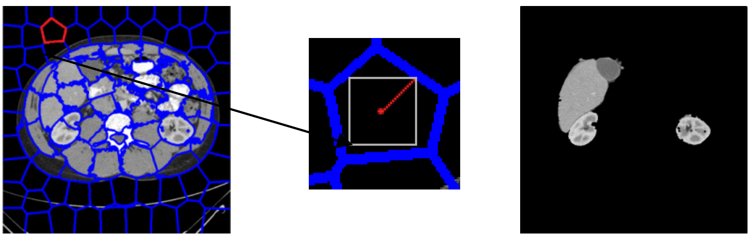
xy is the pixel spatial position, and g is the intensity value of pixel xy in grayscale image. Spatial proximity d_{xy} is calculated by Euclidean distance in two-dimension of image plane, and intensity distance d_g is regarded as feather similarity. D_s is the distance of the two pixels. μ is the parameter of the pixel compactness. The greater it is the spatial proximity is more important, and the clusters are more closely. Therefore, μ effectively balances the spatial proximity in grayscale space. It is known to us all that the standard deviation is an important method to measure the dispersion of numerical data, which also weighs the volatility of samples. Thus, we take the standard deviation of gray feature as the parameter μ which controls the compactness of pixels.

After above, select the minimum gradient pixel of each visual patch as the initial cluster center, which will avoids the impact of the boundary and noisy pixels. And then, use D_s incorporated the intensity similarity and spatial proximity as the clustering condition to cluster the pixels. Figure 2(a) shows the visual patches of the cluster result.

3 Multi-object Segmentation Based on Visual Patches Classification with Spatial Relationship

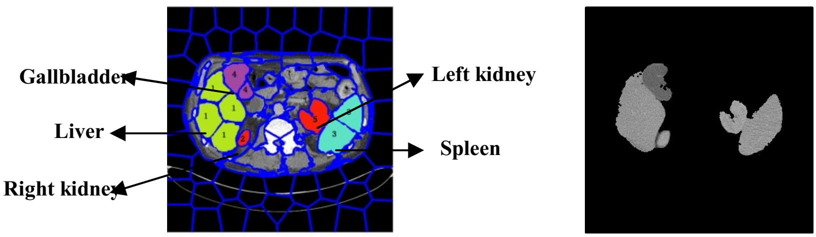
Classification is the essential methods of statistical analysis in the field of pattern recognition, which utilizes a known training sample to achieve segmentation purposes in the feature space of image. Without considering the spatial information, traditional classification segmentation algorithms are not sensitive to inhomogeneous intensity images, which will lead to large error in segmentation. There, we propose a method that exploits the supervised learning classification algorithm based on visual patches to segment the abdominal CT image to achieve the multi-object regions, in which we take the spatial relationship into consideration implemented by establishing an undirected adjacency graph of visual patches. In this way, even though the visual patches, which belong to the different organs, have the similar feathers, we will classify the patches exactly by the mutual positional relationship of organs.

Above all, the prior knowledge of the supervised classification algorithm in our method is the texture and the intensity of visual patches. Thus, it is crucial to extract the feather of the irregular patches. We put forward a measure that looking for the point in the edge of visual patch which is nearest to the cluster center of the patch, and take the distance as the half of the diagonal of rectangle, whose focus is the cluster center, and finally obtain the square shown in Figure 2(b). The gray level co-occurrence matrix method is exploited to extract the visual patches textural feathers [11]. Meanwhile, the intensity of visual patch is the average of all pixels' in patch.



(a) Visual patches (b)Texture region (c) Classification segmentation

Fig. 2. The visualization result of the steps



(a) Undirected adjacency graph of visual pitches (b) Classification segmentation

Fig. 3. The visualization result of the undirected adjacency graph

Through a large number of experiments, we discover an implicit spatial regularity among the visual patches of abdominal organs. First, we sign the visual pitches of the liver, right kidney, spleen, gallbladder, left kidney and background as a label (1,2,3,4,5,0). Then, we find that the visual pitches marked the same label are connected with each other. In addition to, the right kidney is close to the liver, and so as the gallbladder. And the spleen is near to the left kidney, both of which far away from the liver. Equation (5) shows the close and far distance calculated by equation (4), which are taken as the constraint condition of the classification. According to [5,7], liver and spleen is stable and easy to be located. So liver and spleen will be calculated more accurate. Besides, Right kidney and gallbladder are decided by liver. Left kidney is related to spleen. Then, we establish an directed adjacency graph of visual pitches shown in Figure 3(a), and define the distance calculated by equation (4) as the weight of two visual pitches to implement the spatial relationship, which also contains the location information.

$$d_{pitch} = \|P_1 - P_2\|^2 \tag{4}$$

$$\begin{cases} d_{close} < 2S \\ d_{far} > 3S \end{cases} \tag{5}$$

d_{pitch} is the distance of two visual pitches. P is the position (x,y) of the center pixel of the visual pitch. $\|.\|$ is L_2 norm. The side-length of the cluster region is 2S.

When constructing the training sample set for classification, which contains the abdominal organs like liver, kidney, gallbladder, spleen as well as the backgrounds, we should abandon the visual pitches whose intensity is 0 because of the black background of the medical image, such as the red pitch in Figure 2(a). When all the preparations are ready, it is time to start the classification stage, which results in labels of the visual patches. Finally, we utilize the spatial constraints to revise the label of patches. The result of classification segmentation based on visual patch is shown in Figure 2(c) and Figure 3(b), which are filled and isolated.

4 Experimental Results

4.1 Experimental Results of Visual Patches Generation

In order to verify the good performance of visual patches generation, we compare the result with ground truth and the traditional clustering methods in American TCIA database. Experiment shows that the result of the visual patches generation segmentation is closest to the ground truth than the other state-of-the-art algorithms, which is shown in Figure 6. Figure 4 is the result of the visual patches generation and Figure 6(d) is the regions of right kidney which is filled and isolated. DP method took advantage of the Bayesian methodologies and Markov chain to enhance model flexibility, which ignores the intensity overlap and spatial relationship [8]. Although C-GMMs algorithm made use of the powerful probabilistic statistical theory, it also spent time with the convergence function, which was proposed for image segmentation using the feather function to estimate the parameters of GMMs in each iterative [9]. Moreover, when compared to the traditional clustering segmentation methods, our algorithm select the lowest gradient pixel as the initial cluster centers of each visual patch, which no longer rely on the random and manual selection, and largely avoid the effects of noise and intensity inhomogeneity. It also considers the spatial relationship of pixels during the clustering, which inhibits the phenomenon of under-segmentation and over-segmentation significantly, and at the same time retains all of the information of the target area in the original image. Therefore, visual patches generation provides a good basis for clinical medicine diagnosis.

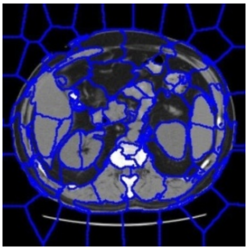


Fig. 4. Visual patches generation

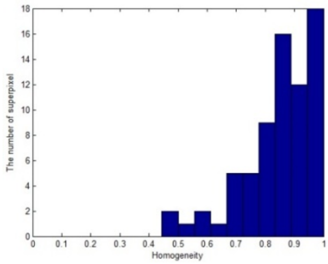
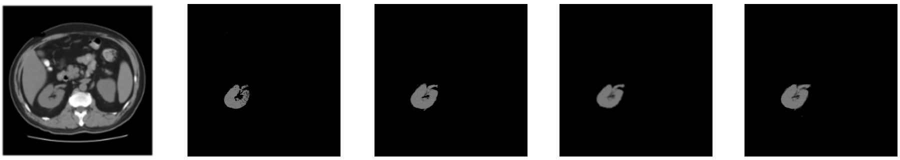


Fig. 5. Homogemeity



(a) Original image (b) DP[8] (c) C-GMM[9] (d) Visual patch (e) Ground turth

Fig. 6. Results of kidneys segmentation by visual patch generation

Besides, we put forward a new evaluate criteria for visual pitch. As known to us all, homogemeity is a measure of the change of texture, and the greater its value the texture is more uniform. Due to the homogemeity ranges from 0 to 1, we can calculate from Figure 5 that almost 80% of the homogemeity of visual pitches, regardless of the black background patches, are more than 0.8, and the rest are resulting from the uneven background clustering by spatial proximity. From this we can also prove the good performance of visual patches when used in small feather space of medical image.

4.2 Result of Multi-object Classification Segmentation

Traditional classification segmentations rely on the whole organ region feathers, whereas the method in this paper regards the organ patch as the classification unit. From the visual patch generation we can achieve several kinds of patches, each of which is likely to be a small part of the organs. Thus, it is easy to be high quality feathers without background, and the classification will be benefit from them. During the experiment, the method is validated on 200 CT sequence totally about 14000 slices which are opened in American TCIA database. Figure 7 shows the average classification accuracy of each organ. As the number of sample growing, the accuracy is gradually increasing and finally stabilized. Because the frequency of occurrence of gallbladder and spleen is lower than the liver and kidney in abdominal CT series, and their change is relatively large, the segmentation accuracy of them is lower. Table 1 shows the result of the segmentation of organs compared to the state-of-art method tested in American TCIA database, which is calculated by equation (6). T is the number of is the true positive pixels, and F is the number of the false negative pixels.

$$Accuracy = \frac{\mathrm{T}}{\mathrm{T+F}} \tag{6}$$

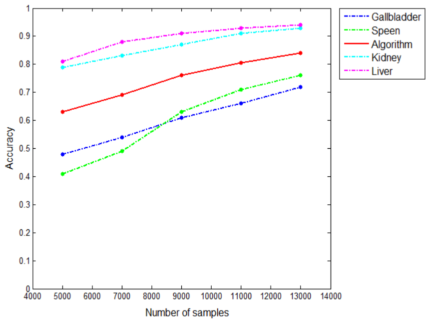
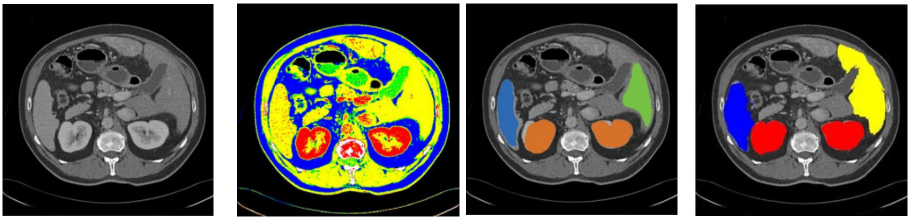


Fig. 7. The accuracy of our method.

Organs /Method	TG [6]	LVP [7]	Our method
Liver	0.80	0.89	0.92
Spleen	0.55	0.72	0.78
Kidney	0.74	0.82	0.90
Gallbladder	0.38	0.56	0.73

Table 1. The segment accuracy of each organ

From Table 1 we can see that classification based on visual pitch considering the spatial relationship of visual pitches is proved to be an excellent measure for segmentation. TG [6] using the threshold and gradient is inevitably affected by noise and intensity inhomogeneity. Besides, it causes over-segmentation with the intensity as the only feather when happened to grayscale overlapping. Although the LPV [7] takes the spatial relationship into consideration, it is highly depend on the inter-organ modeling by a minimum volume overlap constraint, which does not apply to individual diversity and also result in over- and under-segmentation. Our multi-object segmentation method, which combines the image low-level features and the spatial location information, is the fusion of visual features and physical characteristics. We test these methods in American published TCIA database, and the performance of our method is proved higher than the other state-of-art algorithms. Figure 8 shows the visualization result of the comparative algorithms.



(a)Original image (b)TG[6] (c)LPV[7] (d)Our method

Fig. 8. Results of comparative algorithms segmentation

5 Conclusion

The method we presented is the combination of top-down method clustering and bottom-up method classification. In the visual patches generation phase, the pixels are clustered by intensity similarity and spatial proximity, which finally obtains the compact and uniform visual patches and inhibits the phenomenon of under-segmentation and over-segmentation. At next stage, with the difference of the traditional classification segmentation algorithms, our method regards the visual patch as classification unit instead of the pixels, which improves the speed of the segmentation, and takes the spatial relationship of visual patches into consideration to increase the accuracy of the classification. Therefore, multi-object segmentation based on visual patches classification is not only designed for multi-object segmentation, but also greatly increases the accuracy of segmentation. In conclusion, our image segmentation techniques have improved the segmentation accuracy segmentation speed, quality, multi-object and general applicability.

References

1. Withey, D.J., Koles, Z.J.: A Review of Medical Image Segmentation: Methods and Available Software. International Journal of Bioelectromagnetism **10**(3) (2008)
2. Selver, M.A., Kocaoglu, A., Akyar, H., Dicle, O., Guzelis, C.: Patient oriented neural networks to overcome challenges of abdominal organ segmentation in CT angiography studies. Electrical and Electronics Engineering, pp. 5–8, November 2009
3. Freedman, D., Radke, R.J., Zhang, T., Jeong, Y., Chen, G.T.Y.: Model-based multi-object segmentation via distribution matching. In: Computer Vision and Pattern Recognition Workshop (2004)
4. Wolz, R., Chu, C., Misawa, K., Fujiwara, M., Mori, K., Rueckert, D.: Automated Abdominal Multi-Organ Segmentation With Subject-Specific Atlas Generation. Medicine and Biology Society, 1723–1730 (2013). doi:10.1109
5. Okada, T., Linguraru, M.G., Hori, M., Suzuki, Y., Summers, R.M., Tomiyama, N., Sato, Y.: Multi-organ segmentation in abdominal CT images. doi:10.1109/EMBC.2012
6. Liu, Y., Liang, G., Saha, P.K.: A new multi-object image thresholding method based on correlation between object class uncertainty and intensity gradient. Medical Physics, January 2012

7. Liu, X., Linguraru, M.G., Yao, J., Summers, R.M.: Abdominal multi-organ localization on contrast-enhanced CT based on maximum a posteriori probability and minimum volume overlap. Biomedical Imaging (2011). doi:10.1109
8. Adelino, R.F.D.S.: Bayesian mixture models of variable dimension for image segmentation. Compute. Meth. Prog. Biomed. **94**, 1–14 (2009)
9. Luo, S., Hu, Q., He, X., Li, J., Jin, J.S., Park, M.: Automatic liver parenchyma segmentation from abdominal CT images using support vector machines. Complex Medical Engineering, 1–5 (2010). doi:10.1109
10. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: Slic superpixels compared to state-of-the-art superpixel methods. Pattern Analysis and Machine Intelligence (2012)
11. Bo, H., Ma, F.L., Jiao, L.C.: GLCM of Image Texture. Acta Electronica Sinica 01, 155–158 (2006)