Active Scheduling of Organ Detection and Segmentation in Whole-Body Medical Images

Yiqiang Zhan, Xiang Sean Zhou, Zhigang Peng, and Arun Krishnan

Siemens Medical Solutions USA, Inc., Malvern, PA 19355

Abstract. With the advance of whole-body medical imaging technologies, computer aided detection/diagnosis (CAD) is being scaled up to deal with multiple organs or anatomical structures simultaneously. Multiple tasks (organ detection/segmentation) in a CAD system are often highly dependent due to the anatomical context within a human body. In this paper, we propose a method to schedule multi-organ detection/segmentation based on information theory. The central idea is to schedule tasks in an order that each operation achieves maximum expected information gain. The scheduling rule is formulated to embed two intuitive principles: (1) a task with higher confidence tends to be scheduled earlier; (2) a task with higher predictive power for other tasks tends to be scheduled earlier. More specifically, task dependency is modeled by conditional probability; the outcome of each task is assumed to be probabilistic as well; and the scheduling criterion is based on the reduction of the summed conditional entropy over all tasks. The validation is carried out on two challenging CAD problems, multi-organ detection in whole-body CT and liver segmentation in PET-CT. Compared to unscheduled and *ad hoc* scheduled organ detection/segmentation, our scheduled execution achieves higher accuracy with faster speed.

1 Introduction

Whole-body Computed Tomography (CT), Positron Emission Tomography (PET) and Magnetic Resonance (MR) scanning is being accepted for various clinical applications across different organs, *e.g.*, assessment of cancer metastasis in lymph nodes [1] or bones [2], evaluation of the extent and distribution of polymyositis [3], and detection of ankylosing spondylitis [4]. While whole-body scans reveal comprehensive anatomical/functional information of human body, the vast amount of image data makes the detection of potential disease burdensome. Accordingly, computer aided detection/diagnosis (CAD) at a whole body level becomes more desirable to provide useful "second opinions" to radiologists. In practical clinical applications, whole-body CAD is required to detect/segment multiple organs or anatomical structures in limited time. (For simplicity, we denote organ detection/segmentation as "*task*" in the remainder of this paper.) Due to the anatomical context within a human body, the dependency between tasks can be exploited to increase the efficiency and performance of CAD systems. For example, the relatively easier task of femoral head localization in CT
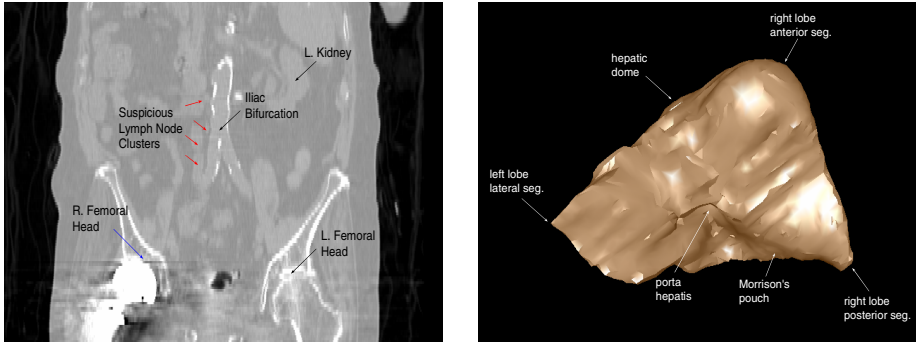
Fig. 1. (Left) A representative abdominal CT image. The blue arrow points to an artificial metal "femoral head". The red arrows point to suspicious abdominal lymph node clusters, which are often close to the iliac bifurcation of the aorta. (Right) Liver model with critical anatomical landmarks.

(bone is very bright in CT) will facilitate a quick and accurate localization of the iliac bifurcation of the aorta, which in turn greatly help the detection and identification of abdominal lymph node clusters .

One way to principally exploit such dependency is to model it as a *scheduling problem*. Due to the unique nature of human anatomy, the scheduling problem of whole-body CAD has three characteristics. *First*, the scheduling problem is highly flexible. The accuracy and speed of whole-body CAD, however, is significantly different with different schedules. *Second*, due to missing data, artifacts or diseases, the scheduler of whole-body CAD must be an *active* one. In other words, the scheduling must be adaptive to the specific patient data at the *runtime*. Refer to the previous example, in general cases, the detector of iliac bifurcation should be fired next to the "femoral head localization". However, as shown in Fig. 1, for a patient who has an *artificial* metal femoral head, the femoral head detector might not detect it correctly. In this situation, instead of firing the "iliac bifurcation detector", the scheduler should trigger the detectors of other organs, *e.g.*, kidneys, which can be localized accurately without the inference of femoral heads. *Third*, the probabilistic factors influence organ detection/segmentation in two aspects: (1) tasks are often statistically dependent, as the relative locations of organs are not deterministic; and (2) the outcome of tasks usually embeds uncertainties.

Although scheduling problems have been extensively studied in different research areas (A brief review will be presented in section 2.), the existing methods can not be directly borrowed to schedule organ detection/segmentation in whole-body CAD, due to the aforementioned unique characteristics. In this paper, we propose to study the scheduling problem of whole-body CAD from an information theoretic view. In this framework, tasks are modeled as a set of *measurements* that aim to achieve the diagnostic information from medical images. (In this paper, "task", "measurement" and "operation" share the same meaning by default.) The principle is to schedule tasks in an order that is optimal in an

information-theoretic sense. More specifically, we explore the gauge of *informa-tion gain* to define the scheduling criterion. Based on this criterion, a sequential decision making process is employed to schedule tasks in whole-body CAD.

2   Related Work

In the last several decades, the topics of scheduling have been extensively stud-ied in the areas of operation research [5] and theoretical computer science [6]. Many scheduling rules/methods were proposed to deal with scheduling problems in various applications, including manufacturing, service industries, transporta-tion and practical computer systems, *etc.*. While earlier studies mainly focus on deterministic systems, more researchers move to flexible and stochastic systems recently. In [7], the scheduling policies for flexible systems are investigated. That paper analyzes an open processing network model with discretionary routing and showed, in general, unbalanced workload routing with priority sequencing gives better performance than a balanced one. In [8], Chou *et al.* studied a stochas-tic single machine problem, where the actual processing time of tasks are not known until processing is completed. They proved that when task processing times are mutually independent random variables, weighted shortest expected processing time among available jobs heuristic is asymptotically optimal for the single-machine problem. Although these scheduling problems share one or sev-eral features with that of whole-body CAD, neither of them account all the aforementioned characteristics of whole-body CAD.

3   Method

3.1   Problem Statement

From an information theoretic view, the scheduling of whole-body CAD is akin to an extensively studied topic in computer vision, *active object recognition*. Recall the previous example, the whole-body CAD aims to gain diagnostic *information* (Do the abdominal lymph node clusters exist? Where are they?) through a set of *measurements* (femoral heads localization, iliac bifurcation localization and lymph node clusters detection). It is in analogy to active object recognition, which aims to identify objects (*information*) by collecting pictures with dif-ferent sensor parameters (*measurements*). Indeed, information theory has been successfully employed in active object recognition. In Denzlers *et al.*'s pioneer work[9], an information theoretic formalism is proposed to select optimal sensor parameter during iterative state estimation. The benefits of the method were demonstrated in an object recognition application using an active camera. Al-though this method is not an off-shelf method to schedule tasks in whole-body CAD, it inspired us to consider our problem from an information theoretic way.

A whole-body CAD system aims to obtain diagnostic information, by execut-ing a set of organ detection/segmentation. The diagnostic information is pre-sented by a set of variables $\{x_i\}$, *e.g.*, the locations of the organs under study.

Each task (organ detection/segmentation) in whole-body CAD actually delivers a measurement y_j to decrease the ambiguity of $\{x_i\}$. Notably, y_j usually belongs to $\{x_i\}$. In a more general sense, however, y_j can be out of $\{x_i\}$. For example, y_j can be a variable representing the center of a template with multiple organs, which is not interested by the CAD system but is helpful to organ localization.

From an information theoretic view, the optimal schedule becomes a task sequence that maximally gains diagnostic information, *i.e.*, maximally decreases the uncertainty of $\{x_i\}$.) Therefore, we model the scheduling problem as a sequential decision making process. (Although the sequential process might not obtain global optimal solution, it has run-time efficiency that is important to CAD.) At each step, the decision is: "Given the current measurements, what is the next y_j, upon measurement, gains most diagnostic information?"

3.2   Scheduling Criterion Based on Information Gain

Assume that $\{x_i\}$ are the variables of interest for a CAD system. The distribution of y_j is Ψ prior to the measurement process. After the measurement, its distribution shrinks, or *changes* in general, to Φ. According to the information theory [10], the information gain after this particular measurement of y_j is:

$$IG_{y_j} = \sum_i \left(H(x_i|y_j \in \Psi) - \int_{y_j \in \Psi} H(x_i|y_j \in \Phi)p(y_j)dy_j \right) \qquad (1)$$

Here we use the expression $y_j \in \Psi$ to mean "y_j has the support Ψ" or "y_j has the distribution Ψ". And $H(x_i|y_j \in \Psi)$ and $H(x_i|y_j \in \Phi)$ are conditional entropies defined in the following form:

$$H(x_i|y_j \in \Phi) = -\int_{y_j \in \Phi} p(y_j) \int_{x_i \in X_i} H(x_i|y_j)dx_i dy_j \qquad (2)$$

$$= -\int_{y_j \in \Phi} p(y_j) \int_{x_i \in X_i} p(x_i|y_j) \log p(x_i|y_j)dx_i dy_j \qquad (3)$$

If y_j is taken from the set $\{x_i\}$, the first term in Eq. (1) goes away because it becomes constant for all y_j. In general, however, we can have y_j's outside of $\{x_i\}$. Then, Eq. (1) is meaningful in its complete form.

The basic principle of our *IG*-based scheduling rule is that a particular measurement operation $y_{j*} \in \{y_j\}$ will be preferred over others if it delivers a maximal value for *IG*. The justification behind this principle is described as follows. According to Eq. (1), information gain is determined by three factors: (1) the support of y_j before measurement, Ψ, (2) the measurement uncertainty of y_j, Φ, and (3) the dependency between y_j and $\{x_i\}$, $p(x_i|y_j)$. Indeed, it is the interplay of all these three factors that determine the speed and the accuracy of multiple organ detection/segmentation. Importantly, since information gain embeds all probabilistic factors that influence the speed and performance of whole-body CAD, the *IG*-based scheduling method is expected to achieve better performance than *ad hoc* strategies, such as "pick the most confident operation first"
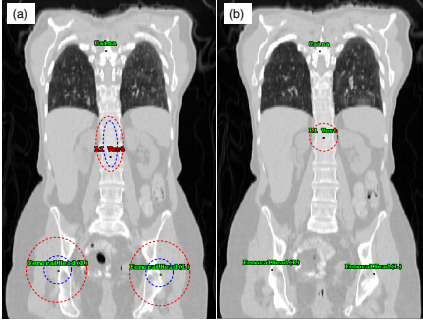
Fig. 2. An example to show the effectiveness of the proposed scheduling method in multi-organ localization. For illustration purpose, detected results are projected to the same coronal view.
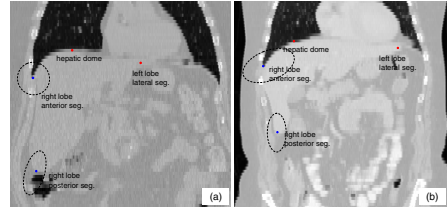

Fig. 3. Active scheduling using the information gain-based scheduling method. The red and blue points denote the detected and undetected landmarks, respectively. The black dashed ellipses denote the support of the prior distribution of the landmark location before measurement. Results are projected to the same coronal view.

or "pick the task that other tasks mostly depend on". In Fig. 2, we present a rather simplified but intuitive example to show the effectiveness of the IG-based scheduling criterion. In this example, the system aims to localize four organs: carina of trachea, left femoral head, right femoral head and L1 vertebra. The dependency between different organ localization is modeled by the relative spatial locations between different organs. More specifically, the positions of the localized organs are used to estimate the positions of the remaining ones to reduce the search range of organ localizers. Let us assume the carina of trachea has been localized. As shown in Fig.2 (a), the estimated position of the L1 vertebra has the minimum support (denoted by the red dashed ellipses). If we use an *ad hoc* schedule strategy that prefers the task having the minimum support, the next organ to be localized should be the L1 vertebra. However, since the neighboring anatomical structures, *e.g.*, the L2 and the T12 vertebra, usually have similar appearance as L1, the L1 localizer has large uncertainty in the vertical direction (Gaussian-fitted uncertainty is denoted by the blue dashed ellipses in Fig.2 (a)) and gets the wrong result (denoted by the red point in Fig. 2 (a)). In other words, the "uncertain" measurement of L1 is not expected to deliver large information gain. According to our IG-based scheduling criterion, instead, the two femoral heads that have stronger "shrink" from Ψ to Φ are preferred as the next organs to be localized. (The supports of Ψ and Φ in Eq. (1) are defined by the red and blue dashed ellipses in Fig. 2 (a), respectively.) After localizing the two femoral heads, the support of the "un-measured" L1 vertebra is significantly reduced (denoted by the red dashed ellipses in Fig.2 (b)) and the localizer is able to successfully localize it (denoted by the green point in Fig. 2 (b)) without being confused by L2 or T12 vertebra.

To effectively evaluate Eq. (1), we employ a Monte-Carlo simulation method. Instead of estimating the conditional probability density function, we sample the conditional entropy using $p.d.f$s that describes detection/segmentation uncertainty and probabilistic relations between different tasks. As all of these $p.d.f$s can be learned off-line, the Monte Carlo simulation method has run-time efficiency in evaluating information gain.

4   Results

We start the validation from a relatively uniform task set, the localization of multiple organs in whole-body CT images. In this system, each organ is localized by a generic, learning-based localizer. (The learning-based localizer is a 3D extension of Viola and Jones's detection method [11] with expanded feature sets. The details of the localizer are omitted due to the limited space.) The dependency between organ localization is modeled by the spatial relations between different organs. More specifically, the positions of the localized organs are used to estimate the positions of the remaining ones to reduce the search range of organ localizers. The uncertainty of the organ localizers ($\Psi(.)$ in Eq. (1)), and the spatial relations between different organs are learned from 40 training samples. Multi-variant Gaussian distribution is used to model the estimation and localization uncertainty.

The experiment is carried out to localize of six organs (carina of trachea, L1 vertebra, left kidney, right kidney, left femoral head, right femoral head) from 18 whole-body CT scans (resolution:$0.927mm \times 0.927mm \times 2.5mm$). We compare the speed and the accuracy using three different scheduling methods: (1) Unscheduled independent organ localization, (2) *Ad hoc* scheduled organ localization (The *ad hoc* scheduling rule prefers the organ whose location is most correlated with other organs.), and (3) *IG*-based scheduled organ localization. As shown in Fig. 4 and Table 1, our method achieves the best accuracy with the fast speed.

In the second set of experiments, our scheduling method is carried onto another CAD problem: liver segmentation in PET-CT scans. To initialize the deformable model, our method detects seven anatomical landmarks as shown in Fig. 1). Here, each task is the detection of an individual landmark. Again, the

Table 1. Quantitative comparison of organ localization error

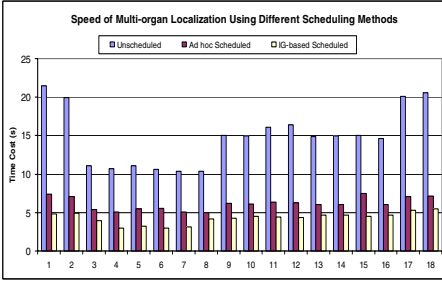|	Unscheduled organ localization		*Ad hoc* Scheduled organ localization		*IG*-based Scheduled organ localization	
	Avg. Err. (mm)	Max Err. (mm)	Avg. Err. (mm)	Max Err. (mm)	Avg. Err. (mm)	Max Err. (mm)
Trachea Carina	1.97	4.20	2.02	7.14	1.97	4.20
Femoral Head	4.47	10.40	4.67	11.08	4.60	9.96
Kidney	9.98	19.66	9.15	21.09	8.97	19.00
L1 Vertebra	**5.58**	**36.00**	**5.47**	**36.85**	**3.37**	**7.03**

Fig. 4. Comparison of the speed of multi-organ localization using different scheduling methods (on 18 whole-body CT scans)
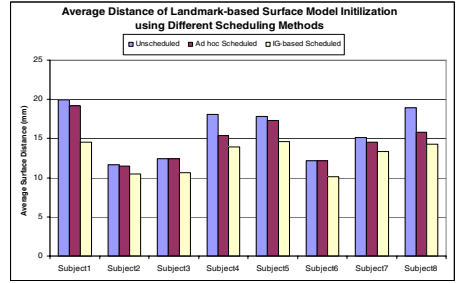
Fig. 5. Quantitative comparison of the landmark-based surface initialization error (on 8 PET-CT cases) using different scheduling methods

generic learning-based landmark detectors and the modeling of the dependency between different tasks are similar to the first experiment.

The experiment is carried on 40 whole-body CT scans (CT resolution: $1.36mm \times 1.36mm \times 5mm$, PET resolution: $5.3mm \times 5.3mm \times 5mm$). We found the *IG*-based scheduling method is actually adaptive to image data. As shown in Fig. 3, the relative locations between the *detected* hepatic dome and left lobe lateral seg. are different in two scans. Based on the prediction from two detected landmarks, the estimated position (prior distribution) of right lobe posterior seg. has more compact support in case 1 (Fig. 3a). "Localization of right lobe anterior seg." is thus expected to achieve less information gain in case 2 than in case 1. Therefore, while the detection order in case 1 is "...→**right lobe anterior seg.→right lobe posterior seg.**", the detection order in case 2 becomes "...→**right lobe posterior seg.→right lobe anterior seg.**". The quantitative comparison is shown in Fig. 4 and 6. Again, our method is superior to other two in terms of detection accuracy and speed.
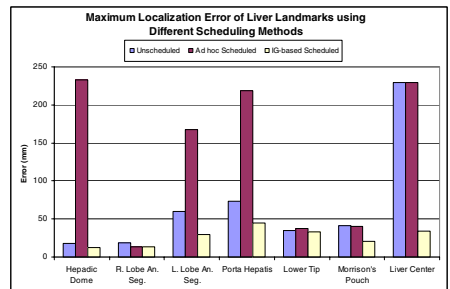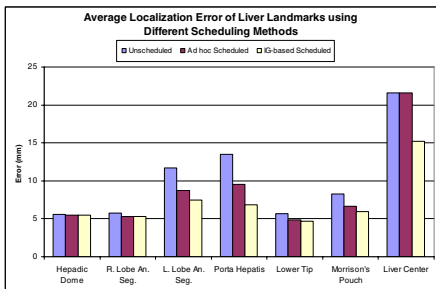

Fig. 6. Quantitative comparison of the localization of multiple liver landmarks using different scheduling methods. (Left) Average landmark detection error of 40 PET-CT cases. (Right) Maximum landmark detection error of 40 PET-CT cases.

5   Conclusions

In this paper, we proposed a rigorous formula to address the scheduling problem of multi-organ detection/segmentation in whole-body CAD. The key idea is to schedule tasks in an order that each operation achieves maximum expected information gain over all the tasks. Our method has two major advantages in scheduling multiple organ detection/segmentation. *First*, various probabilistic factors that influence the performance and speed of whole-body CAD are incorporated in the scheduling criterion. Therefore, the scheduled system is able to achieve more accurate results with less computational cost. *Second*, in our scheduling method, the *next* task is always determined based on current system status. Accordingly, the whole-body CAD is scheduled in an *active* way and thus adaptive to different patient images. Experimental results showed that our method achieves the best performance with fastest speed. Due to the generality of this framework, we plan to extend this method to CAD systems with more complicated task set.

References

1. Visvikis, D., Ell, P.J.: Impact of technology on the utilisation of positron emission tomography in lymphoma: current and future perspectives. European Journal of Nuclear Medicine and Molecular Imaging 30, S106–S116 (2002)
2. Niitsu, M., Takeda, T.: Solitary hot spots in the ribs on bone scan: value of thin-section reformatted computed tomography to exclude radiography-negative fractures. J. Comput. Assit. Tomogr. 27, 469–474 (2003)
3. O'Connell, M., Powell, T., Brennan, D., Lynch, T., McCarthy, C., Eustace, S.: Whole-body MR imaging in the diagnosis of polymyositis. AJR Am. J. Roentgenol. 179, 967–971 (2002)
4. Weber, U., Pfirrmann, C.W., Kissling, R.O., Hodler, J., Zanetti, M.: Whole body mr imaging in ankylosing spondylitis: a descriptive pilot study in patients with suspected early and active confirmed ankylosing spondylitis. BMC Musculoskeletal Disorders 8 (2007)
5. Brucker, P.: Scheduling algorithms (2004)
6. Pruhs, K., Sgall, J., Torng, E.: Handbook of Scheduling: Algorithms, Models, and Performance Analysis. CRC Press, Boca Raton (2003)
7. Nam, I.: Dynamic scheduling for a flexible processing network. Operations Research 49, 305–315 (2001)
8. Chou, M.C., Liu, H., Queyranne, M., Simchi-Levi, D.: On the asymptotic optimality of a simple on-line algorithm for the stochastic single-machine weighted completion time problem and its extensionsbrownian models of open processing networks:canonical representation of workload. Operations Research 54, 464–474 (2006)
9. Denzler, J., Brown, C.M.: Information theoretic sensor data selection for active object recognition and state estimation. IEEE Trans. PAMI 24, 145–157 (2002)
10. Cover, T., Thomas, J.: Elements of information theory (1991)

11. Viola, P., Jones, M.J.: Robust real-time face detection. International Journal of Computer Vision 57, 137–154 (2004)
12. Florin, C., Paragios, N., Funka-Lea, G., Williams, J.: Liver segmentation using sparse 3D prior models with optimal data support. IPMI (2007)
13. Zhan, Y., Shen, D.: Deformable segmentation of 3-D ultrasound prostate images using statistical texture matching method. IEEE Trans. Med. Imaging 25, 256–272 (2005)
14. Bookstein, F.: Principal warps: thin-plate splines and the decompotion of deformations. IEEE Trans. PAMI 11, 567–585 (1989)