Inter Programming Models for the Target Visitation Problem

Achim Hildenbrandt and Gerhard Reinelt
Institut für Informatik, Universität Heidelberg, Im Neuenheimer Feld 368
D-69120 Heidelberg, Germany

The target visitation problem (TVP) is concerned with finding a route to visit a set of targets starting from and returning to some base. In addition to the distance traveled a tour is evaluated by taking also preferences into account which address the sequence in which the targets are visited. The problem thus is a combination of two well-known combinatorial optimization problems: the traveling salesman and the linear ordering problem. In this paper we point out some polyhedral properties and develop a branch-and-cut algorithm for solving the TVP to optimality. Some computational results are presented.

Povzetek: Prispevek obravnava iskanje poti v grafu, kjer je potrebno obiskati več ciljev v najboljšem vrstnem redu.

1 Introduction

Let $D_{n+1} = (V_{n+1}, A_{n+1})$ be the complete digraph on $n + 1$ nodes where we set $V_{n+1} = \{0, 1, \ldots, n\}$. Furthermore let two types of arc weights be defined: weights d_{ij} (*distances*) for every arc (i, j), $0 \leq i, j \leq n$, and weights p_{ij} (*preferences*) associated with every arc (i, j), $1 \leq i, j \leq n$. The *target visitation problem (TVP)* consists of finding a Hamiltonian tour starting at node 0 visiting all remaining nodes (called *targets*) exactly once in some order and returning to node 0. Every tour can be represented by a permutation π of $\{1, 2, \ldots, n\}$ where $\pi(i) = j$ if target j is visited as i-th target. For convenience we also define $\pi(0) = 0$ and $\pi(n + 1) = 0$.

So we are essentially looking for a traveling salesman tour, but for the TVP the profit of a tour depends on the two weights. Namely, the value of a tour is the sum of pairwise preferences between the targets corresponding to their visiting sequence minus the sum of distances traveled, i.e., it is calculated as

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} p_{\pi(i)\pi(j)} - \sum_{i=0}^{n} d_{\pi(i)\pi(i+1)},$$

and the task is to find a tour of maximum value. So we have a multicriteria objective function.

The TVP was introduced in [4] and combines two classical combinatorial optimization problems: the *asymmetric traveling salesman problem (ATSP)* asking for a shortest Hamiltonian tour and the *linear ordering problem (LOP)* which is to find an acyclic tournament of maximum weight. (There is an obvious 1–1 correspondence between acyclic tournaments and linear orders of the node). Computational results of a genetic algorithm for problem instances with up to 16 targets have been reported in [1]. The original application of the TVP is the planning of routes for UAVs (un-

armed aerial vehicles). But there is a wide field of applications, e.g. the delivery of relief supplies or any other routing problem where additional preferences should be considered (town cleaning, snow-plowing service, etc.).

Obviously, the TVP is NP-hard because it contains the traveling salesman problem ($p \equiv 0$) and the linear ordering problem ($d \equiv 0$) as special cases.

In this paper we present first polyhedral results for the TVP and develop an algorithm for solving it to optimality. In section 2 we introduce an integer programming model. Section 3 discusses some structural properties of the associated polytope. A branch-and-cut algorithm based on these results is described in section 4. The algorithm is then applied to a set of benchmark problems and the computational results are presented in section 5. A few remarks conclude the paper.

2 An integer programming model for the TVP

For convenience we first transform the problem to a Hamiltonian path problem and also get rid of the special base node. This transformation is well-known for the ATSP [7] and can be adapted for the TVP as follows.

The key idea is to exploit the fact that each tour has to start at the base and return to it and that no preferences are to be taken into account for the base. In the *TVP-path model* we leave out this node and just search for a Hamiltonian path which visits all targets exactly once.
Following [7] we make the following modifications.

(i) Transform the distance matrix by setting $d'_{ij} = d_{ij} - d_{i0} - d_{0j}$, for all pairs i and j of nodes, $1 \leq i, j \leq n$ $i \neq j$.

(ii) Change the computation of the distance part of the objective function to

$$\sum_{i=1}^{n-1} d'_{\pi(i)\pi(i+1)} - \sum_{i=1}^{n} d_{i0} - \sum_{i=1}^{n} d_{0i}.$$

The preferences are not affected by this change. From now on we consider the TVP as finding a Hamiltonian path in the complete digraph $D_n = (V_n, A_n)$ with additional preference costs to be taken into account. The path is described by a permutation π of $\{1, \ldots, n\}$ where $\pi(k)$ is the node at position k.

We introduce two types of variables. The sequence in which the targets are visited is represented by binary *ATSP variables* x_{ij} for $1 \le i, j \le n$, $i \ne j$, with the interpretation

$$x_{ij} := \begin{cases} 1 & \text{if } i = \pi(k) \text{ and } j = \pi(k+1) \\ & \quad \text{for some } 1 \le k \le n-1, \\ 0 & \text{otherwise.} \end{cases}$$

The fact that some target i is visited before some target j is modeled with binary *LOP variables* w_{ij} for $1 \le i, j \le n$, $i \ne j$, with the definition

$$w_{ij} := \begin{cases} 1 & \text{if } i = \pi(k) \text{ and } j = \pi(l) \\ & \quad \text{for some } 1 \le k < l \le n, \\ 0 & \text{otherwise.} \end{cases}$$

An obvious idea for obtaining an IP model of the TVP is to combine well-known IP formulations for the ATSP and the LOP. This combination gives the following integer programming model.

$$\max \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \ne i}}^{n} p_{ij} w_{ij} - \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \ne i}}^{n} d_{ij} x_{ij} \tag{1}$$

$$\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \ne i}}^{n} x_{ij} = n-1, \tag{2}$$

$$\sum_{i \in S} \sum_{\substack{j \in S \\ j \ne i}} x_{ij} \le |S| - 1,$$

$$\qquad S \subseteq \{1, \ldots, n\}, 2 \le |S| \le n-1, \tag{3}$$

$$\sum_{\substack{i=1 \\ i \ne j}}^{n} x_{ij} \le 1, \qquad 1 \le j \le n, \tag{4}$$

$$\sum_{\substack{j=1 \\ j \ne i}}^{n} x_{ij} \le 1, \qquad 1 \le i \le n, \tag{5}$$

$$w_{ij} + w_{jk} + w_{ki} \le 2, \qquad 1 \le i, j, k \le n, \ i \ne j \ne k, \tag{6}$$

$$w_{ji} + w_{ij} = 1, \qquad 1 \le i, j \le n, \ i \ne j \tag{7}$$

$$x_{ij} - w_{ij} \le 0, \qquad 1 \le i, j \le n, \ i \ne j \tag{8}$$

$$x_{ij} \in \{0, 1\}, \qquad 1 \le i, j \le n, \ i \ne j \tag{9}$$

$$w_{ij} \in \{0, 1\}, \qquad 1 \le i, j \le n, \ i \ne j \tag{10}$$

Constraints (2)–(5) model the directed Hamiltonian paths where inequalities (3) are the *subtour elimination constraints*. Acyclic tournaments are modelled by the *3-dicycle inequalities* (6) and the *tournament equations* (7). Inequalities (8) connect the solutions of both problems. Together with the integrality conditions (9) and (10) this obviously constitutes a 0/1 model of the TVP.

At first want to prove the correctness of the model.

Lemma 1. *The model presented in (1) - (10) is a correct IP model for the TVP).*

Proof. At first we have to prove that every feasible solution fulfills the model. Since (2) - (5), (9) is a well known model for the ATSP and (6) - (7), (10) is a well known model for the LOP it is sufficient to show that the values of x_{ij} do match with the values w_{ij} or in equal that both types of variables describe the same TVP-path. To assure this it is sufficient to prove the following two facts:

a) $x_{ij} = 1 \Rightarrow w_{ij} = 1$

b) $w_{ij} = 1 \Rightarrow i$ must be visited before j in the path describe by the x-variables

Because (8) must be fulfilled a) is obvious. To prove b) we assume j is visited before i in the path. That means the are existing indizees $k_1, \ldots, k_l$ so that $j, k_1, \ldots, k_n, i$ is a part of the path. So it follows that $x_{j,k_1} = x_{k_1,k_2} = \cdots = x_{k_l,i} = 1$. With a) we get that $w_{j,k_1} = w_{k_1,k_2} = \cdots = w_{k_l,i} = 1$. Because of (6) and (7) we can than iteratively conclude that $w_{j,k_2} = 1$, $w_{j,k_3} = 1, \ldots, w_{j,i} = 1$. But this is a contradiction to our assumption.

It remains to show that every feasible solution of (1) - (10) is a correct TVP-path. It is clear that every feasible integer solution must induce a feasible linear ordering and a feasible TSP tour. Because of the facts we mentioned above it is clear that the two feasible solutions must match which each other.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

As an interesting fact we note that the subtour elimination constraints are actually not needed. If (w, x) satisfies (2) and (4)–(10), but not all inequalities (3) then there is some subtour on $k \ge 2$ nodes. W.l.o.g. we can assume that the node set is $\{1, 2, \ldots, k\}$ and the subtour is given as $\{(1, 2), (2, 3), \ldots, (k-1, k), (k, 1)\}$. Hence $x_{12} = x_{23} = \ldots = x_{k-1,k} = x_{k1} = 1$, implying because of (8) that $w_{12} = w_{23} = \ldots = w_{k-1,k} = w_{k1} = 1$. This is a contradiction to the requirement that the w-variables represent an acyclic tournament.

So we can eliminate the exponentially many constraints (3) and obtain a TVP formulation with a polynomial number (cubic in n) of constraints.

For our algorithm it will be useful to calculate the position of a node i in the path. This can easily be done using

the LOP variables. The value $n - \sum\limits_{\substack{j=1 \\ j \neq i}}^{n} w_{ij}$ gives the position of node i.

Note that because of the tournament equations we can substitute an LOP variable w_{ij}, $j > i$, by $1 - w_{ji}$. Now the 3-dicycle inequalities are turned into $1 \geq w_{ij} + w_{jk} - w_{ik} \geq 0$ for all $1 \leq i < j < k \leq n$ and the part of the objective function for the LOP variables reads $\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} [(p_{ij} - p_{ji})w_{ij} + p_{ji}]$.

2.1 Extended formulation of the basic model

The use of extended formulations is a common technique with is used to strengthen the LP Formulation of a combinatorial optimization problem. The key idea of this approach is to add new variables and constraints to a given IP formulation so that the gap between the solution of the LP relaxation and the optimal integral solution becomes much smaller.

In the case of TVP we can obtain an extended formulation for the TVP by adding three-indexed variables, which are a generalization of the linear ordering variables to the standard model. In detail this new variables w_{ijk} are than defined as follows:

$$w_{ijk} := \begin{cases} 1 & \text{if } i = \pi(a),\, j = \pi(b) \text{ and } k = \pi(c) \\ & \text{for some } 1 \leq a < b < c \leq n - 1\,, \\ 0 & \text{otherwise.} \end{cases}$$

So as one see this new type of variables is a straight forward extension of the w_{ij}-variables. In the objective function we assign zero coefficients to the new variables. In order to extend our standard model we also need to introduce two new classes of constraints to make sure that the solution of the new variables match with the old x_{ij} and w_{ij} variables. In detail the extended formulation looks a follows:

$$\max \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} p_{ij} w_{ij} - \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} d_{ij} x_{ij} \tag{11}$$

$$s.t.$$

$$\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} x_{ij} = n - 1, \tag{12}$$

$$\sum_{i=1}^{n} x_{ij} \leq 1,\; 1 \leq j \leq n, \tag{13}$$

$$\sum_{j=1}^{n} x_{ij} \leq 1,\; 1 \leq i \leq n, \tag{14}$$

$$\sum_{i \in S} \sum_{\substack{j \in S \\ j \neq i}} x_{ij} \leq |S| - 1,$$

$$S \subseteq \{1, \ldots, n\},\, 2 \leq |S| \leq n - 1, \tag{15}$$

$$w_{ij} + w_{jk} + w_{ki} \leq 2,$$

$$1 \leq i, j, k \leq n,\; i < j, i < k\, j \neq k, \tag{16}$$

$$w_{ij} + w_{jik} + w_{jki} + w_{kji} = 1$$

$$1 \leq i, j, k \leq n, i < j \tag{17}$$

$$x_{ij} - w_{ijk} - w_{kij} \leq 0\; 1 \leq i, j, k \leq n, i < j \tag{18}$$

$$x_{ij} - w_{ij} \leq 0,\; 1 \leq i, j \leq n, \tag{19}$$

$$x_{ij} \in \{0, 1\},\; 1 \leq i, j \leq n, \tag{20}$$

$$w_{ij} \in \{0, 1\},\; 1 \leq i, j \leq n. \tag{21}$$

$$w_{ijk} \in \{0, 1\},\; 1 \leq i, j, k \leq n. \tag{22}$$

3 The edge-node-formulation

The key idea of the next model is to combine the w and x variables of the -Model to new three index variables which states the relation between a node n and an fixed edge (i, j). More precisely we define:

$$w_{ij}^{k} := \begin{cases} 1 & \text{if } k = \pi(a),\, i = \pi(b) \text{ and } j = \pi(b+1) \\ & \text{for some } 1 \leq a < b \leq n - 1\,, \\ 0 & \text{otherwise.} \end{cases}$$

and analogously :

$$w_{k}^{ij} := \begin{cases} 1 & \text{if } i = \pi(a),\, j = \pi(a+1) \text{ and } k = \pi(b) \\ & \text{for some } 1 \leq a < b \leq\,, \\ 0 & \text{otherwise.} \end{cases}$$

A first IP-Model can then be develop out of the basic model by transforming the inequalities/equations to inequalities/equations with the new variables:

$$max \sum_{i=1}^{n} \sum_{j=1, i \neq j}^{n} p_{ij} \Big(\sum_{m=1, m \neq j}^{n} w_{mj}^{i} \Big) + \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} (w_{ij}^{\Omega} + w_{\Omega}^{ij}) \tag{23}$$

$$s.t.$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} (w_{ij}^{\Omega} + w_{\Omega}^{ij}) = n - 1 \tag{24}$$

$$\sum_{i=1}^{n} (w_{ij}^{\Omega} + w_{\Omega}^{ij}) \leq 1\;\; j \in V \tag{25}$$

$$\sum_{j=1}^{n} (w_{ij}^{\Omega} + w_{\Omega}^{ij}) \leq 1\;\; i \in V \tag{26}$$

$$\sum_{l=1}^{n} w_{lj}^{i} + \sum_{l=1}^{n} w_{lk}^{j} + \sum_{l=1}^{n} w_{li}^{k} + (w_{ik}^{\Omega} + w_{\Omega}^{ik}) \leq 2\;\; i, j, k \in V \tag{27}$$

$$\tag{28}$$

Please note that that $\Omega \in V$ and it could be chosen arbitrarily for each summand in (23) -(26) but $\neq j$ and $\neq i$. It is the same in (27) but here Ω must not be equal i or k

4 Three distance model

Another idea for constructing an IP-model for the TVP has been made by Prof. E. Fernandez from the UPC Barcelona. The key idea of this approach is the use of distance variables. In Detail we define Variables z_{ij}^t which describe the fact whether there is as path of length t between i an j or not. More formally we state:

$$z_{ij}^t = \begin{cases} 1 & \text{if } i \text{ the solution contains a path with } t \text{ arcs} \\ & \text{from } i \text{ to } j, \\ 0 & \text{otherwise.} \end{cases}$$

The advantage of this model is that we only have to deal here with one type of variables. Since we are not longer distinguish between distance and ordering variables we have to adjust the coefficients in the following way:

$$w_{ij}^t = \begin{cases} c_{ij} - d_{ij} & \text{if } t = 1, \\ c_{ij} & \text{otherwise.} \end{cases}$$

With are now able to formulated a TVP model with distance variables:

The formulation is:

$$\max \sum_{i \in N} \sum_{j \in N \setminus \{i\}} \sum_{t \in N \setminus \{n\}} \overline{w}_{ij}^t z_{ij}^t - \sum_{i \in N} (d_{i0} + d_{0i}) \quad (29)$$

$$\sum_{i=1}^{n} z_{ij}^1 \leq 1 \qquad j \in N, \quad (30)$$

$$\sum_{j=1}^{n} z_{ij}^1 \leq 1 \qquad i \in N, \quad (31)$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} z_{ij}^k = n - k \qquad k \in V \quad (32)$$

$$z_{ij}^{t_1} + z_{jk}^{t_2} \leq z_{ik}^{t_1+t_2} + 1,$$
$$i, j, k \in N, t_1, t_2 \in N \setminus \{n\},$$
$$i \neq j, j \neq k, i \neq k, t_1 + t_2 < n, \quad (33)$$

$$\sum_{t=1}^{n-1} z_{ij}^t + z_{ji}^t = 1 \qquad i, j \in N, i \neq j, \quad (34)$$

$$z_{ij}^t \in \{0, 1\} \qquad i, j \in N, i \neq j, t \in N \setminus \{n\}. \quad (35)$$

Also we only have one Type of variable now the $z_{i,j}^1$ variables still play a special role, for example in the objective function. On the other hand we again have a cubic number of variables

5 Conclusions

The target visitation problem turned out to be a very difficult and therefore challenging problem. The present paper gives some first results. More research is needed. An improvement of the simple heuristic used here can be accomplished along well-known lines. It is more interesting to find ways for improving the upper bound. The IP model already seems to be at its limits for fairly small problem instances unless some additional insight into the polyhedral structure can be obtained. Alternate optimization approaches like branch-and-bound with combinatorial bounds, dynamic or semidefinite programming should be devised and their limits should be explored. Furthermore it should be investigated further how the balance between the distance and the preference part of the object function influences the difficulty of problem instances.

References

[1] A. ARULSELVAN, C.W. COMMANDER and P.M. PARDALOS: *A hybrid genetic algorithm for the target visitation problem*, Technical report, University of Florida.

[2] T. CHRISTOF and A. Loebel: *PORTA: POlyhedron Representation Transformation Algorithm, Version 1.4*, http://comopt.ifi.uni-heidelberg.de/software/PORTA/index.html.

[3] M. GRÖTSCHEL, M. JÜNGER and G. REINELT: *Facets of the linear ordering polytope*, Math. Programming 33 (1985), 43–60.

[4] D.A. GRUNDEL and D.E. JEFFCOAT: Formulation and solution of the target visitation problem, In: Proceedings of the AIAA 1st Intelligent Systems Technical Conference, 2004.

[5] A. HILDENBRAND: *Benchmark instances for the TVP*, http://comopt.ifi.uni-heidelberg.de/people/hildenbrandt/TVP/template.html.

[6] M. JÜNGER and S. THIENEL: *The ABACUS System for branch-and-cut-and-price algorithms in integer programming and combinatorial optimization*, Software: Practice and Experience, 30 (2000), 1325–1352.

[7] M. QUEYRANNE and Y. WANG: *Hamiltonian path and symmetric travelling salesman polytopes*, Math. Programming 58 (1993), 89–110.

[8] G. REINELT: *TSPLIB – Benchmark instances for the TSP*, http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/.

Cervix Cancer Spatial Modelling for Brachytherapy Applicator Analysis

Peter Rogelj
University of Primorska, Faculty of Mathematics, Natural Sciences and Information Technologies
Glagoljaška 8, 6000 Koper, Slovenia
E-mail: peter.rogelj@upr.si

Muhamed Baraković
University of Verona, Department of Biotechnology
Ca' Vignal 1, Strada Le Grazie 15, 37134 Verona, Italia
E-mail: muhamed.barakovic@studenti.univr.it

Standard applicators for cervix cancer brachytherapy (BT) do not always enable a sufficient radiation dose coverage of the target structure (HR-CTV). The aim of this study was to develop methodology for building models of the BT target from a cohort of cervix cancer patients, which would enable BT applicator testing. In this paper we propose two model types, a spatial distribution model and a principal component model. Each of them can be built from data of several patients that includes medical images of arbitrary resolution and modality supplemented with delineations of HR-CTV structure, reconstructed applicator structure and eventual organs at risk (OAR) structures. The spatial distribution model is a static model providing probability distribution of the target in the applicator coordinate system, and as such provides information of the target region that applicators must be able to cover. The principal component model provides information of the target spatial variability described by only a few parameters. It can be used to predict specific extreme situations in the scope of sufficient applicator radiation dose coverage in the target structure as well as radiation dose avoidance in OARs. The results are generated 3D images that can be imported into existent BT planning systems for further BT applicator analysis and eventual improvements.

Povzetek: Razvita sta dva modela za izboljšanje brahoterapije.

1 Introduction

Applicators for cervix cancer brachytherapy (BT) enable cancer treatment that in comparison with external beam radiotherapy (EBRT) provides better radiation coverage of the high risk clinical target volume (HR-CTV) and better avoidance of organs at risk [1]. During the last decade remarkable progress has been made in radiotherapy, including cervix cancer BT [2]. Standard BT applicators for cervix cancer, as shown in Fig. 1, however still do not always enable a sufficient radiation dose coverage of the target, especially in cases of locally advanced cervical cancer. Improvements are searched in the direction of incorporating additional application needles. A development of new applicators that would enable better target dose coverage requires knowledge of cervix cancer spatial distribution and variation. Furthermore, as the applicators should be able to avoid organs at risk, the information of their variability should also be taken into account. In this work we aimed to develop methodology to obtain this information statistically using available data of past and present cervix cancer patients. The information required from each patient includes BT planning medical image, delineated HR-CTV structure, reconstructed BT applicator structure, and

organs at risk (OAR) structures. HR-CTV and OAR structures are in each 3D image delineated on each image slice, wherever the specific structure is present and, thus, available as a set of closed planar contours. BT applicators are reconstructed such that an applicator model is positioned in the 3D image inside the BT planning system. The applicator models consist of a ring, applicator tandem and needles, which are reconstructed independently. The actual position of the applicator is evident from the position of the applicator ring. Because the applicator must be positioned directly to the cervix and because the purpose of the models is applicator analysis, the spatial distribution and variation must be defined in the applicator coordinate system.

The significance of tumor distribution depend on the tumor type. It can help in development of tumor treatment and biopsy strategies and techniques [3, 4]. In the case of the cervical cancer it is also important due to BT applicator design.

Representation of 3D structures by sets of closed planar contours is not convenient for further spatial analysis. Other representations can be used instead, e.g., by tensors [5], Gaussian random spheres [6], signed distance maps [7] and others. However, due to eventual high complexity of BT target structures, we have selected the most common
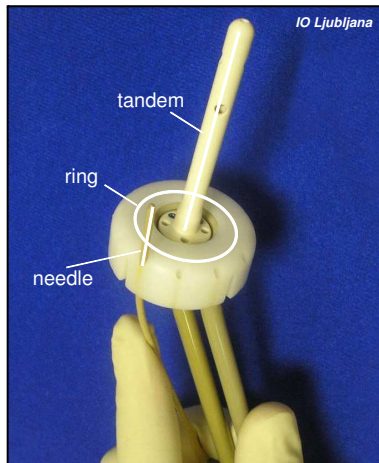
Figure 1: An image of a standard BT cervix cancer applicator with indicated parts: tandem, ring and optional needles.

representation of structures by binary images.

In the following sections our approach to model the spatial configuration of cervix cancer is described first. We describe the proposed methods of constructing the spatial distribution model and the principal component model. Then we show some test results; for the spatial distribution model based on real patient data, while the principal component model is illustrated using synthetic data. We conclude with discussion that includes the analysis of provided benefits and limitations.

2   Methods

Our approach to build spatial models of cervix cancer consists of the following processing steps that are described below: data input, applicator coordinate system definition, structure processing, modelling and data export.

2.1   Data input

The input data for building the models consists of patient medical data sets that comprise all the required information of each patient, i.e., a 3D medical image, delineated HR-CTV structure, OAR structures and a reconstructed applicator structure. This data is typically provided in the DICOM file format, which can be imported using DICOM libraries, e.g., the GDCM library [8], or by Matlab using Image processing toolbox. Medical images are needed only to obtain the image configuration, i.e., transformation of image coordinate system according to the patient coordinate system and image slice positions, which are required to correctly interpret the structures. Structures are given in a form of structure sets that include all the structure data required for BT treatment. The target structure (HR-CTV), OAR structures and the applicator ring structure can be identified from all the structures according to their names that must be known in advance for each individual data set; the structure naming is not standardized. The target and OAR
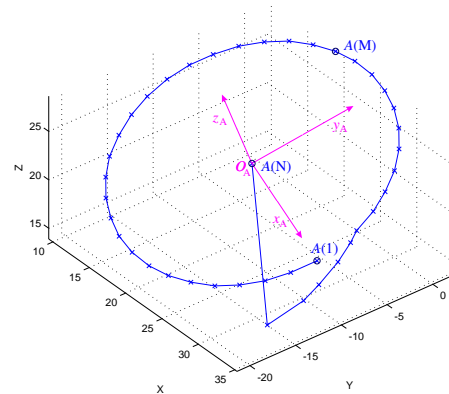
Figure 2: The applicator coordinate system is defined according to the applicator ring structure contour, with origin in the applicator ring center defined by the last point of the contour $A(N)$, xy plane in the ring plane with x axis pointing towards the contour starting point $A(1)$ and z axis in the direction of the tandem.

structures are defined as sets of closed planar contours, i.e., each contour is positioned on one slice of the corresponding 3D image. The applicator ring structures are described with a single open nonplanar contour. All contours are defined in the patient coordinate system.

2.2   Applicator coordinate system definition

Because the spatial configuration of cervix cancer needs to be defined according to the applicator perspective, an applicator coordinate system needs to be defined. The applicator reconstruction [9, 10] is performed on radiotherapy planning systems by importing predefined geometry structures. The applicator consists of tandem, ring and eventual additional needles, see Figure 1, which are all reconstructed independently. The ring structure, when inserted, tightly fits to the cervix anatomy, and provides a good base for defining the applicator coordinate system. Different applicator types may have different ring diameter, may be described with different number of contour points, however in practice the point ordering is always the same. For the illustration see Fig. 2. We propose that the applicator coordinate system is defined with origin in the ring center (the last point of the contour), xy plane in the ring plane, x axis in the direction towards ring contour starting point and z axis in the direction of the tandem.

The transformation that defines the applicator coordinate system according to the patient coordinate system can be for each applicator type computed from its ring contour coordinates. Actually, only three noncollinear contour points are required to compute the applicator coordinate system transformation $\mathbf{T}_A$, i.e., the last point, $A(N)$ that is positioned at the ring center as the coordinate system origin, $A(1)$ that defines the applicator coordinate system x axis, and any other point in the applicator contour circumference, $A(M)$ for defining the xy plane. The procedure is

the following:

$$O_A = A(N) \tag{1}$$
$$V_1 = A(1) - O_A \tag{2}$$
$$V_2 = A(M) - O_A \tag{3}$$

$$V_z = \frac{V_1 \times V_2}{\|V_1 \times V_2\|} \tag{4}$$

$$V_y = \frac{V_z \times V_1}{\|V_z \times V_1\|} \tag{5}$$

$$V_x = \frac{V_y \times V_z}{\|V_y \times V_z\|} \tag{6}$$

$$\mathbf{T}_A = \begin{bmatrix} V_x(x) & V_y(x) & V_z(x) & O_A(x) \\ V_x(y) & V_y(y) & V_z(y) & O_A(y) \\ V_x(z) & V_y(z) & V_z(z) & O_A(z) \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{7}$$

where, V and O are three dimensional vectors with components x, y, and z, such that O_A represents applicator coordinate system origin while V_x, V_y, and V_z are applicator coordinate system axes. Vector (cross) products assured coordinate axis perpendicularity, defining a Cartesian coordinate system.

The obtained transformation matrix $\mathbf{T}_A$ is needed for transforming BT structures to the applicator coordinate system in which the cervix cancer needs to be modelled.

2.3 Structure processing

For each image the corresponding BT target structure and OAR structures must be mapped into the applicator coordinate system. These structures are created by drawing contours on individual image slices and are provided as point sequences in the patient coordinate system. Such vector definition of structures is difficult to process statistically in coordinate system that is not parallel to the coordinate system of the originating image. Our solution is to present the structures in bitmap instead of vector format and process them as 3D (binary) images with voxel values 1 representing regions inside structures and 0 representing the surrounding. The approach is illustrated in Figure 3. The binary images cover the same region as the original medical image, except that, they may have different resolution in x and y image direction to control discretization error and data size. The resolution in z image direction must remain unchanged in order to preserve location of slices on which contours are defined.

The process of converting certain structure into a binary image starts with mapping the structure to the coordinate system of the original image. The transformation $\mathbf{T}_I$, form patient to image coordinate system can be obtained from image meta information, i.e., from DICOM tags *Image Position Patient* (0020,0032) and *Image Orientation Patient* (0020,0037). Thus, each contour C of structure S gets defined in its image coordinate system as C_I:

$$C_I = \mathbf{T}_I^{-1}C. \tag{8}$$

All points of the the same contour gets equal image coordinate z_I that is equivalent to the position of the image slice on which the contour was defined. The obtained structure can, as such, be drawn to the binary image, contour by contour. The process initiates by initializing all voxel values of the binary structure image to 0, followed by drawing the contours by checking each slice voxel if it is positioned inside of a polygon of contour points. Voxels inside the polygon gets negated to correctly interpret even complex structure shapes, e.g., shapes that include holes. Binary structure images enable further data integration towards spatial cervix cancer models.

To integrate the structure binary images of all patients into a spatial model, they all need to be mapped into the common applicator coordinate system (A), because patient coordinate systems (P) and image coordinate systems (I) are specific for each study/patient. Transformation between the coordinate systems are illustrated in Fig. 4. The data defined in image coordinate system (I) can be transformed to the applicator coordinate system (A) through the patient coordinate system (P) using transformation $\mathbf{T}_{IA}$:

$$\mathbf{T}_{IA} = \mathbf{T}_A^{-1}\mathbf{T}_I. \tag{9}$$

Structure binary images do not differ only according to their coordinate systems, but also according to image size and voxel size (resolution). For further analysis they need to be unified. The target region of interest and required precision define configuration of the resulting model (image) size and voxel size. All binary images must be resampled into this common spatial configuration. We recommend resampling by linear interpolation in reverse direction such that intensity corresponding to each voxel in the model configuration is interpolated from voxel intensities in the binary image. Note, that linear interpolation transforms a binary image into an image with real voxel values in interval $[0, 1]$. The result of structure processing is, therefore, a set of structure images S_A in a common applicator coordinate system and with common size and voxel size, i.e., each structure of each patient results in one structure image with common applicator (ring) position:

$$S_A = interp(\mathbf{T}_{IA}^{-1}S), \tag{10}$$

Where S denotes a structure binary image in the coordinate system of the original image, $interp$ a linear image interpolation, and S_A a structure image in the applicator coordinate system.

2.4 Spatial distribution model

The purpose of the spatial distribution model is to provide an overview of BT target spatial extent. It is given in a form of a spatial distribution image D, i.e., an image of the region of interest whose voxel values denote probability of voxel being inside the BT target region. It is obtained from images of the HR-CTV structure by averaging:

$$D = \frac{1}{n}\sum_{p=1}^{P} S_{A,p,HR-CTV}, \tag{11}$$
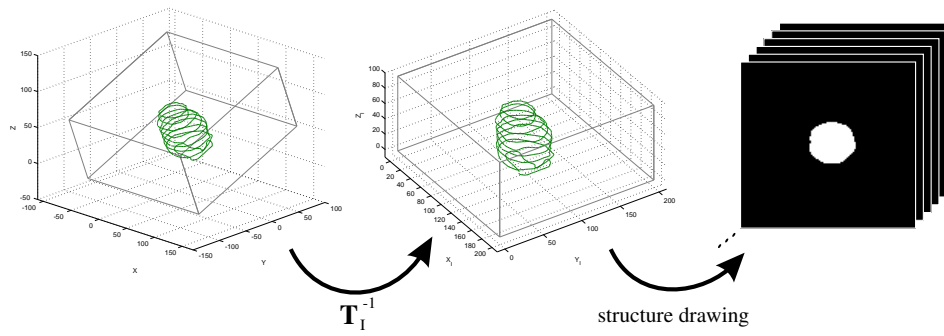
Figure 3: Illustration of structure processing: first, contours provided in the patient coordinate system (left) are transformed to the image coordinate system (center). Then, contours are drawn on image slices in 2D, which resulted in a 3D binary image of the structure (right).
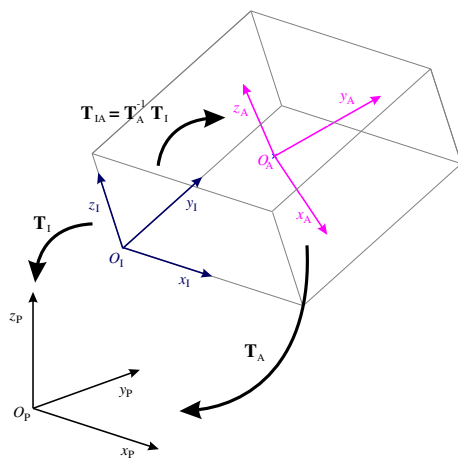
Figure 4: Illustration of patient (P), image (I) and applicator (A) coordinate systems and their transformations: $\mathbf{T}_{IA} = \mathbf{T}_A^{-1}\mathbf{T}_I$.

where $S_{A,p,HR-CTV}$ represents HR-CTV structure data of p-th patient resampled into an applicator coordinate system, and P is a total number of patients included in the analysis.

2.5 Principal component model

The principal component model provides information of the BT target spatial variability expressed by only a small number of parameters. The general idea is to be able to reconstruct any target configuration, i.e., position and extent of HR-CTV as well as OAR structures, by correctly setting the model parameters. As such the principal component model can be used to predict various target configurations, e.g., extreme situations in the scope of sufficient applicator radiation dose coverage in the target structure as well as radiation dose avoidance in OARs. Such situations may be crucial for testing real applicator efficiency. The principal component model tends to extract a minimal set of orthogonal components of spatial variations in the region of interest using the principal component analysis (PCA). PCA projects the data into a lower dimensional linear space such that the variance of the projected data is maximized,

or equivalently, it is the linear projection that minimizes the mean squared distance between the data points and their projections. PCA provides a full set of components that enable perfect data reconstruction, however, it also orders the components according to their importance, i.e., according to their contribution to the data description. It turns out that majority of the components have low importance and only a small error is made when only a few most important components are used. In this case the important components can be computed more efficiently using singular value docomposition (SVD) [11].

Our input data for the PCA analysis of the BT target are the HR-CTV structure images in the applicator coordinate system $S_{A,p,HR-CTV}$. The data of each image is reordered into a row vector and joined for all the patients into a matrix $\mathbf{X}_{P \times L}$, with L being the number of pixels in the image. Then the mean vector $\overline{\mathbf{X}}$ is computed and subtracted from each data row to obtain the matrix $\mathbf{X}_0$ representing the zero-mean data variation. Here, the mean vector $\overline{\mathbf{X}}$ corresponds to the reordered data of the spatial distribution model D. SVD decomposes $\mathbf{X}_0$ into three matrices; matrix $\mathbf{V}$ with orthogonal columns that represent principal components, diagonal matrix $\mathbf{S}$ with singular values that represent importance of the components, and matrix $\mathbf{U}$ providing component weights for reconstructing the input data:

$$\mathbf{X}_0 = \mathbf{U}\mathbf{S}\mathbf{V}^T \qquad (12)$$

The efficient SVD implementations, e.g., Matlab *svds* function, enable computation of only a given number of principal components R, and as such provide approximate solutions:

$$\mathbf{X}_{0P \times L} \approx \mathbf{U}_{P \times R}\mathbf{S}_{R \times R}\mathbf{V}_{L \times R}^T \qquad (13)$$

The obtained matrices $\mathbf{S}$ and $\mathbf{V}$ represent a principal component model of the HR-CTV structure, such that HR-CTV structure of any patient can be represented with the R components, i.e., the columns of $\mathbf{V}$, with weights:

$$\mathbf{U}' = \mathbf{X}_0'\mathbf{V}\mathbf{S}^{-1} \qquad (14)$$

where $\mathbf{X}_0' = \mathbf{X}' - \overline{\mathbf{X}}$ represents deviation of the data from the average. Similarly, BT target data can also be simulated

by manual setting of component weights in $\mathbf{U}$, following equation (13) and adding the mean vector $\overline{\mathbf{X}}$. Component weights form a low dimensional linear space with a certain region around the origin that corresponds to realistic data variation. The limits of this realistic subspace can be estimated by analyzing large amount of data, i.e., large number of patients. Values at the border of the realistic subspace can be used in $\mathbf{U}$ to simulate specific extreme situations suitable for BT applicator analysis.

Testing of BT applicators on their ability to radiate the HR-CTV regions may be biased, as applicators should also be able to avoid radiation of OAR structures. It is important that HR-CTV and OAR structures cannot overlap. This property can be used to simultaneously model both structure types , i.e., HR-CTV as well as OAR structures, without increasing the amount of data in the PCA analysis. For this purpose the input vector $\mathbf{X}$ must be constructed from all the structure images, such that positive values represent target regions and negative values represent OAR regions:

$$\mathbf{X} = \mathbf{X}_{\text{HR-CTV}} \check{} \sum \mathbf{X}_{\text{OAR}} \qquad (15)$$

Here, $\mathbf{X}_{\text{HR-CTV}}$ and $\mathbf{X}_{\text{OAR}}$ are constructed from structure images with reordering into row vectors as described earlier, using HR-CTV and available OAR structures. Typically, OAR structures include bladder, rectum and sigmoid colon.

The simulated or reconstructed data that results from the principal component model, as well as principal components themselves, can be reordered back into 3D images. Due to interpolations and approximations the reconstructed structures are not presented only with values 1, -1 and 0 for target structures, OAR structures and surrounding respectively. Consequently, we recommend completing the reconstruction procedure with thresholding using thresholds -0.5 and 0.5.

2.6 Data export

The resulting images, i.e., an image of the spatial distribution model and simulated or reconstructed BT target configurations can be used in BT planning systems, e.g., Brachyvision©, for further analysis. BT planning systems include functionalities that enable radiation simulations using different radiation plans and can be used to test the efficiency of different applicators. To enable this procedures the images shall be exported to DICOM image format, which can be done using DICOM libraries, e.g. GDCM library [8], or Matlab image processing toolbox.

3 Results

We have tested the proposed methods on real and simulated data. First a spatial distribution model has been created from real data of 264 consecutive cervix cancer patients. Due to relatively large number of patients, the obtained estimate of spatial cervix cancer distribution was named a
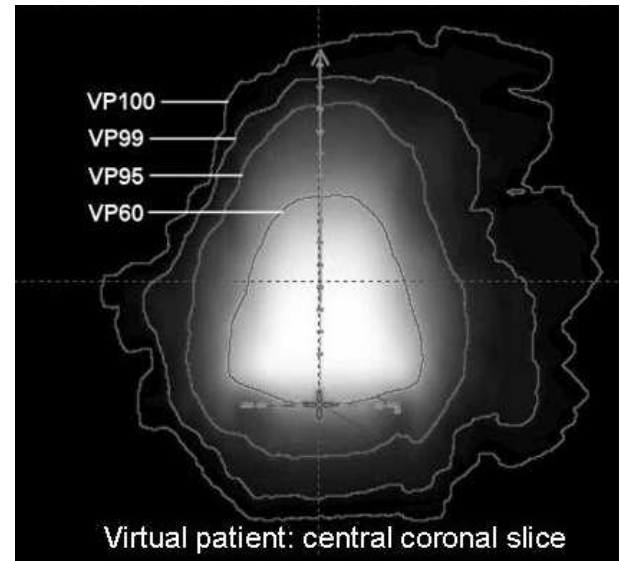

Figure 5: Illustration of the cervix cancer spatial distribution representing a virtual patient, the central coronal slice.

virtual patient (VP). Imported to the BT planning system isosurfaces that connect voxels with the same values were created and labeled as percentage of encompassed voxels. VPn was defined as VP subvolume, encompassed by the n% isosurface, see the illustration in Figure 5. The obtained VP data was used for analysis and development of BT applicators for cervix cancer [12]. It was found out, that standard tandem and ring (T&R) applicator enables adequate treatment of VP60 subvolume, additional needles parallel to tandem extend adequate treatment to VP95 and additional oblique needles, inserted at points, angles and depths extend adequate treatment to VP99 subvolume. The principal component model was not built for this dataset, such that applicators were tested only for their general capability of radiating the HR-CTV, not considering the capability of avoiding the radiation of OAR structures.

The principal component model was tested using a simulated dataset that we have created for this purpose. Note that the simulated structure images presented here do not realistically simulate the BT target configuration, however enable illustration of the concept and testing of its suitability for creating a realistic model.

The simulated data was generated using four random parameters where three of them were used to simulate variability of the HR-CTV structure and the additional one for the variability of one OAR structure. The HR-CTV structure was simulated as an ellipsoid with the three parameters representing the semi-axes lengths while its center was always in the applicator coordinate system origin. The OAR region was simulated as a sphere with the given parameter representing its radius, while its center was defined such that the distance between the edges of BT target and the OAR structure was constant. For the illustration see Figure 6.

A principal component model was generated from a
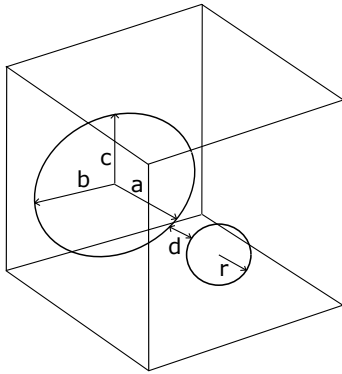
Figure 6: Illustration of the simulated dataset configuration. The HR-CTV was simulated with an ellipsoid and one additional OAR structure as a sphere with a constant distance d from the HR-CTV.

dataset of 400 simulated 3D images with $100 \times 100 \times 50$ voxels. The simulation parameters were selected randomly in the following ranges: $a \in [40, 73]$, $b \in [35, 59]$, $c \in [35, 49]$, $r \in [15, 20]$ and $d = 5$. The computation of the principal component model was restricted to 11 principal components. The mean image $\overline{\mathbf{X}}$ and the components are illustrated in Figure 7. The singular values that represent the distribution of the dataset's energy among the principal components indicate that the component energy gradually decreases with the component number, see Figure 8. However, although not all of the energy was considered, the reconstructed images did not differ considerably from the images from the training set as shown in Figure 9, where a randomly selected input structure image is compared with its reconstructed approximations obtained using three and eleven principal components. We can notice minor differences even when reconstructing from three components only.

If we observe the component weights (the values of matrix U), we can see that they are spread over a limited PCA subspace, see Figure 10, which corresponds to valid structure images. According to the shape of the subspace, we can conclude that component weights of valid images are not fully independent, although the components are orthogonal. By selecting weights manually, additional structure images can be simulated. If the selected weights are from the subspace of valid structure images, the simulated images follow the concepts of the input dataset, else the results may include major deviations as demonstrated in Figure 11.

The possibility to simulate structure images and have control over its validity offers good opportunity to generate specific synthetic images of the BT target region that represent extreme situations for BT applicator testing. In that case the principal component model should be created from real patient data and the test cases selected at the border of the populated PCA subregion.

The realistic model has not been created, yet, however we are looking forward to create it in collaboration with medical institutions that maintain large databases of their cervix cancer patients.

4 Discussion and conclusion

Cancer spatial distributions must be considered whenever cancer treatment tools and procedures are being developed. Unfortunately, statistical analysis of spatial distributions related to specific organs is in general tedious due to difficulties defining the reference coordinate systems because of their complex shapes and their high variability. In the specific case of cervix cancer the organ geometry enables unambiguous coordinate system definition that agrees with the applicator ring structure. Analysis of other cancer types would require definition of analysis coordinate system according to organ geometry and data integration performed by image registration with a reference or atlas [13]. Similarly, image registration has already been used for analyzing interfraction variation of high dose regions of OARs [14], and could be extended to intersubject analysis of cancer distributions.

The spatial distribution model provides useful information about target region that needs to be radiated, and has already been used for development of novel applicator types [12]. However, this model does not consider OARs and difficulties of restricting radiation dose in these structures. If a distribution model would be made for OARs as well, it would most probably overlap with the cancer distribution model due to closeness of some OAR structures to HR-CTV and due to anatomical variability. Better applicator testing must, therefore, take into account the BT target variability, e.g., by testing on diverse specific target configurations, which can correspond to real patients or obtained by modelling. The proposed principal component model has advantages over using the real patients' data, because of the established control over the specificity of the cases, a possibility to simulate the non-existent cases and depersonalization.

The limitation of the principal model is in its high computational cost. Computation of all the PCA components would require enormous amount of memory, only the $\mathbf{V}$ matrix would have the size of 500k $\times$ 500k elements (assuming $2 \times 2 \times 2$ mm voxel size), which in float data format requires 1TB of memory. Using the SVD approach with computation of the most important components only, drastically reduces the memory requirements; in our simulated case matrix $\mathbf{V}$ occupied only 22MB. Such reduction of components is possible due to final thresholding, which is applicable due to binary nature of the structures. When the computational cost remains a problem, high efficient PCA solutions [15] or alternative structure representations could be used.

A principal component model of real cervix cancer has not been made, yet. A large number of patient datasets is required and in contrast to the spatial distribution model the OAR structures must be included. The preparation of such data is tedious due to non-standardized structure naming.
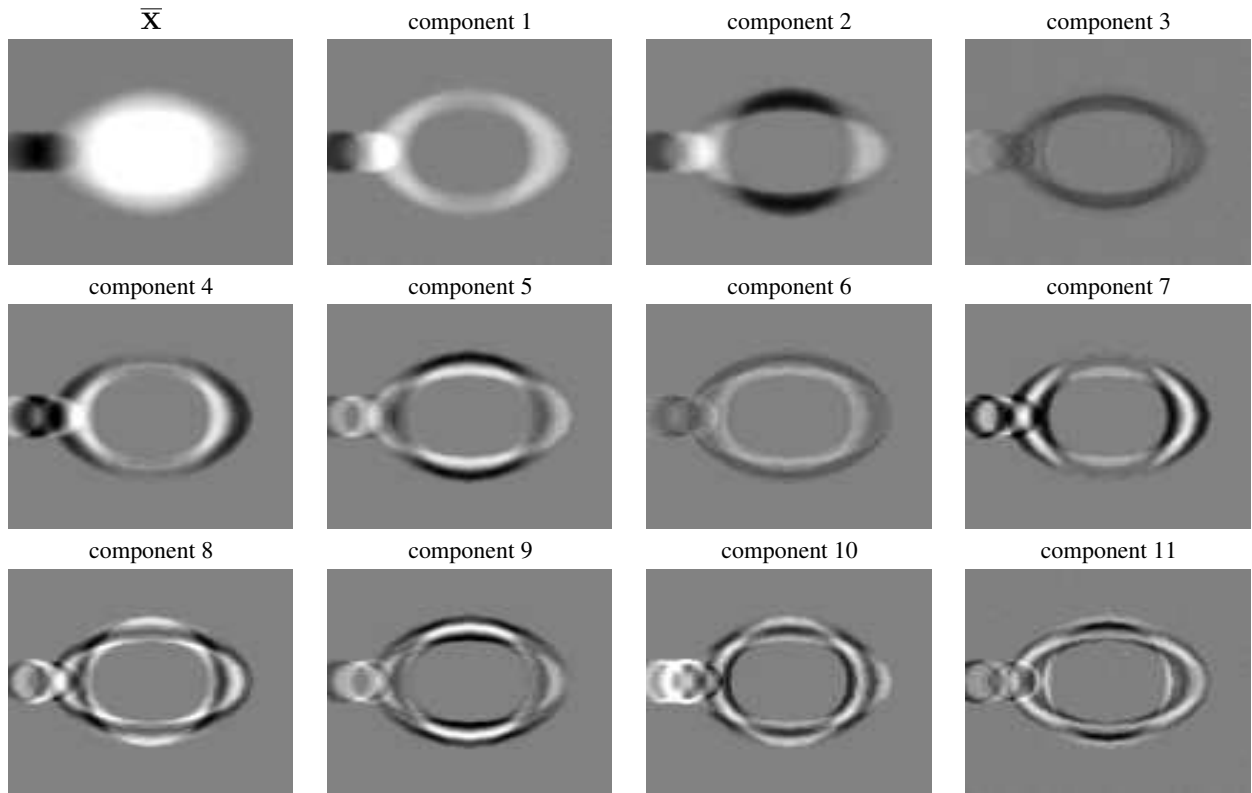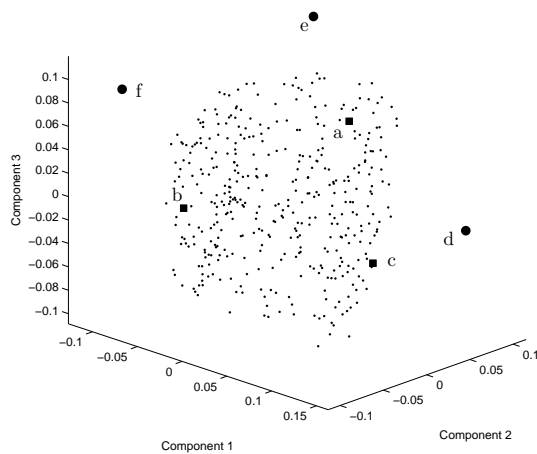
Figure 7: Components of the simulated dataset (central slices only). The mean image $\overline{\mathbf{X}}$ is presented in a scale from -1 (black) to +1 (white) and components with a scale from -0.01 (black) to +0.01 (white).
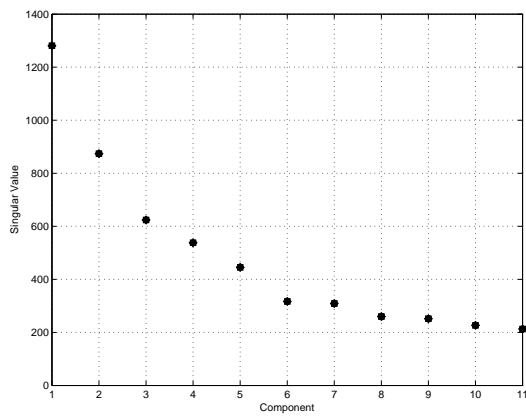
Figure 8: Singular values corresponding to the first 11 components; singular values represent the distribution of the dataset's energy.
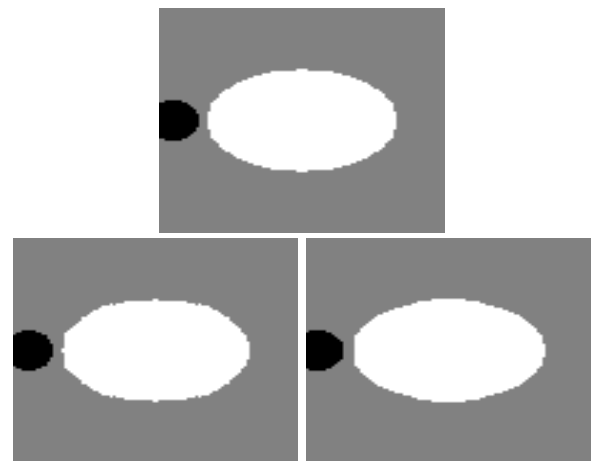
Figure 9: The central slice of an input structure image (top) and its reconstruction using 3 and 11 components (bottom left and right respectively).

Figure 10: Weights for the first three components of the simulated structure images. The small dots correspond to images from the input dataset, squares and large dots represent selected values inside and outside the populated subspace for further simulations.

However the benefits of such dataset are not only in the support of applicator development, but also in outcomes of further statistical analysis that could support clinical process, e.g., structure delineation or radiation planing, as well as making of clinical decisions.

To conclude, it may be widely accepted that reducing dose at organs of risk is difficult without reducing dose at large tumors [16], we believe that applicator improvements based on spatial modelling could provide better alternatives.

References

[1] R. Potter, "Image-guided brachytherapy sets benchmarks in advanced radiotherapy." *Radiother Oncol*, vol. 91, no. 2, pp. 141–146, May 2009.

[2] A. H. Sadozye and N. Reed, "A review of recent developments in image-guided radiation therapy in cervix cancer." *Curr Oncol Rep*, vol. 14, no. 6, pp. 519–526, Dec 2012.

[3] M. B. Opell, J. Zeng, J. J. Bauer, R. R. Connelly, W. Zhang, I. A. Sesterhenn, S. K. Mun, J. W. Moul, and J. H. Lynch, "Investigating the distribution of prostate cancer using three-dimensional computer simulation." *Prostate Cancer Prostatic Dis*, vol. 5, no. 3, pp. 204–208, 2002.

[4] Y. Ou, D. Shen, J. Zeng, L. Sun, J. Moul, and C. Davatzikos, "Sampling the spatial patterns of cancer: optimized biopsy procedures for estimating prostate cancer volume and gleason score." *Med Image Anal*, vol. 13, no. 4, pp. 609–620, Aug 2009.
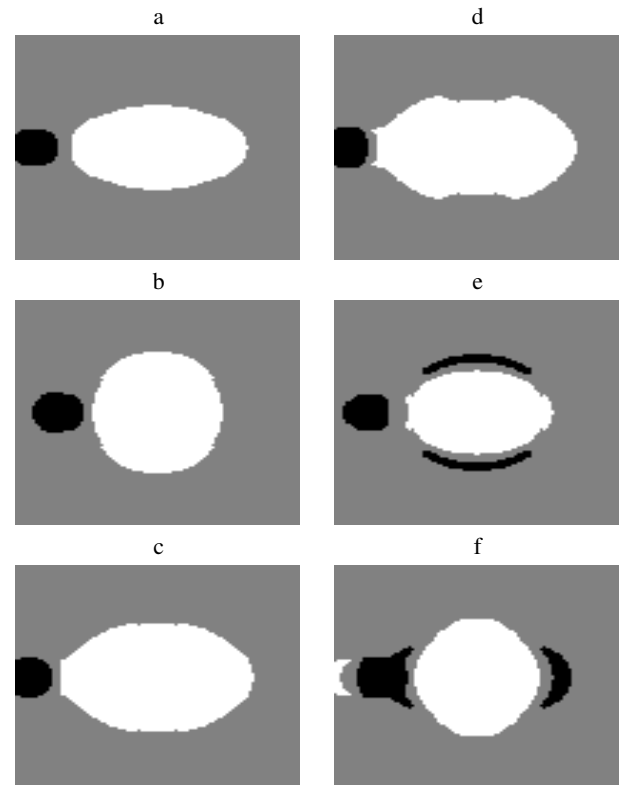
Figure 11: Central slices of simulated structure images using selected component weights from the subspace of valid structure images (left) and from other parts of the PCA space (right).

[5] R. Xu and Y.-W. Chen, "Appearance models for medical volumes with few samples by generalized 3d-pca," in *Neural Information Processing*, ser. Lecture Notes in Computer Science, M. Ishikawa, K. Doya, H. Miyamoto, and T. Yamakawa, Eds. Springer Berlin Heidelberg, 2008, vol. 4984, pp. 821–830.

[6] R. C. Conceicao, M. O'Halloran, E. Jones, and G. M., "Investigation of classifiers for early-stage breast cancer based on radar target signatures," *Progress In Electromagnetics Research*, vol. 105, pp. 295–311, 2010.

[7] K. M. Pohl, S. K. Warfield, R. Kikinis, W. L. Grimson, and W. M. Wells, "Coupling statistical segmentation and pca shape modeling," in *Medical Image Computing and Computer-Assisted Intervention MICCAI 2004*, ser. Lecture Notes in Computer Science, C. Barillot, D. Haynor, and P. Hellier, Eds. Springer Berlin Heidelberg, 2004, vol. 3216, pp. 151–159.

[8] "GDCM library home page (version 1.x)." [Online]. Available: http://www.creatis.insa-lyon.fr/software/public/Gdcm/

[9] D. Berger, J. Dimopoulos, R. Potter, and C. Kirisits, "Direct reconstruction of the vienna applicator on mr

images." *Radiother Oncol*, vol. 93, no. 2, pp. 347–351, Nov 2009.

[10] S. Haack, S. K. Nielsen, J. C. Lindegaard, J. Gelineck, and K. Tanderup, "Applicator reconstruction in mri 3d image-based dose planning of brachytherapy for cervical cancer." *Radiother Oncol*, vol. 91, no. 2, pp. 187–193, May 2009.

[11] M. E. Wall, A. Rechtsteiner, and L. M. Rocha, "Singular Value Decomposition and Principal Component Analysis," *ArXiv Physics e-prints*, Aug. 2002.

[12] P. Petric, R. Hudej, P. Rogelj, J. Lindegaard, K. Tanderup, C. Kirisits, D. Berger, J. C. A. Dimopoulos, and R. Potter, "Frequency-distribution mapping of hr ctv in cervix cancer : possibilities and limitations of existent and prototype applicators." *Radiother. oncol.*, vol. 96, suppl. 1, p. S70, 2010.

[13] K. Diaz, B. Castaneda, M. L. Montero, J. Yao, J. Joseph, D. Rubens, and K. J. Parker, "Analysis of the spatial distribution of prostate cancer obtained from histopathological images," in *Proc. SPIE 8676, Medical Imaging 2013: Digital Pathology, 86760V*, March 2013.

[14] J. Swamidas, U. Mahantshetty, D. Deshpande, and S. Shrivastava, "Inter-fraction variation of high dose regions of oars in mr image based cervix brachytherapy using rigid registration," *Medical Physics*, vol. 39, no. 6, p. 3802, JUN 2012.

[15] V. Zipunnikov, B. Caffo, D. M. Yousem, C. Davatzikos, B. S. Schwartz, and C. Crainiceanu, "Multilevel functional principal component analysis for high-dimensional data," *Journal of Computational and Graphical Statistics*, vol. 20, no. 4, pp. 852–873, 2011.

[16] R. Kim, A. F. Dragovic, and S. Shen, "Spatial distribution of hot spots for organs at risk with respect to the applicator: 3-d image-guided treatment planning of brachytherapy for cervical cancer," *International Journal of Radiation Oncology, Biology, Physics*, vol. 81, no. 2, S, p. S466, 2011.