Introduction of a Combination Vector to Optimise the Interpolation of Numerical Phantoms

Julien Henriet [(1)], Pascal Chatonnay [(2)]

[(1)] Laboratoire Chrono-Environnement, UMR CNRS 6249, Université de Franche-Comté, 4 Place Lucien Tharradin, 25200 Montbéliard, France. julien.henriet@univ-fcomte.fr

[(2)] Institut FEMTO-ST, UMR CNRS 6174, Université de Franche-Comté, 4 Place Lucien Tharradin, 25200 Montbéliard, France. pascal.chatonnay@univ-fcomte.fr

Abstract

Phantoms are 3-dimensional (3D) numerical representations of the contours of organs in the human body. The quality of the dosimetric reports established when accidental overexposures to radiation occur is highly dependent on the phantom's reliability with respect to the subject. EquiVox is a Case-Based Reasoning platform which proposes an interpolation of the 3D Lung Contours (3DLC) of subjects during its adaptation phase. This interpolation is conducted by an Artificial Neural Network (ANN) trained to learn how to interpolate the 3DLC of a Learning Set (LS). ANN is a well-suited tool when known results are numerous. Since the cardinality of our learning set is restrained, the imperfections of each 3DLC have a great impact on interpolations. Thus, we explored the possibility of ignoring some of the 3DLC of LS via implementation of a new learning algorithm which associated Combination Vectors (CV) to LS. The results proved that this method could optimise interpolation accuracy. Furthermore, this study highlights the fact that some of the 3DLC were harmful for some interpolations whereas they increased the accuracy of others.

Keywords:

Interpolation, Adaptation, Optimisation, Case-Based Reasoning, Artificial Neural Network.

I – Introduction

In the case of accidental exposure to radiation, a dosimetry evaluation must be established as soon as possible for each subject. In most cases, this evaluation is based on available 3D voxel phantoms, numerical models created from medical images to represent the subject's organs with maximum realism. Examples of voxel phantoms for dosimetric assessment following internal contamination or external exposure can be found (1) (2). However, even when medical images are available, the subject's specific phantom is not always accessible since its construction is delicate. Moreover, medical images are avoided so as to prevent any additional exposure to radiation. Thus, existing models are used even if their characteristics differ from the subject's biomedical data. Dosimetry assessment accuracy and the resulting decontaminating medical actions are nevertheless highly dependent on the similarity between phantom and subject. Hence, the actual work aims at assisting the physician in choosing the fittest phantom from the existing ones available.

EquiVox is a platform based on Case-Based Reasoning (CBR) which is a problem solving method that uses similar solutions from similar past problems in order to solve new problems (3). CBR is a tool for retrieving, adapting, revising and storing experiences, and many adaptation strategies can be found

in the literature. Adaptation by Generalisation/Specialisation requires a hierarchical organisation of the CBR source cases according to generalisation/specialisation relations. Some characteristics are hidden in the generalisation process whereas special ones are added to the general case during the specialisation process. Adaptation using Adaptation Rules (4) consists of computing a solution for a target case applying a function which takes as its parameters the target case, a source case that presents some similarities and its solution. Differential Adaption (5) is based on the evaluation of the variations between the source and target cases: an approximate solution of the target case is computed by applying the variations between the two cases to the solution for the source case under consideration. Conservative Adaptation (6) is based on the Revision Theory which considers knowledge updates. This kind of adaptation consists of minimising the modifications to be applied to the knowledge. A cost for the possible adaptations must be computed. The EquiVox adaptation phase is based on rules known from experience. After having retrieved the phantom the most similar to the subject's thorax (7) (8), EquiVox proposes an original tool based on Artificial Neural Networks (ANN) (9) to create the 3D contour of the subject's lungs (10) during the CBR adaptation phase (11). The present study goes a step further since it introduces a new concept capable of determining the best subsets of phantoms for the construction of Contours in 3 Dimensions of the Lungs (3DLC) of a given subject with the greatest accuracy.

II – Requirements, hypothesis and method

A large number of phantoms can be found in the literature (12) (13) (14) (15) (16) (17) (18), and radiation protection is also divided into numerous sub-domains. Indeed, some phantoms are commonly used by experts for external radiotherapy, while different ones are used by other physicians for evaluation of internal doses received. In fact, each expert has his own set of 10 to 20 phantoms. When a physician's usual phantoms are all too distant from the subject, the expert must create a new one. Using interactive 3D dilatations and contractions, physicians modify the contours of the 3D organs of their phantoms until they correspond to those of the subject. They then put them together and obtain the final phantom on which the computations will be based (19). Thus, adaptation rules are guided by the experience and knowledge of the experts. EquiVox is able to produce the same transformation process automatically, without human intervention, using an ANN (11). ANN is an interpolation tool which requires a training phase. For the 3DLC construction of EquiVox, the training set was the entire set of known 3DLC (20).

Nevertheless, we assumed that if the subject is a baby, for example, it is relevant to learn how to create 3DLC using an ANN trained on a set of known 3DLC of other babies and to exclude adult ones. Thus, in this study we propose to optimise the subjects' 3DLC construction, taking into account their specific characteristics after the ANN training. We introduced a vector to express whether or not it was relevant to include each known 3DLC in the learning set of subject characteristics: the Combination Vector for Interpolation Optimisation (CVIO).

II – A – The EquiVox platform

Figure 1 presents the technologies that were used and the data flows over the EquiVox architecture. All the phantoms are stored in Rhino3D files (21) and their characteristics kept in a database (data flow #0 in Figure 1). The lung contours are extracted (data flow #1) and then transmitted to the ANN training module (data flow #2) which creates the ANN (data flow #3). When a new phantom is required, the target case description is transmitted to the retrieval module (data flow #4) which determines the similitude and confidence indices taking into account the source case (data flow #5).

If required by the experts, the lung adaptation module sends the characteristics of the source cases (data flow #6) to the ANN interpolation module (data flow #7) which loads the trained ANN (data flow #8) and the coordinates of the lung contour in question (data flow #9) in order to create interpolated contours suited to the target case (data flow #10).

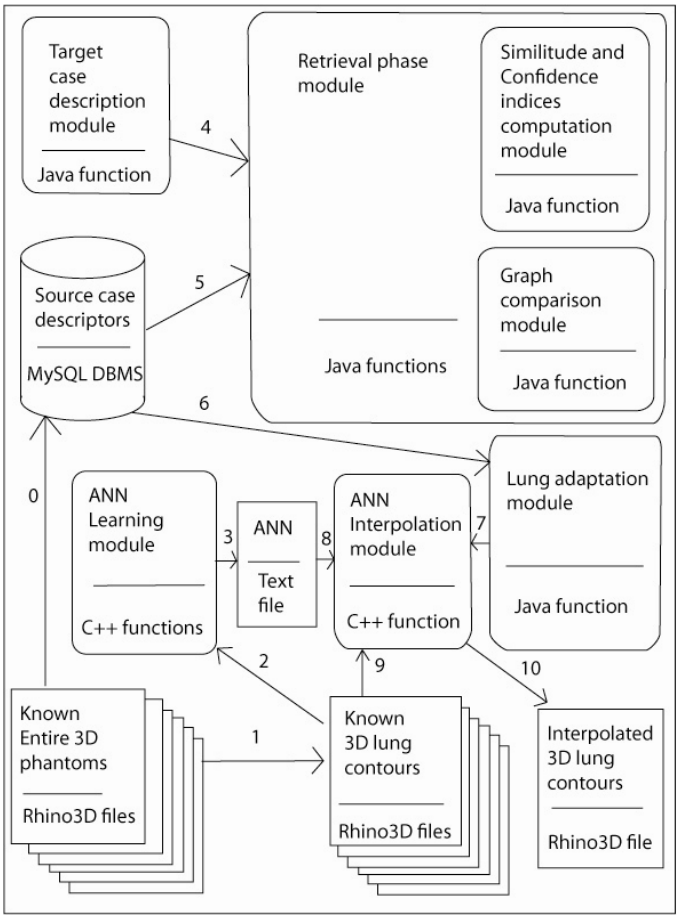

Figure 1. Data flows over EquiVox Architecture.

II – A – 1 – Case modelisation

When radiation overexposure occurs, a dosimetric report must be established for all subjects. For each one, the experts' first task is to choose the most accurate 3D phantom considering the information known about the subject. Each phantom has its own characteristics and is chosen by comparing subjects' available measurements and information to their characteristics. The phantom is thus chosen by analogy.

We exhausted the list of useful characteristics furnished by the physicians of the French Institute of Radiation and Protection (IRSN).

Thus, in EquiVox, a problem is described as a set of *r* descriptors $\{d_1, \dots, d_r\}$.

Each expert has his own set of *n* phantoms: $SP = \{P_1, , \dots, P_n\}$.

Each P_i is the solution part of a case and represents the contours of *m* organs: $P_i=\{P_i^1, \dots, P_i^m\}$.

Each organ *O* is a set of *q* points joined by a Delaunay mesh (22): $P_i^o=\{C_1^{i,o}, \dots, C_q^{i,o}\}$ where $C_j^{i,o}$ denotes the 3D coordinates of point *j* of organ *O* of phantom P_i. $O \in$ *{lung, heart, liver, sternum, ribs, scapulae, spine, breasts, skin, oesophagus and thorax}*.

Finally, a case *i* is: $i = \{ \{d_1^i, \dots, d_r^i\}, P_i\}$. We will note the target case as *t*.

II – A – 2 – Adaptation of 3DLC

Once a matching case is retrieved (7), the expert can decide either to use the the most similar source-case phantom, or to require the EquiVox platform to generate a new phantom, adapting the source cases to the target one. Indeed, if some available phantom measurements are too different from those of the subject, the expert may decide to adapt one of them or even to create a new phantom which may be reused for other problems later. Thus, when the expert requires the generation of a new phantom, the contours of the *m* organs are expected.

Actually, the first organs that experts create in such a personalised process are the lungs. The positions and volumes of the other organs are deduced from those of the lungs. Thus, we first considered the adaptation of 3DLC.

II – A – 2 – a – Solution space modelisation for 3DLC

As previously presented, the 3DLC of phantom P_i are defined in 3D by a set of *q* points joined by a Delaunay mesh: $P_i^{lung} = \{C_1^{i,lung}, \dots, C_q^{i,lung}\}$ where $C_k^{i,lung}$ denotes the 3D coordinates of point *k*: $C_k^{i,lung} = \{x_k^{i,lung}, y_k^{i,lung}, z_k^{i,lung}\}$. For each 3DLC, *q* is equal to 26 723 points. The points were plotted in the same order and in the same Cartesian coordinate system. Thus, the task of the lung contour-adaptation phase of EquiVox consists of interpolating the 3D coordinates of the points of *t* in the same order and in the same Cartesian coordinate system. A Delaunay mesh can then be applied so as to create the contours of the lungs of *t*.

II – A – 2 – b – Adaptation rules

In fact, lung contours and volumes depend mostly on the height of the subject. Indeed, for the lungs, I. Clairand *et al.* (23) proved that the height of a person prevailed for their geometry and volume.

Thus, when experts decide to create the 3DLC of a subject, they choose the one from the stored phantom whose height is the closest without taking into account any other characteristic. The adaptations are usually done manually, applying mathematical transformations (2D and 3D contractions and dilations (19)). These transformations are carried out through 3D modelling tools (such as Rhinoceros (21) or CATIA (24)).

II – A – 2 – c – Method

Since the mesh and the number of points are not variable, the adaptation must be carried out on the point coordinates of the lung contours, point by point. Since no formal equation exists, we had to discover through a learning method the rules that transform the coordinates of the points on one lung contour into other coordinates.

Consequently, data-driven methods using inductive reasoning are the most suitable approaches; ANN and Fuzzy-ANN respond to these requirements. We chose ANN as the tool for this step, assuming this could serve as the basis for further work with Fuzzy-ANN if the first results were not convincing.

We explored the possibility of using a multi-layer perceptron trained with a backpropagation-based method. Other interpolation methods were tested: polynomial and Spline ones (cf Subsection II-A-2-e).

II – A – 2 – d – ANN inputs, outputs and topology

To interpolate the 3DLC, the patient's height must be known. Actually, this is one of the descriptors of the EquiVox target and source cases. Let us note h_i the descriptor corresponding to the height of the case *i* and h_t, the height of the target case *t*.

An ANN with 9 inputs and 3 outputs was designed. Two phantoms were considered:

- The source case *inf* for which h_{inf} is inferior and the closest to h_t;
- The source case *sup* for which h_{sup} is superior and the closest to h_t.

The trained ANN interpolates the 3 coordinates of each point of the lung contours separately. Thus, the 9 inputs permitting interpolation of the coordinates C_k^{lung} of point *k* of *t* are:

- The 3 coordinates of point *k* of the lung contours of *inf*: $C_k^{inf,lung} = \{x_k^{inf,lung}, y_k^{inf,lung}, z_k^{inf,lung}\}$;
- The height of *inf*: h_{inf};
- The 3 coordinates of point *k* of the lung contours of *sup*: $C_k^{sup,lung} = \{x_k^{sup,lung}, y_k^{sup,lung}, z_k^{sup,lung}\}$;
- The height of *sup*: h_{sup};
- The height of the target case: h_t.

The designed ANN is perceptron having one hidden layer. Ten neurons are on the hidden layer with a sigmoid activation function. The activation function of the neurons belonging to the output layer is linear.

Such topologies were also chosen and successfully tested on the NEMOSIS platform (25) for a similar issue: considering a point inside the patient's lung, at both the initial and final position (maximum and minimum respiration respectively), a similar ANN interpolated the positions of the point during an entire breathing cycle with an error inferior to the spatial resolution of the medical images on which the point had been plotted. For NEMOSIS, the number of neurons on the hidden layer was optimised using a validation set in addition to the learning set used for the learning step. In the case of EquiVox, regarding the small number of 3DLC (12 3DLC), we decided not to consider a validation set. Thus, the learning set was composed of 9 3DLC, while the 3 remaining 3DLC belonged to the test set. We assumed that such a topology would deliver sufficiently accurate results. Nevertheless, such a strategy to optimise the number of neurons on the hidden layer would have to be implemented in later work.

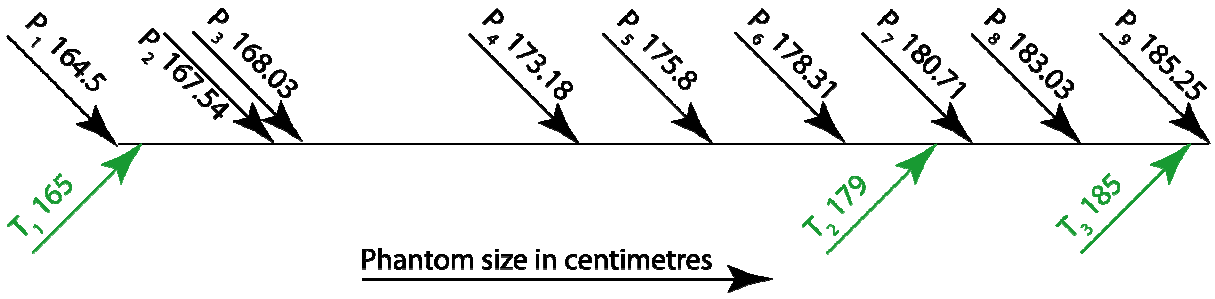

Figure 2. Available 3DLC.

In Figure 2, the 9 heights of the 3DLC P_1 to P_9 used for the training are reported on the axis. The 3 heights of the 3 new 3DLC T_1, T_2 and T_3 are also reported on the same axis. All the thorax organs are represented in P_1 to P_9 whereas only the lungs were drawn in T_1, T_2 and T_3.

II – A – 2 – e – ANN learning set and training step

Training ends when the difference between the expected and the obtained values is minimised. W. Hsieh (26) distinguished four algorithms based on the backpropagation method:

- The BFGS method (Broyden-Fletcher-Goldfarb-Shanno) is a quasi-Newton method, which approximates the value of the Hessian matrix of the second derivatives of the function to be minimised;
- The L-BFGS method (Limited memory – BFGS) is an adaptation of the BFGS method which optimises the computational resources to use. Both of these methods must be coupled with a Wolfe linear search in order to determine an optimal step size between two iterations;
- The Rprop (Resistant backpropagation) method proposes a first order algorithm, but its complexity increases linearly with network topology;
- The iRpropPlus method is one of the fastest and also one of the most accurate algorithms (27). This evolution of the Rprop method allows some synaptic weight updates to be cancelled in the neural network if a negative effect is observed.

	Phantom height	Required precision	Best Learning method
ANN interpolations	1783.1	1E-006	BFGS
	1807.1	1E-006	BFGS
	1830.3	1E-006	BFGS

Table 1. ANN configuration (learning method and required precision) obtaining the best preliminary results.

All of these methods were previously implemented and tested in the EquiVox adaptation phase of 3DLC. Different required precisions were also tested. The coordinates of 10 points were randomly extracted from the 3DLC of P_1 to P_9 and a cross validation was performed. Table 1 shows that the algorithm giving the best interpolations is that which used BFGS as backpropagation method and that obtained a precision equal to 10^{-6}. Thus, the chosen ANN configuration has been compared to a polynomial (Newton, of degree 2) and a Spline interpolation method. The Newton interpolation function proposed by J. Ponce and R. Brette (28) and the Spline one proposed by Scilab (29) were

implemented with Scilab 5.3.2. For each method, a cross-validation for the same 10 points was undertaken using the same 3DLC of P_1 to P_9. Figure 3 presents the mean distances between interpolated and expected coordinates. This figure shows that the polynomial interpolation produced the greatest errors among the three tested interpolations. A factor nearly equal to 10 can be observed between the polynomial interpolation and that of the Spline or the ANN. The Spline and the ANN interpolations gave closer errors. Nevertheless, for all the tested cases, the ANN interpolation errors were inferior to the Spline ones 6 times and were equal only once. These results prove the superiority of the ANN interpolations over the other methods since the ANN interpolation gave a more accurate result in all the tested cases.

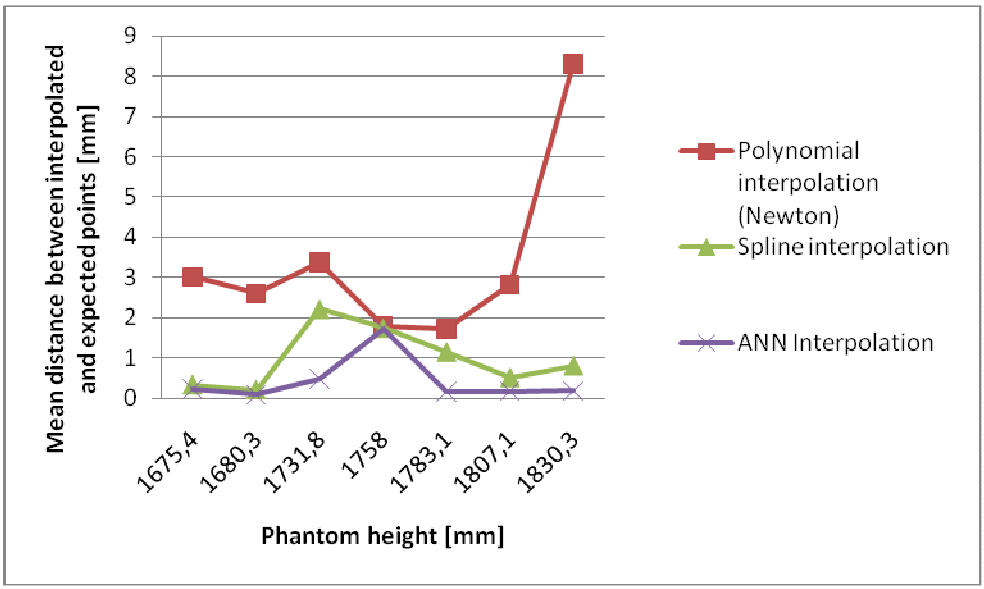

Figure 3. Mean distances obtained between interpolated and expected coordinates for 10 points and 3 interpolation algorithms.

II – B – The Combination Vector for Interpolation Optimisation

II – B – 1 – Previous performance of the ANN interpolation of EquiVox

Tests were previously performed with the entire Learning Set $LS=\{P_1, \ldots, P_9\}$ to interpolate the Test Set $TS = \{T_1,\ldots, T_3\}$. The heights of each T_i has been carefully chosen to test all the possible cases: as shown in Figure 2, the height of T_1 is just above the smaller stored one (P_1), the third one (T_3) is just below the higher stored one (P_9), and the second (T_2) is in the middle of the stored panel of heights. Since the 3DLC alone were designed for the 3DLC of *TS*, these ones were used only for these last tests and not stored in the EquiVox case-base, nor were they used during the ANN learning. In addition, the same manual creation process was followed for all 3LDC of *LS* and *TS*.

Value of the descriptor "height" of the target case (h_t) [mm]	Mean error [mm]	Std error [mm]
1650	1.65	0.53
1790	0.55	0.18
1850	0.80	0.37

Table 2. Mean error between ANN and expected outputs associated with target case height.

Table 2 shows the mean distances observed between the interpolated and expected points for each h_t. The mean distances vary from 0.55 mm to 1.65 mm. The calculations of dosimetric reports are usually computed using a voxelised phantom. The commonly used voxel dimensions are 1.8 mm by 1.8 mm by 4.8 mm (19) (20). The largest mean error is equal to 1.65 mm (inferior to the spatial resolution of commonly used phantoms). Thus, all the 3DLC generated by this ANN can be used to establish dosimetric reports.

II – B – 2 – Optimisation of the ANN learning set

Nevertheless, this adaptation strategy is based on 3DLC that may contain errors in comparison with the expected lung contours of a real subject: these contours are already representations of reality with uncertainties. Thus, biases might be introduced by one or more incorrectly designed 3DLC. The ANN implemented in the EquiVox adaptation phase for 3DLC may reduce the impact of these errors since an ANN is an interpolation tool, but the purpose of this study is to verify the accuracy of *LS* to construct one particular 3DLC. We explored the possibility that a sub-set of *LS* could give more accurate results.

We introduced *u* vectors V_i to stipulate weather each 3DLC was used or excluded from *LS:*

$$V_i \in V = \{(\{0,1\}^9)\}$$

We also introduced two functions, for the local *c()* and global *C()* combinations respectively and defined as:

$$c: LS \times \{0,1\} \to LS \mid \forall i \in \{1,\dots,9\} \begin{cases} c(P_i,0) = \emptyset \\ c(P_i,1) = P_i \end{cases}$$

$$C: V \to LS \mid C(v) = \bigcup_{i=1}^{9} \{c(P_i, v_{i-1})\}$$

II – B – 3 - Cardinality

As explained in Sub-section II-A-2-d, not all the possible learning sets obtained with *C()* allowed the interpolation of the 3DLC of *TS*. Let's note $Card_i$ the cardinality of $V_{T_i} \subset V$ which allows the construction of T_i, and *L* the cardinality of *LS*.

Proposition

$$If\ L \geq 3\ and\ \exists\ j \in \{1,\dots,L\} \mid h_j < h_t < h_{j+1}\ then\ Card_t = 2^L - \sum_{i=0}^{2} C_L^i - \sum_{i=3}^{j} C_j^i - \sum_{i=3}^{L-j} C_{L-j}^i$$

.

Proof

On the one hand, in order to interpolate T_k, a minimum of 3 3DLC must be included in the learning set, and a minimum of one 3DLC in $LS_{inf} = \bigcup_{i=1}^{j}\{P_i\} \mid h_j < h_t < h_{j+1}$ and one in $LS_{sup} = \bigcup_{i=(j+1)}^{L}\{P_i\} \mid h_j < h_t < h_{j+1}$.

- There are 2^L vectors in *V*;
- There are $\sum_{i=0}^{2} C_L^i$ possibilities such that fewer than 3 3DLC are chosen among *LS*;
- $Card(LS_{inf}) = j$;
- $Card(LS_{sup}) = L - j$;
- There are $\sum_{i=3}^{j} C_j^i$ possibilities such that exactly 3 to *j* 3DLC of LS_{inf} are chosen and no LS_{sup}.
- There are $\sum_{i=3}^{L-j} C_{L-j}^i$ possibilities such that exactly 3 to (*L-j)* 3DLC are chosen of LS_{sup} are chosen and no LS_{inf}.

Consequently,

$$Card_t = 2^L - \sum_{i=0}^{2} C_L^i - \sum_{i=3}^{j} C_j^i - \sum_{i=3}^{L-j} C_{L-j}^i$$

.

II – B – 4 – New ANN learning algorithm

A new algorithm for the learning phase was also implemented. Indeed, in the first algorithm LA_{global}, the backpropagation ended when the global mean square error between expected and computed coordinates considering all the points of *all* the 3DLC was inferior to 10^{-6} (cf. sub-section II-A-2-e). Thus, one mean square error was found for the entire learning set in LA_{global}.

In the new version of the learning phase, the algorithm LA_{CVIO} computed the mean square error between expected and computed coordinates considering all the points of *each* 3DLC. Thus, in this new version, if the vector was v, there were $Card(C(v))$ mean square errors, and the backpropagation algorithm was applied until each error was inferior to 10^{-6}.

III – Results

The results presented in Figures 4, 5 and 6 show the accuracy obtained according to the Combination Vector (CV). In these figures, the red lines are the accuracies obtained using $CV_1 = (1,1,1,1,1,1,1,1)$ and LA_{global}, and the green lines the accuracies obtained using CV_1 and LA_{CVIO}. Only accuracies inferior to those obtained with CV_1 and LA_{global} are reported.

III – A - Results for TC_1

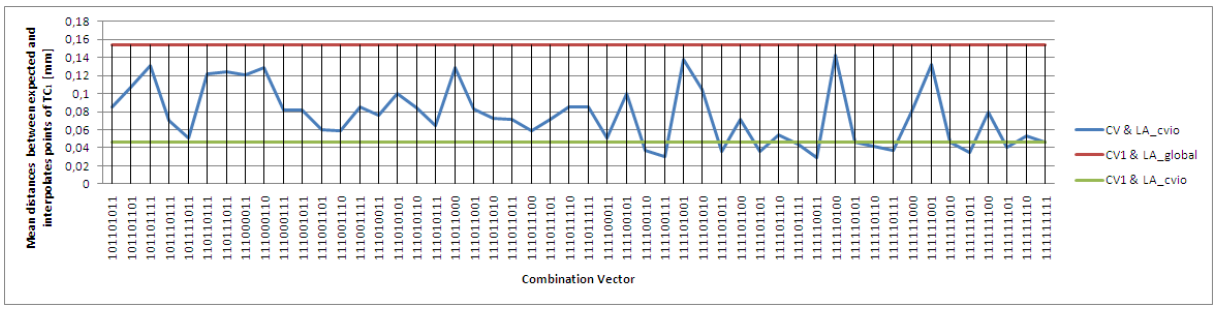

Figure 4. Best results obtained interpolating the 3DLC of T_1.

Figure 4 shows the lesser errors obtained and the CV used with LA_{CVIO} to interpolate T_1. The error obtained with CV_1 and LA_{global} is about 0.16 mm (the red line). The error obtained with CV_1 and LA_{CVIO} is about 0.04 mm (the green line). Figure 4 shows that 51 CV showed errors inferior to CV_1 and LA_{global}. We can also note that 10 CV among these 51 allowed interpolation of T_1 with smaller errors than CV_1 and LA_{CVIO}. These CV excluded some of the 3DLC from *LS*.

Rank	CV P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9
1	(1,	1,	1,	1,	1,	0,	0,	1,	1)
2	(1,	1,	1,	1,	0,	0,	1,	1,	1)
3	(1,	1,	1,	1,	1,	1,	0,	1,	1)
4	(1,	1,	1,	1,	0,	0,	1,	1,	0)
5	(1,	1,	1,	1,	1,	0,	1,	1,	1)
6	(1,	0,	1,	1,	1,	0,	1,	1,	1)

Table 3. CV of the « Top 6 » for T_1.

The CV of the "top 6" are reported in Table 3. The best CV for the interpolation of T_1 are the ones that excluded P_6 and P_7, and used all the others. Also noteworthy is that the second configuration excluded P_5 and P_6, and included all the others. This last CV is also one of the best for T_2 and T_3. In addtion, we note that CV_1 does not figure in this "Top 6": it was not necessary to learn how to construct all the 3DLC to obtain the most accurate 3DLC for T_1. 3 CV excluded 2 3DLC, 2 CV excluded 1 3DLC, and 1 CV excluded 3 3DLC. Finally, only one CV of this "Top 6" included P_6.

III – B – Results for TC_2

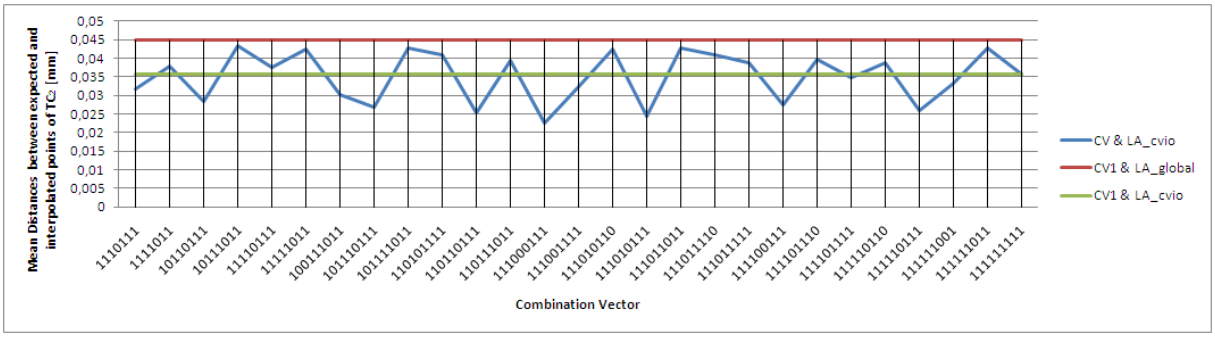

Figure 5. Best results obtained interpolating the 3DLC of T_2.

Figure 5 reports that the CV with which the interpolated 3DLC of T_2 were more accurate than the 3DLC interpolated using CV_1 and LA_{global}. The error obtained using CV_1 and LA_{global} was superior to 0.04 mm (the red line) whereas the error obtained using CV_1 and LA_{CVIO} was inferior to 0.04 mm (the green line). The error with 27 CV were inferior to CV_1 and LA_{global}, and 10 CV among them excluded some of the 3DLC of *LS*.

Rank	CV P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9

1	(1,	1,	1,	0,	0,	0,	1,	1,	1)
2	(1,	1,	1,	0,	1,	0,	1,	1,	1)
3	(1,	1,	0,	1,	1,	0,	1,	1,	1)
4	(1,	1,	1,	1,	1,	0,	1,	1,	1)
5	(1,	0,	1,	1,	1,	0,	1,	1,	1)
6	(1,	1,	1,	1,	0,	0,	1,	1,	1)

Table 4. CV of the « Top 6 » for T_2.

Table 4 presents the 6 best CV: 4 of them excluded 2 3DLC, 1 CV excluded 1 3DLC, and 1 CV excluded 3 3DLC. The CV which excluded P_5 and P_6 appears also in this Table ranked #6. Furthermore, P_6 is systematically excluded from the CV of this "Top 6".

III – C – Results for TC_3

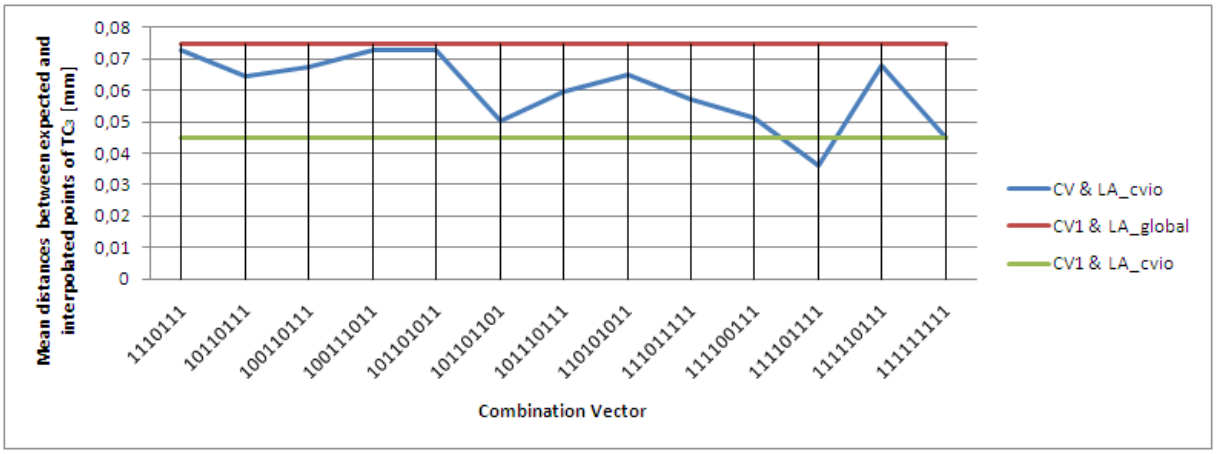

Figure 6. Best CV for the interpolation of T_3.

Figure 6 shows the 13 CV that allowed greater accuracy than CV_1 and LA_{global} for the interpolation of T_3. The mean square error obtained interpolating T_3 using CV_1 and LA_{global} was equal to 0.07 mm (the red line) whereas 0.04 mm using CV_1 and LA_{CVIO}. One subset of *LS* permitted greater accuracy than CV_1 and LA_{CVIO}.

Rank	CV P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9
1	(1,	1,	1,	1,	0,	1,	1,	1,	1)
2	(1,	1,	1,	1,	1,	1,	1,	1,	1)
3	(1,	0,	1,	1,	0,	1,	1,	0,	1)
4	(1,	1,	1,	1,	0,	0,	1,	1,	1)
5	(1,	1,	1,	0,	1,	1,	1,	1,	1)
6	(1,	0,	1,	1,	1,	0,	1,	1,	1)

Table 5. CV of the « Top 6 » for T_3.

Table 5 shows the "Top 6" best CV: 2 CV excluded 2 3DLC, 2 CV excluded 1 3DLC, and 1 CV excluded 3 3DLC. CV_1 ranks #2 and the CV that excluded P_5 and P_6 is at rank #4. For this 3DLC, P_6 was excluded twice from the CV of this "Top 6".

These results prove that it was possible to optimise the ANN interpolation excluding some of the 3DLC from the learning set. The CVIO improved the accuracy of the construction of personalised 3DLC.

III – D – 3DLC inclusion and exclusion

Nevertheless, the example of P_6 is interesting: P_6 was excluded from most of the best learning sets for T_1 and T_2 whereas it was required in most of the best learning sets for T_3. Thus, we studied the impacts of the 3DLC on the interpolation accuracy of T_1, T_2, and T_3. For each 3DLC of *TS*, we extracted the "Top 20" best and the "Top 20" worst CV. We then counted the number of times each 3DLC of *LS* appeared in the "Top 20" best and in the "Top 20" worst.

Rank	In the "Top 20" best CV	In the "Top 20" worst CV
1	P_1, P_3	P_1
2	P_2	P_2
3	P_4	P_7, P_8
4	P_8	P_3
5	P_7	P_6, P_9
6	P_6, P_9	P_5
7	P_5	P_4

Table 6. 3DLC rankings for T_1.

Table 6 shows the ranking for T_1. P_1 and P_3 were almost always used in the best learning sets. P_2 ranks #2, before P_4, P_8, P_7, P_6 (ex-aequo with P_9), and finally P_5. This ranking has to be put in the perspective of the ranking for the worst learning sets. For the former, P_1 was also always used, then come P_2, P_7 ex-aequo with P_8, before the others in the following order: P_3, P_6 and P_9, P_5 and P_4. Since it is not possible to interpolate T_1 without P_1, it is absolutely normal for $P_1$to be first for these 2 rankings. We can note that P_2 and P_4 are in good positions in the "Top 20" best and in the worst position in the "Top 20" worst. In contrast, P_7 is among the worst ranks of the bests and the best ranks of the worsts. Thus, P_7 seems to introduce a bias in the learning set. Consequently, we can deduce from Table 6 that P_1, P_3 and P_4 are required whereas P_7 is to be excluded from the learning set in order to optimise the interpolation of T_1.

Rank	In the "Top 20" best CV	In the "Top 20" worst CV
1	P_8	P_3
2	P_9	P_1, P_2, P_7
3	P_7	P_9
4	P_4, P_3	P_4, P_8
5	P_1, P_2, P_3	P_5
6	P_6	P_6

Table 7. 3DLC rankings for T_2.

Table 7 presents the rankings for T_2. P_8 is at the top of the best, before P_9, P_7, P_3 and P_4. Then come P_1, P_2, P_5 and P_6, whereas P_3 is at the top of the worst before P_1, P_2 and P_7. Then come P_9, P_4 and

P_8 (ex-aequo), and finally P_5. Thus, we can note the positive influence of P_8 whereas P_1, P_2 and P_3 seem to perturb the accuracy of the interpolations of T_2.

Rank	In the "Top 20" best CV	In the "Top 20" worst CV
1	P_9	P_9
2	P_1	P_3
3	P_8	P_2
4	P_4	P_1, P_4
5	P_6, P_7	P_5
6	P_3	P_6
7	P_2, P_5	P_7, P_8

Table 8. 3DLC rankings for T_3.

As it is always required for the interpolation of T_3, P_9 is at the head of the rankings in Table 8. For the best learning sets, this 3DLC comes before P_1, P_8, P_4, P_7 ex-aequo with P_6, followed by P_3, and finally P_2 and P_5 (ex-aequo). For the "Top 20" worst, P_9 is followed by P_3, P_2, P_1 and P_4, P_5, P_6, P_7 and P_8. Thus, P_8 is required whereas P_3 and P_2 should be excluded in order to optimise the interpolation of T_3.

Rank	In the "Top 20" best CV	In the "Top 20" worst CV
1	P_3	P_9
2	P_2, P_4, P_8	P_6, P_8
3	P_5	P_2, P_7
4	P_7	P_1, P_4, P_5
5	P_1, P_6	P_3
6	P_9	

Table 9. 3DLC rankings for T_1, T_2, and T_3.

At last, Table 9 presents the global ranking for all the 3DLC of *TS*. P_3 was used in most of the best learning sets, then come P_2, P_4 and P_8, before the others. P_1, P_6 and P_9 are the last of this ranking. P_9 was used in most of the worst learning sets, then P_6 ex-aequo with P_8. At last P_4, P_5 and P_3 appear. These last rankings tend to prove that the use of P_3 in the learning phase ensures good interpolations, but it is the opposite when considering the rankings of T_2 and T_3. Furthermore, P_6 and P_9 seem to introduce a bias in the general case. This tends to prove that it is necessary to divide the definition domain into sub-domains since the learning of each 3DLC of *LS* has a very different influence over the interpolation of each 3DLC of *TS*.

IV – Discussion

This study enables us to imagine a new kind of adaptation strategy between the one based on rules and the Conservative one: the *Fuzzy Adaptation*. Indeed, in our application domain, there are different and (sometimes) conflicting rules (described by different trained ANN) that can be applied to compute a solution. Moreover, some of the source cases introduced distortions for the

interpolation of one particular case, whereas they were strongly required in order to obtain an accurate result in other cases.

Noting $sol(pb^s)$ as the solution to a problem of the source case pb^s, and pb^t the problem of the target case, an adaptation rule is commonly defined as a set of 2 elements (r, A_r) where *r* is a relation between pb^s and pb^t, and $A_r(pb^s, pb^t, sol(pb^s))$ a function that computes a solution corresponding to pb^t. The adaptation using the rules consists of factorising the solution of the target case into solutions for elementary problems.

In our application, more than one rule can be used for adapting a problem: considering $(pb^s, pb^t, sol(pb^s))$, the set of rules R, a rule $R_i \in R$ is a set of 2 elements $(r^i, A_r{}^i)$. Our ambition is to explore the possibility of combining and creating new rules from the set of rules specified by experts $R^{exp} \subset R$.

For this purpose, we define the operator $+_R$ as the operator of *combination* for 2 rules and the operator $._R$ as the operator of *restriction*.

$+_R : R \times R \rightarrow R$ and $._R : FuzzyCV \times R \rightarrow R$, where $FuzzyCV$ is a set of *Fuzzy Combination Vectors.*

It is now possible to define $Card(R^{exp})$ weights w_i associated to each $R_i \in R^{exp}$.

Consequently, the adaptation of a problem in the Fuzzy Adaptation will be resolved through a rule defined as a combination of a set of rules specified by experts: $R_c = +_{R}{}_{(i=0)}^{(Card(R^{exp}))}(w_i ._R R_i)$.

A future study will have to specify in greater detail the terms of the combination and restriction operators.

V – Conclusion

We proposed a study of the construction of patient 3DLC interpolating a subset of the entire learning set (*LS*) available to us. The results obtained show the value of this method since it stresses the optimisation of accuracy when the interpolation tool is learned using a subset of *LS*. These subsets depended on the person's height. This new algorithm (LA_{CVIQ}) forced the learning phase to continue until the mean square error between the coordinates expected and interpolated of each 3DLC of the subset was inferior to a determined precision, whereas the older version of the learning algorithm (LA_{global}) took into consideration the global mean square error (considering all the 3DLC).

The best subset of *LS* for each 3DLC of our Test Set (*TS*) was different from the others: one 3DLC guaranteed optimal precision for the interpolation of one 3DLC of *TS* when it was used in the learning phase, whereas the same one introduced a bias for the interpolation of another 3DLC of *TS*. Nevertheless, the sub-learning set which used all the 3DLC of *LS*, without P_5 and P_6, seemed to guarantee optimal precision for most of the 3DLC of *TS*.

Consequently, we must continue to study the influence of each 3DLC on the definition domain first. Secondly, a particular effort will be devoted to the improvement of LA_{CVIQ} and turning the *Combination Vector* (CV) into a *Fuzzy CV* (a vector of pertinence for each 3DLC of *LS*). The

determination of the components of the Fuzzy CV becomes a more complex problem that could be treated through the Genetic Algorithm and/or use of metaheuristics.

Acknowledgments

The authors want to thank the experts of IRSN David Broggio, Jad Farah and Didier Franck for their help with the phantoms, the SFRP and the *Ligue Contre le Cancer* for their financial help, and Michel Salomon, Marc Sauget, Brigitte Chebel-Morello, Libor Makovicka, Rémy Laurent, Emmanuelle Fontaine, Hamza Bouhelal and the Mésocentre of the University of Franche-Comté for their help during the implementation of the concepts.

References

1. D. Broggio, B. Zhang, L. de Carlan, A. Desbrée, S. Lamart, B. le Guen, C. Bailloeuil, D. Franck. *Analytical and Monte Carlo assessment of activity and local dose after a wound contamination by activation products.* Health Physics, Vol. 96, pp. 155-163, 2009.

2. C. Huet, A. Lemosquet, I. Clairand, J.B. Rioual, D. Franck, L. de Carlan, I. Aubineau-Lanièce, J.F. Bottollier-Depois. *SESAME: a software tool for the numerical dosimetric reconstruction of radiological accidents involving external sources and its application to the accident in Chile in December 2005.* Health Physics, Vol. 96, pp. 76-83, 2009.

3. J. Kolodner. Case-Based Reasoning. *Morgan Kaufmann Publishers.* 1993.

4. E. Melis, J. Lieber, A. Napoli. *Reformulation in Case-Based Reasoning.* Forth European Workshop on Case-Based Reasoning, Lecture Notes in Artificial Intelligence, Springer, Vol. 1488. pp. 172-183, 1998.

5. B. Fuchs, J. Lieber, A. Mille, A. Napoli. *An Algorithm for Adaptation in Case-Based Reasoning.* Proceedings of the 14th European Conference on Artificial Intelligence, Berlin, pp. 45-49, 2000.

6. J. Lieber. *Application of the Revision Theory to Adaptation in Case-Based Reasoning: the Conservative Adaptation.* Proceedings of the 7th International Conference on Case-Based Reasoning, Lecture Notes in Artificial Intelligence, Springer, Vol. 4626. pp. 239-253, 2007.

7. M. Bopp, J. Henriet, B. Chebel-Morello, L. Makovicka, D. Broggio. *Preliminary Study of a New CBR-Based Application for Voxelised Phantom Creation: ReEPh*. Singaporean-French IPAL Symposium (SinFra'09), pp. 107-115, 2009.

8. J. Henriet, J. Farah, B. Chebel-Morello, M. Bopp, D. Broggio, L. Makovicka. *Feasability Study of a New Platform Based on the Case-Based Reasoning Principles to Efficiently Search and Store Voxel Phantoms*. Radioprotection vol. 45(1), pp. 67-82. 2010.

9. W. McCulloch, W. Pitts. *A logical calculus of ideas immanent in nervous activity.* Bulletin of Mathematical Biophysics. Vol. 5, pp. 115-133, 1943.

10. J. Henriet, L. Makovicka, M. Salomon, M. Sauget, D. Broggio. *EquiVox.* Software declaration #4047-01, 2011.

11. J. Farah, J. Henriet, D. Broggio, R. Laurent, E. Fontaine, B. Chebel-Morello, M. Sauget, M. Salomon, L. Makovicka, D. Franck. *Development of a New CBR-Based Platform for Human*

Contamination Emergency Situations. Radiation Protection Dosimetry (RPD), Vol. 144. pp. 564-570, 2011.

12. ICRP89**.** *Basic anatomical and physiological data for use in radiological protection*. International Commission on Radiological Protection Publication 89*.* 2002.

13. ICRUReport48**.** *Phantoms and Computational Models in Therapy, Diagnosis and Protection*. International Commission on Radiation Units and Measurements Report 48*.* 1992.

14. R. Kramer, M. Zankl, G. Williams, G. Dexter. *The calculation of dose from external photon exposures using reference human phantoms and Monte Carlo methods. Part I: the male (Adam) and female (Eva) adult mathematical phantoms*. Report GSF-Bericht S-885*.* 1982.

15. G.I. Tanaka, H. Kawamura, Y. Nakahara. *Reference Japanese man-I. Mass of organs and other characteristics of normal Japanese*. Health Physics, vol. 36, issue 3, pp. 333–346, 1979.

16. X.G. Xu, T.C. Chao, A. Bozkurt. *VIP MAN, an imaged-based wholebody adult male model constructed from color photographs of the visible human project for multi-particle Monte Carlo calculations*. Health Physics, vol. 78, issue 5, pp. 476-486, 2000.

17. M. Zankl, R. Viet, G. Williams, K. Schneider, H. Fendel, N. Petoussi, G. Dexler. *The construction of computer tomographic phantoms and their application in radiology and radiation protection*. Radiation Environmental Biophysics, vol. 27, pp. 153–164, 1988.

18. I.G. Zubal, C.R. Harrell, E.O. Smith, Z. Rattner, G. Gindi, P.B. Hoffer. *Computerized three dimensional segmented human anatomy*. Medical Physics, vol. 21, issue 2, pp. 299–302, 1994.

19. J. Farah, D. Broggio, D. Franck. *Examples of Mech and NURBS phantoms to study the morphology effect over in vivo lung counting.* Radiation Protection and Dosimetry Special Issue, Vol. 144, pp. 344-348, 2011.

20. J. Farah, D. Broggio, D. Franck. *Creation and use of adjustable 3D phantoms: application for the lung monitoring of nuclear workers*. European Conference on Individual Monitoring of Ionizing Radiation*.* Athens, Greece, 2010.

21. McNeel**.** *Rhinoceros Modeling tools for designers*. http://www.rhino3d.com*.* [Online]

22. G.E. Christensen. *Deformable shape models for anatomy.* Washington University. PhD Thesis, 1994.

23. I. Clairand, L. Bouchet, M. Ricard, M. Durigon, M. Di Paola, B. Aubert. *Improvment of internal dose calculations using mathematical models of different adult heights.* Phys. Med. Biol., Vol. 45, pp. 2771-2785, 2000.

24. Dassault Systems**.** *CATIA - Virtual Design for Product Excellence*. http://www.3ds.com/products/catia/welcome*.* [Online]

25. R. Laurent, J. Henriet, M. Salomon, M. Sauget, R. Gschwind, L. Makovicka. *Respiratory Lung Motion using an Artificial Neural Network.* Neural Computing and Applications. DOI 10.1007/s00521-011-0727-y, 2011.

26. W. Hsieh. *Learning Methods in the Environmental Sciences - Neural Networks and Kernels.* Cambridge University Press, 2009.

27. A. Vasseur, L. Makovicka, E. Martin, M. Sauget, S. Contassot-Vivier, J. Bahi. *Dose calculations using artificial neural networks: A feasibility study for photon beams.* Nuclear Instruments and Methods in Physics Research Section B: Beam interactions with materials and atoms, Vol. 266, pp. 1085-1093, 2008.

28. J. Ponce, R. Brette. *Polynomial interpolation*. *Introduction to scientific computing and its applications.* http://audition.ens.fr/brette/calculscientifique/2006-2007/lecture2.pdf [Online].

29. Digiteo**.** *Scilab Home Page*.http://www.scilab.org [Online].