Reducing false discovery rates for on-line model checking based detection of respiratory motion artifacts

Sven-Thomas Antoni[1] Xintao Ma[2] Sibylle Schupp[2] Alexander Schlaefer[1]

Abstract: Compensating respiratory motion in radiosurgery is an important problem and can lead to a more focused dose delivered to the patient. We previously showed the negative effect of respiratory artifacts on the error of the correlation model, connecting external and internal motion, for meaningful episodes from treatments with the Accuray CyberKnife®. We applied on-line model checking, an iterative fail safety method, to respiratory motion. In this paper we vary its prediction parameter and decrease the previously rather high false discovery rate by 30.3%. In addition, we were able to increase the number of detected meaningful episodes through adaptive parameter choice by 452%.

Keywords: On-line model checking, respiratory motion compensation, prediction, fail safety, event detection, radiosurgery, stereotactic body radiation therapy

1 Problem statement

In radiosurgery typically high radiation doses are used to treat cancer. To reduce side effects, the treatment is planned in advance to tightly follow the tumor's contour. Some tumors can move substantially due to respiratory motion and hence approaches for motion compensation have been proposed. Currently, the tumor's motion can not be tracked directly but instead the external respiratory motion is measured. Several methods to predict external and correlate internal motion exist [SGB+00, SSA04, SBN+07, DHV+10, EDSS13].

The data used in this work was recorded during treatments with the CyberKnife® system (CK, Accuray Inc.). The CK is an image guided radiosurgery system. Using the additional Synchrony™ add on the CK is able to track and compensate for respiratory motion in real time. Therefore, gold fiducials are placed within or near the tumor in advance of the treatment. Throughout a treatment session the fiducials are tracked at discrete times using diagnostic X-ray. In addition, external markers are tracked using an optical system. A correlation model (CM) is used to estimate the internal from the external motion. The CM is initialized at the beginning of a treatment session. Over time, changes in the correlation between internal fiducials and external markers, e.g., due to baseline drift may appear [HNLH08]. Therefore, the CM needs to be updated regularly. Still, sudden respiratory

[1] Hamburg University of Technology, Institute of Medical Technology, Am-Schwarzenberg-Campus 3E, 21073 Hamburg, {antoni,schlaefer}@tuhh.de
[2] Hamburg University of Technology, Institute for Software Systems, Am-Schwarzenberg-Campus 3E, 21073 Hamburg, {xintao.ma,schupp}@tuhh.de

artifacts may lead to distinct deviations between tumor and beams that lead to recreation of the CM or may not be detected by the system [SGB+00, SSA04, SBN+07, ARSS15b].

To overcome latencies in the system, prediction of external motion is used. New methods are able to predict even irregular data with little error [EDSS13]. While this reduces the prediction error our research indicates that the CM may become invalid [ARSS15b].

On-line model checking (OMC) is a new iterative fail safety method. At discrete time intervals OMC validates the input data against a previously defined model which parameters are derived from the data history. Recently, OMC was introduced as a validation method for respiratory motion. While in principle OMC for artifact detection is feasible, a rather high false discovery rate (FDR) was observed [RSG14a, RSG14b, ARSS15a, ARSS15b].

For this work we advance the detection of meaningful episodes showing high correlation errors in the data. In addition, we focus on reducing the FDR of OMC, which we achieve without sacrificing the overall accuracy. Due to the improved episode detection, we are able to work on a significantly smaller dataset than previous works.

2 Materials and methods

2.1 CyberKnife treatment data

Our data contains time series of the surrogate and fiducial marker positions s_i and f_i, as well as the expected fiducial marker positions g_i. The error $e_i = \|f_i - g_i\|$ and the mean error $\overline{e}$ with respect to the fiducial position are used to select interesting episodes. Particularly, we consider episodes with three subsequent small errors followed by one large error. Let $d_i = |e_i - \overline{e}|$ and σ_d be the absolute and standard deviation from the average error, respectively. Using the 90% quantile $q = Q_{0.9}(d)$ we define thresholds $t_l = \overline{d_i}|d_i < q$ and $t_u = t_l + \sigma_d$, where errors below t_l are considered small and errors above t_u are considered large.

We use the two thresholds to derive physiologically meaningful episodes $[i-3, i]$ that exhibit an amplitude of at least 2mm and 6mm in the first principal component (PC) of external and internal motion, respectively, and fulfill $\{i\colon\ e_{i-3}, e_{i-2}, e_{i-1} \leq t_l \wedge e_i > t_u\}$.

2.2 On-line Model Checking

During OMC the breathing motion is predicted and the respiratory model is periodically evaluated.

The prediction model of the chest movement is generated based on a limited history of data which we assume to indicate the movement in the future. The prediction combines discrete Fourier series and linear regression

$$x(t) = d \cdot t + \sum_{i=0}^{4} c_i \cos(i \cdot f \cdot t) + s_i \sin(i \cdot f \cdot t) \pm \beta. \quad (1)$$

Here, β is a deviation factor that allows adopting the prediction to different breathing patterns. To increase the possibility of correctly predicting the respiration, multiple possible movement trajectories are generated. For smaller values of an accuracy parameter $\alpha \in [0,100]$ more trajectories are simulated. The influence of the deviation parameter β is varied in two ways: by varying the actual value of β and by varying the interval $\tau_{lb} = \{[lb, 1000] \colon lb \in \mathbb{Z}\}$ at which a deviation is applied. Previous implementations of OMC did not allow changing lb, which was always set to 1 (Fig. 1a). For this work we focus on varying the parameter lb. For small lb, deviations are added earlier during prediction. For episodes of regular breathing, however, the frequency of irregularities in motion is small. Hence, a larger value of lb is appropriate as it increases the probability to correctly predict regular motion. On the other hand, the value of lb should not be too large. Otherwise, too few trajectories are generated, making it harder to account for natural small irregularities in respiratory motion.

For validation we estimate how likely we can predict the actual value x_0 at time t_0 with the probability

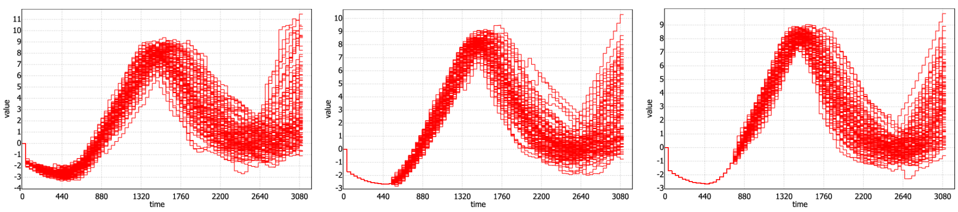
$$Pr[t_0 - t_I \leq t_p \leq t_0 + t_I \wedge x(t_p) - x_I \leq x(t_p) \leq x(t_p) + x_I], \tag{2}$$

where $x(t_p)$ is the predicted value at time t_p and x_I and t_I define a rectangular region around the actual value and the measured time. The probability is estimated for the three PCs of the marker coordinates placed on the patient's chest. In a second step, we capture how often in a row the probability is lower than a threshold θ. The validation result is positive if the probability falls below the safety threshold for three times in a row, indicating there is an artifact of the patient. Otherwise, the result is negative, indicating normal breathing.

3 Results

For the modified selection of episodes we used data of 194 sessions that showed 23 episodes before [ARSS15b]. With the improved episode selection we are now able to identify 104 episodes; an increment of about 452%.

For the OMC parameter variation we investigated treatment sessions of 6 patients of a total mean duration of 35 minutes. Values of t_l ranged from 0.83mm to 3.86mm, t_u from 1.41



(a) (b) (c)

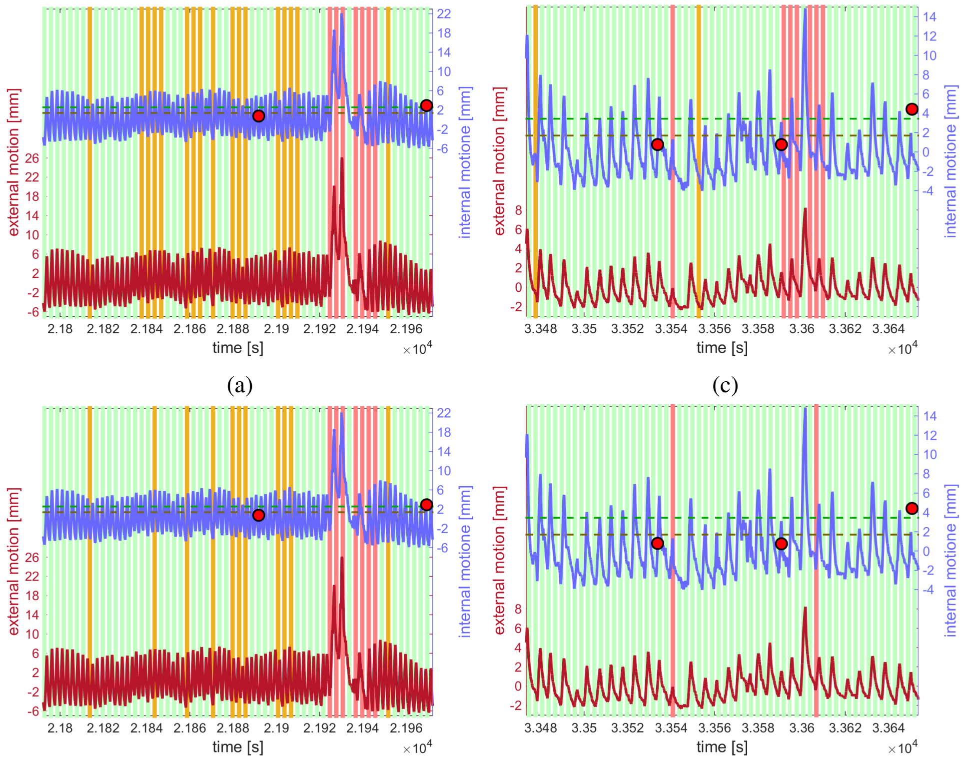
Fig. 1: Examples of predictions with different τ_{lb}, $lb = 1,500,750$ (left to right). A higher lb generates fewer trajectories indicated by a smaller corridor width, resulting in a higher probability to predict the actual value.

to 7.75. OMC parameters were set to $\alpha = 70$, $\theta = 30\%$, $t_I = 0.5ms$ and $x_I = 0.5mm$. We varied the time interval in four steps τ_{lb}, $lb = 1, 250, 500, 1000$.

We identified 6 episodes fulfilling all criteria. Every episode showed a clear artifact prior to the high CM error that is detected by the OMC for every τ_{lb}, see Fig. 2a-b at approximately 21930s and 2c-d at approximately 33540s and 33600s for examples. The mean FDR of OMC validation of those 6 episodes ranged from 50.91% with τ_1 to 35.29% with τ_{750}. The false positives were reduced by 51.14% in mean.

4 Conclusion

Compared to the setting τ_1 used in previous work, we reduced the number of false positives significantly by about 31% using τ_{750} without sacrificing overall accuracy.



(a) (c)

(b) (d)

Fig. 2: Two examples of episodes evaluated with τ_1 and τ_{750} (top to bottom). The PC of external motion is displayed dark red at the bottom, the PC of the CM blue, at the top. The dashed green, upper line denotes t_u, the lower, gray t_l. The red dots give the CM error at the time of measurement. The background shows the result of the OMC validation: green (negative), red (true positive), and orange (false positive). The displayed intervals are limited to 3 minutes.

Even with our modifications, the patient of Fig 2a-b) showed a very high number of false positive OMC validations although we were able to reduce them from 17 with τ_1 to 11 with τ_{750}. Previous results of the same session did not show this behavior. We assume this is due to timing issues and it will need to be addressed in the future. Leaving this patient out, the mean FDR of the remaining 5 episodes ranged from 35.49% with τ_1 to 6.25% with τ_{750} with a mean 89% reduction in false positives.

Overall, we were able to reduce the number of false positives distinctly, making OMC now better suited for artifact detection in respiratory motion.

For future improvements we suggest to include the correlation model directly into the OMC. This would allow to immediately check for errors in the correlation model.

References

[ARSS15a] S.-T. Antoni, J. Rinast, S. Schupp, and A. Schlaefer. Comparing model-free motion prediction and on-line model checking for respiratory motion management. In *Gemeinsamer Tagungsband der Workshops der Tagung Software Engineering 2015, Dresden, Germany, 17.-18. März 2015.*, pages 15–18, 2015.

[ARSS15b] S.-T. Antoni, J. Rinast, S. Schupp, and A. Schlaefer. Evaluation des Einflusses von Artefakten auf den Korrelationsfehler in der bewegungskompensierten Radiochirurgie. In *Tagungsband der 14. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie*, pages 133–138, Bremen, Germany, 2015.

[DHV+10] T. Depuydt, O. C. Haas, D. Verellen, S. Erbel, M. De Ridder, and G. Storme. Geometric accuracy evaluation of the new VERO stereotactic body radiation therapy system. In *UKACC International Conference on Control 2010*, pages 1–6. IET, 2010.

[EDSS13] F. Ernst, R. Dürichen, A. Schlaefer, and A. Schweikard. Evaluating and comparing algorithms for respiratory motion prediction. *Phys Med Biol*, 58(11):3911, 2013.

[HNLH08] M. Hoogeman, J. Nuyttens, P. Levendag, and B. Heijmen. Time dependence of intrafraction patient motion assessed by repeat stereoscopic imaging. *Int J Radiat Oncol Biol Phys*, 70(2):609–618, 2008.

[RSG14a] J. Rinast, S. Schupp, and D. Gollmann. A graph-based transformation reduction to reach uppaal states faster. In *FM 2014: Formal Methods*, pages 547–562. Springer, 2014.

[RSG14b] J. Rinast, S. Schupp, and D. Gollmann. State space reconstruction in Uppaal: An algorithm and its proof. *International Journal On Advances in Systems and Measurements*, 7(1 and 2):91–102, 2014.

[SBN+07] Y. Seppenwoolde, R. I. Berbeco, S. Nishioka, H. Shirato, and B. Heijmen. Accuracy of tumor motion compensation algorithm from a robotic respiratory tracking system: a simulation study. *Med Phys*, 34(7):2774–2784, 2007.

[SGB+00] A. Schweikard, G. Glosser, M. Bodduluri, M. Murphy, and J. Adler. Robotic motion compensation for respiratory movement during radiosurgery. *Comput Aided Surg*, 5(4):263–277, 2000.

[SSA04] A. Schweikard, H. Shiomi, and J. Adler. Respiration tracking in radiosurgery. *Med Phys*, 31(10):2738–2741, 2004.