# Intra-operative tracking of aortic valve calcifications using online learning

LH NGUYEN-DUC[1,2], M CASTRO[1,2], V AUFFRET[1,2,3],
H LE BRETON[1,2,3], JP VERHOYE[1,2,3], M GARREAU[1,2], P HAIGRON[1,2]

[1] INSERM, U1099, Rennes, F-35000, France
[2] LTSI, Université de Rennes 1, Rennes, F-35000, France
[3] CHU Hôpital Pontchaillou, Rennes, F-35000, France

Résumé - Nous proposons une nouvelle méthode basée sur le schéma de détection adaptative afin de suivre des calcifications de la valve aortique visibles dans les images fluoroscopiques. Grâce à une stratégie d'apprentissage, nous utilisons la calcification elle-même et son voisinage dans les images pour générer un modèle statistique permettant de localiser la cible. Cette approche a été évaluée sur dix bases de données de patients. L'erreur absolue moyenne obtenue pour chaque évaluation est inférieure à 1,0mm. Ces valeurs d'erreur restent dans la gamme clinique acceptée.

Abstract - We propose a new method based on the adaptive detection scheme in order to track aortic valve calcifications in intra-operative fluoroscopic images. Thanks to a learning strategy, we use the calcification itself and its surrounding area in live images to generate online a statistical model for localizing the target. This approach has been evaluated on ten patient databases, providing an absolute mean tracking error less than 1.0mm. These error values remain within the clinical accepted range.

## 1 Introduction

Transcatheter aortic valve implantation (TAVI) has emerged as a promising alternative to conventional aortic valve replacement performed by open surgery. TAVI consists in introducing a replacement valve through an artery via a small incision. X-ray angiographic and fluoroscopic imaging is routinely used to guide the TAVI as in figure 1. In the positioning and deployment phase of the transcatheter heart valve (THV), live fluoroscopy displays only 2D projection images of endovascular devices (guide-wire, catheter, THV) and some moving anatomical structures such as aortic valve calcifications, which are denser than the soft tissues. Several angiographic injections are often required due to the lack of navigation tools.

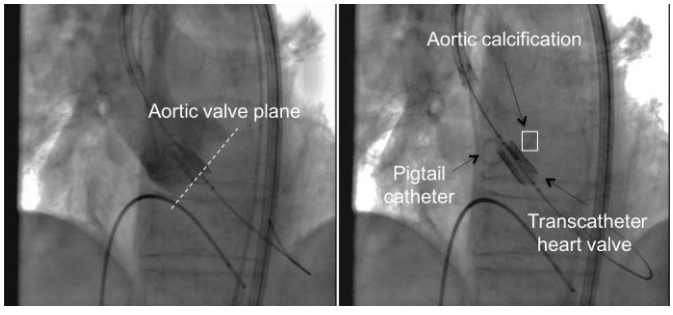


**Figure 1: Transcatheter Aortic Valve Implantation under intra-operative X-Ray image guidance. (LEFT) Angiographic image with contrast medium in the aortic root shape, (RIGHT) Fluoroscopic images during THV positioning.**

Few works on computer assisted TAVI have been reported in the aim to limit the use of X-rays and contrast agent in the course of endovascular intervention. Siemens has prototypically equipped the interventional rotational C-arm with a system for automatic segmentation and overlay of aortic root volume and anatomical landmarks on 2D fluoroscopic images, but without motion correction [1]. Recently, Karar M.E. et al. [2] have proposed a method that integrates a 3D aortic mesh model and landmarks from intraoperative C-arm CT images with tracking of the prosthesis in live fluoroscopic images. This method is based on the tracking of a pigtail catheter, which moves in accordance with the aortic root only when the catheter is locked in one of the native valve's cusp. Therefore, it would be preferable to track the calcifications that are directly linked to the structure of interest, but it arise tracking difficulties related to their sizes and densities. In this paper, we present a novel method for tracking aortic calcifications in order to track the aortic root and the valve plane that represents the moving target for the TVH navigation and deployment during TAVI procedure.

The paper is organized as follows. Section 2 describes the method used for tracking the calcification with the calcification model (Section 2.1), the tracking by detection and the learning process (Section 2.2). Section 3 presents experimental results on real patient data and their quantitative evaluation. Conclusion is given in Section 4.

## 2 Method

The problem of single-target tracking can be expressed as follows: given a sequence of images $f_1, \dots, f_n$, estimate the location $l_k$ of the target for each frame $f_k$. Several approaches for object tracking have been proposed. Recursive tracking methods estimate the current location $l_t$ of an object by applying a transformation on the previous location $l_{t-1}$ [3]. This

approach is sensitive to error accumulation. It is likely to fail and never recover if the object moves out of the research region. Tracking-by-detection methods estimate the object location solely by measurements taken in the current image [4]. This principle remedies the effect of error accumulation. However, unexpected changes of the object appearance cannot be tracked. This drawback is addressed by adaptive methods [5], which update the model during tracking. However, each incorrect update brings error to the model that accumulates over time and causes drift.

These methods rely essentially on the intrinsic features of the object to be tracked (edges, corners, or texture). They cannot be directly applied to track aortic calcifications due to lack of features, frame cuts, occlusions, brightness variations and noise in fluoroscopic images. In the following, we propose a new method based on an adaptive detection scheme (Figure 2) for intra-operatively tracking an aortic valve calcification.

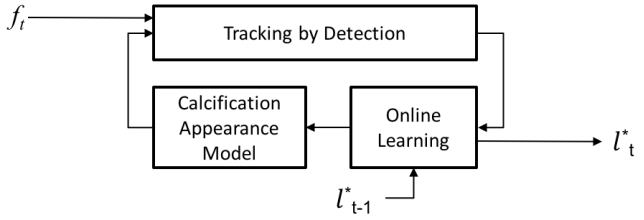


**Figure 2: Overall scheme of the calcification tracking by detection process with the online learning of the calcification model.**

Instead of looking for detecting intrinsic features, we employ the calcification itself and its surrounding area in live images to generate online the statistical model for localizing the target calcification. This model is composed of two sets of image patches representing two different classes (target calcification, other structures). Support Vector Machine (SVM) classifier [6] is training online by considering these two sets of patches. The detection mechanism based on SVM uses an exhaustive search in each frame to determine the target location. The learning process evaluates the current detection errors and updates the model online to improve the tracking performance for the future frames.

### 2.1 Calcification Appearance Model

We employ an image patch to represent a single instance of an object in the image. A patch $x$ is defined as a sub-image of size $(m \times n)$ located at the position $l(x)$ in the image.

The calcification appearance model $M$ is a data set that represents different instances of the target calcification and other structures. This data set $M = \{(X^+, +1), (X^-, -1)\}$ is composed of positive patches $X^+ = \{x_1^+, \dots, x_m^+\}$ considered as the target calcification, and of negative patches $X^- = \{x_1^-, \dots, x_n^-\}$ considered as other structures found in the background, and $\{-1, +1\}$ is a binary label.

At the beginning of the tracking process, the user initializes the first positively labeled patch $x_1^+$ of the model $M$ by manually defining a bounding box of size $(m \times n)$ around the target calcification in the frame $f_0$. Negatively labeled patches are randomly collected from the set of patches $X_0 = \{x | \alpha < \|l(x) - l(x_1^+)\| < r\}$, where r is the search radius and $\alpha$ is the distance from the calcification initial location $l(x_1^+)$. The size $(m \times n)$ of patches remains fixed during tracking.

### 2.2 Tracking by Detection and Online Learning

The tracking by detection process (Figure 3) consists in finding new patches that correspond to the initial calcification patch for every new frame in the sequence. The detection mechanism is separated into four classification stages, each of which enabling early rejection of non-calcification patches. The number of evaluations is thus reduced on average for each frame. At time instant *t*, the tracker maintains the target location $l_t^*$ according to the accepted patch. Each new frame is firstly smoothed by using the Gaussian filter with a kernel standard deviation of 3 pixels to reduce image noise. We generate afterward a set of patches $X = \{x | \|l(x) - l_{t-1}^*\| < r\}$ within the search radius *r* of the last tracked location $l_{t-1}^*$.

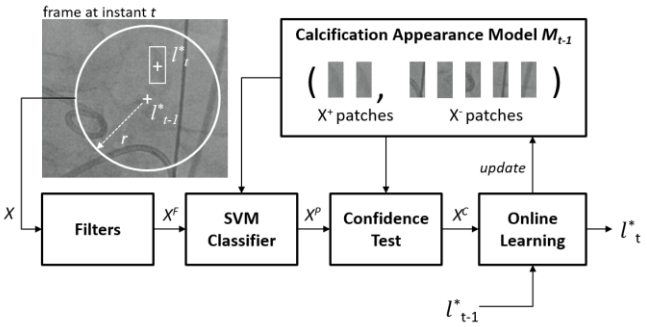


**Figure 3: Detection mechanism and Online learning process.**

**Filters.** In order to eliminate non-relevant patches, we use the variance filter and pixel pair comparisons. The variance filter removes patches, for which gray-value variance is smaller than 50% of variance of the initialized calcification patch. Next, pixel pair comparisons are applied to the remaining patches. Based on the fact that pixels outside the calcification area are brighter than pixels inside the calcification area, each comparison returns 0 or 1 and these measurements are combined to estimate the probability $p(y = +1|x)$ where $y = +1$ indicates the presence of the target calcification in that patch. We crop out the set of image patches and obtain the new sub set $X^F = \{x \in X | p(y = +1|x) \geq 0.5\}$.

**SVM Classifier.** For each patch $x$ of size $(m \times n)$, we concatenate all the rows of $x$ into a vector $v_x$ (of dimension $N = m \times n$). The components of vector $v_x$ correspond to the pixels values of image patch $x$. To express the probability of a patch $x \in X^F$ being a positive calcification candidate $(y = +1)$, we use the discriminative SVM classifier applied to the

corresponding vector $v_x$. SVM is formally defined by a separating hyper-plane and has two main phases. Given already labeled data, the result of the training phase is a discriminant function $h$ which is used in the testing phase to categorize a vector $v_x$, where $x \in X^F$ and $X^F$ is the set of candidate patches accepted by the *Filters* stage. More precisely, the classifier predicts the label of a candidate patch $x$ according to $\hat{y} = sign(h(v_x))$, where $h$ is an optimal hyper-plane.

In the training phase, the input is a set of already labeled vectors $v_{x_i}$ transformed from the calcification model $M_{t-1}$:

$$\left\{\left(v_{x_1^+}, +1\right), \ldots, \left(v_{x_m^+}, +1\right), \left(v_{x_1^-}, -1\right), \left(v_{x_n^-}, -1\right)\right\}$$

The discriminant function $h$ can be found by solving the classical constrained optimization problem [6]. Considering the general case where the training data cannot be separated without error, we use the basic linear kernel $K\left(v_x, v_{x_j}\right) = {v_x}^T v_{x_j}$, where $v_{x_j}$ are the support vectors, combined with a penalty factor $C$. We observed that the value of C was not critical, with similar performances in the range $[100, \ldots, 1000]$. It was set to $C = 100$.

In the testing phase, a new image patch $x$ is labeled as positive calcification candidate ($y = +1$) if $h(v_x) \geq 0$ and as negative calcification candidate ($y = -1$) if $h(v_x) < 0$. These positive candidate patches $X^P = \{x \in X^F | sign(h(v_x) \geq 0\}$ are afterward validated according to a confidence test.

**Confidence Test.** Given two patches $x_1$ and $x_2$, we define a similarity between two patches:

$$s(x_1, x_2) = \frac{1}{2}(ncc(x_1, x_2) + 1) \quad (1)$$

where *ncc* is the Normalized Correlation Coefficient. Formula 1 yields value between 0 and 1, with value close to 1 when the two patches are similar. Given an arbitrary patch $x$ and the object model $M$, we define the confidence value whether the patch belongs to the positive set $X^+$, such as:

$$\theta(x, M) = \frac{s^+}{s^+ + s^-}$$
$$s^+ = \min_{x_i \in X^+} s(x, x_i) \quad (2)$$
$$s^- = \min_{x_j \in X^-} s(x, x_j)$$

We constitute a set of patches $X^C = \{x \in X^P | \theta(x, M_{t-1}) \geq \theta^-\}$, where $\theta^-$ denotes a threshold of confidence. We continue with an online learning strategy by taking into account the tracked location from the previous frame $l_{t-1}^*$.

**Online Learning.** In between the patches $X^C$, Online Learning consists in finding the final patch that contains the target calcification, and also, in finding patches that could be misclassified by both SVM classifier and Confidence Test. If classification errors are detected, the calcification appearance model $M$ is updated to avoid these errors in the future.

To do this, we assume that the displacement of the target calcification between two consecutive frames is limited within a radius $d$. At instant *t*, the patches that are outside of the radius $d$ from the previous tracked location $l_{t-1}^*$ should be considered as the other structures in the background. These patches are added to the negative class $X^-$ to correct the model $M$. In contrast, the patches that are within the radius $d$ from the previous tracked location should be classified as positive calcification candidates. Between these positive candidates, the patch that has the highest confidence value, is considered as the target calcification and is added to the positive class $X^+$ to enhance the calcification model M. The Online Learning process is described according to Algorithm 1.

**Algorithm 1** Online Learning

**Input:** Dataset $X^C = \{x \in X^P | \theta(x, M_{t-1}) \geq \theta^-\}$

1: Divide $X^C$ into two subsets according to $d$
$X^{C1} = \{x | \|l(x) - l_{t-1}^*\| \leq d\}$
$X^{C2} = \{x | \|l(x) - l_{t-1}^*\| > d\}$
2: From $X^{C1}$, find the patch that has the highest confidence value: $x^* = \underset{x \in X^{C1}}{\operatorname{argmax}}\, \theta(x, M_{t-1})$
3: Update the calcification appearance model $M$ with the positive patch, $x^* \rightarrow X^+$, and with the misclassified patches, $X^{C2} \rightarrow X^-$
4: Update the tracker location: $l_t^* = l(x^*)$

In the case where $X^C$ contains only one patch, we consider that the model provides a good result. The calcification model $M$ remains unchanged and the location related to this patch is assigned to the tracker location.

## 3 Experimental Results

Our approach has been tested and evaluated on 10 fluoroscopic sequences that have been acquired on real patients during intra-operative TAVI procedures. In order to limit the radiation to the patient and physicians, fluoroscopic acquisition for the THV deployment phase lasts for only few seconds while fixing the C-arm pose (Artis zeego, Siemens). Each fluoroscopic sequence used for the evaluation contains typically 100 frames representing the fluoroscopic acquisition during the deployment phase. The tested images are of 512x512 to 1024x1024 pixels, with a frame rate of 15 images/s and the pixel size was approximately 0.20mm. We employed the tracking error and the detection rate for assessing performance of the tracking process. For each image of the sequence, the automatically determined location and the manual reference were used to compute the Euclidian distance error. The absolute mean error ± standard deviation was computed over 100 images of the sequence. The detection rate is defined as the ratio between the number of successfully tracked frames and the total number of frames in the tested sequence. All experiments were conducted on an Intel Core i7 processor running at 2.3 GHz. All algorithm parameters were the same for all the experiments, with $r$ = 20 pixels, $\alpha$ = 10 pixels, $d$ = 7 pixels and $\theta^-$ = 0.6, in order to demonstrate the robustness of the proposed method.

Figure 4 depicts an example of the tracking of the aortic calcification. In all of the presented images, the

white bounding box denotes tracked results, the white number denotes the frame number in the sequence. We can observe that the tracker was not affected by the change of illumination and image quality. The calcification was well localized despite the change of image quality at frame 40. At frame 85, the calcification of interest was well tracked in the presence of a similar neighbor calcification. The proposed method outputted good results until the end of the sequence. We obtained the detection rate of 99% and the tracking errors of 0.37±0.16mm computed on 100 frames. The proposed approach tracked precisely the calcification of interest as illustrated in figure 5 with the curve obtained by automatic tracking (red) and the curve obtained by manual tracking (black).

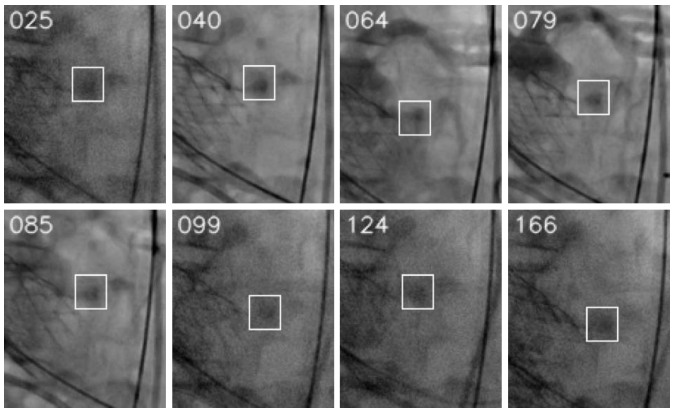


**Figure 4: Qualitative result for the sequence TAVI 02. The white box shows the location of the tracked calcification in the image.**

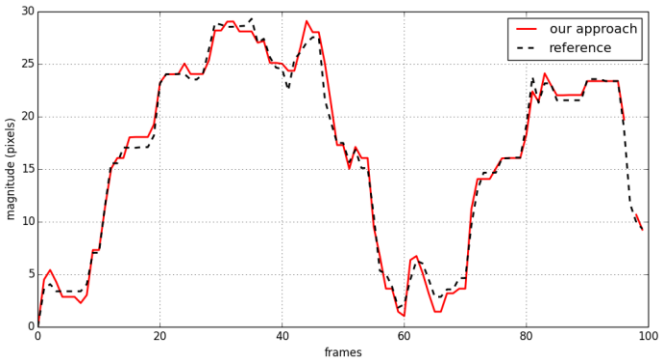


**Figure 5: Displacement magnitude of the tracked calcification in the sequence TAVI 02.**

For all tested sequences, without the proposed Online Learning strategy, the SVM provided an average detection rate of 62%, and the tracking error was less than 1.0mm for only 48.8% of the tracked frames. With Online Learning, the proposed method provided an average detection rate greater than 88%, and the tracking error was less than 1.0mm for 93.3% of the tracked frames. The average processing time was less than 30ms for each frame. In figure 6, mean, standard deviation and maximum tracking errors values that have been computed on the calcification locations are represented for ten real patient sequences. With influences of contrast medium and the guide-wire, sequences 01 and 10 present the highest maximum tracking errors, which are 2.2mm and 1.9mm, respectively. However, all tested fluoroscopic images showed that the mean errors of the calcification tracking were less than 1.0 mm. These error values remain within the clinical accepted range.

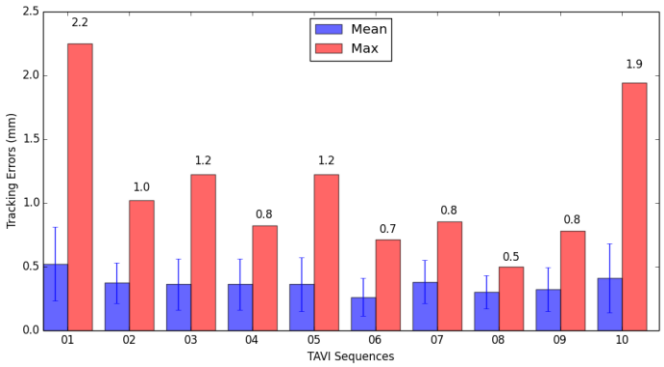


**Figure 6: Tracking errors computed from ten fluoroscopic sequences.**

## 4 Conclusion

We proposed a novel approach to track aortic valve calcifications in fluoroscopic images. In order to limit the use of contrast agent, our method could be used to track the aortic valve plane that represents the moving target for THV deployment. A minimal user-interaction is required to initialize the tracking process. The proposed approach has been tested on 10 patient sequences and a quantitative evaluation has been performed. The tracking method has been developed to track calcifications but it can be applied also on other structures of interest in fluoroscopic imaging. Additional cases are required to confirm the performance of the tracking approach.

### Acknowledgment

This work was partially supported by the ITEA2 MEDIATE project (ITEA 09039), the French ANR within the Investissements d'Avenir program (Labex CAMI) under reference ANR-11-LABX-0004. This work has been partially conducted in the experimental platform TherA-Image (Rennes, France) supported by Europe FEDER.

### References

[1] John, M., et al. "System to guide transcatheter aortic valve implantations based on interventional C-arm CT imaging." Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010. Springer Berlin Heidelberg, 2010. 375-382.

[2] Karar, M.E., et al. "Model-updated image-guided minimally invasive off-pump transcatheter aortic valve implantation." Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011. Springer Berlin Heidelberg, 2011. 275-282.

[3] Lucas, B.D., Kanade T. "An iterative image registration technique with an application to stereo vision." IJCAI. Vol. 81. 1981.

[4] Ozuysal, M., Fua, P., and Lepetit, V. "Fast keypoint recognition in ten lines of code." Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on. Ieee, 2007.

[5] Avidan, S. "Support vector tracking." Pattern Analysis and Machine Intelligence, IEEE Transactions on 26.8 (2004): 1064-1072.

[6] Cortes, C., and Vapnik, V. "Support-vector networks." Machine learning 20.3 (1995): 273-297.