Improving Flutter Localization Performance by Optimizing the Inverse Dower Transform

Muhammad Haziq Kamarul Azman[1 2], Olivier Meste[1], Decebal G. Latcu[3], Kushsairy Kadir[2]

[1] Université Côte d'Azur, CNRS, I3S, France
[2] Universiti Kuala Lumpur, Malaysia
[3] Centre Hospitalier Princesse Grace, Monaco

Abstract

A previous study showed the possibility to localize right or left flutter circuit origin using variability contained in vectorcardiographic loop parameters. The Inverse Dower Transform, used to obtain the vectorcardiograms is based on a very simplistic torso conductor model, and hence not optimized. The present study aims to optimize the transform to maximize classifier accuracy. A parametric optimization model was proposed, as well as an optimization scheme. Model parameters were obtained by iteratively optimizing the linear SVM classifier accuracy until convergence. The goal can be shown to be multimodal and non-smooth. Therefore, a multi-instance and derivative-free method was considered. Previous dataset of 56 flutter recordings (31 right, 25 left) was used, considering only non-overlapped and respiratory motion-corrected F loops. For the SVM classifier, a 3.8% *increase in accuracy was observed (max* 0.95*). When the logistic regression classifier was used, an increase of* 7.8% *was observed (max* 0.98*). Comparison to a targeted transform previously developed showed an improvement by* $17-19\%$*. Observation of the model parameter values showed amplitude reduction applied to Lead X and rotation applied to Lead Z.*

1. Introduction

Atrial flutter (AFL) is an arrhythmia involving a rapid, rotating circular depolarization of the atrium, whose circuit may be located in either right or left atrium. Localization of the circuit is an important objective in radiofrequency catheter ablation therapy, and it may condition the difficulty and efficacity of the procedure. The rotating circular depolarization generates a pseudo-periodic sawtooth waveform on the electrocardiogram (ECG); a full wave representing one complete cycle is known as an F wave. Our previous study showed that by using serial beat-to-beat approach, variability features extracted from the vectorcardiographic (VCG) F loops obtained from the 12-lead ECG using the Inverse Dower Transform (IDT) allowed non-invasive localization of AFL circuits [1].

The VCG is an orthogonal 3-lead system [2], obtained by combination of several signals from a 7-unipolar electrode system. The combination weights were found from geometric interpolations of an image surface spanned by the 7 lead vectors, derived from a homogeneous volume conductor model of the torso (tank filled with saline solution). The Inverse Dower Transform [3] (IDT) utilizes the same information from this image surface. Since the image surface does not account for heterogeneity due to organs and anatomical structures, the IDT is not optimized for a real torso. Furthermore, electrode misplacement is not accounted for.

Thankfully, these can be formulated in a mathematical context, allowing for the use of estimation-optimization techniques for finding the optimized transform coefficients. In addition, the optimization goal can be based on e.g. localization accuracy, which is a more direct endpoint than e.g. similarity to real VCG. However, it is expected that the goal has non-ideal properties, since classifier accuracy depends greatly on the distribution of the data given a set of considered features and classes, as well as the optimization coefficients, and the relation between these elements are not known.

In this paper, we propose a model of the conductor heterogeneity and electrode misplacement, as well as an optimization scheme. A proof of multimodal non-smooth goal is made and an appropriate estimation-optimization scheme is presented. Results were compared with a different transform previously used with atrial signals [4].

2. Methodology

2.1. Dataset and Preprocessing

The dataset used in this study consists of 56 ECG recordings of AFL, acquired during catheter ablation operations at Centre Hospitalier Princesse Grace, Monaco. All

 ISSN: 2325-887X DOI: 10.22489/CinC.2019.290

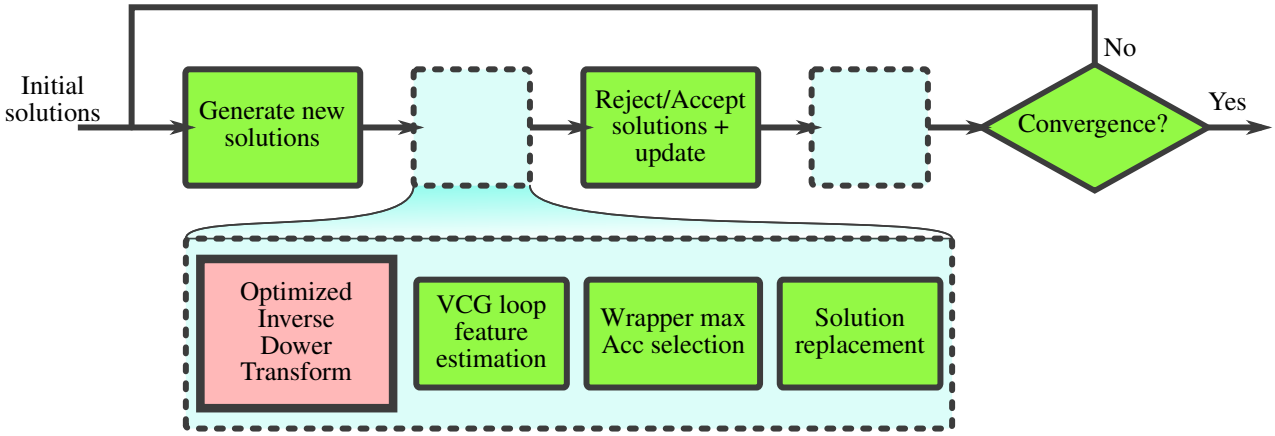

Figure 1. Diagram of the estimation-optimization scheme. Solutions here refer to the set of 6 coefficients of the model parameters (see (2)). The bold-lined red box indicates the novel approach of optimizing the IDT transform matrix.

signals were recorded at $f_s = 2000$Hz. A finite-impulse response notch filter at 50Hz was applied to these signals to remove powerline interference, and were then filtered at $[0.5; 70]$Hz using type II Chebyshev high- and low-pass filters to remove baseline wander and high-frequency noises. Records with missing leads, low F wave amplitudes, low atrioventricular block ratio ($< 2:1$) were excluded from the study. In total, there were 31 recordings with right circuit localization and 25 with left localization.

F waves from each recording were detected, segmented, synchronized, transformed into VCG and removed of respiratory motion using a technique described previously [1]. Only F waves not overlapped within T waves were considered. For this study, only the statistics $\mathrm{Mean}(\cdot)$, $\mathrm{Var}(\cdot)$, $\mathrm{Skewness}(\cdot)$ and $\mathrm{Kurtosis}(\cdot)$ were considered for the 4 loop parameters (total of 16 features, see [1] for details).

2.2. Error Model

The IDT describes a linear combination of a reduced set of leads from the 12-lead ECG to the VCG $\mathbf{Y} = \mathbf{TX}$ where $\mathbf{T}$ is the 3×8 IDT matrix described in [3], $\mathbf{X}$ and $\mathbf{Y}$ are the $8 \times N$ and $3 \times N$ 8-lead ECG and VCG respectively. The 8 leads are arranged as $[\mathrm{V}_1\ \mathrm{V}_2\ \mathrm{V}_3\ \mathrm{V}_4\ \mathrm{V}_5\ \mathrm{V}_6\ \mathrm{I}\ \mathrm{II}]$.

Ideally, the thorax is required to be a homogeneous conductor, and the ECG electrodes should be placed at the same locations originally indicated on the chest. However, this is an impossible feat as physical human anatomy has heterogeneous conductivity and variations from patient to patient, and electrode placement depends on operator proficiency and chest shape. Both these effects can be translated as a gain or attenuation as well as a rotation of each lead vector. Thus, $\mathbf{Y} \rightarrow \tilde{\mathbf{Y}} = \mathbf{AQY}$, where $\mathbf{A} = \mathrm{diag}(a_X,\ a_Y,\ a_Z)$ is a diagonal matrix of gain or attenuation, and $\mathbf{Q} \in \mathbb{R}^{3\times3}$ a rotation matrix, satisfying $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$. We thus have:

$$\mathbf{AQY} = \mathbf{TX} \tag{1}$$

To optimize the IDT, the errors can be compensated:

$$\mathbf{Y} = \mathbf{A}^{-1}\mathbf{Q}^{\mathrm{T}}\mathbf{TX} \tag{2}$$

In practice, $\mathrm{diag}(b_X,\ b_Y,\ b_Z) = \mathbf{B} = \mathbf{A}^{-1}$ and $\mathbf{R} = \mathbf{Q}^{\mathrm{T}}$ were estimated instead. $\mathbf{R}$ can be decomposed into individual 3D rotation matrices $\mathbf{R}_X$, $\mathbf{R}_Y$, $\mathbf{R}_Z$ for each of the 3 axis of the VCG leads, with each matrix computable from the angles ϕ_X, ϕ_Y, ϕ_Z respectively, and $\mathbf{R} = \mathbf{R}_X\mathbf{R}_Y\mathbf{R}_Z$. The 6 parameters capture the essential errors observed on the VCG and avoid models with large number of parameters that would be more difficult to estimate.

2.3. Outline of Optimization Procedure

The goal of parameter estimation: maximize the classifier accuracy, is different than most conventional approaches. This requires us to define it in the scope of optimization, to be able to decide which method is suitable for use. Furthermore, conventional techniques do not account for the nature of data distributions given a certain feature set: a subject particularly associated with machine learning. The scheme presented here constitutes an original approach that combines elements of optimization and machine learning.

2.3.1. Properties of the Goal

The optimization goal is defined as:

$$\max \mathrm{Acc} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FN} + \mathrm{FP} + \mathrm{TN}}$$

Note that the denominator is constant for a given set of M data points contained in $\mathbf{X}$, with a set of features $\mathcal{F}$. The numerator is essentially the number of elements in the union $\{\hat{G} = \text{Right}|G_{\text{true}} = \text{Right}\} \cup \{\hat{G} = \text{Left}|G_{\text{true}} = \text{Left}\} = \mathcal{G}$, with G being a label. The union contains data points whose predicted labels $\hat{G}$ are similar to the true labels G_{true}, for all available classes. Suppose $\mathbf{g}$ a binary M-vector (entry values equal 0 or 1) with elements representing membership or not of each data point in $\mathcal{G}$. Each entry g_m of $\mathbf{g}$ is determined by the classifier through some evaluation of the conditional probability $\mathbb{P}\{G_m|\mathbf{x}_m, \mathbf{B}, \mathbf{R}, \mathcal{E}\}$ that includes not only the optimization parameters, but also the subset of features $\mathcal{E} \subseteq \mathcal{F}$ that give maximum accuracy.

The final form amounts to:

$$\max_{\mathbf{B},\,\mathbf{R},\,\mathcal{E}} \text{Acc} = \frac{1}{\text{M}} \|\mathbf{g}(\mathbf{X}, \mathbf{B}, \mathbf{R}, \mathcal{E})\|_0 \tag{3}$$

It can be shown that this goal formulation has a non-convex and non-smooth form, by noticing that on convergence, the three optimal rotation angles $\hat{\phi}^*_X,\ \hat{\phi}^*_Y,\ \hat{\phi}^*_Z$ can be $\pm 180^\circ$ ambiguous, and that the L_0 norm is not smooth. In practice, class overlap also affects goal convexity and smoothness.

2.3.2. Optimization Scheme

Due to the above-mentioned goal formulation and constraints, a multi-instance derivative-free optimization algorithm should be used, to account for goal non-convexity and non-smoothness. In this study, we employed cuckoo search (CS) as the algorithm of choice [5]. The algorithm requires a small number of tuning parameters and is relatively simple in implementation.

n initial candidate solutions are instantiated randomly across the solution space. For each instance, the goal is first evaluated. Then, an update is performed, moving each solution to another candidate solution via a Lévy random walk. These new candidate solutions are evaluated, and they replace the initial solutions if they have higher goal values.

Next, a random decision is made on whether to reject or accept the solutions. A probability p_a determines the chance of being rejected. Finally, solutions are updated again via a biased random walk. Only accepted solutions are updated. The algorithm iterates again with the final solutions of iteration $k-1$ used as initial candidates of iteration k, until it satisfies a convergence criteria. At the end of each iteration, the best value of the goal indicates the best solution. Figure 1 summarizes the process.

15 instances we used to simultaneously search the solution space. The probability p_a is set to 0.25, as was suggested by the authors of CS. No attempt was made to find the best value of p_a or n. Convergence was set to occur when an accuracy of 1 is achieved or the best goal does not change for 10 iterations. A boundary is imposed after each generation or update of solution to avoid non-sensical solutions. The scale parameters b were limited to values in the range $[0.001;\ 5]$ and rotation parameters ϕ to values in the range $[-179.999;\ 180]^\circ$.

For each evaluation, a set of candidate $\hat{b}_X,\ \hat{b}_Y,\ \hat{b}_Z$ and $\hat{\phi}_X,\ \hat{\phi}_Y,\ \hat{\phi}_Z$ were obtained. For each F wave of each recording, the optimization detailed in (2) is applied, loop features were calculated and an exhaustive wrapper evaluation of feature combination up to a length of 7 features was performed. For each feature combination, the accuracy was calculated. The maximum accuracy was taken as the current solution's goal value. The linear support vector machine (SVM) classifier is due to simplicity and low training time. Blue boxes with dashed lines in Figure 1 summarizes this process.

3. Results and Discussion

3.1. Pre- versus post-optimization comparison of classifier performance

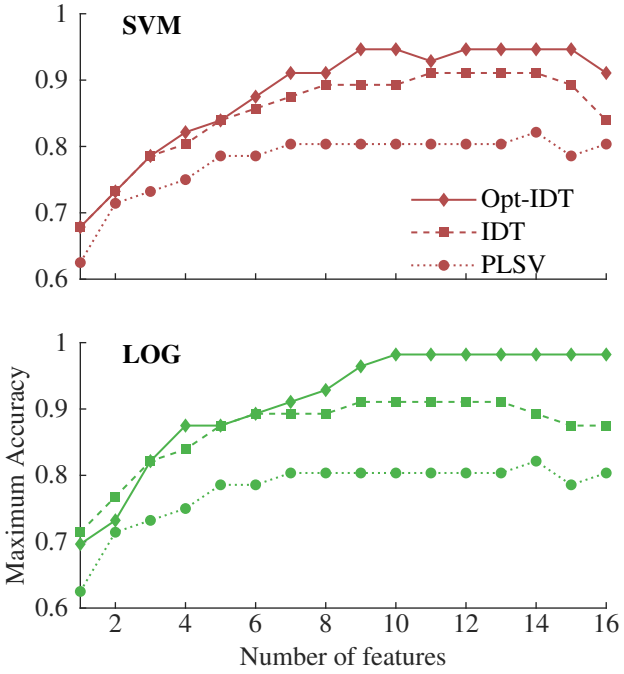

Figure 2. Maximum accuracy of the SVM (top) and LOG (bottom) classifiers pre- and post-optimization of the IDT (IDT and Opt-IDT). Results of the PLSV transform is also shown for comparison.

The maximum SVM classification accuracy is shown in Figure 2 (top diagram). An improvement can be observed, starting as early as 4 combination of features.

The maximum improves from 0.91 to 0.95 ($(\mathrm{Se}, \mathrm{Sp}) = (0.97, 0.92)$). This is an improvement of 3.8% from the previous result ($\mathrm{Acc} = 0.91$, $(\mathrm{Se}, \mathrm{Sp}) = (0.94, 0.88)$). The result illustrates the validity of the approach.

The same optimized dataset was evaluated with the logistic regression (LOG) classifier. The bottom diagram of Figure 2 shows the maximum accuracy of this classifier when operating on the optimized dataset. An even bigger improvement can be seen (7.8% accuracy increase, optimized $\mathrm{Acc} = 0.98$ and $(\mathrm{Se}, \mathrm{Sp}) = (0.96, 1.00)$), suggesting that parameters can be transferred to other classifiers, and that it may also result in better improvements.

3.2. Performance of alternative transform

Alternatives to the IDT are available [4]. These transforms target atrial activity specifically: it is suggestive that they may produce better results than ours. To investigate this matter, classifier performance results using the PLSV transform were also obtained. The processing scheme is similar to the one detailed in [1], except that the IDT is replaced by the PLSV transform.

The result for both SVM and LOG classifiers are shown in Figure 2. As can be seen, performance is low and does not even match the unoptimized IDT. Compared to the optimized IDT, there is a difference of $17-19\%$ in maximum performance. This shows that the IDT is still useful in the context of localization using beat-to-beat approach, and that its optimization results in increased performance as opposed to using alternative transforms.

3.3. Transform coefficients

The optimal parameter values represent the optimal scaling and rotation applied to each VCG lead. After normalization of the scale values by the largest $\hat{b}$, and addition of $\pm 180^\circ$ to the rotation values, the strongest effects were observed on lead X (right-to-left component) and lead Z (front-to-back component; $\hat{b}^*_X = 0.72$ and $\hat{\phi}^*_Z = -12.52^\circ$). The remaining parameters did not present strong effects (scale values close to 1, rotation values close to 0). The two optimization is regarded as sufficient to increase separation between right and left AFL variability. Furthermore, feature selection returned similar relevant features as previously found in addition to other previously irrelevant features, further reinforcing this fact.

3.4. Limitations

The model parameters were obtained using a very small sample size and is not representative of the actual population of AFL circuits. It is expected to be unoptimized on datasets different than the one employed here. Furthermore, no cross-validation was attempted.

The optimization procedure required a very long runtime (total of ~ 90 hours using an 4-core Intel i5 running at 3.3 GHz with parallel processing) due to the use of the exhaustive wrapper evaluation scheme. Shorter runtimes can be achieved using more cores.

4. Conclusion

An original approach of optimizing the Inverse Dower Transform for better localization of right or left flutter circuit is presented. An error compensation model was proposed. Estimation of the model parameters was made considering a more direct endpoint (maximizing localization accuracy). Optimization goals and properties were discussed and a procedure was proposed. This resulted in an increase in classification accuracy (up to 7.8%), also shown to be transferable to other classifier models. Performance using an alternative transform showed inferior performance of $17 - 19\%$ less compared to the optimization proposed here.

References

[1] Kamarul Azman MH, Meste O, Kadir K. Localizing atrial flutter circuit using variability in the vectorcardiographic loop parameters. In Computing in Cardiology, volume 45. September 2018; 4.

[2] Frank E. An accurate, clinically practical system for spatial vectorcardiography. Circulation 1956;13(5):737–749.

[3] Edenbrandt L, Pahlm O. Vectorcardiogram synthesized from a 12-lead ECG: superiority of the inverse Dower matrix. Journal of Electrocardiology 1988;21(4):361 – 367.

[4] Guillem MS, Sahakian A, Swiryn S. Derivation of orthogonal leads from the 12-lead ECG. accuracy of a single transform for the derivation of atrial and ventricular waves. In Computing in Cardiology, volume 33. September 2006; 249–252.

[5] Yang XS, Suash Deb. Cuckoo search via lévy flights. In 2009 World Congress on Nature & Biologically Inspired Computing (NaBIC). IEEE, 2009; 210–214.

Address for correspondence:

Muhammad Haziq Kamarul Azman
Room 416
Universiti Kuala Lumpur, British Malaysian Institute
Batu 8, Jalan Sungai Pusu
53100 Selangor, Malaysia
mhaziq@unikl.edu.my