A Validation Protocol for Assessing Cardiac Phase Retrieval in Intravascular Ultrasound

Aura Hernàndez-Sabaté[1], Monica Mitiko Soares Matsumoto[2], Sergio Shiguemi Furuie[2], Debora Gil[1]

[1]Computer Science Dept, UAB, and Computer Vision Center, Bellaterra, Barcelona, Spain
[2]Faculdade de Medicina and Escola Politécnica da Universidade de São Paulo, Sao Paulo, Brazil

Abstract

A good reliable approach to cardiac triggering is of utmost importance in obtaining accurate quantitative results of atherosclerotic plaque burden from the analysis of IntraVascular UltraSound sequences. Although, in the last years, there has been an increase in research of methods for retrospective gating, there is no general consensus in a validation protocol. In this paper, we propose an objective validation protocol based on the variability of the retrieved cardiac phase and explore the capability of several quality measures for quantifying such variability. We can notice that the residual variance of the regression correlation line is robust against fraction and variabilities as far as one can establish a pair-wise correspondence between candidate and reference.

1. Introduction

An objective evaluation of any technique is a crucial step to ensure the good behavior of algorithms. Checking the accuracy of any method allows the evaluation of their performance, bringing up their strengths and limitations. In this paper we concern for the particular case of retrospective image-based cardiac phase gating methods processed from standard non-gated sequences.

The importance of cardiac phase gating methods falls on quantitatively assessing atherosclerotic plaque burden and accurately predicting plaque rupture. Still, there is no general consensus for defining a suitable "goodness" score providing a reliable measure of the quality of the algorithm we are evaluating. The first methods developed for retrospective ECG-gating [1, 2] reported comparisons between volumetric measures obtained from off-line sampling of sequences and on-line ECG-triggered acquisitions. However, such validation protocol requires ECG-gating systems for prospective image capture, which are not always available and increase acquisition time up to three times [1]. In the absence of ECG-gated acquisitions, a standard well-defined methodology for comparing gating algorithms (image-based or not) is not available. Many methods base on quality assessment of longitudinal cuts appearance [3–6]. However, this inspection is subjective and susceptible of changes depending on the angle chosen for inspection. As well, it does not allow comparison among methods. Quantitative measures are objective but there is no standard validation protocol for assessing the accuracy and robustness of cardiac phase gating techniques [7, 8]. Such heterogeneity in validation protocols makes faithful comparison across methods a difficult task.

Since retrospective ECG-samplings can be delayed from the gold standard by a constant shift and still successfully retrieve cardiac phase, an objective quality measure should only measure the variability in the sampling. In this paper we propose a validation protocol based on the variability of the retrieved cardiac phase. We explore the capability of several quality measures for quantifying such variability. The remaining of the paper is structured as follows: In section 2 we detail the validation protocol we propose. In section 3 we explore the performance of different measures. Finally, conclusions are exposed in section 4.

2. Quality measures

An ideal detector, suitable for its application in clinical practice, should produce stable phases. That is, it should always sample the same cardiac cycle fraction. In this context, one should measure the variability (variance) of a candidate sampling with respect a reference (or gold standard) one, which corresponds to the ground truth. Thus, the variance would indicate how spread we are aiming a target. In order to quantify the deviation between the sampling and the ground truth, we have considered two quality scores reported in the literature: signed distance to the closest reference sample [7] and relative distance to the right of each reference sample [8]. We have also determined residuals of the linear regression correlation of reference against candidate samplings.

We define gs^k the gold standard sampling and $Lgs^k = gs^{k+1} - gs^k$ the gold standard cycle, which corresponds to the length of the interval of each pair of

gold standard positions.

2.1. Signed nearest neighbor

Let c^k be the detected sample positions and consider gs^k the closest gold standard sample position for each c^k. For each pair of sample positions, we define the following distance:

$$EN^k = c^k - gs^k$$

The distances for all frames provide a distance map for each sequence. The interval of standard deviation over all sampled frames detects delayed or forwarded samplings. Thus, we define the Signed Nearest Neighbor as the standard deviation (std) of the former distances:

$$\textbf{Signed Nearest Neighbor} = std(EN^k)$$

Notice that this interval range is $[-\frac{max(Lgs^k)}{2}, \frac{max(Lgs^k)}{2}]$.

2.2. Phase fraction

Consider now, for each gs^k, the right closest detected sample ones, namely r^k. For each pair of sample positions, we define the following distance:

$$ER^k = \frac{r^k - gs^k}{Lgs^k} * 100$$

This measure provides, for each k, the fraction of the gold standard k-cycle. In that sense, the interval of standard deviation for $ER^k, \forall k$, also detects delays samplings. Again, we define the Phase Fraction as the standard deviation of ER distances:

$$\textbf{Phase Fraction} = std(ER^k)$$

Notice that in this case, the interval range is $[0, 100]$.

Figure 1 graphically shows these two distances. The plot on the top shows three examples of EN, while on the bottom, the formula for ER is explained.

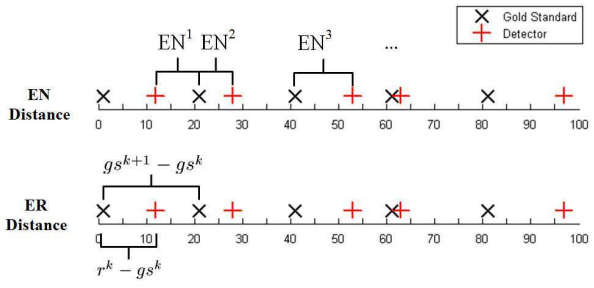


Figure 1. Near and right distances

2.3. Linear regression residuals

The goal of retrospective ECG-gating techniques is to produce detectors sampling at the same cardiac phase. Thus the difference between gold standard samplings and automatic ones should be, at most, a constant shift. Linear regression correlation models this relationship between both signals [9]. If gold standard and detector samplings have the same number of points, we can compute the linear regression correlation of one signal against the other obtaining the following model:

$$Y = aX + \varepsilon$$

where X corresponds to the automatic sampling and Y corresponds to the gold standard. Linear regression estimates the model parameter a and ε corresponds to the adjustment error in the model. Errors $\varepsilon \sim N(0, \sigma^2)$ are assumed to be uncorrelated and distributed with mean 0 and constant (but unknown) variance.

In order to appreciate the differences between both signals and their variation, residuals are useful for detecting failures in model assumptions. Residuals are computed as the differences between the predicted value (extracted from the regression line) and the observed one (extracted from the plot of one signal against the other one). Mathematically, given a set of samples (x_i, y_i) and the regression parameters a_i, the residuals are computed as follows:

$$\varepsilon_i = y_i - x_i * a_i$$

Figure 2 illustrates residuals computation. Black dots correspond to the automatic sampling detections against the gold standard one. In red, we plot the regression line and the residuals computation is illustrated in the detail.

We consider the residual variance since it provides information about the correlation between X and Y:

$$\textbf{Linear Regression Residuals} = std(\varepsilon_i)$$

If they are perfectly related then there is no residual variance.

3. Experiments

This section is devoted to the assessment of the proposed measures as quantitative goodness measures for retrospective ECG-gating.

3.1. Experimental setting

The performance of the different proposed measures has been explored on a set of synthetic samplings covering different cardiac cycle fractions and variabilities. The signal playing the role of reference signal (or gold standard) is

Table 1. Ranges for each quality measure of all the initial shifts and variability from 1 to 5

	Variabilities (1:5)				
Signed Nearest Neighbor	2.20±2.37	3.68±2.27	4.68±1.77	5.24±1.24	5.52±0.78
Phase fraction	2.19±2.38	3.68±2.28	4.61±1.67	5.28±1.25	5.62±0.79
Residuals	1.04±0.01	2.02±0.02	3.03±0.02	4.01±0.03	5.00±0.03

Table 2. Ranges for each quality measure of all the initial shifts and variability from 6 to 10

	Variabilities (6:10)				
Signed Nearest Neighbor	5.66±0.42	5.72±0.21	5.82±0.08	5.74±0.06	5.77±0.03
Phase fraction	5.72±0.43	5.78±0.21	5.72±0.09	5.79±0.06	5.77±0.03
Residuals	6.02±0.03	7.01±0.04	7.98±0.08	9.02±0.07	9.99±0.08

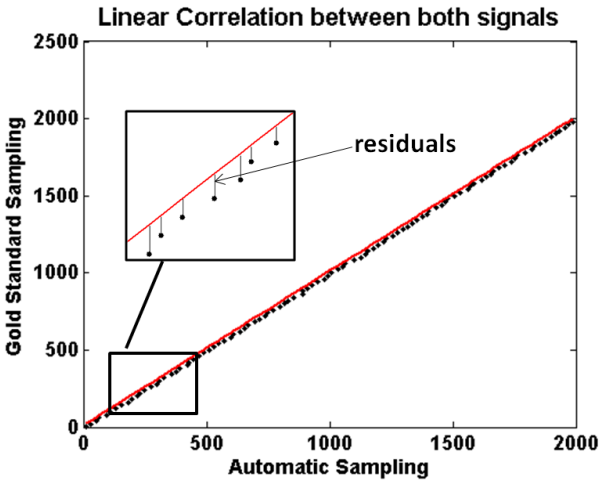


Figure 2. Linear Regression Correlation and residuals illustration

sampled every 20 frames and the candidate signals vary their initial shift from 1 to 20 and the variability from 1 to 10.

If $RSig$ is the reference signal, $S = 1, \ldots, 20$ is the initial shift and $V = 1, \ldots, 10$ corresponds to the variability, we compute synthetic candidate signals $(CSig)$ as:

$$CSig(S, V) = RSig + S + rand * V$$

where $rand$ follows a standard uniform distribution on the open interval $(0, 1)$.

Notice that in the synthetic experiments, the variability corresponds to the error:

$$std(CSig - RSig) = std(S + rand * V) = std(rand)$$

For that, we have compared the standard deviation of the synthetic data to each quality measure. A good quality measure should detect the variability independently on the phase we are sampling.

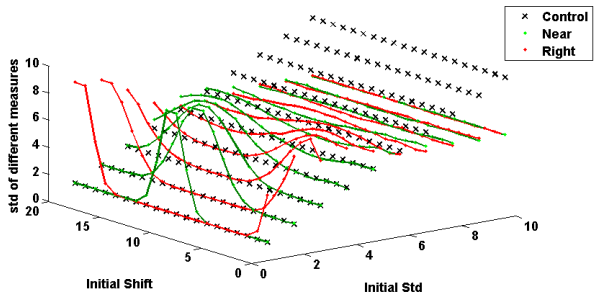
3.2. Results

Figure 3 graphically shows the standard deviation of each proposed measure as the parameters change. Figure 3(a) shows the measures from the literature and figure 3(b) plots the new proposed one. Crosses in black in both plots correspond to the synthetic variability. In fig. 3(a) we present the Signed Nearest Neighbor (Near) quality measure in green and the Phase Fraction (Right) quality measure in red. In fig. 3(b) we show the Linear Regression Residuals (residuals) in magenta. We can observe that Signed Nearest Neighbor and Phase Fraction strongly depend on the initial shift and they are opposed measures. For no shift, Nearest Neighbor properly detects the variation, while worsens for shifted samplings until the middle of the phase. On the contrary, Phase fraction properly detects the variation for shifted samplings while variability affects the detection in no shifted samplings.

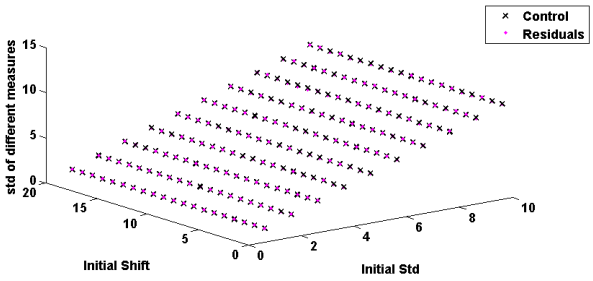
Tables 1 and 2 show the ranges of the above plots. That is, for each variability, ranging from 1 to 5 in table 1 and from 6 to 10 in table 2 we compute the mean and standard deviation among all the initial shifts (from 1 to 20). Notice that the first two quality measures have a high standard deviation for small variabilities and they stabilize with large ones. However, the mean does not correspond to the real variability. On the contrary, the ranges for Linear Regression Residuals cases coincide to the ground truth, that is the mean corresponds to the variability and the standard deviation is very small (up to 0.08).

4. Conclusions

From our simulations, we could conclude that the metrics related to distances are sensitive to the shift considered. Meanwhile, the variance of the residuals are robust against fraction and variabilities as far as one can establish a pair-wise correspondence between candidate and reference. Furthermore, we will investigate these measures re-



(a)



(b)

Figure 3. Standard deviation of different measures against the parameters.

garding how to establish a practical metric and also study false positive and false negative detection effects.

Acknowledgements

This work was supported by the Spanish projects PI071188, TIN2009-13618 and CONSOLIDER-INGENIO 2010 (CSD2007-00018) and by Brazilian research foundation agencies: CNPq and FAPESP 07/53985-3. The last author has been supported by The Ramon y Cajal Program.

References

[1] von Birgelen C, Mario C, Li W, Schuurbiers J, Slager C, de Feyter P, Roelandt J, Serruys P. Morphometric analysis in three-dimensional intracoronary ultrasound: An in vitro and in vivo study performed with a novel system for the contour detection of lumen and plaque. Am Heart Journal 1996; 132:516–527.

[2] Winter S, Hamers R, Degertekin M, K.Tanabe, Lemos P, Serruys P, Roelandt JR, Bruining N. Retrospective image-based gating of intracoronary ultrasound images for improved quantitative analysis: The intelligate method. Catheter Cardiovasc Interv Jan. 2004;61(1):84–94.

[3] Zhu H, Oakeson KD, Friedman MH. Retrieval of cardiac phase from IVUS sequences. In Medical Imaging 2003: Ultrasonic Imaging and Signal Processing, volume 5035. 2003; 135–146.

[4] Nadkarni S, Austin H, et al. A pulsating coronary vessel phantom for two and three-dimensional intravascular ultrasound studies. Ultrasound Med Biol 2003;29 (4):621–628.

[5] Gatta C, Pujol O, oriol Rodriguez Leor, Ferre JM, Radeva P. Robust image-based IVUS pullbacks gating. In Medical Image Computing and Computer-Assisted Intervention MICCAI'2008, 11th International Conference, NY (USA). September 2008; 518–525.

[6] Barajas J, Caballero KL, Rodriguez O, Radeva P. Cardiac phase extraction in IVUS sequences using 1-D gabor filters. In 29th Annual International Conference of the IEEE EMBS. August 2007; 343–346.

[7] Matsumoto MMS, Lemos PA, Yoneyama T, Furuie SS. Cardiac phase detection in intravascular ultrasound images. In Medical Imaging 2008: Ultrasonic Imaging and Signal Processing. February 2008; .

[8] O'Malley SM, Granada J, Carlier S, Naghavi M, Kakadiaris I. Image-based gating of intravascular ultrasound pullback sequences. IEEE Tran Info Tech in BioMed 2008;12(3):299–306.

[9] Seber GAF. Linear regression analysis. John wiley and sons, 1977.

Address for correspondence:

Aura Hernàndez-Sabaté
Computer Vision Center
Edifici O, 08193 Campus UAB, Bellaterra, Barcelona, Spain
aura@cvc.uab.cat