Optimization of contour detection in cardiac MRI using Active Appearance Models: definition of the training set

E. Angelié[1], S. F. Matzken[1], B. P. Lelieveldt[1], R. J. van der Geest[1], J. H. Reiber[1]

[1]Radiology, LUMC, Leiden, Netherlands

Introduction: Active appearance models (AAMs) have been successfully used for image segmentation in cardiac MR (1). Capturing both shape and texture from a training set of manually segmented images the AAM segmentation is highly specific to the task given. However, the cumbersome manual annotation of the images in the training sets makes it difficult to implement the method in clinical routine. Making the AAM approach clinically applicable we addressed two different questions: First what is the minimal the number of images to be included in the training set. Second, the impact of using different distributions of healthy volunteers (N) and patients (P) within the training set was evaluated.

Method: The AAM segmentation procedure consists of two different phases: During a training phase it learns the relationship between model parameter displacements and the residual errors induced between a training image and a synthesized model (2). The matching (detection) phase attempts to find the best fit of the model to the data in a new image. In this work we focused on the training step and the definition of the training set.
We first studied the impact of varying the number of images included in the training set ([10-190] images from the same vendor, randomly chosen between pathological and normal examination) on the accuracy of contour detection. Based on the degree of similarity (3) between the automatically detected and the manually drawn contours, we analyzed the endocardial and epicardial contours in end-diastolic (ED) and end-systolic (ES) phases of the cardiac cycle.
Since the model is defined according to the image characteristics, and the image characteristics varies according to the pathology faced, mixing image characteristics in the training set should make the segmentation more robust. To verify this last hypothesis, we studied the impact of including different images characteristics (P, N, 50%P-50%N, 20%P-80%N, 80%P-20%N) in the training set on the accuracy of the segmentation. We used the ejection fraction (EF) as the criteria to define N and P examinations (EF<40%: P, EF>40%: N). In total, 38 short-axis MR studies (425 images, 50% N and 50% P) were available with manually traced endocardial and epicardial contours in the ED and ES phases. In all experiments AAM training and matching was performed on independent image sets using a leave-n-out procedure.

Results: Figure 1 shows that the degree of similarity increases with the number of images included in the training and 150 images included is sufficient to obtain a good segmentation. Mixing the population (80%N-20%P) to create the model provides the highest degree of similarity. Matching a “model 100% N” on a normal population and a “model 100%P” on a pathological population of images respectively were considered as reference (Figure 2).

Discussion and Conclusion:
-The more images included in the training set, the best the detection is however good matches are achieved when 150 images (80% N-20% P) define the training set.
-The low degree of similarity found using a “model 100%P” can result from the randomly selected pathological examinations; since the shape and texture on the images change accordingly to the pathology, the large variability in image characteristics result in too much freedom for the model to match with a high accuracy on other images.

Reference List

1. Mitchell SC, Lelieveldt BPF, van der Geest RJ., Bosch HG., Reiber JHC, Sonka M., Cardiac segmentation using active appearance models, IEEE Trans. Med. Imag 2001: 20, 415–423.
2. Cootes TF, Taylor CJ, Anatomical statistical models and their role in feature extraction. Br J Radiol. 2004: 77, Spec No 2:S133-9.
3. Angelié E, Westenberg JJM, van der Geest RJ, Danilouchkine M, Koning G, Reiber JHC. How to define the quality of a segmentation: Application for the assessment of left ventricular function from short-axis cardiac MR data sets. Proc. Eur. Card 2002. 1: 9-10.

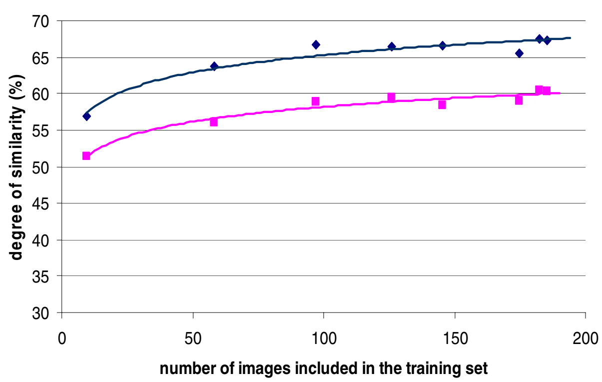

Figure 1: Variation of the averaged degree of similarity of the endocardial and epicardial segmentation (blue and pink lines respectively) as a function of the number of images included in the training set.

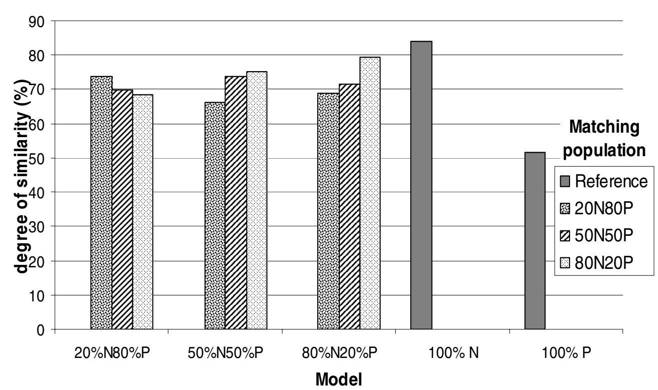

Figure 2: Average degree of similarity (ED and ES phase, endocardial and epicardial contour) when using different models on different matching population. The reference is defined when matching a model built from a population on population of images with the same characteristics.