Robust and Fast Contrast Inflow Detection for 2D X-ray Fluoroscopy

Terrence Chen, Gareth Funka-Lea, and Dorin Comaniciu

Siemens Corporation, Corporate Research,
755 College Road East, Princeton, NJ, USA

Abstract. 2D X-ray fluoroscopy is widely used in computer assisted and image guided interventions because of the real time visual guidance it can provide to the physicians. During cardiac interventions, acquisitions of angiography are often used to assist the physician in visualizing the blood vessel structures, guide wires, or catheters, localizing bifurcations, estimating severity of a lesion, or observing the blood flow. Computational algorithms often need to process differently to frames with or without contrast medium. In order to automate this process and streamline the clinical workflow, a fully automatic contrast inflow detection algorithm is proposed. The robustness of the algorithm is validated by more than 1300 real fluoroscopic scenes. The algorithm is computationally efficient; a sequence with 100 frames can be processed within a second.

Keywords: Contrast detection, fluoroscopy, vessel detection.

1 Introduction

Image-guided interventions have become more and more important for different advanced cardiovascular treatments in recent years. They are widely used for pre-operative planning, intraoperative guidance, and post-operative assessment. In interventional cardiology, 2D X-ray fluoroscopy is still used as the standard imaging modality due to the real time visual guidance it can provide to the physicians. During interventions, it is common to have several angiography acquisitions throughout the entire procedure. An angiography is acquired by injecting a radio-opaque contrast medium into the blood vessel and imaging with X-ray fluoroscopy. Physicians can then visualize the vessels and the blood flow to facilitate the operation.

Although it is rather easy for a physician to tell when the contrast medium appears in a fluoroscopy scene and make judgements accordingly, an automatic contrast inflow detection algorithm is desired for many computer assisted interventions. A few examples include: 1. In a stent enhancement application such as [7], the algorithm needs to discard frames with contrast medium since they could undermine the visibility of stent enhancement. For this purpose, the algorithm needs to determine whether there is contrast injection during the acquisition. If there is, the algorithm has to estimate at which frame the contrast inflow begins to appear and only pre-contrast frames should be processed. 2. In order to

G. Fichtinger, A. Martel, and T. Peters (Eds.): MICCAI 2011, Part I, LNCS 6891, pp. 243–250, 2011.

generated a DSA-like enhancement on coronary arteries [13], the algorithm has to know when the contrast inflow begins to appear so that masks are collected from frames which are free of contrast medium. 3. In order to improve the overlay of a 3D model when contrast agent is present [6], the system has to know the presence of contrast medium so that it can register the vessel/aorta with a presegmented vessel/aorta model. 4. In order to provide automatic roadmapping [2], the algorithm needs to know when the contrast agent appears in the fluoroscopy to disable the overlay. All these interventional procedures can be streamlined with a robust and automatic contrast inflow detection algorithm such that the physician can obtain desired information with less interactions or constraints.

Robust contrast inflow detection in a large scale is challenging. Little work has been done in the literature. Ideally, sharp changes of the overall intensity histograms along the temporal domain can be used as potential candidates indicating when contrast inflow appears. Nevertheless, this works poorly in practice since moving diaphragm and other darker regions can often appear in cardiac X-ray images and mislead the algorithm. Condurache [2] et al. proposed a method which applies morphological operations and a difference of Gaussians (DoG) filter to generate a vessel map. A threshold is then applied to the histogram of the vessel map to decide whether contrast medium appears.

Based on our experience, the main challenge of automatic contrast inflow detection lies in the high variability of data, especially when the images are often acquired using low dose radiation where the contrast medium is weakly detectable by a traditional filter. Previous methods such as [2] were evaluated only on a limited number of data sets and are with empirically set parameters. They can hardly be generalized to work in different clinical applications.

In this paper, we present a learning-based framework to overcome this problem. Given an input fluoroscopic sequence, the proposed framework first classifies it into a contrast scene (CS), where there is contrast medium injection during the acquisition, or a non-contrast scene (NCS), where contrast medium is not used during the acquisition. For a contrast scene, the algorithm then estimates at which frame the contrast inflow begins to appear in the image. We call this frame the beginning contrast frame (BCF). In order to achieve high robustness against data variability, more than 1300 real fluoroscopic sequences are collected at clinical sites in US, Asia, and Europe. Promising results are obtained.

2 Methodology

The use of the discriminative learning techniques has been proven in many medical imaging applications [1][11] as well as in interventional applications [13][10] with high accuracy and efficiency. One of the reasons is that in medical imaging, data variability is often high due to different clinical settings, patients, or dose of radiation. Learning from a large labeled database is one of the most reliable and systematic methods to achieve high robustness against data variability.

Given an input sequence of a fluoroscopic scene, the proposed framework has to first determine whether this is a contrast scene (CS). If it is, the framework

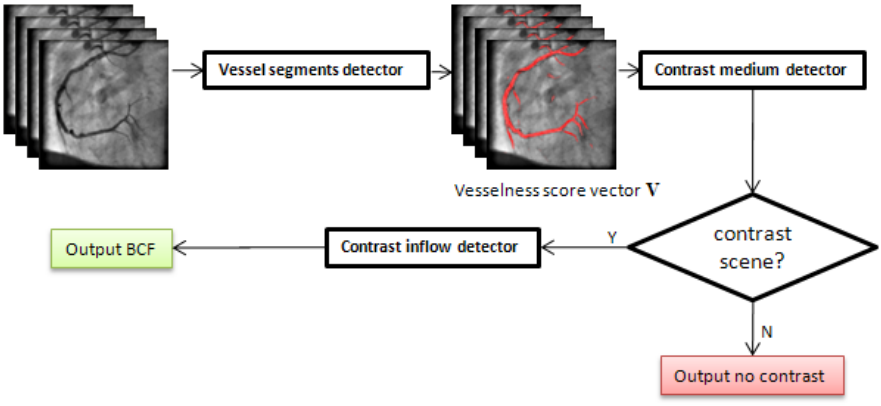

Fig. 1. Workflow of the proposed framework

then estimates the beginning contrast frame. This is more difficult than doing the second step alone, where contrast injection is known for sure to the algorithm.

To solve this problem, we propose a learning-based framework which combines multiple classifiers. Figure 1 illustrates the workflow of the proposed contrast inflow detection framework, where vessel segments detector outputs probability of a patch being inside the vessel or not. The contrast medium detector is a binary classifier to determine whether an input scene is a contrast scene (CS) or a non-contrast scene (NCS). At last, the contrast inflow detector outputs the index of the frame which has the highest probability being the BCF.

2.1 Learning-Based Vessel Detection

T o achieve a robust solution against large data variability, we make use of the discriminative learning technique to facilitate contrast agent detection. Traditional vesselness measurements use different kinds of ridge filters, gradient, edge filters to estimate the vessel-like structures inside an image. Nevertheless, due to the noisy characteristic of the X-ray images, false positives can easily happen. On the other hand, because of preferable low dose radiation in clinics, false negatives on smaller vessels can hardly be avoided. It has been shown that [12] learning based vesselness measurement not only can achieve better performance than conventional filtering based approach [4] but also can be more computationally efficient. For this reason, a learning-based vesselness measurement is applied in our framework. While [12] tried to learn the vesselness in 3D CTA data, we train a learning-based classifier to measure the vesselness for a fluoroscopic image.

To obtain a probability score for a given image to indicate how much contrast (vessel) is present, a binary classifier is implemented as a vessel segment detector, which is learned from a large set of vessel and non-vessel fluoroscopic images. The probabilistic boosting tree (PBT) algorithm [8] is used to learn the classifier. Given an image patch I_{patch}, the PBT classifier calculates the conditional probability that a vessel segment appears in this patch $P(vessel|I_{patch})$. For efficiency,

Fig. 2. Haar wavelet-like feature type examples

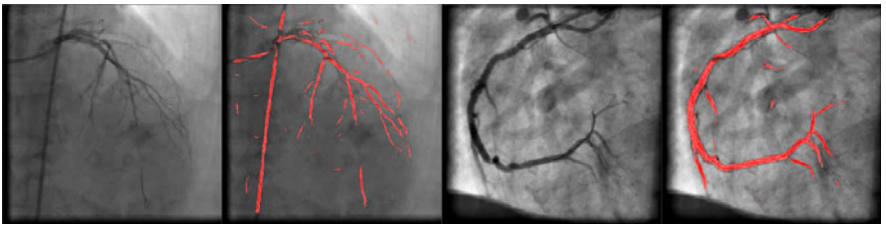

Fig. 3. Input images and the vessel detection results. The darker the red color is, the higher the score is.

the vessel detector is applied to a down-sampled image. Specifically, images are resized to 128×128 for both training and detection. We first apply a set of steerable filters [5] to find ridge points in different orientations in the contrast image. The learned PBT classifier is then applied to each ridge point to compute the probability that the ridge point is from a vessel. A detected vessel region can be defined as the image area where vessel segments are detected with high probabilities. The remaining area of the image contains primarily background structures and is defined as the background region. For vessel segment detection, we choose Haar wavelet-like features generated within a patch size of 8×10. Some examples are shown in Figure 2, which are efficiently calculated using integral image-based techniques [9]. In our implementation, 120000 patches of vessel segments were collected from coronary images to train the PBT classifier. The same number of negative patches were generated automatically by any position which is 10 mm away from the positive patches. The vessel segment detector are tested on 367 images where the coronary vessels were manually annotated. Figure 3 illustrates two examples of the original image and the response of our vessel detector.

2.2 Contrast Medium Detection

With the vessel measure for image, we define a vesselness score vector $\mathbf{V}$ for an input sequence:

$$\mathbf{V} = \{v_1, v_2, \ldots, v_n\}, \tag{1}$$

where n is the number of frames in the sequence. $v_i = \sum P(vessel|I_{patch})$ for all patches in frame i, which represents the vesselness score for the whole image of frame i.

Given the vesselness score vector $\mathbf{V}$, smoothed vectors $\mathbf{G}_{\mathbf{V}_1}, \mathbf{G}_{\mathbf{V}_2}, ..., \mathbf{G}_{\mathbf{V}_m}$ are calculated by applying Gaussian kernels to $\mathbf{V}$ with m different variance σ^2.

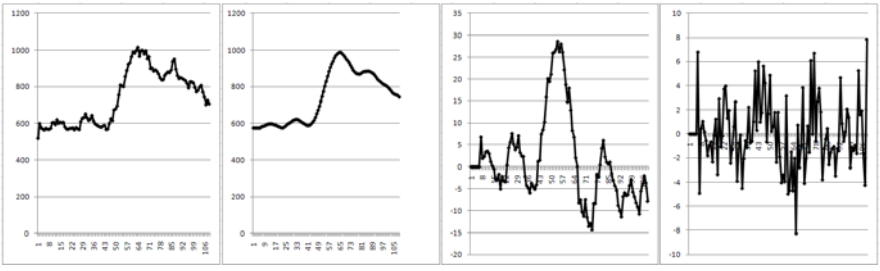

Fig. 4. From left to right: an example of the values of $\mathbf{V}$, and one example of $\mathbf{G_V}$, $\mathbf{F}'$, and $\mathbf{F}''$. The horizontal axis represents the frame index number.

In practice, we set $m = 10$ to include kernel sizes range from 3 to 31 frames. Vectors of the first order and the second order derivatives $\mathbf{F}'_1, \mathbf{F}'_2, ..., \mathbf{F}'_{m+1}$ and $\mathbf{F}''_1, \mathbf{F}''_2, ..., \mathbf{F}''_{m+1}$ for $\mathbf{V}$, $\mathbf{G}_{\mathbf{V}_1}, \mathbf{G}_{\mathbf{V}_2}, ..., \mathbf{G}_{\mathbf{V}_m}$ are then calculated. From each vector of $\mathbf{V}$, $\mathbf{G_V}$s, $\mathbf{F}'$s, and $\mathbf{F}''$s, the following values are calculated as features. The goal is to capture the magnitude and slope of the changes of vesselness measurement throughout the sequence. Let $\mathbf{X}$ be any of the aforementioned vector, $\mathbf{X} \in \{\mathbf{V}, \mathbf{G_V}\text{s}, \mathbf{F}'\text{s}, \mathbf{F}''\text{s}\}$, and let $x_a = min(\mathbf{X}), x_b = max(\mathbf{X}), x_c = median(\mathbf{X}), x_d = mean(\mathbf{X}), x_e = std(\mathbf{X})$. For each $\mathbf{X}$, it is divided into 10 disjoint parts $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{10}$, where $\sum_i |\mathbf{x}_i| = n$, and $|\mathbf{x}_i| = \frac{n}{10}$. The minimum, maximum, median, mean, and standard deviation are also calculated for all $\mathbf{x}_i$ of each $\mathbf{X}$. At last, absolute differences and ratios between each pair of these values are also added into the feature pool. For example, we find one of the most important features is $|x_b - mean(\mathbf{x}_1)|$ for $\mathbf{V}$, which capture the change of vesselness between the frame with the highest vesselness score and the beginning frames, which are contrast free. Since we only need a binary classifier instead of a probabilistic classifier, a support vector machine (SVM) [3] using radial basis function (RBF) as kernel is then applied to train a binary classifier to decide whether a given sequence is a contrast scene or a non-contrast scene. Figure 4 illustrates an example of $\mathbf{V}$, $\mathbf{G_V}$, $\mathbf{F}'$, and $\mathbf{F}''$.

2.3 Contrast Inflow Detection

If the previous classifier reports that there is no contrast injection in the input scene, the algorithm terminates and outputs NCS to the user. Otherwise, the next task is to find the beginning contrast frame (BCF).

In order to reduce the variation due to different acquisition frame rates, given the vesselness score vector $\mathbf{V}$ of an input sequence, $\mathbf{V}$ is sub-sampled using a standard acquisition frame rate. In our case, 15 fps is chosen as the standard acquisition frame rate. By doing so, we try to normalize the sequence to have a closer blood flow rate and facilitate the detection of BCF. Let $\hat{\mathbf{V}}$ be the sub-sampled vector.

Manual labels of BCF for all the sequences in our database are used as gold standard for a supervised learning. Let x be the labeled position in $\mathbf{V}$ and $\hat{x}$ be

its corresponding position in $\hat{\mathbf{V}}$. $\hat{x}$ is marked as a positive example. For other frame $i \in \{0, ..., |\hat{\mathbf{V}}|\}$, if $i < \hat{x} - 3$ or $i > \hat{x} + 3$, then i is marked as a negative example. 1D Haar features generated from a window size of 11 frames at each example are fed into the PBT classifier [8] to train a classifier $P(f|\hat{\mathbf{V}})$, where f indicates the frame where contrast inflow begins. As a result, the BCF is learned by the classifier:

$$BCF = argmax_{f \in \{0,..,|\hat{\mathbf{V}}|\}} P(f|\hat{\mathbf{V}}). \quad (2)$$

3 Numerical Evaluation

In this section, we present the quantitative evaluation conducted on 1348 fluoroscopy sequences acquired during different interventional procedures, including balloon angioplasty, chronic total occlusion (CTO), intravascular ultrasound catheter pullback, and electrophysiology (EP) procedure. The sequences are acquired in clinical sites in US, Asia, and Europe. Our goal is to collect a database with large data variability.

In the first experiment, we evaluate the performance of the contrast medium detection. In the 1348 sequences, there are 956 contrast scenes and 492 non-contrast scenes. We apply a 4-fold cross validation on our data set, evenly distributed the contrast and non-contrast scenes into 4 parts. For each run, three folds of the data are used for training and the remaining one for testing. The performance is validated by missed detection and false detection. Missed detection denotes that a contrast scene is misclassified as a non-contrast scene and missed detection rate (MDR) is the number of missed detection over the total number of contrast scenes. False detection denotes a non-contrast scene misclassified as a contrast scene. False detection rate (FDR) is the number of false detection over the total number of non-contrast scenes. Table 1 columns A1 and A2 show the average training and testing errors, where the missed detection rate is less than 2% and the false detection rate is less than 5%.

Table 1. Training and testing error of contrast medium detection. The unit of columns B1, B2, B3 is frame number.

	A1. MDR	A2. FDR	B1. Mean	B2. Median	B3. Max
Avg. training error	1.5%	3.3%	0.75	1.82	8
Avg. testing error	1.8%	4.1%	0.9	1.98	11

Although both the missed detection rate and the false detection rate are quite low, we further look into the misclassified scenes and notice the following reasons for misclassification. Most of the missed detection is due to extremely low signal to noise ratio (SNR) or poor contrast visibility and when the scene includes only a major vessel branch with very few branches. In such cases, the vesselness scores of contrast-filled images and non-filled images are hardly distinguishable. Figure 5 (A)-1, and (A)-2 shows one of such scenes. On the other hand, the main

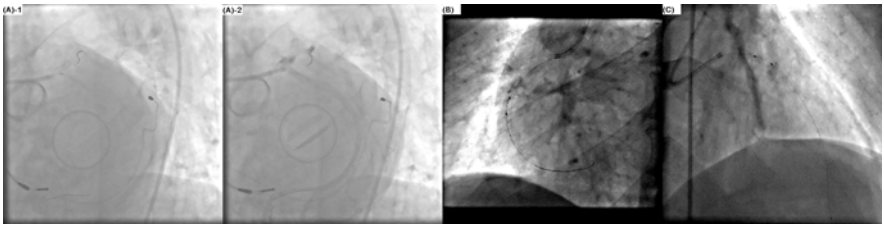

Fig. 5. (A)-1, (A)-2: A missed detection example. (A)-1 shows one frame before contrast injection. (A)-2 shows one of the frames with maximal contrast medium. (B), (C): Two examples of false detection where there is contrast residue moving in the background.

reason for false detection is because in some scenes contrast residue exists in the vessel structures moving in the background, which confuses the detector. Figure 5 (B) and (C) illustrates two examples of false detection. In theory, catheters are hardly distinguishable from the vessels and hence movement of catheters may cause false detection as well. However, this is rarely observed in our experiments.

In the next experiment, we evaluate the accuracy of the beginning contrast frame (BCF), which is measured by the error $|L - D|$, where L is the manual label of BCF, and D is the estimated BCF by the detector. Among the 956 contrast scenes, we again apply a 4 fold cross-validation. Table 1 column B1, B2, and B3 show the average training and testing errors, while Max shows the maximum error across all 4 runs. The mean error is less than 2 frames, which is highly accurate. The accuracy is sufficient for most of real applications. For example, in the stent enhancement application, one can select frames before BCF to generate the enhanced stent.

It is worth noting that another useful output of a contrast inflow detection is a frame with sufficient contrast medium (e.g., presenting one angiography frame to the physician). This can be obtained easily by outputting the frame with the largest v_i in an appropriate $\mathbf{G_V}$ (c.f. $\mathbf{G_V}$ in figure 4)

Lastly, we report the computation time of the proposed framework. Since the vessel detector is only applied to small sub-sampled images (128×128), and the remaining two classifiers are applied to 1D vectors, the computation is very fast. It processes a sequence with 100 frames of original size 1024×1024 within 1 second on a Intel Xeon PC (2.2 GHz) with 3 GB RAM. The average processing time on all the 1348 sequences is about 0.36 seconds, where the average number of frames is 49.3 and the size of images range between 384×384 to 1024×1024.

4 Conclusion

Automatic contrast inflow detection has been found important in many interventional applications. Nevertheless, there is little work in the literature to solve this problem in a large scale. In this paper, we present a learning-based framework which combines three learned classifiers including both SVM and boosting

methods to detect the contrast inflow. Experiments on more than 1300 real fluoroscopic sequences demonstrate that our method is very robust against different clinical settings, dose of radiation, patients, etc. Both missed detection and false detection rates are below 5% and the begin contrast frame can be accurately estimated. In addition, it is with little computational overhead and can be performed within 1 second for most of the fluoroscopic scene. The next step of our work includes but is not limited to extending the current framework for online detection.

References

1. Chen, T., Zhang, W., Good, S., Zhou, S., Comaniciu, D.: Automatic ovarian follicle quantification from 3d ultrasound data using global/local context. In: ICCV (2009)
2. Condurache, A., Aach, T., Eck, K., Bredno, J.: Fast Detection and Processing of Arbitrary Contrast Agent Injections in Coronary Angiography and Fluoroscopy. In: Bildverarbeitung fur die Medizin, Berlin, pp. 5–9 (2004)
3. Cortes, C., Vapnik, V.: Support-Vector Networks. Machine Learning 20 (1995)
4. Frangi, A., Niessen, W., Vincken, K., Viergever, M.: Multiscale vessel enhancement filtering. In: Wells, W.M., Colchester, A.C.F., Delp, S.L. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 130–137. Springer, Heidelberg (1998)
5. Freeman, W.T., Adelson, E.H.: The Design and Use of Steerable Filters. PAMI 13(9), 891–906 (1991)
6. John, M., Liao, R., Zheng, Y., Nöttling, A., Boese, J., Kirschstein, U., Kempfert, J., Walther, T.: System to Guide Transcatheter Aortic Valve Implantations Based on Interventional C-Arm CT Imaging. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010. LNCS, vol. 6361, pp. 375–382. Springer, Heidelberg (2010)
7. Koolen, J.J., Vanhet Veer, M., Hanekamp, C.E.E.: A StentBoost image enhancement: first clincal experience. In: MEDICAMUNDI: Clinical Applications (2005)
8. Tu, Z.: Probablistic Bossting-Tree: Learning discriminative models for classfication, recognition, and clustering. In: ICCV (2005)
9. Viola, P., Jones, M.: Robust Real-time Object Detection. IJCV (2002)
10. Wang, P., Chen, T., Zhu, Y., Zhang, W., Zhou, K., Comaniciu, D.: Robust Guidewire Tracking in Fluoroscopy. In: CVPR (2009)
11. Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D.: Four-chamber heart modeling and automatic segmentation for 3d cardiac ct volumes using marginal space learning and steerable features. IEEE Trans. on Med. Img. 27(11), 1668–1681 (2008)
12. Zheng, Y., Loziczonek, M., Georgescu, B., Zhou, S., Vega-Higuera, F., Comaniciu, D.: Machine Learning Based Vesselness Measurement for Coronary Artery Segmentation in Cardiac CT Volumes. In: SPIE Med. Img. (2011)
13. Zhu, Y., Prummer, S., Wang, P., Chen, T., Comaniciu, D., Ostermeier, M.: Dynamic Layer Separation for Coronary DSA and Enhancement in Fluoroscopic Sequences. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5762, pp. 877–884. Springer, Heidelberg (2009)