Dynamic Layer Separation for Coronary DSA and Enhancement in Fluoroscopic Sequences

Ying Zhu[1]*, Simone Prummer[2], Peng Wang[1],
Terrence Chen[1], Dorin Comaniciu[1], Martin Ostermeier[2]

[1] Siemens Corporate Research Inc., Princeton, NJ, USA
[2] Siemens AG, Health Care, MED AX PLM-I, Forchheim, Germany

Abstract. This paper presents a new technique of coronary digital subtraction angiography which separates layers of moving background structures from dynamic fluoroscopic sequences of the heart and obtains moving layers of coronary arteries. A Bayeisan framework combines dense motion estimation, uncertainty propagation and statistical fusion to achieve reliable background layer estimation and motion compensation for coronary sequences. Encouraging results have been achieved on clinically acquired coronary sequences, where the proposed method considerably improves the visibility and perceptibility of coronary arteries undergoing breathing and cardiac movements. Perceptibility improvement is significant especially for very thin vessels. Clinical benefit is expected in the context of obese patients and deep angulation, as well as in the reduction of contrast dose in normal size patients.

1   Introduction

Digital subtraction angiography (DSA) is a fluoroscopy technique to clearly visualize blood vessels by subtracting a pre-contrast image called *mask* from later images once the contrast medium has been introduced. In this work, we introduce a new technique called coronary DSA (cDSA) to better visualize coronary vessels in 2D dynamic fluoroscopic sequences of the heart. Using a small number of pre-contrast masks, cDSA produces sequences of dynamic coronary arteries by separating and subtracting sequences of moving background layers. cDSA is an important technique with broad applications in image guided cardiovascular intervention. Fig. 1 shows two applications of cDSA. First, the separation of background and coronary layers enables the function of fade-in and fade-out of the dynamic background structures, thus giving clinicians more options in displaying the coronary arteries in motion during cardiac interventions or for diagnosis purpose. Second, with the coronary layer extracted from fluoroscopic sequences, we are able to virtually enhance the contrast medium for improved visibility and perceptibility of coronary arteries, which brings clinical benefits in the context of obese patients and deep angulation.

A main challenge for cDSA is to deal with complex motion caused by cardiac, breathing and patient table movements. In cardiac fluoroscopic sequences, both static bone tissues and tissues undergoing a mixture of cardiac and respiratory

---

* Corresponding author: yingzhu@siemens.com
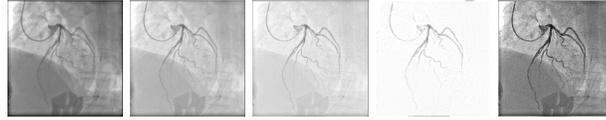
Fig. 1. cDSA applications. From left to right: original image, two images with fading background layer, coronary layer, coronary enhanced image.
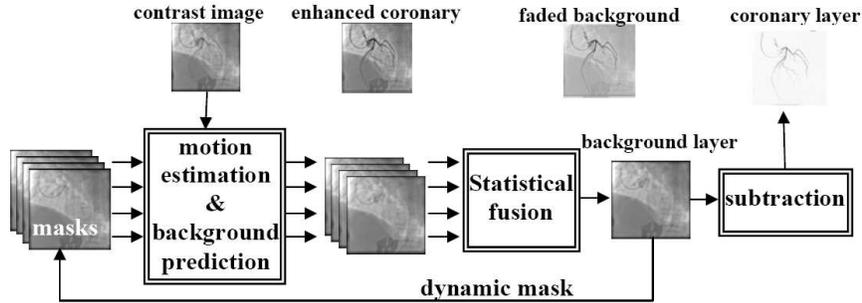
Fig. 2. Bayesian framework of dynamic layer separation.

movements can be seen in transparency. Existing techniques of motion correction [4] remain largely insufficient in dealing with such complex motion. Related work has been reported in dealing with transparent motion [1, 2, 5, 6]. In [7], a technique based on non-parametric motion estimation has been proposed, where a dense motion field is used for motion correction between a mask and a contrast image, and learning-based method is used to facilitate motion estimation.

We present a Bayesian framework for tracking the moving layer of dynamic background structures to achieve coronary subtraction in cardiac fluoroscopic sequences. Dense motion estimation between mask images and a contrast image are used to predict the background layer of the contrast image, and predictions from multiple masks are statistically fused to obtain the final estimation of the background layer. Compared to the method in [7] which selects one mask image for motion compensation, the Bayesian framework improves the accuracy of background layer estimation through uncertainty propagation and statistical fusion of motion compensation from multiple masks.

2 Method

In X-ray imaging, the intensity of the energy flux undergoes exponential attenuation through layers of tissues, resulting in multiplicative transparency [4]. With logarithmic postprocessing, fluoroscopic images are represented by an additive model consisting of multiple layers. In cDSA, only two layers are considered to simplify the problem, a coronary layer defined as the transparent layer containing coronary arteries filled with contrast medium, and a background layer defined as the transparent layer containing background structures. Denote $I_t(\mathbf{x})$, $I_{C,t}(\mathbf{x})$

and $I_{B,t}(\mathbf{x})$ as the contrast-filled frame, its coronary layer and background layer at time t respectively, where $\mathbf{x}$ is the pixel location. The additive layer composition model is expressed as $I_t(\mathbf{x}) = I_{C,t}(\mathbf{x}) + I_{B,t}(\mathbf{x})$. The goal is to remove the background layer to obtain the layer of coronary arteries while both layers are undergoing cardiac, respiratory and other types of movements. The proposed Bayesian framework is illustrated in Fig. 2. First, prior to contrast injection, a small number of images are acquired at different cardiac and breathing phases to serve as static masks for background estimation. Second, once the contrast medium has been introduced, motion estimation is performed between each mask and a contrast image, and the resulting motion field is used to predict the background layer of the contrast image. Predictions from multiple masks are fused statistically to obtain a final estimate of the background layer. At last, the coronary layer is estimated by subtracting the background estimate from the contrast image. In processing a fluoroscopic sequence, layer estimates obtained from previous frames are also used as dynamic masks to predict the background layer of a current frame. In this framework, we assume that the C-arm remains still. New masks have to be reacquired for cDSA when change of angulation occurs.

2.1   Background motion estimation

We use the non-parametric approach introduced in [7] to estimate the motion between a mask image I_m and a contrast image $I_t(\mathbf{x})$. First, a technique of learning-based vessel segment detection is applied to the contrast image to roughly separate the image areas of vessels from the region of background structures and to exclude most of the vessel areas from motion estimation. Second, the Lucas-Kanade-Fusion algorithm is applied to estimate a dense motion field $\mathbf{v}(\mathbf{x})$ between the mask image and the background region of the contrast image. The algorithm combines the Lucas-Kanade algorithm which iteratively estimates incremental motion and the covariance-based filtering technique to retain spatial smoothness and consistency of the motion field. For every pixel $\mathbf{x}$, the algorithm computes an estimate of the displacement vector $\widehat{\mathbf{v}}(\mathbf{x})$ locally. In addition, the algorithm also estimates its covariance $C(\widehat{\mathbf{v}}(\mathbf{x}))$ to characterize the uncertainty in the motion estimation. In homogeneous image areas with lack of textures or areas with vessel pixels excluded from motion calculation, the motion estimates tend to be unreliable and their covariance matrices have large eigenvalues.

2.2   Background layer prediction with uncertainty propagation

Given the motion estimation $\widehat{\mathbf{v}}(\mathbf{x})$ and its covariance $C(\widehat{\mathbf{v}}(\mathbf{x}))$, the probability distribution of the motion vector $\mathbf{v}(\mathbf{x})$ can be approximated as a Gaussian distribution with mean $\widehat{\mathbf{v}}(\mathbf{x})$ and covariance $C(\widehat{\mathbf{v}}(\mathbf{x}))$.

$$\mathbf{v}(\mathbf{x}) \sim N(\widehat{\mathbf{v}}(\mathbf{x}), C(\widehat{\mathbf{v}}(\mathbf{x}))); \quad E[\mathbf{v}(\mathbf{x})] = \widehat{\mathbf{v}}(\mathbf{x}), \quad Cov[\mathbf{v}(\mathbf{x})] = C(\widehat{\mathbf{v}}(\mathbf{x})) \quad (1)$$

Pixel values in the background layer $I_{B,t}(\mathbf{x})$ are predicted from the mask image.

$$I_{B,t}(\mathbf{x}) = I_m(\mathbf{x} + \mathbf{v}(\mathbf{x})) \quad (2)$$

In contrast to the method in [7] which only takes into account the mean of the motion estimates, we incorporate second order statistics and derive the prediction probability density functions (PDFs) of pixel values in the background layer $p(I_{B,t}(\mathbf{x})|I_m)$. In general, the transformation function $I_m(\mathbf{x} + \mathbf{v}(\mathbf{x}))$ is a nonlinear function of $\mathbf{v}(\mathbf{x})$ and techniques such as linearization and unscented transformation [3] are required to parameterize the means and covariances of the probability distribution. Due to the computational complexity of the unscented transformation, we choose to linearize the transformation function as follows.

$$
\begin{aligned}
I_m(\mathbf{x} + \mathbf{v}(\mathbf{x})) &\approx I_m(\mathbf{x} + \widehat{\mathbf{v}}(\mathbf{x})) + \bigtriangledown^T I_m(\mathbf{x} + \widehat{\mathbf{v}}(\mathbf{x}))[\mathbf{v}(\mathbf{x}) - \widehat{\mathbf{v}}(\mathbf{x})] \\
\bigtriangledown I_m(\mathbf{x} + \widehat{\mathbf{v}}(\mathbf{x})) &= [\partial_x I_m(\mathbf{x} + \widehat{\mathbf{v}}(\mathbf{x})), \partial_y I_m(\mathbf{x} + \widehat{\mathbf{v}}(\mathbf{x}))]^T
\end{aligned}
\tag{3}
$$

where $\bigtriangledown I_m(\mathbf{x} + \widehat{\mathbf{v}}(\mathbf{x}))$ denotes the gradient vector of the transformed image $I_m(\mathbf{x} + \widehat{\mathbf{v}}(\mathbf{x}))$. The mean and variance of $I_{B,t}(\mathbf{x})$ (2) are approximated as

$$
\begin{aligned}
E[I_{B,t}(\mathbf{x})|I_m] &= I_m(\mathbf{x} + \widehat{\mathbf{v}}) \\
Var[I_{B,t}(\mathbf{x})|I_m] &= \bigtriangledown^T I_m(\mathbf{x} + \widehat{\mathbf{v}}(\mathbf{x})) \cdot C(\widehat{\mathbf{v}}(\mathbf{x})) \cdot \bigtriangledown I_m(\mathbf{x} + \widehat{\mathbf{v}}(\mathbf{x}))
\end{aligned}
\tag{4}
$$

Through linearization of the transformation function, the uncertainties in motion estimation are propagated to the prediction of background pixel values. The prediction PDF is approximated by a Gaussian distribution.

$$
p(I_{B,t}(\mathbf{x})|I_m) = N(I_{B,t}(\mathbf{x}); E[I_{B,t}(\mathbf{x})|I_m], Var[I_{B,t}(\mathbf{x})|I_m])
\tag{5}
$$

2.3   Statistical fusion with multiple mask images

In cardiac interventional procedures, sequences of fluoroscopic images showing cardiovascular structures in motion are acquired to provide real-time image guidance. Multiple image frames are often captured before a contrast medium flows into coronary arteries. These pre-contrast frames capture the background layer from different cardiac and respiratory phases and are used as static mask images. To deal with large image motion caused by deep breathing, we also include the estimated background layers from previous contrast frames as dynamic mask images.

Denote $\{I_{m,i}(\mathbf{x}) : i = 1, \cdots, n_s\}$ as the static mask images acquired at time $t_1, \cdots, t_{n_s}$, and $\{I_{D,k}(\mathbf{x}) = I_{B,t-k}(\mathbf{x}) : k = 1, \cdots, n_d\}$ as the dynamic mask images coming from the estimated background layers of frames $t - 1, \cdots, t - n_d$. Through motion estimation and uncertainty propagation, we obtain multiple prediction PDFs of the background layer.

$$
\begin{aligned}
p(I_{B,t}(\mathbf{x})|I_{m,t_i}) &= N(I_{B,t}(\mathbf{x}); m_{t,t_i}(\mathbf{x}), \sigma^2_{t,t_i}(\mathbf{x})) \quad (i = 0, \cdots, n_s) \\
p(I_{B,t}(\mathbf{x})|I_{D,k}) &= N(I_{B,t}(\mathbf{x}); m_{t,t-k}(\mathbf{x}), \sigma^2_{t,t-k}(\mathbf{x})) \quad (k = 1, \cdots, n_d)
\end{aligned}
\tag{6}
$$

where $m_{t,t_i}(\mathbf{x}) = E[I_{B,t}(\mathbf{x})|I_{m,i}]$, $\sigma^2_{t,t_i}(\mathbf{x}) = Cov[I_{B,t}(\mathbf{x})|I_{m,i}]$, $m_{t,t-k}(\mathbf{x}) = E[I_{B,t}(\mathbf{x})|I_{D,k}]$, $\sigma^2_{t,t-k}(\mathbf{x}) = Cov[I_{B,t}(\mathbf{x})|I_{D,k}]$ are the estimated mean and covariance of background pixel values. Fusing multiple estimates of the background

layer, we obtain the linear minimum-mean-square-error (MMSE) estimate as

$$\widehat{I}_{B,t}(\mathbf{x}) = \frac{\sum\limits_{i=0}^{n_s} \sigma_{t,t_i}^{-2}(\mathbf{x})m_{t,t_i}(\mathbf{x}) + \sum\limits_{k=1}^{n_d} \sigma_{t,t-k}^{-2}(\mathbf{x})m_{t,t-k}(\mathbf{x})}{\sum\limits_{i=0}^{n_s} \sigma_{t,t_i}^{-2}(\mathbf{x}) + \sum\limits_{k=1}^{n_d} \sigma_{t,t-k}^{-2}(\mathbf{x})} \tag{7}$$

and the estimation of coronary layer is obtained through subtraction

$$\widehat{I}_{C,t}(\mathbf{x}) = I_t(\mathbf{x}) - \widehat{I}_{B,t}(\mathbf{x}) \tag{8}$$

With the background layer separated from the coronary layer, it is straightforward to fade out the background layer or to enhance the coronary layer by layer composition.

$$\alpha_C \widehat{I}_{C,t} + \alpha_B \widehat{I}_{B,t} \quad (\alpha_C \geq 1, 0 \leq \alpha_B \leq 1) \tag{9}$$

To fade out the background layer, we set $\alpha_C = 1$ and decrease α_B. To virtually enhance the contrast, we set $\alpha_B = 1$ and increase α_C.

3 Experimental Results

Fluoroscopic sequences of 30 patients acquired during cardiovascular intervention have been used to evaluate the proposed cDSA method. The sequences were acquired on Angiographic C-arm systems (AXIOM Artis, Siemens Medical Solution) from different rotational angles and included cases of patients holding breath, deep breathing as well as table movements. Since the proposed cDSA technique was planned at the end of the imaging chain for general use cases, the test sequences were not selected particularly by disease phenotypes. Nevertheless, they contain cases of stenosis, lesions and stent placement. Each image frame has either 512×512 pixels or 1024×1024 pixels, and the pixel size is either $0.17mm$ or $0.28mm$. Frames at the beginning of each sequence and before the contrast medium starts to flush into the coronaries are sampled to define static mask images used in processing the following frames. The number of mask images ranges from 3 to 9 frames in each sequence, and they are uniformly sampled from half to one cardiac cycle. There are between 18 to 150 frames per sequence showing intra-coronary flow of the contrast medium, and in total there are 1829 such frames used to compute performance metrics. All testing frames are scaled to 8-bit images with gray values between 0 and 255.

To evaluate the performance of the background layer estimation, we computed the mean squared error (MSE) of the background estimation in each frame, i.e. the mean squared difference between estimated background pixels and the actual background pixels of a testing frame I_t in the background region Ω_t of the frame, $MSE = \frac{1}{|\Omega_t|} \sum\limits_{\mathbf{x} \in \Omega_t} ||\widehat{I}_{B,t}(\mathbf{x}) - I_t(\mathbf{x})||^2$. The histogram of the MSE over 1829 testing frames is shown in Fig. 3-1. MSEs of the $50th$, $60th$, $70th$, $80th$ and $90th$ percentiles are 11.77, 16.02, 18.07, 21.34 and 24.40 respectively, which corresponds to 3.43, 4.00, 4.25, 4.62 and 4.94 of gray value difference. The mean MSE

(a)

(b)

(c)

(d)

(e)
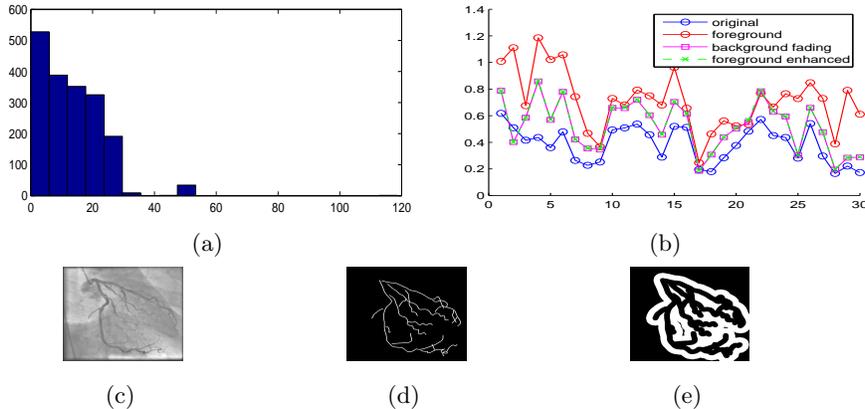
Fig. 3. Performance metrics. (a) Histogram of MSEs over 1829 testing frames. (b) JM distances measured in the original images, images with faded background layers and images with enhanced coronary layers. (c) A test frame. (d) Annotated vessel pixels (in white) and (e) Background pixels (in white) used to compute JM distance.

Table 1. Mean, median and standard deviation (std) of JM distances over original images, coronary layers, images with faded background layers, images with enhanced coronary layers.

cDSA	mean	median	std
original ($\alpha_C = \alpha_B = 1$)	0.3845	0.4262	0.1366
coronary layer ($\alpha_C = 1, \alpha_B = 0$)	0.7170	0.7289	0.2244
faded background ($\alpha_C = 1, \alpha_B = 0.5$)	0.5239	0.5604	0.1865
enhanced coronary ($\alpha_C = 2, \alpha_B = 1$)	0.5271	0.5706	0.1869

is 13.06 and the standard deviation is 10.00. To evaluate how cDSA improves the visibility conditions around coronary arteries, we use Jeffries-Matusita (JM) distance to measure the difference in the gray values between coronary arteries and surrounding background areas. In each testing sequence, we manually annotated coronary arteries in a contrast-filled frame. The distribution of pixel values in the areas occupied by coronary vessels (Fig. 3-(d)) was computed as p_C. The distribution of the gray values of background pixels in the areas surrounding coronary vessels (Fig. 3-(e)) was computed as p_B. The JM distance is defined as $JM(p_C, p_B) = [\int_{\mathbf{z}} (\sqrt{p_C(\mathbf{z})} - \sqrt{p_B(\mathbf{z})})^2]^{1/2}$. The JM distance measures how well the two gray value distributions are separated from each other. It is bounded between 0 and $\sqrt{2}$. Higher values of the JM distance is related to better visibility conditions around coronary arteries. The JM distances measured on the original images, the coronary layers, the images with faded background layers and the images with enhanced coronary layers are plotted in Fig. 3-(b). Their mean, median and standard deviation are further compared in Table 1.The JM distances measured on the coronary layers are consistently higher compared to the JM distances measured on the original images. In 90% of the cases, the JM

distances are improved by both fading out the background layers and enhancing the coronary layers, suggesting improved visibility conditions achieved by cDSA. Fig. 4 shows several image results of layer separation, fade-out of background layers and enhancement of coronary layers. The visibility and perceptibility of the coronary arteries is considerably improved through enhancing coronary layers. In particular, thin vessels are made more visible by the cDSA method. We have also observed that the use of dynamic masks compensates table motion considerably due to the fact that image motion between adjacent frames is small even though over time the accumulated image motion can be large.

4 Discussion

We have presented a novel method for coronary digital subtraction angiography in $2D$ dynamic fluoroscopic sequences. Through dense motion estimation and statistical fusion, a Bayesian framework is proposed to estimate the moving layers of background structures in cardiac fluoroscopic sequences and to obtain the layer of coronary arteries through subtraction. Using this method to separate coronary layers from background structures, we are able to fade out the background layer or virtually enhance the contrast by enhancing the coronary layers to improve the image quality. Encouraging results have been obtained in terms of visibility and perceptibility improvement of coronary vessels and thin vessels in particular. The clinical benefits are expected for cardiac intervention in the context of obese patients and deep angulation. In addition, the ability of coronary enhancement also allows for the reduction of the contrast medium used in normal size patients. Our future study includes the evaluation of cDSA on cases involving diluted contrast medium.

References

1. Vincent Auvray, Jean Lienard and Patrick Bouthemy, "Multiresolution Parametric Estimation of Transparent Motions and Denoising of Fluoroscopic Images, *MIC-CAI*, 352-360, 2005.
2. S. X. Ju, M. J. Black and A. D. Jepson, "Skin and Bones: Multi-layer, Locally Affine, Optical Flow and Regularization with Transparency", *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 307-314, 1996.
3. S. J. Julier and J. K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems", pp. 182-193, 1997.
4. E. H. W. Meijering, W. J. Niessen and M. A. Viergever, "Retrospective Motion Correction in Digital Subtraction Angiography: A Review", *IEEE Trans. Medical Imaging*, Vol. 18, No. 1, pp. 2-21, Jan. 1999.
5. B. Sarel and M. Irani, "Separating Transparent Layers Through Layer Information Exchange", *Proc. European Conf. on Computer Vision*, pp. 328-341, 2004.
6. R. Szeliski, S. Avidan, and P. Anandan, "Layer Extraction from Multiple Images Containing Reflections and Transparency", *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 246-253, 2000.
7. Y. Zhu, S. Prummer, T. Chen, M. Ostermeier and D. Comaniciu, "Coronary DSA: Enhancing Coronary Tree Visibility through Discriminative Learning and Robust Motion Estimation", SPIE Medical Imaging, 2009.

JM=0.5097	MSE=11.96	JM=1.1109	JM=0.4015	JM=0.4024
JM=0.5124	MSE=15.38	JM=0.6572	JM=0.6139	JM=0.6180
JM=0.2640	MSE=17.66	JM=0.7428	JM=0.4210	JM=0.4221
JM=0.4786	MSE=27.65	JM=1.0577	JM=0.7795	JM=0.7807
JM=0.5715	MSE=17.93	JM=0.7693	JM=0.7833	JM=0.7904
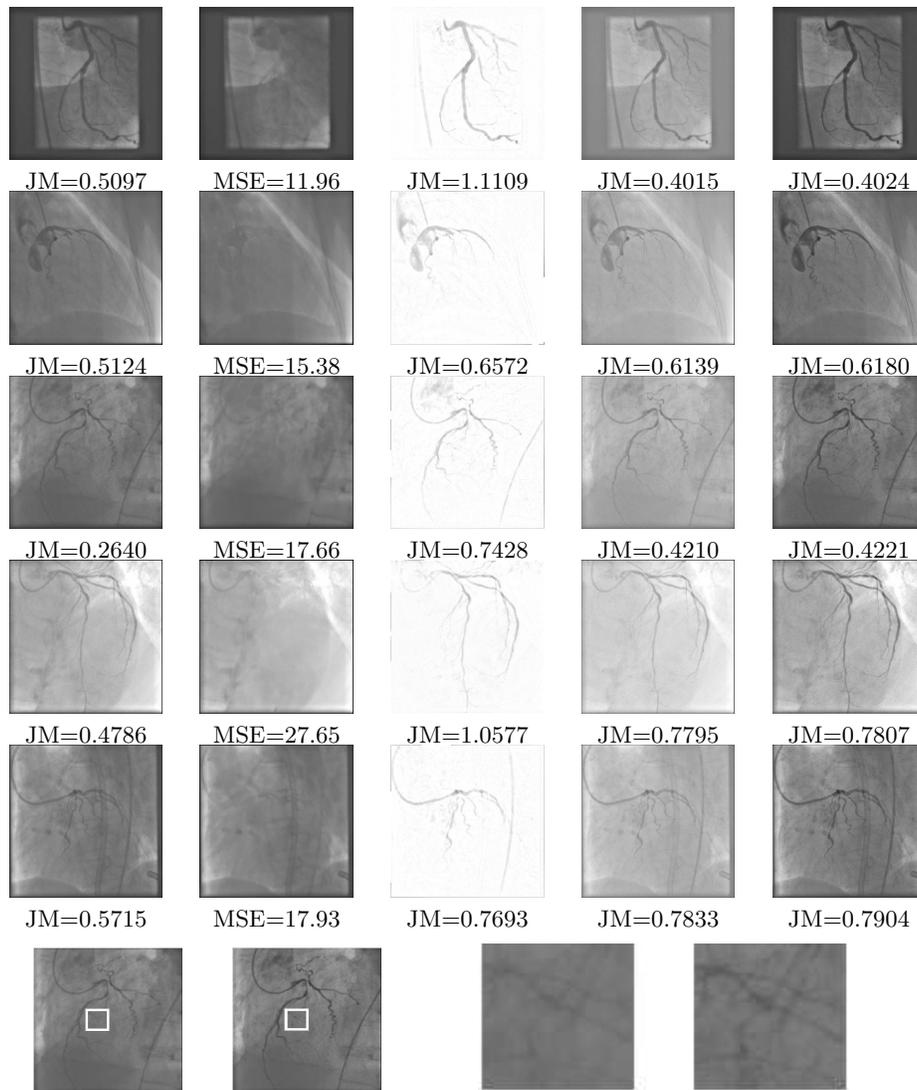
Fig. 4. cDSA results. Column 1: original images; column 2: background layer estimation; column3: coronary layer estimation; column 4: images with faded background; column5: images with enhanced coronary layers. Last row from left to right: original image and coronary enhanced image with white boxes enclosing thin vessels, zoom in on patches of thin vessels from original image and coronary enhanced image.