Alvarez-Gomez Julio*, Spieß Arne, Roth Hubert and Wahrburg Jürgen

A Deep-Learning Approach to Detect Fiducials in Planar X-Ray Images for Undistortion of Conventional C-Arm Images

Abstract: In some applications such as 2D-3D registration, undistorted images are required to achieve optimal results. These types of images can be obtained from a distortion-free C-arm (flat-panel detector) or by undistorting the images given from a conventional C-arm (analogue image intensifier.) Undistorting images require a plate with fiducials connected to the C-arm detector. Detecting fiducials is affected by differences in the image contrast due to elements in the background. Therefore, the results vary from image to image and could require manual tuning of parameters. We propose a deep-learning approach for detecting undistortion-plate-fiducials in X-ray images to overcome the drawbacks previously stated. With an undistortion plate, we took 1120 X-Rays using a C-arm in different poses. Every X-ray is afterward cut into 60 sub-images. We used these sub-images for training a convolutional neural network (CNN). Comparing the CNN and a traditional image processing method based on the Hough Circle algorithm, we found that the detected fiducials using the traditional method give a similar fiducial positioning error. Nevertheless, the fiducial detection rate goes from 89.7% using the traditional method to 100% with the developed CNN. The results show that the detection rate and precision of our deep-learning approach guarantee the undistortion of conventional C-Arm images.

Keywords: Deep-Learning, Fiducial Detection, X-Ray, Image Undistortion, Convolutional Neural Network, C-arm.

_____
*Corresponding author: Alvarez-Gomez Julio: Institute of Control Engineering (RST) and Center for Sensor Systems (ZESS), University of Siegen, Siegen, Germany, e-mail: julio.agomez@uni-siegen.de
Spieß Arne: Center for Sensor Systems (ZESS), University of Siegen, Siegen, Germany.
Roth Hubert, Wahrburg Jürgen: Institute of Control Engineering (RST) and Center for Sensor Systems (ZESS), University of Siegen, Siegen, Germany.

1 Introduction

In conventional C-arms, i.e., C-arms equipped with analog image intensifiers, distortions are a mixture of barrel, S-shape, and spiral distortions, and they are caused due to gravitational effects and the C-arm pose. That makes it difficult to create a general model for the C-arm distortion because it changes with time and position.

Undistorted images are required in some applications, which need patient anatomy to be represented as precise as possible. In the case of 2D-3D registration, for example, it gives optimal results only when the used images are undistorted. In our research of minimal invasive computer-assisted spine surgery, we register Computer Tomography scans (CT-Scan) with X-Ray images. Today, most of the commercial computer-assisted surgery systems use distortion-free C-arms, i.e., C-arms with flat-panel detector, for the registration process. However, we want to increase the reach of these systems by offering the possibility to incorporate conventional C-arms.

A typical way to undistort C-arm images uses a plate with fiducials in a grid shape [1], which are later detected and used to unwarp the image [2]. Detecting the fiducials can be a challenging problem because of the contrast variation depending on the elements in the background. Some Robust algorithms make assumptions when there is no enough information on the image for proper detection [3]. Other studies use statistical models that predict the position of all of the distorted fiducials in a calibration plate just by knowing the position of some of them [4]. This last approach aims to reduce the number of detected fiducials and the amount of obstructed area on the X-ray image. Nevertheless, a robust fiducial detector is still required. With the idea of having a parameterless and robust detector, we propose to use a deep-learning approach for detecting undistortion-plate-fiducials on X-ray images.

2 Method

We aim to incorporate in our workflow a standalone procedure, which makes an inference over an image with fiducials and later undistort an X-ray image. This requires an undistortion plate, an image process unit to detect fiducials, an image undistortion process, and optionally an inpainting procedure to erase the fiducials from the undistorted image. This work is focused on the image process unit to detect fiducials, but we will describe as well the other required components.

Figure 1: Undistortion Plate with 60 Fiducials

2.1 Undistortion Procedure

For our setup, we have an undistortion plate, which consists of 60 fiducials in a grid of 8x8 without corners, as the Figure 1 shows. The fiducials are visible on the X-Ray image, so they are used to determine the distortion of a particular image.

From the construction of the plate, we know the X and Y coordinate of every fiducial. Then we detect the X and Y coordinate of the fiducials on the X-ray image, which are distorted. For the proper performance of the undistortion procedure, all of the fiducials have to be detected. After that, we calculate a 5th-order polynomial transformation to find the distortion function of the image, and using an inverse mapping procedure, we undistort the image [2]. Once the image is undistorted, we execute an inpainting procedure to remove the fiducials from the image [5].

2.2 Fiducial Detection

We have searched for some robust image processing algorithms to detect fiducials. We created our implementation using the filtering stage used by Livyatan et al. [3], and later using the Hough transform algorithm for fiducial detection. Our implementation was automated as much as possible, but the Canny edge detector parameter included in the Hough algorithm profoundly impacts the detection performance. Therefore, we kept the Canny value as a variable.

Additionally, elements on the X-ray image, i.e., different patient anatomy, change the image contrast, so finding an exact Canny value is not possible. In order to avoid errors in the undistortion process, we decided to implement a CNN that helps us with the fiducial's detection. However, we use this conventional image processing approach with two main objectives, one is to establish a ground truth, and second is to decrease the time for labeling the data required to develop the CNN.
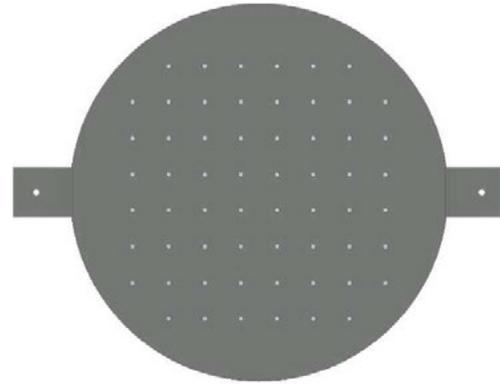
2.3 CNN Approach for Fiducial Detection

For simplifying the deep-learning approach, we exploit a particular feature of the image distortion. We have noticed that regardless of the distortion in an image, we can define a zone around each ideal fiducial position, where each fiducial will lay regardless of the distortion type. For this reason, we divide every image into 60 sub-images, and we know that every one of these sub-images will always contain one fiducial. This area is about 27x27px, with the center on the ideal fiducial position. However, we expand the area to 32x32px to increase the security factor and to facilitate the max pooling operations on the CNN. With this approach, our CNN will focus on detecting fiducials of 3mm diameter within a grayscale image of 32 by 32 pixels. In a previous stage, we store the offset of each sub-image respecting the image. Therefore, we add to the inference its offset to calculate the fiducial prediction respecting the full image. We repeat this process sixty times per image, and then all of the fiducials are detected.

2.4 Data Collection and Labelling

We used a Ziehm Vision C-arm and took 140 images in each of the eight scans we carried out, giving us 1120 images in total. The images were taken using the automatic dose rate control of the C-Arm. In each scan, the C-arm rotates 130°, which covers its entire movement range. Every X-ray image was afterward cut into 60 sub-images, and each of the sub-images labeled with the X and Y coordinate of the fiducial position. These 67.200 sub-images were used for training a convolutional neural network (CNN). During the label processing, we used our traditional image processing method for fiducial detection, which gave us some inaccuracies that were corrected manually. Later we split the dataset in 60% for training, 20% for validating, and 20% for testing.
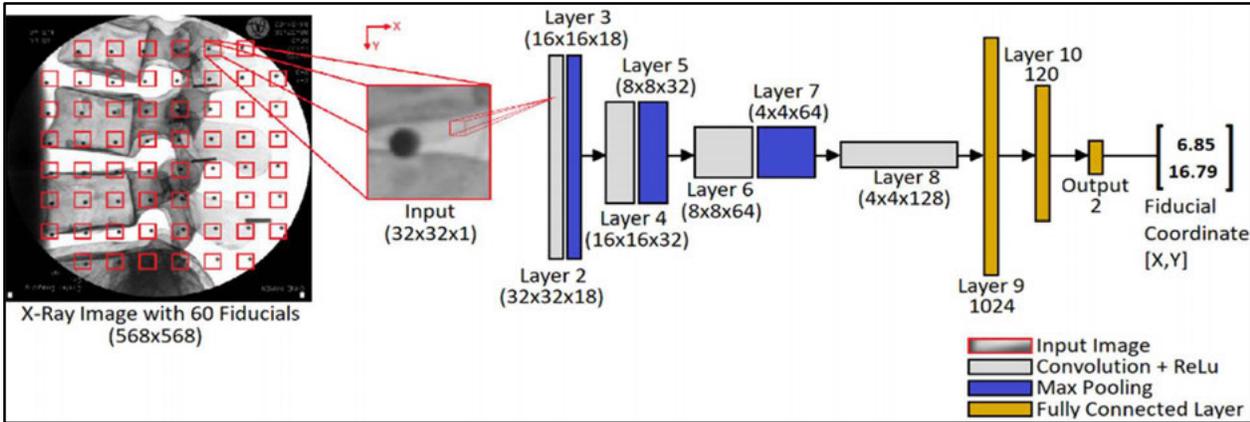
Figure 2: Fiducial detection based on Hugh transformation

2.5 Network Topology

After testing different topologies, we found a topology that give us results close to the ground truth. It consists of three layers of convolutions followed by max-polling and at the end, a fully connected neural network (FCNN) with three hidden layers, as can be seen in **Figure 2** and described in **Table 1.**

Table 1: CNN Topology.

Layer	Type	(Out)-Size	# Channels	Kernel Size	Activation
In	Image	32x32	1	-	-
2	Conv.	32x32	18	5x5	ReLu
3	Max Pool.	16x16	18	2x2	ReLu
4	Conv	16x16	32	5x5	ReLu
5	Max Pool.	8x8	32	2x2	ReLu
6	Conv	8x8	64	5x5	ReLu
7	Max Pool.	4x4	64	2x2	ReLu
8	Conv	4x4	128	5x5	ReLu
9	FCNN	1024	-	-	ReLu
10	FCNN	120	-	-	ReLu
Out	FCNN	2	-	-	-

We found that smaller topologies are not able to generalize the problem, given a high training error. Bigger CNNs showed us a clear overfitting, where the training set showed some promising results while the results over validation dataset were poor. The inference time on an image of 568x568 is around 1.2s.

3 Results

We use the testing dataset to compare the result of the accuracy given by our implemented traditional method using Hough Transform and the CNN. The results are shown in **Table 2**. The error of each fiducial is calculated as the geometrical distance between the predicted and the actual fiducial position. The prediction error is the average of individuals fiducial errors. We find that the detected fiducials using the traditional method, where we can change the Canny threshold value, gives an error of 0.8% when using a threshold of 100. Nevertheless, it gives a detection rate of 89.03%, which, in our case, is equivalent to detect 53 out of 60 fiducials. The detection rate increases to 97.71% when we set the Canny value to 20, but at the cost of increasing the prediction error to 2.99%. On the other hand, the CNN gives us a prediction error of 1.31%, which does not reach the accuracy of the Canny threshold in 100, but it does increase the fiducial detection rate from 97.71%, when using the Canny value of 20, to 100% by inferencing with our CNN. It is worth to clarify under the detection error context that the inference is made on sub-images of 32x32pixels.

Table 2: Accuracy results of tested methods for fiducial detection.

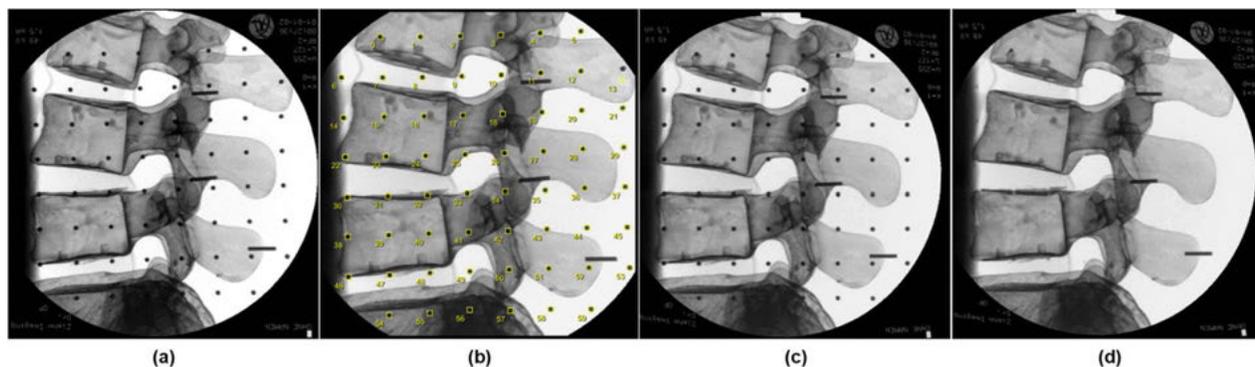| Method	Canny Threshold	Prediction Error [%]	Detection Rate [%]
Image Processing	20	2.99	97.71
	50	1.52	91.9
	80	1.02	89.89
	100	0.87	89.03
CNN	-	1.31	100

Figure 3. Original image with fiducials in (a). Result of inference with CNN in (b). Result of the undistortion in (c) and (d) shows the undistortion after inpainting.

That means the maximum inaccuracy due to the CNN inference is around 0.4192 pixels per coordinate, which guarantees sub-pixel. The entire workflow of having an X-ray image, detecting fiducials, undistorting, and inpainting the image can be seen in **Figure 3**.

4 Discussion and Conclusion

We present a robust approach for detecting fiducials in planar X-ray images for undistortion of conventional C-arm images. We move from a conventional image processing algorithm for detecting fiducials to a deep-learning approach. The implementation of the CNN shows that the detection error can be kept as low as possible compared with a robust image processing method such as 1.31% and 0.87% respectively, but with an increase of the detection rate moving from 97.71% with image processing to 100% using our CNN. For undistorting images, it is required to know the positions of the entire set of fiducials. Our previous implementation using the conventional method lacks this property, and it requires user intervention to carry out the process successfully. However, the CNN results show that detecting fiducials is robust and free of user assistance. The precision of the detection, on the other hand, will impact the quality of the undistortion. In our case, the detection is executed by dividing one X-ray into 60 sub-images of 32 by 32 pixels. One fiducial is detected in each of the sub-images. An inference with a position error of 1.31% in a sub-image of this size creates a maximum position error of ± 0.4192 pixels per coordinate, which gives subpixel precision.

The results indicate the feasibility of our approach for the undistortion of conventional C-arm images. In applications like 2D-3D registration, pre-operative CT-scans are matched with higher accuracy respecting intraoperative undistorted X-ray images. In our case, we found that registering with X-rays without undistortion leads to accuracy errors using our surgical robot of about 8mm compared to 2mm with undistorted images. Our robotic system's accuracy by using undistorted images is close to top-notch spine robot systems with image modalities, as described by Overley [6].

Author Statement

Conflict of interest: Authors state no conflict of interest. Informed consent: Informed consent has been obtained from all individuals included in this study. Ethical approval: The research related to human use complies with all the relevant national regulations, institutional policies and was performed in accordance with the tenets of the Helsinki Declaration, and has been approved by the authors' institutional review board or equivalent committee.

References

[1] Chintalapani G, Jain A, Taylor R (2007) Statistical characterization of C-arm distortion with application to intra-operative distortion correction. Medical Imaging 2007: Visualization and Image-Guided Procedures.

[2] Wolberg G (1994) Digital Image Warping. IEEE Computer Society Press Los Alami-tos, CA, USA.

[3] Livyatan H, Yaniv Z, Joskowicz L (2002) Robust Automatic C-Arm Calibration for Fluoroscopy-Based Navigation: A Practical Approach. MICCAI 2002.

[4] Schumann S, Dong X, Puls M, Nolte L, Zheng G (2012) Calibration of c-arm for or-thopedic interventions via statistical model-based distortion correction and robust phantom detection, Proc. IEEE 9th Int. Symp. Biomed. Imag.

[5] Telea A (2004) An Image Inpainting Technique Based on the Fast-Marching Meth-od, Journal of Graphics Tools.

[6] Samuel C. Overley, Samuel K. Cho, Ankit I. Mehta, Paul M. Arnold, Navigation and Robotics in Spinal Surgery: Where Are We Now?, Neurosurgery, Volume 80, Issue 3S, March 2017, Pages 86–99.