

Anneke Meyer¹, Alireza Mehrtash^{2,3}, Marko Rak¹, Daniel Schindele⁴, Martin Schostak⁴,
Clare Tempany², Tina Kapur², Purang Abolmaesumi³, Andriy Fedorov², Christian Hansen^{1,2}

Automatic High Resolution Segmentation of the Prostate from Multi-Planar MRI

Pre-print Version

¹ Computer-Assisted Surgery Group, Faculty of Computer Science, University of Magdeburg, Germany

² Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, USA

³ Robotics and Control Laboratory, University of British Columbia, Vancouver, Canada

⁴ Clinic of Urology and Pediatric Urology, University Hospital Magdeburg, Germany

AUTOMATIC HIGH RESOLUTION SEGMENTATION OF THE PROSTATE FROM MULTI-PLANAR MRI

Anneke Meyer¹, Alireza Mehrtash^{2,3}, Marko Rak¹, Daniel Schindele⁴, Martin Schostak⁴,
Clare Tempany², Tina Kapur², Purang Abolmaesumi³, Andriy Fedorov², Christian Hansen^{1,2}

¹ Computer-Assisted Surgery Group, Faculty of Computer Science, University of Magdeburg, Germany

² Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, USA

³ Robotics and Control Laboratory, University of British Columbia, Vancouver, Canada

⁴ Clinic of Urology and Pediatric Urology, University Hospital Magdeburg, Germany

ABSTRACT

Individualized and accurate segmentations of the prostate are essential for diagnosis as well as therapy planning in prostate cancer (PCa). Most of the previously proposed prostate segmentation approaches rely purely on axial MRI scans, which suffer from low out-of-plane resolution. We propose a method that makes use of sagittal and coronal MRI scans to improve the accuracy of segmentation. These scans are typically acquired as standard of care for PCa staging, but are generally ignored by the segmentation algorithms. Our method is based on a multi-stream 3D convolutional neural network for the automatic extraction of isotropic high resolution segmentations from MR images. We evaluated segmentation performance on an isotropic high resolution ground truth (n = 40 subjects). The results show that the use of multi-planar volumes for prostate segmentation leads to improved segmentation results not only for the whole prostate (92.1% Dice similarity coefficient), but also in apex and base regions.

Index Terms— MRI, prostate segmentation, deep convolutional neural networks, multi-stream architecture, multi-planar segmentation, 3D planning, 3D model generation

1. INTRODUCTION

With 180,890 new cases and more than 26,000 deaths in the US in 2016, PCa is the most common type of cancer and is the second most deadly cancer among men [1]. Due to its high soft-tissue contrast, MRI is often used for clinical purposes in PCa assessment and therapy. Segmentation of the prostate extracted from such data is commonly needed in research and clinical applications. For example, it is often used to enable MRI-TRUS (transrectal ultrasound) fusion guided biopsy [2], radiation dose planning in brachytherapy and external beam radiotherapy [3]. Shah *et al.* [4] utilized segmentation of the prostate to correlate MRI findings with the prostatectomy specimen.

A variety of MRI prostate segmentation algorithms have been proposed as summarized in [5]. In the last three years,

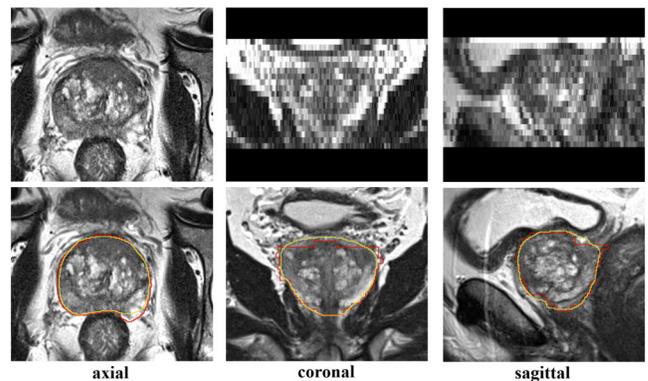


Fig. 1. Illustration of the effect of poor inter-plane resolution on segmentation. The top row illustrates the difference between in-plane resolution (left) and inter-plane resolution for an axial volume. No kind of interpolation was performed on this volume. In the bottom row, the manual segmentation, based only on the axial slices, is depicted by a red contour. A superior contour is visualized in yellow, whereby we fused the axial, coronal and sagittal segmentation by shape-based interpolation.

deep convolutional neural networks (CNN) were introduced into the context of medical segmentation. The fully convolutional network by Long *et al.* [6] enables segmentation in an end-to-end fashion. Furthermore, Ronneberger *et al.* [7] developed the U-Net which is frequently used for medical image segmentation. These architectures have been adapted for prostate segmentation. The fully connected CNN from [6] was fine-tuned for the purpose of prostate segmentation in [8]. Zhu *et al.* [9] applied a deeply supervised 2D U-Net to prostate segmentation. Recently, 3D CNNs were established for the prostate segmentation. Yu *et al.* [10] applied 3D CNNs with a combination of short and long residual connections to the prostate. So far, this work has performed best at the PROMISE12 prostate segmentation challenge [11].

All these methods rely purely on the axial T2-weighted MRI scan of the prostate. Prostate MRI is highly anisotropic

for the typical acquisition protocols, resulting in a factor of 6-10 difference between the out-of-plane and in-plane resolution. As can be seen in the top row of Figure 1, this leads to a strong partial volume effect. Consequently, the prostate gland boundary can be challenging to accurately localize in the axial image in the apex and base regions, where prostate tissue can not be clearly distinguished from surrounding structures like seminal vesicles or neurovascular bundles. Additionally, extracting surfaces from axial volumes can cause step artifacts in 3D prostate models due to the low inter-plane resolution. Shah *et al.* [4] addressed this problem by manually segmenting the prostate from three orthogonal MRI scans (axial, sagittal and coronal) and fusing them with shape-based interpolation in order to generate a high-resolution surface. Similarly, Cheng *et al.* [12] automatically segmented the prostate for the three views by holistically-nested edge detector networks. For each volume, a 2D network is trained separately and subsequently the three segmentation outcomes (with low out-of-plane resolution) are used for surface extraction with ball pivoting, followed by Poisson surface reconstruction to obtain a hole-free and smooth surface.

Since standard prostate imaging protocol includes multi-planar (axial, coronal, sagittal) T2-weighted images [13], we aimed to improve the fidelity of the prostate gland segmentation by utilizing all of the available scan directions to overcome the problems that arise from the low inter-slice resolution of the MRI scans. Our contributions in this work are:

- Instead of using three 2D networks as in [12], we propose obtaining a smooth segmentation from the orthogonal MRI scans with one single multi-stream volumetric network, which greatly simplifies training.
- The resulting segmentation has isotropic voxels with high resolution and enables an easy and fast extraction of the smooth surface for such applications as MRI-TRUS fusion and therapy planning.
- We created a high resolution ground truth that better reflects boundaries in critical apex and base regions. In addition to axial segmentations, sagittal and coronal manual segmentations are acquired and fused to an isotropic volume (yellow contour in Figure 1). The ground truth dataset will be released publicly for other researchers.¹

2. METHODS

In this work, we propose a method that simultaneously uses orthogonal scans for a high resolution segmentation of the prostate with 3D-CNN. First, we describe the architecture of the network, its training and the pre- and postprocessing of the data. Second, we give details on the high resolution ground truth generation.

¹<http://isgwww.cs.uni-magdeburg.de/cas/isbi2018>

2.1. Network Architecture

Our method is based on the 3D U-Net [14]. The input image is processed in an analysis (downsampling) and synthesis (upsampling) path, each containing four resolution steps. In the analysis path, the image is downsampled to increase the receptive field of the network. Each layer in this path consists of two $3 \times 3 \times 3$ volumetric convolution filters followed by a rectified linear unit (ReLU). A $2 \times 2 \times 2$ maxpooling operation with a stride of 2 along each dimension downsamples the volume to half its size. In the synthesis path, the convolutions are preceded by a deconvolution (transposed convolution) with a $2 \times 2 \times 2$ kernel and a stride of 2 that upsamples the features back to the input resolution. For each resolution layer in the analysis path, the filter size is doubled. Respectively, at each resolution step in the synthesis path, the filter size is halved. The last layer is a $1 \times 1 \times 1$ convolution filter followed by a sigmoid function that computes the binary label of the voxel. By means of skip connections, high resolution information is passed from the analysis path to the same level of the synthesis path.

The described network is modified for the multi-planar segmentation with a multistream-architecture as illustrated in Figure 2. Our network has three analysis streams, where the streams process axial, coronal and sagittal scans individually. The streams are then concatenated at the lowest resolution level. Skip connections are established from each stream to the respective level of the synthesis path.

2.2. Preprocessing

The orthogonal scans of the prostate are acquired in the same frame of reference one after another, and in most cases do not require registration. When misalignment was observed due to motion, we applied manual rigid registration. As part of preprocessing, the three scans are resampled by linear interpolation into a common coordinate system with a resolution of $0.5 \times 0.5 \times 0.5$ mm. This resembles the best in-plane resolution of the single scans, which usually have slice thickness of ≈ 3 mm. Next, the per-plane images are cropped to the volume corresponding to their intersection. Prior to normalizing the intensity of the images to an interval of $[0,1]$, the intensities are cropped to the 1st and 99th percentiles.

2.3. Training of the Network

We selected the negative Dice similarity coefficient (DSC) loss function as the objective function of the network:

$$loss = -\frac{2 \sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2}$$

with N being the total number of voxels and p_i the predicted voxels and g_i the ground truth binary voxels. Adam Optimizer [15] was used with default parameters except for the learning rate. The network was trained on a NVIDIA Titan Xp

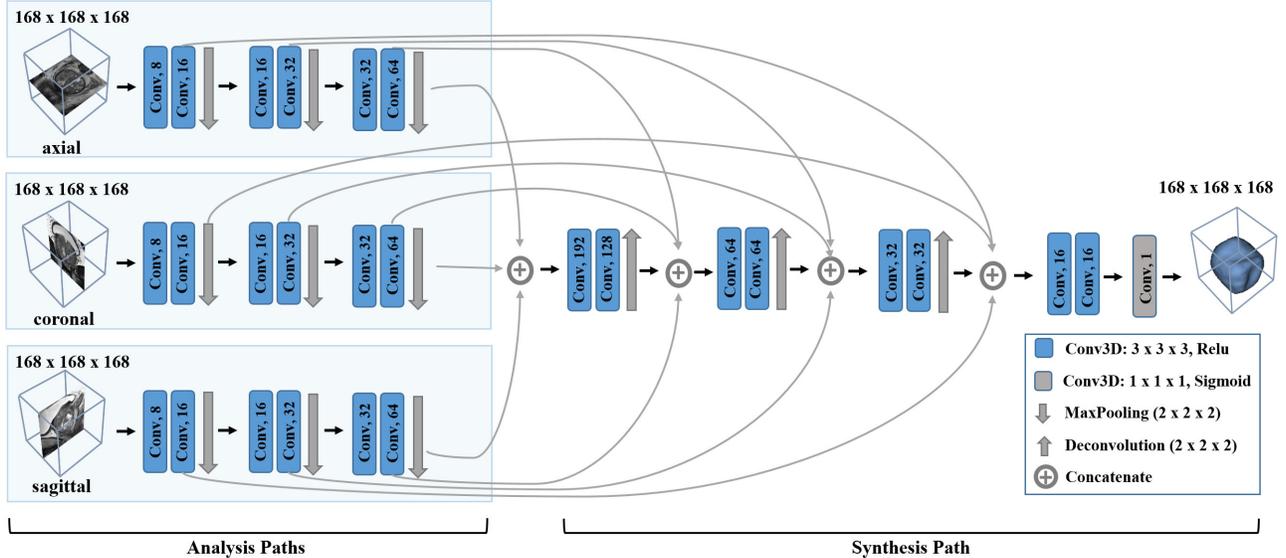


Fig. 2. Architecture of the proposed multi-stream 3D U-Net. Numbers inside the layers correspond to number of feature maps. The full volumes ($168 \times 168 \times 168$) are used as network input. For reasons of simplicity the original U-shape is not maintained.

for 1000 epochs, with an early stop mechanism if the validation loss does not change by $\delta = 0.001$ for 100 iterations. The batch size was set to 1 due to graphical memory capacity. In order to increase the number of training images, data augmentation is introduced by left-right-flipping of the input volumes. Top-bottom and front-back flips were not considered since the orientation of the prostate bears valuable clues for segmentation. As postprocessing of the prediction, a connected components analysis was applied to the segmentation mask. The isosurface is extracted with the marching cubes algorithm. The total computation time for preprocessing, prediction and surface extraction of an unseen image is less than 10 seconds.

For the evaluation we set aside 15 images, each containing axial, coronal and sagittal scans for the test. We used 5-fold cross validation on 25 images for selecting the best hyperparameters such as learning rate and number of epochs and re-trained the network again on the whole training set with these hyperparameters. An unbiased performance estimate is given by evaluating this model on the set-aside test data. Since we were interested in the effect of improvement by use of the additional orthogonal planes, we also evaluated the simple single-stream 3D U-Net on the axial volumes. The procedure was the same as for the multi-stream multi-planar network.

2.4. Dataset and Ground Truth

As the PROMISE12 dataset provides only axial volumes, we tested the presented approach on the SPIE-AAPM-NCI PROSTATEx challenge dataset [16]. We selected 40 T2-weighted MR scans which contained axial, coronal and sagittal volumes. The alignment was checked visually us-

ing 3D Slicer [17] (<http://slicer.org>). For three cases, we had to register the volumes manually. The segmentations were obtained manually for each view on the original resolution of the image by a medical student and later corrected by an expert urologist. Then, the signed distance transformation of the three segmentations is computed. Subsequently, the anisotropic distance volumes are transformed into an isotropic high resolution representation with linear interpolation. By averaging the distances and thresholding them with zero, we obtain the fused segmentation (shape-based interpolation as in [18]). The resulting segmentations were manually verified and corrected further if necessary.

3. RESULTS

We evaluated our approach with the following metrics that were also used in the PROMISE12 challenge: DSC, absolute relative volume difference (aRVD), 95% - Hausdorff-Distance (95-HD) and the average of shortest distances between boundary points of both volumes (ABD). These metrics represent both boundary-based and volume-based evaluation and are calculated in 3D. A detailed definition of these metrics can be found in [11]. Evaluation was performed for the whole volume as well as for apex, base and mid-gland regions, each of which contains 1/3 of the prostate. The quantitative results of the experiments are presented in Table 1 and qualitative results can be found in Figure 3.

A mean DSC of 92.1% was obtained for the whole prostate by use of the proposed multi-planar segmentation with the multi-stream network, whereas the original architecture based on only the axial volume obtained a DSC of 90.7% (p-value of 0.0085 in a one-sided paired t-test). While aRVD

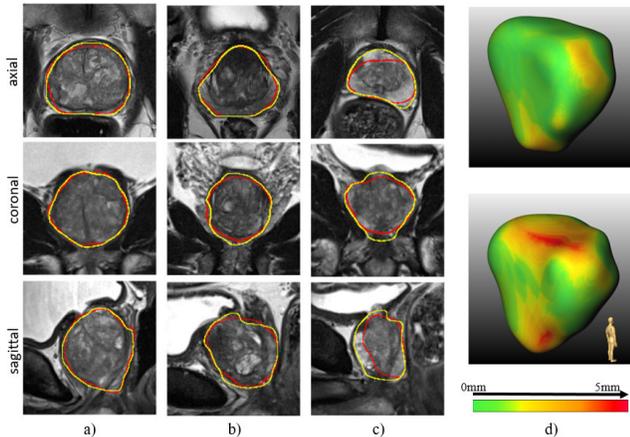


Fig. 3. Qualitative results of a) the best (DSC=94.8%), b) an average (DSC=92.5%) and c) the worst segmentation (DSC=85.4%). The ground truth is depicted by the yellow contour, the prediction of the multi-stream network is shown in red. d) 3D visualization of the surface distance from multi-planar segmentation (top) and axial-only segmentation (bottom) of another volume.

did not change by the multistream network, 95-HD improved by more than 22% to 3.44mm and the ABD improved by 16%. Hence, we reduced leaking boundaries by use of the multi-plane approach and produced a result that is closer to the ground truth boundary. Improvement was also consistent considering the apex and base region. Only the mid-gland region could be equally well segmented by both techniques. One can assume that this is because of the good contrast the prostate boundaries have in the mid-gland region on axial slices. Comparing our results to Cheng *et al.* [12] who also use orthogonal volumes (DSC = 88.6%), we obtain higher accuracy. However, this has to be considered with caution, because we used different datasets.

In a second experiment, we evaluated the effect of skip-connections that pass high resolution information from the analysis path to the synthesis path of the network, by repeating the experiment with removed skip connections. For multi-planar segmentation, performance was dropped by removing

Table 1. Evaluation results for axial and multi-planar networks (distances in mm).

		DSC	aRVD	95-HD	ABD
Multi-Planar	Apex	85.1%	23.0%	4.20	1.33
	Mid-gland	95.5%	5.3%	2.30	0.79
	Base	90.2%	9.9%	3.14	0.99
	Whole	92.1%	8.3%	3.44	1.00
Axial Only	Apex	82.8%	25.2%	5.26	1.56
	Mid-gland	95.2%	4.1%	2.87	0.85
	Base	87.5%	12.3%	4.16	1.30
	Whole	90.7%	8.3%	4.43	1.20

the skip connections: the DSC decreased to 91.6% and the 95%-HD distance increased to 3.71mm. Thus, the transfer of high resolution information to the synthesis path of the network does improve segmentation performance.

4. CONCLUSION

The aim of this work is to provide a simple yet powerful framework for the accurate segmentation and surface extraction of the prostate by use of axial, sagittal and coronal volumes. We proposed a 3D multi-stream CNN that simultaneously computes an isotropic high resolution segmentation from the three orthogonal volumes that can be easily used for surface extraction. We evaluated the segmentations against an isotropic manual ground truth that takes all three orthogonal volumes into account. Results show that an accurate segmentation can be obtained with this proposed method, improving the outcome compared to a single stream network using only axial volumes of the prostate. The created ground truth will be made available online for other researchers.

To the best of our knowledge, the only approach that utilizes the multi-planar volumes in prostate segmentation with deep learning is the one proposed by Cheng *et al.* [12]. Our method offers several advantages over this approach. First, by using a single network, as compared to the three networks, we can simplify the training process. Second, computationally expensive surface extraction methods, such as ball pivoting and Poisson surface reconstruction, are not required: the surface can be easily generated from the high resolution segmentation. Consequently, our approach leads to faster prostate model generation while obtaining accurate results.

Because orthogonal volumes are usually included in the PCa MRI protocols, this work does not require any modification to the imaging procedures, and should encourage efforts to incorporate multi-planar volumes into the segmentation process for improved precision of prostate model generation for individualized therapy planning. Future work should include the enhancement of the networks architecture by use of residual connections. Additionally, automatic registration should be employed for cases where the volumes are not well aligned. Finally, a more clinical-relevant evaluation should be conducted, incorporating the effect of segmentation on the registration accuracy for targeted fusion biopsy procedures.

Acknowledgements This work has been funded by the EU and the federal state of Saxony-Anhalt, Germany under grant number FuE 74/16 and ZS/2016/04/78123 as part of the initiative 'Sachsen-Anhalt WISSENSCHAFT Schwerpunkte', and grants from NIH (U24 CA180918 and P41 EB015898). The Titan Xp used for this research was donated by the NVIDIA Corporation. Data used in this research were obtained from The Cancer Imaging Archive (TCIA) sponsored by the SPIE, NCI/NIH, AAPM, and Radboud University.

5. REFERENCES

- [1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2016," *CA Cancer J Clin*, vol. 66, no. 1, pp. 7–30, 2016.
- [2] A. Fedorov, S. Khallaghi, A. C. Sánchez, A. Lasso, S. Fels, K. Tuncali, E. S. Neubauer, T. Kapur, C. Zhang, W. Wells, P. L. Nguyen, P. Abolmaesumi, and C. Tempany, "Open-source image registration for MRI-TRUS fusion-guided prostate interventions," *Int J Comput Assist Radiol Surg*, vol. 10, no. 6, pp. 925–934, 2015.
- [3] M. A. Schmidt and G. S. Payne, "Radiotherapy planning using MRI," *Phys Med Biol*, vol. 60, no. 22, pp. R323–61, 2015.
- [4] V. Shah, T. Pohida, B. Turkbey, H. Mani, M. Merino, P. A. Pinto, P. Choyke, and M. Bernardo, "A method for correlating in vivo prostate magnetic resonance imaging and histopathology using individualized magnetic resonance-based molds," *Rev Sci Instrum*, vol. 80, no. 10, pp. 104301, 2009.
- [5] S. Ghose, A. Oliver, R. Martí, X. Lladó, J. C. Vilanova, J. Freixenet, J. Mitra, D. Sidibé, and F. Meriaudeau, "A survey of prostate segmentation methodologies in ultrasound, magnetic resonance and computed tomography images," *Comput Methods Programs Biomed*, vol. 108, no. 1, pp. 262–287, 2012.
- [6] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.
- [7] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Med Image Comput Comput Assist Interv*, 2015, pp. 234–241.
- [8] Z. Tian, L. Liu, and B. Fei, "Deep convolutional neural network for prostate MR segmentation," in *Medical Imaging 2017: Image-Guided Procedures, Robotic Interv., and Modeling*, 2017, vol. 10135, p. 101351L.
- [9] Q. Zhu, B. Du, B. Turkbey, P. L. Choyke, and P. Yan, "Deeply-supervised CNN for prostate segmentation," in *International Joint Conference on Neural Networks (IJCNN)*, 2017, pp. 178–184.
- [10] L. Yu, X. Yang, H. Chen, J. Qin, and P.-A. Heng, "Volumetric ConvNets with mixed residual connections for automated prostate segmentation from 3D MR images," in *AAAI Conference on Artificial Intelligence*, 2017, pp. 66–72.
- [11] G. Litjens, R. Toth, W. van de Ven, C. Hoeks, S. Kerckstra, B. van Ginneken, G. Vincent, G. Guillard, N. Birbeck, J. Zhang, et al., "Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge," *Med Image Anal*, vol. 18, no. 2, pp. 359–373, 2014.
- [12] R. Cheng, N. Lay, F. Mertan, B. Turkbey, H. R. Roth, L. Lu, W. Gandler, E. S. McCreedy, T. Pohida, P. Choyke, M. J. McAuliffe, and R. M. Summers, "Deep learning with orthogonal volumetric HED segmentation and 3D surface reconstruction model of prostate MRI," in *IEEE 14th International Symposium on Biomedical Imaging ISBI*, 2017, pp. 749–753.
- [13] J. V. Hegde, R. V. Mulkern, L. P. Panych, F. M. Fennessy, A. Fedorov, S. E. Maier, and C. M. Tempany, "Multiparametric MRI of prostate cancer: An update on state-of-the-art techniques and their performance in detecting and localizing prostate cancer," *J Magn Reson Imaging*, vol. 37, no. 5, pp. 1035–1054, 2013.
- [14] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," in *Med Image Comput Comput Assist Interv*, 2016, pp. 424–432.
- [15] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [16] G. Litjens, O. Debats, J. Barentsz, N. Karssemeijer, and H. Huisman, "Computer-aided detection of prostate cancer in MRI," *IEEE Trans Med Imaging*, vol. 33, no. 5, pp. 1083–1092, 2014.
- [17] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, et al., "3D Slicer as an image computing platform for the quantitative imaging network," *Magnetic resonance imaging*, vol. 30, no. 9, pp. 1323–1341, 2012.
- [18] G. T. Herman, J. Zheng, and C. A. Bucholtz, "Shape-based interpolation," *IEEE Computer Graphics and Applications*, vol. 12, no. 3, pp. 69–79, 1992.