Comparison of Texture Synthesis Methods for Content Generation in Ultrasound Simulation for Training

Oliver Mattausch[1], Elizabeth Ren[1], Michael Bajka[2], Kenneth Vanhoey[1], Orcun Goksel[1]
[1]Computer Vision Lab, ETH Zurich, Switzerland,
[2]Departments of Obstetrics and Gynecology, University Hospital Zurich, Zurich, Switzerland

Ultrasound Simulation

- Ultrasound: radiation-free, low-cost, real-time
- Training difficult, e.g., of rare pathologies
- Few volunteers for transvaginal, transrectal, biopsy
- Huge potential for interactive virtual-reality based ultrasound training simulator
- Goal: Training simulation that convinces experts
- In this work, we evaluate simulation of US speckle using methods of texture synthesis

Ultrasound Speckle

- Typical noise patterns of tissue under ultrasound
- Essential for realistic ultrasound simulation
- Ultrasound scattered by microscopic entities in tissue
- Speckle formation: Convolve point-spread function (PSF) with distribution of point scatterers [1, 2]
- Exhibits properties of Markov Random Fields (MRF)

By-Example Texture Synthesis

- Generation of larger output texture from input image
- MRF assumption [3]: Input is local and stationary
- Locality: pixel colors depend on local neighbors only
- Stationarity: two pixels with similar surrounding have similar color
- Parametric vs. non-parametric texture synthesis

Our Contribution

We conducted a perceptual study of example-based synthesis of locally homogeneous US image regions among 19 ultrasound experts, comparing 4 representative texture synthesis methods in 3 questionaires. This study examines the following questions:

- Can texture synthesis generate plausible US speckle?
- Which is the texture synthesis method of choice?
- Can human judgement be predicted from data using machine learning?

Conclusions

- Ratings show that US speckle can indeed be emulated by texture synthesis
- **PIXEL** is the superior choice for creating plausible ultrasound speckle
- Correlation coefficients indicate feasibility of predictor from texture features
- In future work, predictor could be used as metric for judging new methods

References

[1] J C Bamber and R J Dickinson, "Ultrasonic b-scanning: a computer simulation," *Physics in Medicine and Biology*, vol. 25, no. 3, pp. 463, 1980.

[2] J. Meunier, M. Bertrand, and G. Mailloux, "A model for dynamic texture analysis in two-dimensional echocardiograms of the myocardium," *SPIE*, vol. 0768, pp. 193–200, 1987.

[3] Li-Yi Wei, Sylvain Lefebvre, Vivek Kwatra, and Greg Turk, "State of the art in example-based texture synthesis," in *Eurographics State of the Art Reports*, 2009.

[4] Alexei A. Efros and Thomas K. Leung, "Texture synthesis by non-parametric sampling," in *Proceedings of Conference on Computer Vision*. 1999, ICCV '99, pp. 1033–, IEEE Computer Society.

[5] Javier Portilla and Eero P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *Journal Computer Vision*, vol. 40, no. 1, pp. 49–70, Oct. 2000.

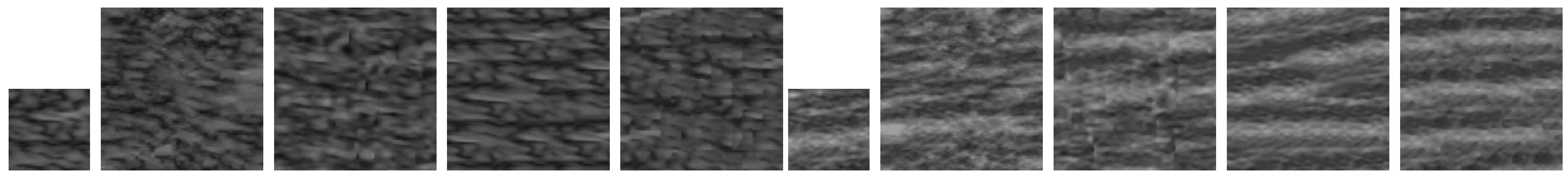
[6] Vivek Kwatra, Irfan Essa, Aaron Bobick, and Nipun Kwatra, "Texture optimization for example-based synthesis," *ACM Transactions on Graphics, SIGGRAPH 2005*, August 2005.

[7] Alexei A. Efros and William T. Freeman, "Image quilting for texture synthesis and transfer," in *Proceedings of Conference on Computer Graphics and Interactive Techniques*. 2001, SIGGRAPH '01, pp. 341–346, ACM.

Acknowledgments

Department of Obstetrics, Dir. Prof. R. Zimmermann, Department of Gynecology, Dir. Prof. D. Fink, Swiss Commission of Technology and Innovation (CTI), ERC grant "VarCity" (#273940)

Methods

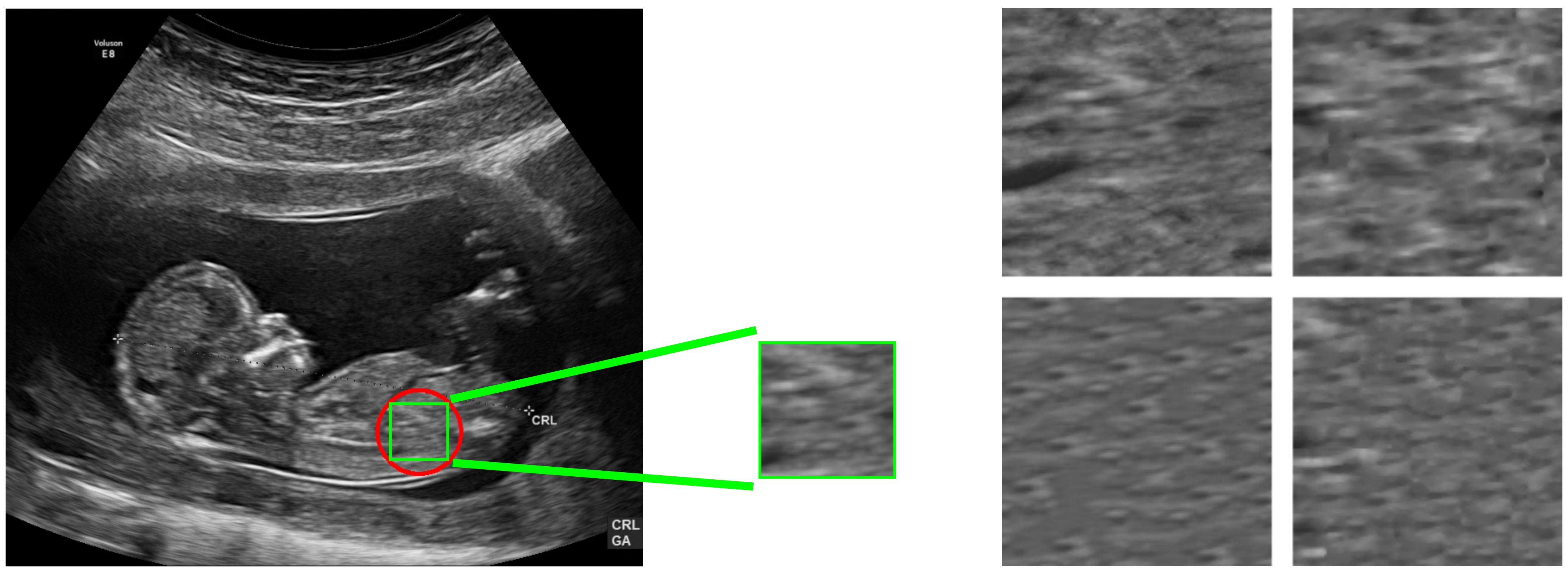


a) Tissue-mimicking phantom b) In-vivo muscle tissue

Figure 1: Texture synthesis applied to US imagery: a homogeneous US patch (smallest images) forms the exemplar from which a larger texture is synthesized. From left to right, the four techniques we considered: **PIXEL**, **PARAM**, **OPTIM** and **PATCH**.

Our study compares 4 methods, 3 non-parametric (**PIXEL**, **OPTIM**, **PATCH**), 1 parametric (**PARAM**) (Fig. 1):

- **PIXEL** uses pixel-wise growing based on neighborhood similarity in the input image [4]
- **PARAM** learns set of statistical parameters to generate new textures [5]
- **OPTIM** minimizes *global* energy based on similarity between output and best-matching neighborhoods in input [6]
- **PATCH** sequentially copies whole neighborhoods that slightly overlap in the output [7]



a) Input image b) Sample c) Synthesized patches

Figure 2: From a US image (a), we identify a homogeneous region (red circle) and crop a square sample (b), which is used as input for the different texture synthesis techniques (c). Each sample is used to synthesize four different images, one per technique.

We selected 11 US images of different anatomical structures, including liver, uterus, breast, muscle, fetus, uterus phantom. From a homogeneous region, samples were cropped and used as input to the texture synthesis methods (Fig. 2), producing 4×11=44 synthesized images. The participants were asked to compare the region in the red circle (Fig. 2(a)) with the synthesized textures (Fig. 2(c)), judging how well the texture mimicks the target texture.

(6/66): To continue, select the patch that best mimicks the tissue inside the red circle. Please don't use the back button of your browser.

(9/44): On a scale from '--' (implausible) to '++' (plausible), please rate how well a patch mimicks the tissue inside the red circle. Click 'Next' to submit and continue. Please don't use the back button of the browser.

(3/11): Please rank the patches according to how good they mimick the tissue inside the red circle ('1' is best, '4' is worst). Click 'Next' to submit and continue. Please don't use the back button of your browser.

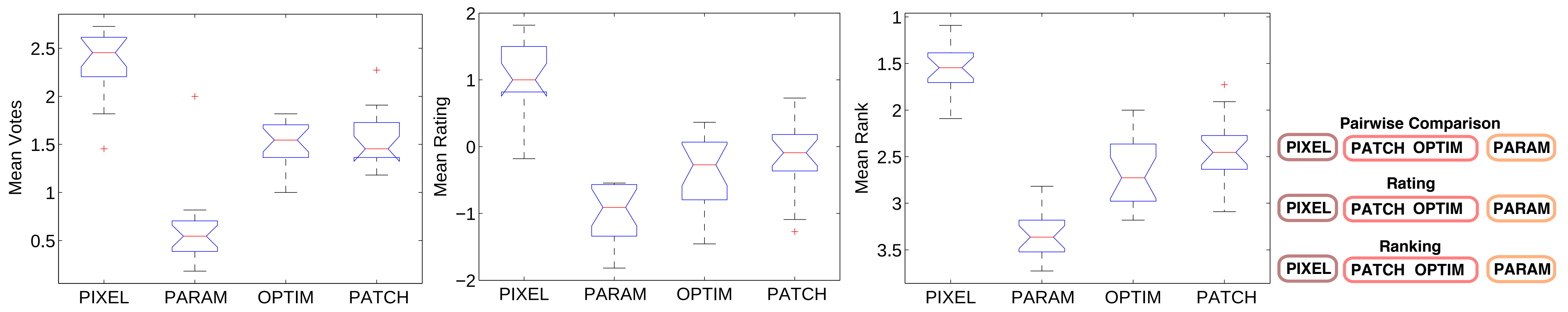
a) Paired Comparison b) Rating c) Ranking

Figure 3: A question page of each of our three ultrasound questionnaires.

We asked participants three different types of *questionnaires*, each answering a slightly different objective (Fig. 3):

- **Paired Comparison** of two patches (which better mimicks source?) (66 questions)
- **Rating** of synthesized patch using 5-point Likert scale (44 questions)
- **Ranking** of synthesized patches from all 4 methods, from best to worst fit (11 questions)

Results



a) Paired Comparison b) Rating c) Ranking d) Posthoc analysis

Figure 4: (a)-(c): Results of the 3 questionnaires. Note that for Ranking, a lower rank is better. Statistical significance analysis (d).

From 19 ultrasound experts, 9 participants were specialized in Obstetrics/Gynecology, one radiologist, and 9 technical experts. The plots in Fig. 4 show that **PIXEL** performs better in all questionaires (a)-(c) and the difference is shown to be significant in the posthoc analysis (Tukey's HSD) (d). Attributes like experience (indicated by the number of ultrasound scans taken in a lifetime) did not lead to statistically different results.

Feature-based Predictor

To learn the realism of speckle appearances in generated US texture patches from results of the **Rating** questionaire, a regression forest of 50 trees was trained on a selection of image features, using least-squares boosting, with leave-one-patch-out experiments.

Satisfactory maximum correlation coefficients of **0.64 Pearson** and **0.66 Spearman** was achieved by using classic features like **LBP distances**, **texton distances**, and **Bhattacharyya histogram distances** for training.