A conversation with CHCCS 2022 achievement award winner Prof. Hao (Richard) Zhang

Hao (Richard) Zhang
Simon Fraser University (SFU)

Abstract

The 2022 CHCCS Achievement Award from the Canadian Human-Computer Communications Society is presented to Hao (Richard) Zhang from Simon Fraser University (SFU) for his numerous high-impact contributions to computer graphics. His diverse research addresses fundamental problems in geometric modeling, shape analysis, geometric deep learning, and computational design and fabrication. CHCCS invites a publication by the award winner to be included in the proceedings, and this year we continue the tradition of an interview format rather than a formal paper. This permits a casual discussion of the research areas, insights, and contributions of the award winner. What follows is an edited transcript of that conversation between Richard and Danny Cohen-Or, from Tel Aviv University, that took place on February 29, 2022, via Zoom. The transcript was then expanded to include more content.

The Interview

Danny: Richard, congratulations on being recognized as the 2022 CHCCS Achievement Award winner. We have known each other for many years. When I first met you, I thought you were more of a mathematician who was trying to make his way into computer graphics. Is that a fair way to describe you at the start? Where are you now? How has your research evolved over the years?

Richard: Thank you, Danny! These are all big questions to answer (chuckle)! Let me first say where I am now in terms of my core research. Most of my recent works are on geometric deep learning, including learning neural implicit fields that started with IM-Net [1], deep learning of 3D structures that started with GRASS [10], and unsupervised domain transform following DualGAN [22]. I think we will come back to these later. Looking back at my research journey, the start certainly had an interesting twist.

I guess the surprising thing is that graphics was *not* my first choice when I was applying for PhD. My master's thesis from Waterloo was on formal methods, where a big part of the work was a proof of equivalence between two behavioral properties of asynchronous circuits. I was more into math indeed, and my top choice for doing a PhD was with either David Kirkpatrick or Jack Snoeyink at UBC, working on computational geometry.

Danny: Yes, computational geometry would be quite a bit more theoretical. Then how did you end up going to Toronto?

Richard: Here is what I can recall. I was told that Jack Snoeyink was going to leave for UNC and David Kirkpatrick would become the Dean and be very busy. I was then recruited to the University of Toronto to do theory, working on lower bounds, specifically.

Danny: You were getting far from computer graphics!

Richard: Absolutely. Graphics was not even on my radar at the time, partly because I only got a 68 in my undergraduate graphics class at Waterloo (lol). To continue the story, I quickly found that lower bounds was not my cup of tea, so I decided to switch topic after one term. However, I could not find anyone doing computational geometry at Toronto. So I ended up "settling" for computer graphics. I was happy and lucky that Eugene (Fiume) took me in and was so supportive. In the early years, I was indeed still holding a more mathematical mindset while trying to "work my way" into graphics. So your first impression was spot on.

Danny: Your first well-known works were on shape analysis and spectral geometry processing. In fact, you introduced spectral clustering for mesh segmentation into graphics [11]. I think we had two mesh segmentation papers in the same session, in Korea!

Richard: Thanks for the credit, Danny. With my predisposition to computational geometry, I naturally got interested in geometric modeling and shape analysis, especially 3D segmentation at first.

Yeah, I remember Korea, and that boat party. That was where I first met Arik Shamir, in person. I do not remember whether you and I met there. My work on spectral methods was inspired by Gabriel Taubin's SIGGRAPH 1995 paper on a signal processing approach to mesh smoothing [18]. However, instead of filtering and signal processing, I was more interested in directly applying eigenvalues and eigenvectors of mesh Laplacians to solve various problems such as shape segmentation and correspondence.

Danny: What was your most memorable result from all these works that you did on spectral analysis?

Richard: Hmmm, that is a fun question (lol)! Let me think. I guess "most memorable" does not necessarily mean the best results. Amazingly, what comes first to my mind is more of a "negative result." I still remember all the visuals about this vividly.

Danny: "Negative result"?

Richard: Yes, an issue that we discovered when working on spectral correspondence [8]. I called it "eigenvector switching", where if we use the magnitudes of eigenvalues, which were considered as representing frequencies, to sort the Laplacian eigenvectors of two shapes, these eigenvectors do not always match up; they can switch orders. The cause of this is stretching of shape parts.

Danny: Did you find a solution?

Richard: Partially. Remember our "part-aware metric" paper [12]? The key was to reason about shape parts by looking inside, to analyze visibility from within the shape interiors. Since the visibility only changes dramatically when a viewpoint crosses a part boundary, it is not sensitive when the viewpoint is within a part, no matter how large or small it is; it is insensitive to part stretching. It was one of my favourite works as it connects visibility, a prominent concept from computational geometry, to shape understanding.

Danny: I also recall the use of "diffusion maps" soon after that.

Richard: Yes, that work was the conduit to the next phase of my research, *co-analysis*, and the transition started around 2009 to 2010

Graphics Interface Conference 2022
16-19 May

and had a lot to do with your influence. Our collaboration started when you did your sabbatical at UBC in 2008. We first worked on point cloud processing, focusing on incomplete scans. I published my first SIGGRAPH paper on ROSA, rational symmetry axis to compute curve skeletons of incomplete point clouds. And then we worked extensively on co-analysis. We applied diffusion maps and spectral clustering in feature space for 3D co-segmentation [16].

Danny: Unsupervised learning! I loved it. We started to work more on machine learning (ML) for shape analysis then.

Richard: Indeed. I got into ML, mostly on unsupervised or weakly supervised learning related to clustering problems, in early 2010's, and never looked back. However, it took me some self-convincing to get on board with the "deep learning train" (lol). Was it around 2016 that you told me, "Richard, from now on, all we do will be about deep learning." Something like that. But I worried that I could not contribute since I was just a novice in that field.

Danny: But you are "converted" now. What changed it?

Richard: You "dragged" me half way, but I still needed to learn the field. So I decided to do my sabbatical in Leo Guibas' lab at Stanford in 2016, so that I could immerse myself in the "hub" of geometric deep learning. It was definitely the right choice.

Danny: That was the year you came up with DualGAN [22]?

Richard: Right, but it was not a product of my sabbatical. During that time, I travelled back home from Palo Alto every 2-3 weeks to spend weekends with family. During one of those trips, a visiting PhD of mine came to me with an idea that was inspired by *dual learning* [3] from natural language processing (NLP). I was fascinated by the possibility of training English-to-French translation without paired sentences. With GANs into the mix, it became DualGAN [22], and I take credits for giving the network its name (chuckle)! In retrospect, NLP has generated some great ideas that made strong impact on visual computing. Seq2seq, and of course, Transformers, but dual learning is not that well-known.

In any case, we submitted DualGAN to ICCV 2017 and I told you about it right away, and you loved it! Then you messaged me one day, telling me that you saw a Facebook post about a piece of work that had exactly the same idea, using a *cycle consistency* loss in a GAN for unpaired image-to-image translation. That work was CycleGAN [23], first posted by Jun-Yan Zhu on Facebook on March 31, 2017. This post prompted us to upload DualGAN to arXiv on April 8; it was my first ever pre-publication post on arXiv!

After both papers were accepted to ICCV, I wrote to Alyosha Efros. He was very courteous and we cross-cited each other's works. Now I always tell people that CycleGAN and DualGAN have a combined citation count of more than 13K (wink)!

Danny: I love this story! This work is truly amazing, revolutionary! Even more amazing is that you were working on something else at the same time that was also significant.

Richard: I guess you are right (chuckle). The problem I brought to Leo at Stanford stemmed from our work on "Fit and Diverse" [20]. The mutation and crossover we defined for shape evolution took inspiration from genetic programming which operates on DNA codes. I had always wondered what could be a true DNA code for shapes, that is *generative* and *editable*. Yes, there was the "Shape DNA" paper [15], which used Laplacian spectrum as a shape signature.

Leo and I named our project "Shape DNA" anyways. Collaborating with Sid Chaudhuri and others, we developed a *recursive autoencoder* to learn symmetry hierarchies [19], an organization of shape parts. This symmetry hierarchy paper was one of my favourites. It organizes the semantic parts of a man-made shape, like a chair or a handcart, using a tree that is built on symmetry and connectivity relations between the parts. By the way, Sid was the first to bring up recursive neural networks, that he learned from an NLP work (again), by Richard Socher [17]. Our innovation was the recursive autoencoder that encodes and decodes symmetry hierarchies.

So from the shape parts, the encoder network forms a tree and finally produces a root code, which would be our best attempt at coming up with a shape DNA. The really nice thing is that the codes can be sampled again and used to generate new symmetry hierarchies, using the decoder. That is GRASS, generative recursive autoencoder of shape structures [10]. However, until this day, I still cannot figure out how to interpret the root codes, e.g., to make them editable, like real DNA codes. There is still more to do.

Danny: GRASS was probably your first major work on geometric deep learning, a topic that is really flourishing. I think by now, you are best known for your recent works on neural implicits.

Richard: It was around early 2018 that I started to give talks on "Computer Graphics in the Age of AI and Big Data", a follow-up of my earlier series of talks on "Why is Computer Graphics Hard?" A core talking point was several 3D challenges that I identified: the data challenge, the representation challenge, and the functionality challenge. Specifically, the representation challenge is due to a lack of universally accepted representation for 3D shapes, as we have point clouds, meshes, voxels, structural graphs, multi-view images, and more! I kept asking what would be the "best" neural representation for 3D shapes. What motivated me to look into implicit functions was merely the fact that I had not seen neural networks, in particular convolutional neural networks (CNNs), designed for them at the time. So I gave this problem to a master student.

The student came back with as simple of a network architecture as it can be, few layers of multi-level perceptrons (MLPs), which takes a 3D point coordinate and a shape code and predicts an occupancy value. That was IM-Net [1], which follows exactly how implicit shapes are represented. My key observation was that shapes are defined by their boundaries, i.e., inside vs. outside; they are not about voxel intensity distributions in these pre-set boxes, that the CNNs use. IM-Net learns shape boundaries as continuous functions.

Danny: Of course, in the same year, there were DeepSDF [9] and Occupancy Networks [13], all at CVPR 2019!

Richard: Yeah, it was both amazing and scary, that after CycleGAN and DualGAN appearing in ICCV 2017, I was involved with yet another paper (IM-Net) that was built on exactly the same idea as another two papers, at the same venue! I believe both DualGAN and IM-Net are high-impact works.

Danny: Well, to me, these implicit networks are like basic NeRFs [14], no? They appeared one year earlier than NeRFs! Of course, we all know now how influential NeRF has been recently.

Richard: Right on! IM-Net inputs a 3D point coordinate and produces a scalar, such as a distance or an in/outside flag. NeRF takes a 3D point and two directions and produces a color and density value. Both networks are built with MLPs! A recent article by Frank Dallert [2] called the three neural implicit papers (IM-Net, DeepSDF, OccNet) as the "immediate precursors to NeRF".

Danny: I want to say again that IM-Net, GRASS, and DualGAN are all ground-breaking works! Aside from these, what other line of works from you are as important to you?

Richard: I will have to bring up our series of works from 2015 and on, all related to understanding and predicting functionality of 3D objects [7, 5, 4, 6, 21]. After having worked on shape analysis for so long, I came to realize that the ultimate goal of 3D shape understanding is at the functional level. Think about it, for a robot to do the things that humans usually do, it is ultimately about acting on and interacting with 3D environments, with the objects therein and other humans or robots. For generative modelling, why would we want to generate something in 3D? I think the ultimate goal is for the 3D object or 3D scene to perform its intended functions. This is a really important topic, that I will continue to work on.

Danny: That is great. Well, I think this is probably a good spot to end. Thanks so much for answering all of these questions.

Richard: My pleasure! Of course, it goes without saying that none of the works mentioned would have been possible without the very talented students and dedicated collaborators that I had the privilege to work with. I will thank them all in my keynote talk.

Danny: And congratulations again on the award.

Richard: Thanks Danny!

References

[1] Z. Chen and H. Zhang. Learning implicit fields for generative shape modeling. In *CVPR*, 2019.

[2] F. Dellaert. Technical perspective: Neural radiance fields explode on the scene. *Communications of the ACM*, 65(1):98, 2022.

[3] D. He, Y. Xia, T. Qin, L. Wang, N. Yu, T.-Y. Liu, and W.-Y. Ma. In *Advances in Neural Information Processing Systems 29 (NIPS)*, 2016.

[4] R. Hu, W. Li, O. van Kaick, A. Shamir, H. Zhang, and H. Huang. Learning to predict part mobility from a single static snapshot. *ACM Transactions on Graphics (SIGGRAPH Asia)*, 36(6):227:1–13, 2017.

[5] R. Hu, O. van Kaick, B. Wu, H. Huang, A. Shamir, and H. Zhang. Learning how objects function via co-analysis of interactions. *ACM Transactions on Graphics (SIGGRAPH)*, 35(4):47:1–13, 2016.

[6] R. Hu, Z. Yan, J. Zhang, O. van Kaick, A. Shamir, H. Zhang, and H. Huang. Predictive and generative neural networks for object functionality. *ACM Transactions on Graphics (SIGGRAPH)*, 37(4):151:1–13, 2018.

[7] R. Hu, C. Zhu, O. van Kaick, L. Liu, A. Shamir, and H. Zhang. Interaction context (icon): Towards a geometric functionality descriptor. *ACM Transactions on Graphics (SIGGRAPH)*, 34(4):83:1–12, 2015.

[8] V. Jain and H. Zhang. Robust 3d shape correspondence in the spectral domain. In *Shape Modeling International*, pages 118–129, 2006.

[9] J. Joon Park, P. Florence, J. Straub, R. Newcombe, and S. Lovegrove. DeepSDF: Learning continuous signed distance functions for shape representation. In *CVPR*, 2019.

[10] J. Li, K. Xu, S. Chaudhuri, E. Yumer, H. Zhang, and L. Guibas. GRASS: Generative recursive autoencoders for shape structures. *ACM Transactions on Graphics (SIGGRAPH)*, 36(4):52:1–10, 2017.

[11] R. Liu and H. Zhang. Segmentation of 3D meshes through spectral clustering. In *Proc. of Pacific Graphics*, pages 298–305, 2004.

[12] R. Liu, H. Zhang, A. Shamir, and D. Cohen-Or. A part-aware surface metric for shape analysis. *Computer Graphics Forum (Eurographics)*, 28(2):397–406, 2009.

[13] L. Mescheder, M. Oechsle, M. Niemeyer, S. Nowozin, and A. Geiger. Occupancy networks: Learning 3D reconstruction in function space. In *CVPR*, 2019.

[14] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *ECCV*, 2020.

[15] M. Reuter, F.-E. Wolter, and N. Peinecke. Laplace-beltrami spectra as "Shape-DNA" of surfaces and solids. *Computer Aided Design*, 38(4):342–366, 2006.

[16] O. Sidi, O. van Kaick, Y. Kleiman, H. Zhang, and D. Cohen-Or. Unsupervised co-segmentation of a set of shapes via descriptor-space spectral clustering. *ACM Transactions on Graphics (SIGGRAPH Asia)*, 30(6):126:1–9, 2011.

[17] R. Socher, C. C.-Y. Lin, A. Y. Ng, and C. D. Manning. Parsing natural scenes and natural language with recursive neural networks. In *ICML*, 2011.

[18] G. Taubin. A signal processing approach to fair surface design. In *SIGGRAPH*, 1995.

[19] Y. Wang, K. Xu, J. Li, H. Zhang, A. Shamir, L. Liu, Z. Cheng, and Y. Xiong. Symmetry hierarchy of man-made objects. In *Eurographics*, 2011.

[20] K. Xu, H. Zhang, D. Cohen-Or, and B. Chen. Fit and diverse: Set evolution for inspiring 3d shape galleries. *ACM Transactions on Graphics (SIGGRAPH)*, 31(4):57:1–57:10, 2012.

[21] Z. Yan, R. Hu, X. Yan, L. Chen, O. van Kaick, H. Zhang, and H. Huang. RPM-NET: Recurrent prediction of motion and parts from point cloud. *ACM Transactions on Graphics (SIGGRAPH Asia)*, 38(6):240:1–15, 2019.

[22] Z. Yi, H. Zhang, P. Tan, and M. Gong. DualGAN: Unsupervised dual learning for image-to-image translation. In *ICCV*, 2017.

[23] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks, 2017.