SIGGRAPH Asia 2020 Course Notes
Learning 3D Functionality Representations

Ruizhen Hu[1*] Manolis Savva[2*] Oliver van Kaick[3*]
[1]Shenzhen University [2]Simon Fraser University [3]Carleton University

Figure 1: *3D objects and scenes exhibit a variety of functionalities. The three examples here (coat rack, mug, and nightstand drawer) show the challenge of learning relations between functional parts (in orange) and other entities.*

Abstract

A central goal of computer graphics is to provide tools for designing and simulating real or imagined artifacts. An understanding of functionality is important in enabling such modeling tools. Given that the majority of man-made artifacts are designed to serve a certain function, the functionality of objects is often reflected by their geometry, the way that they are organized in an environment, and their interaction with other objects or agents. Thus, in recent years, a variety of methods in shape analysis have been developed to extract functional information about objects and scenes from these different types of cues.

In this course, we discuss recent developments involving functionality analysis of 3D shapes and scenes. We provide a summary of the state-of-the-art in this area, including a discussion of key ideas and an organized review of the relevant literatures. More specifically, we first present a general definition of functionality from which we derive criteria for classifying the body of prior work. This definition facilitates a comparative view of methods for functionality analysis. Moreover, we connect these methods to recent advances in deep learning, computer vision and robotics. Finally, we discuss a variety of application areas, and outline current challenges and directions for future work.

Keywords: Shape analysis, functionality analysis, geometric modeling, 3D representations, deep learning

1 Introduction

Functionality-aware processing is relatively new in graphics. However, the number of papers published at graphics conferences explicitly addressing functionality is growing. Moreover, incorporating models of functionality in an effective manner is a fundamental problem, important for applications such as fabrication and product design, which benefit from an explicit understanding of the function or purpose of an object. Moreover, the representation of functionality for 3D data has been of interest in related research fields such as computer vision and robotics, particularly in recent work that has been increasingly using 3D representations obtained with deep learning. We believe that a course on functionality analysis for computer graphics and adjacent research fields will be valuable and serve to delineate exciting directions for multi-disciplinary work bridging these fields.

The goal of our course is to give the necessary background to researchers wishing to enter this particular area of research by discussing the state-of-the-art works, organizing them in a structured scheme, and pointing out connections and opportunities for future work. In particular, we hope to motivate further work at the intersection of deep learning, computer graphics, and computer vision.

The course requires the audience to be familiar with basic concepts of geometry, computer graphics, and geometric modeling in 3D. Knowledge of machine learning is also helpful but not necessary.

2 Relation to previous courses

This course is related but complementary to a few earlier courses. The SIGGRAPH Asia 2016 course on “Directions in Shape Analysis towards Functionality” summarizes some of the earlier work in computer graphics, mostly in computer graphics, and does not provide the organizational framework of this course to classify prior work and connect it with other research areas. The SIGGRAPH Asia 2017 course on “Modeling and Remodeling 3D Worlds” focuses on 3D shape and scene modeling, with functionality-assisted modeling being one downstream application. The Eurographics 2018 tutorial on “Functionality Representations and Applications for Shape Analysis” does not cover more recent work in learned 3D representations for functionality, and does not address connections to deep learning, computer vision, and robotics. The Eurographics 2019 tutorial on “Learning Generative Models of 3D Structures” focuses on 3D shape and 3D scene generation, whereas we focus on 3D functionality representations with generation being only one of many other tasks such as analysis and 3D scene understanding.

3 Content of the course

- **Introduction:** introduction, description of what is covered in the course, rationale and motivation for the course.

*All the authors have equal contribution.

- **Part 1:** definition of functionality and organization of prior work along several axes.
- **Part 2:** geometry-only functionality analysis, summary of early work.
- **Part 3:** geometry + interaction–based functionality analysis, connections with learned 3D representations.
- **Part 4:** geometry + agent–based functionality analysis, connections with computer vision and robotics.
- **Conclusion**: summary of the state-of-the-art, and discussion of future challenges and possible future work.

4 Course materials

The course consists of talks given by three speakers, aided by a set of slides functioning as course notes. The complete slides are available at https://learn3dfunc.github.io/.

To complement the slides, Table 1 provides a list of references covered in our discussion of existing work. The references are organized according to a set of criteria derived from our definition of functionality [HSvK18]. The criteria and acronyms used in the table are explained as follows.

5 Classification of relevant literature

In our definition [HSvK18], the functionality of an entity is revealed by its geometry and a set of interactions between the entity in question (which we call the *functional entity*) and *interacting entities*. We classify existing methods based on the characteristics of the components of this definition.

Works. The methods in the literature are divided into three categories, based on whether a method only analyzes the geometry of the functional entity (*Geometry-only (G)*), or also considers interactions of the functional entity with other entities (*Geometry+interaction (GI))*, where we group the methods in which the functional entity is a human(-oid) agent in a separate category (*Geometry+agent (GA)*).

Functional entity. The level of organization at which the functional entity appears, which can be at the level of a *part*, *object*, *multi-object*, or *scene*.

Component / interacting entity. Denotes either the component of the geometry analyzed (part geometry or object geometry) or the type of interaction considered (static, dynamic, or agent interaction).

Dynamicity. The relation can be either static (*stat*) or dynamic (*dyn*).

Relations. Denotes how the relation between the entities is represented, which can be one of: Spatial arrangement (SA), Boundary representation (BR), Dense volume feature (VF), Gestalt and symmetry grouping (SG), Mechanical relations (MR), or Human(-oid) actions (HA).

Input. Denotes the representation of the input data, which can be one of: RGB-D image (rgbd), point cloud (pcl), polygonal mesh (mesh), or voxelized volume (voxels).

Approach. Denotes the nature of the approach used to analyze the functionality of the entities, which can be one of: supervised learning, unsupervised learning, or an approach handcrafted to the problem at hand.

Model type. Denotes the type of the model constructed, which can be either generative or discriminative.

6 Lecturers' biographies

Ruizhen Hu is an Assistant Professor at Shenzhen University, China. She received her Ph.D. from the Department of Mathematics, Zhejiang University. Before that, she spent two years visiting Simon Fraser University, Canada. Ruizhen's research interests are in shape analysis, geometry processing and fabrication. More details at: http://csse.szu.edu.cn/staff/ruizhenhu/

Oliver van Kaick is an Associate Professor at Carleton University, Ottawa, Canada. He received a Ph.D. from the School of Computing Science at Simon Fraser University (SFU). Oliver was then a postdoctoral researcher at SFU and Tel Aviv University. Oliver's research is concentrated in shape analysis and geometric modeling. More details at: http://www.scs.carleton.ca/~olivervankaick/index.html

Manolis Savva is an Assistant Professor at Simon Fraser University. He obtained his PhD at the Stanford Graphics Lab advised by Pat Hanrahan. His research focuses on analysis, organization and generation of 3D content. He has also worked in data visualization, grounding of natural language to 3D content, and on establishing large-scale 3D datasets such as ShapeNet and ScanNet. More details at: http://msavva.github.io

References

AKIZUKI S., AOKI Y.: Tactile logging for understanding plausible tool use based on human demonstration. In *BMVC Workshop on Vision for Interaction and Behaviour Understanding (VIBE)* (2018).

ANDRIES M., DEHBAN A., SANTOS-VICTOR J.: Automatic generation of object shapes with desired affordances using voxelgrid representation. *Frontiers in Neurorobotics 14* (2020), 22.

BAR-AVIV E., RIVLIN E.: Functional 3d object classification using simulation of embodied agent. pp. 307–316.

FISH N., AVERKIOU M., VAN KAICK O., SORKINE-HORNUNG O., COHEN-OR D., MITRA N. J.: Meta-representation of shape families. *ACM Trans. on Graph (SIGGRAPH) 33*, 4 (2014), 34:1–11.

FU Q., FU H., YAN H., ZHOU B., CHEN X., LI X.: Human-centric metrics for indoor scene assessment and synthesis. *Graphical Models 110* (2020), 101073.

FISHER M., RITCHIE D., SAVVA M., FUNKHOUSER T., HANRAHAN P.: Example-based synthesis of 3D object arrangements. *ACM Trans. on Graph (SIGGRAPH Asia) 31*, 6 (2012), 135:1–11.

FISHER M., SAVVA M., HANRAHAN P.: Characterizing structural relationships in scenes using graph kernels. *ACM Trans. on Graph (SIGGRAPH) 30*, 4 (2011), 34:1–12.

FISHER M., SAVVA M., LI Y., HANRAHAN P., NIESSNER M.: Activity-centric scene synthesis for functional 3D scene modeling. *ACM Trans. on Graph (SIGGRAPH Asia) 34*, 6 (2015), 179:1–13.

GELFAND N., GUIBAS L. J.: Shape segmentation using local slippage analysis. pp. 214–223.

Works	Functional entity	Representation of geometry or interactions: Component / interacting entity	Dynamicity	Relations	Additional classification criteria: Input	Approach	Model type
Geometry-only (G)							
Xu et al. [XSF02]	scene	object-geo	stat	SA	mesh	handcrafted	generative
Merrell et al. [MSL*11]	scene	object-geo	stat	SA	mesh	handcrafted	generative
Yu et al. [YYT*11]	scene	object-geo	stat	SA	mesh	supervised	generative
Fisher et al. [FSH11]	scene	object-geo	stat	SA	mesh	handcrafted	discriminative
Fisher et al. [FRS*12]	multi-object	object-geo	stat	SA	mesh	supervised	generative
Zhao et al. [ZWK14]	multi-object	object-geo	stat	BR	pcl	handcrafted	discriminative
Zhao et al. [ZHG*16]	multi-object	object-geo	stat	BR	mesh	supervised	generative
Zheng et al. [ZCOM13]	object	part-geo	stat	SG	mesh	handcrafted	generative
Mitra et al. [MYY*10]	object	part-geo	stat	SG	mesh	handcrafted	discriminative
Xu et al. [XLX*16]	object	part-geo	stat	SG	rgbd	handcrafted	discriminative
Fish et al. [FAvK*14]	object	part-geo	stat	SA	mesh	supervised	generative
Yumer et al. [YK14]	object	part-geo	stat	SA	mesh	supervised	generative
Pechuk et al. [PSR08]	part	part-geo	stat	SA	rgbd	supervised	discriminative
Gelfand et al. [GG04]	part	–	–	–	mesh	handcrafted	discriminative
Andries et al. [ADSV20]	object	–	stat	–	voxels	supervised	generative
Geometry+interaction (GI)							
Hu et al. [HZvK*15]	object	stat-inter	stat	BR	pcl	handcrafted	discriminative
Hu et al [HvKW*16]	object	stat-inter	stat	BR	pcl	supervised	discriminative
Pirk et al. [PKH*17]	object	dyn-inter	dyn	VF	mesh	handcrafted	discriminative
Myers et al. [MTFA15]	part	stat-inter	stat	SA	rgbd	supervised	discriminative
Kim et al. [KS14]	part	stat-inter	stat	SA	rgbd	supervised	discriminative
Laga et al. [LMS13]	part	stat-inter	stat	SA+SG	mesh	supervised	discriminative
Hu et al. [HLK*17]	part	stat-inter	dyn	SA+BR	pcl	supervised	discriminative
Xiang et al. [XQM*20]	part	stat-inter	dyn	SA	mesh	supervised	discriminative
Hu et al. [HYZ*18]	object	stat-inter	stat	SA+BR	vol	supervised	generative
Yi et al. [YHL*18]	part	stat-inter	dyn	SA	pcl	supervised	discriminative
Wang et al.[WZS*19]	part	stat-inter	dyn	SA	pcl	supervised	discriminative
Yan et al. [YHY*19]	part	stat-inter	dyn	SA	pcl	supervised	discriminative
Li et al. [LWY*20]	part	stat-inter	dyn	SA	pcl	supervised	discriminative
Kokic et al. [KSHK17]	part	stat-inter	dyn	SA	pcl	supervised	generative
Li et al. [LSK20]	part	stat-inter	dyn	SA	pcl	supervised	generative
Krs et al. [KMG*20]	object	stat-inter	stat	BR	mesh	supervised	generative
Geometry+agent (GA)							
Grabner et al. [GGVG11]	scene	agent-inter	stat	HA	mesh	supervised	generative
Savva et al. [SCH*14]	scene	agent-inter	stat	SA+HA	mesh	supervised	discriminative
Zhu et al. [ZJZ*16]	scene	agent-inter	stat	SA	mesh	supervised	generative
Jiang et al. [JKS13]	multi-object	agent-inter	stat	SA	rgbd	supervised	discriminative
Wang et al. [WLY17]	multi-object	agent-inter	stat	SA+HA	mesh	supervised	discriminative
Fisher et al. [FSL*15]	multi-object	agent-inter	stat	SA+HA	mesh	supervised	generative
Savva et al. [SCH*16]	multi-object	agent-inter	stat	SA+HA	mesh	supervised	generative
Ma et al. [MLZ*16]	multi-object	agent-inter	dyn	SA+HA	mesh	unsupervised	generative
Zheng et al. [ZLDM16]	object	agent-inter	stat	SA	mesh	handcrafted	generative
Kim et al. [KCGF14]	object	agent-inter	stat	SA	mesh	supervised	generative
Bar-Aviv & Rivlin [BAR06]	object	agent-inter	stat	SA+HA	mesh	handcrafted	discriminative
Zhu et al. [ZZCZ15]	object	agent-inter	dyn	SA+HA	rgbd	supervised	discriminative
Zhao et al. [ZCK17]	object	agent-inter	dyn	SA+HA	mesh	handcrafted	discriminative
Lee et al. [LCL06]	object	agent-inter	dyn	SA	mesh	supervised	generative
Li et al. [LLK*19]	scene	agent-inter	stat	SA+HA	rgbd	supervised	generative
Zhang et al. [ZHN*20]	scene	agent-inter	stat	SA+HA	rgbd	supervised	generative
Mao et al. [MZX*19]	object	agent-inter	stat	SA	mesh	supervised	generative
Fu et al. [FFY*20]	scene	agent-inter	stat	SA+HA	mesh	supervised	discriminative
Monszpart et al. [MGC*19]	scene	agent-inter	stat	SA	rgbd	supervised	generative
Ruiz et al. [RMC19]	scene	agent-inter	stat	SA+BR	mesh	supervised	generative
Starke et al. [SZKS19]	object	agent-inter	dyn	SA	vol	supervised	generative
Akizuki et al. [AA18]	object	agent-inter	dyn	SA+HA	rgbd	supervised	discriminative

Table 1: *Prior work classified according to our definition of functionality [HSvK18].*

GRABNER H., GALL J., VAN GOOL L.: What makes a chair a chair? In *Proc. Conf. on Computer Vision and Pattern Recognition (CVPR)* (2011), IEEE, pp. 1529–1536.

HU R., LI W., KAICK O. V., SHAMIR A., ZHANG H., HUANG H.: Learning to predict part mobility from a single static snapshot. *ACM Trans. on Graph (SIGGRAPH Asia) 36*, 6 (2017), 217:1–13.

HU R., SAVVA M., VAN KAICK O.: Functionality representations and applications for shape analysis. *Computer Graphics Forum (Eurographics State-of-the-art Report) 37*, 2 (2018), 603–624.

HU R., VAN KAICK O., WU B., HUANG H., SHAMIR A., ZHANG H.: Learning how objects function via co-analysis of interactions. *ACM Trans. on Graph (SIGGRAPH) 35*, 4 (2016), 47:1–13.

HU R., YAN Z., ZHANG J., VAN KAICK O., SHAMIR A., ZHANG H., HUANG H.: Predictive and generative neural networks for object functionality. *ACM Trans. on Graph (SIGGRAPH) 37*, 4 (2018), 151:1–151:13.

HU R., ZHU C., VAN KAICK O., LIU L., SHAMIR A., ZHANG H.: Interaction context (ICON): Towards a geometric functionality descriptor. *ACM Trans. on Graph (SIGGRAPH) 34*, 4 (2015), 83:1–12.

JIANG Y., KOPPULA H., SAXENA A.: Hallucinated humans as the hidden context for labeling 3D scenes. In *Proc. Conf. on Computer Vision and Pattern Recognition (CVPR)* (2013), IEEE, pp. 2993–3000.

KIM V. G., CHAUDHURI S., GUIBAS L., FUNKHOUSER T.: Shape2pose: Human-centric shape analysis. *ACM Trans. on Graph (SIGGRAPH) 33*, 4 (2014), 120:1–12.

KRS V., MECH R., GAILLARD M., CARR N., BENES B.: PICO: procedural iterative constrained optimizer for geometric modeling. *IEEE Trans. Visualization & Computer Graphics early access* (2020).

KIM D. I., SUKHATME G. S.: Semantic labeling of 3D point clouds with object affordance for robot manipulation. In *Int. Conf. Robotics and Automation (ICRA)* (2014), pp. 5578–5584.

KOKIC M., STORK J. A., HAUSTEIN J. A., KRAGIC D.: Affordance detection for task-specific grasping using deep learning. In *Proc. Conf. on Humanoid Robotics* (2017), pp. 91–98.

LEE K. H., CHOI M. G., LEE J.: Motion patches: building blocks for virtual environments annotated with motion data. *ACM Trans. on Graph (SIGGRAPH) 25*, 3 (2006), 898–906.

LI X., LIU S., KIM K., WANG X., YANG M., KAUTZ J.: Putting humans in a scene: Learning affordance in 3D indoor environments. In *Proc. Conf. on Computer Vision and Pattern Recognition (CVPR)* (2019), pp. 12360–12368.

LAGA H., MORTARA M., SPAGNUOLO M.: Geometry and context for semantic correspondences and functionality recognition in man-made 3D shapes. *ACM Trans. on Graph 32*, 5 (2013), 150:1–16.

LI Y., SCHOMAKER L., KASAEI S. H.: Learning to grasp 3d objects using deep residual u-nets. *arXiv preprint arXiv:2002.03892* (2020).

LI X., WANG H., YI L., GUIBAS L. J., ABBOTT A. L., SONG S.: Category-level articulated object pose estimation. In *Proc. Conf. on Computer Vision and Pattern Recognition (CVPR)* (2020).

MONSZPART A., GUERRERO P., CEYLAN D., YUMER E., MITRA N. J.: iMapper: interaction-guided scene mapping from monocular videos. *ACM Trans. on Graph (SIGGRAPH) 38*, 4 (2019), 92:1–92:15.

MA R., LI H., ZOU C., LIAO Z., TONG X., ZHANG H.: Action-driven 3D indoor scene evolution. *ACM Trans. on Graph (SIGGRAPH Asia) 35*, 6 (2016), 173:1–13.

MERRELL P., SCHKUFZA E., LI Z., AGRAWALA M., KOLTUN V.: Interactive furniture layout using interior design guidelines. *ACM Trans. on Graph (SIGGRAPH) 30*, 4 (2011), 87:1–10.

MYERS A., TEO C. L., FERM¨ULLER C., ALOIMONOS Y.: Affordance detection of tool parts from geometric features. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on* (2015), IEEE, pp. 1374–1381.

MITRA N. J., YANG Y.-L., YAN D.-M., LI W., AGRAWALA M.: Illustrating how mechanical assemblies work. *ACM Trans. on Graph (SIGGRAPH) 29*, 4 (2010), 58:1–12.

MAO A., ZHANG H., XIE Z., YU M., LIU Y., HE Y.: Automatic sitting pose generation for ergonomic ratings of chairs. *IEEE Trans. Visualization & Computer Graphics early access* (2019).

PIRK S., KRS V., HU K., RAJASEKARAN S. D., KANG H., YOSHIYASU Y., BENES B., GUIBAS L. J.: Understanding and exploiting object interaction landscapes. *ACM Trans. on Graph 36*, 3 (2017), 31:1–14.

PECHUK M., SOLDEA O., RIVLIN E.: Learning function-based object classification from 3D imagery. *Computer Vision and Image Understanding 110*, 2 (2008), 173–191.

RUIZ E., MAYOL-CUEVAS W. W.: Scalable real-time and one-shot multiple-affordance detection. In *ICRA Workshop on Computational Models of Affordance in Robotics* (2019).

SAVVA M., CHANG A. X., HANRAHAN P., FISHER M., NIESSNER M.: SceneGrok: Inferring action maps in 3D environments. *ACM Trans. on Graph (SIGGRAPH Asia) 33*, 6 (2014), 212:1–10.

SAVVA M., CHANG A. X., HANRAHAN P., FISHER M., NIESSNER M.: PiGraphs: Learning Interaction Snapshots from Observations. *ACM Trans. on Graph (SIGGRAPH) 35*, 4 (2016), 139:1–12.

STARKE S., ZHANG H., KOMURA T., SAITO J.: Neural state machine for character-scene interactions. *ACM Trans. on Graph (SIGGRAPH Asia) 38*, 6 (2019), 209:1–209:14.

WANG H., LIANG W., YU L.-F.: Transferring objects: Joint inference of container and human pose. In *Proc. Conf. on Computer Vision and Pattern Recognition (CVPR)* (2017), pp. 2933–2941.

WANG X., ZHOU B., SHI Y., CHEN X., ZHAO Q., XU K.: Shape2Motion: joint analysis of motion parts and attributes from 3D shapes. In *Proc. Conf. on Computer Vision and Pattern Recognition (CVPR)* (2019).

XU M., LI M., XU W., DENG Z., YANG Y., ZHOU K.: Interactive mechanism modeling from multi-view images. *ACM Trans. on Graph (SIGGRAPH Asia) 35*, 6 (2016), to appear.

XIANG F., QIN Y., MO K., XIA Y., ZHU H., LIU F., LIU M., JIANG H., YUAN Y., WANG H., YI L., CHANG A. X., GUIBAS L. J., SU H.: SAPIEN: A simulated part-based interactive environment. In *Proc. Conf. on Computer Vision and Pattern Recognition (CVPR)* (2020).

XU K., STEWART J., FIUME E.: Constraint-based automatic placement for scene composition. In *Proc. Graphics Interface* (2002), vol. 2.

YI L., HUANG H., LIU D., KALOGERAKIS E., SU H., GUIBAS L.: Deep part induction from articulated object pairs. *ACM Trans. on Graph (SIGGRAPH Asia) 37*, 6 (2018), 209:1–209:15.

YAN Z., HU R., YAN X., CHEN L., VAN KAICK O., ZHANG H., HUANG H.: RPM-Net: recurrent prediction of motion and parts from point cloud. *ACM Trans. on Graph (SIGGRAPH Asia) 38*, 6 (2019), 240:1–240:15.

YUMER M. E., KARA L. B.: Co-constrained handles for deformation in shape collections. *ACM Trans. on Graph (SIGGRAPH Asia) 33*, 6 (2014), 187:1–11.

YU L.-F., YEUNG S. K., TANG C.-K., TERZOPOULOS D., CHAN T. F., OSHER S.: Make it home: automatic optimization of furniture arrangement. *ACM Trans. on Graph (SIGGRAPH) 30*, 4 (2011), 86:1–12.

ZHAO X., CHOI M. G., KOMURA T.: Character-object interaction retrieval using the interaction bisector surface. *Computer Graphics Forum (Eurographics) 36*, 2 (2017), 119–129.

ZHENG Y., COHEN-OR D., MITRA N. J.: Smart variations: Functional substructures for part compatibility. *Computer Graphics Forum (Eurographics) 32*, 2pt2 (2013), 195–204.

ZHAO X., HU R., GUERRERO P., MITRA N., KOMURA T.: Relationship templates for creating scene variations. *ACM Trans. on Graph (SIGGRAPH Asia) 35*, 6 (2016), to appear.

ZHANG Y., HASSAN M., NEUMANN H., BLACK M. J., TANG S.: Generating 3D people in scenes without people. In *Proc. Conf. on Computer Vision and Pattern Recognition (CVPR)* (2020).

ZHU Y., JIANG C., ZHAO Y., TERZOPOULOS D., ZHU S.-C.: Inferring forces and learning human utilities from videos. In *Proc. Conf. on Computer Vision and Pattern Recognition (CVPR)* (2016), IEEE, pp. 3823–3833.

ZHENG Y., LIU H., DORSEY J., MITRA N. J.: Ergonomics-inspired reshaping and exploration of collections of models. *IEEE Trans. Visualization & Computer Graphics 22*, 6 (2016), 1732–1744.

ZHAO X., WANG H., KOMURA T.: Indexing 3D scenes using the interaction bisector surface. *ACM Trans. on Graph 33*, 3 (2014), 22:1–14.

ZHU Y., ZHAO Y., CHUN ZHU S.: Understanding tools: Task-oriented object modeling, learning and recognition. In *Proc. Conf. on Computer Vision and Pattern Recognition (CVPR)* (2015), pp. 2855–2864.