Towards Artificial Systems: What Can We Learn from Human Perception?

Heinrich H. Bülthoff[1,2] and Lewis L. Chuang[1]

[1] Max Planck Institute for Biological Cybernetics, Spemannstraße 38, 72076 Tübingen, Germany

[2] Department of Brain and Cognitive Engineering, Korea University, Anam-dong, Seongbuk-gu, Seoul, 136-713, Korea

{Heinrich.Buelthoff,Lewis.Chuang}@Springer.com

Abstract. Research in learning algorithms and sensor hardware has led to rapid advances in artificial systems over the past decade. However, their performance continues to fall short of the efficiency and versatility of human behavior. In many ways, a deeper understanding of how human perceptual systems process and act upon physical sensory information can contribute to the development of better artificial systems. In the presented research, we highlight how the latest tools in computer vision, computer graphics, and virtual reality technology can be used to systematically understand the factors that determine how humans perform in realistic scenarios of complex task-solving.

Keywords: perception, object recognition, face recognition, eye-movement, human-machine interfaces, virtual reality, biological cybernetics.

The methods by which we process sensory information and act upon it comprise a versatile control system. We are capable of carrying out a multitude of complex operations, in spite of obvious limitations in our biological "hardware". These capabilities include our ability to expertly learn and identify objects and people by effectively navigating our eyes and body movements in our visual environment. This talk will present the research perspective of the Biological Cybernetics labs at the Max Planck Institute, Tübingen and the Department of Brain and Cognitive Engineering, Korea University. Key examples will be drawn from our research on face recognition, the relevance of dynamic information and active vision; in order to convey how perceptual research can contribute towards the development of better artificial systems.

To begin, our prodigious ability to learn and remember recently encountered faces – even from only a few instances - reflects a multi-purpose pattern recognition system that few artificial systems can rival, even with the availability of 3D range data. Unintuitively, this perceptual expertise relies on fewer, rather than more, facial features than state-of-the-art face-recognition algorithms typically process. Our visual field of high acuity is extremely limited (~2°) and experimental studies indicate that we have an obvious preference for selectively fixating the eyes and noses of faces that we inspect [1]. These facial features inhabit a narrow bandwidth of spatial frequencies (8 to 16 cycles per face), that face-processing competencies are specialized for [2]. Therefore, perceptual expertise appears to result from featural selectivity, wherein

B.-T. Zhang and M.A. Orgun (Eds.): PRICAI 2010, LNAI 6230, pp. 1–3, 2010.

sparse coding by a dedicated system results in expert discrimination. The application of the same principles in artificial systems holds the promise of improving automatic recognition performance.

Self-motion as well as moving objects in our environment dictate that we have to deal with a visual input that is constantly changing. Automated recognition systems would often consider this variability to be a computational hindrance that disrupts the stable retrieval of recognizable object features. Nonetheless, human recognition performance on objects [3] and faces [4] is better served by moving rather than static stimuli. Understanding why this is so, could allow artificial recognition systems to function equally well in dynamic environments. First, dynamic presentations present the opportunity for associative learning between familiar object views, which could result in object representations that are robust to variations in pose [5, 6]. Furthermore, dynamic presentations could allow the perceptual system to assess the stability of different object features, according to how they tend to appear and disappear over rigid rotations. This could offer a computationally cheap method for determining the minimal set of object views that would be sufficient for robust recognition [7, 8]. Finally, characteristic motion properties (e.g., trajectories, velocity profile) could even serve as an additional class of features to complement a traditional reliance on image and shape features by automated recognition systems [9, 10].

Purposeful gaze behavior indicates a perceptual system that is not only capable of processing information, but proficient in seeking out information, too. We are capable of extracting a scene's gist within the first few hundred milliseconds of encountering it [11]. In turn, this information directs movement of our eyes and head for the joint purpose of fixating information-rich regions across a large field of view [12]. In addition, we use our hands to explore and manipulate objects so as to access task-relevant information for object learning or recognition [13, 14, 15]. Careful observations of how we interact with our environments can identify behavioral primitives that could be modeled and incorporated into artificial systems as functional (and re-usable) components [16]. Furthermore, understanding how eye and body movements naturally coordinate can allow us improve the usability of artificial systems [17].

This perspective of the perceptual system as an active control system continues to be insightful at a higher level, when we consider the human operator as a controller component in dynamic machine systems. Take, for example, a pilot who has to simultaneously process visual and vestibular information, in order to control helicopter stability. Using motion platforms and immersive graphics, it is possible to systematically identify the input parameters that are directly relevant to a pilot's task performance and thus, derive a functional relationship between perceptual inputs and performance output [18]. Such research is fundamental for the development of virtual environments that are perceptually realistic. This is especially important when designing artificial systems (e.g., flight simulators) that are intended to prepare novices for physically dangerous situations that are not easily replicable in the real world [19].

Until now, we have discussed how findings from perceptual research can contribute towards improving artificial systems. However, the growing prevalence of these systems in our daily environs raises an imperative to go beyond this goal. It is crucial to consider how perceptual and artificial systems may be integrated into a coherent whole by considering the "human-in-the-loop". Doing so will lead towards a new generation of autonomous systems that will not merely mimic our perceptual competencies, but will be able to cooperate with and augment our natural capabilities.

References

1. Armann, R., Bülthoff, I.: Gaze Behavior in Face Comparison: The Role of Sex, Task, and Symmetry. Attention, Perception & Psychophysics 71, 1107–1126 (2009)
2. Keil, M.S.: "I Look in Your Eyes, Honey": Internal Face Features Induce Spatial Frequency Preference for Human Face Processing. PLOS Computational Biology 5, 1–13 (2009)
3. Chuang, L., Vuong, Q.C., Thornton, I.M., Bülthoff, H.H.: Recognizing novel deforming objects. Visual Cognition 14, 85–88 (2006)
4. Knappmeyer, B., Thornton, I.M., Bülthoff, H.H.: The use of facial motion and facial form during the processing of identity. Vision Research 43, 1921–1936 (2003)
5. Wallis, G.M., Bülthoff, H.H.: Effects of Temporal Association on Recognition Memory. Proceedings of the National Academy of Sciences of the United States of America 98, 4800–4804 (2001)
6. Wallis, G., Backus, B.T., Langer, M., Huebner, G., Bülthoff, H.H.: Learning Illumination- and Orientation-Invariant Representations of Objects Through Temporal Association. Journal of Vision 9, 1–8 (2009)
7. Wallraven, C., Bülthoff, H.H.: Automatic Acquisition of Exemplar-Based Representations for Recognition from Image Sequences. In: IEEE Conference on Computer Vision and Pattern Recognition (2001)
8. Bülthoff, H.H., Wallraven, C., Graf, A.: View-Based Dynamic Object Recognition Based on Human Perception. In: Proceedings of the 16th International Conference on Pattern Recognition, pp. 768–776 (2002)
9. Newell, F.N., Wallraven, C., Huber, S.: The role of characteristic motion on object categorization. Journal of Vision 4, 118–129 (2004)
10. Curio, C., Bülthoff, H.H., Giese, M.A.: Dynamic Faces: Insights from Experiments and Computation. MIT Press, Cambridge (2010)
11. Vogel, J.A., Schwaninger, A., Wallraven, C., Bülthoff, H.H.: Categorization of natural scenes: Local versus Global Information and the Role of Color. ACM Transactions on Applied Perception 4, 1–21 (2010)
12. Chuang, L.L., Bieg, H.-J., Fleming, R.W., Bülthoff, H.H.: Measuring Unrestrained Gaze on Wall-Sized Displays. In: Proceedings of the 28th European Conference on Cognitive Ergonomics (2010)
13. Chuang, L.L., Vuong, Q.C., Thornton, I.M., Bülthoff, H.H.: Human Observers Use Personal Exploration Patterns in Novel Object Recognition. Perception 36(Suppl.), 49 (2007)
14. Blanz, V., Tarr, M.J., Bülthoff, H.H.: What Object Attributes Determine Canonical Views? Perception 28, 575–600 (1999)
15. Gaißert, N., Wallraven, C.: Perceptual Representations of Parametricaly-Defined and Natural Objects Comparing Vision and Haptics. In: Proceedings of the Haptics Symposium, pp. 1–8 (2010)
16. Sprague, N., Ballard, D., Robinson, A.: Modeling Embodied Visual Behaviors. Transactions on Applied Perception 4, 1–26 (2007)
17. Bieg, H.-J., Chuang, L.L., Fleming, R.W., Reiterer, H., Bülthoff, H.H.: Eye and Pointer Coordination in Search and Selection Tasks. In: Proceedings of the Symposium on Eye Tracking Research and Applications, pp. 89–92 (2010)
18. Nieuwenhuizen, F.M., Zaal, P.M.T., Teufel, H., Mulder, M., Bülthoff, H.H.: The Effect of Simulator Motion on Pilot Control Behaviour for Agile and Inert Helicopter Dynamics. In: 35th European Rotorcraft Forum, pp. 1–13 (2009)
19. Niccolini, M.L., Pollini, L., Innocenti, M., Robbuffo Giordano, P., Teufel, H., Bülthoff, H.H.: Towards Real-Time Aircraft Simulation with the MPI Motion Simulator. In: Proceedings of the 2009 AIAA Modeling and Simulation Technologies Conference, pp. 1–10 (2009)