Abstract:

By far, most of the bits in the world are image and video data. YouTube alone gets 300 hours of video uploaded every minute. Adding to that personal pictures, videos, TV channels and the gazillion of security cameras shooting 24/7 one quickly sees that the amount of visual data being recorded is colossal. In the first part of this talk I will discuss the problem of “saliency prediction” - separating between the important parts of images/videos (the “wheat”) from the less important ones (the “chaff”). I will review work done over the last decade and its achievements. In the second part of the talk I will discuss one particular application of saliency prediction that our lab is interested in: making images and videos accessible to the visually impaired. Our plan is to convert images and videos into tactile surfaces that can be "viewed" by touch. As it turns out, saliency estimation and manipulation both play a key factor in this task.