

How much can we infer about a person's looks from the way they speak? In this paper, we study the task of reconstructing a facial image of a person from a short