This paper presents a novel method to manipulate the visual appearance (pose and attribute) of a person image according to natural language descriptions. Ou