In the following sections, we first introduce the problem formulation and denotation for motion transfer between human and anime character, and then propose a three-stage strategy that exploit structure and content information separately from real human video and anime character image and finally generate the target anime video.

A. definition of the problem.

In this work, given an arbitrary real human video {V^r\_x} with specifc motion ‘x’, and a single 2D anime image {I^a}, we aim to synthesis a video sequence {V^a\_x} in which anime character ‘a’ performing motion ‘x’. Therefore, we can clearly formulate the total objective of our project as:

F(V^r\_x, I^a) = V^a\_x,

where F(·) is an abstraction function of our entire system.

However, ready-made pair (V^r\_x, I^a, V^a\_x), which will serve as ground true, cannot be collected neither online nor from real world. In detailed, we don’t have anime video performing the same motion ‘x’ as real video we collected. To fix this common problem in synthetic tasks, the previous work either generate

Given a video sequence of a human performing an action(e.g. running), and a single image of an anime character, we aim to generate a new video sequence of that anime character performing the same action.