VS: Reconstructing Clothed 3D Human from Single Image via Vertex Shift

Leyuan Liu[a], Yuhan Li [a], Yunqi Gao [a], Changxin Gao [b], Yuanyuan Liu [c], Jingying Chen [a]

[a] Central China Normal University , [b] Huazhong University of Science and Technology, [a] China University of Geosciences (Wuhan)

Introduction

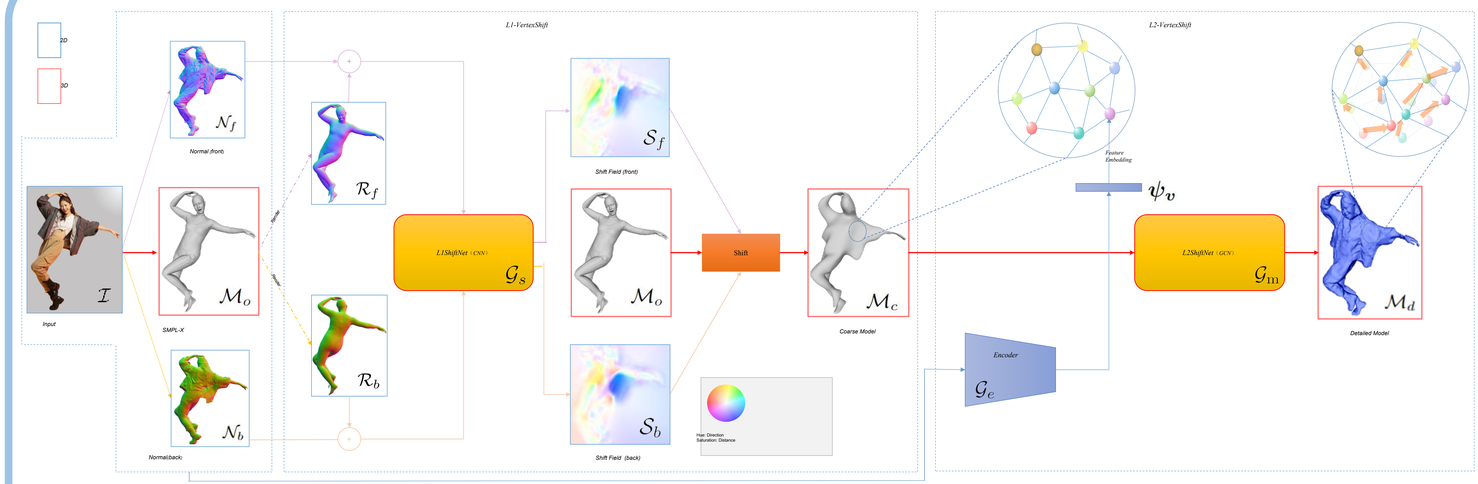
Facing the challenge of adapting to diverse human poses and reconstructing loose clothing， we propose VS, which stands for vertex shift, to reconstruct high-fidelity and artifact-less clothed 3D humans from single images. VS is a two-stage deformation method that uses a “stretch-refine” strategy to reconcile the contradiction between large deformations required for reconstructing loose clothing and delicate deformations for recovering details.Our contributions are summarized as follows:

(1) We propose a two-stage deformation method that uses a “stretch-refine” strategy for clothed 3D human re_x0002_construction, contributing to reconciling the contradiction between large deformations for reconstructing loose cloth_x0002_ing and delicate formations for recovering surface details.

(2) We introduce shift fields inferred from normal maps for stretching the coarse model to align well with the input image, allowing our deformation method to handle loose clothing and correct inaccurate pose estimates.

(3) We combine implicit-function-learned features with a graph convolutional network, making VS not only recover surface details but also suppress artifacts.

Method



VS employs a “stretch-refine” strategy to stepwise deform the SMPL-X (Mo) into a coarse human model (Mc) and a detailed human model (Md) using StretchVS and RefineVS modules, respectively.

(1) First, two shift fields (S = {Sf , Sb}) are inferred by warping the body normals (R = {Rf , Rb}) into clothing normal maps (N = {Nf , Nb}) via a CNN (ShiftEstNet, Gs). Then, StretchVS shifts vertices of the SMPL-X to from the coarse model using the shift fields.

(2) Taking the coarse model as input, RefineVS employs a GCN (DeformNet, Gm) embedded with implicit-function-learned features (Ψ) to infer vertex locations of the detailed model

Results