Technical Report of Image based Hair Modeling

Tianyu Wang

2019.11.20

1 How we do it now

1.1 Semi-automatic hair modeling tool based on single image

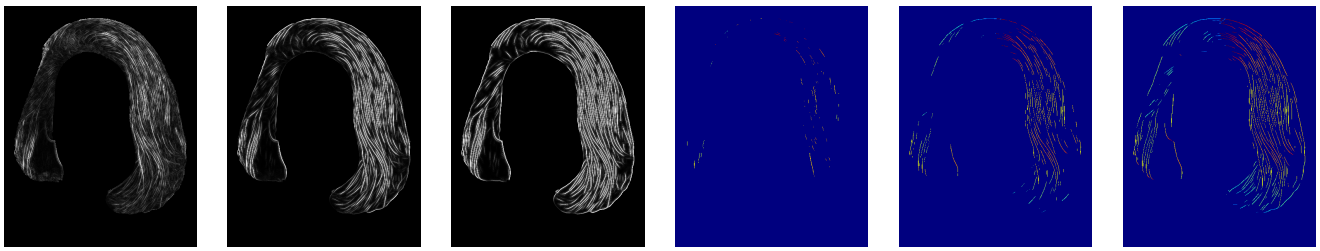
After simple user interaction annotation on the input image, the system will generate hair model in strand level without more user interaction. This solution is mainly based on the pipeline of [Chai et al., 2016] and we referenced some other papers to add more flexibility to artists' refinement and control. The generated hair modeling finally can be used in physically based simulation. Some early papers' results can not be used for simulation, such as [Chai et al., 2012, Chai et al., 2013].

Given a portrait I,as shown in figure 1(a):
a.We need artists to select the hair region and this can be done by many interactive segmentation methods like [Liu et al., 2009] and we get the hair mask M,as shown in figure 1(b);

(a) I (b) M (c) $D_{dense} = ((cos + 1)/2, (sin + 1)/2, 0)$

图 1:

b.After getting the hair region M, according to the paper [Chai et al., 2012]'s iterative direction filter method which is a modified version of the paper [Paris et al., 2004], we get the direction map of hair region named as theta map D and the confidence map to this direction estimation C;



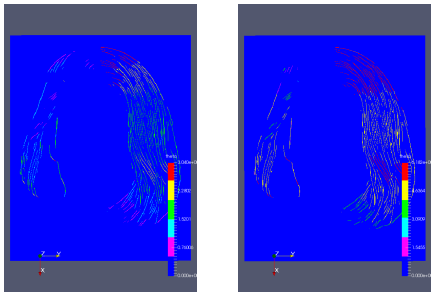
(a) C_0 (b) C_3 (c) C_6 (d) $D_{sparse0}$ (e) $D_{sparse3}$ (f) $D_{sparse6}$

图 2: As shown, we get refined sparse direction field and confidence map iteratively

c.Because of the image's noise in generation or filming, our system will truncate the theta map D according to the confidence map C, that we regard the direction estimation whose confidence map value is higher that a given threshold is reliable. Then we can get the useful sparse direction estimation D_{sparse},as shown in figure 2;

(a) semantic annotation by user interaction

图 3: semantic annotation by user interaction



(a) D_{sparse} (b) $D_{sparse/correct}$

图 4: We correct the direction map (a) from filtering in the range [0,pi] by annotation and get the direction map (b) in the range [0,2pi] without ambiguity.

d.Because the direction map from filtering is ambiguous as shown in [Chai et al., 2013], that we can not judge whether the direction is right or its opposite direction is right(we define the direction from the scalp to outside is the right direction, but the filter can not get this high-level semantic meaning), we give the artist or other user a interaction interface,as shown in figure 3, and let the user to segment the hair region into some sub-regions and annotate the general direction from scalp to outside by their observation. According to this annotated results, we check whether the sparse direction D_{sparse} from filtering is right or not, and get the corrected version $D_{sparse/correct}$,as shown in the figure 4;

e.For getting the dense direction field in hair region for each pixel as D_{dense}, we diffuse the sparse corrected direction map from last step $D_{sparse/correct}$ into the whole hair region(A quadratic optimization is needed as shown in [Fu et al., 2007]). Then we get the D_{dense}, as shown in the figure 1(c);

(a)The 3D hair model database

图 5:

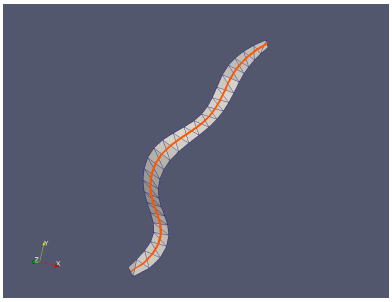


图 6: A strip from a hair model can be represented by polyline for simplicity of distance computation.

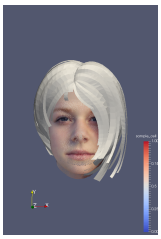
f.We collect 305 hair model represented by some strips, as shown in the figure 5, and we recombine these models to augment the hair database into about 40k, and we need to project them into 2D to get their direction map D^* and hair region mask M^* for retrieval;

In details, we get the strip set of each model by Union-find Set, then represent each strip using polyline, as shown in the figure 6, compute the distance between each pair of polyline and get $n*n$ distance matrix. According to Agglomerative Clustering algorithm(There are alternative methods (SpectralClustering, AffinityPropagation, DBSCAN, OPTICS) , but according to the experiment results, the quality of Agglomerative Clustering is best), we cluster the model into several large strip sets, and for each two models, we recombine them by adding new cluster and removing conflict cluster on the cluster level (From two point clouds' hausdorff distance, we can judge the possibility of conflict) and generate new hair model to augment the hair database, as shown in figure 7.

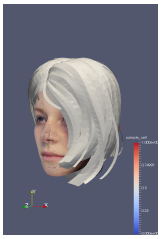
图 7: In each sub figure, we cluster the medium and right shown by different colors and do recombination between clusters to generate new hair style as left.

g.We can retrieval in the hair database to find the best candidates matching the given image's direction field D_{dense} or regress by deep learning; we show one of the candidates h_i in figure 8(a) matching the given image I. For the next step, we need to fit a face model and get the camera parameter first to align the selected hair model with the given image I.

In details, we first get 68 face landmarks from the give image I as shown in figure 8(d) using the state of the art 2D or 3D face landmark detection method [Bulat and Tzimiropoulos, 2018] and fit a 3D face morphing model and estimate the camera parameter based on basel face model 2009 database [Paysan et al., 2009] by solving a non-linear least square problem using trust region method. We project our reconstructed 3D face model and selected hair model into 2D, showing the difference of the given landmarks and our projected landmarks in figure 8(e). We further render hair model and face model by OpenGL in the given image's pose to get the corresponding segmentation area as shown in figure 8(f).

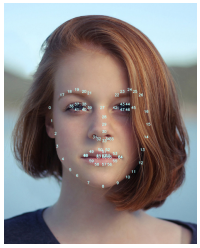


(a) selected model

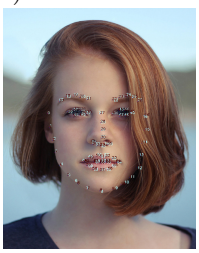


(b)show in another pose

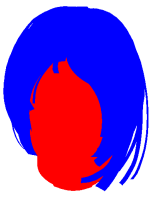
(c) the given image



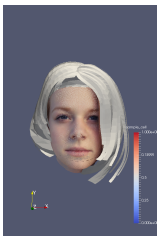
(d) the given landmarks



(e) the estimated landmarks



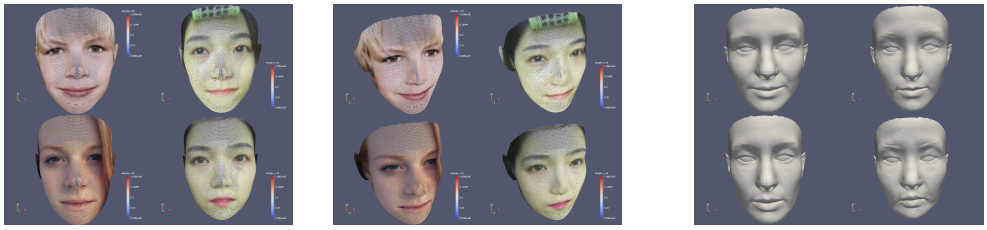
(f) render to segment



(g)mesh after deformation

图 8: As shown, we select the hair model best matching the given image and fit a face model to the given image and estimate the camera parameter of the given image.

We also show some other cases of the image based face fitting results to verify the estimated camera parameter is right as shown in figure 9.



(a) face from image (b) show in another pose (c) without color from image

图 9: We fit the face based on BFM 2009 first for estimating the camera parameter and we show some other cases in this step.

h.For each hair model h_i in the best candidate set S_h,first we do boundary matching between it and the given image's hair boundary by solving a dynamic programming as shown in [Forney, 1973].

In details, we get the hair region from last step as shown in figure 10(a) and apply Gaussian filter to smooth the result to get the contour line of hair model. Now we have the contour line of the hair model and the contour line of the given image's hair mask as show in figure 10.

(a) hair region (b) after smoothing (c) hair mask of image (d) contour line comparison

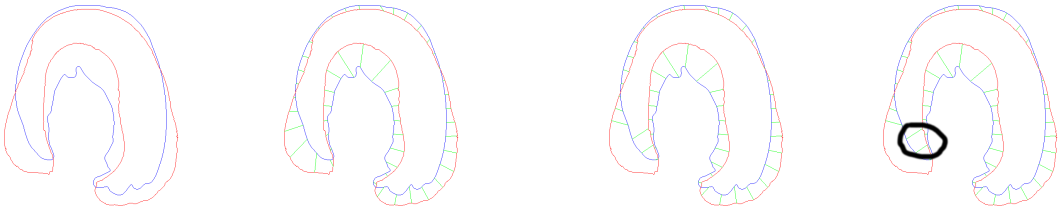
图 10: We get the hair region of the selected model and smooth it for contour line extraction.

We sample some points on two contour lines and get their best correspondence by minimizing this energy 1 using a dynamic programming as shown in [Forney, 1973]. In [Chai et al., 2016], the energy term 3 is to penalize the difference of the hair boundary edge and image boundary edge in Euclidean distance measure, but clearly, we should penalize the difference of the edge length in geodesic distance measure. As shown in 11, we can see that [Chai et al., 2016]'s method can not get the right correspondence and after our modification, the correspondence is right.

$$\min_{M} \sum_{P_i^H} (E_p(P_i^H) + E_e(P_i^H, P_{i+1}^H)) \tag{1}$$

$$E_p(P_i^H) = |P_i^H - P_{M(i)}^I|^2 + \lambda_n (1 - n_i^H . n_{M(i)}^I)^2 \tag{2}$$

$$E_e(P_i^H, P_{i+1}^H) = (|P_i^H - P_{i+1}^H| - |P_{M(i)}^I - P_{M(i+1)}^I|)^2 \tag{3}$$

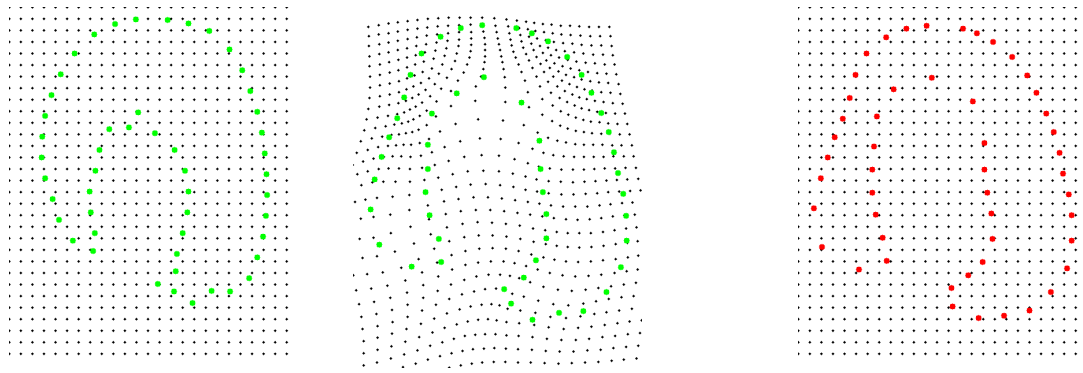


(a) contour lines (b) by geodesic distance (c)by [Chai et al., 2016] (d) wrong area

图 11: We get sparse hair boundary correspondence and we modify the wrong formulation in [Chai et al., 2016].

i.Then we can do warping deformation for the hair model with the image information using

the method in [Donato and Belongie, 2002, Zhou et al., 2005]. This post processing is correcting the hair model for better similarity;



(a)2D hair before warping (b)2D hair after warping (c)corresponding points in the given image

图 12: We can warping the whole hair domain into the image by thin plain spline method.

In details, we further interpolate the boundary correspondence to every pixel, using the Thin-Plate-Spline (TPS) method as shown in figure 12. Then we only need to deform the 3D hair mesh by this correspondence information using some mesh deformation method. Here we select the Laplacian deformation for keeping local feature. The result is shown in figure 8(g)

Actually, this system has some limitations and there are some possible solutions for that:
The data augmentation and searching algorithm are not optimal combination. Searching can only use limited information but actually the local similarity but global dissimilarity will let the searching algorithm reject. So it is hard to find a good initial candidate that marching the given image well always. And in this framework, the 2D direction field is not being used well that only step(j)and(g) use it for selecting candidates but this is too soft to constrain the final generation's projection as similar to the given direction field as possible. A deep learning framework is more useful for local similarity exploration and using 2D direction field as constraint, but on the other hand, the training will smooth the 3D generation as shown in [Yang et al., 2019]. If we can not only use the 3D information to supervise the network's generation but also use the projected 2D direction field to supervise its 2D projected generation which can be facilitated by differentiable rendering techniques as [Li et al., 2018, Tewari et al., 2018], the results may be more realistic. There exist some works on face modeling like [Tewari et al., 2018] using differentiable rendering, but how to extend it into hair modeling is not clear. Hair is much easier to deform and its visibility is more complex compared with face which may be a challenge in this idea.

j.We then select the hair model after warping, selecting the model whose 2D projected direction field D^* matches the given image's direction field D_{dense} best as the final solution h_{best};

k.After getting the best hair model, we voxel it and diffuse it to get dense 3D direction field $D_{dense3d}$ and then we generate hair strand from scalp according to [Chai et al., 2013].

Here is a almost complete survey of the recent state of the art papers related to image based hair modeling, from 2012 to 2019, published on SIGGRAPH(Asia) or CVPR. In Chinese now.

2 Survey of Image based Hair Modeling

1.[Luo et al., 2012]: 使用多张不同角度拍摄的头发获得的照片重建外层的头发模型 (一层表面三角网格, 不可用于直接的物理模拟): 因为头发的拍摄视角依赖的高光, 因此颜色不适合作为特征应用于多视角的 correspondence 中。[Luo et al., 2012] 对每张图像抽取发丝方向特征，根据该特征通过 MRF 优化获得部分深度信息，最后根据部分深度信息进行泊松重建并 refine 获得最终的 hair mesh。The key contribution is the orientation map's importance.

2.[Chai et al., 2012]: 使用单张正面的肖像照片和少量的用户交互标注（标注脸部区域和头发区域）重建 2.5D 的正面头发和脸部模型用于图像的替换发型和 relighting（不可用于直接的物理模拟）。对于头发区域通过 Gabor 滤波获得稀疏的可靠的发丝方向并追踪这些方向生成稠密的 2D 头发 strand(并不是从头皮开始)；最后结合拟合的 3D 脸部模型，对每个 2D 头发顶点估计深度获得对应的 3D 头发，并生成更多的 3D 头发填充头发区域。annotation to solve the direction ambiguity.

3.[Chai et al., 2013]: 使用单张正面的肖像图片和少量用户标注（标注头发的发丝从头皮开始的走向）重建可用于物理模拟的但不完整的头发和脸部模型。通过用户标注的走向结合 Gabor filter 获得的发丝方向，生成 2D 的无歧义的发丝方向，并使用 [Chai et al., 2012] 的方法估计其深度信息获得稀疏的 3D 方向，并进行 diffuse 获得稠密的 3D 头发方向，最后使用一种迭代式的头发生成算法，生成从头皮生长出的头发。The key contribution is the novel hair generation algorithm.

4.[Luo et al., 2013]: 使用多张不同角度拍摄的头发获得的照片生成可用于模拟的完整的头发模型（结果看上去很好）。1. 通过 PMVS algorithm 重建点云，2. 根据 [Luo et al., 2012] 计算 2D 方向图，3. 根据 1,2 结果重建 3D 方向场，4. 获取 Ribbons 表示的 3D 头发模型（相互连接，连接到头皮，方向性分析（避免方向的二义性），this is useful to convert a 3D direction field into a hair model expressed by ribbons，5. 生成头发。

5.[Hu et al., 2014b]: 使用多张不同角度拍摄的头发获得的照片生成受约束的特定发型-辫子。1. 通过 [Luo et al., 2013] 的方法获得点云模型和 3D 方向场，2. 通过过程式的方法构造一个辫子数据库，3. 将输入点云和数据库做 fitting（部分匹配，相互连接），4.diffuse and hair generation as[Luo et al., 2013]. data driven strategy

6.[Xu et al., 2014]: 动态头发建模 from video。时空约束。

7.[Hu et al., 2014a]: 使用多张不同角度拍摄的头发获得的照片生成可用于模拟的完整的头发模型。其使用模拟生成的头发作为初始样例，解决 [Luo et al., 2013] 的方向性分析方法的问题，可以在不对初始点云进行清洗的情况下，获得比 [Luo et al., 2013] 更好的结果。它首先将模拟生成的头发和点云对应的 sparse hair strand 进行 ICP fitting，然后根据点云对应的 sparse hair strand 生成头发模型的 ribbons 表示，但是和 [Luo et al., 2013] 不同，它使用 fitting 后的模拟的头发发丝作为指引连接 ribbon 片段，提高鲁棒性，但是受算法策略限制，相比较 [Luo et al., 2013] 在局部细节上表现不好。The key idea is using the simulated examples as the direction guidance.

8.[Chai et al., 2015]: 使用一张单张正面的肖像图片和少量用户标注重建出精确的深度图。shape from shading. 1. 构造一个包含头发区域和脸部区域的基本形状，2. 使用 SFS 估计精确的头发和脸部的法向量，3. 通过构造螺旋线表示的头发来拟合图像头发区域，获得该螺旋线的法向量,4. 数值优化求解 per pixel's depth 满足基本形状的深度约束，SFS 法向量约束和拟合获得的螺旋线法向量的约束，获得深度图表示的头发模型（并不能用于物理模拟，和工业界的常规需求不是很一致）。

9.[Hu et al., 2015]: 使用单张正面的肖像图片重建出可用于物理模拟的完整的头发模型。1. 给定图像的头发区域一些标注表示典型的头发走向，在一个三维头发模型中查找和给定标注最相似的作为候选，combine 候选，将 combine 以后的候选根据给定图像的 2D 方向场进行头发丝级别的 deform。和 [Chai et al., 2016] 的做法比较相似，但是在 retrieval 步使用的表示不一样，combine 和 retrieval 的顺序不一样。

10.[Chai et al., 2016]: 使用单张正面的肖像图片重建出可用于物理模拟的完整的头发模型。通过构造一个肖像图图片和对应标注的头发区域的 segmentation mask and 2D direction map, 训练一个神经网络，从而可以输入一个肖像获得对应的 hair region segmentation mask and its 2D direction map；另一方面构造一个三维头发模型数据集，通过肖像的 segmentation 和 2D direction 信息在数据集进行搜索候选模型并进行形变，最终根据 [Chai et al., 2013] 的方法进行 diffuse and hair generation。Some previous papers' algorithm's combination(1.orientation from Gabor filter from[Luo et al., 2012, Chai et al., 2012] / 2.annatation for solving the ambiguity from [Chai et al., 2013]/ 3.wisp's construction from [Luo et al., 2013]/ 4.hair database idea from [Hu et al., 2015]/ 5.hair generation algorithm from [Chai et al., 2013] / 6.hair strand deformation algorithm from [Hu et al., 2015]) and a novel hair wisp level deformation processing from itself.

11.[Zhang et al., 2017]: 使用前后左右四张头发照片重建出可用于物理模拟的完整的头发模型。1. 对四张图片进行头发区域分割和相机参数估计，2. 利用估计的相机参数绘制头发模型数据库中每个模型的 2 维投影，3. 通过投影的轮廓线距离进行搜索获得最优模型，转成 water tight 闭合 hair mesh,4. 对 hair mesh 根据 4 张给定图像进行 deform，similar to [Chai et al., 2016],5. 对 hair mesh 贴四张图片的纹理，根据纹理估计其表面头发走向并 diffuse 至内部，6. 生成头发。data driven stategy. sparse multi-view based hair modeling.

12.[Saito et al., 2018]: 使用单张正面的肖像图片重建出可用于物理模拟的完整的头发模型。1. 训练一个 VAE,encode the hair feature into the latent space,input is the hair occupancy volume and flow volume and the output is the reconstructed hair occupancy volume and flow volume. 2. 训练一个 hair embedding network 预测 latent code from input image.

13.[Zhou et al., 2018]: 使用单张正面的肖像图片重建出可用于物理模拟的完整的头发模型。1.

训练一个 AE,encode the 2D orientation info and mask info into a latent space named as hair feature, decode the hair feature as hair strand feature. hair strand feature 直接表示为一个 N*M matrix, N 根头发，每根头发有 M 个有序顶点。训练数据生成方法在 [Chai et al., 2016] 基础上有所变动。[Zhou et al., 2018] and [Saito et al., 2018] 很相似，只不过 [Zhou et al., 2018] 用了一个网络直接从图像特征到 hair strand 特征，[Saito et al., 2018] 用了两个网格，一个 encode/decode ，另一个用来从图像特征预测。

14.[Zhang et al., 2018]: 通过 RGB-D 相机重建出可用于物理模拟的完整的头发模型。1. 对 color images 进行头发区域和方向估计 as [Chai et al., 2016],2. 利用深度 images 获得 3Dfusion model,3. 对于前后两个视角下的 RGB 图片估计的 hair mask 和 2D direction field 在一个预定义的 hair 数据库寻找局部细节相似的模型，4. 构建 3 中寻找到的模型和深度相机获得的 fusion model 之间的 NNF 映射，5. 构建 3D 方向场生成头发。The key idea is combining the global feature given by the fusion model and the local feature given by the searched model with 2D direction clues.

15.[Liang et al., 2018]: 通过一段人物的多个角度的视频重建出三维头部和头发。1.SFM 重建出 visual hull,2. 获取逐帧的 2D hair direction and strand ,3. 通过一帧正面肖像重建人脸,4. 通过 2D strands 检索 hair database, 获取多个 candidate，5.global and local deformation of the hair candidate at strand level.

16.[Zhang and Zheng, 2018]: 使用单张正面的肖像图片重建出可用于物理模拟的完整的头发模型。非常直接的 idea: 通过训练 single image's 2D feature(hair mask and 2D direction field) 到 volumetric 3D hair model feature(space occupancy and 3D direction field) 的映射 using GAN。of course, some needed post-processing(smooth and hair strand deformation)

17.[Nam et al., 2019]: 使用多张不同角度拍摄的头发获得的照片生成不可用于模拟的完整的头发模型。MVS->strand generation-> refine.1. 一个新的 MVS 算法（line-based PatchMatch multi-view stereo）获得更精确的点云 (每个点代表一个有向线段（包含位置和方向）)，2. 重建 hair strand from 点云，3. 利用原始的多视角 image refine hair strand，获得更长的头发。The key contribution is the high accuracy of the reconstructed hair strand but the weakness is that the hair strand is not generated from the scalp so that it can not be used in physical simulation.

18.[Yang et al., 2019]: 动态头发建模 from video by DCNN。训练两个网络，一个负责将生成静态头发的方向场，一个负责生成相邻两帧视频之间的 warping field, 数据集包含静态头发和动态模拟生成的头发，最后通过一个和 [Xu et al., 2014] 相似的时空约束优化进行 correction 获得最终的动态头发结果。其结果过于 smooth，缺少 local details.

According to the strategy difference in these papers, we can classify them into:

1.dense multi-view based static hair modeling:
a.point cloud from MVS -> fitted ribbons -> strands:[Luo et al., 2012, Luo et al., 2013, Hu et al., 2014a, Nam et al., 2019] complex hardware needed, good result
b.(data-driven)point cloud from MVS -> search database -> strands:[Hu et al., 2014b] braid，a special topic，complex hardware
c.(data-driven)2D strand similar to [Chai et al., 2016] and visual hull from SFM-> search database-> strand deform:[Liang et al., 2018] need hair database, lightweight hardware needed, complex pipeline

2.single view or sparse multi-view based static hair modeling:
a.shape from shading strategy:[Chai et al., 2015] can not be used in simulation
b.estimate 2D direction and depth from single image: [Chai et al., 2012, Chai et al., 2013] can not be used in simulation
c.(data-driven)2D image feature(direction/ segmentation) -> search(retrieval) candidate in hair database -> deform hair candidate and generate hair strand: [Hu et al., 2015, Chai et al., 2016, Zhang et al., 2017, Zhang et al., 2018] need hair database, lightweight hardware needed, can be used in simulation
d.(DNN based data-driven method):[Zhang and Zheng, 2018, Zhou et al., 2018, Saito et al., 2018] need hair database, not stable

3.dynamic hair modeling:(a special topic)
a.sparcetime constraint problem:[Xu et al., 2014]
b.DCNN:[Yang et al., 2019]

Summary of the representation of 2D information and 3D information in these papers:

2D image feature representation ： RGB color / hair segmentation mask / hair orientation field/ depth field

3D hair model representation：ribbons mesh/ strands/ uniform sample strands as matrix/ watertight mesh/ volumetric occupancy and volumetric direction / point cloud

参考文献

[Bulat and Tzimiropoulos, 2018] Bulat, A. and Tzimiropoulos, G. (2018). Super-fan: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with gans. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 109–117.

[Chai et al., 2015] Chai, M., Luo, L., Sunkavalli, K., Carr, N., Hadap, S., and Zhou, K. (2015). High-quality hair modeling from a single portrait photo. *ACM Transactions on Graphics (TOG)*, 34(6):204.

[Chai et al., 2016] Chai, M., Shao, T., Wu, H., Weng, Y., and Zhou, K. (2016). Autohair: Fully automatic hair modeling from a single image. *ACM Transactions on Graphics*, 35(4).

[Chai et al., 2013] Chai, M., Wang, L., Weng, Y., Jin, X., and Zhou, K. (2013). Dynamic hair manipulation in images and videos. *ACM Transactions on Graphics (TOG)*, 32(4):75.

[Chai et al., 2012] Chai, M., Wang, L., Weng, Y., Yu, Y., Guo, B., and Zhou, K. (2012). Single-view hair modeling for portrait manipulation. *ACM Transactions on Graphics (TOG)*, 31(4):116.

[Donato and Belongie, 2002] Donato, G. and Belongie, S. (2002). Approximate thin plate spline mappings. In *European conference on computer vision*, pages 21–31. Springer.

[Forney, 1973] Forney, G. D. (1973). The viterbi algorithm. *Proceedings of the IEEE*, 61(3):268–278.

[Fu et al., 2007] Fu, H., Wei, Y., Tai, C.-L., and Quan, L. (2007). Sketching hairstyles. In *Proceedings of the 4th Eurographics workshop on Sketch-based interfaces and modeling*, pages 31–36. ACM.

[Hu et al., 2014a] Hu, L., Ma, C., Luo, L., and Li, H. (2014a). Robust hair capture using simulated examples. *ACM Transactions on Graphics (TOG)*, 33(4):126.

[Hu et al., 2015] Hu, L., Ma, C., Luo, L., and Li, H. (2015). Single-view hair modeling using a hairstyle database. *ACM Transactions on Graphics (TOG)*, 34(4):125.

[Hu et al., 2014b] Hu, L., Ma, C., Luo, L., Wei, L.-Y., and Li, H. (2014b). Capturing braided hairstyles. *ACM Transactions on Graphics (TOG)*, 33(6):225.

[Li et al., 2018] Li, T.-M., Aittala, M., Durand, F., and Lehtinen, J. (2018). Differentiable monte carlo ray tracing through edge sampling. In *SIGGRAPH Asia 2018 Technical Papers*, page 222. ACM.

[Liang et al., 2018] Liang, S., Huang, X., Meng, X., Chen, K., Shapiro, L. G., and Kemelmacher-Shlizerman, I. (2018). Video to fully automatic 3d hair model. In *SIGGRAPH Asia 2018 Technical Papers*, page 206. ACM.

[Liu et al., 2009] Liu, J., Sun, J., and Shum, H.-Y. (2009). Paint selection. *ACM Transactions on Graphics (ToG)*, 28(3):69.

[Luo et al., 2012] Luo, L., Li, H., Paris, S., Weise, T., Pauly, M., and Rusinkiewicz, S. (2012). Multi-view hair capture using orientation fields. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1490–1497. IEEE.

[Luo et al., 2013] Luo, L., Li, H., and Rusinkiewicz, S. (2013). Structure-aware hair capture. *ACM Transactions on Graphics (TOG)*, 32(4):76.

[Nam et al., 2019] Nam, G., Wu, C., Kim, M. H., and Sheikh, Y. (2019). Strand-accurate multi-view hair capture. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 155–164.

[Paris et al., 2004] Paris, S., Briceño, H. M., and Sillion, F. X. (2004). Capture of hair geometry from multiple images. In *ACM transactions on graphics (TOG)*, volume 23, pages 712–719. ACM.

[Paysan et al., 2009] Paysan, P., Knothe, R., Amberg, B., Romdhani, S., and Vetter, T. (2009). A 3d face model for pose and illumination invariant face recognition. In *2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*, pages 296–301. Ieee.

[Saito et al., 2018] Saito, S., Hu, L., Ma, C., Ibayashi, H., Luo, L., and Li, H. (2018). 3d hair synthesis using volumetric variational autoencoders. In *SIGGRAPH Asia 2018 Technical Papers*, page 208. ACM.

[Tewari et al., 2018] Tewari, A., Zollhöfer, M., Garrido, P., Bernard, F., Kim, H., Pérez, P., and Theobalt, C. (2018). Self-supervised multi-level face model learning for monocular reconstruction at over 250 hz. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2549–2559.

[Xu et al., 2014] Xu, Z., Wu, H.-T., Wang, L., Zheng, C., Tong, X., and Qi, Y. (2014). Dynamic hair capture using spacetime optimization. *To appear in ACM TOG*, 33:6.

[Yang et al., 2019] Yang, L., Shi, Z., Zheng, Y., and Zhou, K. (2019). Dynamic hair modeling from monocular videos using deep neural networks. *ACM Transactions on Graphics (TOG)*, 36(4):156.

[Zhang et al., 2017] Zhang, M., Chai, M., Wu, H., Yang, H., and Zhou, K. (2017). A data-driven approach to four-view image-based hair modeling. *ACM Transactions on Graphics (TOG)*, 36(4):156.

[Zhang et al., 2018] Zhang, M., Wu, P., Wu, H., Weng, Y., Zheng, Y., and Zhou, K. (2018). Modeling hair from an rgb-d camera. In *SIGGRAPH Asia 2018 Technical Papers*, page 205. ACM.

[Zhang and Zheng, 2018] Zhang, M. and Zheng, Y. (2018). Hair-gans: Recovering 3d hair structure from a single image. *arXiv preprint arXiv:1811.06229*.

[Zhou et al., 2005] Zhou, K., Huang, J., Snyder, J., Liu, X., Bao, H., Guo, B., and Shum, H.-Y. (2005). Large mesh deformation using the volumetric graph laplacian. In *ACM transactions on graphics (TOG)*, volume 24, pages 496–503. ACM.

[Zhou et al., 2018] Zhou, Y., Hu, L., Xing, J., Chen, W., Kung, H.-W., Tong, X., and Li, H. (2018). Hairnet: Single-view hair reconstruction using convolutional neural networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 235–251.