

Supplemental Material

In the Supplemental Material, we present the detailed user-study results. In specific, our user study was conducted on the FF++

	User Information		
	sex	age	vision
user1	female	20	normal
user2	male	25	normal
user3	male	20	normal
user4	female	26	normal
user5	male	22	normal
user6	female	26	normal
user7	female	29	normal
user8	male	28	normal
user9	male	29	normal
user10	male	26	normal
user11	female	26	normal
user12	male	23	normal
user13	male	36	normal
user14	male	31	normal
user15	male	24	normal
user16	female	33	normal
user17	male	20	normal
user18	female	20	normal
user19	female	20	normal
user20	female	21	normal
user21	female	52	normal
user22	female	29	normal
user23	male	20	normal
user24	male	25	normal
user25	male	20	normal
user26	female	31	normal
user27	male	32	normal
user28	male	27	normal
user29	male	20	normal
user30	male	34	normal
user31	male	28	normal
user32	male	36	normal
user33	male	39	normal
user34	male	23	normal
user35	female	33	normal
user36	female	26	normal
user37	female	30	normal
user38	female	31	normal
user39	female	20	normal
user40	male	22	normal
user41	male	20	normal
user42	female	27	normal
user43	male	23	normal
user44	female	23	normal
user45	female	28	normal
Mean/Std	-	26.91 ± 6.36	-

Table 1: Users' basic information table.

	3DMM		
	ID Similarity	Att.Preservation	Naturalness
user1	3.00	2.60	2.10
user2	2.50	2.20	1.90
user3	2.60	3.00	2.10
user4	2.80	2.20	2.30
user5	2.60	2.60	2.30
user6	2.60	2.40	2.10
user7	2.60	2.70	2.80
user8	2.80	2.60	2.00
user9	2.60	2.60	2.20
user10	2.80	2.60	2.40
user11	2.40	2.50	2.30
user12	2.70	2.50	2.30
user13	2.90	2.60	2.40
user14	2.60	2.70	2.40
user15	2.50	2.40	2.30
user16	2.60	2.50	2.40
user17	2.70	2.70	2.10
user18	2.70	2.80	2.30
user19	2.60	2.60	2.50
user20	2.80	2.80	2.30
user21	2.50	2.60	2.60
user22	2.60	2.60	2.20
user23	2.70	2.40	2.50
user24	2.60	2.30	2.10
user25	2.50	2.60	2.10
user26	2.60	2.60	2.30
user27	2.70	2.60	2.20
user28	2.90	2.10	2.60
user29	2.40	2.20	2.30
user30	2.70	2.80	2.30
user31	2.50	2.50	2.40
user32	2.80	2.60	2.20
user33	2.90	2.70	2.40
user34	2.60	2.60	2.10
user35	2.90	2.60	2.30
user36	2.70	2.60	2.30
user37	2.60	2.50	2.20
user38	2.60	2.60	2.20
user39	2.70	2.80	2.20
user40	2.50	2.30	2.50
user41	3.00	2.50	2.40
user42	2.80	3.00	2.00
user43	2.70	2.60	2.30
user44	2.80	2.50	2.90
user45	2.60	2.60	2.10
Mean/Std	2.67 ± 0.56	2.56 ± 0.57	2.29 ± 0.55

Table 2: Detailed User Study's Ranking Scores of 3DMM.

dataset with 45 participants with normal vision (25 male, 20 female; aged 20 to 52, average age 26.91 ± 6.36) to further compare

the synthesised image quality of the face swapping methods. Participants were informed about the purpose of the study but not about

	FaceSwap		
	ID Similarity	Att.Preservation	Naturalness
user1	2.70	2.50	2.50
user2	3.00	2.80	2.70
user3	2.80	2.90	2.90
user4	2.90	2.80	2.70
user5	2.70	2.50	2.60
user6	2.80	2.70	2.60
user7	2.70	2.70	2.50
user8	2.60	2.40	2.70
user9	2.70	2.80	2.60
user10	2.70	2.80	2.60
user11	2.60	2.70	2.70
user12	2.70	2.50	2.50
user13	2.90	2.40	2.90
user14	2.40	2.90	2.80
user15	2.90	2.90	2.80
user16	2.80	2.80	2.70
user17	2.80	2.80	2.80
user18	2.60	2.40	2.90
user19	2.70	3.00	2.80
user20	2.70	2.70	3.00
user21	2.40	2.50	2.80
user22	2.70	2.70	2.50
user23	3.00	2.40	2.50
user24	2.30	2.80	2.60
user25	2.70	2.50	2.90
user26	2.90	2.70	2.70
user27	2.90	2.80	2.60
user28	2.70	2.40	2.90
user29	2.60	2.70	2.70
user30	2.70	2.70	2.80
user31	2.90	2.60	2.80
user32	2.40	2.90	2.60
user33	2.90	2.50	2.70
user34	2.50	2.70	2.60
user35	2.90	2.80	2.80
user36	2.60	2.80	2.60
user37	2.70	2.70	2.30
user38	2.80	2.50	2.40
user39	2.90	2.60	2.40
user40	2.80	2.40	2.50
user41	2.70	2.70	2.50
user42	2.90	2.50	2.50
user43	3.00	2.70	2.60
user44	2.70	2.80	2.80
user45	2.70	2.70	2.60
Mean/Std	2.73 ± 0.55	2.67 ± 0.56	2.67 ± 0.57

Table 3: Detailed User Study's Ranking Scores of Faceswap.

	Deepfakes		
	ID Similarity	Att.Preservation	Naturalness
user1	2.60	2.70	2.40
user2	2.80	2.70	2.70
user3	2.80	2.50	2.60
user4	2.70	2.90	2.80
user5	2.90	2.50	2.40
user6	2.70	2.80	2.80
user7	3.00	2.70	2.50
user8	2.70	2.70	2.70
user9	3.10	2.50	2.50
user10	2.60	3.00	2.60
user11	3.00	2.90	2.40
user12	2.90	2.60	2.60
user13	2.60	2.60	2.90
user14	3.00	2.70	2.90
user15	2.90	2.80	2.80
user16	2.70	2.60	2.80
user17	3.10	2.90	2.80
user18	3.10	2.90	2.70
user19	2.90	2.70	2.80
user20	2.60	2.90	2.60
user21	2.70	2.70	2.80
user22	2.90	3.10	2.80
user23	2.80	2.80	2.80
user24	3.00	2.80	2.60
user25	2.60	2.90	2.40
user26	2.90	3.10	2.80
user27	2.80	2.90	2.70
user28	3.00	2.80	2.60
user29	3.10	2.40	2.70
user30	2.80	2.80	2.90
user31	2.90	2.70	2.40
user32	2.80	2.70	2.40
user33	3.10	3.00	3.00
user34	2.90	2.60	2.50
user35	2.90	2.60	2.60
user36	2.70	2.90	2.40
user37	2.90	3.10	2.40
user38	2.90	2.50	2.20
user39	2.80	3.00	2.50
user40	2.60	2.60	2.70
user41	2.80	2.60	2.50
user42	3.10	2.90	2.60
user43	2.60	2.90	2.30
user44	2.70	2.70	2.60
user45	2.80	2.70	2.30
Mean/Std	2.84 ± 0.50	2.76 ± 0.55	2.62 ± 0.56

Table 4: Detailed User Study's Ranking Scores of Deepfakes.

any of our hypotheses. They were given image tuples randomly selected from a synthesised image pool which contains source im-

ages, target images, and the face swapping images. Note that we show 10 image tuples for each benchmark method and ours. The

	HiFiface		
	ID Similarity	Att.Preservation	Naturalness
user1	3.10	3.00	3.20
user2	2.90	2.90	3.50
user3	3.40	3.30	3.00
user4	2.80	3.20	3.40
user5	3.40	3.20	3.10
user6	3.20	2.70	3.60
user7	3.10	3.30	3.30
user8	3.30	3.10	3.10
user9	3.10	3.20	3.50
user10	3.20	3.40	3.70
user11	3.30	3.30	3.40
user12	2.80	3.10	3.30
user13	3.20	3.30	3.30
user14	3.10	3.10	3.30
user15	3.10	3.30	3.10
user16	3.40	3.00	3.20
user17	3.20	3.20	3.60
user18	3.40	3.10	3.40
user19	3.30	3.20	3.30
user20	2.90	3.10	3.20
user21	2.90	3.20	3.20
user22	3.40	3.20	3.30
user23	3.10	3.40	3.50
user24	3.00	2.80	3.60
user25	3.10	2.90	3.40
user26	3.20	3.10	3.30
user27	3.10	3.20	3.30
user28	3.20	3.20	3.40
user29	3.10	3.30	3.40
user30	3.10	3.30	3.50
user31	3.10	3.00	3.00
user32	2.90	3.20	3.40
user33	3.20	3.10	3.10
user34	2.90	3.20	3.40
user35	3.20	3.20	3.30
user36	3.20	3.20	3.30
user37	3.30	3.40	3.50
user38	3.20	3.50	3.50
user39	3.30	3.20	3.30
user40	3.10	3.40	3.70
user41	3.10	3.30	3.20
user42	3.20	3.20	3.30
user43	3.00	3.30	3.20
user44	3.10	3.20	3.10
user45	2.80	2.90	3.30
Mean/Std	3.13 ± 0.52	3.18 ± 0.54	3.33 ± 0.60

Table 5: Detailed User Study's Ranking Scores of HiFiface.

	SimSwap		
	ID Similarity	Att.Preservation	Naturalness
user1	3.30	3.30	3.50
user2	3.40	3.40	3.70
user3	3.60	3.20	3.70
user4	3.70	3.00	3.60
user5	3.70	3.60	3.60
user6	3.70	3.50	3.50
user7	3.50	3.50	3.60
user8	3.80	3.40	3.70
user9	3.80	3.30	3.60
user10	3.70	3.30	3.80
user11	3.20	3.60	3.80
user12	3.40	3.50	3.60
user13	3.70	3.70	3.40
user14	3.30	3.20	3.90
user15	3.90	3.10	3.40
user16	3.70	3.20	3.40
user17	3.40	3.70	3.60
user18	3.80	3.40	3.70
user19	3.90	3.00	3.70
user20	3.60	3.30	3.40
user21	3.60	3.20	3.80
user22	3.80	3.40	3.60
user23	3.70	3.60	3.60
user24	3.70	3.50	3.50
user25	3.50	3.20	3.80
user26	3.80	3.30	3.50
user27	3.40	3.10	3.70
user28	3.30	3.50	3.80
user29	3.60	3.50	3.70
user30	3.80	3.10	3.90
user31	3.70	3.20	3.60
user32	3.60	3.20	3.60
user33	3.80	3.40	3.50
user34	3.60	3.10	3.70
user35	3.50	3.30	3.60
user36	3.70	3.50	3.60
user37	3.60	3.60	3.60
user38	3.60	3.30	3.40
user39	3.30	3.70	3.70
user40	3.40	3.40	3.70
user41	3.50	3.20	3.80
user42	3.20	3.70	3.50
user43	3.50	3.20	3.40
user44	3.60	3.70	3.50
user45	3.40	3.20	3.50
Mean/Std	3.58 ± 0.61	3.36 ± 0.58	3.62 ± 0.57

Table 6: Detailed User Study's Ranking Scores of SimSwap.

basic information of the participants, including their gender, age, vision condition (normal or abnormal), is shown in table 1. We pro-

vide the detailed user study results in the tables 2-8, and the average scores of each individual user are presented.

	FSGANv2		
	ID Similarity	Att.Preservation	Naturalness
user1	3.30	3.40	3.80
user2	3.30	2.90	3.50
user3	3.10	3.40	3.60
user4	3.30	3.30	3.80
user5	3.30	3.60	3.80
user6	3.20	3.50	3.80
user7	3.00	3.30	3.50
user8	3.20	3.50	3.60
user9	3.30	3.30	3.60
user10	3.40	3.20	3.80
user11	3.40	3.40	3.60
user12	3.10	3.40	3.40
user13	3.00	3.20	3.80
user14	3.30	3.50	3.80
user15	3.10	3.10	3.80
user16	3.10	3.40	3.90
user17	3.10	3.50	3.80
user18	3.20	3.60	3.60
user19	3.60	3.40	3.70
user20	3.20	3.30	3.90
user21	3.20	3.50	3.70
user22	3.20	3.30	3.60
user23	3.20	3.40	3.60
user24	3.10	3.50	3.50
user25	2.90	3.50	3.80
user26	3.10	3.50	3.20
user27	3.30	3.40	3.40
user28	3.10	3.40	3.90
user29	2.90	3.30	3.80
user30	3.40	3.40	3.60
user31	3.20	3.30	3.50
user32	3.20	3.20	3.50
user33	3.20	3.70	3.80
user34	3.40	3.70	3.60
user35	3.40	3.40	4.00
user36	3.20	3.60	3.80
user37	3.60	3.30	3.70
user38	3.40	3.50	3.70
user39	3.40	3.40	3.60
user40	3.40	3.30	3.30
user41	3.50	3.30	3.70
user42	3.10	3.00	3.70
user43	3.50	3.10	4.00
user44	3.10	3.40	3.50
user45	3.10	3.40	3.60
Mean/Std	3.24 ± 0.58	3.38 ± 0.58	3.67 ± 0.54

Table 7: Detailed User Study's Ranking Scores of FSGANv2.

	Ours		
	ID Similarity	Att.Preservation	Naturalness
user1	3.90	3.50	3.90
user2	3.60	3.20	3.70
user3	3.70	3.70	4.10
user4	3.60	3.40	4.10
user5	3.80	3.60	3.90
user6	3.80	3.50	4.30
user7	3.70	3.20	3.80
user8	3.70	3.80	3.80
user9	3.70	3.50	3.80
user10	3.80	3.30	4.00
user11	3.30	3.40	3.80
user12	3.40	3.30	4.00
user13	3.60	3.50	3.80
user14	4.00	3.60	4.00
user15	3.70	3.50	3.80
user16	3.60	3.70	4.00
user17	4.00	3.40	3.80
user18	3.60	3.60	3.80
user19	4.00	3.90	3.90
user20	3.80	3.60	4.00
user21	3.90	3.40	3.80
user22	3.70	3.80	4.00
user23	3.80	3.30	4.10
user24	3.40	3.50	3.70
user25	3.30	3.50	3.90
user26	3.60	3.70	3.90
user27	3.50	3.30	3.80
user28	3.60	3.10	3.90
user29	3.60	3.70	3.80
user30	3.80	3.30	4.00
user31	3.60	3.50	4.00
user32	3.40	3.80	3.70
user33	3.70	3.50	3.70
user34	3.50	3.60	3.90
user35	3.50	3.90	3.80
user36	3.50	3.50	3.90
user37	3.70	3.40	3.70
user38	3.70	3.70	3.90
user39	3.80	3.40	4.10
user40	3.80	3.60	3.80
user41	3.80	3.40	3.80
user42	3.70	3.10	3.80
user43	3.60	3.90	3.90
user44	3.70	3.50	4.00
user45	3.80	3.40	3.90
Mean/Std	3.67 ± 0.56	3.51 ± 0.58	3.89 ± 0.43

Table 8: Detailed User Study's Ranking Scores of Our Method.



Figure 1: Additional qualitative results. The third column represents the results constructed by 3DMM, the fourth column represents the results generated without guidance, the fifth column represents the results guided by identity only, the sixth column represents the results guided by both identity and expression, and the seventh column represents the results guided by identity, expression, and pose. The guidance effectively embeds identity information of source faces and preserves other attributes of target faces in the face swapping results.

Further, we illustrate more diverse qualitative results to demon-

strate the superiority of our method by using the two-stage condi-

tional guided DDPM framework. The additional qualitative results are shown in Figure 1.

As is evident, our designed modules, including the image condition, the identity guidance, the expression guidance and the pose guidance, work effectively on achieving high-quality face image synthesis.