Appendix

A Animal Motion Quantative Evaluation

We evaluated our Motion Avatar on the Zoo-300K dataset, and the results are presented in table 2. The comprehensive evaluation demonstrated that our method achieves high-quality animal motion generation. The results indicate that our approach produces highly realistic and promising animations, highlighting the effectiveness and potential of our technique in generating detailed and lifelike animal motions.

B User Study

In this work, we conduct an exhaustive evaluation of the effectiveness of our motion avatar generation, leveraging both qualitative and quantitative assessments. This user study assesses the real-world applicability of 4 motion video generated from the Motion Avatar platform by the input prompt, examined by 50 participants through a Google Forms interface as in Fig 5.

Participants were presented with four videos labeled Video A, Video B, Video C, and Video D. Each video showcased a unique motion generated using Motion Avatar with different input prompt. The participants evaluated these videos by responding to a series of targeted questions aimed at assessing the motion's accuracy, the mesh's visual quality, the integration of motion and mesh, and their overall emotional response to the animations.

The evaluation was structured around several key aspects:

1. Motion Accuracy: Participants rated the naturalness and accuracy of the motions on a scale from 1 ('Very Inaccurate') to 5 ('Very Natural'). The average score was 4.2, indicating a high fidelity in motion portrayal.

2. Mesh Quality: The visual quality and detail of the mesh were rated from 1 ('Poor Quality') to 5 ('Excellent Quality'), with an average score of 4.0, highlighting the superior visual appeal of our models.

3. Motion and Mesh Integration: The integration of motion and mesh was assessed, with most participants rating this aspect a 4.5 on average, reflecting seamless integration that enhances fluidity and realism.

4. User Engagement and Appeal: Participants reflected on their feelings towards the animations, rating their overall engagement and appeal from 1 ('Not Engaging') to 5 ('Highly Engaging'). The average engagement score was 4.3, suggesting that the animations were highly engaging and appealing to the audience.

Results suggested that:

- 92% of participants believed the animations could be directly utilized in real-world applications without significant modifications.

- Only 8% felt that minor adjustments were necessary before deployment.

These findings underscore the high quality, impressive results, and broad usability of the animations. Nearly all participants found the quality of the generated videos to be very high,

Table 2: The table shows that our **Motion Avatar** achieves high-quality animal motion generation on the Zoo-300K dataset. It demonstrates the effectiveness of our method, producing realistic and promising animal motions.

Method	R-Prec Top 1 ↑	R-Prec Top 2 ↑	R-Prec Top 3 ↑	FID ↓	MultiModal-Dist ↓	Diversity →
Anaconda (GT)	0.400	0.591	0.713	0.015	2.870	10.233
Anaconda (Ours)	**0.042**	**0.080**	**0.128**	**0.51067**	**9.053**	**5.165**
Ant (GT)	0.413	0.650	0.768	0.12	2.000	9.700
Ant (Ours)	**0.056**	**0.110**	**0.166**	**68.220**	**9.380**	**4.380**
Bat (GT)	0.406	0.677	0.841	0.011	1.805	12.084
Bat (Ours)	**0.193**	**0.536**	**0.086**	**79.403**	**9.430**	**0.984**
Bear (GT)	0.641	0.851	0.916	0.012	1.827	13.587
Bear (Ours)	**0.156**	**0.287**	**0.382**	**35.735**	**8.890**	**4.364**
Bird (GT)	0.569	0.762	0.851	0.004	2.679	12.294
Bird (Ours)	**0.059**	**0.091**	**0.127**	**54.735**	**9.726**	**4.389**
Buffalo (GT)	0.547	0.750	0.852	0.014	3.876	16.405
Buffalo (Ours)	**0.070**	**0.140**	**0.227**	**167.900**	**13.378**	**4.352**
Buzzard (GT)	0.423	0.639	0.778	0.007	2.654	11.045
Buzzard (Ours)	**0.034**	**0.084**	**0.111**	**65.274**	**8.645**	**1.824**
Camel (GT)	0.297	0.469	0.609	0.014	3.381	11.025
Camel (Ours)	**0.078**	**0.188**	**0.250**	**92.963**	**11.654**	**5.249**
Cat (GT)	0.141	0.281	0.422	0.041	2.274	9.060
Cat (Ours)	**0.063**	**0.188**	**0.219**	**56.405**	**7.195**	**1.969**
Centipede (GT)	0.386	0.580	0.712	0.026	2.256	11.279
Centipede (Ours)	**0.123**	**0.223**	**0.317**	**65.076**	**8.590**	**5.664**
Chicken (GT)	0.094	0.203	0.313	0.075	3.503	9.196
Chicken (Ours)	**0.063**	**0.141**	**0.172**	**70.271**	**7.953**	**2.752**
Cobra (GT)	0.386	0.583	0.700	0.009	2.881	10.901
Cobra (Ours)	**0.217**	**0.324**	**0.406**	**0.309**	**6.397**	**10.823**
Komodo (GT)	0.260	0.375	0.456	0.064	5.208	9.968
Komodo (Ours)	**0.073**	**0.125**	**0.161**	**43.677**	**7.074**	**3.047**
Coyote (GT)	0.523	0.765	0.869	0.034	2.042	12.985
Coyote (Ours)	**0.052**	**0.102**	**0.158**	**89.865**	**9.834**	**2.396**
Crab (GT)	0.367	0.570	0.698	0.019	2.373	12.260
Crab (Ours)	**0.066**	**0.109**	**0.175**	**79.233**	**9.862**	**2.737**
Cricket (GT)	0.429	0.658	0.777	0.003	2.015	14.790
Cricket (Ours)	**0.036**	**0.075**	**0.111**	**111.337**	**12.126**	**3.810**
Crocodile (GT)	0.567	0.879	0.891	0.016	2.177	13.694
Crocodile (Ours)	**0.054**	**0.094**	**0.133**	**86.886**	**10.049**	**2.827**
Crow (GT)	0.390	0.644	0.796	0.018	2.682	12.131
Crow (Ours)	**0.034**	**0.058**	**0.079**	**61.148**	**10.915**	**5.021**
Deer (GT)	0.497	0.722	0.816	0.010	2.700	11.116
Deer (Ours)	**0.092**	**0.174**	**0.231**	**12.931**	**8.106**	**8.577**
Dog (GT)	0.609	0.850	0.940	0.025	2.001	15.927
Dog (Ours)	**0.039**	**0.086**	**0.111**	**80.711**	**12.049**	**4.249**
Eagle (GT)	0.551	0.788	0.896	0.003	1.607	14.266
Eagle (Ours)	**0.057**	**0.121**	**0.169**	**65.728**	**10.753**	**4.716**
Elephant (GT)	0.544	0.750	0.851	0.006	2.331	14.017
Elephant (Ours)	**0.044**	**0.091**	**0.136**	**91.635**	**10.651**	**2.837**
Fire Ant (GT)	0.399	0.587	0.692	0.004	3.602	9.441
Fire Ant (Ours)	**0.027**	**0.055**	**0.081**	**48.101**	**7.988**	**2.065**
Flamingo (GT)	0.078	0.250	0.375	0.023	2.971	8.589
Flamingo (Ours)	**0.031**	**0.062**	**0.109**	**47.836**	**8.396**	**3.119**
Fox (GT)	0.380	0.566	0.723	0.009	2.807	11.209
Fox (Ours)	**0.061**	**0.128**	**0.179**	**30.368**	**9.234**	**7.117**

Table 3: Continued from Table 2.

Method	R-Prec Top 1 ↑	R-Prec Top 2 ↑	R-Prec Top 3 ↑	FID ↓	MultiModal-Dist ↓	Diversity →
Gazelle (GT)	0.462	0.700	0.825	0.024	2.401	12.907
Gazelle (Ours)	**0.019**	**0.041**	**0.075**	**95.267**	**11.966**	**4.100**
Giant Bee (GT)	0.353	0.555	0.703	0.029	2.700	12.992
Giant Bee (Ours)	**0.097**	**0.167**	**0.234**	**63.874**	**9.436**	**5.031**
Goat (GT)	0.394	0.663	0.825	0.013	1.975	11.550
Goat (Ours)	**0.033**	**0.064**	**0.094**	**102.515**	**10.767**	**2.069**
Hamster (GT)	0.297	0.531	0.672	0.039	2.251	12.088
Hamster (Ours)	**0.052**	**0.104**	**0.141**	**108.243**	**10.693**	**2.000**
Hermit Crab (GT)	0.439	0.680	0.811	0.009	2.677	10.232
Hermit Crab (Ours)	**0.046**	**0.085**	**0.113**	**45.140**	**8.103**	**2.704**
Hippopotamus (GT)	0.542	0.773	0.879	0.044	2.201	11.970
Hippopotamus (Ours)	**0.052**	**0.102**	**0.139**	**86.651**	**9.931**	**1.684**
Horse (GT)	0.518	0.686	0.779	0.006	2.816	10.935
Horse (Ours)	**0.053**	**0.090**	**0.137**	**31.290**	**9.131**	**6.203**
Hound (GT)	0.639	0.822	0.909	0.014	2.519	11.514
Hound (Ours)	**0.047**	**0.107**	**0.146**	**68.021**	**9.198**	**2.529**
Isopetra (GT)	0.498	0.708	0.824	0.020	2.710	11.242
Isoptera (Ours)	**0.035**	**0.088**	**0.127**	**56.617**	**9.598**	**3.903**
Jaguar (GT)	0.538	0.762	0.871	0.009	2.354	11.287
Jaguar (Ours)	**0.043**	**0.074**	**0.105**	**71.047**	**9.314**	**2.442**
Leopard (GT)	0.558	0.758	0.884	0.017	1.989	13.777
Leopard (Ours)	**0.058**	**0.115**	**0.152**	**67.369**	**9.686**	**4.035**
Lion (GT)	0.542	0.756	0.881	0.010	2.357	11.772
Lion (Ours)	**0.042**	**0.066**	**0.107**	**79.524**	**9.760**	**2.419**
Lynx (GT)	0.526	0.751	0.869	0.019	2.158	12.317
Lynx (Ours)	**0.059**	**0.096**	**0.140**	**77.358**	**9.666**	**2.858**
Mammoth (GT)	0.551	0.775	0.860	0.004	2.518	12.963
Mammoth (Ours)	**0.055**	**0.112**	**0.155**	**80.510**	**9.973**	**2.614**
Monkey (GT)	0.361	0.551	0.683	0.015	3.425	11.441
Monkey (Ours)	**0.053**	**0.082**	**0.118**	**31.473**	**8.819**	**2.880**
Ostrich (GT)	0.508	0.691	0.817	0.011	2.507	12.145
Ostrich (Ours)	**0.084**	**0.145**	**0.189**	**59.510**	**9.215**	**3.988**
Parrot (GT)	0.524	0.754	0.859	0.040	2.091	14.021
Parrot (Ours)	**0.068**	**0.149**	**0.202**	**27.881**	**10.990**	**9.306**
Pigeon (GT)	0.431	0.670	0.810	0.047	2.151	11.879
Pigeon (Ours)	**0.208**	**0.328**	**0.426**	**0.431**	**7.187**	**11.446**
Piranha (GT)	0.287	0.499	0.684	0.010	2.262	11.645
Piranha (Ours)	**0.038**	**0.069**	**0.096**	**81.586**	**9.930**	**1.719**
Polar Bear (GT)	0.519	0.746	0.867	0.024	2.162	12.509
Polar Bear (Ours)	**0.031**	**0.065**	**0.102**	**112.532**	**11.160**	**3.155**
Pteranodon (GT)	0.510	0.730	0.815	0.021	2.704	10.532
Pteranodon (Ours)	**0.037**	**0.076**	**0.121**	**26.197**	**8.329**	**1.424**
Puppy (GT)	0.344	0.563	0.729	0.036	1.421	12.181
Puppy (Ours)	**0.094**	**0.240**	**0.292**	**55.135**	**8.730**	**5.339**
Raptor (GT)	0.531	0.781	0.881	0.066	1.896	13.216
Raptor (Ours)	**0.077**	**0.151**	**0.216**	**105.263**	**10.961**	**4.023**
Rat (GT)	0.396	0.639	0.802	0.057	2.609	11.625
Rat (Ours)	**0.243**	**0.385**	**0.482**	**0.495**	**6.247**	**11.592**
Reindeer (GT)	0.580	0.833	0.918	0.020	2.177	14.077
Reindeer (Ours)	**0.043**	**0.073**	**0.106**	**130.645**	**11.970**	**2.953**

Table 4: Continued from Table 3.

Method	R-Prec Top 1 ↑	R-Prec Top 2 ↑	R-Prec Top 3 ↑	FID ↓	MultiModal-Dist ↓	Diversity →
Rhino (GT)	0.349	0.575	0.690	0.031	2.971	10.070
Rhino (Ours)	**0.055**	**0.106**	**0.147**	**60.834**	**8.572**	**3.423**
Roach (GT)	0.516	0.693	0.828	0.014	2.456	12.001
Roach (Ours)	**0.047**	**0.120**	**0.182**	**56.589**	**9.980**	**5.248**
Sabre-toothed tiger (GT)	0.625	0.795	0.867	0.029	2.804	10.963
Sabre-toothed Tiger (Ours)	**0.035**	**0.071**	**0.098**	**18.381**	**8.898**	**5.478**
Sand mouse (GT)	0.373	0.595	0.742	0.007	2.878	11.316
Sand Mouse (Ours)	**0.035**	**0.067**	**0.118**	**65.683**	**9.128**	**3.479**
Scorpion (GT)	0.548	0.772	0.888	0.022	1.702	12.928
Scorpion (Ours)	**0.048**	**0.809**	**0.108**	**63.289**	**10.349**	**4.709**
Shark (GT)	0.412	0.627	0.760	0.022	3.118	9.744
Shark (Ours)	**0.210**	**0.335**	**0.387**	**0.568**	**7.090**	**9.415**
Skunk (GT)	0.484	0.712	0.786	0.021	2.615	11.999
Skunk (Ours)	**0.022**	**0.056**	**0.094**	**81.971**	**10.250**	**3.155**
Spider (GT)	0.373	0.618	0.752	0.007	2.279	12.354
Spider (Ours)	**0.060**	**0.103**	**0.150**	**81.980**	**10.375**	**2.722**
Stegosaurus (GT)	0.371	0.573	0.706	0.049	2.976	10.145
Stegosaurus (Ours)	**0.042**	**0.073**	**0.113**	**59.360**	**7.915**	**1.001**
T-Rex (GT)	0.125	0.250	0.344	0.163	4.367	5.915
T-Rex (Ours)	**0.000**	**0.031**	**0.063**	**31.338**	**7.166**	**2.823**
Tricera (GT)	0.316	0.502	0.625	0.004	3.626	10.532
Tricera (Ours)	**0.047**	**0.090**	**0.137**	**62.515**	**8.280**	**1.059**
Toucan (GT)	0.489	0.723	0.842	0.012	2.457	13.668
Toucan (Ours)	**0.205**	**0.364**	**0.464**	**0.915**	**7.441**	**12.437**
Turtle (GT)	0.426	0.665	0.790	0.016	2.266	11.510
Turtle (Ours)	**0.036**	**0.078**	**0.123**	**64.208**	**8.913**	**2.299**
Tyrannosaurus Rex (GT)	0.425	0.681	0.799	0.007	1.606	11.587
Tyrannosaurus Rex (Ours)	**0.054**	**0.087**	**0.144**	**68.680**	**8.512**	**1.638**
Wyvern (GT)	0.424	0.643	0.750	0.077	2.693	11.574
Wyvern (Ours)	**0.098**	**0.196**	**0.290**	**43.339**	**8.184**	**4.517**
Average (GT)	0.437	0.650	0.767	0.024	2.559	11.790
Average (Ours)	**0.069**	**0.144**	**0.173**	**62.785**	**9.371**	**4.158**

with no significant criticisms regarding their quality. The results were notably impressive, with many participants expressing enthusiasm for widespread use of these animations.

This comprehensive user study confirms that our animations not only meet but exceed user expectations in terms of quality, realism, and engagement, making them highly suitable for varied practical applications. The insights from this study will guide further enhancements to ensure our animation generation remains at the forefront of technological and artistic innovation.

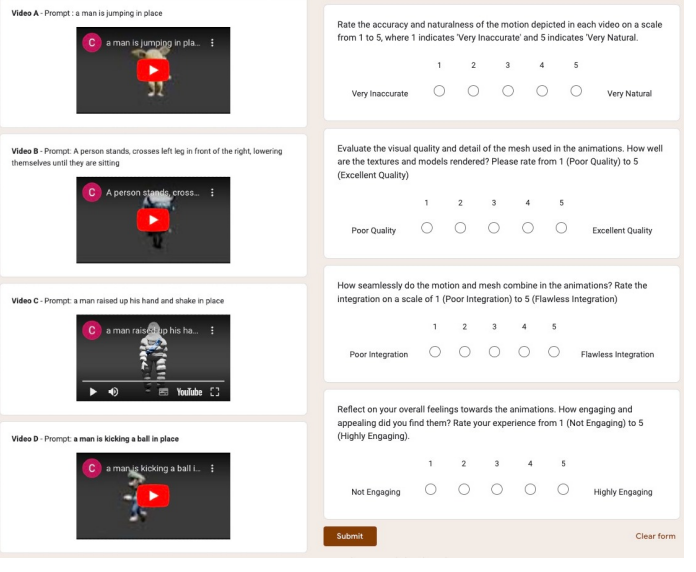

Figure 5: This figure displays the User Interface (UI) used in our User Study, showcasing four videos (Video A to D) each with distinct motion animations from various models. Participants evaluate these animations on aspects such as motion accuracy, mesh quality, integration of motion and mesh, and overall user experience. They rate each aspect from 1 (low) to 5 (high) to assess how the animations mirror real-world movements, the visual appeal of the animations, their integration quality, and their engagement level. This evaluation aims to determine the realism and engagement effectiveness of each animation model.

C Avatar Q&A Dataset Example


```
[
    {
        "instruction": "I saw an animal attacking something. A closer look makes it clear that
         a monkey is attacking",
        "input": "",
        "output": "The animal is {Monkey}, and motion is {Attack}.",
        "history": []
    },
    {
        "instruction": "A chicken walked quickly from my line of sight",
        "input": "",
        "output": "The animal is {Chicken}, and motion is {Walk Quick}.",
        "history": []
    },
    {
        "instruction": "A fox, seeing the food ahead, lowers its body and slowly approaches
        its prey",
        "input": "",
        "output": "The animal is {Fox}, and motion is {Low Bite}.",
        "history": []
    },
    {
        "instruction": "A fox walked out of the woods.",
        "input": "",
        "output": "The animal is {Fox}, and motion is {Walk Out}.",
        "history": []
    },
    {
        "instruction": "The rabbit hopped across the meadow, its fluffy tail bouncing in the
        sunlight.",
        "input": "",
        "output": "The animal is {Rabbit}, and motion is {Hop}.",
        "history": []
    }
]
```


Listing 1: Examples from the **Avatar Q&A Dataset**