



# Duet Dance Motion Dataset Annotation

Welcome, 

[Logout](#)

Please select the genre you want to annotate:

[Fox Trot](#)

[West Coast Swing](#)

[Cha Cha](#)

[Salsa](#)

[Samba](#)

[Jive](#)

[Rumba](#)

[Merengue](#)

[Paso Doble](#)

[Quick Step](#)

[Argentine Tango](#)

[Tango](#)

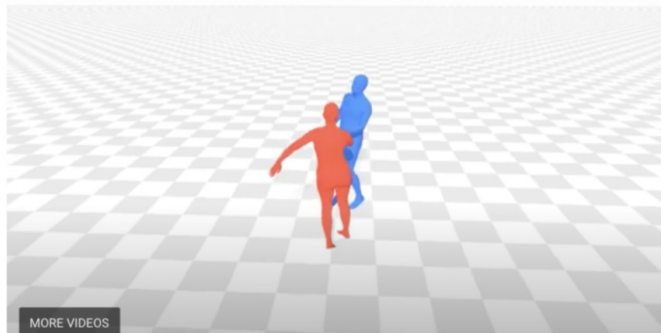
[Sensual Bachata](#)

[Traditional Bachata](#)

You selected: Jive

JV M10 F13 012

0 people have annotated this video



[MORE VIDEOS](#)

0:08 / 0:41

Your Previous Annotations

No previous annotations found

[Previous](#)

[Next](#)

## Annotation

Enter your annotation here...

Start Time:

End Time:

[Get Current Time](#)

[Go to End](#)

[Add Annotation](#)

[Submit All](#)

[Clear All](#)

## Pending Annotations

0:00.000 - 0:05.696

The dancers do a basic Jive move in place rhythmically as the song starts



Figure 4. Annotation Tool

## C. LLM-based Annotation Refinement

We employed a structured prompting framework to generate refined annotations using an LLM. This process included:

1. **System Prompt:** The model is instructed with the following directive: *“Below is a list of human-written annotations for a {genre} duet dance routine between a lead and a follow. As a professional dance movement analyst, refine each segment to be independently coherent, descriptive, and under 75 words. Avoid relying on prior context—each annotation must clearly describe the movement of both partners. Use the delimiter ‘delimiter’ to separate annotations, and omit the original timestamps. Follow the Annotation Guidelines {D1}, use Duet Dance Terminology & Vocabulary {D2}, and refer to Genre-specific Move Descriptions {D3} and provided examples {D4}.”*
2. **Annotation Guidelines (D1):** A detailed set of instructions to ensure consistency between raw and refined annotations.
3. **Duet Dance Terminology (D2):** A reference document listing essential duet dance terms and structured keyword clusters organized by spatial relations, body movement, and rhythm.
4. **Genre-specific Move Descriptions (D3):** A curated compilation of genre-wise dance moves and their execution, sourced from reliable online references.
5. **Few-shot Examples (D4):** Three hand-crafted examples per genre demonstrating how raw descriptions are refined, based on expert analysis of diverse samples from the MDD dataset.

We also use GPT-4o to extract move names from each refined annotation using the following prompt:

*“You are a professional dance movement analyst. For each of the following refined duet dance annotations, extract a concise key dance move name mentioned or implied. If a movement corresponds to a known dance move (e.g., Underarm Turn, Hammerlock, Cross Body Lead), prefer to use that name. Use domain-specific terminology from here: {D2} and Genre-specific Move Descriptions{D3} Avoid full-sentence descriptions. Separate the move names for each annotation using a delimiter: ‘delimiter’. Do not include timestamps or any additional commentary.”*

We have included some examples for the LLM-processed descriptions and LLM-processed move names in Tab. 1

## D. Ablation Study

Tab. 3 shows the ablation results of InterGen Model for the Text-to-Duet task under different text conditions. The InterGen model trained on GPT-refined prompts outperforms

both raw text and move-name versions, highlighting its effectiveness.

## E. Qualitative Visualization

We visualize one of the samples from each genre in each category: Latin (Fig. 5), Ballroom (Fig. 6) and Social (Fig. 7).

## F. Generalization Experiments

We design a set of qualitative experiments to evaluate the model’s ability to generalize across varying text prompts and dance genres. Specifically, we assess how well the model captures stylistic diversity within a genre (intra-genre generalization) and adapts to stylistic shifts across genres (inter-genre generalization), under controlled text and music variations. For each experiment, we do a paired comparison between two generated sample controlled by the experimental condition.

We examine the following scenarios:

1. **Intra-genre Generalization:** These experiments test the model’s ability to generate diverse motions within the same genre when conditioned on different textual prompts, with music held constant.
  - (a) **Same music, different text:** We analyze how semantically distinct text prompts influence the resulting motions within the same genre and music context. This illustrates the model’s capacity to produce choreographic diversity while maintaining genre fidelity.
  - (b) **Same music, subtly different text:** We evaluate the model’s sensitivity to fine-grained textual nuances—such as tempo cues, directional hints, or partner interaction specifics—demonstrating its precision in interpreting subtle prompt variations.
2. **Inter-genre Generalization:** These experiments evaluate how well the model generalizes motion generation across different dance genres while controlling for text or music.
  - (a) **Same music, different text from different genres:** We investigate how the model adapts to genre and prompt changes while maintaining synchronization with the same music. Here, the text prompts are sampled from different genres leading to prompting the model with move descriptions that were not present in the training data for that genre. This highlights genre-aware style transfer and robustness to stylistic shifts.
  - (b) **Different music, same text:** We assess how genre conditioning influences the interpretation of an identical text prompt when paired with genre-specific music. Testing this shows how the same move can be adapted to music belonging to different

Table 1. Some examples showing the LLM processed text descriptions and LLM processed move names for raw text annotations belonging to different genres

Genre	Original Description	LLM-processed description	LLM-processed Move Name
Jive	fallaway throwaway from closed, maintaining body contact, shifting back fro controlled release into open hold fluidly. Also, their left-to-right handhold goes through a synced transition	The leader initiates a Fallaway Throwaway from a Closed Position, maintaining body contact while shifting weight back for a controlled release, guiding the follower into an Open Hold with a fluid extension, while their left-to-right hand connection ensures a synced transition.	Fallaway Throwaway
Rumba	Leader faces follower in separated position facing each other, both style their arms, leader reconnects in open double handhold leading Open-to-Fan move after which they raise their arms in left-to-right handhold	With independent arm styling in a separated position while facing each other, the leader reconnects with the follower in an open double handhold, smoothly guiding her to his left executing Open-to-Fan move, after which they raise their free arms being in left-to-right handhold	Open-to-Fan
Foxtrot	leader steps back from right in controlled QQ rhythm, they both go back synchronously in closed hold	The leader initiates the Back Feather, stepping backward from the right foot with a controlled yet dynamic Quick-Quick rhythm, maintaining a good flow as both partners travel backward synchronously in closed position.	Back feather
Waltz	both the dancers are back to back in shadow, follower in shadow. They both (follower mirrors) go into c/w rotation ending with descent	The dancers gracefully execute a Waterfall, with the follower positioned in Shadow Position, seamlessly mirroring the leader’s movement as they both ascend into an elegant clockwise rotation with a suspended, flowing descent.	Waterfall
Sensual Bachata	Reverse turn: leader rotates the follower until she faces him and they both enter cross handhold	The leader initiates a Reverse turn, smoothly guiding the follower to rotate until she faces him, transitioning into a cross-handhold connection.	Reverse turn

dance genres. This reveals how genre knowledge modulates movement vocabulary and rhythmic interpretation.

Examples from each scenario, generated by our adapted version of InterGen for the Text-to-Duet task, are showcased in the supplementary video. We observe that InterGen generally adheres well to the provided text prompts and music in intra-genre settings, demonstrating consistency and responsiveness to prompt variation. In inter-genre scenarios, the model exhibits a degree of stylistic adaptivity, though genre-specific fidelity may vary. Future work could focus on enhancing cross-genre generalization and enabling the synthesis of novel or unseen movement styles.



Category	Genre	Time (min)	$\bar{T}$	$\hat{T}$	$T$	Annotations
Ballroom	Waltz (WZ)	61.87	151.78	7.70	302.48	906
	Tango (TG)	32.80	89.45	6.71	192.90	554
	Foxtrot (FT)	31.80	83.68	9.14	200.91	548
	Quickstep (QS)	32.30	69.78	8.64	190.34	551
Latin	Rumba (RB)	38.83	131.92	9.51	242.52	672
	Cha Cha (CC)	32.82	99.91	5.08	260.58	574
	Samba (SB)	34.20	90.94	2.28	149.80	603
	Paso Doble (PD)	32.42	96.07	3.67	140.67	562
	Jive (JV)	33.21	96.60	3.49	148.13	581
Social	Salsa (SS)	56.49	58.08	3.13	217.86	992
	Traditional Bachata (TB)	41.38	119.86	2.96	221.77	712
	Sensual Bachata (ST)	43.52	128.46	7.23	208.34	762
	Merengue (MR)	36.39	92.36	3.16	229.73	642
	Argentine Tango (AT)	61.87	91.03	6.47	272.92	634
	West Coast Swing (WC)	50.70	97.40	3.31	120.14	894
<b>Total</b>		<b>620.60</b>				<b>10,187</b>

Table 2. Genre-wise data distribution in the dataset, showing the total duration and the number of descriptions for each dance style.

Methods	R-Precision $\uparrow$			FID $\downarrow$	MM Dist $\downarrow$	Diversity $\rightarrow$	MModality $\uparrow$	BED $\uparrow$	BAS $\uparrow$
	Top 1	Top 2	Top 3						
Ground Truth	0.231	0.3984	0.5219	0.065	0.0770	1.387	-	0.327	0.170
InterGen (no text, music)	0.023	0.067	0.088	2.014	2.526	1.300	<b>1.768</b>	0.364	0.163
InterGen (move name, music)	0.061	0.173	0.216	0.721	2.211	1.288	1.521	0.355	<b>0.192</b>
InterGen (raw text, music)	0.911	0.192	0.244	0.511	1.722	<b>1.392</b>	1.411	0.381	0.190
InterGen (GPT-4o, music)	<b>0.105</b>	<b>0.206</b>	<b>0.302</b>	<b>0.426</b>	<b>1.532</b>	1.380	1.352	<b>0.385</b>	0.185

Table 3. Text Ablation study for InterGen (Text-to-Duet task)

## MDD: A Dataset for Text-and-Music Conditioned Duet Dance Generation

Jive	 <p>The dancers execute an <b>Overturned Change of Place</b> from <b>Left to Right</b>, maintaining their <b>left-to-right hand connection</b> as the follower <b>rotates</b> under the leader's arm, ensuring a <b>smooth directional shift</b> and <b>rhythmic flow</b> with the beat.</p>
Rumba	 <p>The leader <b>pulls</b> the follower in from an <b>Open Fan Position</b>, <b>seamlessly</b> executing <b>Open-to-sweetheart transition</b> with lead's <b>hand connection change</b> from left to right, <b>guiding</b> the follower <b>perpendicularly</b> into a <b>Closed Hold</b>, while <b>supporting</b> his free left hand on her shoulder blade to establish a strong yet <b>fluid frame</b>.</p>
Cha Cha	 <p>The leader <b>connects</b> with the follower's <b>left hand wrist</b> from his right hand, then <b>swiftly</b> executes a <b>sharp New Yorker to Free Spin</b>, using the same <b>right-to-left handhold</b> to <b>rapidly</b> toss her into an <b>outside turn</b>.</p>
Samba	 <p>The leader executes multiple <b>Promenade Runs</b>, dynamically <b>connecting</b> between the <b>left and right sides</b> of the follower, maintaining good speed ensuring the <b>high-energy Samba flow</b>.</p>
Passo Doble	 <p>The leader <b>pulls</b> the follower <b>backward</b> from an <b>Open Position</b> into <b>Closed Hold</b>, then initiates a <b>Clockwise rotation of the frame</b> from the right foot, ending it by <b>sharply</b> stepping onto the left foot with a <b>bent knee</b>, maintaining an <b>Extended Closed Frame</b> with bold, flamenco-infused precision.</p>

Figure 5. Samples visualized for each Latin Category





Waltz	 <p>The leader initiates a <b>Natural Turn</b>, stepping forward with the right foot in Closed Position, guiding the follower into a smooth clockwise rotation while maintaining a strong frame; as the turn progresses, both partners execute a rise on the first beat, a controlled side step on the second beat, and a graceful lowering action on the third beat, repeating on second phrase too for a complete turn.</p>
Tango	 <p>The leader pulls the follower inward from a right-to-right handhold, smoothly leading her into a left-side outside turn, after which they seamlessly transition into an <b>Open Fan</b> position in a left-to-right handhold, both dancers extending their free arms outward in an expressive finish.</p>
Fox Trot	 <p>The leader smoothly executes <b>Progressive Chassé to Right</b> by elevating onto their left foot before stepping forward with the right, initiating a continuous flow, while the follower reciprocates by stepping back with the left and then doing a Chassé, both maintaining a fluid Closed Hold, transitioning through controlled rise and fall dynamics, executing a Slow Quick Quick (SQQ) rhythm.</p>
Quick Step	 <p>The leader executes a <b>Progressive Chassé to Left</b>, stepping outside the follower on his left side while maintaining a closed embrace, smoothly transitioning into a forward-moving Chassé in a Quick-Quick rhythm, ensuring the dynamic flow characteristic of Quickstep.</p>

Figure 6. Samples visualized for each Ballroom Category


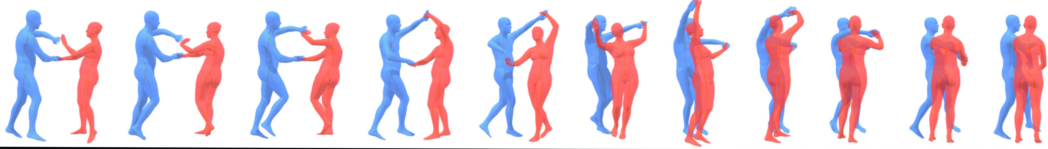




Salsa	 <p>In an open-handhold, the leader swings both of the follower's hands towards his left side, leading her into a <b>Right turn</b> as she quickly transitions into a catapult handhold position, with her right arm pointing upwards.</p>
Traditional Bachata	 <p>In a left-over-right cross handhold, the leader initiates a <b>Follower's Left-side turn</b>, then smoothly guides her to his side, transitioning into a left-to-right arm connection in a couple handhold.</p>
Sensual Bachata	 <p>The leader guides the follower's left arm around her neck for a hair caress, placing her right arm on his right shoulder, before transitioning into a Left-over-right cross handhold, then smoothly leading her into a <b>Right turn</b>, opening into a perpendicular position with a fluid transition.</p>
Merengue	 <p>From an open two-hand connection, the leader gently pulls the follower toward his left side, guiding her to pass across smoothly, before turning to face her, doing a <b>Position exchange</b> that concludes with a looped handhold.</p>
Argentine Tango	 <p>The dancers remain in a closed embrace, where the leader initiates a <b>Forward-backward swinging</b> move with the left leg, establishing a subtle yet expressive cadence while the follower attentively mirrors the movement, maintaining a delicate balance between anticipation and responsiveness with the rhythm of the music.</p>
West Coast Swing	 <p>The dancers begin in an open position, where the leader initiates an <b>Open Whip</b>, maintaining a forward-facing stance while directing the follower into a controlled outside turn, ensuring a steady frame and fluid motion.</p>

Figure 7. Samples visualized for each Social Category