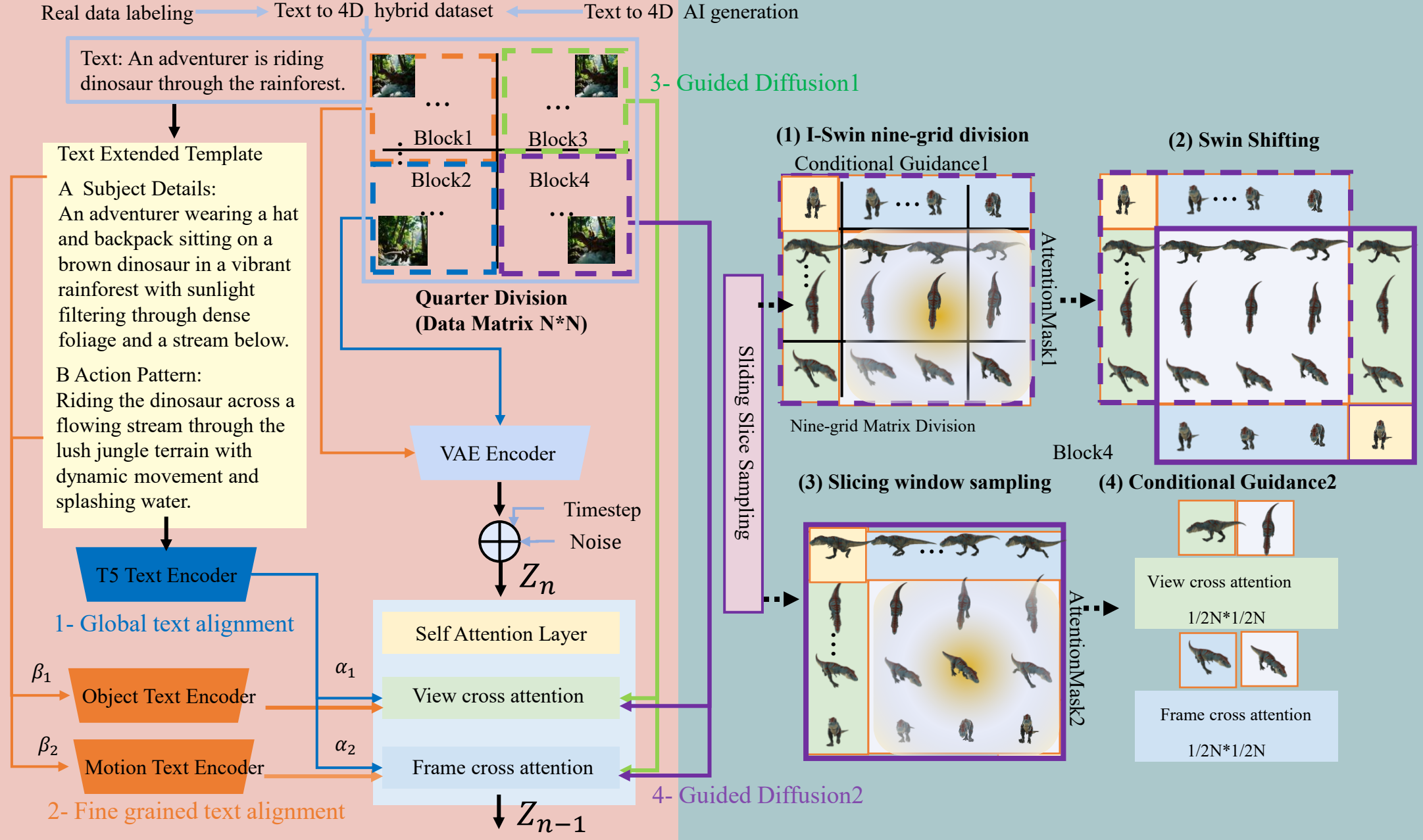
Real data labeling
Text to 4D hybrid dataset
Text to 4D AI generation
Text: An adventurer is riding dinosaur through the rainforest.
Text Extended Template
A Subject Details:
An adventurer wearing a hat and backpack sitting on a brown dinosaur in a vibrant rainforest with sunlight filtering through dense foliage and a stream below.
B Action Pattern:
Riding the dinosaur across a flowing stream through the lush jungle terrain with dynamic movement and splashing water.
T5 Text Encoder
1- Global text alignment
Block1
Block2
Block3
Block4
Quarter Division
(Data Matrix N*N)
3- Guided Diffusion1
VAE Encoder
Timestep
Noise
Z_n
Sliding Slice Sampling
(1) I-Swin nine-grid division
Conditional Guidance1
AttentionMask1
Nine-grid Matrix Division
(2) Swin Shifting
Block4
(3) Slicing window sampling
AttentionMask2
(4) Conditional Guidance2
View cross attention
1/2N*1/2N
Frame cross attention
1/2N*1/2N
β_1
Object Text Encoder
β_2
Motion Text Encoder
α_1
α_2
Self Attention Layer
View cross attention
Frame cross attention
2- Fine grained text alignment
Z_{n-1}
4- Guided Diffusion2