Output:

Short form content

10 sec - 30 sec --- video

Input

* Zeroshot (not starting from any guidance)
* **Text guidance (news article) -- generation**
* Video guidance (video guidance) -- refinement

Setup

* Plots and stories:
  + Plot description and script (Text form)
  + Intermediate steps:
    - Plot DSL (Domain Specific Language) Generation
      * LLM based plot DSL generator
    - Visual Object Generation (Scene -> Visual) Image Gen
      * Theme (real, animated, pastels, toys)
      * Audience, sub aspects
    - Summarizer (Plot -> 30 sec plot -> pick visuals)
      * Conditioned on query
    - Animator (Connects visuals and scenes through animations)
      * Animation DSL
    - Narration over final thing – audio component
  + Models:
    - LLMs (GPT4o, Claude, Llama)
    - Visual Gen (DallE, StableDiffusion v3)

Experiments:

* Little red riding hood -- script

Business Applications:

* Video editing – podcasts – Lex clips (we can bench against actual clips)
  + Condition on topic
* Insta reel – turn your gallery into reels
  + Looks at your pictures
  + Generate visual

Plot DSL:

Plot : Scene | Collection(Scene)

Scene: Setup(Characters, Setting, Objects, Actions)

Character: Char(body type, emotion…)

Setting: Setting(time, background, location, …)

Action: Map(char X setting X object)

Object: Object(description, PoI)

Plot := Collection(Scene1(…), Scene2(…), )

Animation DSL:

Animation : Crop | Pan | Filter | …

Eval:

* YouTube shorts and Insta Reels
* Use engagement metrics to refine generation policy