

MINIMAX НЗ

ПРОМТЫ ДЛЯ ВИДЕО ИИ

От идеи и референса
до кино-трейлера за 60 секунд
и масштабируемого видеопроизводства

ПРОМТЫ • КАМЕРА • РЕФЕРЕНСЫ • COMFYUI
ЗВУК • МОНТАЖ • ПРОИЗВОДСТВО • MASTERCLASS



Андрей Леонов

Андрей Леонов

**MiniMax H3: Промты
для видео ИИ**

«Автор»

2026

Леонов А.

MiniMax H3: Промты для видео ИИ / А. Леонов — «Автор»,
2026

Практическое руководство по работе с MiniMax H3 и связанным видеопроизводственным процессом. Книга объясняет понятным языком, как формулировать промты, использовать текстовые и референсные режимы, строить сцены, удерживать персонажа и визуальный стиль, создавать рекламные ролики, Shorts, трейлеры и мини-сериалы. Отдельные разделы посвящены зелёному рир-фону, ComfyUI, аудиту, препроектному обследованию, требованиям, референсам, звуку, голосам, управлению производством и Video Production Scaling.

© Леонов А., 2026

© Автор, 2026

Содержание

Содержание	6
1. MiniMax H3 как режиссёрский инструмент	8
2. Главный принцип промтинга	9
3. Структура базового промта	10
4. T2VA — видео из текста	11
5. I2VA — точный первый кадр	12
6. FL2VA — путь между двумя кадрами	13
7. L2VA — движение к финальному кадру	14
8. Ref2VA — работа с несколькими референсами	15
9. Камера	16
10. Композиция	17
11. Свет и ощущение премиального изображения	18
12. Персонаж и continuity	19
13. Предметная реклама	20
14. Реклама с человеком	21
15. Кино-трейлер за 60 секунд	22
16. Мини-сериал	23
17. Диалог и voiceover	24
18. Звук	25
19. Монтажное мышление	26
20. Ограничения и negative prompt	27
21. Библиотека практических формул	28
22. Перевод с русского на английский	29
23. Вирусные Shorts	30
23.1. Режимы MiniMax H3 простыми словами	31
23.2. Зелёная рир-проекция: герой отдельно, мир отдельно	32
Пример 1. Герой на зелёном фоне для последующей замены окружения.	32
Пример 2. Герой идёт на зелёном фоне, а камера имитирует будущий рекламный проход.	32
Конец ознакомительного фрагмента.	34

Андрей Леонов

MiniMax H3: Промты для видео ИИ

MiniMax H3: Промты для видео ИИ

От идеи и референса до кино-трейлера за 60 секунд и масштабируемого видео-производства

Автор: Андрей Леонов

Producer Capybara Cinema Studio

Содержание

1. MiniMax H3 как режиссёрский инструмент **3**
2. Главный принцип промтинга **3**
3. Структура базового промта **3**
4. T2VA — видео из текста **4**
5. I2VA — точный первый кадр **5**
6. FL2VA — путь между двумя кадрами **6**
7. L2VA — движение к финальному кадру **7**
8. Ref2VA — работа с несколькими референсами **8**
9. Камера **9**
10. Композиция **10**
11. Свет и ощущение премиального изображения **10**
12. Персонаж и continuity **11**
13. Предметная реклама **11**
14. Реклама с человеком **11**
15. Кино-трейлер за 60 секунд **12**
16. Мини-сериал **13**
17. Диалог и voiceover **13**
18. Звук **13**
19. Монтажное мышление **14**
20. Ограничения и negative prompt **14**
21. Библиотека практических формул **15**
22. Перевод с русского на английский **17**
23. Вирусные Shorts **17**

24. Производственный pipeline	26
25. Video Production Scaling	27
26. Свой IP и авторская вселенная	27
27. MASTERCLASS	27
28. Контент-инвентарь запуска	28
29. Запуск за семь дней	28
30. Локальный компьютер и GPU	28
31. ComfyUI и workflow	29
32. Контроль качества	29
33. Библиотека промтов	30
34. Как работать с Андреем Леоновым	30
35. Финальная стратегия	30
Об авторе	31
Источники и актуальность	31
Атлас практических промтов MiniMax H3	32
Практический playbook MASTERCLASS	81
Практический playbook Video Production Scaling	82
20 упражнений для тренировки промтов	82

1. MiniMax H3 как режиссёрский инструмент

В этой книге английский промт всегда сопровождается русским объяснением. Сначала формулируйте сцену понятным русским языком, затем переводите её в режиссёрскую структуру. Такой подход позволяет работать даже без глубоких технических знаний.

2. Главный принцип промтинга

Сравните: «женщина идёт по красивому городу» и «женщина в длинном тёмном пальто идёт по мокрой улице; камера плавно движется рядом на уровне груди; она останавливается перед витриной; тёплый свет витрины отражается на её лице». Во втором случае модель получает последовательность наблюдаемых событий.

3. Структура базового промта

Практический каркас:

integrated_multimodal_description:

[style]

[Shot 1]

[subject]

[action]

[camera]

[lighting]

[Shot 2]

overall_soundscape:

[physical sounds]

non_diegetic_music:

[instrumentation, rhythm, dynamics]

Не перегружайте короткую сцену десятками действий. Один основной ход камеры и одно заметное изменение обычно надёжнее.

4. T2VA — видео из текста

T2VA подходит, когда нет обязательного исходного кадра. Вы описываете полную сцену и её временную последовательность.

РУССКОЕ ЗАДАНИЕ: 15-секундный фрагмент вертикального трейлера. Ночной город после дождя. Мужчина в длинном пальто идёт по пустой улице. Камера начинает с общего плана, затем плавно движется рядом. В конце герой останавливается перед светящимся экраном; отражение экрана видно в его глазах. Дождь, шаги и далёкий транспорт слышны реалистично.

ENGLISH PROMPT:

integrated_multimodal_description:

A cinematic 15-second vertical trailer segment set in a rain-soaked city at night. A lone man in a long dark coat walks through an almost empty street. [Shot 1] A wide establishing shot shows wet asphalt and practical lights. [Shot 2] At 5.00 seconds the camera slowly tracks beside the man. [Shot 3] At 10.00 seconds he stops in front of a luminous screen and the camera makes a subtle push-in. The screen reflection is visible in his eyes. Realistic human motion, controlled cinematic lighting, shallow depth of field.

overall_soundscape:

Light rain, footsteps on wet pavement, distant traffic and quiet city ambience.

non_diegetic_music:

A restrained electronic cinematic score with a slow pulse and gradual build.

5. I2VA — точный первый кадр

I2VA начинается с изображения. Сначала зафиксируйте, что должно остаться неизменным: лицо, одежда, мебель, свет, композиция. Затем опишите только развитие.

РУССКОЕ ЗАДАНИЕ: использовать изображение как точный первый кадр; сохранить внешность героя и интерьер; герой медленно поднимает голову; камера делает плавный push-in; за окном начинается дождь.

ENGLISH PROMPT:

integrated_multimodal_description:

For the target video, at 0.00 seconds into the target video, Picture 1 is fully referenced. Preserve the subject identity, wardrobe, room layout and lighting direction. The subject remains seated briefly, then slowly raises his head and looks toward the window. The camera performs a very slow push-in. Rain gradually becomes visible outside. Preserve natural anatomy and realistic facial motion.

overall_soundscape:

Quiet room tone, subtle fabric movement and distant rain.

non_diegetic_music:

Minimal piano texture with a slow sustained pad.

6. FL2VA — путь между двумя кадрами

FL2VA нужен, когда есть точное начало и точный конец. Главное — описать наблюдаемое преобразование между двумя точками, а не просить модель придумать отдельный фильм.

ENGLISH PROMPT:

How the reference pictures align with the target video — Picture 1 aligns with the 0.00-second mark; Picture 2 aligns with the 15.00-second mark.

integrated_multimodal_description:

A single continuous cinematic shot. The empty hall from Picture 1 gradually becomes populated as people enter naturally. The camera slowly tracks forward. Architecture and furniture remain stable. Lighting becomes slightly warmer. The final composition must match Picture 2 exactly.

overall_soundscape:

The hall begins almost silent. Footsteps and distant conversation gradually increase.

non_diegetic_music:

A restrained orchestral pulse grows from sparse to full.

7. L2VA — движение к финальному кадру

L2VA полезен, когда важнее точный финальный образ. Изображение является пунктом назначения.

ENGLISH PROMPT:

How the reference pictures align with the target video — Picture 1 aligns with the 15.00-second mark.

integrated_multimodal_description:

Begin in a compatible wider environment. A woman walks through a softly lit gallery toward a sculpture. The camera tracks behind her, then moves to her side. The composition gradually converges toward Picture 1. The final second settles into the exact subject placement, lighting and framing of Picture 1.

8. Ref2VA — работа с несколькими референсами

Full-reference режим позволяет разделить функции источников. Один референс может отвечать за внешность, другой за одежду, третий за интерьер, видео — за движение камеры, аудио — за голос. Не следует писать «используй всё из всех файлов». Назначайте каждому источнику конкретную роль.

ENGLISH PROMPT:

subject_definitions:

<Subject 1> is the presenter, with facial identity derived from Picture 1.

<Subject 2> is the wardrobe derived from Picture 2.

<Environment 1> is the studio derived from Picture 3.

<Video 1> provides camera movement.

<Audio 1> provides vocal performance.

summary:

reference generation + audio reference.

retention_analysis:

<Subject 1>: fully_preserved.

<Subject 2>: fully_preserved.

<Environment 1>: fully_preserved.

<Video 1>: attribute_transfer for camera movement.

<Audio 1>: fully_copy for speech.

detailed_description:

Premium studio cinematography. [Shot 1] The presenter stands in the center. [Shot 2] At 5.00 seconds the presenter turns toward camera. [Shot 3] At 10.00 seconds the camera reaches a medium close-up and holds.

overall_soundscape:

Quiet studio room tone.

non_diegetic_music:

Minimal cinematic texture.

9. Камера

Основные команды: static shot, push in, pull out, pan left/right, tilt up/down, tracking shot, truck left/right, arc shot, POV, zoom in/out, roll. Не складывайте много движений в короткую сцену. Для премиальной рекламы часто работают static shot и very slow push-in; для динамики — tracking shot; для раскрытия пространства — pull-out или arc shot.

Вместо «камера делает всё сразу» пишите: «камера медленно движется вперёд, удерживая героя в центре; в конце слегка поднимается, открывая город». Простота движения повышает предсказуемость.

10. Композиция

Определяйте положение главного объекта: centered composition, subject on the left third, subject on the right third, foreground, midground, background, negative space, symmetrical composition. Композиция должна служить действию. Если место нужно оставить под последующую графику, заранее создайте negative space и не просите модель рисовать сложный текст.

Для вертикального видео главный объект должен оставаться читаемым на маленьком экране. Для рекламы полезно заранее разделять чистую сцену и графику: текст и логотипы часто надёжнее добавлять на монтаже.

11. Свет и ощущение премиального изображения

Слово «дорого» нужно переводить в физические признаки. Используйте controlled studio lighting, soft key light, subtle rim light, realistic skin texture, restrained highlights, clean reflections, natural shadow transitions. Для предмета: large soft source, controlled specular highlights, clean product edges. Для ночного города: wet pavement, practical lights, soft haze, controlled neon reflections.

Не перегружайте сцену эффектами. Один понятный источник света и один цветовой акцент часто выглядят убедительнее десятка бликов и частиц.

12. Персонаж и continuity

Создайте Character Bible: возраст, внешность, волосы, одежда, аксессуары, манера движения, голос и характер. Используйте одну и ту же формулировку в серии сцен. Для референсного режима полезны стабильные labels: <SUBJECT_MAIN>, <WARDROBE_MAIN>, <LOCATION_MAIN>.

Если в одной сцене герой в charcoal suit, а в следующей — в navy business suit, модель может воспринять это как изменение костюма. Continuity начинается с текста.

13. Предметная реклама

ENGLISH PROMPT:

Premium vertical product commercial. A glass bottle stands on a dark wet stone surface. A narrow controlled light travels across the glass, revealing realistic reflections. The camera performs a very slow push-in. The bottle remains stable and centered, with clean geometry and consistent label placement. Dark background, subtle atmospheric haze, premium commercial cinematography.

overall_soundscape: Subtle water droplets and quiet studio ambience.

non_diegetic_music: Minimal premium electronic texture with a soft low-frequency pulse.

14. Реклама с человеком

Человек должен выполнять микродействие: открыть дверь, взять предмет, сделать глоток, повернуться, посмотреть в окно, провести рукой по ткани. Микродействие создаёт правдоподобие.

ENGLISH PROMPT:

overall_soundscape: Footsteps, fabric movement, cup touching the counter and quiet morning ambience.

non_diegetic_music: Warm acoustic rhythm with light percussion.

15. Кино-трейлер за 60 секунд

Трейлер на 60 секунд лучше собирать из нескольких коротких генераций, а не пытаться заставить одну генерацию выполнить весь фильм. Практическая схема: четыре фрагмента примерно по 15 секунд. Первый создаёт загадку, второй знакомит с героем и проблемой, третий усиливает конфликт, четвёртый даёт кульминационный образ и финальный вопрос. После генерации фрагменты соединяются монтажом, а звук, музыка, титры и переходы доводятся отдельно.

ENGLISH PROMPT:

A cinematic 15-second mystery trailer segment. [Shot 1] 0.00–3.00: an abandoned apartment at night, an old radio turns on by itself. [Shot 2] 3.00–7.00: a young man enters and hears his own voice from the future. [Shot 3] 7.00–11.00: the radio becomes silent and the lights flicker. [Shot 4] 11.00–15.00: the camera pushes toward a blinking red indicator. Cut to black.

overall_soundscape: electrical hum, footsteps, radio static, room tone and sudden silence.

non_diegetic_music: dark cinematic pulse that stops before the final image.

16. Мини-сериал

H3 создаёт короткие эпизоды, а монтаж превращает их в сериал. Масштаб возникает в системе повторяемых сцен, а не внутри одного гигантского промта.

17. Диалог и voiceover

Реплика должна быть короткой. Для речи в официальном формате используется <d> с языком и точным текстом. Например: <d>[Russian] «Я думал, что знаю этот город.»</d>. Для закадрового голоса явно указывайте off-screen voiceover и закрытые губы.

ENGLISH:

The woman walks toward the elevator and says in an off-screen voiceover, with her lips completely closed: <d>[Russian] «Иногда решение появляется раньше вопроса.»</d>

Если нужен lip sync, не совмещайте длинную реплику с большим количеством сложного действия.

18. Звук

Разделяйте физический звук, голос и музыку. Физические звуки: шаги, дождь, дверь, ткань, стекло, двигатель, вода, город. Музыка описывайте через инструменты, ритм и динамику: low strings, deep percussion, slow pulse, gradual crescendo. Слово «эпично» слишком абстрактно.

Если ролик предназначен для диктора, оставляйте место в звуковом спектре и не делайте музыку слишком плотной.

19. Монтажное мышление

Каждый кадр должен отвечать на один вопрос: кто, где, что происходит, что изменилось. В коротком ролике лучше четыре ясных состояния, чем множество хаотичных действий.

Используйте принцип столкновения планов: предмет → реакция героя; результат → причина; деталь → общий план. Генерация становится сильнее, когда вы строите монтаж из отдельных смысловых блоков.

20. Ограничения и negative prompt

Ограничения должны исправлять конкретную проблему. Для анатомии: avoid extra fingers, malformed hands, duplicated limbs. Для продукта: preserve product geometry, stable proportions, consistent label placement. Для визуального шума: avoid excessive particles, random objects, unnecessary camera shake. Для текста: avoid distorted letters and malformed logos.

Не используйте огромный список запретов без причины. Ограничение не должно противоречить основному заданию.

21. Библиотека практических формул

1. Product: product + surface + light movement + push-in + hero frame.
2. Portrait: character + micro-action + tracking + reaction + close-up.
3. City: location + weather + subject crossing + camera + ambience.
4. Luxury: object + dark environment + single key light + slow camera + reflections.
5. Mystery: ordinary place + anomaly + reaction + escalation + final question.
6. Transformation: start state + continuous change + camera continuity + exact final state.
7. Tutorial: presenter + demonstration + close-up + voiceover + result.
8. Fashion: model + wardrobe + controlled walk + tracking + fabric movement.
9. Food: dish + ingredient action + macro detail + steam + hero plate.
10. Architecture: building + time of day + camera reveal + light transition + wide final.
11. Technology: device + interaction + interface glow + orbit + hero frame.
12. Sport: athlete + preparation + movement + impact + reaction.
13. Automotive: vehicle + environment + tracking + detail + hero pass.
14. Travel: place + traveler + movement + reaction + wide reveal.
15. Education: expert + problem + demonstration + result + clean ending.
16. App: person + phone + one interaction + reaction + beauty shot.
17. Book: book + hand + page detail + push-in + atmosphere.
18. Jewelry: jewelry + controlled highlight + macro + refraction + hero.
19. Restaurant: dish + chef action + steam + texture + table reveal.
20. Event: venue + arrival + interaction + speaker + audience reaction.
21. Horror: ordinary room + anomaly + reaction + silence + reveal.
22. Comedy: normal situation + unexpected action + reaction + timing beat.
23. Romance: two characters + small gesture + eye contact + slow camera.
24. Historical: period environment + character + action + camera + sound.
25. Fantasy: character + world rule + visual event + reveal + symbol.
26. Science fiction: technology + human interaction + system response + scale reveal.
27. Music video: performer + performance + camera rhythm + lighting rhythm.
28. Documentary: realistic subject + observational camera + natural light + human moment.
29. Corporate: professional + workspace + action + detail + wide context.
30. Social short: hook + transformation + reaction + proof + final image.

22. Перевод с русского на английский

Не переводите литературные эпитеты буквально. Переводите режиссёрскую функцию. «Крупный план» — close-up. «Плавное приближение» — slow push-in. «Съёмка вслед» — tracking shot. «Движение по дуге» — arc shot. «Мягкий ключевой свет» — soft key light. «Контрольный свет» — rim light/backlight. «Физические звуки» — physical sounds.

Русский текст сначала должен быть наблюдаемым. Только после этого его удобно превращать в английский промт.

23. Вирусные Shorts

Вирусный ролик начинается с события, а не с объяснения. Hook может быть необычным объектом, невозможным масштабом, неожиданным действием или резким визуальным контрастом.

ENGLISH PROMPT:

A cinematic 15-second vertical social video. In the first second, an enormous coffee cup stands in a city square like a monumental installation. A tiny human walks along its rim. The camera quickly pulls out to reveal the impossible scale and the crowd. The final shot reveals the complete installation from a high wide angle. Realistic materials, believable human motion, premium commercial lighting.

overall_soundscape: city ambience, footsteps, crowd reaction.

non_diegetic_music: fast modern rhythmic pulse with a final accent.

23.1. Режимы MiniMax H3 простыми словами

Официальная документация использует несколько названий режимов. По-народному их можно понимать так.

T2VA — «текст в видео со звуком». Вы словами описываете сцену, а модель создаёт видео и аудиосоставляющую.

I2VA — «картинка становится первым кадром видео». Вы даёте исходное изображение и объясняете, что должно произойти дальше.

FL2VA — «из первого кадра прийти во второй». Вы задаёте начало и конец, а модель строит движение между ними.

Ref2VA — «несколько референсов с разными ролями». Один файл может задавать внешность героя, другой — одежду, третий — место, видео — движение камеры, а аудио — характер речи. Главное правило: каждому референсу заранее назначается одна понятная функция.

Практический смысл всех этих названий очень простой: сначала определите, что именно вы хотите зафиксировать — текст, первый кадр, последний кадр, переход между кадрами или набор характеристик из нескольких референсов.

23.2. Зелёная рир-проекция: герой отдельно, мир отдельно

Зелёный фон нужен не для красоты. Он позволяет сначала снять или сгенерировать героя на ровном зелёном фоне, а затем отделить его от фона и поставить в другой мир. Это особенно удобно для рекламных роликов, трейлеров, ведущих, персонажей и серийных сцен.

Референс А — герой в полный рост на ровном зелёном фоне. Ноги полностью видны, вокруг тела есть запас пространства, свет равномерный, одежда не зелёная.

Референс С — при необходимости отдельный образец света: например, направление контрового света, который должен совпасть с фоном.

Пример 1. Герой на зелёном фоне для последующей замены окружения.

РУССКОЕ ЗАДАНИЕ:

Герой в полный рост стоит на равномерном чистом зелёном фоне. Он смотрит прямо в камеру, затем делает два шага вперёд и останавливается. Камера статична. Руки и ноги полностью видны. Никаких зелёных деталей в одежде. Равномерный мягкий свет без зелёного оттенка на коже. Нейтральное студийное изображение, чистые края силуэта.

ENGLISH PROMPT:

A full-body performer stands on an evenly lit pure green screen. He looks directly into camera, then takes two natural steps forward and stops. Static camera. Full body and feet remain visible throughout. No green clothing or green accessories. Soft even studio lighting with neutral skin tones and minimal green spill. Clean silhouette edges, realistic anatomy, natural body mechanics, production-ready chroma key plate.

Пример 2. Герой идёт на зелёном фоне, а камера имитирует будущий рекламный проход.

РУССКОЕ ЗАДАНИЕ:

Герой в полный рост идёт слева направо по зелёному фону. Камера плавно следует за ним на уровне груди. Шаги естественные, одежда сохраняет форму, руки двигаются естественно. Зелёный фон остаётся равномерным, без рисунка и теней.

ENGLISH PROMPT:

A full-body performer walks naturally from left to right across a seamless evenly lit green screen. The camera performs a smooth lateral tracking movement at chest height, matching the performer's pace. Natural walking cycle, stable wardrobe, realistic arm movement. Uniform chroma green background, no texture, no visible wrinkles, no hard shadows, clean edges and consistent exposure.

РУССКОЕ ЗАДАНИЕ:

Герой стоит на зелёном фоне, но освещён так, будто позади него ночной город: мягкий холодный контровой свет с правой стороны и слабый тёплый свет на лице. Он медленно поворачивает голову. Камера делает очень медленный push-in.

ENGLISH PROMPT:

A performer stands full-body on a clean green screen, lit as if he were standing in a night city: a soft cool rim light from camera-right and a subtle warm key on the face. He slowly turns his head toward camera. The camera performs a very slow push-in. Preserve realistic skin, clothing and anatomy. Keep the chroma background evenly green and free of texture.

Как собрать это в ComfyUI стандартными средствами:

1. Video Load — загрузите видео героя на зелёном фоне.
2. Chroma Key — выделите зелёный цвет и получите маску героя.
3. Background — подайте изображение или видео нового окружения.
4. Soft Edge — слегка смягчите край маски, чтобы не получить жёсткую вырезку.
5. Video Combine — соедините героя и новый фон в итоговый видеоряд.

Важный принцип: сначала исправляйте саму маску. Если по контуру остаётся зелёная кайма, не пытайтесь скрыть её цветокоррекцией. Проверьте равномерность зелёного фона, экспозицию, мягкость света и параметры Chroma Key. Для сложных сцен полезно отдельно контролировать отражения и полупрозрачные детали.

Конец ознакомительного фрагмента.

Текст предоставлен ООО «Литрес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на Литрес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.