AI 뉴스와 용어
월·수·금 오전 7시

LAB METAL
어려운 AI, 쉽고 재미있게

무료 · 언제든 수신거부
metallab.ai

제 8 호 | 2026년 8월 21일 금요일 | 오늘의 AI

크레아, 카메라 경로·3D 장면 제어 붙인 Seedance Studio 공개

프롬프트 프리셋과 카메라 패스, 3D 씬 컨트롤을 한 화면에 모은 영상 작업대. 모델은 Seedance 2.5



· 크레아, 카메라 경로·3D 장면 제어 붙인 Seedance Studio 공개

영상 생성 AI를 써 본 사람이 가장 자주 부딪히는 벽은 화질이 아니라 조작이다. 카메라가 어디서 출발해 어디로 가는지, 인물과 배경이 어느 위치에 서 있는지를 결과물에 반영시키려면 결국 글로 설명해야 한다. 크레아(Krea)가 8월 19일 공개한 Seedance Studio는 그 지점을 화면 위의 컨트롤로 옮겨 놓은 도구다. 크레아는 자사 공식 X 계정에 올린 게시물에서 Seedance Studio를 "Seedance 2.5로 만드는 데 필요한 모든 것을 한곳에 모은 도구"라고 소개했다. 구체적으로 이름을 댄 구성 요소는 세 가지다. 프롬프트 프리셋(prompt presets), 카메라 패스(camera paths), 그리고 3D 씬 컨트롤(3D scene controls).

게시물은 소개 영상과 함께 올라왔고, 크레아는 같은 글에서 "어떤 기능을 더 넣었으면 좋겠느냐"고 이용자에게 물었다. 발표 몇 시간 만에 댓글이 붙었고, 한 이용자는 모델 훈련 기능을 언급하며 크레아를 자신의 주력 플랫폼으로 쓰고 있다고 답했다.
각 컨트롤이 화면에서 어떤 모양으로 동작하는지, 세부 파라미터가 몇 개인지는 이번 게시물에 담기지 않았다. 확인된 것은 세 축이 한 도구 안에 함께 들어갔다는 사실이다.

이름이 헷갈리기 쉬운 대목부터 정리하는 게 낫다. Seedance는 틱톡을 만드는 바이트댄스의 AI 연구 조직 바이트댄스 시드(ByteDance Seed)가 개발한 영상 생성 모델이다. 이 모델은 보통 독립 앱이 아니라 다른 제작 도구 안에서 "고를 수 있는 모델" 형태로 등장한다.

크레아는 두 가지 일을 동시에 하는 회사다. 2022년 샌프란시스코에서 시작해 그림과 영상을 실시간으로 만들어 보여주는 도구로 이름을 알렸고, 자체 이미지·영상 모델을 직접 훈련한다. 동시에 구글의 비오, 오픈AI의 소라, 콰이쇼우의 클링처럼 남이 만든 모델 수십 개를 한 화면에 모아 골라 쓰게 해 준다. 이번 Seedance Studio는 후자 쪽 작업이다. 모델의 산지는 바이트댄스이고, 그 모델을 다루는 작업대를 만든 쪽이 크레아다.

Seedance를 쓴 작업물은 최근 국내 지면에도 여러 번 올랐다. 힉스필드가 8월 19일 유튜브에 공개한 90년대 홍콩영화풍 AI 단편 '레드 플래그' 역시 Seedance를 힉스필드 플랫폼에서 돌려 만든 결과물이다.

영상 생성 모델의 결과물을 좌우하는 요소는 크게 둘로 나뉜다. 하나는 무엇이 보이느냐, 다른 하나는 그것을 어떻게 보느냐다. 앞쪽은 글로 설명하기 비교적 쉽다. 인물의 옷차림, 시간대, 배경의 재질 같은 것들이다. 뒤쪽이 어렵다. 카메라가 인물 뒤에서 앞으로 돌아 나오는 움직임, 벽을 스치듯 지나가는 각도, 피사체와의 거리 변화는 문장으로 옮기는 순간 해석의 여지가 커진다. 같은 프롬프트를 두 번 넣어도 카메라가 다르게 움직이는 이유가 여기에 있다.

카메라 패스와 3D 씬 컨트롤이 별도 항목으로 발표에 등장한 것은 이 두 번째 축을 화면 조작으로 다루겠다는 신호로 읽힌다. 프롬프트 프리셋은 첫 번째 축, 즉 반복해서 쓰는 스타일 지시를 저장해 두고 꺼내 쓰는 쪽에 가깝다.

어디서 시작하나 — Seedance Studio는 크레아 서비스 안에 새로 붙은 도구다. 발표 게시물에는 도구로 연결되는 링크가 함께 올라왔고, 앱 안의 정확한 메뉴 위치는 이번 발표에 적히지 않았다. 크레아 계정에 로그인한 뒤 영상 도구 목록에서 Seedance Studio 항목을 찾는 것이 출발점이다.

작업 순서 — 발표에서 확인된 세 구성 요소를 기준으로 보면 흐름은 이렇게 잡힌다.

무엇을 해볼 수 있나 — 제품 소개 영상을 만든다고 하면, 프리셋으로 조명과 질감 톤을 고정해 놓고 카메라 경로만 바꿔가며 같은 제품을 여러 각도로 뽑아볼 수 있다. 짧은 내러티브 영상이라면 장면마다 프롬프트를 새로 쓰는 대신 프리셋 하나를 공유하고 카메라 움직임으로 컷을 구분하는 방식이 가능하다. 반복 작업에서 시간이 줄어드는 쪽은 대체로 이런 지점이다.

현재 공개된 범위는 여기까지다. 크레아는 추가 기능 요청을 받고 있고, 원하는 기능이 있다면 크레아 공식 X 계정의 해당 게시물에 댓글로 남기면 된다.

영상 생성 도구의 경쟁축이 모델에서 조작부로 옮겨가고 있다. 1~2년 전만 해도 새 소식의 대부분은 "몇 초까지 만들 수 있다", "손가락이 안 뭉개진다" 같은 출력 품질 이야기였다. 지금은 다르다. 품질은 여러 회사가 비슷한 수준에 올라섰고, 남은 차이는 원하는 그림을 몇 번 만에 얻느냐에서 갈린다. Seedance Studio가 프롬프트 프리셋·카메라 경로·3D 장면 제어라는 세 축을 한 화면에 모은 것도 그 흐름 위에 있다.

오늘 지면

1 크레아, 카메라 경로·3D 장면 제어 붙인 Seedance Studio 공개
2 크리에이티브 웹캠에 손 흔들면 테레민 소리, 브라우저 악기 공개
3 메타 AI 맥용 데스크톱 앱 출시, 앱 화면 통째로 물어본다
4 크리에이티브 일레븐랩스, 유튜브 썸네일 자동 번역 템플릿 공개

이 지면은 2026년 8월 21일 금요일 아침 구독자 메일함으로 나간 레터를 신문 판형으로 다시 조판한 것입니다. 기사 전문은 metallab.ai 에서 읽을 수 있습니다.

크리에이티브



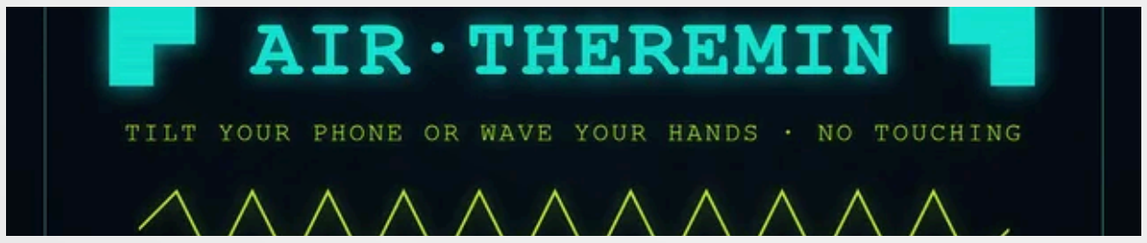
일레븐랩스, 유튜브 썸네일 자동 번역 템플릿 공개

유튜브는 영상 하나에 언어별로 다른 썸네일을 지정할 수 있는 기능을 지원한다. 스페인 시청자에게는 스페인어 문구가 적힌 썸네일이, 프랑스 시청자에게는 프랑스어 문구가 적힌 썸네일이 뜨는 방식이다. 번역된 제목과 현지어 자막, 더빙 오디오까지 붙이면 영상 전체가 시청자의 언어로 도착한다. 문제는 썸네일 디자인이었다. 번역된 헤드라인은 원본 레이아웃에 그대로 맞아떨어지지 않는 경우가 많아서, 언어를 하나 늘릴 때마다 디자인 툴에서 새로 작업을 반복

▶ 2면에 계속

크리에이티브

웹캠에 손 흔들면 테레민 소리, 브라우저 악기 공개

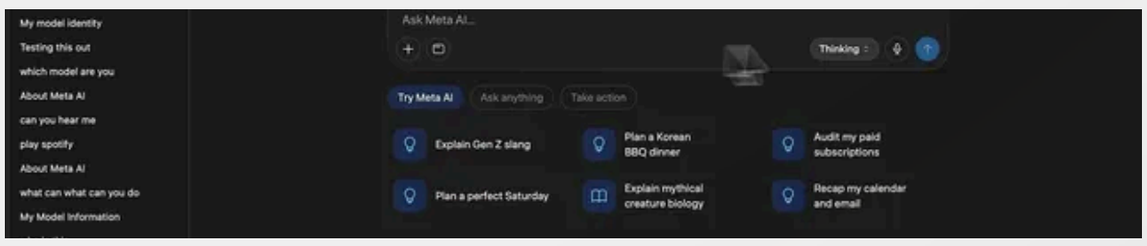


노트북 웹캠 앞에서 두 손을 활짝 벌리면 음량이 커지고, 손바닥을 맞대면 소리가 뚝 끊긴다. 개발자 Pavel Gurov가 만든 브라우저 악기 'Air Theremin'은 이 동작만으로 연주가 이뤄진다. 별도 프로그램 설치나 로그인 없이 웹페이지에 접속해 시작(START) 버튼만 누르면 된다. 이 프로젝트는 theremin.site를 기반으로 만들어졌다고 소개 페이지에 적혀 있다.

테레민은 러시아 발명가 레온 테레민이 1920년대에 만든 전자악기로, 연주자가 악기에 손을 대지 않고 두 개의 안테나 사이 손 위치만으로 음높이와 음량을 조절하는 방식이 특징이다. 손이 안테나에 가까워지거나 멀어질 때 발생하는 전자기장 변화를 소리로 바꾸는 원리인데, 이번 브라우저 버전은 실제 안테나 대신 웹캠의 손 인식 기술이나 스마트폰의 자이로 센서로 같은 경험을 재현한다.

▶ 2면에 계속

메타 AI 맥용 데스크톱 앱 출시, 앱 화면 통째로 물어본다



엑셀 시트를 띄워 놓고 "이 표에서 뭐가 이상한지 봐 줘"라고 말하면, 캡처를 뜨거나 파일을 올리지 않아도 된다. 메타 AI의 맥용 데스크톱 앱이 실행 중인 특정 앱의 화면을 그대로 대화에 붙일 수 있게 됐기 때문이다. 여기에 키보드 대신 말로 질문을 넣는 받아쓰기까지 함께 들어갔다.

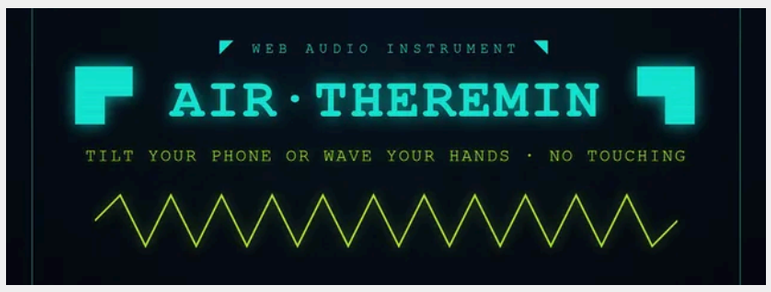
테스팅카탈로그가 전한 바에 따르면 메타는 macOS용 데스크톱 앱을 일부 지역에서 정식으로 내놨다. 핵심은 둘이다. 하나는 화면 공유로, 사용자가 지정한 앱 화면을 대화의 맥락(context)으로 첨부한다. 다른 하나는 받아쓰기 입력이다.

▶ 2면에 계속

◀ 1면에서 계속

크리에이티브

웹캠에 손 흔들면 테레민 소리, 브라우저 악기 공개



1 개발자 Pavel Gurov가 웹캠으로 손 동작을 인식해 연주하는 브라우저 테레민 'Air Theremin'을 공개했다

2 웹캠 손 인식(HANDS), 폰 자이로(GYRO), 마우스 대체 입력 3가지 방식을 지원한다

3 파형 4종과 비브라토·트레몰로·에코 등 이펙트, 녹음 기능까지 브라우저 안에서 제공한다

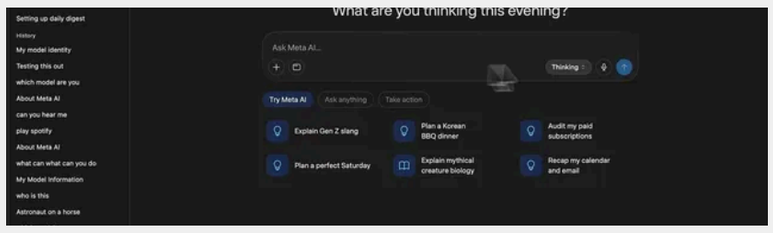
브라우저에서 프로젝트 페이지에 접속하면 화면 하단에 시작 버튼이 뜬다. 노트북이라면 내장 웹캠, 스마트폰이라면 자이로 센서를 그대로 활용하며 둘 다 없는 환경에서는 마우스로도 조작할 수 있다.

웹캠이 달린 노트북이나 스마트폰만 있으면 별도 앱 없이 브라우저에서 바로 실행된다. 두 조건이 모두 없는 데스크톱 환경에서는 마우스 입력으로 대체된다.

예를 들어 화상회의 전 잠깐 짬이 날 때 웹캠 앞에서 손을 흔들며 즉흥 멜로디를 만들어볼 수 있고, 스마트폰을 들고 앉은 자리에서 자이로 모드로 조용히 음정을 실험해 볼 수도 있다. 시소처럼 기기를 좌우로 흔들면 비브라토가 걸리고, 몸을 뒤로 기울이면 톤이 어둡고 공간감이 커지는 식으로 반응이 즉각적이라 짧은 소리 스케치용으로도 쓸 만하다.

◀ 1면에서 계속

메타 AI 맥용 데스크톱 앱 출시, 앱 화면 통째로 물어본다



1 메타 AI의 macOS용 데스크톱 앱이 일부 지역에서 정식 출시됐다. 설치 파일은 공식 페이지 ai.meta.com/meta-ai/assistant 의 「Download for Mac (Beta)」에서 바로 받는다

2 macOS 15 이상·애플 실리콘 맥에서 돌고, 파일은 Meta-AI-1.0.dmg(약 16MB)다. 요금은 무료지만 연산이 무거운 기능에는 사용량 한도가 시험 중이다

3 실행 중인 특정 앱 화면을 대화 맥락으로 붙이는 화면 공유와 받아쓰기가 핵심이고, 사이드바에 Media·Artifacts·Scheduled, 입력창에 Thinking 선택기와 마이크가 붙어 있다

공식 소개 페이지에도 맥 앱 자리가 생겼다. 페이지에 실린 설명은 이렇다. "데스크톱의 어떤 창이든 메타 AI가 보게 해서 작업을 처리하거나, 어떤 앱에서든 목소리로 받아쓰기를 할 수 있다."

메타 AI는 페이스북·인스타그램을 만드는 메타의 AI 조직이다. 라마(Llama)라는 내려받아 쓸 수 있는 모델로 이름을 알렸고, 2026년 8월에는 라마4의 실패를 인정한 뒤 후속 모델 뮤즈(Muse)를 오픈소스로 풀었다. 어시스턴트를 돌리는 모델은 메타 초지능연구소(Meta Superintelligence Labs)의 Muse Spark 라고 공식 FAQ에 적혀 있다.

설치 파일은 메타 AI 어시스턴트 공식 페이지에서 바로 받는다. 맥 앱스토어를 거치지 않고 메타가 직접 배포하는 방식이다.

◀ 1면에서 계속

크리에이티브

일레븐랩스, 유튜브 썸네일 자동 번역 템플릿 공개



1 일레븐랩스가 ElevenCreative Flows 안에서 유튜브 썸네일을 언어별로 자동 번역하는 워크플로 템플릿을 공개했다

2 GPT Image 2로 헤드라인 텍스트만 바꾸고 레이아웃·로고·디자인은 그대로 유지하며, LLM 노드가 번역 정확성을 다시 확인한다

3 썸네일 업로드와 언어 슬롯을 입력으로, 번역된 썸네일을 출력으로 묶어 재사용 가능한 템플릿으로 제공된다

번역을 눈으로 확인할 수 없는 언어라도 그대로 올리지 않도록 검증 단계가 하나 더 붙는다. LLM 노드가 생성된 썸네일 이미지를 다시 읽어 번역이 맞는지 확인하고, 실제로 어떤 문구가 적혔는지 알려준다. 여기에 언어 자체를 입력값으로 바꾸는 텍스트 노드를 더하면, LLM이 프롬프트를 그때그때 다시 써 주는 재사용 가능한 템플릿이 완성된다. 최종 결과물은 썸네일 업로드와 언어 슬롯을 입력으로, 번역된 썸네일들을 출력으로 묶은 템플릿 형태로 제공된다.

이 워크플로는 일레븐랩스의 ElevenCreative Flows 안에서 만들어졌다. 일레븐랩스는 튜토리얼과 함께 완성된 템플릿을 공개해, 직접 노드를 만들지 않고도 바로 가져다 쓸 수 있게 했다.

원문은 별도의 요금제나 지역 제한을 언급하지 않았다. 완성된 템플릿은 팀과 공유하거나 개인 채널에 바로 적용할 수 있는 형태로 소개됐다.

오늘 나온 어려운 말

풀이 전문은 metallab.ai/glossary

일레븐크리에이티브 플로우스
ElevenCreative Flows
일레븐랩스가 만든 노드 연결형 자동화 도구로, 여러 작업 단계를 이어 붙여 콘텐츠 제작 과정을 반복 실행한다

GPT Image 2
GPT Image 2
오픈AI가 만든 이미지 생성 모델로, 프롬프트에 따라 이미지를 새로 그리거나 일부만 수정한다.

씨댄스 스튜디오
Seedance Studio
크레아가 시드댄스 영상 생성 모델용으로 내놓은 작업 도구로, 카메라 경로와 3D 장면 배치를 화면에서 직접 조작한다.

Thinking 드롭다운
Thinking Dropdown
메타 AI 데스크톱 앱 입력창 옆에서 답을 얼마나 깊이 생각해서 내놓을지 고르는 선택 메뉴

오늘 해볼 것

오늘 지면에 나온 것 중 지금 바로 해볼 수 있는 6가지

1 일레븐크리에이티브 플로우스

ElevenCreative Flows

일레븐랩스가 만든 노드 연결형 자동화 도구로, 여러 작업 단계를 이어 붙여 콘텐츠 제작 과정을 반복 실행한다

이게 뭔데

일레븐크리에이티브 플로우스는 여러 개의 작업 상자를 화살표로 이어 붙여 콘텐츠 제작을 자동으로 돌리는 도구다. 번역, 이미지 생성, 검토처럼 반복되는 작업 하나하나를 상자(노드)로 만들어 순서대로 연결해두면, 다음부터는 처음 재료만 새로 넣어도 같은 과정을 그대로 반복해준다.

예를 들어 유튜브 썸네일을 여러 나라 말로 바꾸는 작업을 생각해보자. 원래는 언어가 하나 늘 때마다 디자인 프로그램을 열어 글자를 새로 앉히고 배치를 다시 맞춰야 했다. 이 도구 안에서는 번역과 이미지 생성을 맡는 상자, 번역이 맞는지 확인하는 상자를 미리 만들어두고 언어만 바꿔가며 복제하면 끝난다. 한 번 만든 흐름은 틀로 저장해두고 다음에도 재료만 갈아 끼우면 다시 쓸 수 있다.

즉 이 도구 자체가 그림을 그리거나 글을 옮기는 것은 아니다. 그런 일을 하는 다른 인공지능 도구들을 순서대로 배치해 자동으로 돌아가게 만드는 조립대에 가깝다.

해 보려면

1. 반복하고 싶은 작업을 정하고, 그 작업에 필요한 원본 재료(이미지나 문서)를 준비한다.
2. 번역이나 이미지 생성을 맡을 노드를 하나 만들고, 원하는 비율·해상도·언어 같은 조건을 프롬프트로 지정한다.
3. 결과물이 원본의 레이아웃이나 디자인을 그대로 유지하도록 지시문에 명확히 적는다.
4. 같은 노드를 복제해 언어나 조건만 바꾸면 여러 버전을 한 번에 만들 수 있다.
5. 결과를 다시 확인하는 검증 노드를 추가해, 의도한 대로 나왔는지 점검하게 한다.
6. 완성된 흐름을 템플릿으로 저장해두면 다음번엔 원본 재료만 새로 넣어 같은 과정을 반복할 수 있다.

오늘 지면에서 — 일레븐랩스, 유튜브 썸네일 자동 번역 템플릿 공개

2 GPT Image 2

GPT Image 2

오픈AI가 만든 이미지 생성 모델로, 프롬프트에 따라 이미지를 새로 그리거나 일부만 수정한다.

이게 뭔데

GPT Image 2는 오픈AI가 만든 이미지 생성 인공지능 모델이다. 글로 적은 장면을 그림으로 그려주거나, 이미 있는 이미지의 일부만 골라 다시 그려주는 역할을 한다.

비유하자면 지시를 정확히 알아듣는 그래픽 디자이너를 옆에 앉혀둔 것과 비슷하다. "이 썸네일에서 글자만 프랑스어로 바꾸고 나머지는 그대로 둬줘" 같은 구체적인 주문을 넣으면, 로고와 레이아웃은 건드리지 않고 헤드라인 문구만 바꿔서 새 이미지를 내놓는다.

이런 특성 덕분에 원본 디자인을 유지하면서 언어별 버전을 여러 개 만드는 반복 작업에 특히 잘 맞는다. 사람이 디자인 툴을 열어 매번 새로 손보는 대신, 같은 지시를 언어만 바꿔 넣으면 여러 버전이 순서대로 만들어진다.

해 보려면

GPT Image 2를 쓸 수 있는 도구에서 이미지를 하나 업로드하고 다음과 같은 프롬프트를 넣어볼 수 있다.

"이 이미지에서 헤드라인 텍스트만 스페인어로 번역해서 바꿔줘. 레이아웃, 로고, 색상, 나머지 디자인은 모두 그대로 유지해줘."

언어 이름만 바꿔가며 같은 요청을 반복하면 프랑스어판, 독일어판 등 여러 언어 버전을 순서대로 만들 수 있다.

오늘 지면에서 — 일레븐랩스, 유튜브 썸네일 자동 번역 템플릿 공개

3 씨댄스 스튜디오

Seedance Studio

크레아가 시드댄스 영상 생성 모델용으로 내놓은 작업 도구로, 카메라 경로와 3D 장면 배치를 화면에서 직접 조작한다.

이게 뭔데

씨댄스 스튜디오는 영상 생성 AI에게 '어떻게 찍을지'를 글이 아니라 손으로 지시하게 만든 작업대다. 영상 생성 AI를 써 본 사람이 흔히 겪는 어려움은 화질이 아니라 조작이다. 인물이 무엇을 입었는지, 배경이 어떤 재질인지는 글로 설명하면 대체로 잘 전달되지만, 카메라가 인물 뒤에서 앞으로 돌아 나오는 움직임이나 벽을 스치는 각도는 문장으로 옮기는 순간 애매해진다. 같은 문장을 두 번 넣어도 카메라가 매번 다르게 움직이는 이유가 여기 있다.

이 도구는 그 애매한 부분을 화면 위의 조작판으로 옮겨 놓았다. 자주 쓰는 지시문을 저장해뒀다가 다시 불러오는 프롬프트 프리셋, 카메라가 지나갈 경로를 화면에 직접 그려 넣는 카메라 패스, 인물과 배경을 입체 공간에 배치하는 3D 씬 컨트롤, 이렇게 세 가지 조작부가 한곳에 모여 있다.

여기서 개발한 쪽과 만든 쪽을 구분해서 볼 필요가 있다. 영상을 실제로 생성하는 모델 자체는 틱톡을 만드는 바이트댄스의 AI 연구 조직이 만들었고, 씨댄스 스튜디오는 그 모델을 다루기 편하게 조작판을 얹은 크레아의 작업 도구다. 크레아는 이 모델 말고도 여러 회사의 영상 생성 모델을 한 화면에 모아 골라 쓸 수 있게 해주는 회사이기도 하다.

해 보려면

1. 크레아의 영상 생성 화면에서 모델 목록 중 시드댄스를 선택한다.
2. 장면을 설명하는 프롬프트를 입력한다. 예시: "노을이 지는 좁은 골목, 우산을 든 사람이 카메라 쪽으로 걸어온다"
3. 카메라 패스 조작부에서 카메라가 인물 앞에서 시작해 뒤로 물러나며 골목 전체를 보여주는 경로를 그려 넣는다.
4. 3D 씬 컨트롤로 인물과 배경 사이의 거리, 배치를 조정해본다.
5. 생성 버튼을 눌러 결과를 확인하고, 카메라 경로만 다르게 바꿔 같은 프롬프트로 다시 생성해 차이를 비교해본다.

오늘 지면에서 — 크레아, 카메라 경로·3D 장면 제어 붙인 Seedance Studio 공개

4 Thinking 드롭다운

1. 노트북이나 웹캠이 달린 기기에서 카메라 접근을 요구하는 웹페이지(예: 브라우저 기반 화상회의, 웹캠 필터 사이트)에 접속해본다.
2. 브라우저가 카메라 권한을 묻는 팝업을 띄우는지 확인하고 허용을 눌러본다.
3. 권한을 허용한 뒤 별도 프로그램 설치 없이 곧바로 화면에 카메라 영상이 뜨는지 살펴본다.
4. 브라우저 설정 메뉴에서 특정 사이트에 부여한 카메라 권한을 나중에 취소할 수 있는지도 확인해본다.

오늘 지면에서 — 웹캠에 손 흔들면 테레민 소리, 브라우저 악기 공개

6 손 추적 모델

hand tracking model

카메라 영상에서 손의 위치와 손가락 모양을 실시간으로 찾아내는 AI 기술

이게 뭔데

손 추적 모델은 카메라로 찍힌 영상을 보고 손이 어디에 있는지, 손가락이 어떤 모양을 하고 있는지를 컴퓨터가 알아내는 기술이다. 사진 속에서 사람 얼굴을 자동으로 찾아 네모 박스를 그려주는 것과 비슷한데, 대상이 얼굴이 아니라 손이고 손가락 관절 하나하나의 위치까지 짚어낸다는 점이 다르다.

예전에는 이런 걸 하려면 장갑에 센서를 달거나 특수 카메라, 별도 하드웨어가 필요했다. 하지만 지금은 노트북이나 스마트폰에 원래 달려 있는 평범한 웹캠 영상만으로도, 소프트웨어가 화면 속 손의 움직임을 계산해낸다. 손을 벌리면 벌어졌다는 것을, 오므리면 오므렸다는 것을 좌표값으로 뽑아내는 셈이다.

이렇게 뽑아낸 손의 위치 정보는 다른 프로그램이 바로 활용할 수 있다. 손이 벌어진 정도를 음량으로, 손의 높낮이를 음높이로 바꾸는 식으로 연결하면 손짓만으로 연주하는 악기도 만들 수 있고, 게임 조작이나 화상회의 중 손동작 인식 같은 데도 같은 원리가 쓰인다.

해 보려면

웹캠이 달린 노트북이나 스마트폰으로 손 추적 기반 웹 서비스나 앱을 열어본 뒤, 카메라 프레임 안에서 손을 벌렸다 오므리거나 좌우로 움직이며 반응 속도와 정확도를 직접 확인해보자. 손을 프레임 밖으로 뺐을 때 인식이 즉시 끊기는지, 조명이 어두울 때도 잘 작동하는지 비교해보면 손 추적 모델의 한계를 체감할 수 있다.

오늘 지면에서 — 웹캠에 손 흔들면 테레민 소리, 브라우저 악기 공개

오늘 나온 말, 조금 더

풀이 전문은 metallab.ai/glossary

트레몰로

tremolo

트레몰로란 소리의 크기를 규칙적으로 떨리게 만드는 효과다. 형광등이 지지직거리며 밝아졌다 어두워졌다를 반복하는 모습을 떠올리면 된다. 음이 켜졌다 꺼졌다를 아주 빠르게 반복하면서 마치 소리가 떨리는 것처럼 들리는데, 이 떨림의 대상이 음량이라는 점이 핵심이다.

비슷한 이름의 효과로 비브라토가 있는데, 이건 떨리는 대상이 다르다. 비브라토는 음의 높낮이가 미세하게 흔들리는 것이고, 트레몰로는 음의 크기가 흔들리는 것이다. 두 효과 모두 오래된 기타 앰프나 신시사이저, 요즘의 브라우저 기반 악기 앱에서도 버튼 하나로 켜고 끌 수 있는 기본 이펙트로 자주 등장한다.

연주자가 직접 손이나 페달로 음량을 반복해서 조절하지 않아도, 기기가 알아서 일정한 속도로 음량을 오르내리게 해주는 것이 이 효과의 쓸모다. 그래서 잔잔한 배경음을 긴장감 있게 만들거나, 단조로운 소리에 생동감을 더하고 싶을 때 많이 쓰인다.

비브라토

vibrato

비브라토란 소리의 음높이를 아주 짧은 주기로 살짝 올렸다 내렸다 반복해서, 음을 떨리는 느낌으로 들리게 만드는 효과다. 가수가 긴 음을 뽑을 때 목소리가 미세하게 흔들리거나, 바이올린 연주자가 현을 누른 손가락을 좌우로 잘게 떠는 모습을 떠올리면 된다. 음이 딱딱하게 고정돼 있지 않고 살짝 살아 움직이는 것처럼 들리게 해준다.

요즘 나오는 브라우저 기반 악기나 음악 생성 도구에서는 이 효과를 사람이 실제로 손을 떨지 않아도 버튼 하나로 켜고 끌 수 있게 만들어 둔다. 프로그램이 정해진 주기로 음높이를 자동으로 미세하게 흔들어주는 식이라, 초보자도 클릭 한 번으로 진짜 악기를 흔들어 연주한 듯한 소리를 낼 수 있다.

음성 입력

Voice Input

음성 입력은 키보드로 글자를 치는 대신 목소리로 말한 내용을 그대로 텍스트로 바꿔 AI에게 전달하는 방식이다. 손이 자판 대신 입을 대신하는 셈이다.

비유하자면 비서에게 편지 내용을 불러 주면 비서가 받아 적어 주는 것과 같다. 다만 여기서는 AI 앱 자체가 그 받아쓰기를 해서, 말한 문장이 곧바로 채팅창의 질문으로 들어간다. 손으로 긴 문장을 치기 귀찮을 때, 혹은 다른 작업을 하면서 동시에 질문을 던질 때 편리하다.

다만 이건 말한 것을 글자로 옮겨 질문 한 번을 넣는 기능일 뿐, AI와 계속 목소리로 대화를 주고받는 기능과는 다르다. 후자는 음성 모드라고 따로 부른다.

화면 공유

Screen Sharing

화면 공유란 지금 화면에 띄워 놓은 창을 사진 찍듯 그대로 AI에게 보여주는 기능이다. 원래 화상 회의에서 옆 사람과 자료를 같이 보려고 쓰던 말인데, 이제는 AI 비서한테도 똑같이 쓴다. 엑셀 표를 열어 놓고 "여기서 뭐가 이상해?"라고 물으면, 캡처를 떠서 붙여넣거나 파일로 저장해 올릴 필요 없이 AI가 그 창을 곧바로 들여다본다.

비유하면 옆자리 동료를 불러 모니터를 손가락으로 가리키며 "이거 좀 봐줘"라고 하는 것과 비슷하다. 파일을 이메일로 보내고 상대가 열어볼 때까지 기다리는 대신, 지금 화면에 떠 있는 그대로를 즉시 넘겨주는 셈이다.

중요한 건 화면 전체가 늘 자동으로 넘어가는 게 아니라는 점이다. 보통은 사용자가 어떤 창(앱)을 보여줄지 직접 고른다. 그래야 열어 둔 다른 창이나 알림까지 원치 않게 AI에게 노출되는 일을 막을 수 있다.

카메라 경로

camera paths