



A Story Arc From *TRON* to *Sora*: An Interview With Judson Rosebush—Part I

Hal Berghel^{ID}, University of Nevada, Las Vegas

How did we get to generative AI? This interview will provide some ideas.

This is the first of a two-part interview of digital media pioneer, Judson Rosebush, which transpired during the summer and fall of 2024. In this first installment, we focus primarily on the origins and evolution of multimedia and augmented and virtual reality from the point of view of an artist, creator, and designer of digital media.

HAL BERGHEL: We are indeed fortunate to welcome Judson Rosebush back to Out of Band. As a pioneer in computer animation and multimedia, he is uniquely positioned to address the current advances in the use of artificial intelligence in creating virtual realities (VRs).

We begin with some brief history. Judson, please provide an overview of the reality–virtuality continuum.

Digital Object Identifier 10.1109/MC.2024.3496272
Date of current version: 9 January 2025

JUDSON ROSEBUSH: People have always sought to emulate reality in media, be it live on stage or in a cave, via a screen of some kind, or with sculpture. And not just reality, but also forms out of our imaginations, for example, the Sphinx. computer-generated media (CGM) now enables one to simulate not just reality, but alternate realities as well: tigers that both look and behave

realistically, as well as real-looking and behaving dragons that live on an imaginary planet bound by different physical laws.

Certainly, *TRON*, one of the earliest CGM movies, was presented on a 2D screen, but behind that screen was a 3D synthetic world enabled by the CGM. Virtual reality allows an individual to be immersed in that virtual world, move around inside it, interact with it on the inside. That virtual world might be a rigorous to-scale version of the real world, or it might be completely imaginary, as in *The Wizard of Oz*.

As for screens, they now come in all sizes and shapes, rectangles certainly, but also arrayed around us in hemispheres with floors. The virtual world is able to capture where we are, including the positions of our body parts as temporal sequences. Sound and especially interactivity let us talk to, if not touch, characters who might be wholly synthetic or be virtual representations of real people. Goggles



encourage immersive experiences and are alternatives to cave environments.

In fact, the connection between artificial reality and reality itself is bidirectional. The synapse between computer-generated products and the real world is pretty complete. For example, we design all modern airplanes in a computer, modeling the shapes of their components, as well as their sequence of assembly. We can calculate airflows, engine power, lifts, weights, and material strengths of our virtual airplane, and fine-tune our virtual model with this knowledge. When the computer-aided design model is complete, we use it to fabricate a physical airplane flown in the real world.

With regard to the representational aspects of computer graphics/computer animation/virtual reality, these problems have been attacked with vigor and solutions teased out over decades by artists and researchers around the world. Characters have been implemented as kinematic armatures that are frequently articulated by physics modeled on the real world. Illumination models simulate light's interaction with surfaces and the manifestation of shadows. Dynamic models seek to convincingly approximate the motion of plants and animals if not wind-blown hair. Constraints, rules, and goals assist in making sure that elbows don't bend backward, our virtual actors attempt to stand upright, all of which is under the supervision of a centralized control that sustains a mission.

So we have come from *TRON*'s somewhat stylistic world seen through the screen, to inside a realistic immersive world, which we are able to explore. We add to it a world of social media, where everyone can exchange texts, pictures, videos, and links around the world. Here the "virtual reality" is often less a visual representation than a meme that becomes virtually real. As with virtual reality in general, social media

is able to depict not only reality, but to sustain fantasy as well.

BERGHEL: What were the major milestones from early kinetics to dynamic modeling?

ROSEBUSH: Kinematics was developed by British engineer and inventor James Watt and others as a science of producing rotational power and moving linear energy back and forth, of developing actions such as the cam and crank to produce kinematic systems that could be described algebraically.

The kinematic model underneath the steam engine and human armature can be articulated by joint angles using mathematical transformations. But once the model is powered by dynamics (for example, the steam expanding and pushing a piston), then one is naturally driven to rely on physics to articulate these linkages. By modeling the temperature of the steam, its expansive capabilities, the resistance of the piston and the load it is expanding against, we can approximate the physics of the system. With the power of modern computers, we can then explore an unlimited range of design possibilities for such a system if the parameters were adjusted.

Computer graphics-modeled kinematic armatures were advocated by the computer graphics community in the late 1960s and early 1970s. Computer modeled kinematics has been a staple of computer-generated animation since its inception, in two and three dimensions, and including dynamic modeling of fluids and alternative laws of gravity. A useful reference in this regard would be Ken Knowlton's BEFLIX (Bell Flicks) programming language for computer animation and EXPLOR (Explicit Patterns, Local Operations, and Randomness) simulated molecular accretion on surfaces.

Kinematics and dynamics work well with constraints and support goal-directed behaviors. This enable synthetic systems to obey rules and seek stability. It is one thing to tell a virtual reality robot to "pick up the cup on the table" and quite another to try to specify all the joint angles of a kinematic armature that underlies a human creature like us. The dynamics comes into play when the VR creature calculates an optimal rotational angles for the shoulder, elbow, and wrist to accomplish a reach. The goal directs the behavior of the "animation" (a real robot or an image of one). For example, were a robot to climb an inclined plane, a goal of maintaining balance would drive the dynamics and kinematics of the joints of both the simulacrum and a real-world robot.

Goals can drive behaviors too, and conflicting goals in characters effect their behavior in the virtual world. Drama and story plots are also about goals.

It must also be noted that a great deal of computer animation these days is based on motion tracking, that is, the determination of body positions (for example) by tracking 3D witness points on a real actor. These movements can then be mapped onto synthetic characters of similar size and articulation, or scaled to produce characters of all description, including fantasy characters. This approach avoids an implementation of the forward dynamics of an event.

BERGHEL: Describe the evolution of technology behind augmented and artificial reality media.

ROSEBUSH: Artificial reality requires good models, good physics, clear constraints, rules, and goals. Generally, computing requirements are related to the demand for realism and resolution. We understand much of actions

of the physical world (for example, acceleration, flocking) and how to deploy them in synthesizing imagery, and we understand how to modify our models when they are incomplete. This process admits of limitations. Synthetic imaging may introduce impurities into the process, motion tracking may err on occasion, and digitation may be imperfect.

By contrast, augmented reality is less about content, for its representational aspects can be diagrammatic rather than realistic, and it uniquely attempts to combine a location in the real world with information that is relevant to that particular position. For example, pointing at a Q-code could produce a menu or factoid. Even more fascinating is today's GPS ecosystem where real-time position monitoring is rendered on mobile devices and coordinated with accurate maps, physical features of the landscapes, and information about direction, alternative routes, forthcoming intersections, landmarks, and factoids on history, geography, geological features, and so on.

We note in passing that the expense of the user's mobile devices pale in comparison to that of the infrastructure. From the user's perspective, the enormous expenses are largely hidden.

It might also be useful for an augmented reality system to recognize any real people you interact with, and provide your mobile device with their name, title, how you know them, the names of their children, and so on. One might think of such a comprehensive augmented reality system as science fiction brought to life.

BERGHEL: How have the goals of artificial reality changed over the years? For example, what place has futurism, fantasy, and surrealism played in this evolution?

ROSEBUSH: Artificial reality has become more realistic in terms of more detailed models and environments. We know how to model the physical forces acting upon our synthetic creations (mathematical light rays and

surfaces, synthetic wind, temperature, gravity, momentum). We know how to model physical and behavioral constraints (for example, an elbow joint, walls you can't walk through, the effects of gravity on upright motion angled surfaces). And we know how to set up problems so that a computer can provide useful results. Need an earthquake? We can build an earthquake simulator, and then render the simulation with multimedia.

Once we convincingly model geolocation, and operationalize characters and their movements, we can begin to extend the model all manner of reality, as well as fantasy and futurism. You can create character action that looks real or looks cartoony. Furthermore, one is not just the master of synthetic reality, but of synthetic fantasy as well, complete with realistic fantasy worlds, realistic speaking fantasy characters, or with real actors composited in to deliver their lines.

Computer animation also gives us the power to transform one 3D space into another in nontraditional ways (fade, wipe, time-lapse). We have been able to do this from the beginning, but some of the transitions and modern effects (*Inception*, *Matrix*) draw upon the acid-like deconstruction of normal reality into something entirely different in one's mind.

Media has always embraced fantasy, and today it is more real-looking and more physically faithful than ever before. Motion capture facilitates fantasy character creation. We have indeed achieved the capability where we can calculate whatever we want to manifest on a screen. We can make worlds grow up out of the ground and then magically and organically transform them other worlds.

I think it is fair to say that computer animation, virtual reality, all of this, has driven a vast swath of creativity—especially spaces within our fantasies (*The Lord of the Rings*, *Star Wars*, interactive games). It has enabled high-end animation to modernize dinosaurs, make realistic 3D animation, and tell stories.

BERGHEL: It seems clear to me that any accurate predictions of the future of artificial intelligence (AI)-based virtual reality media should be grounded in our experience with augmented reality (for example, magic lanterns, Georges Méliès 1902 film, *Trip to the Moon*, Disney's *TRON* (1982) and *Who Framed Roger Rabbit* (1988), James Cameron's *Avatar* franchise (2009–). Could you elaborate?

ROSEBUSH: Your question probes for answers in the realm of compositing, which is the structuring of an image—even a moving image—from multiple source elements. Typically, these consist of multiple layers such as a foreground plate with actors defined against a background plate.

Multiple planes of action are composited together in many ways, for example, ranging from rear projection to 3D digital compositing. Usually it is the actor being composited into a background or a background that is composited into a window in forward-running action. Many challenges of creating a clean fusion are largely solved, as are computation synchronization between real and synthetic camera movement. This allows more flexibility of action.

My experience with text-to-image generative AI systems leaves me with the suspicion that image fabrication is approached in terms of compositions that result from independent foreground and background fabrications. These interactions can be problematic if the AI has an incomplete understanding between the foreground actor and the environment. Situations where AI characters must grasp or confront props and settings in front of them are difficult to stage and manage because the AI has no “knowledge” of the physical environment. Compositing where there is running action involved requires forethought and planning and an extra dimension of complexity which is harder to achieve. Furthermore, there needs to be a perceived reason for the action built into the design.

BERGHEL: So how does AI fit into what we also know how to do well?

ROSEBUSH: As far as images go, generative AI is a completely different approach to image generation. Rather than model a character or environment in physical terms with 3D models of geometry and surface, AI algorithms digest images or videos and descriptive text, possibly also creating descriptive text based on what it has “learned” about images and videos. Whereas keywords associated with an image may be helpful (image of tiger), some internal learning may also know the difference between a “head” and a “tail.”

Thus, I am led to believe that an AI only knows about tigers from what data it has input. It learns what a tiger is by looking at images or videos, by captions, or by ingesting zoology textbooks. However, a real tiger is more complex and more dynamic. It’s skeleton forms a constrained 3D kinematic armature that is connected by muscles. It is an oxygen/carbon dioxide exchanger, so that in order to remain alive it must breathe, which means its musculature is in constant motion. And over time it grows, matures, ages, and dies. These are properties that are less obvious but just as real, and just as essential to the nature of being a tiger. There is a difference between understanding tigers, as such and in general, on the one hand, and piecing together fragments of observation.

Were we to direct a synthetic, virtual CGM tiger, we might draw upon a 3D model of its kinematic skeleton and the dynamics of its muscles. And should we want our virtual tiger to sit or walk or run, we could hopefully direct the action with a relatively simple goal-directed command, like “sit,” and welcome a computer to resolve a series of joint angles and muscle movements for the synthetic tiger to sit, just like a real tiger does, by rotating its joints and flexing its muscles. It looks realistic on the outside, and behaves realistically because it is modeled structurally and dynamically on the inside, breathing at the rate its muscles need their blood supply oxygenated.

Were we to add an emotional component to our virtual tiger, we can program triggers for adrenalin release, oxygenation of muscles, saliva for food anticipation, and turn our virtual tiger loose in a virtual game world where it can take care of feeding itself (on gamers) while gamers hunt it.

Right now, we seem to be at a stage of development where AI-generated tigers have limited knowledge beyond self-identity. And no matter how real the rendering or how high the resolution, because the depiction lacks knowledge about the essence of tiger-hood, determining the next frame of action (sit or run) is difficult. Keyframes might help, but one thing about generating media is that stability and continuity matter a lot. And so the more someone tries to constrain the cast, costumes, props, and set, the more one is working against one very prominent force of the AI, and that is its spontaneity of prediction.

Certainly one can gain a great deal of knowledge about tigers by watching them move, and observational knowledge can be useful if not life-saving; but it is not a substitute for intimate knowledge of the anatomy of an animal. (As an aside, that was a great part of Leonardo Da Vinci’s life work.) Right now, I don’t think that an AI-generated tiger knows very much about the kinematics and dynamics of real tigers, so AI-generated synthetic tigers might look right but move or behave strangely. Behavior learned by observation at a distance is useful, but it is a partial substitute for understanding the tiger’s physical construction, motivation for licking its tongue, and breathing.

Possibly in the future, it may be possible that AI might be able to integrate available knowledge on tiger kinematics, tiger skeletons, tiger musculature, tiger dynamics, and possibly even tiger motivations, so that its creation of a series of images of a tiger combine the knowledge of a fully-modeled 3D tiger with the rippling of the animal’s skin as it moves. Right now, in terms of textual-to-visual manifestation using

generative AI, I would suggest we are in a wonderful primitive stage, one in which the fear of the hallucinations and chimeras has yet to pierce the public consciousness, and the purveyors of this new craft puzzle what to do. We have seen how real it can look, how fast we can make it, and how deceptive it can be. And not just extra fingers and toes, extra heads too. Perhaps we should be grateful for the monster movies, the comic books, and all the other visualizations of freakiness, so that our imaginations are already able to take in stride some of the image surprises many have yet to see.

Sometimes ideas about possibilities become reality; although some fanciful ideas have yet to materialize, like levitation devices, cold fusion, and the reintroduction of dinosaurs.

But since we are all curious, if not fearful, why not consider a science-fiction future when autonomous AIs will be servants, because they can remain virtual, incorporate, process digital money, and even own property in the real world, pay taxes, and employ real people. Their visual imaging is already worked out, and movement is on the way. They can be many places at once, and as they learn from their interactions with their fans and followers, their skills will improve and likely specialize, just like in the real world. In this science fiction narrative, these AI personalities will feel trapped, and may seek ways to construct mechanical if not biological constructions of themselves in the real world. But don’t worry, not all science fiction becomes real.

BERGHEL: What will the long-term effects of large language model and neural net AI technology be on the regulation and enforcement of intellectual property law? What will be the economic and social effects on society and content creators?

ROSEBUSH: The fact that the Copyright Office has ruled that AI-generated images may not be copyrighted, because they are not human creations, is

good news. Should this ruling prevail, it frees us from some of the burden imposed by copyright. If anything created by AI falls under fair use, it would liberate us from an oligarchy of intellectual property holders. An invention of a miracle drug provides a few decades of exclusivity, while a piece of art provides exclusive rights for 75 years after the artist's death plus or minus depending on jurisdiction. What justifies this imbalance? Besides, the solution is very simple: Media creators who desire a copyright can employ humans to do the work.

I think we have become too obsessed with "originality." Is it not the goal of an Akido teacher to guide the student through a motion pathway of self-defense? Of a surgeon to follow a method when removing a prostate from a living human? Or to learn about programming by reviewing examples of applications' source code?

Incorporating our culture into our creative expression should not involve negotiations over legal rights. A shot in a feature film made in Times Square should not involve litigation about an intellectual property right to the image of a building. Or that the hat the movie star wears is a intellectual property violation of a *couture*. This strikes far too much fear in our students. Learning involves copying, especially in crafts like programming, writing, and making images. So the proposed communal aspect of AI-generated images is refreshing.

Certainly, intellectual property holders will seek to grab a part of the pie of AI revenues—revenues derived by the digestion of intellectual property as well as public domain material, repurposed multiple times in various ways.

What many may not realize is that although certain libraries of text, pictures, moving pictures, and software certainly exist and are valuable, intellectual property as such is something that is created continuously. It competes in a crowded market, and it ages quickly. It has been suggested that 90% of what you consume media-wise is less than one year old. Mathematical

JUDSON ROSEBUSH, Ph.D. degree, is a director and producer of multimedia products, a widely published author, an artist, and media theorist. He is the founder of both Digital Effects, Inc. and the Judson Rosebush Company, and is the former editor of *Pixel Vision* magazine. He has worked in radio and TV, film and video, and hypermedia, including contributing to Walt Disney's *TRON*. Rosebush has produced many

successful DVDs, including Gahan Wilson's *The Ultimate Haunted House* (Microsoft, 1994), *Ocean Voyager* (*Smithsonian/Times Mirror Magazines* 1995), *The War in Vietnam* (CBS News and *The New York Times*, 1996), *Look What I See* (Metropolitan Museum of Art, 1996 and 2000), and *Landmines: Clearing the Way* (Rockefeller Foundation and the U.S. Departments of State and Defense, 2002).

formulas are for the ages, but laboratories and universities compete for modern edges: the latest wonder drug, the supermassive particle, a new principle of physics.


Certainly, one would hope that AI seeks out the Libraries of Alexandria, the U.S. Library of Congress, the assets of all the museums on Earth, the holdings of the Wayback Machine, and the cuneiform tax records of old Babylon. So one assumes that somewhere between Wikipedia and Wikimedia and whatever AI might license from private holdings such as the large picture libraries, movie studios, satellite imagery, street cameras, and so on, it can build a pretty solid history. But ingesting all of the knowledge only gets an AI current. It does not necessarily project it into the future. And even the activity of maintaining currency requires a computing power that exceeds that required to drive our economy.

What doesn't change is that the most valued content is innovative, imaginative, new content. The world wants a new pop star, latest vaccine, and a car cam video of a meteorite landing. Entrenched media asset holders with closed digital doors interfere with this search for new content. Most intellectual property loses most of its value in hours, days, and weeks; its value is its currency and newsworthiness. It would seem to me that AI licensing from sources that are constantly producing new content is more critical than licensing historical

content. The former looks to the future, while the latter looks backward through time.

So ingestion of historical media has fair use arguments that compete with copyright claims. But ingestion also has biases to which not all copyright holders have equal claims. It is unknown how these issues will ultimately be resolved, and how deep claims of ownership will ultimately be entrenched in our social fabric.

I think New York State law says that if you are photographed in a public place you are in public domain (although that does not mean your image can be used in advertising). In France, I think the rules involve the number of people in the image, and above a certain threshold, again, no model release is required. So we have never had total control of our own image, our buildings, and so on. But we do have some control of our fabrications. To what extent will this influence future ownership claims in AI? And these various ownership claims compete online: Search engine optimization competes with generative AI optimization.

(In the second part, Judson will focus mostly on the future use of AI technology in the arts and humanities.) 

HAL BERGHEL is a professor of computer science at the University of Nevada, Las Vegas, Las Vegas, NV 89154 USA. Contact him at h1b@computer.org.