

Breakout 1.2: Program Considerations – ARPA-E Questions

Tuesday, April 26, 14:10-15:55

Objective: Help us imagine what an ARPA-E investment in telepresence avatars would look like. What must projects achieve to de-risk telepresence avatar technology? How will ARPA-E validate that projects are on a path to realistic avatars? How else can ARPA-E enable success?

The Ultimate Goal

The Avatar Turing Test: How would we test avatars as a realistic representations of the remote person, for the purpose of business partnerships and personal relationships?

- The Avatar Turing Test is sufficient but not necessary test for fungible alternatives to travel intensive in-person conversation. What is a clear necessary and sufficient test?
- Prioritize a list of test constraints/scenarios: bandwidth, latency (avg & stdev), packet loss, number of participants, computing power, memory, etc.
- How does the test evaluate subjective metrics, e.g. trust, ease of use, comprehension?

Structure of an ARPA-E Program

The goal of ARPA-E funding is to de-risk new technologies, not create final commercial products. What can ARPA-E do with \$30M in three years that will sufficiently de-risk the tech?

Low-Hanging Fruit Strawman: Users go to a local store where a \$30,000 machine creates their base avatar model. Current real-time Avatar performance capture technology has been made easier to use, without makeup, and available for an at-home price. The Wi-Fi connected clients connect to meetings with 2-30 participants, having conversations as if they were in person. The avatars are rendered on state-of-the-art HMDs or 2D high-definition screens.

- Estimate the strawman development costs. Describe cost drivers: hardware, labor, etc.?
- Does the strawman sufficiently de-risk a technology that would eventually pass the Avatar Turing Test? If not, what additional investments would be required?

Breakout Session 1.2: Feedback

Objectives:

- We should make interoperability a goal to both provide guidance and enticement for participating groups as well as for transition, but enforcing standards is really tough.
- Body tracking will be done by gaming industry, but they will not focus on facial feature tracking (good white space for ARPA-E). Solution should be markerless.
- We could go for a portfolio approach and let teams propose different research areas.
- In general, group couldn't get past a need for use cases, and could not contribute what tech would be common to a large portion of use cases.
 - Good to have a set of 4-7 use case scenarios in which we could evaluate success. Attending a meeting? Speaking at a conference? Broad class of representative use cases.
- Size of a group that would be tough?
 - Minimum is 3. Double digits is really messy
- Real time capture – path is not clear today.

Avatar Turing Test:

- Is a "turing test" necessary? Maybe look at if avatar is realistic enough to inspire empathy.
- Ignore network constraints first, first see if you can make it happen at all (with consideration that one day it will have to go over a network).
- Would be really difficult to put different team's avatars into a single competitive environment due to software integration.

FOA vs Competition:

- FOA would produce more basic research, while a competition would produce more integration work.
- FOA will be the best bang for the buck, and the most creative freedom will come from that.
- Could do a hybrid FOA/competition, give funding to do basic research, then pull fundamental technologies out and do a competition based on their integration.
- ARPA-E Hackathon?
- Competition:
 - Stuff that you can also do in physical reality would be good to do because you can compare across and show virtual > physical
 - Summarize knowledge into cohesive whole
 - Sharing objects
 - Bring digital pieces together to do something with
 - Assemble the puzzle
 - Crowd sourcing the judging? Youtube viewers can see the difference? If you crowdsource, subjective becomes compelling.

Suggested In-Scope vs Out-of-Scope

In scope:

- Facial Tracking
- User Interfaces
- Full body motion capture in a room (not chair)
- Audio/acoustic environment capture and reproduction
- High quality motion from low quality sensors
- Compensating for latencies algorithmically
- Proposing an evaluation methodology
- Hand Gesture Tracking
- Interacting with an object
- Representing a common artifact that the group wants to share (e.g. a device, document, etc.)

Out of scope:

- HMD
- GPU/CPU processing
- Rendering from a graphical sense
- New networking protocols
- General hardware
- Social Implications (some disagreement about this) & Policy
- Security
- Haptics (leave for Telelabor)