Synthetic Humans for Entertainment & Accessibility
Sunday 17th September
13:30-14:30

ROBERTO IACOVIELLO
Lead Research Engineer
RAI

ERINROSE WIDNER
Global Head, Business Strategy Emerging & Creative Technologies
VERIZON

MICHAEL DAVEY
Technologist

MATT KIRBY
Actor/Sign Language Translator

CLIVE SANTAMARIA
Chief Architect
ITV

MODERATOR

MUKI KULHAN
Innovation Co-Lead
IBC ACCELERATORS

Associate sponsors: AMD, Microsoft
Media partner: 365
Leverage commodity tools to generate photorealistic digital twins of humans that could be used for TV productions and other platform environments.

Subsequently, improve entertainment and accessibility going beyond traditional production workflows.

We have chosen to explore both these ideas within one IBC accelerator due to their similar themes and the opportunity to share and leverage the learnings.
DEVELOPMENT OF VIDEO AND VOICE

- Real image
  - No voice
- Real person
  - Real voice
- Digital animation
  - Real voice
- Synthetic human
  - Real voice
- Synthetic human
  - Synthetic voice
Two different use cases based on synthetic humans:

- **Workstream 1: Entertainment** - create a realistic synthesized, Maria Callas, both in her likeness and voice
- **Workstream 2: Accessibility** - Look at new ways to scale accessibility including voice synthesis, and tackling lip-sync

Both use Synthetic Humans through motion capture technology to replicate human movements and facial expressions, creating faithful representation of characters.

Deployment of a pipeline for the integration of the synthetic humans to support Multiplatform publishing (e.g., TV programme, Volumetric LED wall, Headset display)
WORKSTREAM 1:
SYNTETIC HUMANS
for ENTERTAINMENT
SYNTHESISED MARIA CALLAS
Original Objective:

- Leverage Broadcaster’s archives: video (face & body model) & audio (voice cloning)

Challenges:

- **Automated** reconstruction of photorealistic 3D models
- Realistic interaction between animated 3D models and pre-recorded digital twin captured using volumetric studio
- Integration into led studio production (realtime)
MANY EMERGING SKILLS NEEDED

- Motion capture
- 3D model clean up
- Skin
- Face Building
- Digital Clothes Animation
- Photo realistic Rendering
- Body Animation
- Audio to Lip-sync
- Audio synthesis
- Training from video source
- Face Animation

HOW HAVE THESE BEEN USED IN THESE PROTOTYPES?
IBC2023 AUTOMATION: Body orchestrator

SYNTHETIC HUMANS FOR ENTERTAINMENT AND ACCESSIBILITY

<table>
<thead>
<tr>
<th>Title</th>
<th>Started at</th>
<th>Process status</th>
</tr>
</thead>
<tbody>
<tr>
<td>f3b763:3-330-4839-5:939-348464:bebf5a</td>
<td>3 days ago</td>
<td></td>
</tr>
<tr>
<td>2f5d3737-3323-4835-a240-1e65202d14a3</td>
<td>7 days ago</td>
<td></td>
</tr>
<tr>
<td>ff8a315-3444:ae91-9671-ee458a30573c</td>
<td>10 days ago</td>
<td></td>
</tr>
<tr>
<td>2d430d49-3165-402b-5959-cf1313a5d19c7</td>
<td>14 days ago</td>
<td></td>
</tr>
<tr>
<td>862f389d-4d4b-45f3-af81-a609b70f161ff</td>
<td>15 days ago</td>
<td></td>
</tr>
<tr>
<td>0fbbf5a61-c8de-4740-a720-2658e6def8f19</td>
<td>15 days ago</td>
<td></td>
</tr>
<tr>
<td>Process 4</td>
<td>1 month ago</td>
<td></td>
</tr>
<tr>
<td>Process 3</td>
<td>1 month ago</td>
<td></td>
</tr>
<tr>
<td>Process 2</td>
<td>1 month ago</td>
<td></td>
</tr>
</tbody>
</table>
WORKFLOW FOR ENTERTAINMENT
WORKFLOW FOR ENTERTAINMENT

Model Creation
- Archive exploitation
- Image restoration
- Face builder
- Meta-Human

Body and face Animation
- Facial Animation
- Body Animation
- Clothes Animation
- Voice Cloning

Virtual production

2D video
- 3D Theatre

Volumetric capture

IBC2023 #ACCELERATORS2023
Zoom-in CLOTHES ANIMATION  

Zoom-in BODY VOLUMETRIC CAPTURE
WORKFLOW FOR ENTERTAINMENT
THE GREAT MARIA CALLAS
WORKSTREAM 2: SYNTHETIC HUMANS for ACCESSIBILITY
Original objective:

• Develop ways to scale creation of accessibility to accompany the growth in content

Challenges:

• **Automated** text-to-voice synthesis for enabling voice & audio description
• Develop lip-sync to be able to enable lip reading
• BSL-presented weather forecasts
• Standardisation and Commercial considerations
ACCESSIBILITY WORKFLOW

Avatar creation
- Synthetic human creation
- 3D model clean up
- Face Building

Text to Voice
- Transcript → Audio synthesis

Voice to Lip-sync animation
- Face Animation
- Motion capture
- Photo realistic Rendering
- Audio to Lip-sync
Matt Kirby
Synthetic Voice
Darren Altman: Voice Actor
Matteus: Virtual Matt Kirby
TEAM LEARNING POINTS & FUTURE WORK:

SYNTHETIC HUMANS

for ENTERTAINMENT & ACCESSIBILITY
How was the quality of audio and video from the archives?
Quite a challenge to create 3D models and upscale video

How accurate was the animation and cleanup?

*Entertainment:* Motion capture is mature and body movement is natural and required limited cleanup

*Accessibility:* Body, facial expression and lip patterns are good enough for specific use cases.

Hand and finger mocap requires lots of manual correction.

Production time: Cleaning up a mocap is x3 as long as vs capture.
How was the text-to-voice synthesis?
Workflows are pretty common now and voices are realistic, and can have ranges of emotion and variation.
Production time: Negligible

How was the modelling and clean-up?
Avatar modelling: Rapid advances but still challenging, requiring manual sculpting. Model validation.
Production time: 80/20 rule applies

How good was the Lip-sync?
Accuracy depends on accent, depends on trained voices
Lip-sync to music or sung voice challenging as lip-sync technology optimised for speech.
Lip-sync not as accurate for accessibility, not always detailed in an avatar. Breaking down domain silos.
Production time: 1.5x real time
BUT WAIT, THERE’S MORE!

TALENT ID: RESPONSIBLE UTILITY & VALUE
TALENT ID WORKFLOW FOR VIRTUAL AND FICTIONAL HUMANS & SYNTHETIC HUMANS

1. Talent Data Ingestion / Collection
2. Talent Data Deduplication
3. Application of Citation-Backed Notability™, Survivorship, and Validation Rules
4. ID Resolution & Registration Service

- **Provenance**: Guard Against Deep Fakes
- **Accounting**: Talent Appearance Tracking for Royalty, Residuals and Participations
- **Security**: Drive personal identifiable information out of Transactions (e.g. Gov. Issued ID’s)

#IBC2023 #ACCELERATORS2023
FUTURE POTENTIAL OF SYNTHETIC HUMANS

• Current **production quality is improving** for our broadcast applications (compared to gaming, Hollywood, etc)

• **Multiplatform and multiple devices** can be addressed

• **Reusable, sustainable 3D assets**, for all media: (eg: new formats, remote production/comms, fan & avatar interaction etc)

• Further **automation & integration** is required to improve democratization and scale production

• Develop a general-purpose **speech to lip-readable and/or signing avatars** service
Industry calls to action

- Reusability
- Standardisation
- Ecosystem Maturity
- Talent Identification
Come to the Accelerator Zone project POD, to...

- Meet the Team!
- Workflow challenges in detail & live demos
- Future accessibility discussions and technologies
- Tracking Talent ID
- White Light demo (at D&B stand across from Pod)

Go to the IBC website project page for:

- Highlight Videos
- White Paper
- Sign Language Avatar videos
- Workflow PDFs
Synthetic Humans
for Entertainment
& Accessibility

Thank You