

Tracking Character Diversity in the Animation Pipeline

Mara MacMahon
mara@pixar.com
Pixar Animation Studios
Emeryville, USA

Gordon Cameron
gordon.cameron@epicgames.com
Epic Games
Cary, USA

Emily Wilson
emilywilson@pixar.com
Pixar Animation Studios
Emeryville, USA

Jessica Heidt
jheidt@pixar.com
Pixar Animation Studios
Emeryville, USA

Paul Kanyuk
pkanyuk@pixar.com
Pixar Animation Studios
Emeryville, USA

Peter Nye
peter@pixar.com
Pixar Animation Studios
Emeryville, USA

Joshua Minor
joshm@pixar.com
Pixar Animation Studios
Emeryville, USA

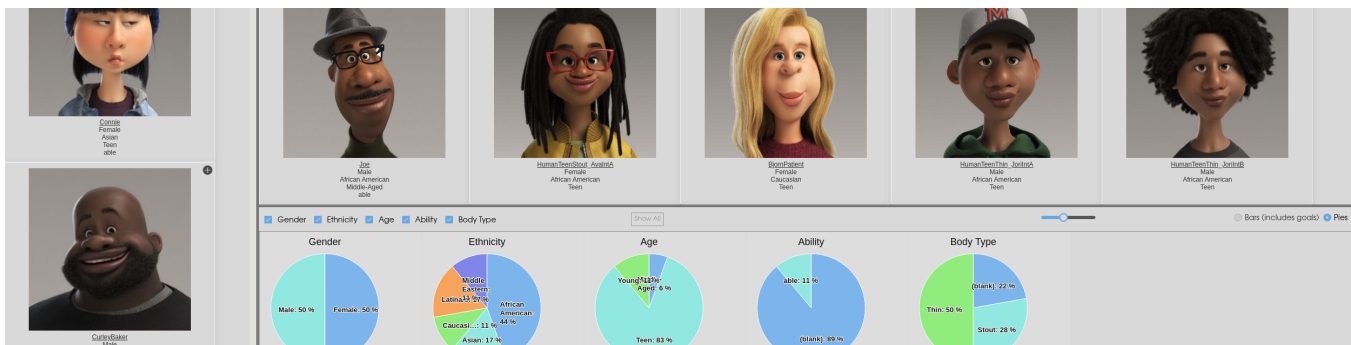


Figure 1: The main UI for this project is a web-based application called *Character Linker* that lets you populate a sequence or shot with characters and immediately view the resulting composition in terms of diversity characteristics, such as ethnicity, gender, age, body type and ability. It is used both in the early stages of a sequence to populate the sequence or shot, and the later stages, when the shots have been constructed and you want to view the result in some detail as described above. ©Pixar.

ABSTRACT

As we explore a broad range of characters and stories in our films, it has become increasingly valuable to view breakdowns of our character pools and selections by demographic: to build and use our assets efficiently, reinforce storytelling and world building choices, and ensure consistent decision-making across the pipeline. With the *Character Linker* App within *Traction* (*Traction* is Pixar’s asset and shot-tracking tool), production is able to see a live breakdown of the character pool as assets are built, and sequence/shot composition, as they are populated—with the ability to visualize by a range of categories, including gender, ethnicity, body-type, and age, among others. Each film can define and populate these categories specific

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
SIGGRAPH '22 Talks, August 07-11, 2022, Vancouver, BC, Canada
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9371-3/22/08.
<https://doi.org/10.1145/3532836.3536257>

to their story, set breakdown goals to measure progress against, and iterate on crowd asset selections to ensure each character is utilized to the fullest.

ACM Reference Format:

Mara MacMahon, Emily Wilson, Peter Nye, Gordon Cameron, Jessica Heidt, Joshua Minor, and Paul Kanyuk. 2022. Tracking Character Diversity in the Animation Pipeline. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Talks (SIGGRAPH '22 Talks)*, August 07-11, 2022. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3532836.3536257>

1 BACKGROUND

When starting production on *Soul*, a priority when discussing character demographics was to be true to the film’s real-world location: New York City. The team put together a favorable character pool to achieve this goal, made up of 80% BIPOC and a 50% split between male and female. This ensured that any randomization of crowds would result with satisfactory representation.

Featured background characters, however, had different results. At the time, there was no official statistical tracking method in



Figure 2: A scene from a sequence in *Turning Red* where demographic statistics helped better match the character selection with the director’s vision. ©Disney/Pixar.

place to ensure the selections were representative of our diversity goals. The character selections were being made on a sequence-by-sequence basis, with no opportunity to see the bigger picture. After production had begun on the first three sequences, we decided to manually track the selections that had already been made. They showed that without any simultaneous tracking, the results amounted to 50% BIPOC, 50% White, 70% male, and 30% female. Not only was this far below our diversity goals, but it also meant that 20% of our characters were being overused, showing up on screen 50% of the time. Once this data was made available, character selections were swapped out and the immediate issue was resolved. We continued to track the data throughout production, being mindful of the statistics when making character selections. The final film’s diversity statistics were 80% BIPOC, 20% White, 48% male, and 52% female, which was much more in line with the representation we were aiming for.

The lesson learned here was that when the production crew had diversity data readily available to them they had the opportunity to make intentional decisions and stay in line with our goals, making us far less susceptible to unconscious bias [Kim et al. 2021]. We immediately started working to embed demographic tagging into our pipeline so future shows would be able to track more easily.

2 IMPLEMENTATION

Our goal was to be able to track the true realized diversity of our character pool, using real metrics on the screen presence of visible characters, along with statistics for how that character was vocalizing in each shot and sequence of the film. We can use this information to both check that we are honoring the diversity aspiration in our scripts, art and stories, but also to make agile character selection decisions, and create a closer working relationship between script and screen. We do this in four stages: tagging, identifying, aggregating, and visualizing.

The art department determines the values for demographic characteristics - such as, but not limited to: gender, race, age and body type - associated with each member of our character pool. This information allows us to create a tagging dictionary of art assets to character variant names - which is consulted whenever the digital character models are built, and used to populate a models table entry in our production database during the install (mapping model name to demographic breakdown).

When digital characters are placed into shots of the film, the shots are rendered using USD imaging [Universal Scene Description [n.d.]] and processing of the results allows us to efficiently identify which characters are visible for each frame of the shot - we store the proportion of the screen space the projection of the character’s bounding box covers in an entry of a database tracking which models are used in which shots.

We can then aggregate the data from the tagging and identification passes to extract statistics on not only which models are present in which shots, but what kinds of diversity characteristics are featured, filtering and sorting these results using a variety of criteria in a web based front end where one can visualize, for example, what proportion of Asian, Male, Seniors are featured in a particular sequence, shot or indeed across the entire film.

We also use knowledge of “who said what” from the original dialogue sessions, encoded in OTIO metadata [OpenTimelineIO [n.d.]] that travels with the edit of the film, to track the vocalization duration and word count of each of the characters in each shot - thus we can use audio cues as another means to track and visualize how “present” different diversity subpopulations of our character pool are represented in the audio cut of our film.

3 USE CASES

On the film *Turning Red*, our demographic tagging proved valuable for hitting directorial notes concerning the authenticity for the look and feel of a Toronto Chinatown middle school. The number of East Asian students seemed low to the director and wanted to see at least twice as many. When the crowds team consulted the data from *Character Linker*, the screen presence weighted percentage for the East Asian category was about 20%. The team then replaced key groups of characters with significant screen presence, regenerated the metrics and showed the results to the director when the numbers were above 40%. This new character placement was approved and helped validate the utility of our system. Another useful aspect of monitoring our character tags and screen presence in the character linker was that anomalies would stand out and indicate mistakenly placed characters. When characters would show up without tags, it would often indicate an obsolete character asset that was accidentally present and could have caused unforeseen issues. By catching these mistakes early, the crowds team was able to head off problems before they compromised expensive renders.

4 FUTURE WORK

While our screen space presence metric is well correlated with perceived presence, it’s an approximation that can be improved. We hope to try applying recent work in Machine Learning models to predict which pixels are most salient to the viewer [Doggett et al. 2019] and weight our presence metrics accordingly.

REFERENCES

- Erika Doggett, Anna Maria Wolak, P. Daphne Tsatsoulis, and Nicholas McCarthy. 2019. Neural pixel error detection. *ACM SIGGRAPH 2019 Talks* (2019).
- Theodore Kim, Holly Rushmeier, Julie Dorsey, Derek Nowrouzezahrai, Raqi Syed, Wojciech Jarosz, and A. M. Darke. 2021. Countering Racial Bias in Computer Graphics Research. arXiv:2103.15163 [cs.GR]
- OpenTimelineIO. [n.d.]. Project Website: <https://opentimeline.io/>. Title = Project Website: <https://openusd.org> Universal Scene Description. [n.d.].