(un)BiasViz: An Interactive Tool for Balancing Visual Biases in Video Streams

Naimul Hoque
Stony Brook University
mdhoque@cs.stonybrook.edu

Nazmus Saquib
MIT Media Lab
saquib@mit.edu

Syed Masum Billah
Pennsylvania State University
skb5969@psu.edu

Klaus Mueller
Stony Brook University
mueller@cs.stonybrook.edu

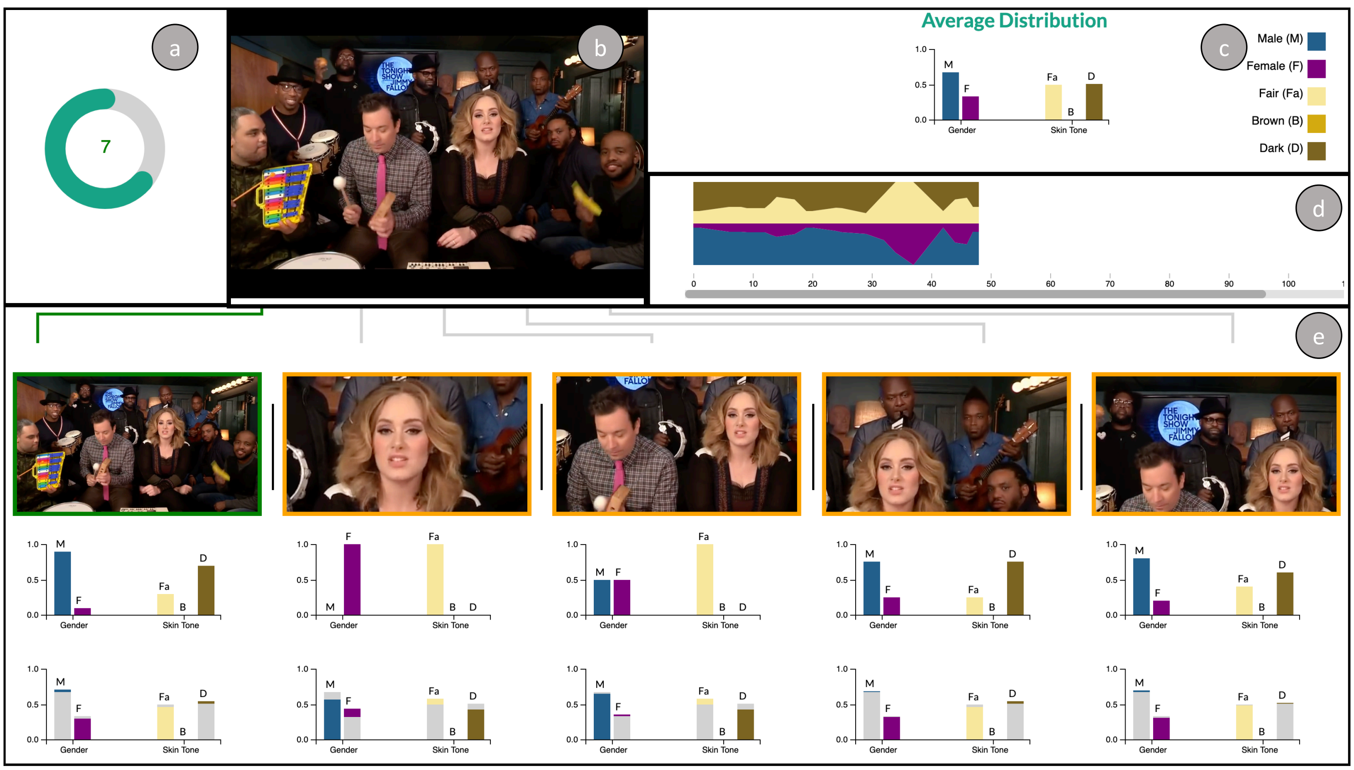

Figure 1: Components of *(un)BiasViz*: **(a)** a countdown timer showing when the charts and graphs will be updated next; **(b)** the output stream; **(c)** bar-charts and **(d)** a stream graph showing the cumulative distribution and the timeline distribution of visual biases (e.g., gender, skin-tone) in output stream until now; **(e)** input streams (e.g., cameras with different angles and views) along with two bar-charts each, one showing visual biases in that stream 10 seconds in advance, and another showing how choosing that particular stream will affect the overall distribution of visual biases in (c) in the next 10 seconds.

ABSTRACT

Visual media such as movies, TV shows, live streams pass on gender and skin-color biases through the portrayals of the characters appearing in the video content. Researchers and human right organizations have long been voicing concerns over such biases in visual media. As a result, media producers have become increasingly aware of these biases. However, they struggle to balance the distribution of gender and skin-color in visual media during the production cycle, as the production system they typically use lack support in this regard. In this paper, we propose (un)BiasViz, an interactive, visual tool that assists media producers to evaluate gender and skin-tone distribution in realtime. A user study with 20 participants shows 43% reduction of the gender bias and 44% reduction of skin-color bias while producing live streams from multiple camera-feeds—all of which are indicative of the promise and potential of (un)BiasViz in commercial production system.

CCS Concepts

•**Human-centered computing** → **Visual analytics;**

INTRODUCTION

Visual media play an important role in our society – not only visual content in media is seen as a reflection of societal beliefs, but also societal beliefs can be influenced by the media. Researchers have shown that if visual content mostly revolves around characters of any specific gender or race, children may learn to think less of other marginalized groups in society [7]. Meanwhile, several other studies have reported that visual media is afflicted toward a certain gender or race [29, 27].

Such concerns have brought forth research on detecting visual biases in media. For example, the *Bechdel Test* measures gender bias in movies by simply asking whether a movie features at least two female characters talking to each other about something other than a man [1]. Alarmingly, many Hollywood movies do not pass this test. Big organizations, such as UNESCO, the Geena Davis Institute, and UN Women have made an enormous effort to bring awareness to these implicit biases and stereotypes through rigorous research [32, 26, 25, 2], in an attempt to make content-creators more conscious. Recently, researchers have developed a number of artificial intelligence (AI) based techniques to automatically detect and mitigate biases in automated domains, such as texts [4, 19, 16, 21].

However, using AI to automatically mitigating biases in creative task such as video production is hard, mainly because content creation is a form of art, often directed by scripts and the artistic view of the creators. For example, when a director makes a male-dominated film this might be because he or she is not biased toward male characters, but because the plot requires showing more men than women.

In this paper, we present ***(un)BiasViz***, an interactive visualization tool for content creators that (i) automatically extracts visual biases from a video in production; (ii) displays these biases both cumulatively and historically; and (iii) offers visual cues on how to reduce biases in the remaining part of that video (see Figure 1). While we specifically consider two types of visual biases, namely, gender bias and skin-color bias, (un)BiasViz can be generalized to other types biases, such as racial bias, political bias, and others. We use off-the-shelf computer vision packages for the first step of our workflow.

The concept of (un)BiasViz is rooted in the observation that during the production of a live show professional studios typically use multiple cameras capturing different perspectives of the scene, such as zooming into a character or keeping a master angle that shows all of the characters or only one individual in a frame. The producers can then choose any of those camera-feeds at any time. Consequently, these camera feeds are expected to have distributions of characters of varied gender and skin-color constellations. (un)BiasViz proactively displays visual biases for each camera-feed using easy-to-compare bar-graphs. It also presents the impact of choosing an individual camera-feed in the near future, enabling producers to make an informed decision as they choose the next camera-feed.

In a user study with 20 participants, we found (un)BiasViz to be effective in reducing gender bias by 43% and skin-color bias by 44%, without increasing the mental workload of the study participant.

Our paper makes the following research contributions:

- The design and development of *(un)BiasViz* as the first interactive visual tool for moderating visual biases in live video streams.
- An evaluation of both performance and user satisfaction that can be achieved with (un)BiasViz.

RELATED WORK

Gender Bias in Media

Several researchers studied the relationship between gender and media, and how gender bias affects general societal beliefs [35, 28]. Researchers also studied the effect of media-imposed gender biases on particular demographics. Sutton et al. (2002), Brown et al. (2005), and Bogt et al. (2010) identified that children or teenagers are more susceptible to inherit gender stereotypes at an early age than any other demographic [30, 6, 31]. Gadassi et al. (2009) found that the gender portrayal in media can affect an individual's career choice [11]. Other studies reported the effects of media in making violence against women [24], and low self-esteem of women [10, 3] socially more acceptable.

Skin-Color Bias in Media

As gender bias adversely affects women, discrimination based on skin-color disadvantages dark-skinned people [15]. Ben-Zeev et al. (2014) found that educated black men appear lighter in the mind of their peers [5]. Other researchers linked darker skin color to smaller incomes, lower marriage rates, longer prison terms, and fewer job prospects [15, 17, 18].

Computer-Vision Aided Visual Analysis of Video Streams

There has been recent work in which researchers employed computer vision techniques to automatically analyze gender biases in visual media. One such tool is GDIQ [25], which showed that men appeared more often than females in almost all types of movies, even in female-led movies. RTFace [34] is another live video analysis platform that uses face recognition system to blur peoples' faces in live video. Our work is in line with these efforts; however, unlike GDIQ and RTFace, which either report historical gender stereotypes or blur faces in video, our tool assists content-creators to actively moderate gender bias directly during the production cycle.

FIELD STUDY

In order to understand how live videos are produced, we visited two professional production studios. We observed that the live studios had two main devices: (i) a *Controller*, and a *Switcher*. The controller had multiple input ports that were connected to

CHI'20, April 25–30, 2020, Honolulu, HI, USA

ISBN 978-1-4503-6708-0/20/04. . . $15.00

DOI: https://doi.org/10.1145/3313831.XXXXXXX

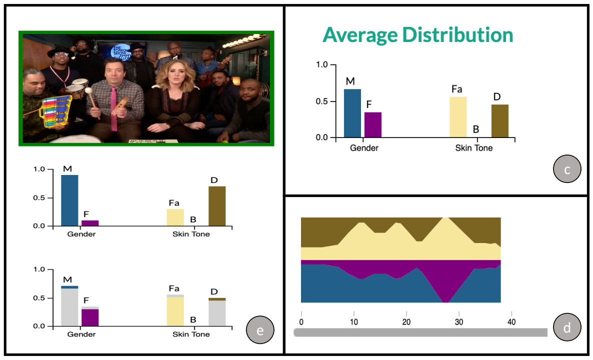

Figure 2: The regions ***(c)***, ***(d)***, and ***(e)*** of (un)BiasViz, as shown in Figure 1.

multiple cameras, and output ports to route the input streams to the output stream. The Controller was connected to the Switcher, a visual interface, which was used to select one of the available input streams as the output stream. An input stream could be selected by pressing a designated button on the Switcher, or by clicking on the input stream on the Switcher interface. The combination of Controller and Switcher is commonly referred to as *Video Switcher*.

Another feature is *Broadcast Delay (BD)*, a 7 to 10 seconds temporal delay deliberately put into the signal path from Controller to Switcher. It enables producers to moderate the video content at the Controller before broadcasting. One application of this delay is the bleeping of child-inappropriate content.

Design Guidelines

These insights gave way to the following design guidelines:

- G1. The interface would emulate the Switcher functions.
- G2. The controller would run available computer vision packages to extract visual biases from an input video stream.
- G3. The broadcast delay would allow users to select the input video stream most suited to mitigate evolving bias.
- G4. Due to the short broadcast delay. our visual information displays would allow for quick insight and comparisons.
- G5. The interface would update the status of all visualization components in sync with the broadcast delay.

(UN)BIASVIZ

We use Python as a server side language to process the video streams in our system. A frame rate of 1 fps was chosen to analyze the video frames. The rationale behind this is that the faces in the screen will very rarely change within 1 second.

Face Detection

Face detection is the most important part of our analysis pipeline as most of our analyses are based on the attributes extracted from faces. We used dlib, an open source Python/C++ library, to detect faces from a video frame. Dlibs facial landmark detector is built on the work of Kazemi and Sullivan [20] and returns a total of 68 points which in turn can be used to detect mouth, eyes, jaws, and nose.

Gender Detection

For the gender detection from face images, we have used a trained convolutional neural network model as proposed by Levi and Hassner [22]. It is claimed to have a 86% accuracy rate on the Adience benchmark [8]. We created our own test set of 1000 faces extracted from our video sources and found our accuracy rate to be 89%.

Skin Tone Detection

The task of detecting a person's race solely from a face image can be challenging even for humans, especially when one is not familiar with the race of the characters appearing in the video. Neural networks have shown good performance on detecting race from images for people from some specific geographic regions [33] but not in general. We thus opted for skin-tone as a feature instead of race because we wanted our tool to be generally applicable – we cannot guarantee that a user will be accustomed to the races that can appear in the video and so take an informed decision based on racial perception. Conversely, skin-tone is easy to recognize visually for humans and computers alike and there is very little chance of disagreement among different users in this matter. Skin-tone has been used to identify race in many cultures, albeit this is a method that has seen some debate in the literature.

For detecting skin-tone from face images, we leverage unsupervised learning (k-means) for labelling the face images. Our skin-tone labels are *Fair*, *Brown*, and *Dark* in accordance with the racial categorization defined by the United Kingdom (UK)- *White*, *Asian*, and *Black*. We used the *CelebA* dataset [23] to train our skin-tone detection model. *CelebA* contains $202,599$ face images of $10,177$ different celebrities. For each image we used dlib's landmark detection framework to draw a bounding box around a face excluding the hair area. The pixels around eyes and mouths were also excluded as those pixels might disrupt the skin-tone estimation if the person wears sunglass or shade, or has lipstick in the mouth or has a mustache.

We apply k-means on the selected pixels with $k = 3$. Since, the images only contain face pixels and we have excluded possible disrupting pixels from the faces, the biggest cluster among these 3 clusters should contain the pixels that will indicate the skin-tone of the faces. We take the centroid of the biggest cluster as an estimation of the skin-tone of the faces. After we get the estimations of the skin-tones (hexadecimal values), we again apply a k-means with $k = 3$ on these skin-tone estimations to find the skin-tone class of each image. Based on the predefined classes and the clustering result, we assign appropriate skin-tone label (*Fair, Brown, and Dark*) to each image. To evaluate the performance of the labelling task, one of the authors measured the accuracy on one thousands images.The accuracy rate for the labelling task was 98%. Finally, we train this labeled *CelebA* dataset on a two hidden layered Convolutional Neural Network. The network achieved an accuracy rate of 93% on the testing set of *CelebA*.

Visual Interface

Figure 1 represents (un)BiasViz on a simulated live-stream setup. The visual interface is divided into five different regions: (a), (b), (c), (d), and (e).

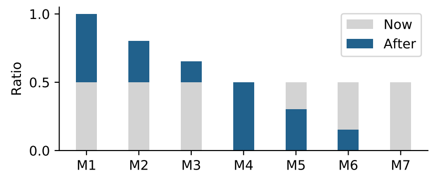

Figure 3: Examples of the Difference or Delta plot used in (un)BiasViz. This plot is used to show comparison between current and future distribution of a bias variable.

Region (a) includes a 10 second countdown timer (consistent with the broadcast delay) to indicate when the charts will be updated. The timer resets to 10 upon reaching a value of 0. Region (b) consists of a simple video interface that allows the user to see the current broadcast of the live-stream. There are no interactions or additional information associated with this region.

Bias Distribution

Region (c) and (d) (figure 1 and 2) provides two charts (bar chart and stream graph) to present the overall bias distribution of the live-stream. These two charts update every 10 seconds. The short duration of this time span requires a data representation that allows users to quickly compare the distributions visually at minimal mental load. These type of visual comparison task is often referred as *Comparative Information Visualization* [14, 13]. According to Gleicher et al. [14, 13], designing uniform guidelines for comparative visualization is difficult as it often depends on the application of the system and the actions required by the user. We chose bar charts since they allow easy comparisons based on height. Pie charts, om the other hand, require users to make comparisons based on angle which can be difficult when the sectors have similar values [9]. Using stacked bar charts in place of a horizontal distribution of bars have similar problems than pie charts. They also make it difficult to compare stack segments with similar values. We hence concluded that presenting the bias variables in form of simple bar charts, placing the categories such as Male, Female side by side would allow users to quickly evaluate the overall distribution, assess their values visually, and take an informed decision with minimal cognitive load.

The color legends for each of the bias variables are also presented in region c. Here we avoided the use of green and red color for a sensitive binary attribute like *gender* since these colors often elicit a sense of danger and safety, and we did not want to inject this type of bias into our visualizations.

Timeline

We used a stream graph as an alternative to the commonly used line charts to show the trend of the bias variables over time. While line charts are good at showing trends over time, they are not suitable to visualize non-binary categorical variables such as skin-tone. In order to eliminate small-scale jitter from the display, we applied *Moving Average* filtering to smooth the stream graph. That is, for any time t, the stream graph has a value that is the average of the values from time $t-3$ to time $t+3$. The horizontal axis of the stream graph dynamically increases as the time increases in the live-stream.

Input Camera Feeds

The bottom row of (un)BiasViz (region e of figure 1) shows the set of input camera feeds in a horizontal arrangement. At any moment during the live-stream a user can click on any of the camera feeds which will connect it to the main video (region b). The selected feed is highlighted with a green border while the other feeds maintain their orange borders. To visually convey the flow of video data we draw a green-colored line between the selected stream and the output stream.

An example of the bar charts associated with one camera feed is shown in figure 2. The first chart associated with a feed shows the average distributions of the bias variables in the next 10 seconds calculated from that feed.

The second chart associated with each camera feed is a stacked bar chart, which shows the after effect of selecting a camera feed. The main objective of this chart is to present how the selection of a feed effects the average distribution in figure 2(c). For this we provide a new variant of a stacked bar chart, which we call *Delta Plots*. An example of different conditions that can appear in *Delta Plots* is shown in figure 3. In *Delta Plots*, the current value of a category is always shown in grey color. For example, in figure 3 the current value of the male ratio is the same in all 7 bars (0.5). The after effect (the value in the next bd seconds) is shown in *blue* color. The first 3 bars (F1, F2, and F3) in figure 3 show the case when the after effect increases from the current value. In F1, the current value is 0.5 and then after effect is an increase of 0.5. To visualize a reduction, the order of the stacked bar chart is reversed –grey bar over blue rather than blue bar over grey bar. The bars F5, F6, and F7 represent a reduction in the after values from the current values. The now and after value pairs for these 3 bars are $(0.5, 0.30)$, $(0.5, 0.15)$, and $(0.5, 0)$. We can imagine a line that touches the tops of the colored bar portions which clearly shows the downward trend implied by these set of bars. The *Delta Plots* are designed for quick comparisons – at any moment the user needs to only see the colored bar to determine the after effect, and it direction of impact, when choosing a camera feed. Figure 4 presents an example scenario to demonstrate how (un)BiasViz can be used to balance visual bias in the video streams.

EVALUATION

We conducted a user study to assess the effectiveness and usability of (un)BiasViz. Specifically, we aimed at validating the following 2 hypotheses:

- **H1**: A live video production software with (un)BiasViz will be more effective in balancing visual biases, than without it.
- **H2**: (un)BiasViz tool will be simple and easy to use.

Participants

We recruited 20 participants (11 males, 9 females) through local mailing lists, social networks, and word-of-mouth. The participants varied in age from 19 to 35 ($M = 25.9, SD = 3.81$), gender ($Male = 11$, $Females = 9$). Our inclusion criteria

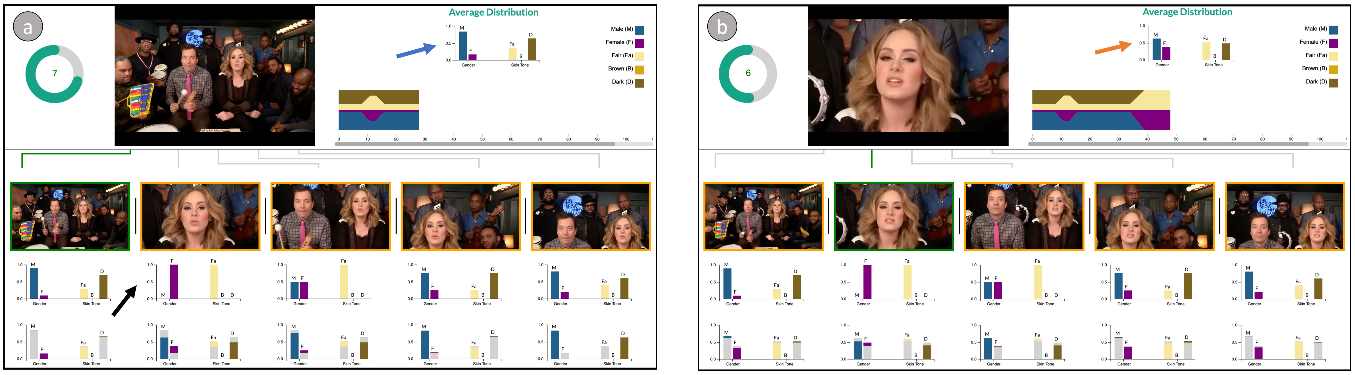

Figure 4: Interaction with *(un)BiasViz*. (a) During live-streaming the producer observes that both the gender and skin-tone distributions are unbalanced (blue arrow). The producer decides to change the feed to balance the distribution. By glancing over all the distributions and the imminent effect associated with each camera-feed, the producer decides to select the second camera-feed, as it will balance both the gender and the skin-tone distribution (orange arrow). (b) The distribution of gender and skin-tone twelve seconds later, which are more balanced than (a) (orange arrow).

included familiarity with video editing, content creation, post-production, news broadcasting, and live streaming. Among 20 participants, there were 2 professional filmmakers, 1 camera operators, 1 video advertisement makers, 3 journalist, and the rest were aspiring YouTubers. All of them had undergraduate degrees in different majors.

User Task and Study Setup

The participants were required to produce a 3-minute video clip from 5 different camera feeds, while balancing the distribution of male and female actors in the final clip, as well as their skin-colors.

To acquire real-world, pre-production videos for our study, we sought to collect videos that were available online and were captured from different camera angles to simulate a live-stream setup. Finding such videos was difficult, however, because content creators already edited and compiled these videos from different camera angles. We specifically searched for videos (i) that were shot as a single camera angle, or (ii) had multiple angles stacked together in a single frame. We found several musical shows meeting our first criterion; we cropped out different parts of such videos to simulate different camera angles. We also found several news shows meeting our second criterion, in which we cropped each portion showing only one character to re-crate the live studio setting. While cropping, we ensured that the gender and skin-tone distribution of in each cropped feed was different, and no single feed provided absolute parity to the average distribution of gender and skin-tone. In total, we prepared 6 videos from 3 different genres, such as talk-shows, news, and live music performance.

To facilitate remote participation, we deployed our prototype on the web, hosted on a MacBook Pro laptop running an Apache web server. Five participants conducted the experiment remotely, and the rest conducted it on the aforementioned MacBook Pro laptop in our lab. When administrating remote studies, the experimenter communicated over Skype.

Design

We conducted a repeated measures within-subject experiment.

C1. Simple Video Editor: This prototype represented a basic video editing software (e.g., iMovie), where users can choose a video-clip from a list of available clips at any time, to produce the final video-clip. This was our *baseline*.

C2. Simple Video Editor + (un)BiasViz-Lite: This prototype only visualized the cumulative and timeline distributions of biases in the output clip, as shown in Figure 2.c and 2.d).

C2. Simple Video Editor + (un)BiasViz: This was the full prototype, as shown in Figure 1.

To minimize the learning effect, we counterbalanced the ordering of study conditions and task videos. We allotted ∼15 minutes for each practice session. At the end of each condition, we administrated a NASA-TLX and a SUS questionnaire about the recently used video editor. The experimenter took notes during the session. All sessions were video recorded and transcribed. Each session lasted for 60 minutes, and culminated with participants making suggestions and recommendations.

Data Collection and Analysis

We analyzed the experimenter's notes, logs and recorded videos to measure the following metrics: (i) the distribution of male and female in the produced video clips; (ii) the distribution of skin-tones (e.g., white, brown, and dark); and (iii) the number of times participants changed camera feeds. We also reported the following subjective measures: (i) coherency of the produced video clips when using (un)BiasViz, and (ii) perceived workload and SUS scores.

We performed a repeated-measures ANOVA with study condition as the independent variables. For our statistical analyses, we used the Greenhouse-Geisser correction for correcting violations of sphericity, and post-hoc tests using a paired t-test with the Bonnferroni correction.

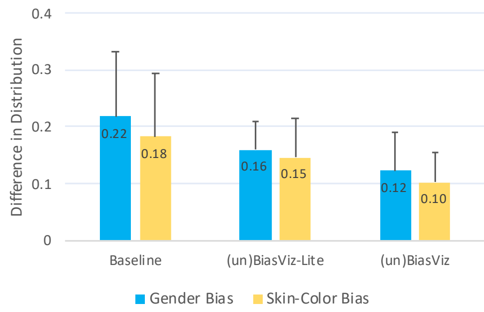

Figure 5: The average differences in distributions of gender and skin-colors in the final clip produced using 3 study editors. Error-bars show +1 SD.

Distribution of Gender

We found a significant effect of study condition on the differences in gender distribution, i.e., gender bias, in the produced video clips, $F(1.392, 26.453) = 6.602, p < .010$. On average, these biases were 0.218 ($SD = 0.112$) for baseline, 0.159 ($SD = 0.049$) for (un)BiasViz-Lite, and 0.123 ($SD = .066$) for (un)BiasViz, as shown in Figure 5. While using (un)BiasViz, the participants reduced gender bias by 43% compared with baseline, and by 22% compared with (un)BiasViz-Lite. However, only the former was found to be statistically significant ($p < .032$). Even though using (un)BiasViz-Lite yielded 27% reduction of bias compared with baseline, this reduction was not statistically significant.

Distribution of Skin-Color

We also found a significant effect of study condition on the differences in skin-color distribution, i.e., skin-color bias, in the produced video clips, $F(2, 38) = 5.683, p < .007$. As shown in Figure 5, the average skin-color biases were 0.183 ($SD = 0.111$) for baseline, 0.145 ($SD = 0.069$) for (un)BiasViz-Lite, and 0.102 ($SD = .0518$) for (un)BiasViz. Between baseline and (un)BiasViz, the bias was reduced by 44%, which was found to be statistically significant ($p < .027$), as expected. No other comparisons were significant.

Results

Switching Camera

We anticipated that the participants would frequently switch cameras when using (un)BiasViz. Surprisingly, we found that switching camera occurred most frequently when they used baseline ($M = 15.250$, $SD = 8.503$), followed by (un)BiasViz-Lite ($M = 11.550$, $SD = 5.443$), and (un)BiasViz ($M = 13.650$, $SD = 7.169$). However, these differences were not significant, as reported by a one-way repeated measures ANOVA.

In our post-study discussion, several participants mentioned that while using the baseline, they tried to remember what camera-feed they had chosen in the past. But at some point they became clueless and started to choose a camera-feed randomly and more frequently.

Video Coherency

To assess the quality and coherency, we asked three human evaluators to rate each video produced by using (un)BiasViz on a scale of 3 (1=incoherent, 2=partially-coherent, and 3=coherent). The inter-annotator agreement was substantial (Fleiss' $\kappa = 0.63$). The final verdict was made via majority voting. Out of 20 video clips, 16 clips were rated as coherent, 3 as partially-coherent, and only 1 as incoherent.

One of our participants who as a professional film-maker commented on the video quality of the produced clip as follows: "*It is very difficult to measure how good a video is. In terms of coherency, it is people's natural instinct to go with a video that has activity in it. Having said that, you often see live shows that frames one person and someone else is actually talking in the show. No one complains about them. Yes, I saw some minor inconstancy in the videos, but that is probably because the gender and skin-tone distributions among the videos were itself disproportional.* "

Perceived Workload and Subjective Feedback

We found a significant difference among the workload under three conditions, $F(2, 38) = 3.262, p < .049$. The participants reported that balancing visual biases in videos without any aid was mentally demanding ($M = 44.542$, $SD = 16.583$). Surprisingly, they reported that their workload decreased the most ($p < .040$) when using (un)BiasViz-Lite ($M = 35.747$, $SD = 13.312$), rather than using (un)BiasViz($M = 37.272$, $SD = 17.269$).

We observed similar trend in SUS scores – SUS score increased the most in (un)BiasViz-Lite ($M = 85.416$, $SD = 5.103$), followed by (un)BiasViz ($M = 77.083$, $SD = 5.685$), and baseline ($M = 51.250$, $SD = 17.084$).

In the post-study discussion, the participants attributed glancing over multiple bar-chars at once as the reason why their workload marginally increased and SUS score marginally decreased in (un)BiasViz compared with (un)BiasViz-Lite. However, they mentioned that they would get used to it with practice.

DISCUSSION

While all participants mentioned that (un)BiasViz would benefit professional studios in balancing visual biases, some of them were concerned about the potential misuse of this tool. One of them asked *"what if someone uses (un)BiasViz to increase bias in the produced video"?* There is no straightforward answer to this question, but probably the best answer is to increase the consciousness among content creators, give them a tool, and then hold them accountable afterwards.

Although we designed (un)BiasViz as an assistive tool that can be used to balance bias in any way the creator seems fit, some expert participants concerned about (un)BiasViz hindering their artistic freedom. Some other participants mentioned that there was a certain danger that people could "game" the system when it became too transparent. In the literature this is known as *transparency-gameability trade-off*. However, as discussed by Ghani [12], gameability is acceptable as long as it incentivizes people to modify their behavior in positive ways.

Limitations

In this work, we loosely defined gender and skin-color biases as the differences in distributions of male and female characters, and white, brown, and dark skin-colors, respectively. However, it is easy to make a live stream that is perfectly balanced, yet biased towards a target population. Another limitation is we evaluated our tool on a simulated live-streaming setup, using 3-minute videos. While the results were promising, it still needs to be evaluated on an actual production studio.

CONCLUSION

(un)BiasViz tackles a hard problem—how to moderate visual biases in live video streaming or broadcasting. Towards that, (un)BiasViz is a computer-vision based interactive visual analytics tool that assists content moderators or producers. This tool can be integrated into any production system. (un)BiasViz can be also be generalized to moderating other types of visual biases. To the best of our knowledge, there is no production system that provides feedback on visual biases as the video is being produced. In other words, visual bias is a video metric that has not yet been acknowledged by production systems in use. (un)BiasViz can bridge this gap.

REFERENCES

[1] 2019. Bachdel Test. (2019). https://bechdeltest.com/ Accessed: 2019-04-03.

[2] 2019. UN Women. (2019). https://www.unwomen.org/en Accessed: 2019-04-03.

[3] Jennifer Stevens Aubrey. 2006. Exposure to sexually objectifying media and body self-perceptions among college women: An examination of the selective exposure hypothesis and the role of moderating variables. *Sex Roles* 55, 3-4 (2006), 159–172.

[4] Solon Barocas, Moritz Hardt, and Arvind Narayanan. 2017. Fairness in machine learning. *NIPS Tutorial* (2017).

[5] Avi Ben-Zeev, Tara C. Dennehy, Robin I. Goodrich, Branden S. Kolarik, and Mark W. Geisler. 2014. When an “Educated” Black Man Becomes Lighter in the Mind’s Eye: Evidence for a Skin Tone Memory Bias. *SAGE Open* 4, 1 (2014), 2158244013516770. DOI: http://dx.doi.org/10.1177/2158244013516770

[6] Jane D Brown, Carolyn Tucker Halpern, and Kelly Ladin L’Engle. 2005. Mass media as a sexual super peer for early maturing girls. *Journal of adolescent Health* 36, 5 (2005), 420–427.

[7] Jane D Brown and Carol J Pardun. 2004. Little in common: Racial and gender differences in adolescents’ television diets. *Journal of Broadcasting & Electronic Media* 48, 2 (2004), 266–278.

[8] Eran Eidinger, Roee Enbar, and Tal Hassner. 2014. Age and gender estimation of unfiltered faces. *IEEE Transactions on Information Forensics and Security* 9, 12 (2014), 2170–2179.

[9] Stephen Few and Perceptual Edge. 2007. Save the pies for dessert. *Visual Business Intelligence Newsletter* (2007), 1–14.

[10] BL Fredrickson and T Roberts. 1997. Toward understanding womenâĂŹs lived experiences and mental health risks. *Psychology of Women Quarterly* 21, 2 (1997), 173–206.

[11] Reuma Gadassi and Itamar Gati. 2009. The effect of gender stereotypes on explicit and implicit career preferences. *The Counseling Psychologist* 37, 6 (2009), 902–922.

[12] Rayid Ghani. 2019. You Say You Want Transparency and Interpretability? (2019). http://www.rayidghani.com/you-say-you-want-transparency-and-interpretability

[13] Michael Gleicher. 2017. Considerations for visualizing comparison. *IEEE transactions on visualization and computer graphics* 24, 1 (2017), 413–423.

[14] Michael Gleicher, Danielle Albers, Rick Walker, Ilir Jusufi, Charles D Hansen, and Jonathan C Roberts. 2011. Visual comparison for information visualization. *Information Visualization* 10, 4 (2011), 289–309.

[15] Arthur H Goldsmith, Darrick Hamilton, and William Darity Jr. 2006. Shades of discrimination: Skin tone and wages. *American Economic Review* 96, 2 (2006), 242–245.

[16] Moritz Hardt, Eric Price, Nati Srebro, and others. 2016. Equality of opportunity in supervised learning. In *Advances in neural information processing systems*. 3315–3323.

[17] Matthew S Harrison and Kecia M Thomas. 2009. The hidden prejudice in selection: A research investigation on skin color bias. *Journal of Applied Social Psychology* 39, 1 (2009), 134–168.

[18] Joni Hersch. 2011. The persistence of skin color discrimination for immigrants. *Social Science Research* 40, 5 (2011), 1337–1349.

[19] Matthew Joseph, Michael Kearns, Jamie H Morgenstern, and Aaron Roth. 2016. Fairness in learning: Classic and contextual bandits. In *Advances in Neural Information Processing Systems*. 325–333.

[20] Vahid Kazemi and Josephine Sullivan. 2014. One millisecond face alignment with an ensemble of regression trees. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1867–1874.

[21] Matt J Kusner, Joshua Loftus, Chris Russell, and Ricardo Silva. 2017. Counterfactual fairness. In *Advances in Neural Information Processing Systems*. 4066–4076.

[22] Gil Levi and Tal Hassner. 2015. Age and Gender Classification Using Convolutional Neural Networks. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) workshops*. http://www.openu.ac.il/home/hassner/projects/cnn_agegender

[23] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. 2015. Deep Learning Face Attributes in the Wild. In *Proceedings of International Conference on Computer Vision (ICCV)*.

[24] Sarah K Murnen, Carrie Wright, and Gretchen Kaluzny. 2002. If “boys will be boys,” then girls will be victims? A meta-analytic review of the research that relates masculine ideology to sexual aggression. *Sex roles* 46, 11-12 (2002), 359–375.

[25] Geena Davis Institue on Gender in Media. 2015. The Reel Truth: Women Aren't Seen or Heard. (2015). https://seejane.org/research-informs-empowers/data/

[26] Geena Davis Institue on Gender in Media. 2017. The Geena Benchmark Report: 2007-2017. (2017). https://seejane.org/wp-content/uploads/geena-benchmark-report-2007-2017-2-12-19.pdf

[27] James A Rada. 1996. Color blind-sided: Racial bias in network television’s coverage of professional football games. *Howard Journal of Communications* 7, 3 (1996), 231–239.

[28] Karen Ross. 2010. *Gendered media: Women, men, and identity politics*. Rowman & Littlefield.

[29] Stacy L Smith, Marc Choueiti, and Katherine Pieper. 2014. Gender bias without borders. (2014). https://seejane.org/wp-content/uploads/gender-bias-without-borders-full-report.pdf

[30] MJ Sutton, Jonathon D Brown, KM Wilson, and JD Klein. 2002. ‘Shaking the Tree of Knowledge for Forbidden Fruit. *Sexual teens, sexual media: Investigating media’s influence on adolescent sexuality* (2002), 25–55.

[31] Tom FM Ter Bogt, Rutger CME Engels, Sanne Bogers, and Monique Kloosterman. 2010. “Shake it baby, shake it”: Media preferences, sexual attitudes and gender stereotypes among adolescents. *Sex roles* 63, 11-12 (2010), 844–859.

[32] UNESCO. 2019. Gender-Sensitive Indicators for Media. (2019). http://www.unesco.org/new/en/communication-and-information/crosscutting-priorities/gender-and-media/gender-sensitive-indicators-for-media Accessed: 2019-04-03.

[33] Thanh Vo, Trang Nguyen, and C Le. 2018. Race recognition using deep convolutional neural networks. *Symmetry* 10, 11 (2018), 564.

[34] Junjue Wang, Brandon Amos, Anupam Das, Padmanabhan Pillai, Norman Sadeh, and Mahadev Satyanarayanan. 2017. A scalable and privacy-aware IoT service for live video analytics. In *Proceedings of the 8th ACM on Multimedia Systems Conference*. ACM, 38–49.

[35] Julia T Wood. 1994. Gendered media: The influence of media on views of gender. *Gendered lives: Communication, gender, and culture* 9 (1994), 231–244.