

US Patent & Trademark Office

Patent Public Search | Text View

United States Patent Application Publication

20250259446

Kind Code

A1

Publication Date

August 14, 2025

Inventor(s)

Patel; Koninika et al.

SYSTEM AND METHOD FOR DETECTING AND VISUALIZING DEMOGRAPHICS, DIVERSITY AND DISPARITY IN USER- GENERATED VIDEOS

Abstract

A system for detecting and visualizing demographics, diversity and disparity in user-generated videos includes a collection of user-generated content, such as videos, having demographic information, a target collection of content having defined diversity information, creator content having diversity information, and a processor capable of performing a diversity analysis on the user-generated content, providing an indication of disparity between the user-generated content and the target collection of content, and taking an action based on the analysis to modify the collection to more closely reflect the target collection of information such as providing the disparity analysis to a vendor so that the vendor can prepare a promotional campaign to a target group of customers more closely identified with its demographic profile.

Inventors: Patel; Koninika (Sunnyvale, CA), Rosario; Barbara (Berkeley, CA), Reed; Daina (Santa Cruz, CA)

Applicant: Vyrill Inc. (San Francisco, CA)

Family ID: 85038517

Appl. No.: 19/183721

Filed: April 18, 2025

Related U.S. Application Data

parent US continuation 17877685 20220729 ABANDONED child US 19183721

us-provisional-application US 63227321 20210729

Publication Classification

Int. Cl.: G06V20/40 (20220101); **G06F3/0481** (20220101); **G06Q30/0201** (20230101)

U.S. Cl.:

CPC G06V20/41 (20220101); **G06F3/0481** (20130101); **G06Q30/0201** (20130101);
G06V20/48 (20220101);

Background/Summary

CROSS-REFERENCE TO RELATED APPLICATIONS [0001] This application is a continuation application of application Ser. No. 17/877,685, filed Jul. 29, 2022, which claims the benefit of U.S. Provisional Application No. 63/227,321, filed Jul. 29, 2021.

BACKGROUND

Field of the Invention

[0002] This invention relates to computer systems for analyzing media content using Artificial Intelligence (AI) or machine learning, and more particularly, to a system for detecting and visualizing demographics, diversity, and disparity in user-generated videos.

Discussion of the Prior Art

[0003] User-generated content (UGC), alternatively known as user-created content (UCC), is any form of content, such as images, videos, text and audio, that has been posted by users on online platforms such as social media.

[0004] For consumers, UGC, particularly user-generated videos (UGVs), has become an increasingly common source of information on brands and their related products and services. UGVs can include reviews, tutorials and demonstrations that directly educate consumers about products or services, or they can indirectly expose consumers to brands through product placement, wherein products or services are incorporated into content such as home videos or web series. On YouTube, the largest online video platform, over 500 hours of videos are uploaded every minute, with over 5 billion videos watched every day, many of which are UGVs featuring brands, products or services.

[0005] There is a growing need for companies to better understand and control the influence UGVs have on consumer perception of their brands. Companies are particularly interested in monitoring demographics (e.g., race, gender and age) of individuals appearing in UGVs alongside their brands and in finding ways to improve upon the diversity of those demographics.

[0006] While prior art methods exist for quantifying diversity, these methods fall short of providing meaningful and actionable insights to brand managers or content creators. In addition, there is a need for a scalable system that can score and visualize diversity among large and ever-growing catalogs of UGVs hosted on platforms such as YouTube.

SUMMARY OF THE INVENTION

[0007] The present invention provides a system for detecting and visualizing demographics, diversity and disparity in UGVs.

[0008] The system is configured to generate diversity and disparity scores for UGVs, where diversity is a measure of the inclusion of different types of people (e.g., race, gender and age) in a video or collection of videos and disparity is a measure that compares the types of people in one video or collection of videos to those in another video or collection of videos. The system visualizes diversity and disparity scores, along with other demographic information, in meaningful and actionable ways, and it can be used by both brand managers and content creators to analyze and take strategic action based on the diversity and disparity scores.

Description

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] Exemplary embodiments of the invention will now be described with reference to the accompanying drawings, in which:

[0010] FIG. 1 shows a user interface or dashboard of a system according to the invention, the dashboard providing a home page for a brand manager to access the system's functionality.

[0011] FIG. 2 shows another view of a user interface or dashboard of a system according to the invention, wherein all videos of a single product are being compared to all videos of a product category.

[0012] FIG. 3 shows another view of a user interface or dashboard of a system according to the invention, wherein diversity and demographic information are shown for all videos of a single product.

[0013] FIG. 4 shows another view of a user interface or dashboard of a system according to the invention, wherein diversity, disparity and demographic information are shown for a single video.

[0014] FIG. 5 shows another view of a user interface or dashboard of a system according to the invention, wherein diversity, disparity and demographic information are shown for a single content creator.

[0015] FIG. 6 shows a user interface component that enables the user to take actions based on diversity and demographic information.

[0016] FIG. 7 shows a user interface component that enables the user to take control the weighted averaging of a cumulative diversity score.

[0017] FIG. 8 shows another view of a user interface or dashboard of a system according to the invention, wherein a content creator can license a video to a company or brand manager.

[0018] FIG. 9 shows another view of a user interface or dashboard of a system according to the invention, wherein a brand manager can create a brand management campaign.

[0019] FIG. 10 shows another view of a user interface or dashboard of a system according to the invention, wherein a content creator can connect with ongoing brand management campaigns.

DETAILED DESCRIPTION OF THE ILLUSTRATED EMBODIMENTS

[0020] There are several known methods of quantifying diversity in a collection, such as a collection of individuals in a video. One of the most common measurements of diversity is the Herfindahl-Hirschman Index (HHI). The HHI typically has a value ranging from 1 to $1/N$, where N is the number of different demographic categories being analyzed in a collection. For example, in calculating the racial diversity of individuals in a video across five different racial groups ($N=5$), a video entirely composed of individuals of just one of the racial groups, such as Asian, would have an HHI of 1, while a video that is made up equally of people from all five racial groups (most diverse), would have an HHI of 0.2. The HHI is calculated using the following formula:

$$[00001] H = \frac{1}{N} \sum_{i=1}^N s_i^2$$

where N is the number of different categories in a demographic being analyzed and s_{i} is the ratio of individuals that belong to a given one of those categories (i).

[0021] The most obvious shortcoming of the HHI is that the index is unable to effectively convey the difference between two collections having the same HHI, but different demographic distributions. For example, a video containing only Asian individuals would have the same HHI as a video containing only black individuals ($HHI=1$). In addition, the range of HHI values (1 to $1/N$) can be confusing to understand and to visualize.

[0022] The system of the present invention utilizes an improved scoring of both diversity and disparity in videos. Compared to prior art indexes, such as the HHI, the improved scoring can be more easily understood and visualized in a software user interface, and the combination of both

diversity and disparity scoring can provide a more complete picture of demographic distribution. [0023] The system's improved diversity scoring is generally calculated in a manner similar to the HHI, but the values are normalized to a range of 0 to 100 (or, alternatively, 0 to 1 or 0% to 100%), such that 0 means not diverse and 100 is most diverse, for example:

$$[00002] \text{diversity} = (1 - \frac{\text{HHI} - \frac{1}{N}}{1 - \frac{1}{N}}) * 100$$

Again, where N is the number of different categories in a demographic being analyzed.

[0024] This normalized diversity scoring can be applied to each demographic of interest in a video or collection of videos, for example, age, gender and race, such that a video or collection of videos can have a separate age diversity score, gender diversity score and race diversity score. In addition, the scoring of different demographics can be weighted averaged to calculate an overall or cumulative diversity score for a video or collection of videos. It should be understood that the diversity scoring can be performed on any demographic of interest, such as relationship status, nationality, education, profession, and as to any parameter such as product type, topic, language, and country, and that any suitable method of calculating an overall or cumulative diversity score can be performed.

[0025] The system's disparity scoring quantifies the difference between the types of individuals featured in a first video or collection of videos and the types of individuals featured in another video or collection of videos. The disparity scoring can be used, for example, to quantify the “uniqueness” of individuals featured in a product video found on YouTube (first video) compared to individuals found in all other videos on YouTube featuring the same product or category of product (collection of videos). The disparity scoring is generally performed using the following formula:

$$[00003] \text{Disparity} = (\sum_{i=1}^N (1 - c_i) * s_i) * 100$$

where c.sub.i represents the number of people per category in a whole collection (e.g., all videos in a product category), s.sub.i represents the number of people per category in an individual video (e.g., first video) and N is the number of demographic categories being considered.

[0026] The disparity scoring allows certain demographic categories to be weighed more heavily than others, such that greater weight can be given to demographic categories that are more “unique” or that occur less regularly in the videos. For example, if we consider only three categories in the race demographic, white, black and Asian, and if the distribution of race within those categories is 80% white, 15% black and 5% Asian among a collection of videos, the disparity scoring of a single video could place more emphasis on the number of Asian individuals featured in the video than on the number of white individuals featured in the video, thus resulting in a higher disparity score for videos featuring “unique” or less commonly occurring types of individuals.

[0027] Therefore, disparity scoring involves a summation of (1-c.sub.i)*s.sub.i where c.sub.i is the percentage of people in a category of a first video or collection of videos and s.sub.i is the percentage of people in the same category in a second video or collection of videos. It should be understood that any suitable alternative method of weighting demographic categories can be applied to the disparity scoring.

[0028] The disparity scoring can be performed for any demographic of interest in a video or collection of videos, and the scoring of different demographics can be weighted averaged to calculate an overall or cumulative disparity score for comparing a video or collection of videos to another video or collection of videos.

[0029] The diversity and disparity scoring allow for many different types of comparisons to be made. Diversity scoring, for example, can be performed on a single video featuring a single product (e.g., Conair Ceramic Hair Dryer), on all the videos on YouTube featuring a single product, on all the videos on YouTube featuring products in the same category (e.g., all Hair Dryers) or on all videos from a single content creator (e.g., YouTube personality or influencer). Disparity scoring can be performed on videos or collections of videos to compare, for example, a single video

featuring a product against all videos on YouTube featuring that product, a single video featuring a product against all videos on YouTube featuring products in the same category, or a single video featuring a product against all videos from a single content creator.

[0030] In one implementation of the system, a vendor, Company X, plans to target customers from demographic groups A and B for a marketing campaign. The goal of the company is to generate 50,000 click throughs from videos related to a selected product assuming an average conversion rate of 5% from the click throughs and an average order size of \$40. That would generate revenue of \$100,000 ($50,000 \times 0.05 \times \40). It is assumed that each video generates 5,000 click throughs, thus requiring 10 videos featuring the product directed at the target demographic. After analyzing 10,000 videos featuring the selected product, the system determines that only 4 videos include content featuring the targeted demographic. The system can then recommend content creators and the type of content to produce the 6 additional videos that are needed to meet the company's revenue goal. The system can automatically or manually launch a campaign by contacting the identified content creators to produce the required videos. Content creators may be fans, social media influencers or shoppers who have purchased the product.

[0031] In another implementation of the system, when a customer having a known profile, e.g., membership in a defined demographic of a population, searches for a product or a product feature, the system can recommend videos culled from a collection of videos that may be uploaded by a vendor of the product for viewing by the customer based on the demographic and diversity content of each of the videos in the vendor's collection. The video recommendations might be across a product category, across one brand or multiple categories and products across multiple brands. In addition to video demographics and disparity scoring, the system can analyze other dimensions such as sentiment, topics, scenes, type of creator, and a video's visual quality. These are merely examples of the types of diversity and disparity scoring that can be performed and the types of comparisons that can be made. Further examples will now be explored with respect to certain embodiments of the system of the invention.

[0032] Referring to FIG. 1, the system of the invention is provided as a software application, such as a web application, software as a service (SaaS), mobile application or the like, having a user interface or dashboard (100). The dashboard (100) allows a user to see diversity and disparity scores, as well as insights gleaned from the scores, and to take actions based on those scores and insights.

[0033] The system of the invention is configured to allow a user to analyze one or more videos, for example, a user can monitor and analyze all of the videos on YouTube that feature one or more products from a single company. Alternatively, a vendor may upload to the system a portfolio of videos related to one or more products for analysis. The exemplary dashboard (100) of FIG. 1 provides a view (120) for a vendor's brand manager (121) to see the demographic (122) and diversity information (124) for all of the vendor's videos (125) featuring products (126). Demographic information (122) can be detected in the videos using any known computer vision or machine learning algorithms or can be manually extracted from the videos. The view (120) displays diversity scores (124) for age, gender and ethnicity across all of the videos being monitored by the brand manager. The demographic and diversity information can be filtered to relate to any one or more products or services.

[0034] In one implementation of the invention a vendor may upload one or more collections of user-generated content relating to a product or video. A diversity analysis may then be performed on the uploaded content to determine the diversity of information in the collection, such as demographic diversity. The collection may then be assigned a diversity score. Based on demographic information on each of the vendor's prospective customers, a disparity analysis can be performed to measure the difference in diversity between the vendor's uploaded content and the diversity of a target group of the vendor's customer base. A subset of the user-generated content may be selected using the disparity analysis that is more closely identified with the vendor's target

group of customers. The subset of the content is then provided to the vendor which can send it to the vendor's customers as part of a promotional campaign that provides the identified customer group with a more personalized experience regarding the product or service.

[0035] FIG. 2 shows another view (220) of the dashboard that enables a brand manager to compare videos she is monitoring of a single product (in this example, an ionic hairdryer being sold under her brand) against a collection of videos featuring products of the same category (in this example, all hairdryer videos on YouTube featuring her brand and the brand of her competitor's). The view (220) shows comparisons of cumulative or overall diversity scoring (222) of individuals in the videos, diversity scores for age (231), gender (232) and ethnicity (233) in the videos, and the distributions of age (241), gender (242) and ethnicity (243) amongst individuals in the videos. The information is displayed using various bar graphs, pie charts and shapes, and a dropdown menu (250) allows the brand manager to change what styles of charts are used to display the information.

[0036] FIG. 3 shows another view (320) of the dashboard that enables a brand manager to see detailed demographic and diversity information for a single product (in this example, all videos on YouTube featuring the Magna Ionic Hairdryer). As described in more detail with reference to FIG. 7, the view (320) also includes a user interface component (700) that enables a brand manager to take certain actions (350) based on the demographic (331), diversity scoring (332) and product information (333) that are displayed for the product and to view high-level insights (340) gleaned from the information.

[0037] FIG. 4 shows another view (420) of the dashboard that enables a brand manager to see detailed demographic information (431), diversity scoring information (432) and disparity scoring information (433) for a single video (450) of a product, including comparisons of that information against that of all videos featuring the same product.

[0038] The system of the invention is configured to include a database of profiles for various content creators or influencers (i.e., users that upload content to one or more social media or content hosting platforms). Each content creator profile contains information about videos uploaded to one or more social media or content hosting platforms by that content creator. FIG. 5 shows another view (520) of the dashboard that enables a brand manager to see information retrieved from the profile for a single content creator (525), including demographic information (531), diversity scoring information (532) and disparity scoring information (533) for videos uploaded by that content creator, and comparisons of that information against other collections of videos. The view (520) also includes a user interface component (550) that enables the brand manager to take action based on information included in the content creator profile, including, for example, searching for profiles for similar content creators, messaging the content creator and adding the content creator profile to a favorites list.

[0039] FIG. 6 shows in further detail a user interface component (600) of the dashboard. The component (600) enables a user, such as a brand manager, to take certain actions based on the demographic information, diversity scoring information and disparity scoring information made available by the system and to view high-level insights gleaned from the information. In the example of FIG. 6, the user can start a promotional campaign, message content creators, export contact information for a content creator, find profiles of content creators, find videos indexed by the system, create video compilations based on videos indexed by the system, create collages of images take from videos indexed by the system, search for SEO ("search engine optimization") terms based on demographics or download detailed reports. It should be understood that many other user actions can be enabled by the system based on the demographic information, diversity scoring information and disparity scoring information made available to the user by the system. In the example of FIG. 6, component (600) displays insight gleaned from diversity information for a collection of videos, including a general assessment of the diversity of individuals in the collection of videos and indications that the videos have low gender and ethnicity diversity scores. It should be understood that many other high-level insights can be displayed by the system based on the

demographic information, diversity scoring information and disparity scoring information made available to the user by the system.

[0040] FIG. 7 shows a user interface component (700) of the dashboard that enables the user to adjust the relative weights assigned to age, gender and ethnicity when the system scores the overall cumulative diversity of a video or collection of videos. It should be understood that the system can be configured to generate overall or cumulative diversity scores based on a weighted averaging of various other demographic categories and that the system can enable a user to adjust the weighting based on any of those other demographic categories.

[0041] FIG. 8 shows another view (820) of the dashboard that enables a content creator (810) to see diversity information (832) for one of their videos (850) and to license that video to one or more brand managers representing various brands within various product categories or verticals. In the example of FIG. 8, the content creator can see recommended verticals and brands to whom they may license their video (850). These recommendations are generated based on the demographic information, diversity scoring information and demographic scoring information the system has generated or stored for videos of the various brands.

[0042] The system of the invention is configured to enable a user, such as a vendor, to create a promotional campaign based on the demographic information, diversity scoring information or disparity scoring information detected, generated and/or made available by the system. For example, in another view (920) of the dashboard as shown in FIG. 9, brand managers can create brand management campaigns to invite content creators to create and upload new UGVs based on their company's products and brands and can target those brand management campaigns towards content creators with profiles that have desirable demographic, diversity scoring or disparity scoring information.

[0043] FIG. 10 shows another view (1020) of the dashboard that enables a content creator to connect to the ongoing brand management campaigns created and managed by brand managers using the system of the invention. In the example of FIG. 10, the content creator can search, sort and connect to various brand management campaigns created and managed by various brands.

[0044] It should be understood that the system of the invention can be configured to incorporate improvements or alternatives to the diversity and disparity scoring described herein, and that the system of the invention can be configured to analyze diversity and disparity in UGC other than recorded videos, including text, images, audio, video games, virtual reality or augmented reality applications, and live audio or video streams.

[0045] There have thus been described and illustrated certain embodiments of a system for detecting and visualizing demographics, diversity and disparity in UGVs according to the invention. These embodiments are merely example implementations of the invention and are not intended to limit the scope of the invention to their particular details. Alternative embodiments of the invention not expressly disclosed herein will be evident to persons of ordinary skill in the art.

Claims

1. A method for detecting demographics, diversity and disparity in user-generated videos to personalize customer experience, the method comprising: receiving from a vendor one or more collections of user-generated content relating to a defined product or service, performing a diversity analysis to determine the diversity of said information of at least one of said one or more collections of user-generated content based upon demographic information associated with a target group of customers, the diversity analysis performed using the formula

$$\text{diversity} = (1 - \frac{\text{HHI} - \frac{1}{N}}{1 - \frac{1}{N}}) * 100$$
 where N is the number of different parameters in a collection being analyzed, selecting a subset of user-generated content more closely identified with said demographic information based on said diversity analysis of said one or more collections of user-

generated content, exporting said selected subset to said vendor, and sending said subset of user-generated content by said vendor to said target group to provide a more personalized experience related to the product or service.

2. The method of claim 1 wherein the user-generated content includes videos.

3. The method of claim 1 wherein the one or more collections of user-generated content includes demographic information.

4. The method of claim 1 further comprising: displaying the one or more collections of user-generated content on a graphical user interface.

5. The method of claim 1 further comprising: said information including information regarding one or more defined parameters.

6. The method of claim 5 wherein the diversity analysis includes quantifying one or more of said parameters of the one or more collections to refine said diversity analysis.

7. A system for detecting demographics, diversity and disparity in user-generated videos, the system comprising: one or more collections of user-generated content relating to a defined product or service, the content having information regarding one or more defined parameters, one or more collections of customer information regarding one or more target groups of customers, demographic indicia associated with each of said one or more groups of customers, a processor configured for: receiving said one or more collections of user-generated content for said defined product or service from a vendor, performing a diversity analysis of the diversity of the information associated with the user-generated content of each of said one or more collections, the diversity analysis performed using the formula $\text{diversity} = (1 - \frac{\text{HHI} - \frac{1}{N}}{1 - \frac{1}{N}}) * 100$ where N is the number of different parameters in a collection being analyzed, assigning a diversity score to the user-generated content of at least one of said one or more collections of user-generated content, selecting a subset of said user-generated content more closely identified with the demographic indicia of at least one of said one or more target groups of customers, and exporting said subset to said vendor for presentation by said vendor to said one of said one or more customers.
