IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 16, 2020, accepted April 1, 2020, date of publication April 6, 2020, date of current version April 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2985673

Exploration of the Characteristics of Emotion Distribution in Korean TV Series: Common Pattern and Statistical Complexity

YUNHWAN KIM[1] AND SUNMI LEE[2]
[1]Division of Media Communication, Hankuk University of Foreign Studies, Seoul 02450, South Korea
[2]Department of Applied Mathematics, Kyung Hee University, Yongin 17104, South Korea

Corresponding author: Sunmi Lee (sunmilee@khu.ac.kr)

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A5B5A07072016).

ABSTRACT The series is one of the most popular formats of television programs, but the emotional aspects of its content have not attracted much attention from researchers. Furthermore, the cinemetric approach, in which move data are quantitively examined, has not been used to analyze affective content in television series. Therefore, this study focused on emotion distribution in Korean television series. In all, 337 episodes from 20 series were divided into shots, then keyframes were selected from the shots. Facial emotions on the keyframes were measured using an online artificial intelligence service. Emotion distributions were obtained from each episode and then were observed for common patterns. In addition, the statistical complexity of the distributions were calculated and compared by the temporal phase of the episode, channel type, genre, ratings, and season. The results show that the emotion distributions of all episodes in all series had the same form: an asymmetric U-shaped distribution biased toward zero. The statistical complexity of emotion distribution was greater in early episodes than in later episodes of a series, and it was higher in episodes with higher ratings. In addition, the complexity of emotion distribution was higher in hard genres and during the summer season. There was no significant difference between terrestrial television channels and cable/general programming channels.

INDEX TERMS Affective content analysis, cinemetrics, emotion distribution, microsoft azure cognitive services, statistical complexity, TV series.

I. INTRODUCTION

The series is one of the most popular television (TV) program formats. This type of TV program is more about delivering stories with affection than about logical reasoning. The stories are delivered mainly through the characters' verbal and nonverbal behaviors, which manifest who they are, what they are doing, and how they feel. Thus, studies of the TV series have been centered around the story it delivers or its content [1]–[4].

Despite the abundance of literature on various aspects of TV series content, relatively little attention has been paid to the expressed emotion. This stands in contrast to the literature on other types of video content. For example, movies mainly consist of human verbal and nonverbal behaviors, as TV series does; thus, it is crucial to analyze the emotions expressed implicitly and explicitly in order to understand a certain aspect of the movie content. In this regard, much effort has been devoted to analyzing the emotions in video content using quantitative methods. Affective content analysis [5] is one of these methods. Rather than annotating emotions by human coders, affective content analysis measures emotions directly from a video using various features. This approach has been adopted to analyze various kinds of video data [6]–[8]; however, the TV series has not been actively investigated in terms of the expressed emotions.

This approach can be considered in the broader context of *cinemetrics*, which is a method to analyze video content quantitatively. Statistical analysis, which has drawn relatively less attention in video research, can be an important aspect of movie analysis [9], and another kind of knowledge of movie can be obtained by utilizing a variety of low-level

The associate editor coordinating the review of this manuscript and approving it for publication was Yu-Huei Cheng.

features, such as shot duration, temporal shot structure, visual activity, luminance, and color [10]. Cinemetrics may also be referred to as *statistical style analysis* [11], [12] because these low-level features manifest the style of a movie. Although cinemetric analysis has been applied to movies, it has not been widely used for analyzing TV programs [13]. Because the styles of TV programs tend to follow the trajectory of development for movie styles [14], cinemetric analysis has much unrealized potential for the analysis of TV programs, including TV series. A small number of studies have adopted cinemetric analysis to TV news [15], [16], TV sitcoms [17], [18], and TV series [13]. This study attempts to extend the approach in the line of these studies.

In cinemetric studies, one of the central issues is determining which metric to use. In previous research, shot length (duration) was the most widely used metric [19], [20]. Metrics such as shot scale [21], light [22], and dominant color [23] have also been used for analysis, but other quantifiable metrics have not been actively developed [24]. Another issue is determining which statistical quantity can be observed from the distribution of a certain metric. A video is just a temporal array of photos; thus, a measurement at the video level usually generates a distribution of measurements at the photo level. This distribution can be examined by a certain statistical quantity, which reflects the characteristics of the video; it is comparable to examining the distribution of word frequency when analyzing textual data. In this regard, statistical quantities such as mean [10], median [25], dispersion [16], and 1/f characteristics [26] have been used to identify the characteristics of the distribution generated from video data.

Little attention, however, has been paid to the characteristics of emotion distribution in video data. In particular, it is difficult to find previous studies that quantitatively measured facial emotions from video data and investigated the characteristics of their distribution. The techniques for facial emotion recognition, most of which are based on *deep learning*, have been developing rapidly. Online artificial intelligence services also provide emotion recognition functions, which make it much easier to use them for research purposes. A few studies have used facial emotion recognition to analyze social media photographs [27] or newspaper photographs [28]. However, it is difficult to find studies that measured facial emotions in TV series or examined the emotion distribution obtained from TV series data.

This study is intended to fill the gaps in the literature. First, it aims to obtain emotion distributions by measuring facial emotions from the keyframes of each shot in TV series and explore what the distributions generally look like. The second aim is to compare the distributions by the characteristics of the TV series. For this purpose, this study examines the statistical complexity of a distribution, which is a metric of how complex a given distribution is [29]. When applied to the emotion distribution of a video, it can exhibit a certain characteristics of the video; in extreme cases, a video with uniform emotion distribution and a video with random emotion distribution would look much different. The emotion distributions of real-world videos would be located between these two extremes. This study investigates how the statistical complexity of emotion distribution differs based on the characteristics of TV series—more precisely, by the temporal phase of an episode, channel type, genre, ratings, and season. Among the huge variety of TV series in many countries, Korean TV series were selected for analysis in the present research because of their recent popularity among global viewers.

The research questions of this study are as follows;

RQ1. What common patterns can be observed in the emotion distributions of Korean TV series?

RQ2. How does the statistical complexity of emotion distribution differ based on the characteristics of each Korean TV series?

II. RELATED WORKS

A. AFFECTIVE CONTENT ANALYSIS OF VIDEO

According to Wang and Ji [5], the first attempt of direct video affective content analysis was made by Hanjalić and Xu [8]. They based their analysis on a three-dimensional approach [30], [31], in which human emotion was modeled as a geographical region on a space with the axes of pleasure-arousal-dominance or valence-arousal-control. The authors reduced this into a two-dimensional model (arousal-valence) for efficiency and used low-level features, including motion, vocal effects, and shot length, to extract affective curves on the dimensional space from a video.

This method was used in many studies to analyze various kinds of video content. Wang and Cheong [6] conducted an affective content analysis of films. They used shot duration, visual excitement, lighting key, and color energy as the visual features for emotional classification by a support vector machine and showed the validity of the features. Canini, Benini, and Leonardi [32] also analyzed the emotion of movies. The authors used 24 features on the audio, visual, and film-grammatical aspects of a film and showed their validity in predicting the emotion of the film. Music videos were analyzed in terms of their emotion by Zhang *et al.* [7], who extracted various features for arousal and valence, then analyzed 552 music videos in terms of their emotion. Yazdani, Kappeler, and Ebrahimi [33] also analyzed 40 music videos using 11 visual features; the emotions that the videos could induce in people were determined by comparing the results of the analysis with the assessments of respondents. Jaimes *et al.* [34] analyzed the affective content of meeting videos. By comparing the low-level features extracted from videos and the manual labeling, the authors showed that the low-level features are promising for analyzing meeting videos in terms of emotion.

Little research, however, has been conducted on measuring emotions from the facial expressions of characters in videos. Various low-level features have been used for affective content analysis, but the emotions revealed in human faces were not actively utilized in previous research. To name a few exceptional studies, Hakim, Marsland, and Guesgen [35] used data on human faces, in which experts labelled the revealed emotion on them as training data for video classification. In addition, Ryoo and Chang [36] associated the color features of videos with the expressed emotion on human faces. However, relatively little attention has been paid to the facial emotions on human faces in affective content analysis of video, and the emotional differences according to the attributes of the video have been particularly understudied.

B. CINEMETRICS: QUANTITATIVE ANALYSIS OF FILMS

Cinemetrics, as a quantitative approach to film data, is a common method that uses metrics to quantify a video. The most widely used metric is shot length [19], [20], which is considered to be the ''single definable element'' of quantitative attempts in film analysis [24]. Salt [11] compared various Hollywood movies in terms of their shot lengths. Since then, the characteristics of Hollywood films in terms of their shot lengths have been analyzed in many studies. Cutting *et al.* [19] demonstrated that the shot lengths in Hollywood movies were shortened during the period from 1935 to 2010. Cutting, Brunick, and DeLong [37] showed that a film can be divided into quarters in terms of the shot length dynamics; shots at the quarter boundaries are longer, whereas the shots in the middle of each quarter are shorter. Cutting and Candan [38] found a trend in both English and non-English movies that shot lengths were shortened in the sound movie era (1930–2015). Cutting [39] inspected the difference in mean shot duration among three periods: 1930–1955, 1960–1985, and 1990–2015. Baxter, Khitrova, and Tsivian [40] explored the cutting structure, based on shot lengths, in the movies of three directors. Svanera *et al.* [41] showed that, with other metrics, shot length can be used to predict the director of a given film.

Other metrics have also used in cinemetrics research. Cutting and Armstrong [21] measured shot scale, which indicates the degree of close-up, in Hollywood movies; they suggested that the average shot scale has increased. In another study, the same authors showed that shot scale was negatively associated with shot length in Hollywood movies [42]. Cutting, Brunick, and Candan [43] demonstrated that shot scale, along with other metrics, can be used as a tool for parsing events from Hollywood movies. Motion, the difference in pixels between neighboring shots, and the luminance, the degree of brightness of a video, are other metrics used in cinemetrics research [19], [44]. Canini *et al.* [22] showed that the trajectories of motion and luminance can be a strong characterization of a movie. Cutting and DeLong *et al.* [45] invented the visual activity index (VAI), which couples motion with the amount of camera movement; they demonstrated that the VAI has increased in Hollywood movies from 1935 to 2005. Narrative shift style, which is a way to transition from one frame to the next (e.g., cut, dissolve, fade, and wipe), was investigated in popular movies by Cutting [46]; the study identified four kinds of narrative shift pattern—general, historical, genre-specific, and film-specific. Avgerinos *et al.* [23] extracted certain features, including dominant color, motion, contrast, tempo, and face-to-frame ratio, for YouTube video analysis. Cutting [47] used features such as shot transition, motion, shot type, and shot duration to observe the narrative dynamics of popular movies. Álvarez *et al.* [48] used 24 features in three categories—image, pace, and motion— to predict the genre, production year, and popularity of movies.

As can be seen from the above, cinemetric analysis has been widely applied to movie data. However, that relatively little research has used cinemetric analysis for TV content. Film and TV content are both in video format, and the development of TV styles tends to follow the development of movie styles [14]. Thus, the cinemetric analysis has much potential for application in TV content. However, only a limited number of studies have realized this potential. Salt [14] analyzed the shot scale of TV series drama, and Butler [17] analyzed the shot length of TV sitcoms. Schaefer and Martinez [15] and Redfern [16] analyzed the shot length of TV news. Out study attempts to extent the cinemetric approach to Korean TV series. Specifically, it uses the emotion expressed in the faces of characters in a video (see Methods section). To the best of our knowledge, the emotional facial features have not yet been used in the cinemetric analysis of TV series.

C. STATISTICAL COMPLEXITY AS A CHARACTERISTIC OF EMOTION DISTRIBUTION

The statistical structure of a cinemetric distribution can be an indicator of the characteristics of a video. For example, if we measure the shot length of every shot of a video, the statistical structure of the shot length distribution can be a characteristic quantity of the video. In this regard, many measurements of cinemetric distribution have been examined to determine the characteristics of a video.

The most commonly used statistics in cinemetric studies are average and dispersion. Average shot length has been widely used as the key metric in statistical film studies [49]. Salt [11] showed that the average shot length in silent films was shorter than that in sound films. Cutting *et al.* [19] measured the average shot lengths of 160 films from 1935 to 2010 and found that shot lengths have decreased over time. Cutting, Brunick, and Delong [49], [50] measured the shot lengths in 143 Hollywood films and demonstrated that acts in films are different in terms of their average shot lengths. Avgerinos *et al.* [23] analyzed the average of other metrics, including rhythm, motion, and face-to-frame ratio. On dispersion of a distribution, Redfern [16], [20], [25] used Q_n, which measured the dispersion based on the distance between each pair of data points to characterize shot length distributions. Standard deviation was used in Avgerinos *et al.* [23] and Moorthy, Obrador, and Oliver [51] to measure the dispersion of various cinemetric quantities.

A group of studies delved further and tried to identify the probability function of cinemetric distribution. Cutting, DeLong, and Nothelfer [26] investigated the shot lengths of 150 Hollywood films to determine whether the shot length distribution followed a 1/f pattern, which was observed in the distribution of viewer reaction time. Cutting, DeLong, and Brunick [52] expanded the sample size to 295 movies and affirmed the previous result of a 1/f pattern. Vasconcelos and Lippman [53] conducted a statistical modeling of the shot length distributions of 23 promotional movie trailers. They fitted the shot length distributions to Erlang and Weibull distributions and argued that these are more realistic models than Poisson distribution. Redfern [54] examined the shot length distributions of 134 Hollywood movies and suggested that, in contrast to the previous claims, the distributions of the overwhelming majority of examined movies do not follow log-normal distribution.

As can be seen from the above, the statistical properties of cinemetric distribution have been examined in the literature. In the line of this research, our study focuses on the *statistical complexity* of cinemetric distribution. A system is considered to be complex if it does not display patterns that are considered to be simple [29]. Because there is no univocal definition of complexity, many methods of measurement have been suggested [55]. The simplest measure of complexity is entropy [56], which defines complexity as the amount of information contained in a system; a system with higher disorder, which is equivalent to a larger amount of information, has a higher level of complexity. In other words, because disorder can be regarded as random, a fully random system whose possible states have equal probability has maximum entropy, whereas a perfectly ordered system that has a single possible state has minimum entropy [57]. Measuring complexity using entropy only, however, is not perfect; a system in an equilibrium state, in theory, would have the maximum level of entropy [29], but this is not the case in most real-world systems. In addition, entropy automatically increases as the number of possible states increases [56]. Thus, the complexity of a system can be measured using both entropy and the degree of how far the system is from the equilibrium state [29], [56]. In this way, the two extremes where a system is fully random (e.g., ideal gas) or fully ordered (e.g., crystal) have minimum complexity, whereas a system that is disordered but far from equilibrium has maximum complexity.

In the literature, there has been recent interest in the complexity of media content in various forms. Image complexity has been measured by a variety of metrics, including fractal dimension [58], algorithmic specified complexity [59], compressed file size [60], the degree of flatness in the profile of gray-scale mean variance [61], and the entropy of luminosity [48]. In addition, the complexities of text discourse [62], movie narration [63], and the rhythm of films [64] have been investigated. However, statistical complexity has not been actively used to measure complexity in media content. Lopes and Machado [65] suggested several metrics, including statistical complexity, to characterize musical information; the authors analyzed the music of four popular musicians, but this work is one of the few exceptions.

Our study used statistical complexity to measure the complexity of emotion distributions in Korean TV series. To the best of our knowledge, the complexity of emotion distribution has not been measured in previous research. In particular, the distribution of facial emotion expressions has not been investigated in terms of its complexity. In this study, a video was divided into a series of shots; then, the keyframe of each shot was measured in terms of facial emotions (see Methods section). Only frequency distribution was considered, not the temporal order of emotions. This is equivalent to an investigation of word frequency distribution [66], [67] in a *bag of words* approach [68], in which text is considered to be a collection of constituting words and the order of the words is ignored. In this regard, our approach can be referred to as a *bag of shot emotions*, which we used to investigate Korean TV series in terms of their statistical complexity.

III. METHODS

A. DATA AND PREPROCESSING

1) RESEARCH SAMPLE

The research samples were selected from a pool of all TV series that began in 2017 on terrestrial TV channels and cable/general programming channels in Korea. Any series with too small (less than 12) or too large (more than 30) number of episodes were excluded from the samples. We considered two episodes that were broadcasted consecutively and had lengths of approximately 30 minutes each to be a single episode with a commercial break, which was not allowed on terrestrial TV channels in Korea at the time of writing. The top five and bottom five series in terms of viewer ratings (according to TNmS media data agency) on terrestrial TV channels and cable/general programming channels for each (20 series in total) were included in the research samples. The genres of the series were categorized as soft or hard: soft genres included comedy, romance, youth, and fantasy, whereas hard genres included crime, mystery, and thriller. The season of each series was also determined. The research samples are presented in Table 1. (KBS2, MBC, and SBS are terrestrial TV channels, whereas JTBC, tvN, and OCN are cable/general programming channels.)

2) SHOT BOUNDARY DETECTION AND KEYFRAME SELECTION

Every episode in the samples was divided into shots. Only the keyframes selected from shots, rather than all frames, were analyzed; this method has been widely adopted in the literature for efficiency purposes [6]. Since the aim of this study is to analyze video content, so we used an existing shot boundary detection method rather than proposed a new method. The fundamental principle of shot boundary detection is that neighboring frames in the same shot may differ less in terms of pixels, whereas neighboring frames in different shots may differ more [69]. The RGB (red, green, and blue, each of which lies between 0 and 255) histograms of each frame were generated; the middle of two neighboring frames was determined as the shot boundary if the histogram difference between the two frames was larger than a certain threshold [70]. The histogram difference (D_f) between the k^{th}

TABLE 1. Research sample.

Name	Channel	Genre	Episode	Ratings (%)	Period	Season
Defendant	SBS	hard	18	17.7	1.23-3.21	winter
Manager Kim	KBS2	soft	20	14.0	1.25-3.30	winter
Whisper	SBS	hard	17	13.7	3.27-5.23	spring
Ruler: Master of The Mask	MBC	hard	20	11.8	5.10-7.13	summer
Stealer	MBC	hard	30	10.9	1.30-5.16	winter/spring
Manhole	KBS2	soft	16	2.7	8.9-9.28	summer
I'm Not a Robot	MBC	soft	16	3.63	12.6-1.25	winter
The Best Hit	KBS2	soft	16	3.7	6.2-7.22	summer
20th Century Boy and Girl	MBC	soft	16	3.93	10.9-11.28	fall
The Cyborg Mom	MBC	soft	12	3.93	9.15-12.1	fall
Strong Woman DoBongSoon	JTBC	soft	16	7.9	2.24-4.15	spring
Prison Playbook	tvN	soft	16	7.69	11.22-1.18	winter
Woman of Dignity	JTBC	soft	20	5.99	6.16-8.19	summer
Avengers Social Club	tvN	soft	12	5.23	10.11-11.16	fall
Stranger	tvN	hard	16	4.56	6.10-7.30	summer
Duel	OCN	hard	16	1.3	6.3-7.23	summer
The Liar and His Lover	tvN	soft	16	1.51	3.20-5.9	spring
Introverted Boss	tvN	soft	16	1.65	1.6-3.14	winter
The Package	JTBC	soft	12	1.78	10.13-11.18	fall
Rain or Shine	JTBC	soft	16	1.88	12.11-1.3	winter

frame and $(k+1)^{\text{th}}$ frame was calculated by the chi-square method [71]:

$$D_f(k, k+1) = \sum_{i=1}^{3} \frac{[H(k,i) - H(k+1,i)]^2}{H(k,i)}$$

In the above equation, i is the red, green, and blue. $H(k, i)$ is the histogram of red, green, or blue for the k^{th} frame.

An adaptive thresholding method [72], [73] was used to determine the threshold of the histogram difference. In the continuum of histogram differences, a symmetric sliding window of size 21 (before and after 10 differences of a certain histogram difference) was used. The middle difference was determined to be a shot boundary if the following conditions were simultaneously satisfied:

(1) The middle difference is the maximum in the window.

(2) The middle difference is greater than $\max(\mu_{left} + T_d\sqrt{\sigma}_{left}, \mu_{right} + T_d\sqrt{\sigma}_{right})$.

Here, $\boldsymbol{\mu}$ is the mean of the differences; $\boldsymbol{\sigma}$ is the standard deviation of the differences; and T_d is 5 (constant) [73].

After the shot boundaries were determined, a keyframe was selected from each determined shot. The keyframe is meant to be representative of the content of a shot [74]. Since neighboring frames in the same shot may differ less in terms of color histograms, whereas neighboring frames in different shots may differ more. Thus, frames that constitute a shot would be similar with each other and any frame would be eligible for key frame. For the sake of simplicity, the middle frame of each shot was selected as the keyframe.

Finally, these selected keyframes were manually inspected. If multiple keyframes were extracted from a single shot due to an error of the algorithm, all but one keyframe were discarded.

B. MEASUREMENTS

1) EMOTION MEASUREMENT USING A PRETRAINED ONLINE AI SERVICE

The emotion in each keyframe was measured from the faces in the keyframes using the Face API from Microsoft Azure Cognitive Services. The pretrained online AI services enable researchers focus more on the purpose of the study and have been used in used in many previous studies especially for analyzing visual content [27], [28], [75]–[77]. The pretrained AI detects faces on a given photo and determines the relative strengths of 8 emotion categories: anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise; the sum for these categories is 1 for each detected face [78]. The emotion of each keyframe was calculated as the sum of 7 of these emotion categories (excluding neutral), each of which was averaged across all faces detected in the keyframe.

2) GENERATING EMOTION DISTRIBUTION

The emotion distribution of each episode was obtained by dividing the range of emotion ([0, 1]) into 100 intervals and counting the number of keyframes with emotion falling into each interval. Each distribution was normalized so that the total area of the histogram bars was 1. For RQ1, the emotion distributions of all episodes in the sample were investigated to determine whether any common patterns were observed.

3) STATISTICAL COMPLEXITY OF EMOTION DISTRIBUTION

For the emotion distribution of each episode in all samples, statistical complexity was calculated. We adopted the calculation method in Lopez-Ruiz, Mancini, and Calbet [29]. The

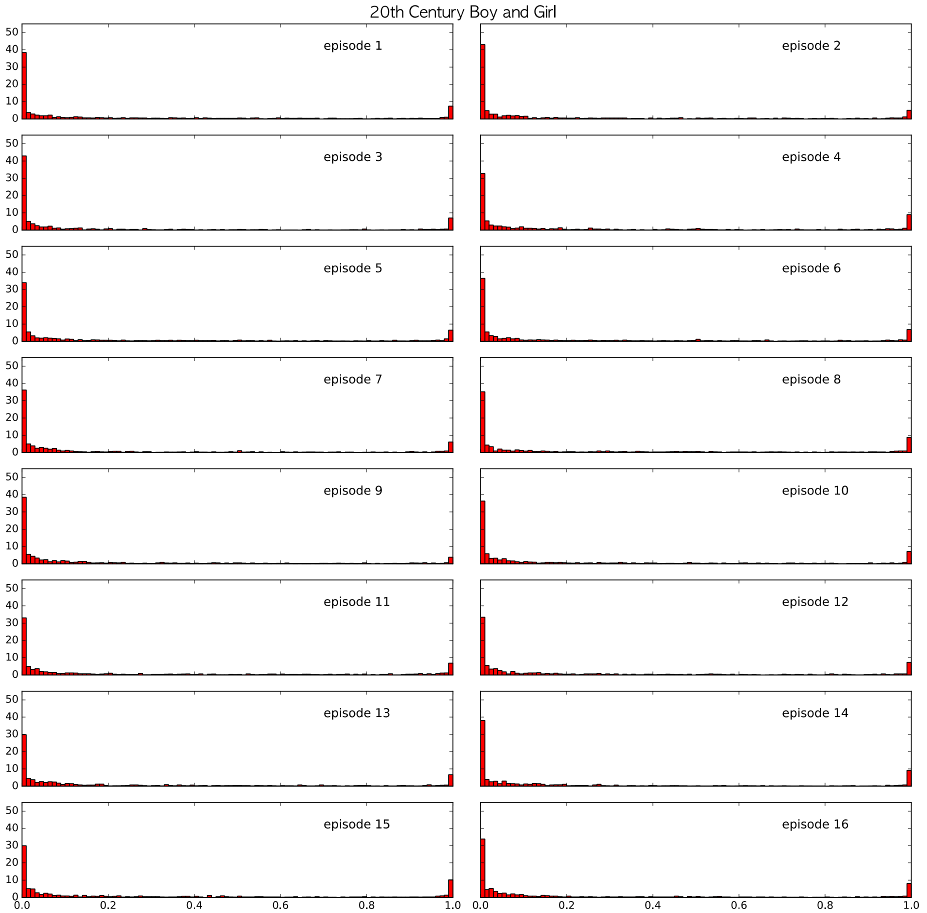

FIGURE 1. Emotion distributions of the Korean TV series, *20th Century Boy and Girl*.

information (H) was calculated by Shannon entropy:

$$H = -K \sum_{i=1}^{N} p_i log p_i$$

Originally in Lopez-Ruiz, Mancini, and Calbet [29], N was the number of possible states of a system, p_i was the probability that the system reaches the state i ($\in N$), and K was constant. Applying it here to this study, N is the number of intervals (100) in total range of emotion ([0, 1]), p_i is the normalized proportion of shots whose emotions in keyframes fall into the interval i, and K is 1. The disequilibrium (D) was calculated as follows:

$$D = \sum_{i=1}^{N} (p_i - 1/N)^2$$

And the complexity (C) is the interplay between the information stored in the system and its disequilibrium:

$$C = HD = -\left(K \sum_{i=1}^{N} p_i log p_i\right)\left(\sum_{i=1}^{N} (p_i - 1/N)^2\right)$$

For RQ2, the means of the complexity were compared in terms of the temporal phase of the episodes, channel type, genre, ratings, and season. For a certain TV series, the total number of episodes were divided into four intervals: episodes in the first quarter were categorized as the early phase, the second and third quarters were the mid phase, and the fourth-quarter episodes were the late phase. The channels, genres, ratings, and seasons of samples are presented in Table 1.

IV. RESULTS

A. PATTERNS OF EMOTION DISTRIBUTIONS

The emotion distributions of all episodes of *20th century Boy and Girl* are presented as an example in Figure 1. Among 100 intervals of emotion range, the lowest emotion interval ([0, 0.01]) had the largest number of shots. The frequency of the shots rapidly decreased from the next interval, but it slightly increased in the highest emotion interval ([0.99, 1]). All episodes of all series showed the same pattern. (The other figures were omitted due to space constraints but are available upon request.) These asymmetric U-shaped distributions biased toward zero suggest that shots that convey a minimum level of emotion make up the largest share in Korean TV series. They also suggest that the emotions are expressed in a bipolar manner: almost all emotions are expressed at either

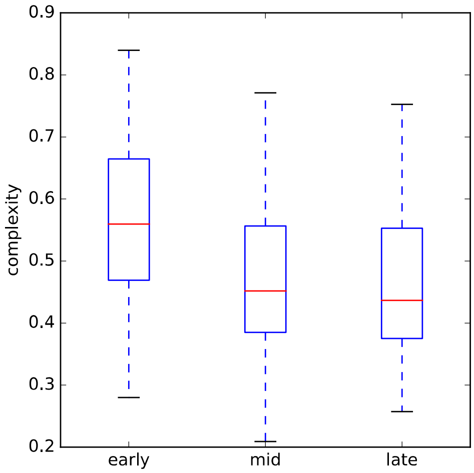

FIGURE 2. Differences in mean complexity by the temporal phase of episodes.

the minimum or maximum level. From these results, it can be inferred that the unique characteristics of each TV series are expressed by the temporal arrangement of shots, considering that emotion distributions showed similar patterns across all series.

B. COMPARISON OF THE MEAN COMPLEXITY OF EMOTION DISTRIBUTIONS

First, it was investigated whether the means of complexity differed according to the temporal phase of the episodes. As Figure 2 shows, the mean complexities of early episodes (0.559) were higher (Kruskal-Wallis $H = 35.488$, $p < .001$) than that of mid (0.470) and late episodes (0.461). This result suggests that emotion distribution is more complex in early episodes and becomes less complex as the story develops in mid and late episodes.

Next, the differences in the mean complexity by channel type were examined. As shown in Figure 3 (a), no significant difference were found (Mann-Whitney $U = 14017.0$, $p = .455$) in the mean complexity of terrestrial TV (0.494) and cable/general programming channels (0.494). This result suggests that Korean TV series, which are broadcast on different types of channels, are not distinct in terms of the complexity of emotion distribution. Differences in the means of complexity by TV series genres were also investigated. Figure 3 (b) indicates that the complexity of emotion distribution in hard-genre series (0.581) was significantly higher (Mann-Whitney $U = 5130.0$, $p < .001$) than that in the soft-genre series (0.448). In addition, Figure 3 (c) shows the means of complexity according to ratings: the Korean TV series with higher ratings (0.505) had greater complexity (Mann-Whitney $U = 12296.0$, $p = .024$) in emotion distribution compared with lower-rated series (0.481).

Finally, it was investigated whether the means of complexity were different by season. Figure 4 indicates that the complexity of emotion distribution was highest in the summer (0.517) and lowest in the fall (0.426), those in the spring (0.508) and the winter (0.494) were in the middle, and they were significantly different (Kruskal-Wallis $H = 19.634$, $p < .001$).

V. DISCUSSION

A shot is the basic meaningful unit of video data, just as a word is the basic meaningful unit of textual data. Thus, the emotion distribution obtained from shots can exhibit a certain aspect of characteristics of a given video, just as the word frequency distribution shows the characteristics of a given text. In this regard, our study focused on the emotion distributions of Korean TV series. In all, 337 episodes from 20 TV series were divided into shots, and keyframes were selected from the shots. Facial emotions on the keyframes were measured using an online artificial intelligence service, and emotion distributions were obtained from each episode and observed for common patterns. In addition, statistical complexity was calculated and compared according to the temporal phase of episode, channel type, genre, ratings, and season. Our key findings are discussed in this section.

First, we found that the emotion distributions of all episodes in all series had the same form. The asymmetric U-shaped distributions biased toward zero indicate that the two extremes—the weakest and the strongest—are the most frequent emotions in Korean TV series. This result suggests that the emotional *ingredients* of Korean TV series are not dramatically different from each other; Korean TV series usually use shots with either neutral or explosive emotion. Thus, it can be inferred that the temporal arrangement of shots is what characterizes each TV series. This temporal order of emotion, or emotional dynamics, is a topic for further investigation. Future studies should investigate how emotions are displayed temporally and how the emotional dynamics are different by director, genre, or ratings.

Next, we found that the statistical complexity of emotion distribution was higher in early episodes than in later episodes. This may seem to conflict with the traditional narrative theory. Narrations in literature or drama are assumed to be divided into a certain number of parts (the number of parts may be three, four, or five, depending on the theorists); the first part is commonly regarded as the beginning, introduction, or setup, in which the narrative conflict is still immature [47]. Thus, it may be expected that the complexity of emotion distribution is lower in early episodes, when the characters and backgrounds are introduced and the story has not unfolded yet. The results of this study, however, suggests that this is not the case for Korean TV series. In Korean TV series, the story seems to develop in a different route from traditional narrative theory; rather than preparing in early episodes for the narrative conflict to mature in later episodes, Korean TV series showed more complexity in emotion distribution in early episodes. This may be attributed to fierce competition in the Korean TV series market: unless a series catches the interest of viewers in early episodes, it is challenging to keep their eyes on in later episodes. This is congruous with another result of this study, which showed

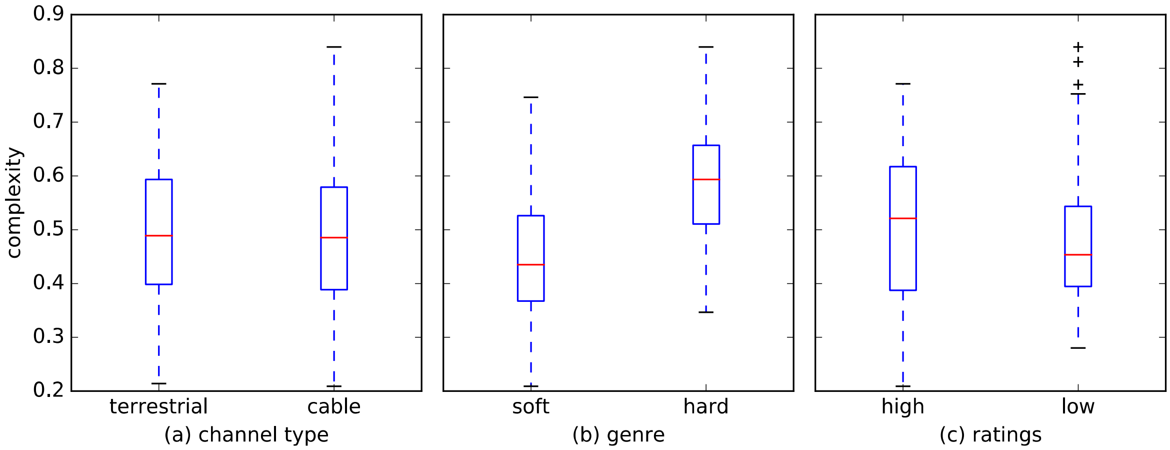

FIGURE 3. Differences in mean complexity by channel type, genre, and ratings.

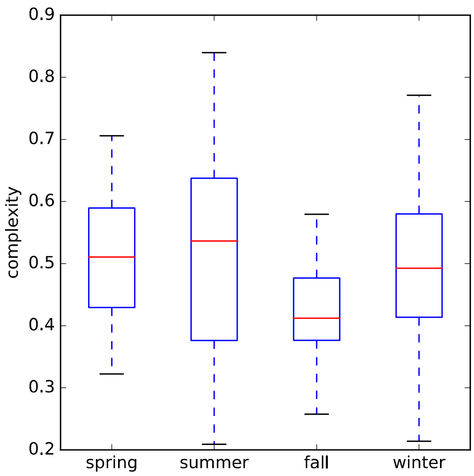

FIGURE 4. Differences in mean complexity by season.

that the complexity of emotion distribution was higher in episodes with higher ratings. Viewers tend to prefer TV series with complex emotion distribution. This preference seems to be implicitly reflected in the unique storytelling approach of Korean TV series, which differs from the traditional narrative theory.

Finally, the statistical complexity of emotion distribution was different by certain characteristics of the TV series. The complexity was higher in the series with hard genres such as crime, mystery, and thriller than in the series with soft genres such as comedy, romance, youth, and fantasy. While the content itself in the soft-genre series may be more about emotion than that in the hard-genre series, this result suggests that the statistical complexity of emotion distribution can exhibit another aspect of TV series. The complexity was defined in this study as the interplay between entropy and being far from disequilibrium. Thus, the complexity of emotion distribution in soft-genre series can be lower if the shots express stronger emotions (equivalent to higher entropy) but the degrees of their strength are more similar with each other (equivalent to being closer to equilibrium); this seems to be the case for Korean TV series. Also, the complexity of emotion distribution was the highest in the series of the summer and the lowest in the series of the fall. This result suggest that the attributes of the seasons were implicitly reflected in the content; characters in the summer series might be more active and changeable in terms of their emotion, whereas characters in the fall series might be the opposite. The closer investigation of the relationships between the expressed facial emotions and the characteristics of the TV series would be another topic for further study, and it also can be examined whether these relationships hold across the TV series in various cultures.

The findings of the present study can have theoretical and practical implications. They showed that the cinemetric approach can be applied to TV series, and they also showed that facial emotion expression can be an another metric for cinemetric research. In addition, the findings showed that TV series has its own way of narrative; this would help us to understand how the media format influence the way of storytelling. In practice, the relationships between the property of emotion distribution and the characteristics of TV series would provide series makers a hint for success in the market.

The framework of this study can be applied to other types of video data, including popular movies and TV series in cultures other than Korean. The result of this study can be a starting point for discussion on the implications of those future studies. However, given the exploratory nature of this study, the lack of theoretical perspective may prevents the further discussion of the results, which is a major limitation of the present research. The quantitative investigation of TV series and other types of video data is still in deficit and the theoretical perspective that can lead to a richer discussion has not actively developed. Based on the results of the present research, more empirical and theoretical studies are expected to be conducted. Also, it can be investigated in future studies how the audience perception of emotion would be different by the emotions on the screen.

REFERENCES

[1] S. W. Head, "Content analysis of television drama programs," *Quart. Film Radio Telev.*, vol. 9, no. 2, pp. 175–194, Dec. 1954.

[2] J. C. McNeil, "Feminism, femininity, and the television series: A content analysis," *J. Broadcast.*, vol. 19, no. 3, pp. 259–271, Jun. 1975.
[3] Y. Ye and K. E. Ward, "The depiction of illness and related matters in two top-ranked primetime network medical dramas in the united states: A content analysis," *J. Health Commun.*, vol. 15, no. 5, pp. 555–570, Jul. 2010.
[4] R. Helles and S. S. Lai, "Travelling or not? A content analysis mapping danish television drama from 2005 to 2014," *Crit. Stud. Telev., Int. J. Telev. Stud.*, vol. 12, no. 4, pp. 395–410, Dec. 2017.
[5] S. Wang and Q. Ji, "Video affective content analysis: A survey of State-of-the-Art methods," *IEEE Trans. Affect. Comput.*, vol. 6, no. 4, pp. 410–430, Oct. 2015.
[6] H. L. Wang and L.-F. Cheong, "Affective understanding in film," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 6, pp. 689–704, May 2006.
[7] S. Zhang, Q. Huang, S. Jiang, W. Gao, and Q. Tian, "Affective visualization and retrieval for music video," *IEEE Trans. Multimedia*, vol. 12, no. 6, pp. 510–522, Oct. 2010.
[8] A. Hanjalic and L.-Q. Xu, "Affective video content representation and modeling," *IEEE Trans. Multimedia*, vol. 7, no. 1, pp. 143–154, Feb. 2005.
[9] N. Redfern, "Film studies and statistical literacy," *Media Edu. Res. J.*, vol. 4, no. 1, pp. 58–73, 2013.
[10] K. L. Brunick, J. E. Cutting, and J. E. DeLong, "Low-level features of film: What they are and why we would be lost without them," in *Psychocinematics: Exploring Cognition at the Movies*, A. P. Shimamura, Ed. Oxford, U.K.: Oxford Univ. Press, 2013, pp. 133–148.
[11] B. Salt, "Statistical style analysis of motion pictures," *Film Quart.*, vol. 28, no. 1, pp. 13–22, Oct. 1974.
[12] W. Buckland, "What does the statistical style analysis of film involve? A review of moving into Pictures. More on film history, style, and analysis," *Literary Linguistic Comput.*, vol. 23, no. 2, pp. 219–230, 2008.
[13] B. Salt, "Review of Jeremy G. Butler, television style," *New Rev. Film Telev. Stud.*, vol. 8, no. 4, pp. 454–458, Dec. 2010.
[14] B. Salt, "Practical film theory and its application to TV series dramas," *J. Media Pract.*, vol. 2, no. 2, pp. 98–113, Jan. 2001.
[15] R. Schaefer and T. Martinez, "Trends in network news editing strategies from 1969 through 2005," *J. Broadcast. Electron. Media*, vol. 53, no. 3, pp. 347–364, Jul. 2009.
[16] N. Redfern, "The structure of ITV news bulletins," *Int. J. Commun.*, vol. 8, pp. 1557–1578, 2014.
[17] J. Butler, "Statistical analysis of television style: What can numbers tell us about TV editing?" *Cinema J.*, vol. 54, no. 1, pp. 25–44, 2014.
[18] T. Arnold, L. Tilton, and A. Berke, "Visual style in two network era sitcoms," *J. Cultural Anal.*, Jul. 2019. [Online]. Available: https://culturalanalytics.org/article/11045-visual-style-in-two-network-era-sitcoms
[19] J. E. Cutting, K. L. Brunick, J. E. DeLong, C. Iricinschi, and A. Candan, "Quicker, faster, darker: Changes in hollywood film over 75 years," *I-Perception*, vol. 2, no. 6, pp. 569–576, Aug. 2011.
[20] N. Redfern, "Shot length distributions in the short films of Laurel and Hardy, 1927 to 1933," *Cine forum*, vol. 14, pp. 37–71, Aug. 2012.
[21] J. E. Cutting and K. L. Armstrong, "Cryptic emotions and the emergence of a metatheory of mind in popular filmmaking," *Cognit. Sci.*, vol. 42, no. 4, pp. 1317–1344, May 2018.
[22] L. Canini, S. Benini, P. Migliorati, and R. Leonardi, "Emotional identity of movies," in *Proc. 16th IEEE Int. Conf. Image Process. (ICIP)*, Nov. 2009, pp. 1821–1824.
[23] C. Avgerinos, N. Nikolaidis, V. Mygdalis, and I. Pitas, "Feature extraction and statistical analysis of videos for cinemetric applications," in *Proc. Digit. Media Ind. Academic Forum (DMIAF)*, Jul. 2016, pp. 172–175.
[24] M. Burghardt, M. Kao, and C. Wolff, "Beyond shot lengths: Using language data and color information as additional parameters for quantitative movie analysis," in *Proc. Digit. Humanities*, Kraków, Poland, 2016, pp. 753–755.
[25] N. Redfern, "The impact of sound technology on the distribution of shot lengths in hollywood cinema, 1920 to 1933," *CINEJ Cinema J.*, vol. 2, no. 1, pp. 77–94, 2012.
[26] J. E. Cutting, J. E. DeLong, and C. E. Nothelfer, "Attention and the evolution of hollywood film," *Psychol. Sci.*, vol. 21, no. 3, pp. 432–439, Mar. 2010.
[27] Y. Kim and J. H. Kim, "Using computer vision techniques on instagram to link users' personalities and genders to the features of their photos: An exploratory study," *Inf. Process. Manage.*, vol. 54, no. 6, pp. 1101–1114, Nov. 2018.
[28] Y. Peng, "Same candidates, different faces: Uncovering media bias in visual portrayals of presidential candidates with computer vision," *J. Commun.*, vol. 68, no. 5, pp. 920–941, Oct. 2018.
[29] R. Lopez-Ruiz, H. L. Mancini, and X. Calbet, "A statistical measure of complexity," *Phys. Lett. A*, vol. 209, no. 5, pp. 321–326, Dec. 1995.
[30] H. Schlosberg, "Three dimensions of emotion," *Psychol. Rev.*, vol. 61, no. 2, pp. 81–88, 1954.
[31] J. A. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions," *J. Res. Personality*, vol. 11, no. 3, pp. 273–294, Sep. 1977.
[32] L. Canini, S. Benini, and R. Leonardi, "Affective recommendation of movies based on selected connotative features," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 4, pp. 636–647, Apr. 2013.
[33] A. Yazdani, K. Kappeler, and T. Ebrahimi, "Affective content analysis of music video clips," *Proc. 1st Int. ACM workshop Music Inf. Retr. User-Centered Multimodal Strategies (MIRUM)*, New York, NY, USA, 2011, p. 7.
[34] A. Jaimes, T. Nagamine, J. Liu, K. Omura, and N. Sebe, "Affective meeting video analysis," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jun. 2005, pp. 1412–1415.
[35] A. Hakim, S. Marsland, and H. W. Guesgen, "Computational analysis of emotion dynamics," in *Proc. Humaine Assoc. Conf. Affect. Comput. Intell. Interact.*, Sep. 2013, pp. 185–190.
[36] S. Ryoo, H. Univerasity, J.-K. Chang, and H. Univerasity, "Emotion affective color transfer using feature based facial expression recognition," in *Proc. Adv. Sci. Technol. Lett.*, Dec. 2013, pp. 131–135.
[37] J. E. Cutting, K. L. Brunick, and J. E. Delong, "The changing poetics of the dissolve in hollywood film," *Empirical Stud. Arts*, vol. 29, no. 2, pp. 149–169, Jul. 2011.
[38] J. E. Cutting and A. Candan, "Shot durations, shot classes, and the increased pace of popular movies," *Projections*, vol. 9, no. 2, pp. 40–62, Jan. 2015.
[39] J. E. Cutting, "The evolution of pace in popular movies," *Cognit. Research: Princ. Implications*, vol. 1, no. 1, pp. 1–21, Dec. 2016.
[40] M. Baxter, D. Khitrova, and Y. Tsivian, "Exploring cutting structure in film, with applications to the films of D. W. Griffith, Mack Sennett, and Charlie Chaplin," *Digit. Scholarship Humanities*, vol. 32, no. 1, pp. 1–16, 2015.
[41] M. Svanera, M. Savardi, A. Signoroni, A. B. Kovacs, and S. Benini, "Who is the Film's director? Authorship recognition based on shot features," *IEEE Multimedia Mag.*, vol. 26, no. 4, pp. 43–54, Oct. 2019.
[42] J. E. Cutting and K. L. Armstrong, "Facial expression, size, and clutter: Inferences from movie structure to emotion judgments and back," *Attention, Perception, Psychophys.*, vol. 78, no. 3, pp. 891–901, Apr. 2016.
[43] J. E. Cutting, K. L. Brunick, and A. Candan, "Perceiving event dynamics and parsing hollywood films.," *J. Experim. Psychol., Hum. Perception Perform.*, vol. 38, no. 6, pp. 1476–1490, 2012.
[44] J. E. Cutting, "How light and motion bathe the silver screen.," *Psychol. Aesthetics, Creativity, Arts*, vol. 8, no. 3, pp. 340–353, 2014.
[45] J. E. Cutting, J. E. DeLong, and K. L. Brunick, "Visual activity in Hollywood film: 1935 to 2005 and beyond," *Psychol. Aesthetics, Creativity, Arts*, vol. 5, no. 2, pp. 115–125, May 2011.
[46] J. E. Cutting, "Event segmentation and seven types of narrative discontinuity in popular movies," *Acta Psychologica*, vol. 149, pp. 69–77, Jun. 2014.
[47] J. E. Cutting, "Narrative theory and the dynamics of popular movies," *Psychonomic Bull. Rev.*, vol. 23, no. 6, pp. 1713–1743, Dec. 2016.
[48] F. Álvarez, F. Sánchez, G. Hernández-Peñaloza, D. Jiménez, J. M. Menéndez, and G. Cisneros, "On the influence of low-level visual features in film classification," *PLoS ONE*, vol. 14, no. 2, 2019, Art. no. e0211406.
[49] J. E. Cutting, K. L. Brunick, and J. E. Delong, "How act structure sculpts shot lengths and shot transitions in hollywood film," *Projections*, vol. 5, no. 1, pp. 1–16, Jan. 2011.
[50] J. E. Cutting, K. L. Brunick, and J. DeLong, "On shot lengths and film acts: A revised view," *Projections*, vol. 6, no. 1, pp. 142–145, Jan. 2012.
[51] A. K. Moorthy, P. Obrador, and N. Oliver, "Towards computational models of the visual aesthetic appeal of consumer videos," in *Proc. Computer Vision*, vol. 6315, no. 1, K. Daniilidis, P. Maragos, and N. Paragios, Eds. Berlin, Germany: Springer, 2010, pp. 1–14.
[52] J. E. Cutting, J. E. DeLong, and K. L. Brunick, "Temporal fractals in movies and mind," *Cognit. Res., Princ. Implications*, vol. 3, no. 1, p. 8, Dec. 2018.
[53] N. Vasconcelos and A. Lippman, "Statistical models of video structure for content analysis and characterization," *IEEE Trans. Image Process.*, vol. 9, no. 1, pp. 3–19, Sep. 2000.

[54] N. Redfern, "The log-normal distribution is not an appropriate parametric model for shot length distributions of hollywood films," *Digit. Scholarship Humanities*, vol. 30, no. 1, pp. 137–151, Apr. 2015.
[55] M. Efatmaneshnik and M. J. Ryan, "A general framework for measuring system complexity," *Complexity*, vol. 21, no. S1, pp. 533–546, Sep. 2016.
[56] J. S. Shiner, M. Davison, and P. T. Landsberg, "Simple measure for complexity," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 59, no. 2, pp. 1459–1464, Feb. 1999.
[57] C. Anteneodo and A. R. Plastino, "Some features of the López-Ruiz-Mancini-Calbet (LMC) statistical measure of complexity," *Phys. Lett. A*, vol. 223, no. 5, pp. 348–354, Dec. 1996.
[58] P. Machado, J. Romero, M. Nadal, A. Santos, J. Correia, and A. Carballal, "Computerized measures of visual complexity," *Acta Psychologica*, vol. 160, pp. 43–57, Sep. 2015.
[59] W. Ewert, W. A. Dembski, and R. J. Marks, "Measuring meaningful information in images: Algorithmic specified complexity," *IET Comput. Vis.*, vol. 9, no. 6, pp. 884–894, Dec. 2015.
[60] D. C. Donderi, "An information theory analysis of visual complexity and dissimilarity," *Perception*, vol. 35, no. 6, pp. 823–835, Jun. 2006.
[61] D. H. Zanette, "Quantifying the complexity of black-and-white images," *PLoS ONE*, vol. 13, no. 11, 2018, Art. no. e0207879.
[62] K. Sun and W. Xiong, "A computational model for measuring discourse complexity," *Discourse Stud.*, vol. 21, no. 6, pp. 690–712, Dec. 2019.
[63] J. E. Cutting, "Simplicity, complexity, and narration in popular movies," in *Narrative Complexity: Cognition, Embodiment, Evolution*, M. Grishakova and M. Poulaki, Eds. Lincoln, NE, USA: Univ. of Nebraska Press, 2019, pp. 200–222.
[64] J. E. Cutting and K. Pearlman, "Shaping edits, creating fractals: A cinematic case study," *Projections*, vol. 13, no. 1, pp. 1–22, Mar. 2019.
[65] A. M. Lopes and J. A. Tenreiro Machado, "On the complexity analysis and visualization of musical information," *Entropy*, vol. 21, no. 7, p. 669, 2019.
[66] S. T. Piantadosi, "Zipf's word frequency law in natural language: A critical review and future directions," *Psychonomic Bull. Rev.*, vol. 21, no. 5, pp. 1112–1130, Oct. 2014.
[67] C. Bentz, D. Alikaniotis, T. Samardži, and P. Buttery, "Variation in word frequency distributions: Definitions, measures and implications for a corpus-based language typology," *J. Quant. Linguistics*, vol. 24, nos. 2–3, pp. 128–162, Jul. 2017.
[68] Y. Zhang, R. Jin, and Z.-H. Zhou, "Understanding bag-of-words model: A statistical framework," *Int. J. Mach. Learn. Cybern.*, vol. 1, nos. 1–4, pp. 43–52, Dec. 2010.
[69] M. Jiang, J. Huang, X. Wang, J. Tang, and C. Wu, "Shot boundary detection method for news video," *J. Comput.*, vol. 8, no. 12, pp. 1–5, 2013.
[70] J. E. Solem, *Programming Computer Vision With Python*. Sebastopol, CA, USA: O'Reilly, 2012.
[71] P. V. Kathiriya, D. S. Pipalia, G. B. Vasani, A. J. Thesiya, and D. J. Varanva, "X2 (Chi-Square) based shot boundary detection and key frame extraction for video," *Res. Inventry: Int. J. Eng. Sci.*, vol. 2, no. 2, pp. 17–21, 2013.
[72] C. Cotsaces, N. Nikolaidis, and I. Pitas, "Video shot boundary detection and condensed representation: A review," *IEEE Signal Process. Mag.*, vol. 23, no. 2, pp. 28–37, 2006.
[73] Y. Yusoff, W. Christmas, and J. Kittler, "Video shot cut detection using adaptive thresholding," in *Proc. Brit. Mach. Vis. Conf.*, 2000, pp. 1–10.
[74] F. Camastra and A. Vinciarelli, *Machine Learning for Audio, Image and Video Analysis: Theory and Applications*, 2nd ed. London, U.K.: Springer, 2015.
[75] K. Dutta, V. K. Singh, P. Chakraborty, S. K. Sidhardhan, B. S. Krishna, and C. Dash, "Analyzing big-five personality traits of indian celebrities using online social media," *Psychol. Stud.*, vol. 62, no. 2, pp. 113–124, Jun. 2017.
[76] M. Kop, P. Read, and B. R. Walker, "Pseudocommando mass murderers: A big five personality profile using psycholinguistics," *Current Psychol.*, vol. 132, no. 2, pp. 90–99, Mar. 2019.
[77] S. Muralidhara and M. J. Paul, "#Healthy selfies: Exploration of health topics on Instagram," *JMIR Public Health Surveill*, vol. 4, no. 2, p. e10150-12, 2018.
[78] L. Larsen, *Learning Microsoft Cognitive Services*, 2nd ed. Birmingham, U.K.: Packt, 2017.

YUNHWAN KIM was born in Seoul, South Korea, in 1978. He received the B.A. degree in political science (communications major), the M.A. degree in communication, and the Ph.D. degree in communication from the Hankuk University of Foreign Studies, Seoul, South Korea, in 2005, 2007, and 2015, respectively.

Since 2009, he has been a Lecturer with the Division of Media Communication, Hankuk University of Foreign Studies in Seoul, South Korea. He is the author of more than 15 articles. His research interests include computational social science, visual data analytics, and social simulation.

SUNMI LEE was born in Seoul, South Korea, in 1972. She received the B.S. and M.S. degrees in mathematics from Kyung Hee University, South Korea, in 1997, and the Ph.D. degree in applied mathematics from the University of California at Los Angeles, USA, in 2005. From 2009 to 2012, she was a Research Assistant Professor with Arizona State University, USA. She is currently an Associate Professor with the Department of Applied Mathematics, Kyung Hee University, South Korea. Her research areas are mathematical modeling and numerical simulations (numerical ordinary/partial differential equations, optimal control theory) for applications in physical and life sciences.