

Affective classification model based on emotional user experience and visual markers in YouTube video

Roshaliza Mohd Rosli¹, Syaripah Ruzaini Syed Aris², Azhar Abd Aziz², Toshio Tsuchiya³ and Anitawati Mohd Lokman^{2*}

Malaysia Digital Economy Corporation (MDEC), Malaysia¹

Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Malaysia²

Faculty of Economics, Shimonoseki City University, Japan³

Received: 24-June-2021; Revised: 20-August-2021; Accepted: 23-August-2021

©2021 Roshaliza Mohd Rosli et al. This is an open access article distributed under the Creative Commons Attribution (CC BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

A video is composed of rendered elements such as text, audio, and visual elements. It may convey messages that emotionally engage viewers via embedded elements that demand visual attention, referred to as Visual Markers (VM). However, little attention has been paid to VM, particularly in terms of determining which VM influences viewers' emotional experience. Lack of understanding of VM and its impact on viewers' emotional experience may result in negative impact and hamper efficient video classification and filtering. This is crucial when, for instance, a YouTube video is used for malicious agenda. To fill this gap, this research was conducted to identify VM in Extremist YouTube Videos (EYV). It is helpful in determining significant viewers' emotional responses upon watching EYV, and to develop an affective classification model based on emotional User Experience (UX) and VM in YouTube videos. The research conducted in Kansei evaluation using 20 YouTube video specimens with 80 respondents. Multivariate analysis was performed to determine the structure of emotions, the relationship between a VM and emotional responses, and classify the emotional responses and influential VM. The result has enabled this research to develop an affective classification model comprising three emotional dimensions; offensive, intrigue and awkward. The model contributes a new understanding of the body of knowledge of emotional evocative video elements and provides insights to authorities, policy makers, and other stakeholders to manage the classification of emotional evocative video. It could be used as a basis for formulating an algorithm to filter video content. Although the model was based on work under certain limitations, they lend some novelty by linking affect to VM in video classification. Future work could explore enhancing its applicability using wider scope and population of subjects and instruments. Additionally, video producers could extend the model in producing videos capable of invoking a targeted emotion to the viewers.

Keywords

Emotional UX, Human computer interaction, Kansei, Visual markers, YouTube video.

1. Introduction

This research explores viewers' emotional responses upon watching videos in the context of visual features. Visuals have the ability to capture viewers' attention because they elicit emotional responses. When people interact with computers, visual elements are found to attract their attention and are capable of motivating their hearts and minds [1, 2]. However, prior research has indicated that the majority of previous work on affective-based video has focused on text features or low-level features of the video's affective content using the researcher's definition of text and affect, which is viewed as researcher bias [2–5].

On the other hand, video classification has been viewed as a critical means of ensuring the efficient retrieval of videos [6–9]. Previous research has indicated that affective computing-based video retrieval is the future direction and presents eminently compelling research issues [7, 10, 11].

Additionally, due to the computer revolution, emerging issues such as an extremist group infiltrating social media to spread propaganda and execute various illicit activities have been widespread more than they were. Extremist groups are becoming tech-savvy and have their own media outlets to achieve their goals [12, 13]. They want to provoke the interest of potential new members, supporters and sympathizers, particularly young adults [14–17]. One way is through video

* Author for correspondence

production of impressive messages embedded with specific representations called Visual Markers (VM) [18]. The VM is capable to attract attention and interest of the viewers. According to [19], who explored the role of emotionally evocative visuals/images and its capacity to attract viewers' attention because they stimulate emotional responses. The extremist groups deliberately embed depictions of violence, mutilated bodies, or war casualties in their effort to engage the hearts and minds of their target viewers [16, 20].

Moreover, the fact that the path to extremism is just one click away is very worrying. Detecting such extremism promoting videos is important, though can be technically challenging [21]. To do so, we need to have knowledge of how viewers are influenced by VM, thus videos can be filtered based on the VM. The absence of this knowledge might lead to more serious problems relating to exposure to such videos, especially on YouTube. More so when it relies on users' feedback for the video takedown.

Hence, this research is motivated by this knowledge gap. As such, the research sets its objectives to identify the VM in Extremist YouTube Videos (EYV), to determine significant viewers' emotional responses upon watching EYV, and to develop an affective classification model based on emotional User Experience (UX) and VM in YouTube videos. Considering the reported negative influence of EYV, this study examined young viewers' emotional responses to EYV videos as a case study, as this type of video is regarded as one of the most emotionally evocative media for viewers. The investigation employs Kansei Engineering (KE) methodology to discover viewers' emotions when interacting with video stimuli. Kansei is a Japanese term that has close meaning of emotion, and KE has in it a well-established methodology for discovering links between the user's emotion and the stimuli. KE could parametrically links viewers' emotional response to the VM of the EYV, and thus enable the research to develop an affective classification model based on emotional UX and VM in the EYV. The model could be very useful for digital natives at large, policy makers and other stakeholders in video classification, filtering, and censoring, as well as a reference for combating extremism.

2.Literature review

Moor [22], a renowned theoretician in the field of computer ethics, proposed that the computer and Internet revolutions can be divided into three stages.

The first two stages, introduction and permeation, have already taken place. The third and final stage, which is currently ongoing, is dubbed the 'power stage.' At this stage, the most serious social, political, legal, and ethical issues relating to information technology will begin to surface on a large scale. Consistent with Moor's assertion, the use of the Internet by extremists to disseminate propaganda and engage in other illegal online activities has since surfaced [13, 23–28].

Consider the following two randomly chosen news stories from the media, each of which demonstrates a distinct type of alleged extremist utilizing the Internet for a variety of activities. Awan wrote in 2014 about far-right extremists in the United Kingdom who had infiltrated the Internet and used it to widely sponsor their brand of anti-Muslim hatred [29]. On a Monday morning in January 2015, when users logged into Malaysia Airlines' website, they were presented with a picture of a Malaysia Airlines Airbus with the messages '404-Plane not found' and 'Hacked by cyber caliphate' [30]. One thing that the two newsworthy incidents have in common is that they would not have been possible without computer and internet technology.

Social media platforms such as YouTube, Facebook, Twitter, and blogs are a subset of internet technology that has been targeted by various extremist groups [18–20, 24, 31–33]. It has developed into a grassroots communication platform used by extremists for a variety of purposes, including propaganda dissemination, recruitment of new members, and fundraising. The use of propaganda is believed to have the greatest effect on audiences [13, 31–34]. Extremist groups appear to leverage internet technology to recruit new members. A typical news report is one about a teenage girl being arrested on her way to join an extremist group and marry one of its members. It is not merely a media frenzy, but a global phenomenon.

Extremist videos broadcast on YouTube, for example, are an ideal propaganda tool due to their information density. The popularity is unsurprising, given that the publishing extremist content on social media is relatively inexpensive, more secure due to its anonymity, and convenient when compared to more traditional methods [32, 33, 35]. Additionally, it enables multimedia-rich content and has a high potential for reaching massive audiences globally with little to no regulation. Social media are typically characterized by user-generated content and social networking. It contributes to the difficulty of policing

activities that occur largely unregulated on the platform. Consider YouTube; despite community guidelines and administrative efforts, the platform maintains a massive archive of malicious and offensive videos [1, 21, 24, 31, 32]. Many of these videos remain online in the absence of a YouTube reviewer's flag request. At the time of writing, a simple keyword search for 'IS burn Jordanian pilot' returned approximately 28,900 results. While some are excerpts from news broadcasts, talk shows, or analyst critiques that do not directly show the original video footage, the remainder is complete clips.

The community guideline website makes no explicit mention of how YouTube videos are reviewed, or the lead time involved. The message is rather generic in nature, stating that flagged videos will be thoroughly reviewed by YouTube staff [36]. There are additional avenues of escape in terms of EYV's accessibility, misleading descriptions, tags, and titles. Additionally, risky behavior, such as lying about one's age, will circumvent the age restriction setting, and deactivating the safe mode setting will allow all inappropriate content to be accessed. Authorities and monitoring agencies are unlikely to spend time and resources monitoring every online activity of members of the public.

Moor [22], in asserting that the world is now in the most critical stage of the computer revolution, termed the 'power stage,' implies that it is no longer difficult for many people to master the mechanics of the technology. With many people armed with technology, he posed a legitimate question about who will ultimately control it. Similarly, to YouTube, which almost everyone with Internet access has probably used (except for people in China or Iran, where it is completely banned). When it comes to videos with violent content, each of us as users has a role to play, as YouTube relies on user feedback before removing those videos. In Malaysia, the Malaysian Communications and Multimedia Commission (MCMC), a regulatory and oversight body, stated that curbing the use of social media for extremist purposes is complicated by the inability to block all content [37]. There appears to be a reliance on the service provider and a lack of jurisdiction. Perhaps community policing is still necessary while a more effective solution is developed. MCMC, too, encourages the public to report directly to the service provider any social media accounts that violate the service provider's terms and conditions of use.

Inappropriate videos, such as those posted by extremists on YouTube, have attracted the attention of researchers, policymakers, and monitoring organizations, who have begun to investigate the impact of this content on viewers [1,5,17, 18]. Studies on the content analysis of extremist videos concluded that video footage is capable of conveying extremely powerful messages to the point of mobilizing members to launch attacks [3, 10, 25, 26]. Additionally, the production of impressive messages is frequently accompanied by the VM. The VM contained within the video enabled viewers to identify those videos, even when the title and categorization did not contain the word extremist. However, the effect of the VM on viewers' emotions is unknown, thus making it difficult to detect.

Three major issues could be identified from the review, and they are described in the following sub-sections.

2.1 Inadequate attention paid to the visual aspects of extremist videos, particularly those that may have an effect on viewers' emotions

There is a substantial body of research demonstrating the power of visual images to influence viewers. However, the majority of studies place a greater emphasis on comprehending the words in extremists' messages than on visuals [3, 12, 38]. They conducted content analysis on extremist videos in order to learn about the different types of videos, their modus operandi, and production characteristics. Two dimensions were used to categorize videos: operational versus non-operational and individual versus group-oriented [12]. Thematic analysis of Swedish right-wing extremist videos revealed qualitative aspects such as the use of music and editing techniques [39]. There are recurring VM throughout the video that serve to identify them [40]. Additionally, the videos are produced with the hearts and minds of young viewers in mind. Due to a lack of knowledge regarding the effects of visual features, it was difficult to comprehend the VM and how they might have influenced viewers' emotions. This results in an incorrect assessment of the video's influence, particularly on young viewers.

2.2 From an affective standpoint, the classification of EYV is overlooked

Extremist video classification was largely based on textual metadata, such as video titles, manual tags, editorial categories, and comments [4, 5, 21, 38, 41, 42]. While the videos hint at the emotional potential of the material, the videos are not classified in a way that

takes affect into account [43]. According to Zhao et al. [2], most previous works on video classification (not specifically pertaining to EYV) were primarily concerned with video content and paid little attention to viewer affect. In addition, a lack of knowledge in this area results in an underestimate of the video's influence on viewers' emotional responses, risking misclassification of the video.

2.3 The absence of existing classifications for EYV may result in a more serious situation if not addressed immediately

It's worrying to learn that extremism is only a click away. Detecting videos promoting extremism is critical, though it can be technically challenging [21]. Now, there is a single category of 'violent' videos that contain elements deemed to be violent and emotionally disturbing. A violent scene in a video can have a variety of visual and spiritual effects [44]. Exposure to the videos may result in some degree of self-radicalisation online [23, 28, 45, 46]. Additionally, high levels of exposure to violence have been linked to emotional desensitisation, a risk factor for more violent behaviour [47]. For instance, after viewing extremist video footage, a Malaysian undergraduate was inspired to wage war against the Thai military [48]. Due to YouTube's reliance on user feedback for video takedown, the lack of existing classifications for EYV may result in more serious problems relating to exposure to the videos.

Based on the above discussions, it is imperative to discover the knowledge about the impact of VM to viewers' emotions, thus motivating the research to examine visual characteristics that have an emotional effect on viewers in YouTube videos. This could contribute to the effort to combat and counter extremism by understanding how VM affects viewers' emotional reactions to EYV.

3. Methods

There are several past research applying KE to the evaluation of emotions in videos, films, or movies [49–51]. The gaining interest in the adoption of KE in research is due to the comprehensive ability to discover, quantify, and link emotions to design characteristics, and then incorporate them into new product design. Other methods, such as Product Emotion (PrEmo) [52], Geneva Emotion Wheel (GEW) [53], and PANAS-X [54] are limited to instruments or tools that address data collection but do not extend to data analysis and beyond. KE is also flexible where other tools such as PrEmo, GEW, and PANAS-X can also be incorporated into its methods.

Thus, KE is adopted in this research to discover, quantify, and link emotions to VM in EYV, in its attempt to develop a classification model based on the viewer's emotions and VM, which has been missing before.

The research was conducted in 4 phases: exploratory study, Kansei evaluation, Kansei analysis, and confirmatory study. *Figure 1* illustrates the phases, which are described in the following sub-sections.

3.1 Exploratory study

There were three stages involved in the exploratory study to determine valid specimens (Refer *Figure 1*).

Identification of initial video specimen

As mentioned previously, this research sought to ascertain young viewers' emotional responses to EYV, as this type of video is regarded as one of the most emotionally evocative forms of media. To determine the specimen's scope, the identification process began with a manual search on YouTube using the terms "extremist", "extremist group", and "violent extremist". The search returned thousands of videos, which had to be filtered as the majority of the videos were irrelevant. A manual review was then conducted to select initial video specimens using a set of control criteria, which included short video (30 seconds – 2 minutes), moving images, and visible violent acts.

Investigation of VM

VM are the visually appealing elements of videos that draw the viewer's attention and help identify the video's nature. The investigation into VM began with document analysis and was confirmed by video analysis. Iterative processes of skimming, reading, and interpretation were used to conduct the document analysis. The interpretation process entails a closer examination of the data and the creation of codes and categories based on the data characteristics in order to discover themes pertinent to the VM. Two documents were referred to in this regard: the Malaysian Communications and Multimedia Content Forum's content code and the Malaysian Film Censorship Board's dossier. The resulting themes, which comprise the VM, the descriptions, and attributes were then analyzed across all initial video specimens. The result has enabled the research to develop a specimen vs. VM classification matrix to be used in finalizing video specimens.

Finalizing video specimen

A set of rules were followed to identify valid specimens using the developed specimen vs. VM

classification matrix in accordance with KE methodology [55, 56]. The rules are: i) Tick 1 attribute under 1 VM, ii) Eliminate redundancy for specimens having exactly the same attributes, iii) Valid specimen must have 2 or more of the same attributes as the other.

The rule was followed to ensure that the intended statistical calculation could be performed correctly during data analysis.

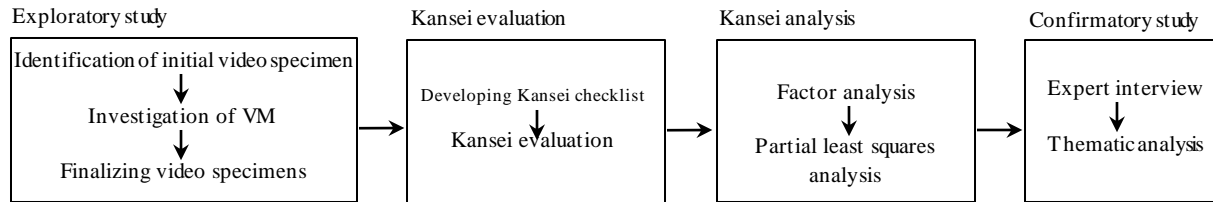


Figure 1 The research phases

3.2 Kansei evaluation

There were two stages involved in Kansei evaluation to obtain emotional responses.

Developing Kansei checklist

This research incorporated 62 emotional descriptors identified from past research related to emotional UX. The emotional descriptors were then arranged into a decreasing 5-point semantic differential scale to form a Kansei checklist. The checklist was used as an instrument in this research to measure the strength of emotional responses towards the valid specimens. The translation of each emotion into the native language was written next to the English word in an italic font style to help respondents to verify its meaning.

Kansei evaluation

The evaluation took place over the course of four separate sessions. It began with a briefing on informed consent and necessary instructions for the respondents. Respondents were asked to watch the specimens one by one and rate their emotional responses based on the checklist. The specimens were displayed via a projector with muted audio to set focus exclusively on the VM. Respondents were then asked to rate their emotional responses based on the Kansei checklist. The evaluation employed 80 young viewers as respondents who were recruited based on their familiarity with video watching experiences. Young viewers were selected as they are the easy target to get influenced by extreme messages [14–16]. Researchers have found that extremist supporters are rising among young people under 35 years of age [14, 16]. Furthermore, YouTube, among other social networking sites, has high engagement among young users [17].

3.3 Kansei analysis

Using the average evaluation result from the Kansei evaluation, two analysis procedures were conducted:

Factor analysis (FA)

The FA was conducted using the average Kansei evaluation result to find significant emotional structure. In determining the number of factors to represent the whole population of emotional dimensions, eigenvalue and cumulative percentage of variance were observed.

Partial least squares (PLS) analysis

In Kansei analysis, the PLS coefficient score is commonly used to identify which concept has significant influence over the domain under study [55, 56]. Thus, it was used by this research to determine which VM has a high influence on the emotional concept. To do this, first, in preparing the PLS data set, average data was appended to the specimen vs. VM classification matrix. Secondly, the PLS analysis was then executed using JMP Software version 13. To determine influence, the PLS range was then calculated for all VM attributes based on the scores. The range represents the distance between the minimum and maximum score within a VM. For instance, in *Table 1*, range for victim profile is calculated as the following:

$$Range = VM_{Max} - VM_{Min} = 0.1561 - (-0.1200) = 0.2761 \quad (1)$$

Subsequently, the mean of range was calculated and used as a baseline to determine significant influence on VM on respondents' emotions. For instance, in *Table 1*, mean range is calculated as the following:

$$\overline{Range} = \frac{1}{n} \sum_{i=1}^n Range_i = (0.2761 + 0.0948 + 0.2799 + 0.4532) / 4 = 0.2760 \quad (2)$$

The PLS result is used to formulate a classification model that enables the encoding of EYV from VM according to the dimension of emotional response.

3.4 Confirmatory study

The purpose of this phase is to evaluate the findings from the Kansei analysis. The ultimate aim is to formulate an effective classification model based on validation of the Kansei analysis results.

Expert interview

Open-ended questions were prepared based on the resulting significant emotions and VM from the Kansei evaluation phase. The questions were derived from the identified themes to structure the answers within the specific subject. Four experts with different professional backgrounds were interviewed to assess the developed model. This is consistent with the recommendation made by [55, 56] to invite one or more experts to collaborate on the development of the kansei product in order to gather rich data reflecting individual experiences, thus justifying the strategy of bringing in experts from diverse work backgrounds. Experts are the professional elite with specific interpretive knowledge (“know-why”) and procedural knowledge (“know-how”) because of their position and power [57]. In this research, the experts involved were a police officer, a cybersecurity specialist, a filmmaker and an internet outreach champion.

To evaluate the findings from the Kansei analysis, specimens that were highly associated with significant emotions were selected. The experts were then asked to view the specimen and then answer the question based on their work experience and knowledge. At the end of the session, the experts were invited to raise any concerns before the closing of the interview. There were also additional questions designed to probe for information if it did not come up. The interview was recorded and transcribed for the purpose of thematic analysis.

Thematic analysis

The research then performed thematic analysis of the interview results. The three-stage procedure used for analysis were (i) data reduction, (ii) data display, and (iii) drawing and verifying conclusions. Data reduction occurred through familiarization of the transcripts, coding and finding patterns. Drawing and verifying conclusions as the final part of the interview analysis involved developing propositions based on the themes that then served as answers to the research questions. When significant disagreement exists between experts in reaching a conclusion, the three-stage procedure is iterated until a firm conclusion which confirms the Kansei analysis result is drawn.

Table 1 Instance of range calculation result

VM	Attributes	Offensive	Range
Victim profile	Child/Woman	0.0421	0.2761
	Man	-0.1200	
	Mix	0.0240	
	Not human	0.1561	
Uniform	Not visible	-0.0474	0.0948
	Visible	0.0474	
Violent act	Weapon-based	0.1795	0.2799
	Physical-based	-0.0469	
	Both	-0.1004	
	Post violent act	-0.0920	
Logo	Not visible	-0.2266	0.4532
	Visible	0.2266	
Mean range			0.2760

4. Results

The results are discussed in the next sub-sections.

4.1 Exploratory study

4.1.1 The valid video specimen

An exhaustive review based on the control criteria has resulted in 65 YouTube videos to be used as initial video specimens. Then, the document analysis and video analysis conducted has enabled the research to identify a total of 12 VM (Refer *Table 2*). Based on the identified VM, this research carefully analysed each specimen against the VM to discover the attributes by identifying an observable pattern. For example, for the victim profile, the research has obtained a pattern where the VM could be classified as a man, a child, or woman, or a mixture. From the analysis, a total of 30 attributes were identified, and the attributes can be found in *Table 2*. The 30 attributes were then screened across 65 specimens, resulting specimen vs. VM classification matrix as shown in *Figure 2*. After a thorough analysis of the attributes based on the specified rules (refer *Figure 2*), twenty valid video specimens were finally determined. *Figure 3* shows an instance of a valid specimen indicating a few attributes. From the figure, we could see that the victim was a man. In snapshot A, highlighted with an arrow for visual marker 1 and 2, can be classified as a violent act of physical based on where the victim was blindfolded, and the perpetrator was holding his head. Snapshot B, highlighted with an arrow for visual marker 3, can be classified as a violent act of weapon based since it was visible that the victim was shot several times, although the weapon was not visible throughout the video duration. In snapshot C, highlighted in arrow for visual marker 4, can be classified as a post violent act as splattering of the brain is observable.

Table 2 Identified VM, the descriptions, and attributes

No. VM	Description	Attributes
1	Victim profile	Appearance of human or non-human victim
2	Victim quantity	Number of victims that are visibly seen
3	Violent act	Appearance of violent act towards victim
4	Weaponry	Weapons that are visibly seen in carrying out violent act
5	Uniform	Appearance of uniforms that are visibly seen being worn by victims or perpetrators
6	Symbol	Presence of symbol (e.g., flag, logo) that could represent identity of victims/perpetrators
7	Dead body	Presence of dead body
8	Bloody injury	Visibility of bloody injury as the result of violent act
9	Turmoil	Visibility of turmoil as the result of violent act (e.g., people running)
10	Infrastructure disruption	Visibility of infrastructure disruption as the result of violent act (e.g., building collapsed)
11	Scene setting	Setting of the video prolonging violent scene or otherwise
12	Facial expression	Visibility of facial expression of victim or perpetrator

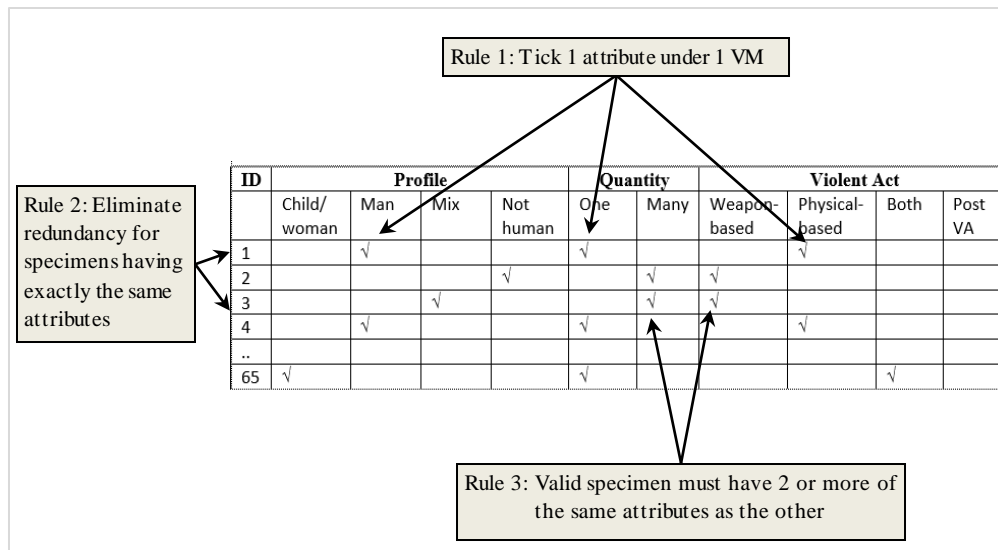


Figure 2 Specimen vs. VM classification matrix for selection of valid specimens

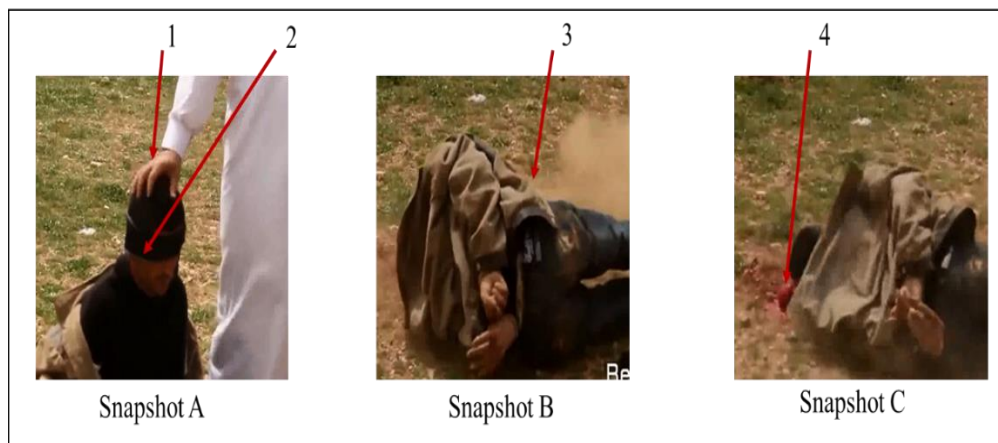


Figure 3 Instance of a valid specimen

4.2 Kansei analysis

Average data was calculated from data obtained from the Kansei evaluation and tested for internal consistency reliability. The analysis yielded an overall Cronbach's alpha value of 0.88, which is higher than the common benchmark value of 0.70.

4.2.1FA

The average evaluation data was used to execute FA to determine the significant emotional structure that was used to establish the concept for the EYV watching experience. *Table 3* shows that Factor 1 explains the majority of the data with a percentage of 48.07%, Factor 2 explains 22.83% and Factor 3 explains 4.52% of the data. The first two factors represent the majority of the factor contributions. This demonstrates that both factors have a dominant influence on the EYV watching experience. In terms of cumulative contribution, the first two factors both represent 70.90% of the variability, while the three factors explain 75.41% of the variability. This research has decided to include Factor 3 to increase the proportion that represents most of the data. The proportion of variability explained by the remaining factors is minimal (4.15% and less), thus considered insignificant and can be safely ignored. Factor loading of all variables in the first 3 factors was then examined.

Table 4 shows factor loading results in descending order, which facilitates observation of emotion concepts for EYV watching experience. Variables that had high factor loading scores were deemed

significant. The approach adopted in this research is to examine both the highest and second highest factor loadings for a particular factor. Significant loading thresholds were set at 0.8, 0.8 and 0.6 accordingly, in deciding significant emotional descriptors. As a result, 27 significant emotional descriptors associated with the 3 significant factors (Refer *Table 7*) were identified. Interpretation in naming the factors took two aspects into consideration – value and direction of each loading. Based on the high variable loading, Factor 1 is labelled as offensive. Factor 2 is labelled as intrigue considering the positive affect towards stimuli commonly sensed as negative influence. Factor 3 is labelled as awkward considering the majority of variables related to shyness towards what is commonly perceived as negative influence. These findings enabled the research to conclude 3 significant emotion concepts; offensive, intrigue and awkward as the most significant to describe the EYV watching experience. The remaining 35 variables can be safely disregarded as having low factor loading scores.

Table 3 Factor contributions

Factor	Contribution	Cumulative contribution
Factor 1	48.07%	48.07%
Factor 2	22.83%	70.90%
Factor 3	4.52%	75.41%
Factor 4	4.15%	79.57%
Factor 5	3.90%	83.46%
Factor 6	3.66%	87.12%

Table 4 Factor loading after varimax rotation

Variable	Factor 1	Factor 2	Factor 3
Angry	0.9457	-0.2317	0.0996
Afraid	0.9239	-0.2503	0.1007
Irritable	0.9213	-0.1570	0.0557
Loathing	0.9167	-0.3198	0.0833
Scared	0.9128	-0.2752	0.2527
Upset	0.9115	-0.2896	0.1975
Nervous	0.8863	-0.3121	0.1386
Distressed	0.8818	-0.3074	0.3021
Jittery	0.8805	-0.3440	0.2129
Surprised	0.8783	-0.0524	0.2018
Disgusted	0.8692	-0.3512	0.1015
Guilty	0.8682	-0.2027	0.1500
Violent	0.8673	-0.4124	0.0467
Confident	-0.1452	0.9318	0.0823
Lively	-0.0600	0.9132	0.1846
Inspired	-0.1640	0.8882	-0.0456
Energetic	-0.0793	0.8848	0.0265
Amazed	-0.3213	0.8817	-0.1616
Enthusiastic	-0.2682	0.8684	-0.2363
Interested	-0.2451	0.8545	-0.1528
Excited	0.0148	0.8464	-0.0950
Proud	-0.4069	0.8418	0.0819

Variable	Factor 1	Factor 2	Factor 3
Strong	-0.2726	0.8385	-0.3193
Timid	0.3145	0.0049	0.6623
Shy	0.1369	-0.0494	0.6596
Tired	0.4898	-0.3871	0.6396
Sheepish	-0.3419	0.0066	0.6282

4.2.2 PLS analysis

PLS analysis was performed to determine the relationships between emotional responses and VM in EYV. It also facilitated the discovery of influential VMs for each emotional response. The result is fundamental for the affective classification model development.

Influence of VM on emotional responses

In Kansei analysis, the PLS coefficient score is commonly used to identify which concept has influence over the domain under study [55, 56]. Thus, it was used by this research to determine which VM has a high influence on the emotional concept. *Table 1* shows an instance of the analysis of offensive emotions. There are columns VM, attributes, the PLS score for each attribute, and the calculated range. Within each VM, attributes that have largest positive score are the most influential VM on EYV to evoke viewers’ emotions. The largest negative score shows VM as the least influential. In the table, maximum and minimum scores are highlighted respectively. For instance, the victim profile that leads to offensive experience is the ‘not human’ as a subject; the least influential is the ‘man’ as a subject.

The mean range was used as a threshold to determine significant influence on VM to respondents’ emotions. According to the table, the mean range is 0.2760, and thus VM that have higher score than the mean range is

considered influential VM to viewers. It could be observed from the table that logo, violent act, and victim profile have a higher range than the mean range according to their influence order. Uniform has a lower score; thus, it is not a significant influence. The research concluded that the most influential VM to evoke offensive experience in EYV is the logo, followed by violent act and victim profile. The least influential visual marker is uniform, and is excluded as the score is below the mean range. The analysis was repeated for the rest of the variables. The overall result can be found in *Figure 4*. VM such as the logo, violent act, victim profile, scene setting, and victim quantity are the most influential in causing viewers to feel intrigue. Whereas VM such as violent acts, scene setting, weaponry, logo and victim profile are influential in causing viewers to feel awkward. It is observed that the set of influential VM is almost similar except that the influence order is different. Further to this result is the detail of the attribute for VM specification towards respective factors. This will be useful for effective video classification model development. The research observed details for all 27 variables identified for the respective factors to allow derivation of more rules that would be meaningful for classification development. Then, the correlations between dependent variables of the actual emotional responses and predicted values were calculated to show the reliability of the PLS and verify the model.

Offensive VM	Intrigue VM	Awkward VM
1. Logo	1. Logo	1. Violent Act
2. Violent act	2. Violent act	2. Scene setting
3. Victim profile	3. Victim profile	3. Weaponry
4. Weaponry	4. Scene setting	4. Logo
5. Scene setting	5. Victim Quantity	5. Victim profile

Figure 4 Influential VM to emotional experience by descending order

4.3 Confirmatory study

Firstly, the open-ended question was guided by themes identified from Kansei analysis, which were VM, offensive EYV watching experience, intriguing EYV

watching experience, awkward EYV watching experience, and towards a safer online experience. The experts were asked to give feedback based on these questions.

Secondly, confirmatory specimens were selected from the highly associated specimens for the three significant emotional concepts, i.e., offensive, intrigue, and awkward, based on the Kansei evaluation results. The research discovered that specimen 13 evoked intrigue, specimen 12 evoked a combination of offensive and awkward, specimen 5 and 11 evoked offensive, while specimen 17 and 20 evoked awkward. For expert evaluation, Specimens 11, 13 and 17 were chosen to represent offensive, intrigue and awkward EYV watching experiences, respectively. The experts viewed these specimens and answered questions based on the above themes. The study then analysed the transcribed data for the confirmatory study as the following.

4.3.1 Interview with experts

Four interview sessions were conducted in the confirmatory study. The first interview was with E1, a police officer; the second interview was with E2, a filmmaker; the third interview was with E3, a chief operating officer, and the fourth with E4, a senior director of the advocacy and outreach division of a regulatory body for communications and multimedia. The findings of the experts are described in the following sub-sections.

4.3.2 VM and its influence on emotion

In this research, VM is defined as an interesting part that calls for visual attention. It also refers to attributes in the video that make it identifiable. An example for better illustration would be a video containing an all-black uniform and a flag printed with Arabic script are VM for Islamic militants. *Table 5* consists of part of the transcript by experts with identified codes such as agreement, VM instances, and experience.

The Agreement is the code for experts' agreement with the statement that the VM indeed influences viewers' emotions. VM instances is the code for instances of VM mentioned by experts. Experience is the code for view on VM by giving examples based on their work experience. E1 commented on videos containing certain VMs as not suitable for young viewers to watch. Weapon usage and victims are among the VMs that viewers will see in videos containing violent content. E2 believes that VM can influence the emotions of the viewers. As a film maker, he had this experience where a smoking scene on a train was removed because such a scene potentially influenced the younger generation. E2 also added that the look and feel of the videos are real, hence they have so much emotional impact. Similarly, E3 agreed that VM is influential in giving emotional experiences to young viewers. E3 also stated that when one sees something visual, it influences one's

cognitive thinking and mind. From an advocacy point of view, E4 affirmed that VM plays a significant role in influencing the emotions of young viewers. E4 understood the use of VM in creating brand identity and the proposition that the malicious group was trying to allure the viewers. Beyond being influential, E4 added that the VM could be designed in such a way that it appeals to the intended target audience exactly as what the producer wanted it to be. Hence, it can be very deceiving. The experts have also elaborated on the concept of offensive, intrigue, and awkward in the EYV watching experience.

Offensive EYV watching experience

The Offensive has been described as an emotion related to highly dissatisfied or disagreeing with the content of EYV. To evaluate the offensive EYV watching experience, experts reviewed specimen 11 showing a soldier being captured and executed with more than 20 bullets. The highlight of the findings for this theme is in the following *Table 6*.

The results of statistical analysis have successfully identified violent acts, victim profile, weaponry, and scene setting as the top 5 influential VMs for this emotion. The table consists of part of the transcript by experts with identified codes such as agreement, violent act, weaponry, scene setting and impression. 'Agreement' is the code for experts' agreement with an offensive theme. 'Violent act', 'Weaponry' and 'Scene setting' are codes following identified visual markers in this research. 'Impression' is the code for views passed on by experts about EYV on this particular theme. The coding has enabled the research to observe patterns where experts highly discussed violent acts. According to E1, the overall visual in the video was offensive. E1 commented on the act of covering the victim's face, forcing him to speak, and the shooting execution. Although the weapon was not visible, the fact that the victim was shot was offensive. E2 also agreed that even without the presence of weapons, visuals can be offensive and highly impactful when viewers see the after effect of violent action. E2 gave an example of an after effect, such as splattering of the brain after the victim was shot. Furthermore, video that showed the execution of a victim would appear as offensive since the viewer just watches the particular scene without knowledge of the story behind it. E3 assured us that people have different opinions about extremist YouTube videos depending on their background. Some believe that the videos are all part of propaganda. From E3's personal view, such an act depicted in the video specimen should not happen. Killing or random shooting is not

normal. For E4, the video specimen shown was offensive in terms of blatant disregard for life. E4 commented that one acted like God and there was a victim whose life was so easily taken away. Similar to E3, E4 too expressed concern over some people who think that all these videos were staged and not real. The danger lies when they do not think it is real. Therefore, they are becoming less sensitive about it.

Intrigue EYV watching experience

Intrigue has been described as an emotion related to arousal of interest or curiosity towards content in the EYV. To evaluate the Intrigue EYV watching experience, experts reviewed specimen 13 showing an extremist group training children soldier. The highlight of findings from interviews for this theme is in Table 7.

The result of statistical analysis has successfully identified violent act, victim profile, scene setting and victim quantity as the top 5 influential VMs for this emotion. The table consists of parts of the transcript with identified codes such as agreement, impression, violent act, victim profile and experience. The coding has enabled the research to observe patterns where experts highly discussed impressions. E1 agreed that the video specimen caused him to feel intrigued. In E1 opinion, intriguing is inquisitive in nature. As for young viewers, they probably feel intrigue because they have never seen it before. Furthermore, E1 believes when children and women are the victims, as shown in the video specimen, that triggers extra

sympathy among the viewers. E2 commented on curiosity among the viewers rising for not knowing the reasons behind why the victims were treated that way. E2 also made an interesting judgement about viewers who felt intrigue upon seeing a child being beaten, probably because the circumstances were relatable to them. One who had past experience of being a bully at school might feel inspired by such videos. E3, however, does not feel comfortable knowing about this emotional dimension. The expert insisted that something is not right about young viewers who have a positive affect over sensitive content such as this. E3 speculated viewers rated the video as intriguing because there was an element of violence but not to the extent of killing. Besides, the involvement of children as victims also contributes towards heightening curiosity among viewers. The E4 point of view has similarity with the rest of the experts. Young viewers rated the video as intriguing because they had no prior knowledge of it and had not seen it before. Hence, they wanted to see and judge for themselves. The expert added, the fact that they know the video contains something sensitive and not suitable for their viewing, makes all the more reason for them to want to view it. Additionally, E4 made subtle comments about the possibility of these young viewers becoming desensitized to such videos. The fact is that the more often they see violent content, the less sensitive they become towards it.

Table 5 Sample coding for VM

Experts' views on VM		Code
E1	'... not suitable for young viewers to watch'	Agreement
	'...use of weapon and crowd element'	VM instances
E2	'...it gives so much emotional impact'	Agreement
	'...VM can influence emotion of the viewers'	Agreement
	'...cut the smoking scene on a train for my movie.'	Experience
E3	'...agree that the VM see in the videos are influential...'	Agreement
	'...visual...influence your cognitive thinking...'	Agreement
E4	'...VM play significant role...not just influencing but actually it is more insidious'	Agreement
	'...help to create brand identity...'	Experience

Table 6 Sample coding for offensive EYV

Experts' views on offensive EYV watching experience		Code
E1	'...because what was done to the victim.'	Agreement
	'...covered the victim's face and forcing the victim to speak...'	Violent act
	'...shot using some weapon though the weapon was not visible...'	Violent act, weaponry
E2	'...scene on execution...without knowing the story behind, that can be violent...'	Scene setting, agreement
	'Even without the presence of weapon...'	Weaponry
	'...shooting and after effect...'	Violet act
E3	'...extremist propaganda.'	Impression
	'It is really offensive...'	Agreement

Experts' views on offensive EYV watching experience		Code
	'...killing people or random shooting...'	Violent act
E4	'...blatant disregard for life.'	Agreement, violent act
	'...life is so easily taken away.'	Violent act
	'...Western propaganda.'	Impression

Table 7 Sample coding for intrigue EYV

Experts' views on intrigue EYV watching experience		Code
E1	'I feel intrigue...'	Agreement
	'...feel more intrigue perhaps because they have never seen it before.'	Impression
	'...norm to have sympathy more on children and women.'	Impression, victim profile
E2	'...cause curiosity by viewers for not knowing why the victims are willing...'	Agreement, impression
	'...feel intrigue because the visual is relatable to them.'	Impression
E3	'...no killing and involve children, maybe that is why they rated is as intriguing'	Agreement, violent act, victim profile
	'...not comfortable about this.'	Impression
E4	'Of course, there are tendency to feel intrigue.'	Agreement
	'...natural curiousness about things that you do not know.'	Impression
	'...more often you see violent content, the less sensitive you are to it.'	Experience

Awkward EYV watching experience

Awkward has been described as an emotion related to shyness or discomfort towards content in the EYV. To evaluate the awkward EYV watching experience, experts reviewed specimen 17 showing torturing acts among groups of different religions. The results of statistical analysis have successfully identified violent acts, scene setting, weaponry, logo and victim profile as the top 5 influential VMs for this emotional dimension. The highlights of the findings from the interview for this theme are in *Table 8*.

The table consists of part of the transcript by experts with identified codes such as disagreement, impression, agreement, violent act, suggestion, experience, and scene setting. After seeing a group of men circling several victims who took turns being whipped on their backs, E1 expressed the opinion that the video is more intriguing rather than awkward, as

the punishment as shown in the video could be due to a small offense committed by the victims. On the contrary, E2 firmly stated that it was awkward because the victims seemed to be willing to get beaten and they did not fight back. E2 presumed that it could probably be an act of sacrifice related to extreme religion ritual. E3 was surprised that young viewers felt awkward about what seemed like an act of extremism. However, E3 did not negate the fact that certain people do have divergent thinking from the common. E3 agreed that the video caused awkward feelings. The punishment of victims, if it was meant to be according to a specific law, did not seem to reflect accordingly. That to E3 was awkward. E4 instead did not feel awkward watching the video, but recognized that viewers would feel awkward for what seems like street justice and no apparent reason why the victims were treated that way.

Table 8 Sample coding for awkward EYV

Experts' views on awkward EYV watching experience		Code
E1	'...more intriguing for me rather than awkward...'	Disagreement
	'...perhaps because it is not a big offense...'	Impression
E2	'Awkward because they seem to be willing...'	Agreement
	'...did not fight back...not being tied nor being held.'	Violent act
E3	'...something needs to be done on these students for having such feeling.'	Impression
	'...agree it is awkward...'	Agreement
	'...certain protocol in executing the punishment...'	Experience
E4	'...did not feel awkward...'	Disagreement
	'...street justice.'	Scene setting
	'...why he was beaten...'	Violent act

5. Discussions

Despite the substantial body of research demonstrating the power of visual images to influence viewers, the majority of past studies placed a greater emphasis on

comprehending the words in extremists' messages than on visuals [3, 20, 38], and paid little attention to the viewer's affect [2]. Past investigations involving extremist video classification largely focused on

textual metadata, such as video titles, manual tags, editorial categories, and user feedback [4, 5, 21,41, 42]. While the videos hint at the emotional influence, the videos were not classified in a way that takes emotion into account [38, 43]. The lack of knowledge in this area results in an underestimate of the video's influence on viewers' emotional responses, risking misclassification of the video. Thus, this research was motivated to fill in the knowledge gap by focusing on emotional elements in the response to EYV, and embarked on developing an affective classification model based on emotional UX and VM in YouTube video.

The research has set its objective firstly to identify VM in EYV. Document analysis conducted on the Malaysian Communications and Multimedia Content Forum's content code and the Malaysian Film Censorship Board's dossier, and video analysis has enabled the research to identify a total of 12 VM in YouTube videos. Further analysis of the specimens has helped the research to identify 30 attributes for the 12 VM. The successful achievement of the objective is demonstrated in *Table 2*.

Secondly, the research embarked on determining significant viewers' emotional responses upon watching EYV. The research initially identified 62 emotional descriptors, of which 27 were found to have a significant influence on EYV watching experience, distributed mostly among three significant factors. These factors, based on the semantics, were represented by the keyword's offensive, intrigue, and awkward respectively. These keywords were concluded as the emotion concepts best suited to describe the EYV watching experience. Prior research revealed that viewers of vivid violence exhibited stronger emotional reactions and higher attention levels [58]. The representation of the semantics also matches with the affect curve as discussed by [59]. This is in accordance with the reconciliation of basic and dimensional perspectives of emotion models, where each basic emotion can be mapped into an n-dimensional perspective. The affect curve extends from pleasant to unpleasant emotions. Investigation thus far has shown that certain specimens evoke a specific emotional response.

The analysis concluded that offensive emotions related to high displeasure or disagree with the content in EYV. Intrigue is an emotion related to the arousal of interest or curiosity about the content of EYV. Awkward is an emotion related to shyness or inconvenient with EYV content. The successful

achievement of the objective is demonstrated in *Figure 4*.

The ultimate target of this research is to develop an affective classification model based on emotional UX and VM in YouTube videos. The PLS analyses conducted have enabled this research to formulate an affective classification model based on emotional UX and VM in YouTube videos. It is a descriptive model that describes the EYV watching experience and the relationships between VM. In developing the model, results from FA were used to conceptualise emotional dimensions and results from PLS were used to derive rules for the affective classification model.

For respective emotional concepts, 5 influential VMs were discovered. Within each VM, the attributes that have the largest positive score are the most influential VMs on EYV to evoke viewers' emotions. The result has enabled the research to classify the most influential attributes of the 5 influential VMs in each emotional concept. This will be a sound guide for recognition of VMs that cause a particular emotional response. *Table 9* shows the concluded emotion vs. VM classification rules. To illustrate an example of classification rules, an EYV is classified as offensive if it matches the VM as classified within the dimension.

Through a confirmatory study, the research has successfully evaluated the classification rules and thus enabled the development of the intended affective classification model based on emotional UX and VM in YouTube videos (*Figure 5*). Under caution, as discovered by the expert study, there is a potential for viewers to become sympathizers and desensitized. Viewers could develop sympathy for the victims and, over time, subconsciously become sympathizers supporting the extremism agenda. EYV are known to contain sensitive content and may not be suitable for public viewing. It is commonly accepted as causing negative effects. This research has found a contrary discovery in the intrigue emotional dimension, which is associated with positive affect. If the positive affect is due to natural curiosity, it is already questionable. Even more so, when viewers are becoming less sensitive to violent content and instead focus on curiosity rather than feeling that 'this should stop now.' This is called desensitization, where less empathy and sympathy are developed for the depicted victims [47]. Awkward, too, has the potential for desensitization. This could be due to viewers becoming less sensitive to the victims and instead focusing on the awkwardness of the situation rather

than despising it. The model is useful to facilitate classification of EYV according to the emotional dimension evoked as part of the watching experience. EYV may contain violent content, not suitable for public viewing and commonly accepted as inappropriate. However, the finding from this research ratifies a rather delicate and worrying truth. With this in mind, the developed model is best used hand-in-hand with the emotion vs. VM classification table, as it provides reference to significant emotional dimensions within the emotion concept. The evaluated model is demonstrated in *Figure 5*.

The model provides new knowledge about how certain VMs in videos could influence viewers' emotions, thus psychologically affecting viewers' sentiments. This fills in the gap of knowledge regarding the effects of VM. Without this knowledge, it was difficult to comprehend the VM and how they might have influenced viewers' emotions, which could have led to an incorrect assessment of the video's influence on viewers.

The findings of this study, particularly on the identification of visual markers and the relationships between emotional responses, could benefit society by providing them with a sound reference to become actively involved in community policing with regards to countering violent extremism. EYV embedded with children as victims (visual markers), for instance, potentially evoked an emotional dimension of 'intrigue'. At home, parents who are in the know could use this information to create awareness among their children. It should also extend to schools, offices, and communities where everyone in society has a role to play in monitoring and educating. Researchers from different fields could use the findings and discussions as a solid premise to extend, enhance or even explore new research areas. The possibilities are massive, from automated affective video classification, violent scene detection, machine learning high-level visual content analysis, behavioural intention induced by extremist videos, and production of counter violent extremism videos. It is hoped that with the achievement of all the above, the spread of EYV will take a different turn.

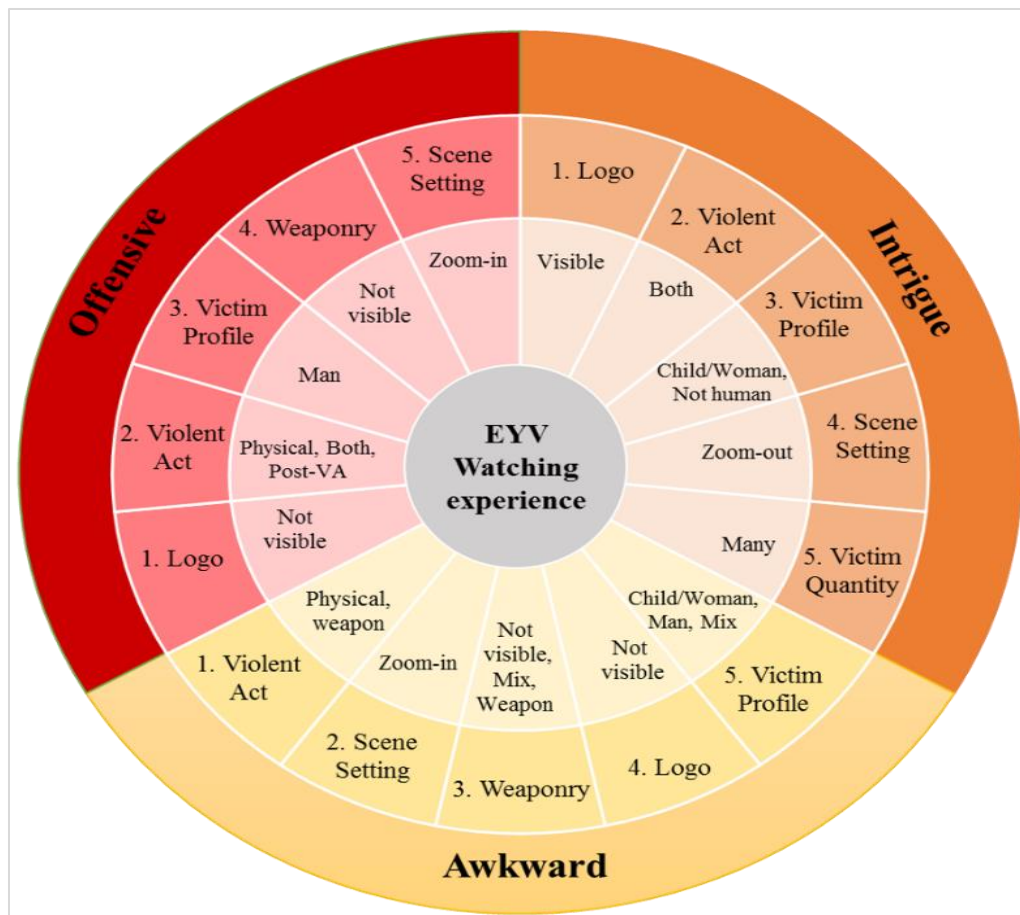


Figure 5 The affective classification model of EYV

Table 9 Emotion vs. VM classification rules

Emotion	Descriptor	VM					
		Victim profile	Victim quantity	Violent Act	Weaponry	Logo	Scene Setting
Offensive	angry	Man	One	Physical-based	Not visible	Not visible	Zoom-in
	afraid	Man	Many	Post VA	Not visible	Not visible	Zoom-in
	irritable	Man	Many	Both	Not visible	Not visible	Zoom-in
	loathing	Man	Many	Physical-based	Not visible	Not visible	Zoom-in
	scared	Man	Many	Post VA	Not visible	Not visible	Zoom-in
	upset	Man	Many	Physical-based	Not visible	Not visible	Zoom-in
	nervous	Man	One	Post VA	Not visible	Not visible	Zoom-in
	distressed	Man	Many	Physical-based	Not visible	Not visible	Zoom-in
	jittery	Man	Many	Post VA	Not visible	Not visible	Zoom-in
	surprised	Man	Many	Post VA	Not visible	Not visible	Zoom-in
	disgusted	Man	One	Physical-based	Not visible	Not visible	Zoom-in
guilty	Man	Many	Post VA	Not visible	Not visible	Zoom-in	
violent	Man	One	Physical-based	Not visible	Not visible	Zoom-in	
Intrigue	confident	Child / woman	Many	Both	Edged	Visible	Zoom-out
	lively	Not human	Many	Both	Edged	Visible	Zoom-out
	inspired	Child / woman	Many	Both	Edged	Visible	Zoom-out
	energetic	Not human	Many	Both	Edged	Visible	Zoom-out
	amazed	Not human	Many	Both	Edged	Visible	Zoom-out
	enthusiastic	Not human	Many	Both	Edged	Visible	Zoom-out
	interested	Not human	Many	Both	Edged	Visible	Zoom-out
	excited	Child / woman	Many	Both	Edged	Visible	Zoom-out
	proud	Not human	Many	Both	Edged	Visible	Zoom-out
strong	Not human	Many	Both	Edged	Visible	Zoom-out	
Awkward	timid	Child / woman	Many	Physical-based	Not visible	Not visible	Zoom-in
	shy	Man	Many	Physical-based	Not visible	Not visible	Zoom-in
	tired	Man	Many	Physical-based	Not visible	Not visible	Zoom-in
	sheepish	Mix	Many	Physical-based	Mix weapon	Visible	Zoom-in

6. Conclusion and future work

It is known that EYV can be offensive, contains violent content and supposedly not allowed for public viewing. However, filtering of the content has been challenging for policy makers and the authorities. With the developed affective classification model of EYV they will have a sound reference in building filtering mechanism for video contents. To society and digital native, the new findings from this study, particularly on the identification of the VM and the relationships of the emotional responses, which has never been discovered before, will benefit them in helping them to actively involved in community policing with regards to countering malicious intentions. For instance, EYV embedded with children as victim (VM) potentially evoked emotional dimension of intrigue. At home, parents could use this information to create awareness among their children. It should also extend to school, offices, and communities at large where everyone in the society has a role to play in monitoring and educating. This is in tandem with the authority's encouragement that the public to act for those social media accounts which defied the terms and conditions of their use [37]. The

main novelty of this research relies on the Kansei (emotional) elements in video classification. It delves into what and how the Kansei is influenced, which give actionable insights to policy makers and authorities. The what is the itemized VM and emotional responses, and the how is the relationship of VM and emotional responses. EYV usually contains violent content, and it affects the viewer's emotion. The research also found that viewers responded positively to certain EYV, which normally is not anticipated for video of such nature. This provides empirical evidence that emotional responses towards sensitive content is rather subjective and requires careful management of the contents filtering.

The developed model, however, was a result of investigation under certain limitations. As described earlier, the research was scoped to visual features in videos that require the action of 'seeing'. Hence, the research results are limited to visceral responses to visual features. This limitation offers an opportunity for further investigation to include other features such as text features and language for better generalization of knowledge contribution. Secondly, the term extremist refers to groups or individuals who are

motivated by religious or political ideology and manifest support or commit violence. Thus, the model is not applicable to other types of extremism or violence, such as against animals or domestics. Thirdly, the evaluation subjects were young Malaysian viewers, thus a generalization to a wider scope of population, background, thus a generalization to a wider population and background requires further investigation.

The discussed scope and limitations are useful as a blueprint for future research to extend the research findings to larger populations and subjects, as well as comparative analysis to enhance the generalizability and establish a more comprehensive classification model. Although the model was based on work under certain limitations, it lends some novelty by discovering emotional elements in video classification. The result contributes to the domain with emphasis on UX for its applicability and methodology used. It outlines information on VM and how it affects the viewer's emotions, in which the knowledge can be used to devise an algorithm to filter videos. Researchers and producers could use the model in producing videos capable of invoking emotions among the viewers. This research could also benefit beyond the UX field. Researchers and scholars could use the model as a solid premise to extend, enhance or even explore new research areas. The possibilities are many, from automated affective video classification, violent scene detection, visual content analysis, and behavioural intention induced by EYV. For example, the findings on emotional dimensions for EYV and the associated influential VM could be used to develop an algorithm for automated emotion-based video classification. This will eventually reduce the spread of EYV, paving the way for safer and healthier communities online as well as in the offline world. A complete list of abbreviations is shown in *Appendix I*.

Acknowledgment

The authors would like to acknowledge Research Initiative Group for Emotion, Kansei and Design Engineering (RIG EKDE), UiTM, as well as Malaysia Association of Kansei Engineering (MAKE), for all their assistance to the research paper. This work was supported by UiTM and the Ministry of Higher Education (600-IRMI/FRGS 5/3 (328/2019).

Conflicts of interest

The authors have no conflicts of interest to declare.

References

[1] Rosli RM, Lokman AM, Aris SR. Identification of influential visual markers in extremist YouTube videos from the viewpoint of younger viewers through partial

least square (PLS) analysis. In international conference on Kansei engineering & emotion research 2018 (pp. 664-70). Springer, Singapore.

- [2] Zhao S, Yao H, Sun X. Video classification and recommendation based on affective analysis of viewers. *Neurocomputing*. 2013; 119:101-10.
- [3] Ahmad S, Asghar MZ, Alotaibi FM, Awan I. Detection and classification of social media-based extremist affiliations using sentiment analysis techniques. *Human-centric Computing and Information Sciences*. 2019; 9:1-23.
- [4] Tulga AY. ISIS religious and extremist propaganda on social media: dictionary-based study of twitter. *Small Wars Journal*. 2020.
- [5] Alfifi M, Kaghazgaran P, Caverlee J, Morstatter F. A large-scale study of ISIS social media strategy: community size, collective influence, and behavioral impact. In proceedings of the international AAAI conference on web and social media 2019 (pp. 58-67).
- [6] Medida LH, Ramani K. An optimized e-lecture video retrieval based on machine learning classification. *International Journal of Engineering and Advanced Technology*. 2019; 8(6):4820-7.
- [7] Ledwich M, Zaitsev A. Algorithmic extremism: examining YouTube's rabbit hole of radicalization. *arXiv preprint arXiv:1912.11211*. 2019.
- [8] Kale A, Wakde DG. Video retrieval using automatically extracted audio. In international conference on cloud & ubiquitous computing & emerging technologies 2013 (pp. 133-6). IEEE.
- [9] Mussiraliyeva S, Omarov B, Bolatbek M, Ospanov R, Baispay G, Medetbek Z, et al. Applying deep learning for extremism detection. In international conference on advanced informatics for computing research 2020 (pp. 597-605). Springer, Singapore.
- [10] Liu Y, Albanie S, Nagrani A, Zisserman A. Use what you have: video retrieval using representations from collaborative experts. *arXiv preprint arXiv:1907.13487*. 2019.
- [11] Fernandez M, Alani H. Artificial intelligence and online extremism: challenges and opportunities. *Predictive Policing and Artificial Intelligence*. 2021:132-62.
- [12] Salem A, Reid E, Chen H. Multimedia content coding and analysis: unraveling the content of jihadi extremist groups' videos. *Studies in Conflict & Terrorism*. 2008; 31(7):605-26.
- [13] Speckhard A, Ellenberg M, Baddorf Z. Breaking the ISIS brand counter narrative project: understanding, preventing, and intervening in militant jihadi terrorism and violent extremism. In *from Territorial Defeat to Global ISIS: Lessons Learned 2021:94-111*. IOS Press.
- [14] Conway M, Khawaja M, Lakhani S, Reffin J. A snapshot of the Syrian jihadi online ecology: differential disruption, community strength, and preferred other platforms. *Studies in Conflict & Terrorism*. 2020:1-7.
- [15] Rieger D, Frischlich L, Bente G. Dealing with the dark side: the effects of right-wing extremist and Islamist extremist propaganda from a social identity

- perspective. *Media, War & Conflict*. 2020; 13(3):280-99.
- [16] Dauber CE, Robinson MD, Baslios JJ, Blair AG, Zeidel R, Al-hashimi H. Call of duty jihad: how the video game motif has migrated downstream from Islamic state. *Perspectives on Terrorism*. 2019; 13(3):1-17.
- [17] Light B, Griffiths M, Lincoln S. 'Connect and create': young people, YouTube and Graffiti communities. *Continuum*. 2012; 26(3):343-55.
- [18] Vergani M, Zuev D. Analysis of YouTube videos used by activists in the Uyghur nationalist movement: combining quantitative and qualitative methods. *Journal of Contemporary China*. 2011; 20(69):205-29.
- [19] Cowen AS, Keltner D. Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences*. 2017; 114(38):E7900-9.
- [20] Winkler CK, Dauber CE. Visual propaganda and extremism in the online environment. Army War College Carlisle Barracks Pa Strategic Studies Institute; 2014.
- [21] Agarwal S, Sureka A. A focused crawler for mining hate and extremism promoting videos on YouTube. In proceedings of the ACM conference on hypertext and social media 2014 (pp. 294-6).
- [22] Moor JH. The future of computer ethics: you ain't seen nothin'yet!. *Ethics and Information Technology*. 2001; 3:89-91.
- [23] Torok R. Social media and the use of discursive markers of online extremism and recruitment. In research anthology on religious impacts on society 2021 (pp. 734-64). IGI Global.
- [24] Moussaoui M, Zaghdoud M, Akaichi J. A possibilistic framework for the detection of terrorism-related Twitter communities in social media. *Concurrency and Computation: Practice and Experience*. 2019; 31(13).
- [25] Khader M. Combating violent extremism and radicalization in the digital era. IGI Global; 2016.
- [26] Bowman-grieve L. A psychological perspective on virtual communities supporting terrorist & extremist ideologies as a tool for recruitment. *Security Informatics*. 2013; 2:1-5.
- [27] Dogrul M, Aslan A, Celik E. Developing an international cooperation on cyber defense and deterrence against cyber terrorism. In international conference on cyber conflict 2011 (pp. 1-15). IEEE.
- [28] Winter C, Neumann P, Meleagrou-hitchens A, Ranstorp M, Vidino L, Fürst J. Online extremism: research trends in internet activism, radicalization, and counter-strategies. *International Journal of Conflict and Violence*. 2020; 14(2):1-20.
- [29] <https://www.independent.co.uk/voices/comment/uk-s-far-right-using-islamic-state-powerful-propaganda-tool-and-it-s-working-9751657.html>. Accessed 30 September 2014.
- [30] <https://edition.cnn.com/2015/01/25/asia/malaysia-airlines-website-hacked/index.html>. Accessed 26 January 2015.
- [31] Ribeiro MH, Ottoni R, West R, Almeida VA, Meira JW. Auditing radicalization pathways on YouTube. In proceedings of the conference on fairness, accountability, and transparency 2020 (pp. 131-41).
- [32] <https://www.buzzfeednews.com/article/ishmaeldaro/YouTube-comments-hate-speech-racist-white-nationalists-super>. Accessed 17 May 2018.
- [33] Grinnell D, Macdonald S, Mair D, Lorenzo-dus N. Who disseminates Rumiya? examining the relative influence of sympathiser and non-sympathiser Twitter users. In proceedings of the 2nd European counter terrorism centre advisory group conference 2018 (pp. 1-18).
- [34] Woo J, Son J, Chen H. An SIR model for violent topic diffusion in social media. In proceedings of international conference on intelligence and security informatics 2011 (pp. 15-9). IEEE.
- [35] Fu T, Huang CN, Chen H. Identification of extremist videos in online video sharing sites. In international conference on intelligence and security informatics 2009 (pp. 179-81). IEEE.
- [36] <https://www.YouTubecom/yt/about/policies/#community-guidelines>. Accessed 1 December 2018.
- [37] <https://www.interpol.int/en/Who-we-are/Member-countries/Asia-South-Pacific/MALAYSIA>. Accessed 30 June 2014.
- [38] Monaci S. Social media campaigns against violent extremism: a new approach to evaluating video storytelling. *International Journal of Communication*. 2020; 14:980-1003.
- [39] Ekman M. The dark side of online activism: Swedish right-wing extremist video activism on YouTube. *MedieKultur: Journal of Media and Communication Research*. 2014; 30(56):79-99.
- [40] Andre V. 'Neojihadism' and YouTube: patani militant propaganda dissemination and radicalization. *Asian Security*. 2012; 8(1):27-53.
- [41] Al-rawi A, Groshek J. Jihadist propaganda on social media: an examination of ISIS related content on twitter. *International Journal of Cyber Warfare and Terrorism (IJCWT)*. 2018; 8(4):1-5.
- [42] Chen LH, Hsu HW, Wang LY, Su CW. Violence detection in movies. In international conference computer graphics, imaging and visualization 2011 (pp. 119-24). IEEE.
- [43] Chen YL, Chang CL, Yeh CS. Emotion classification of YouTube videos. *Decision Support Systems*. 2017; 101:40-50.
- [44] Wang H, Yang L, Wu X, He J. A review of bloody violence in video classification. In international conference on the frontiers and advances in data science 2017 (pp. 86-91). IEEE.
- [45] Vacca JR. Online terrorist propaganda, recruitment, and radicalization. CRC Press; 2019.
- [46] Bertram L. Terrorism, the internet and the social media advantage: exploring how terrorist organizations exploit aspects of the internet, social media and how these same platforms could be used to counter-violent extremism. *Journal for Deradicalization*. 2016:225-52.

- [47] Mrug S, Madan A, Windle M. Emotional desensitization to violence contributes to adolescents' violent behavior. *Journal of Abnormal Child Psychology*. 2016; 44(1):75-86.
- [48] Samuel TK. The lure of youth into terrorism. *SEARCCT's Selection of Articles*. 2011; 2:107-19.
- [49] Abd KS, Lokman AM, Tsuchiya T. Emotional responses towards unity YouTube videos: experts vs. viewers perspectives. *International Journal of Affective Engineering*. 2021.
- [50] Lokman AM, Awang AA, Zaili RA, Fathir MF. Designing racial unity through films: the Kansei engineering approach. *Advanced Science Letters*. 2016; 22(5-6):1368-72.
- [51] Sato N, Anse M, Tabe T. A method for constructing a movie-selection support system based onkansei engineering. In *symposium on human interface and the management of information 2007* (pp. 526-34). Springer, Berlin, Heidelberg.
- [52] Desmet PM. Design for mood: twenty activity-based opportunities to design for mood regulation. *International Journal of Design*, 2015; 9(2):1-19.
- [53] Sacharin V, Schlegel K, Scherer KR. Geneva emotion wheel rating study. *Archive ouverte UNIGE*. 2012.
- [54] Watson D, Clark LA. The PANAS-X: manual for the positive and negative affect schedule-expanded form. University of Iowa.1999.
- [55] Lokman AM. *Pencenalan rekayasa kansei (in Bahasa Melayu)*. KL: Dewan Bahasa dan Pustaka. 2019.
- [56] Lokman AM. KE as affective design methodology. In *international conference on computer, control, informatics and its applications 2013* (pp. 7-13). IEEE.
- [57] Bogner A, Littig B, Menz W. *Interviewing experts*. Springer; 2009.
- [58] Riddle K. Transportation into vivid media violence: a focus on attention, emotions, and mental rumination. *Communication Quarterly*. 2013; 61(4):446-62.
- [59] Hanielic A, Xu LO. Affective video content representation and modeling. *IEEE Transactions on Multimedia*. 2005; 7(1):143-54.



Roshaliza Mohd Rosli received her bachelor's degree in Information Technology from Universiti Teknologi Petronas, Malaysia in 2002 and the MSc. and Phd. degrees also in Information Technology from Universiti Teknologi MARA, Malaysia in 2014 and 2019, respectively. Her research interest includes User Experience, Kansei Engineering and Internet Ethics. She is currently working with Malaysia Digital Economy Corporation (MDEC) to serve the nation in leading the digital economy forward as a manager for Industry Adoption Programme.
Email: roshaliza@gmail.com



Syaripah Ruzaini Syed Aris received her BSc in Information System Engineering and MSc in IT in 2003 and 2004 respectively. She then obtained her PhD degree in Information Technology from Universiti Teknologi MARA. Currently, she is working as Coordinator for Postgraduate Program for Center of Information System Studies, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA. Research interest include IT Management, Risk Management, Information System Strategic Planning and Crowdsourcing. She is a member of IEEE.
Email: ruzaini@tmsk.uitm.edu.my



Azhar Abd Aziz received his BSc MIS from the University of Rhode Island, USA in 1996 and MIT and PhD degrees in Information Technology from Universiti Kebangsaan Malaysia in 2000 and 2020, respectively. Currently is a senior lecturer in the Faculty of Computing and Mathematical Sciences, in the Universiti Teknologi MARA, Malaysia. Research interest includes Computer Ethics, Robotic Intervention for Children with Learning Disability and Community Digital Division. He is a committee member of Malaysian Association of Kansei Engineering (MAKE).
Email: azhar@tmsk.uitm.edu.my



Toshio Tsuchiya was educated as a system engineer at Hiroshima University and is currently a professor of Shimonoseki City University. He received Ph.D. from Hiroshima City University in 2006. Assistant at faculty of engineering, Hiroshima university (1994-1996). Lecturer at Shimonoseki city university (1996-1998). Associate professor at Shimonoseki city university (1998-2007). Professor at Shimonoseki city university (2007-). Guest researcher at Linkoping university (2004). Honorary research fellow at the University of Leeds (2004-2005). Visiting researcher at Universiti Teknologi MARA (2016-2017). His research interests Include Data Mining, Knowledge Engineering, Fuzzy Set Theory and Kansei Engineering.
Email: anita.uitm@gmail.com



Anitawati Mohd Lokman is a professor in Information Technology at University Teknologi MARA Malaysia. She holds a Bachelor of Engineering in Information Engineering in 1994 from a Japanese university, Master of Science (IT) in 2001 from Universiti Teknologi MARA, and Doctor of Philosophy in Science (IT) in 2009. Upon receiving her first degree she served a Japanese multinational semiconductor company for 8 years before joining the academia in 2002 until today. She is a Board Member of Kansei Engineering & Emotion Research International Executive Board, and Editorial Member of the International Journal of Affective Engineering. She has published more than 200 journal articles and conference papers, and several books and book chapters with local and international publisher. Her research interest is in Emotional User Experience, Kansei Engineering, Emotion & Design, and Emotion & Wellbeing. She is a Professional Technologist, and a Senior Member of IEEE.

Email: anita.uitm@gmail.com

Appendix I

S.No.	Abbreviation	Description
1	EYV	Extremist YouTube Videos
2	FA	Factor Analysis
3	GEW	Geneva Emotion Wheel
4	KE	Kansei Engineering
5	MAKE	Malaysia Association of Kansei Engineering
6	MCMC	Malaysian Communications and Multimedia Commission
7	MDEC	Malaysia Digital Economy Corporation
8	PANAS-X	Positive and Negative Affect Schedule
9	PLS	Partial Least Squares
10	PrEmo	Product Emotion
11	RIGEKDE	Research Initiative Group for Emotion, Kansei and Design Engineering
12	UX	User Experience
13	VM	Visual Markers