'Mind the Gap': Evaluating User Physiological Response for Multi-Genre Video Summarisation

Arthur G. Money
Brunel University
London, UK
arthur.money@brunel.ac.uk

Harry Agius
Brunel University
London, UK
harry.agius@brunel.ac.uk

Abstract: Existing video summarisation techniques are often only capable of summarising video from pre-specified content genres and are often not able to produce personalised summaries as they are not able to source relevant user specific data. Because users often experience strong emotions and associated physiological responses whilst watching video, their physiological response to video content may serve as a new and valuable data source for video summarisation. Previously, we developed the Entertainment-Led Video Summarisation (ELVIS) technique that summarises video based on five physiological response measures: electro-dermal response (EDR), heart rate (HR), blood volume pulse (BVP), respiration rate (RR), and respiration amplitude (RA). Here, we report a statistical analysis on a range of data collected from ELVIS in trials with 100 users relating to five distinct video content genres (Action, Drama, Romance, Horror and Comedy). The results show that the ELVIS, EDR, HR, BVP, RR and RA video summaries all consistently match with the most entertaining video sub-segments as self-reported by the user, and that the composite ELVIS video summaries achieve significantly higher level of overlap compared with a RANDOM selection. More generally, users reported that, compared with video summaries produced by another contemporary video summarisation technique, ELVIS video summaries are comparatively 'enjoyable' and 'informative' for all five video content genres. We therefore conclude that video summarisation according to users' physiological responses has great value for future development of video summarisation techniques that are applicable across a wide range of video content genres.

Keywords: video summarisation, personalisation, community, collaborative systems, physiological response, affect, emotion

1. INTRODUCTION

The amount of digital content video available to us is growing at an exponential rate [1] and thus users increasingly require support to access content in more efficient and effective ways. Hence, an important research challenge is to develop new techniques that support efficient and effective storage and retrieval of digital video content [2]. Video summarisation research responds to this need by developing techniques that identify and abstract the most important sub-segments from the content of a full length video. The video summaries that arise are abbreviated representations of the original video. These may subsequently be used within a range of multimedia applications, such as interactive browsing and searching systems, and offer the user a valuable means of managing and accessing digital video content [1].

Feature-based video summarisation techniques have been the most popular approach taken within this research domain over the last two decades or so. They focus on summarising video by analysing features that are present only within the video stream, such as shape, colour, object motion, speech, or on- screen text [3]. Typically, these techniques are designed to summarise specific genres of video, such as sports videos,

home videos, drama, action, news video and so forth. Although these techniques achieve high levels of automation by requiring little or no manual supervision, they pay a heavy price in terms of versatility as they are only capable of summarising video for specific types of content or genres effectively [4]. Some examples of feature-based techniques include the work of Duxans et al. [5] analyse the audio stream of football videos, in particular they analyse two acoustic features; the sound energy level, and an index of sound repetition. Content with the highest level of sound energy and repeated sounds are included in the video summary. Correa and Ma [6] demonstrate a video summarisation technique that is ideally suited to 'performance video' and sportscasts. Initially, the technique distinguishes foreground moving objects from static background objects for each respective frame within a video. Background frames are then assimilated, and snapshots of foreground movement are then transposed onto a candidate set of representative static background images. The result is a series of video mosaics that summarise the narrative presented within video.

Context and user-based video summarisation techniques attempt to overcome the challenge of producing more personalised video summaries by analysis of additional contextual and/or user-based information that is sourced from outside of the

video stream. User-based information is additional information relating to the content of a video that is sourced directly from the user. Contextual information is any additional information relating to the content of a video that is not sourced directly from the video or user. For example, San Pedro et al. [7] present a user-based 'social summarisation' technique, which identifies video scenes uploaded to multimedia oriented social websites such as YouTube. They identify the most important scenes by establishing overlaps in uploaded video content and calculating the number of users that have uploaded the respective video scenes. Joho et al. [2] present a user-based summarisation technique that analyses user facial expressions collected via a Webcam that records the user's facial expressions as they view video content. The latter automatically analyses the user facial expressions and subsequently tags the video content with 'facial expression tags' such as Neutral, Positive, Negative, and Rejective.

Although a range of feature-based, context and user-based video summarisation techniques exist, the majority fail to overcome the trade-off that exists between *Content*, *Automation* and *Personalisation*, which we term the *CAP gap*. If significant progress is to be made within the video summarisation research domain, new information sources must be identified that make progress in overcoming this trade-off. Figure 1 demonstrates the CAP gap trade-off faced by current video summarization techniques.

Figure 1: The CAP gap

While feature-based summaries benefit from providing highly automated video summarisation solutions, such techniques are often domain specific; only offering video summarisation solutions for specific types of video *content*, most notably genre [1]. Furthermore, video summaries are not *personalised* according to the individual's personal comprehension of a video and hence are often inconsistent with the criteria humans perceive and remember video content by [8]. Context and user-based techniques have the advantage of delivering personalised video summaries for a range of video genres, yet the majority achieve this at the cost of only being applicable to specific content genres and/or through low levels of automation by requiring conscious input from the user in the form of manual annotation.

New user-based information sources and associated summarisation techniques are needed that respond positively to all three criteria that represent the CAP gap, i.e. techniques that effectively summarise video content across a wide range of genres, whilst maintaining high levels of automation, and also producing highly personalised video summaries. Consequently, we look towards user physiological responses to video content as a new user-based information source to summarise video content. Physiological response is a user-based information source, since it can be captured directly from the user while requiring no conscious effort from them. Video content has been used for decades to elicit strong physiological responses in the user [9].

With the exception of our own work [3, 4], physiological response is yet to be fully incorporated into existing video summarisation techniques. In this study, we utilise our ELVIS (Entertainment-Led Video Summarisation) technique, a novel automated user-based technique that identifies temporal sub-segments of a given video segment (VS) for inclusion within a video summary, based on real-time user physiological responses to that video content. The remainder of this paper is presented as follows: Section 2 provides an overview of how automated physiological video summarisation may be achieved via physiological response to video content. Section 3 presents the user trial design and statistical analysis methods used to evaluate the effectiveness of ELVIS across five video content genres. Section 4 presents the results of the user trials, Section 5 concludes this study.

2. PHYSIOLOGICAL-RESPONSE-BASED VIDEO SUMMARISATION

For some time, physiological response measures have served as a recognised means of evaluating users' emotional responses to video content [9]. Real-time changes in user physiological responses may also be used to infer the user's *affective state* [10], which is a generic term that refers to the user's underlying emotion, attitude, or mood at a given point in time. Affective state may be considered to comprise of two key dimensions: *valence*, the level of attraction or aversion the user feels toward a specific stimulus, and *arousal*, the intensity to which an emotion elicited by a specific stimulus is felt. A variety of physiological responses have been used to infer changes in a user's affective state, the most common of which are as follows:

- *Electro-Dermal Response (EDR)* measures the electrical conductivity of the skin and is believed to be linearly correlated with the arousal dimension [11].

- *Respiration amplitude (RA)* has been used to indicate arousal and valence levels [12].
- *Respiration rate (RR)* has been used as an indicator of arousal. [13].
- *Blood Volume Pulse (BVP)* measures the extent to which blood is pumped to the body's extremities. This has been used to serve as a measure of a user's valence [14].
- *Heart Rate (HR)* acceleration and deceleration has also been shown to be an indicator of valence [15].

We have developed the ELVIS technique which analyses user physiological response data relating to the video the user is viewing to produce video summaries. ELVIS therefore provides the opportunity to explore whether users' physiological responses may serve as a suitable user-based source of information for producing automated and individually personalised affective video summaries and helping to close the CAP gap. ELVIS makes the assumption that user physiological responses are likely to be most significant and pronounced whilst being exposed to sub-segments of video content that have the most personal relevance to that individual user. Since these sub-segments are likely to have the most impact and will also be the most memorable; it is these sub-segments that become the foremost candidates for inclusion within a summarised version of a given video stream. Figure 2 shows how physiological responses are incorporated into the ELVIS technique and how video summaries are output.

Figure 2: Overview of ELVIS's approach

As a first step, the user engages in the viewing of a full length video whilst their physiological responses are captured. The processing of this captured data is then carried out by ELVIS, from which an internal temporally indexed representation of the video content is produced in terms of the significance of physiological responses throughout the duration of the video. ELVIS then identifies the temporal locations of the most significant physiological responses scaled according to the duration of the video summary requested by the user. A notable point is that currently it is infeasible to map physiological responses onto a full range of specific emotions [10], hence we do not aim to summarise and label the discrete emotions that occur within a

video. The ELVIS technique uses the five measures of physiological response (EDR, HR, BVP, RR and RA) as presented above, and assumes that each of these measures have equal levels of importance in representing the user's response to video content. It is also assumed that the higher the number of individual significant physiological responses to a video sub-segment (VSS), the higher the likelihood that this is indicative of an entertaining VSS. The significant responses identified by ELVIS can then be associated temporally with the viewed video content by a media playback application. The video summary may then be played back to the user. The output of this process is a personalised video summary incorporating the VSSs that elicited the most significant physiological responses during viewing of the video. A full formal description of the ELVIS technique is presented in [3].

3. USER EVALUATION OF ELVIS VIDEO SUMMARIES

A trial protocol was developed in order to evaluate ELVIS video summaries across five video content genres as well as varying levels of abstraction. An overview of the trial protocol is presented in Figure 3.

Figure 3: Trial protocol for evaluating ELVIS video summaries

User trials were carried out with five groups of 20 users (N=100 total users). The primary aim of the user trials was to evaluate the extent to which ELVIS produced personalised video summaries, i.e. identified the most entertaining video sub-segments compared with video sub-segments each individual user personally self-reported as being the most entertaining. In order to establish the effectiveness of ELVIS across a wide range of video content genres, user trials were carried out for content representing five distinct video content genres; Action, Drama, Romance, Horror and Comedy. After collecting physiological response data from each participant (20 viewings of each video segment), video summaries were produced from their physiological responses for each individual physiological measure (EDR, HR, BVP, RR and RA) as well as composite ELVIS summaries based on the combination of all of these measures (ELVIS). In order to establish the feasibility of using physiological response measures as a means of

summarising video content, the extent to which ELVIS video summaries matched self-reported entertaining sub-segments was compared with the extent to which randomly selected video sub-segments (RANDOM) matched self-reported sub-segments. These randomly generated video summaries served as a baseline against which the performance of ELVIS could be compared. Users were also asked to rate ELVIS video summaries in terms of their perceived enjoyability and informativeness. The results of these ratings were compared with user ratings for video summaries produced via an alternative video summarisation technique presented in the literature. In summary, the aims of evaluating video summaries produced by ELVIS may be considered as follows:

- Evaluate the extent to which ELVIS video sub-segment selections match the most entertaining video sub-segments as self-reported by the user, compared with a RANDOM selection and each respective individual physiological measure (EDR, HR, BVP, RR and RA) across a range of video content genres.

- Evaluate the extent to which users find the video summaries enjoyable and informative at 4%, 10% and 25% levels of abstraction, compared with an existing summarisation technique.

4. RESULTS

A statistical analysis of overlap percentages achieved by the seven video sub-segment selection procedures (ELVIS, RANDOM, EDR, HR, BVP, RR, and RA) was carried out for each group of 20 users to evaluate the extent to which each of the respective video sub-segment selections achieved statistically significant overlaps with self-reported video sub-segments. The primary aim of this analysis was to evaluate the performance of ELVIS in identifying the most entertaining video sub-segments compared with RANDOM and each respective physiological measure. For each of the five video segments used, the user trials can be considered as a within subjects one-way repeated measures design with seven treatment conditions: ELVIS selection, RANDOM selection, EDR selection, HR selection, BVP selection, RR selection, and RA selection. Paired t-tests were performed for overlap scores achieved by ELVIS, compared to RANDOM and each of the physiological measures (EDR, HR, BVP, RR, and RA). The P-values produced by the paired t-tests allowed the performance of ELVIS to be evaluated to establish whether it outperforms RANDOM, EDR, HR, BVP, RR and/or RA to a statistically significant degree ($\alpha = 0.05$). The results for 10% summaries are now presented.

4.1 Paired t-tests for 10% summaries

The results of the paired t-tests comparing mean overlap values for ELVIS summaries against six treatment conditions for the Action, Drama and Romance, Horror and Comedy, at the 10% level of abstraction, are presented in Table 1.

Table 1: Paired t- test, for Action, Drama and Romance, Horror, Comedy 10% summaries

Action 10% summaries		Paired Differences				
		Mean	Std. Dev.	t	df	Sig. (1-tailed)
Pair 1	ELVIS_10 - EDR	2.23	4.48	2.23	19	0.019
Pair 2	ELVIS_10 - RA	5.22	4.27	5.47	19	0.000
Pair 3	ELVIS_10 - RR	3.72	3.29	5.06	19	0.000
Pair 4	ELVIS_10 - BVP	4.23	4.86	3.90	19	0.001
Pair 5	ELVIS_10 - HR	1.45	4.19	1.55	19	0.069
Pair 6	ELVIS_10 - RAN	8.18	3.68	9.95	19	0.000
Drama 10% summaries						
Pair 1	ELVIS_10 - EDR	2.11	4.89	1.93	19	0.035
Pair 2	ELVIS_10 - RA	5.37	4.02	5.97	19	0.000
Pair 3	ELVIS_10 - RR	5.28	3.10	7.62	19	0.000
Pair 4	ELVIS_10 - BVP	5.25	4.08	5.75	19	0.000
Pair 5	ELVIS_10 - HR	4.10	3.41	5.38	19	0.000
Pair 6	ELVIS_10 - RAN	8.75	3.62	10.82	19	0.000
Romance 10% summaries						
Pair 1	ELVIS_10 - EDR	1.72	5.79	1.33	19	0.100
Pair 2	ELVIS_10 - RA	4.65	5.81	3.58	19	0.001
Pair 3	ELVIS_10 - RR	3.72	4.53	3.67	19	0.001
Pair 4	ELVIS_10 - BVP	2.99	4.41	3.04	19	0.004
Pair 5	ELVIS_10 - HR	3.61	5.07	3.19	19	0.003
Pair 6	ELVIS_10 - RAN	7.57	5.16	6.56	19	0.000
Horror 10% summaries						
Pair 1	ELVIS - EDR	1.79	3.75	2.13	19	0.024
Pair 2	ELVIS - RA	4.68	3.64	5.75	19	0.000
Pair 3	ELVIS - RR	3.84	3.47	4.95	19	0.000
Pair 4	ELVIS - BVP	1.62	2.54	2.85	19	0.005
Pair 5	ELVIS - HR	4.09	3.48	5.26	19	0.000
Pair 6	ELVIS - RAN	6.49	3.78	7.68	19	0.000
Comedy 10% summaries						
Pair 1	ELVIS - EDR	2.07	2.88	3.21	19	0.003
Pair 2	ELVIS - RA	5.36	2.14	11.19	19	0.000
Pair 3	ELVIS - RR	4.12	2.51	7.34	19	0.000
Pair 4	ELVIS - BVP	6.94	3.30	9.41	19	0.000
Pair 5	ELVIS - HR	1.09	2.27	2.16	19	0.022
Pair 6	ELVIS - RAN	8.26	3.12	11.85	19	0.000
= Statistically significant at the 5% level						

Included in the results are the mean paired differences in percentage overlap measures of significance of the differences in mean overlap (1-tailed), and the mean paired differences in percentage overlap between ELVIS, EDR, HR, BVP, RR, RA, and RANDOM for Action, Romance and Drama, Horror and Comedy VSs respectively.

As can be seen from the "Sig. (1-tailed)" column, ELVIS performed significantly better than RANDOM at the 5% level of significance (which is shown by 100% (0.000) confidence level at three decimal points) for Action, Drama, Romance, Horror and Comedy video segments (VSs). Therefore, in statistical terms, ELVIS performed significantly better than RANDOM for all five VSs. Based on these results there is strong evidence that ELVIS achieves on average significantly higher

mean percentage overlap scores compared to a RANDOM 10% VSS selection.

The one-tailed t-tests comparing mean overlap percentages for Action, Drama, Romance, Horror and Comedy summaries at the 10% level of abstraction revealed that for HR, EDR, RA, RR, and BVP, in absolute terms, ELVIS achieved higher mean percentage overlap scores. EDR, however, was not statistically different for Romance. HR was not statistically different for Action.

4.2 Results: Enjoyability and informativeness

Table 2 reports on enjoyability and informativeness scores for 4%, 10%, 25%, 100% summarised content achieved by ELVIS summaries generated specifically for each individual user (ELVIS) for the five video content genres. ELVIS results are compared with the results of Ngo et al. [16], who also reported 10%, 25%, and 100% summaries.

Table 2: Enjoyability and informativeness scores

	Enjoyability				Informativeness			
	4%	10%	25%	100%	4%	10%	25%	100%
ELVIS - Act.	3.36	3.72	3.98	4.41	3.41	3.61	4.06	4.49
ELVIS - Dra.	3.27	3.42	3.79	4.46	3.35	3.47	3.84	4.51
ELVIS - Rom.	3.32	3.59	3.80	4.29	3.39	3.53	3.90	4.43
ELVIS - Horr.	3.25	3.51	3.75	4.25	3.39	3.50	3.98	4.46
ELVIS - Com.	3.35	3.62	3.93	4.30	3.44	3.57	4.01	4.59
Ngo [2005]		3.19	3.67	4.53		3.25	3.81	4.62

Ngo et al. [16] demonstrated that as the amount of the full-length content included in the video summaries increased (i.e. the total percentage of the video included in the summary), so the self-reported scores for both enjoyability and informativeness increased. This was also the case for all ELVIS video summaries across all genres. ELVIS enjoyability scores ranged between 3.25 (4%) and 3.98 (25%) for summarised content, indicating that users tended to agree with the enjoyability statement. ELVIS Action summaries scored the highest absolute values for 4%, 10% and 25% with 3.36, 3.72 and 3.98 respectively. ELVIS summaries achieved scores between 3.35 (4%) and 4.06 (25%) for informativeness, which also indicated that users tended to agree with the informativeness statement. ELVIS Action summaries once again were highest for 10% and 25% levels of abstraction with 3.61 and 4.06 respectively, however, ELVIS Comedy summaries achieved the highest score for 4% abstraction with 3.44.

Comparing ELVIS scores with Ngo et al. [16] scores, all scores 10% and 25% summary scores across all genres were higher than Ngo et al. [16]. Therefore, this indicates that users found ELVIS summaries more enjoyable and informative, compared with those produced by Ngo et al. [16]. Indeed, even at 4%, both ELVIS summaries achieved comparatively higher enjoyability and informativeness scores than the 10% summaries of

by Ngo et al. [16], despite there being a tendency for users to report higher enjoyability and informativeness for video summaries that contain higher proportions of the original video content. These are encouraging results, especially when considering that the original full-length (100%) video content used by Ngo et al. [16] achieved higher enjoyability scores than the original 100% VSSs used by ELVIS for all genres, with scores ranging between 4.25 and 4.46 compared with 4.53 for Ngo et al. [16]. This was also the case for informativeness scores, with ELVIS 100% VSSs scores ranging between 4.43 and 4.59, whereas the Ngo et al. [16] 100% Vs scored 4.62.

5. CONCLUDING DISCUSSION

The majority of existing video summarisation techniques struggle to overcome the *Content, Automation and Personalisation (CAP) gap*, most notably because they are typically only capable of summarising video content from a pre-specified video content genre due to their high levels of automation. New user-based information sources enable the development of techniques that may make some progress in overcoming the CAP gap and, hence, users' physiological responses to video content were proposed as a potentially valuable source here.

To this end, we evaluated the effectiveness of ELVIS, a physiological-response-based video summarisation technique, in achieving summarisation of video content from multiple genres. The results revealed that, across five video content genres, ELVIS consistently produced significantly more personalised video summaries (i.e. achieved the highest level of overlap with self-reported most entertaining VSSs) compared with a randomly selected VSS selection and VSS selections made up of individual physiological response measures. These results indicate that ELVIS consistently produced the most personalised and entertaining VSSs. This is promising and indicates that exploring new user-based information sources, such as physiological response for video summarisation may serve as a valuable new information source to aid the video summarisation community in overcoming the CAP gap.

Comparing the results of ELVIS summaries in terms of enjoyability and informativeness, ELVIS achieved higher mean scores at the 4% level of abstraction than Ngo et al. achieved at the 10% level. This is noteworthy, particularly when considering that enjoyability and informativeness scores tended to improve as the percentage abstraction level increased, highlighting that ELVIS performed well according to these criteria. In terms of enjoyability and informativeness at 10% and 25%, ELVIS also outperformed Ngo et al. [16] in every case and for every video content genre. This indicates that not only are physiological-response-

based video summaries capable of being versatile and applicable across a range of video content, but are also able to produce video summaries that are enjoyable and informative compared with existing video summarisation techniques.

In more general terms, the findings of this study have a number of implications for the video summarisation research community, and in particular in making progress in overcoming the CAP gap. Some of these implications include:

- ELVIS, an automated user-based technique has been shown to consistently identify the most entertaining video sub-segments for individual users across a range of video content genres. This demonstrates that the physiological-response-based approach shows promise in overcoming the CAP gap and as serving as a new information source that may be incorporated into new video summarisation techniques.
- The findings that physiological response data can be used to produce personalised video summaries also indicates that physiological response data may serve as a new and valuable candidate user-based information source that may be incorporated into future user-based video summarisation techniques.
- Given the high levels of automation that may be achieved when using physiological response data, there is now a real opportunity to develop hybrid video summarisation techniques, i.e. fusing together the typically automated feature-based techniques with physiological response based techniques. Potentially realising much higher levels of summarisation accuracy than either of these approaches being developed in isolation.
- Competitive levels of enjoyability and informativeness were achieved while responding positively to all three CAP gap criteria. This provides further indications that physiological-response-based techniques are capable of making notable progress in overcoming the CAP gap while also producing viewable video summaries as reported by the user.

Our results indicate that user-based video summarisation is increasingly becoming a valuable avenue of research, showing real promise in closing the CAP gap in future video summarisation systems. Future research should explore how applicable physiological-response-based video summarisation techniques are within more applied settings, since the findings of this study were set within the context of laboratory experiments.

3. REFERENCES

1. Money, A.G. and H. Agius, *Video summarisation: A conceptual framework and survey of the state of the art*. Journal of Visual Communication and Image Representation, 2008. **19**(2): p. 121-143.
2. Joho, H., et al., *Exploiting facial expressions for affective video summarisation*, in *ACM International Conference on Image and Video Retrieval2009*, ACM: Santorini, GR. p. 1-8.
3. Money, A.G. and H. Agius, *ELVIS: Entertainment-Led Video Summarisation*. ACM Transactions on Multimedia Computing, Communications and Applications 2010. **6**(3): p. 1-30.
4. Money, A.G. and H. Agius, *Analysing User Physiological Responses For Affective Video Summarisation*. Displays, 2009. **30**(2): p. 59-70.
5. Duxans, H., X. Anguera, and D.B.M.S.a.B. Conejero, 2009. BMSB '09. IEEE International Symposium on, *Audio based soccer game summarization*, in *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB '09)2009*, IEEE: Bilbao, Spain. p. 1-6.
6. Correa, C.D. and K. Ma, *Dynamic Video Narratives*. ACM Transactions on Graphics, 2010. **29**(3): p. 1-9.
7. San-Pedro, J., V. Kalnikaite, and S. Whittaker. *You can play the again: Exploring social redundancy to derive highlight regions in videos*. in *ACM International Conference on Intelligent User Interfaces*. 2009. Sanibel Island, Florida, USA: ACM.
8. Ma, Y., et al., *A generic framework of user attention model and its application in video summarization*. IEEE Transactions on Multimedia, 2005. **7**(5): p. 907-919.
9. Detenber, B.H., R.F. Simons, and G. Bennett, *Roll 'em!: The effects of picture motion on emotional responses*. Journal of Broadcasting & Electronic Media, 1998. **42**(1): p. 113-127.
10. Scheirer, J., et al., *Frustrating the user on purpose: A step toward building an affective computer*. Interacting with Computers, 2002. **14**: p. 93-118.
11. Gomez, P., W. Stahel, and B. Danuser, *Respiratory responses during affective picture viewing*. Biological Psychology, 2004. **67**(3): p. 359 - 373.
12. Philippot, P., C. Chapelle, and S. Blairy, *Respiratory feedback in the generation of emotion*. Cognition and Emotion, 2002. **16**: p. 605-627.
13. Gomez, P. and B. Danuser, *Affective and physiological responses to environmental noises and music*. International Journal of Psychophysiology, 2004. **53**(2): p. 93-103.
14. Healey, J.A., *Wearable and Automotive Systems for Affect Recognition from Physiology.*, in *Department of Electrical Engineering and Computer Science2000*, MIT: Cambridge, MA, USA. p. 158.
15. Cacioppo, J.T., et al., *The psychophysiology of emotion across the lifespan*. Annual Review of Gerontology and Geriatrics, 1997. **17**: p. 27-74.
16. Ngo, C., Y. Ma, and H. Zhang, *Video summarization and scene detection by graph modeling*. IEEE Transactions on Circuits and Systems for Video Technology, 2005. **15**(2): p. 296-305.