Sensory technology for measuring engaged viewership of broadcast television content

Sivakumar Ramachandran

Institute of Art, Design and Technology, Dún Laoghaire,
Republic of Ireland

Abstract

Broadcast television viewership is traditionally measured using a demographically optimised sample, whereby a technological installation at the viewer's home, monitors the channels being watched, when and for how long. Such an approach is used to inform broadcasters of viewing share and in the case of channels carrying advertising, a means to gauge revenue. However, these methods reveal little about the engagement of the viewer in the programmes. At the outset, there is no data collected as to what the viewer might be doing whilst the television set is on, whether they are for instance looking at other platforms on smart phones, or simply the television is on in the background, or emotional company for the lonely. It is not a true reflection of broadcast television consumption.

Sensory technology can be used to monitor and understand viewership more accurately. For instance, cameras can be used to gauge the amount of time viewers are watching the screen, their pupil dilation, emotional response in facial expressions. Sensors can monitor the sweat response on the skin. Body language can be monitored via a combination of sensors and cameras.

It is argued that an enhanced sensory approach to monitoring viewership gives a truer representation of engagement. It is a first quantitative step to then fully appreciate whether viewers are 'hooked' on particular television programmes, how they relate to them personally and how they influence interactions with family, friends, and colleagues. Coupled with an understanding of the psychological factors surround viewership, such an approach it is suggested can give broadcasters and producers a much better rating of the success of televisual content, which in turn may result in a better commercial and economic model of broadcast. Sustainability of broadcast channels can be better gauged.

Keywords: Television, Ratings, Engagement, Facial, Emotional, Technology.

Introduction

This paper discusses the need for improvements in the gauging of television viewership, using advanced technologies to measure emotional response. In this way, the amount of engagement can be determined, to ascertain whether the television device is simply switched on as a background device or not.

The study which forms the basis of this paper considered the traditional methods of capturing viewership since the early days of broadcast, and its limitations.

With the onset of the multi-channel environment, and services offered as Video on Demand (VOD) or internet-based OTT (Over the Top), the culture of engagement changed. The paper then considers what engagement means in this context.

The study then considers the optimum methods of gauging emotional response from non-intrusive, external sensor means; facial expression recognition, body language, pupil dilation and sweat.

The paper finally considers the ramifications of gauging emotional engagement in television programmes from the wider contexts such as screenwriting, production, commercial revenue, and mental health.

Historical aspects of television ratings measurement

Television ratings measurement began in the USA, following a similar device being used for radio ratings in the 1920s. There was an early recognition that it was not necessary to gather data from every household with a radio set, but to collect a sample based upon a cross-section of the listening public (Kelly 2017, 113-132). Selection was based upon factors such as age, earnings, ethnicity, and religious belief. Such data stratification was already familiar to the pollsters, who practiced techniques applied to general elections. elections. A device called an Audimeter (Nielsen 1945) was used to track which station a listener was tuned to. It is interesting however that the criteria largely stayed the same when applied to television broadcast.

In the same way, when a similar device to the Audimeter was applied to television channel monitoring, the programme audio track was the indicator that the viewer had tuned a particular channel (Nielsen 1945, 239-255). By the late 1940s, technology facilitated telephone lines to communicate the data to collection centres, wherein pollsters previously employed for political campaigns manually recorded and analysed 'viewership'. This process was intensive in terms of human activity, however data collection was confined to large urban and metropolitan areas, where television reception was concentrated, at a time when audiences were slowly growing, post-World War 2.

By the 1970s, paper-based methods of data collection and analysis were too difficult and technological development at the time allowed the data to be stored on computer (Bourdon and Méadel, 2014). The principle behind this measurement method has changed little since these early days. In fact, the principle of the audio track being the quantitative indicator of viewership has not changed at all. However, there has been much greater and ongoing development in the analysis of the data.

The Nielsen method of data collection and analysis was adopted worldwide by both commercial broadcasters and national authorities. This dominance, whilst giving unique commercial advantage to one company, does provide a harmonised framework for the method of television rating measurement, which is especially important for content consumed worldwide in a variety of different time and platform related ways. This early establishment of the methodology in a global television landscape, therefore, means that 'Nielsen' is the *defacto* standard in television audience measurement. Including the USA, thirty or so broadcast monitoring agencies around the world directly use a Nielsen based system.

The understanding of the social demographics of viewers in terms of categories such as age, ethnicity, political opinion and attitudes has become more sophisticated, both in terms of statistical technique, but also in conjunction with greater understanding of the social dimensions, their influences and dynamics over time and due to impact events, such as Covid-19. Moreover, these analyses are tailored to the cultural and linguistic settings of the 'home' markets of content; where programmes are first made accessible (Born, 2000).

More sophisticated analyses, however, do not consider whether the viewer is watching the television channel under scrutiny, and is engaged in it.

The agencies employed in television ratings measurement have varied over the growth period of employing the Nielsen techniques. In some countries where television consumption is regulated via a legislative mechanism, government appointed bodies oversee the measurements. In other countries where there is essentially deregulation, industry appointed bodies perform these tasks at the behest of the broadcasters.

In addition to this, broadcasters themselves gauge public opinion by fielding complaints about their services on official websites, or by television programme which themselves handle complaints. Opinionated individuals from time to time make their own views and emotions felt, achieving prominence. In the UK for instance in the 1970s, Mary Whitehouse ran campaigns against what she felt was poor television, at a time when the BBC was countering this by producing television programmes in which the public critiqued its own audio-visual output (Tracey and Morrison, 1979).

One of the most current and well-known viewer critique television franchises in the English speaking and European countries is *Gogglebox* (known for instance as *Aquí mando yo* in Spain). In these shows, viewers watch a diverse social-demographic set of individual households watching the week's television programmes in their country. This is classed as 'fly on the wall' reality television, and whilst the responses from viewers do not necessarily seem spontaneous, the viewers of Gogglebox can see emotional responses, particularly the close-ups on the faces of the programme's participants, whilst they watch television from their sofa sets at home. (https://www.imdb.com/title/tt6078248/). The Gogglebox programmes viewed are a mixture of conventional and internet streamed content. In the age of Covid-19, there is assumed to be a particularly high level of camera automation so that the production company minimise their interaction with the households, for health and safety reasons.

Modern conventional television ratings measurement

The essential Nielsen based principle of monitoring the television channel which is chosen by the viewer is implemented in slightly different ways across the world, with variations of the technology used to achieve the same ends. The Nielsen methodology, to draw upon a voluntary small sample of people from diverse demographics has been maintained.

In the last 15-20 years, technology has progressed, such that there has been an improvement in both the verification of which household member is watching at any one time, but also the user interface, with a degree of confidentiality (Buzeto and Moyana, 2013, 53-62).

Nielsen based technology now includes devices such as UNITAM, a handheld user interface that requires individual viewers to log into the monitoring system before selecting a channel. Thus, the viewer, channel, viewing time and duration are deemed to be reliably recorded. Updated UNITAM based systems can monitor viewership regardless of digital viewing whether via satellite, cable or internet connected computers/smart phones/tablets. VOD (Video on Demand) type services or exclusively OTT (Over the Top) one can all be monitored. It is still the audio track of the television programme which is the elemental data source in viewership.

There is a complex set of measurement criteria and scenarios associated with the Nielsen system, that relate to aspects such as time-shifting, repeated programmes within the week, multi-channel simultaneous broadcast and any other such aspects that need to be accounted for, to process accurate data on viewership.

In the Republic of Ireland, television ratings are performed by an organisation called Television Audience Measurement (TAM) Ireland, which uses the

Nielsen system of data collection and analysis (The Nielsen Group/TAM Ireland, 2017).

The analytical and statistical methodological framework of the Nielsen system is well established. However, the data collected has not changed since even before the birth of television when the audiometer was used in the 1920s for a limited multi-channel radio environment.

In a multi-channel environment where there is a dilution of audience density across so many niche genres, a much more complex relationship between broadcasters and viewers and expectations within a social media framework, that are vastly different from its early days, it is not surprising that the Nielsen system is being challenged in terms of accuracy.

However, there are global hits on the VOD streaming channels, in the English-speaking world, the Queen's Gambit and Tiger King being recent examples. These programmes at a core level, engage the audience, transgressing international boundaries, with imagery and suspense that sustains audiences which then translate into many social conversations and self-actualising promotion (particularly in a Covid-19 lockdown) online. Simple counting of viewership does not capture this level of engagement.

Furthermore, with so many forms of audio-visual, sensor, and physical stimulation available in the twenty-first century, the proposition that the viewer would always be engaged in the television programme being accessed, seems unlikely when one considers the hour-by-hour activity in a home or elsewhere, even during the Covid-19 pandemic. There are a multitude of other activities at home that adult and child viewers could be engaged in while a television programme is running in the background; cooking, chores, playing, minding children, collecting delivered items, or an intimate time with a loved one. Their attention and emotions may not be focused on the television, and the programme might not be engaging enough for the viewer to pause its transmission. Even if someone is continuously watching a television programme, it is possible that they are not engaged in it, that the mind is wandering. Non-invasive, modern technology such as cameras and sensors can reveal with much greater accuracy, whether a viewer is casual or not.

A scientific approach, looking at audience response is necessary. This would involve a small sample of viewers on a demographic basis, in the framework of the Nielsen methodology, being observed by cameras place above the television device, or using wearable, portable sensors to chart emotional response. This will reveal whether there is direct emotional engagement with the content. Such technology must not transgress personal data protection. If the viewer is not stimulated artificially by alcohol or other substances whilst watching television, a realistic assessment of emotional engagement can be made.

How to monitor in real time, emotional response

Whilst social media is a good way to gauge engagement in TV programmes statistically, conclusions cannot be drawn completely, as most peripheral discussion is done publicly through text online, via Facebook, twitter etc., or sometimes through visual representation on Instagram or TikTok. Otherwise through private messages and audio/video conversations such interactions. Since Covid-19, the 'water cooler' moment or physical get togethers of people to chat about their experiences, for instance in pubs and bars has diminished.

Given this change in social interaction, the home becomes the focus of understanding viewership. However, a more sophisticated method of measurement is needed in order to elucidate the true popularity of a TV show.

Facial expression

Facial expression is the best indicator of emotional engagement when watching television.

Outside of the film and television industry, for instance in the security sector, computer algorithms have been developed to identify different facial expressions and defining them as classes of muscle movement. These classes are correlated to types of emotional response. Statistical techniques such as Bayesian Analysis are used to compare different responses. Artificial Intelligence and machine learning techniques produce more sophisticated analysis.

In the scenario of television viewing, culturally different reactions to the same televisual 'beat' can be explored. Such techniques can be used with more widely to understand more critically, audience response, that too on an international basis. A greater understanding of cultural differences in reactions to the same content could be the basis of trials for transnational television formats. There is a need for control or reference data for comparison.

On a trial basis, there would need to be confidential interviews with participants to validate the data collected. This would help standardise what the computer-generated responses mean. A video interview of the participants, consent willing, would also reveal viewer body language and help understand emotional reaction and cultural variation.

One of the most successful methods of facial expression identification, adopted in software is the Facial Action Coding System (FACS) developed in 1997 by P. Ekman. (Ekman and Rosenberg, 1997)

FACS provides a standard set of single action and more animated facial expression identifiers, a subtle

subset of which are applicable to the reaction to television programmes.

Semantic analysis of FACS determines a mathematical/statistical model of implicit emotional conditions, extracted from the raw data, and map this to the brain's neurological structure. Such an analysis allows for aggregation of facial expressions into a representative 'shorthand' for emotional response.

Implementation of mathematical/statistical models in systems has been achieved in a few different ways. A system using the i-motion platform has been developed at Pace University, NY, USA, primarily for facial recognition from cameras, but also for other types of sensor in relation to emotional response (Taggart, Dressler, Kumar, Khan, and Coppola, 2016).

Software developers at Dartmouth Hampshire, NH, USA have developed an open-sourced software toolkit (Py-feat) based on the Python programming language which can be implemented in conjunction with a camera and a computer-based data collection device (Cheong, Xie, Byrne and Chang, 2021).

Any system of facial recognition needs to be customised to the cultural norms relating to each audience across the world. Different cultures visually respond to dramatic triggers in television in different ways. The methodology of analysis, categorisation, and implementation, however, remains the same.

Body language

The expression 'Edge of your seat' in the English language, refers to the visible signs of emotional engagement in television drama. Seat placed sensors can detect movement and thus characterise this in relation to the timecoded point in a television broadcast (Kiforenko and Kraft, 2016).

Eye tracker

The movement of the eyes and pupil dilation can add to the understanding of emotional response, for instance in gauging any form of arousal in the viewer that might not be immediately recognisable from facial expressions or other means. Eye tracker technology, from a camera or another means for specific types of television shows such as those involving romance or non-scripted dating, might provide useful additional data, recognising the need for calibration of the devices for different eye conditions and cultural differences (Renshaw, Stevens and Denton, 2009).

Galvanic Skin Response

A possible indicator of emotional response is Galvanic Skin Response (GSR). Skin acidity held in sweat can indicate specific responses such as anxiety. It may provide a companion source of data alongside more comprehensive methods of emotional

state identification, in relation to the thriller and drama genres for instance. Skin electrical resistance needs calibrating for different skin types, body and environmental conditions (Sharma, Kacker and Sharma, 2016).

Digital Processing and Communications

The data from cameras and sensors will need to be transmitted to a central computer for statistical analysis, in a similar fashion to a Nielsen based system. However, the data received for processing should not be a direct copy of the data received at each home.

The amount raw data received from a sensor, particular a camera, is relatively large compared to that which is collected by a conventional Nielsen based system.

There are several reasons why such data cannot be transmitted as raw, to a remote processing computer.

1. The bandwidth of the internet communications system between a domestic household and a processing computer would be too small
2. A centralised processing computer would have to process data from thousands of sources, which may not be feasible in the short time scales needed to gauge real time viewership or overnight ratings
3. Data from sensor, in particular camera could be highly personal and private, which should not be held at home, let alone communicated to a central computer, and in addition subject to 'hacking', computer viruses and a whole host of unwanted intrusion.

The handling of data from sensors is proposed as follows:

1. The raw data from cameras and other sensors is only held in the dynamic memory of the device, long enough for processing to be carried out.
2. The key parameters of the data are immediately extracted, separated from personal identification details of the viewer, and then stored in a secure portion of the device, long enough for it to be successfully and reliably transmitted to a central processing computer.

The device needed to extract facial expression data could be prototyped on a platform such as a Raspberry Pi if accurate processing speed can be achieved within its typical dynamic memory allocations.

Proposed Prototype

A proposed prototype is to develop a facial expression platform, keyed to the timecodes in a television programme, and then the processing of the results for codification in a data format akin to the Nielsen system.

The data transmitted by each household could be an addition to the Nielsen system of transmission, as an encapsulated data element containing the Channel, user demographic, timecode and facial expression.

An Existing Product

A notable example of a system already developed has been implemented, mainly for application to assess 'eyes on engagement' in television advertising. The broadcasting platform *tvision* includes an addition for customers to volunteer in being assessed for their eyes on screen engagement, to see for instance if a viewer's interest is present for a particular product or the generic products in that class. (www.tvisioninsights.com)

Benefits for Screenwriters and Directors

Screenwriters and directors will benefit directly from the data gathered by such a system, to consider quantified results, in relation to the script and direction of television episodes. In this way, the quality and direction of scripted television shows can be improved and focused more easily.

The audience reaction to the structure of drama screenplay in terms of turning points, midpoints, dramatic and emotional conflict, revelations of mysteries and cliffhangers can be measured in terms of impact with quantification. (Coplan, 2006)

For comedy screenplays, reactions to characters, how well audiences empathise, the growth of interest, surprise, twists and turns and resolutions can be monitored.

The benefits of a quantifiable approach to creatives are such that there is less needed to conjecture audience reaction, prior to a commitment from producers.

Benefits to Producers, Broadcasters and Advertisers

A quantifiable approach to understanding audience reaction can reduce the risk of producers investing in projects that may not bring a return.

A detailed analysis of data received from the proposed system can pinpoint more clearly which genres of television programming are more successful, from a wider perspective than just considering the number of people tuned to the programme. In this way, a much more reliable set of criteria can be used by producers in the selection of pitches for new shows. Within a particular genre, such as a game show or thriller, the aspects of a format or script can be analysed to see what the audience reaction might be. In this way, there is a greater likelihood that future productions will not only secure higher ratings but have higher visibility in social media.

Broadcasters can make smarter investments into their schedules or streaming portfolio by garnering the real engagement to television shows they make available, the way these shows are marketed and the tailoring of linear schedules to ensure maximum engagement.

Before a linear broadcaster schedules a new programme, with the possible trepidation of viewing figures, real engagement as determined by the proposed prototype will give the confidence that initial viewing figures, which may be subject to high publicity, will be sustained over an entire run of for instance six episodes. It should be recognised that whilst a highly marketed show may have good viewing figures out of curiosity, broadcasters do not want to be saddled with such a programme if the figures are destined to tail off by mid-season; a device like the one proposed can spot such problems earlier and avoid costly decreases in advertising revenue or losses of opportunity for other, cheaper shows or repeats to be screened in the same slot.

For live broadcasts, the prototype provides the opportunity for broadcasters to monitor in real time, viewer engagement, and make modifications on air to increase viewership. Such dynamic intervention, made possible by a greater understanding of immediate audience reaction, can save a show whilst it is still on air, and bring about a better public critique for the broadcaster.

Advertising slots, social media engagement, the once referred to 'water cooler moment' in offices, trending TV shows can be gauged in quantifiable terms, related to emotional engagement. (Nabi, Stitt, Halford, Finnerty and Keli, 2009).

Catharsis and mental health

The proposed prototype can be used to better understand the effects of television on mental state and catharsis. It can answer the question as to whether television nowadays offers catharsis or if it reinforces the experiences of an increasingly vivid planet. Given the state of the world, it can reveal as to whether audiences in general need more emotionally comforting viewing or something more urgent and visceral.

Individual psychological conditions affected in different ways by television can also be assessed, for instance anxiety disorders. (Jahangir, Nawaz and Khan, 2014).

Conclusion

This paper in summary has made the case for directly assessing the emotional engagement of television audiences in real-time, using primarily camera-based technology and associated digital processing, and then transmit a confidentially assured, coded version of the information alongside the Nielsen data already collected by regulators from

demographically stratified audience samples. The benefits of such an enhancement to television ratings measurement were highlighted for stakeholders across the industry, as well as mental health specialists.

The technical proposal for such a system should be developed and trialed, using a suitable prototyping software/hardware platform to refine the design and more importantly ascertain as to whether the principle of measuring emotional engagement of television audiences brings for the expected benefits as discussed.

Bibliography

Born, Georgina. 2000. "Inside television: television studies and the sociology of culture." In *Screen 4 1 4 Winter 2000 Reports and debates*.

Bourdon Jérôme and Méadel Cécile. 2014. "*Television audiences across the world: Deconstructing the ratings machine*". Palgrave MacMillan, London.

Buzeta, Cristian and Moyano, Patricio. 2013. "The Television Audience Measurement in the Digital Age" In *Cuadernos.Info*. 33(33):53-62.

Kelly, John P. 2017. "Television by Numbers: The Challenges of Audience Measurement in the Age of Big data." In *Convergence: the international journal of research new media technologies*.

25: 1,113-132.

Nielsen, Arthur C. 1945. "Two years of operation of the audiometer and the Nielsen Radio Index." in *Journal of Marketing* 9:3:239-255.

Cheong, Jin Hyun, Xie, Tiankang, Byrne, Sophie and Chang, Luke J. 2021. *Python Facial Expression Analysis Toolbox*. University of Dartmouth, NH, USA.

Coplan, Amy. 2006. "Caring about Characters: Three Determinants of Emotional Engagement". In *Philosophy of Film and Film Theory*, 1-19.

Ekman, Paul, E and Rosenberg, Erika L. 1997. "What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)" Oxford University Press, UK.

Jahangir, Syeda F, Nawaz, Nazia and Khan Nayab. 2014. "Effects of Media (Television) on Mental Health." In *FWU Journal of Social Sciences*, Summer 2014, Vol.8, No.1, 97-107.

Kiforenko, Lilita and Kraft, Dirk. Emotion Recognition through Body Language using RGB-D Sensor. 2016. In *Proceedings of VISAPP 2016 - International Conference on Computer Vision Theory and Applications*, Rome, Italy.

Nabi, Robin, L, Stitt, Carmen E, Halford, Jeff and Finnerty, Keli, L. 2009. "Emotional and Cognitive Predictors of the Enjoyment of Reality-Based and Fictional Television Programming: An Elaboration of the Uses and Gratifications Perspective." In *Media Psychology*, Vol 8: 4. 421-447.

Renshaw, Tony, Stevens, Richard and Denton, Paul D. 2009. "Towards Understanding Engagement in Games: An Eye-Tracking Study." In *On the Horizon*. Vol.17:4.

Sharma, Mahima, Kacker, Sudhanshu and Sharma, Mohit. 2016. "A Brief Introduction and Review on Galvanic Skin Response Mahima Sharma* in *International Journal of Medical Research Professionals*, 2(5); 254-57.

TAM Ireland, 2017. *Understanding TV Data*. The Nielsen Group.

Taggart, Richard W, Dressler, Michael, Kumar Poonam, Khan Shahroze and Coppola, Jean F. 2016. "Determining Emotions via Facial Expression Analysis Software" Richard W. Taggart, Michael Dressler, Poonam Kumar, Shahroze Khan, Jean F. Coppola. In *Proceedings of Student-Faculty Research Day, CSIS*, Pace University, May 6th, 2016.

Tracey Michael and Morrison David. 1979. "Mary Whitehouse." In *Whitehouse*. *Communications and Culture*. Palgrave Macmillan, London.