

360 CINEMATIC LITERACY: A CASE STUDY

Peter J Passmore¹, Maxine Glancy², Adam Philpot¹ and Bob Fields¹

¹Middlesex University, UK and ²BBC Research and Development, UK

ABSTRACT

360 degree film making necessitates a new language for storytelling. We investigate this issue from the point of view of the user, inferring 360 literacy from what users say about their viewing experiences. The case study is based on material from two user studies on a 360 video profile of an artist. Interviews were analysed using thematic analysis to understand how users made sense of the video. The sense of presence had a strong impact on the experience, while the ability to look around meant new skills had to be developed to try to make sense of 360 video. Viewers had most to say about a few particular shots, and some themes of note emerge: such as being in unusual places, certainty about what should be attended to and focus points, switches between first and third person views, and close-ups and interest.

INTRODUCTION

It is said that 360 film making requires a new language to be developed for film makers to tell stories. Recently wisdom amassed over the last few years has been embodied in guidelines or recommendations, for example, from Jaunt, Google and Vimeo and others. However not many empirical studies have been conducted to ratify or challenge these guidelines. Some guidelines concerning lower level film language elements, such as, about shots, transitions, and camera placement, are relatively easy to assess, and are partially addressed in this study. Other guidelines concerning the philosophy of 360 film making, and are more arbitrary. For example Google's Jessica Brillhart considers that there should be action going on all around the viewer, for which she developed the concept of Probabilistic Experiential Editing (<http://filmmakermagazine.com/96090-look-into-the-cut/#.WQwLCvnyupo>); whereas Jaunt generally recommend the main action should usually occur within 150 degrees in front of the viewer. The video used in the case study follows the latter approach.

In the following sections consideration is given to the most prevalent form of editing, continuity editing, which viewers are probably most literate in understanding, which has distinct rules about how to produce film that viewers will easily understand. Understanding of film is then located in a simple model which allows both viewer experience and sensemaking of film to be highlighted. Guidelines for 360 film making are briefly reviewed and this is then followed by a case study in which user experience is analysed with reference to viewer understanding, leading to the identification of a number of emergent themes that characterise user experience, and elements of 360 literacy.

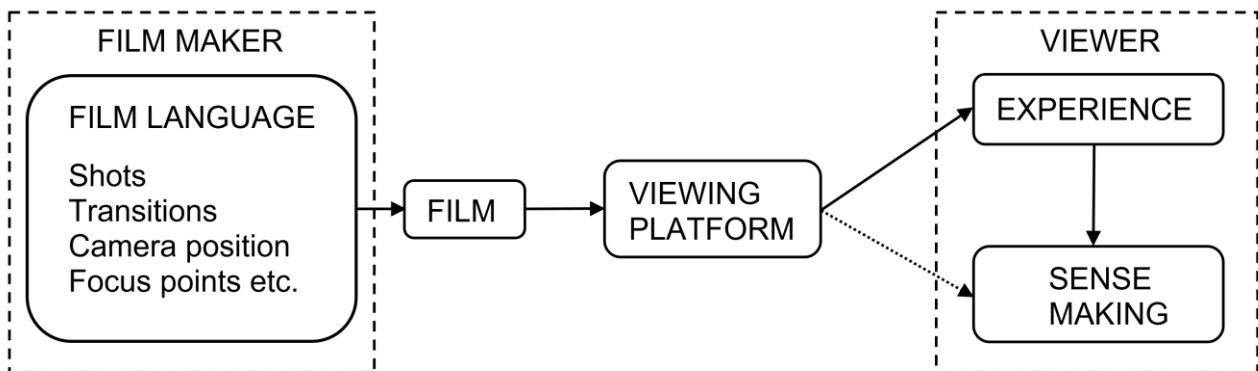


Figure 1 - A Model of Film Sensemaking.

A Model of Film Sensemaking.

A simple language based model of film sense making is shown in Figure 1. In this view the film maker uses the language of film making, which is composed of lower syntactic elements, such as shots, transitions and sequences etc., in order to tell a story which is embodied in a film. The user then views this via a viewing platform, e.g. a TV or virtual reality Head Mounted Display (HMD), leading to an experience they need to make sense of to understand the story. For experienced literate viewers, we may consider that experiential factors require minimal attention and that viewers more or less directly make sense of the film (hence the dotted line from viewing platform to sensemaking).

People ‘make sense’ of the situation they are in as a part of the routine activities they are involved in. Theories of sensemaking (e.g. Klein et al (1)) could be usefully applied to analysing understanding in film, as they emphasise the cognitive tasks that the user actively engages in, in order to create meaning. With traditional viewing, the user is fairly limited in what they can do to actively seek information to aid sensemaking, for example generating and testing hypotheses, and guiding their attention to look for cues in the frame. However, in 360 video this capacity is enhanced by the ability to look around, and as we move into more interactive narrative experiences, this capacity will only increase.

Continuity Editing.

Film making and TV production is dominated by continuity (often referred to as ‘invisible’) editing, which has a set of rules about how shots are composed and edited, the aim of which is to convey narrative by producing an easy to watch, seamless viewing experience. Viewers of such film could give a good account of the story they viewed, without necessarily being able to recount much detail of how shots were composed and edited. Continuity editing aims, by careful composition and editing of a sequence of shots, to direct viewer attention to consistency of story across time and space. Some rules of continuity editing apply to 360 video, such as continuity of sound over shots. However, many others, such as ‘shot/reverse shot’, do not.

360 Video Guidelines.

There is general consensus among various guidelines for 360 video on many points (see <http://www.bbc.co.uk/rd/topics/interfaces> for a review), which are relevant in this case study, such as for camera position, type of transitions, and shot timing. For example there is agreement that care must be taken with moving shots as they can induce motion sickness in 360. In terms of camera height it is generally agreed that the camera should be at roughly head height. Lower shots can make people feel shorter, and floor level shots could make viewers feel embedded in the floor, higher shots can induce vertigo. Higher aerial shots, such as drone shots, can be tolerated if they are stabilised and motion is smooth. Long distance shots do not work well due to the low resolution of current recording and playback resolutions. Close-up shots increase the scale of the subject, and this is particularly powerful when viewed in a HMD. The Jaunt guidelines quote the advice from Oculus that a minimum distance of 0.75 meters be used, a guideline that is probably designed for stereo viewing. On timing it is generally advised to give around at least 20 seconds per shot for the viewer to orientate to the new scene, unlike traditional TV, where shots can be much shorter. In general, the most recommended transition is the fade to and from black or blink transition. Jump cuts can be more difficult for viewers especially if transitioning between different spaces. Although it is not explicit, these guidelines address both literacy and experiential factors, trying to simplify the language while providing a good experience.

CASE STUDY

The study produced a new analysis by re-analysing data from two previous studies which used semi-structured interviews and thematic analysis to identify themes in users responses. The studies were based on viewers watching a beta version of a 360 video: *The Resistance of Honey* (<https://www.youtube.com/watch?v=t6u3opMTCV4>), which is a nine minute profile of an artist who combines his interests in bee keeping and electronic music. Over the two studies, subjects watched the video on a HMD and either on a static monitor, a hand held phone (Passmore et al (2)), or in a CAVE-like system (Philpot et al (3)). In both cases the order of viewing condition was alternated to counter order effects. In the first study eight subjects viewed video on the HMD first, and eight viewed on the screen or phone first. In the second study nine viewed on the HMD first and seven on the CAVE-like system first. In the current study we primarily consider the seventeen viewers who viewed with the HMD first.

Subjects first watched a short 360 video to acclimatise to the task. They then watched the video on their first viewing platform and they were then interviewed. The questions asked viewers generally about their experience, were the same for each participant, and allowed expansion on replies of interest (2). The participants were video recorded and observed throughout the study.

Thematic Analysis

The thematic analysis methodology of Braun and Clarke (4) was used for the analysis of the transcripts of interviews collected during the study. The transcripts were coded using open coding around relevant comments to identify emergent themes. In this study the interviews were re-analysed with a view to understand how directorial devices affect users experience in making sense of, and understanding 360 video. As the focus of this study was language

and literacy the emerging themes largely focussed on the structure of the video, with themes emerging about scenes, camera position, cuts, and transitions. Viewer reaction was inferred by the frequency and range of comments they made about individual items and also from their lack of comments on others.

The video was shot largely at locations in a park. The hive was inside a small shed in the park, and the hive entrance was on the outside of the shed. One shot was taken inside a simulated hive. One shot was taken of the Beeman standing in a polytunnel. Other shots were taken in the Beeman's studio room. In (2) user experience of viewing 360 footage was compared on a static screen, a phone, or a HMD, and user experience was profiled for each viewing platform on the basis of seven emergent themes: presence, certainty (about the viewing experience), concentration on story, attention, engagement, comfort and social ease.

RESULTS

The two themes that are most prominent when considering users comments from a literacy perspective are presence and the users' ability to look around, which is not surprising as these two elements speak directly to the difference between regular film and 360 viewing. Presence is the strong experiential factor that makes people feel they are actually present in the scene, at the point of view of the camera, compared to traditional viewing where the user's point of view is some distance away from the screen, upon which is displayed the cameras point of view. The ability to look around is linked to the users' certainty about what they should be looking at, and their attention to, and concentration on, story. It is related to their ability to make sense of the situation they are in.

Analysis of Interviews

Interviews were analysed to infer user experience and understanding. Although the original prompt questions were not specifically designed to look at literacy, much could still be concluded from viewer transcripts. The first point to note is that the 360 video (probably in common with most 360 videos) followed closely the tradition of continuity editing in sequentially portraying a number of scenes linked in both space and time. Secondly, the sound track, consisting of music, sound effects and intermittent narration by the Beeman, was continuous over the piece, uniting the shots. The piece was shot with the Beeman always in the front view, and he was generally the main moving object. Moreover, the narration was often used to steer viewer attention.

Viewers apparently applied their existent filmic literacy to understand the piece. Their comments indicated that they understood without question, both temporal and spatial aspects of the piece. For example, that the piece was about one individual, who was the same individual appearing from shot to shot, and that he was the source of the narration, or that the shots depicted a sequence over time, was never questioned, and was implicit in viewers' discussion. Their understanding of the general geography of the scenes was also good. For example, viewers managed to work out that the Beeman was in the same shed that he had earlier been standing outside of. Points where the viewers were confused about what was going on, were the same points you would expect them to be confused, if the film were viewed traditionally.

Presence and the ability to look around both mediated and interfered with viewers' ability to make sense of the video. For ordinary TV viewing, the experiential features are different but they do not impact in the same way on sensemaking - there is not the same sense of presence because of the way the film is viewed, and there is not the same ability to look around the film world. One consequence is that viewers find it harder to make sense in 360 because of lack of certainty in what should be attended to and fear of missing out. Having a lot to view can disperse concentration such that viewing is not just about receiving narrative but also about guessing how to direct attention. Viewers may also look around out of curiosity and the need to establish where they are, or could be affected by the context, rather than just focussing on story. The context can help to give focus, for example in football, the focus is generally just tracking the ball, but in other contexts it is less clear, forcing the viewer to try and decide what should be attended to, before they can deal with making sense of the story. Knowing what to attend to thus emerges as a key component of 360 literacy because whereas previously narrative components could be served to order, through framing and zooming etc., now the viewer may have to play the role of the director to some extent, to frame the appropriate shots at the right time for themselves, to get the correct message. Thus there are two main aspects to viewing in 360: dealing with the experience and making sense of the story.

In the next section we analyse specific film elements from the video, noting viewer experience and relating this to literacy and the guidelines. In the subsequent section we enumerate the emergent themes arising from the analysis of the data relating to literacy, before enumerating viewer experience of 360 video in Figure 2.

Film conventions and guidelines.

Framing the shot in traditional video is achieved by a combination of camera placement, zoom and focus. In 360 film making there is generally no zoom or focus available, so that just leaves camera placement, along with devices to try and get the viewer to attend to the content intended by the director. We found camera placement to be critical in determining how the viewer feels in the scene, both the height of the camera and the distance to the subject.

Establishing shot. An establishing wide shot followed by a closer shot is a standard continuity editing technique at the beginning of a sequence. The video started with an opening wide shot of the Beeman walking to his shed and entering it, followed by a closer shot of him in the shed, which worked well, and viewers had no issues following it.

At head height. The general advice is that the camera should be at head height, and we found that viewers either commented positively when this was the case in the video, or they did not comment, suggesting it did not cause them any issues, agreeing with advice from guidelines.

Above head height. For shots inside the Beeman's shed the camera was placed above the Beeman's head about a meter away. As predicted by the guidelines, some viewers complained about the shot being too high, or making them feel vertigo. This also had the effect of accentuating the fact that the viewer has no body in the scene, because they have to look down, which can be disorientating.

A couple of shots were taken in the studio from above head height, near to a wall, of the seated Beeman in his studio. Some viewers commented on feeling like 'a fly on the wall'.

Below head height. One shot, of the mouth of the hive outside the shed, was about 60cm high. As predicted by the guidelines, some viewers said this made them feel small

Close-up shots. The Jaunt guidelines recommend not to use close-ups in general, citing the Oculus guidelines that objects should not be closer than 0.75 m. Such advice is probably meant for situations where stereo 360 video is viewed, but it is quite possible to get closer in monoscopic 360 without apparent eye strain.

Contrary to the guidelines, a number of shots in were taken in close-up, usually 10cm or less from some objects in the scene, leading to the greatest number of comments. One shot simulated being in a hive, with a viewpoint about 10cm away from bees walking on honeycomb all round. Users variously described this as interesting, to get a view from inside a hive; as being too close; and some users said it made them felt free to look around because there was not one obvious focal point.

The low shot at the hive mouth, (viewing distance about 10cms), some users found interesting because they could get a novel view, and be close to bees without being stung.

In one shot, the Beeman was seated at a table in his studio manipulating electronic boards with wires, components, potentiometers, and the like attached, making music, with the camera placed between his face and hands at about 10cms viewing distance. This shot received the most comments from viewers, who generally found it interesting to see what the Beeman was doing with his hands. Some viewers described this as making them feel small (e.g. two inches high), or awkward (sitting on the table), or in an odd pose ("I'd be sitting in his lap or something"). One viewer said she felt that she actually was the Beeman manipulating his electronics. There were some similarly close-up shots (less than 10cms) of bees on flowers for which a couple of viewers mentioned feeling bee sized. It seems that the sense of presence, and lack of embodiment, makes some users try to rationalise about their size in these close-up views – leading to a range of different perceptions. This effect was not apparent when viewers watched a screen based version of the video.

Unexplained shots. In two shots in the piece, the Beeman appears briefly with other beemen dressed in bee keeper suits, this is completely unexplained in the story. Viewers were confused by this and commented on it, for example saying it was spooky. Such comments are not limited to 360 viewing however.

Unexplained shot not commented on. One shot showed a speeded up stop frame like animation of five beekeepers moving around. No viewer at all commented on this incongruous shot, perhaps just interpreting it by the music and visuals as a music video type transition.

Transitions. We consider responses to three types of transitions used in the piece. As predicted by guidelines fade to black transitions appeared to work well, viewers did not comment on them. Jump cuts were used, and they caused some negative comments from some viewers who found them too jarring, however many viewers did not comment on this, perhaps suggesting they can be tolerated. One viewer commented, that he was confused by a jump cut between two shots that were at different positions outside, as he did not at first realise the camera position had changed. Perhaps a fade to black transition would be

better here. One transition between two scenes was based on matching on action. The Beeman was holding up a bee comb with some microphones attached, scanning from right to left, first in a polytunnel, then outside in the garden. This transition worked well only if viewers were focused on the Beeman.

Emerging themes

We identified the following emergent themes in relation to viewer literacy and viewer experience. The first four relate to presence and the second two to looking around.

Being in Unusual Places. Many viewers commented positively about being in, rather than seeing unusual places.

Changes in Distance or Scale. Some viewers commented on a change in perceived size within the scene due to camera placement.

Differences between third and first person view of the character in-world. Viewers commented on a number of perspectives: like being a guest in the Beeman's studio; feeling they were seeing the Beeman's perspective, feeling small, or like a fly on the wall.

Close-ups and Interest. Close-up shots seemed to lead to an increase in viewer interest, but also sometimes changes in perceived size.

Following the action, knowing where to look. There were some instances where the viewer was not sure of where they should be looking.

Multiple and Single focus points. Some scene viewers found that multiple points of focus, or identifying a single one, (e.g. the Beeman) gave them confidence to look around.

Having identified the emergent themes, we can now detail how experience and sensemaking (i.e. the right hand side of Figure 1) play out for 360 in Figure 2. The experience component is largely dominated by presence, while the sensemaking component contains a restricted version of classic literacy (for example some elements such as framing, are not available in 360), augmented by the new literacy elements, largely relating to looking around. Individual 360 films have experimented with custom cues to sensemaking, and no doubt some of these may become new literacy conventions over time.

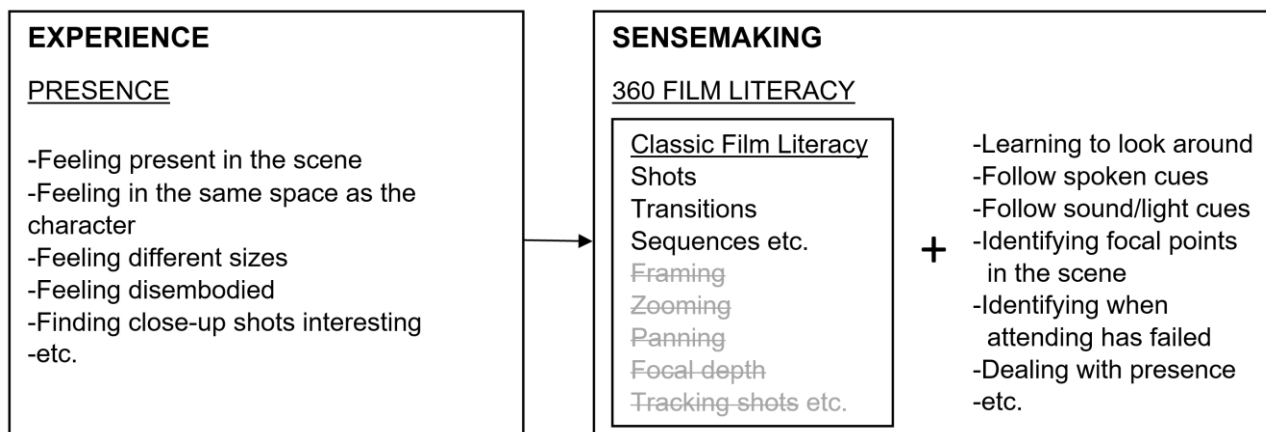


Figure 2 – User Experience in the Sensemaking Model for 360 HMD Viewing.

DISCUSSION

In general, the results of this study agree with the consensus guidelines about basic film language elements such as camera placement and use of transitions. A notable exception is the level of interest shown in close-up shots, along with possible accompanying feelings around perception of placement or size of self in the scene.

Viewers clearly seem to understand elements of the video from a continuity editing point of view, and this is perhaps not surprising given the general level of viewer filmic literacy in the sample population. The two elements that were really different from traditional viewing were the sense of presence and the ability to look around. The sense of presence made many viewers feel they were actually in the scene, giving a visceral experiential component to the activity not encountered in traditional viewing. The sense of presence was found to both mediate and interfere with understanding the story.

The ability to look around left viewers uncertain about how to read the 360 scene. However, they soon developed strategies which involved such things as scanning for focal points of interest, and making judgements about when it was OK to look around, as shown in Figure 2.

The simple sensemaking model of filmic literacy (Figure 1) is general and applies to traditional and 360 filmic literacy. In considering 360 literacy a key aspect is that the experiential components are much stronger than for traditional viewing. It could be said that a language based model for creating 360 experience is inadequate, as it does not cover the experiential components of 360 viewing. Moving to 360 viewing takes traditional viewing a first step into the world of interaction, and the importance of designing the experiential aspect will only increase as experiences become more interactive as film, VR, and gaming continue to converge.

REFERENCES

1. Klein, G. A., Phillips, J. K., Rall, E. L., and Peluso, D. A. (2007) A Data-Frame Theory of Sensemaking. In Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making. pp. 113–155. Expertise: Research and Applications Series. Mahwah, NJ: Lawrence Erlbaum Associates.
2. Passmore, P. J., Glancy, M., Philpot, A., Roscoe, A., Wood, A., and Fields, B. (2016) Effects of Viewing Condition on User Experience of Panoramic Video. In ICAT-EGVE 2016 International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments (Reiners, D., Iwai, D., and Steinicke, F. eds), The Eurographics Association, 10.2312/egve.20161428
3. Philpot, A., Glancy, M., Passmore, P. J., Woods, A., and Fields, B. (2017) User Experience of Panoramic Video in CAVE-like and Head Mounted Display Viewing Conditions. In: ACM International Conference on Interactive Experiences for TV and Online Video 2017, pp 65-75. ACM.
4. Braun, V., and Clarke, V. (2006) Using thematic analysis in psychology. Qualitative Research in Psychology. 3, pp. 77–101.