Accurate Visual Saliency Modelling for 3D Video Initialisation Document

Student: Hamish Sams, Supervisor: Charith Abhayaratne

January 15, 2019

Abstract

This project is based around the research and design of computing saliency from 3D video accurately to predict areas of common attention by the HVS (Human Visual System). The algorithm should use many different models to create an overall more accurate result than each respectively. Development and implementation of saliency in different areas are not defined within the scope of this project.

Contents

1 Project Specification **1**
1.1 Context 1
1.2 Aims 1
1.3 Objectives 2

2 Literature Review **2**

3 Initial Plan **3**
3.1 Gantt Chart 4

4 Risks **5**
4.1 Risk Register 5
4.2 Risk Assessment 6

References **7**

1 Project Specification

1.1 Context

As explained in the abstract, the context of this project is to research and develop methods of saliency mapping specifically for 3D video, but will use spatial methods on top of temporal to improve accuracy. Emphasis is to be put on the saliency maps accuracy as the uses of the maps created are out of scope of this project. As saliency is a fairly new topic, especially with 3D video, research moves fast and therefore it is important to use state of the art information to keep this project up to date. Using the previous context the following specification is designed:

1.2 Aims

- Explore methods of state of the art salience modelling to predict/mimic the HVS.
- Research and develop systems of accurate video saliency.
- Review a range of spatial and temporal salience models using data to validate model accuracy.

1.3 Objectives

- Use spatial and temporal modelling to monitor moving saliency more accurately while predicting which model is more important in each scene.
- Research and implement the current state of the art saliency models to mimic the HVS.
- Use MVD (Multi-view video plus depth) algorithms to increase saliency accuracy with the use of a 3D camera.
- Use a bottom up approach for measuring involuntary saliency over a targeted top down algorithm.
- Use many different saliency models to improve upon each respectively.
- Express uses of salience modelling but not develop nor implement them.
- Use primary and secondary data to review models and their effectiveness.

2 Literature Review

The work in [1] is based on saliency maps to watermark video. It does this through global motion compensating and many other methods. This compensation method is discussed along with the equations and the proposed system is then quantitatively compared using work from [2]. This is a good report for developing a saliency map by implementing global motion compensation, but is certainly aimed at experts in the area rather than beginners. This work is useful after a basic system of saliency is setup and a motion compensation scheme is wanted. The referenced data set ([2]) appears as a good benchmark for testing algorithms.

In terms of video saliency [3] researches into MVD (Multi-view Video plus Depth). Here an emphasis is put on compression of the video using the saliency map. The footage is taken from [4] and is quantitatively compared using different methods. This is a much more beginner friendly report but lacks maths and definitive algorithms to implement depth detection. It explains depth which will be used in my paper well but the compression sections are out of scope completely.

For a more general view of saliency for 3D video [5] is good. Lots of different methods are explained for creating a 3D saliency map. Here different methods of creating a map are explained and modeled together along with explanation of test design. Methods of saliency maps are also briefly touched upon with reference to [6] [7] [8] [9] This literature explains well all different saliency methods with corresponding maths but not much explanation for implementation. The data is comprehensive and shows easily the better results, however the literature is rather bulky compared to other articles. Most of the work here is useful towards my project especially the explanation of the filters used.

[10] is a very basic article showing the basics of these concepts early on from some of the big names in saliency at the time. Unfortunately, most of this information is out of date and is not useful for state of the art designs. Like previously [11] is simply running out of date, this paper uses more data and less concepts than [10] but once again this is not state of the art.

[12] works on displaying and explaining saliency maps in 3D depth video. A lot of the algorithm is based on "Visual comfort" and explains the meaning and implementation. This article explains exactly what saliency methods it uses and shows the maths to implement such a design. The data here is taken from [5] and is used to create a quantitative comparison between schemes and the proposed method. This article uses the exact same recording device and with similar aims. The focus here is on visual comfort and will likely be a good source of information for such methodology.

The work in [13] focuses less on an algorithm to predict saliency but instead on working out what is and isn't eye catching. It goes about this by measuring eye focus to different colors, textures and motion. This report shows well between what colors are best or what motion is most likely to catch vision but doesn't put much data comparing if motion or color is more important to a viewer. This report allows for saliency algorithms to be based off color, texture and motion easily despite not implementing it directly in the literature.

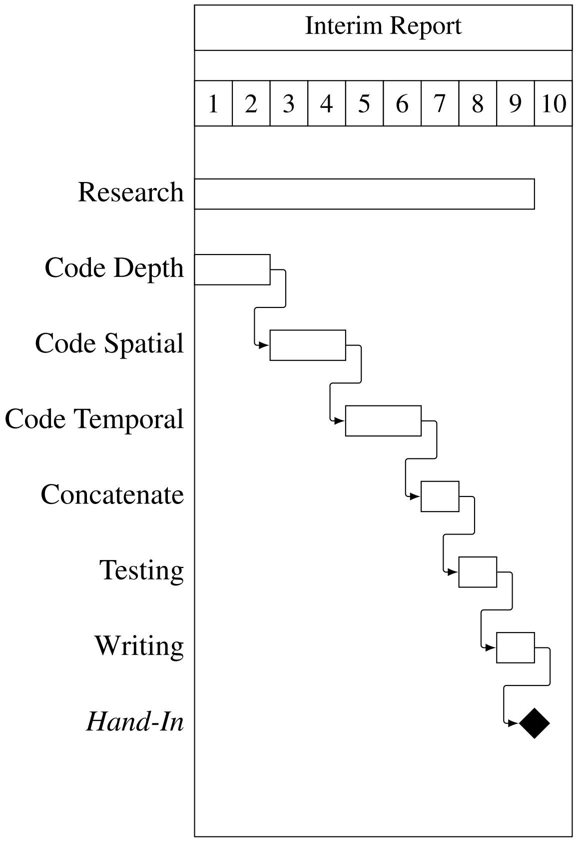
A more mathematical view of a contrast based saliency is shown in [14]. Here a salience map is created by many different methods including dot product and contrast based methods. Here maths is shows clearly allowing for implementation in other systems. Many methods are used but aren't compared to others very clearly. This is a possible method for measuring saliency in my predicted system.

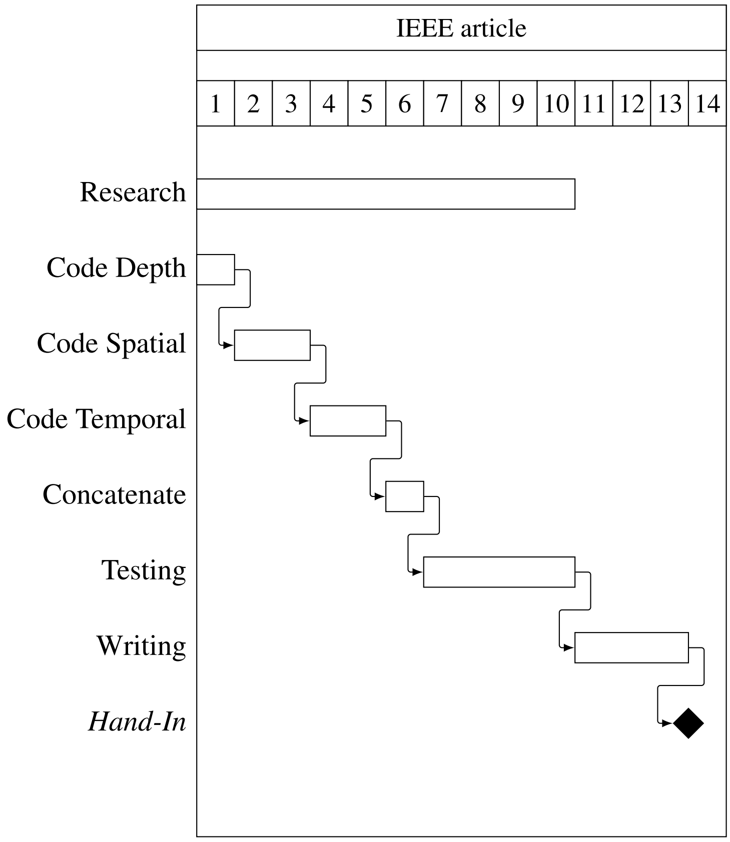
3 Initial Plan

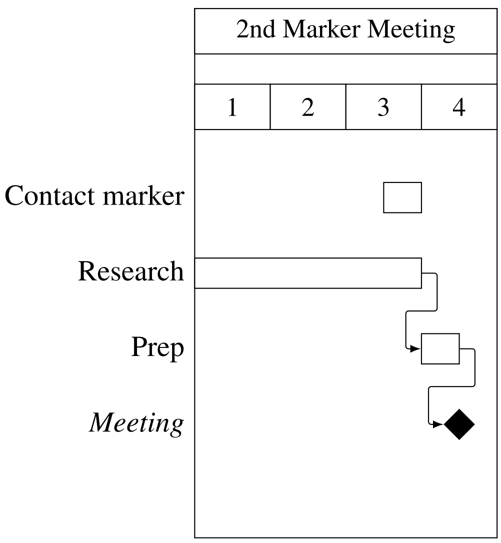
Key Dates	Description
19-10-18	Project Initialization Document
09-11-18	2nd Marker Meeting
14-01-19	Interim technical report.
08-02-19	Second Marker Viva
Unknown	Symposium
03-05-19	IEEE style Article
10-05-19	Public Engagement video

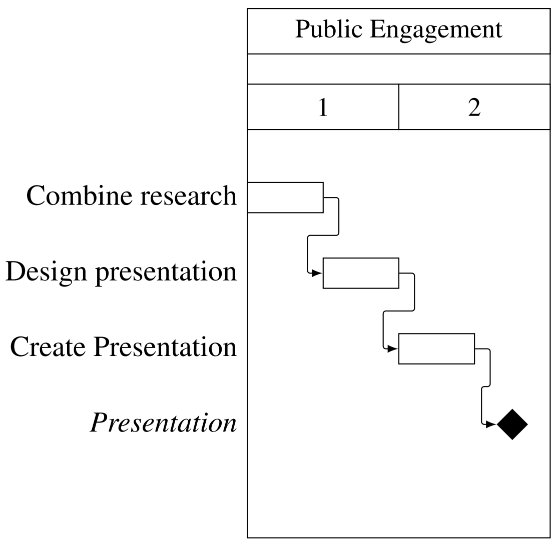
Table 1: Preliminary Risk Register

Currently the Symposium has no date attached and is assumed to be the same date and conjoined with the public engagement video as the public engagement video is most commonly a recording of the symposium. On top of that the Initialization document will not be discussed in further detail as it is already completed.

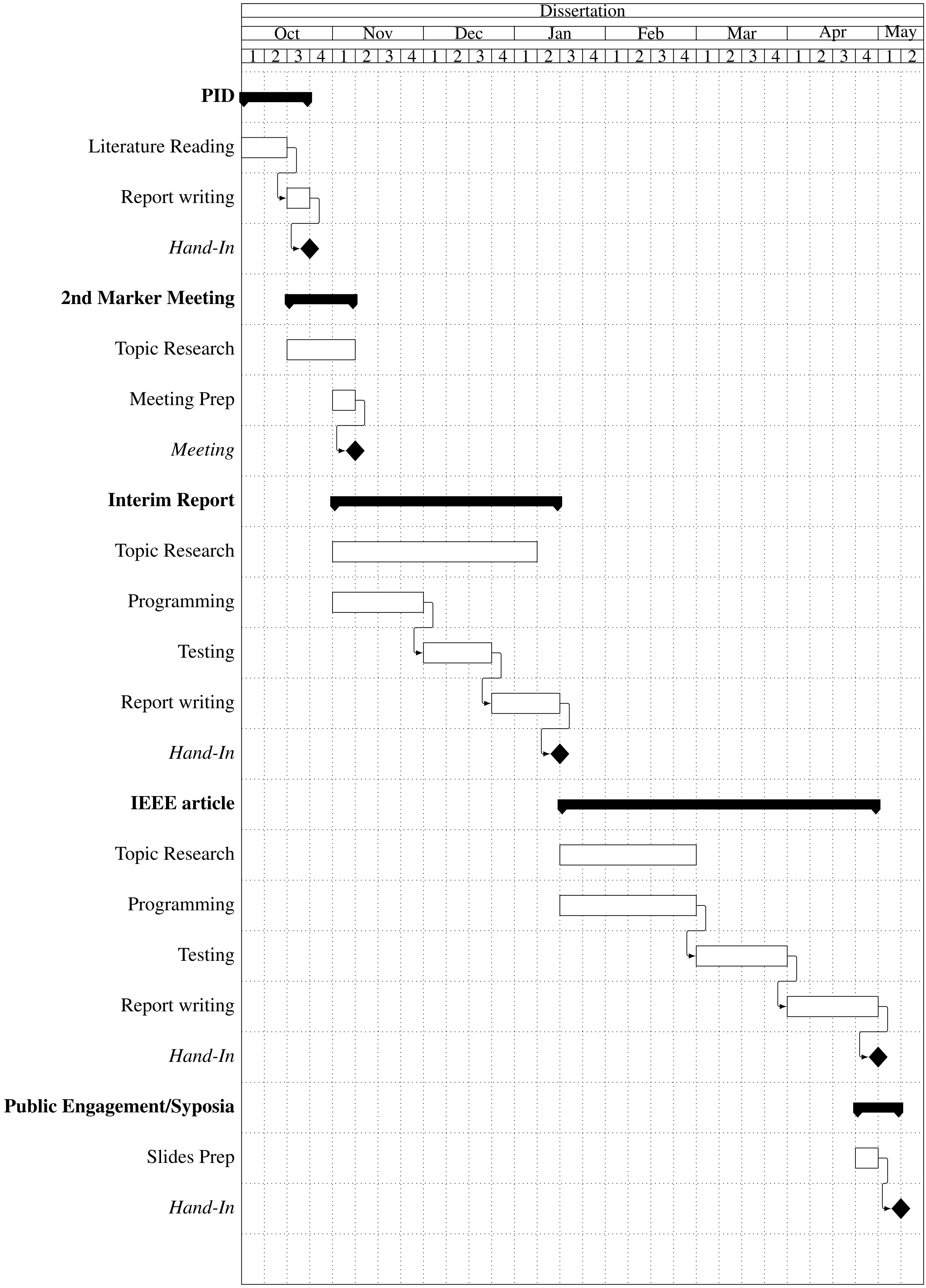








3.1 Gantt Chart



4 Risks

4.1 Risk Register

Hazard summary	Existing Measures	Likelihood	Severity	Risk Rating	Additional Measures	Residual Likelihood	Residual Severity	Residual risk rating
Data Loss	Data stored on a university computer is automatically backed up infrequently and allows for rough document restoration.	4	4	16	Use backup software at home and university such as git in conjunction with overleaf online to allow for transfer between desktops as-well as a backup solution.	2	4	8
Over Budget	University has to approve all purchases to prevent going over budget. However this doesn't prevent bad planning	1	2	2	N/A	N/A	N/A	N/A
Time issues	Gantt chart created to track progress and key dates.	3	2	6	N/A	N/A	N/A	N/A
Software licencing or errors	Sheffield CiCS provide common software free of charge for home PC's. University PCs have software pre-installed and automatically licensed. University has a tech team that can look into further issued with software.	3	3	9	Keep to gantt chart to prevent software errors at bad times increasing stress and pressure.	3	2	6
Hardware issues	Contact the most qualified person to fix the hardware issue.	2	2	4	N/A	N/A	N/A	N/A
Health and safety lost time injury	Risk assessment covers likely injuries and can be used to mitigate the risk. For large lost time a extenuating circumstances form may be handed in.	2	3	6	N/A	N/A	N/A	N/A

Risk Rating Reference

	Likelihood				
Severity	**1**	**2**	**3**	**4**	**5**
1	1	2	3	4	5
2	2	4	6	8	10
3	3	6	9	12	15
4	4	8	12	16	20
5	5	10	15	20	25

Risk Rating	Explanation
1-5	No additional measures needed but can be implemented to reduce further risk.
6-12	Decide whether further measures need to implemented to lower risk rating.
15-25	Stop the corresponding task(s) immediately and seek to reduce risk.

Table 2: Preliminary Risk Register

4.2 Risk Assessment

Potential Hazard	Potential Harm	Existing Measures	Likelihood	Severity	Risk Rating	Additional Measures	Residual Likelihood	Residual Severity	Residual risk rating
Electricity	Electrocution could occur from bad electrical practices. Electrocution could effect a single user or even surrounding users	PAT testing is enforced on all mains equipment on university premises. Food and drink are banned in labs preventing spillages.	4	1-3	4-12	Avoid contact with other dissertation projects in labs to avoid low voltage (<30V) shocks. Only plugging in PAT tested and UK rated equipment while checking pre-plugged in cables for faults to protect from medium voltage (<300V) shocks.	1	1-3	1-3
Surroundings	In a general dissertation lab others may be doing anything and in so lots of small unknown risks exist. This could effect many or few people	All projects must have a risk assessment assigned. The room also has a risk assessment assigned for the risks possible inside.	3	2	6	Avoid interfering with others work especially if unattended. Read the room risk assessment and stay aware of the surroundings as they may change hour on hour	2	2	4
Sitting	Posture can cause back and general muscle issues when sat with the incorrect support for prolonged times.	Computer rooms are installed with adjustable chairs and some with standing desk areas.	3	3	9	Be sure to use a working adjustable chair. Stand often and take regular breaks/walks. Use effective peripheral devices. Report issues as soon as they start to occur and take longer or more frequent breaks.	2	2	4
Keyboard and Mouse	Long use of keyboard can lead to overuse syndrome/RSI(Repetitive strain Injury) and is common in computer users.	Adjustable chairs and methods of propping up keyboards to a comfortable level allow for strain free typing.	3	2	6	If a frequently used computer is uncomfortable equipment may be bought out of the budget to improve workspace.	2	2	4
Monitor	Eye strain can occur from over use of digital screens and bad viewing practice.	Monitors are placed at sensible distances with variable height and angle.	1	2	2	N/A	N/A	N/A	N/A
Heavy Lifting	Heavy lifting can cause back and muscle issues from strain but can also lead to an increased risk of slips and trips.	The building has lifts to avoid lifting on stairs with limited view but also has access to lifting trained porters with reasonable equipment that can carry heavy or awkward loads.	1	3-4	3-4	N/A	N/A	N/A	N/A
Loose wires and spillages	Badly placed wires and un-explained spillages lead to slips and trips that can hurt anybody in the area if un-treated	Room risk assessments explain the correct measures on discovery or treatment of a slips and trips hazard.	3	2	6	N/A	N/A	N/A	N/A
Moving objects	Moving objects used to video test sequences may get out of control and hit other room users or damage equipment.	None	4	2	8	Upon test sequence filming have another ready to catch any moving objects and announce to others in the area of possible hazard	2	2	4

Risk Rating Reference

Severity \ **Likelihood**	**1**	**2**	**3**	**4**	**5**
1	1	2	3	4	5
2	2	4	6	8	10
3	3	6	9	12	15
4	4	8	12	16	20
5	5	10	15	20	25

Risk Rating	Explanation
1-5	No additional measures needed but can be implemented to reduce further risk.
6-12	Decide whether further measures need to implemented to lower risk rating.
15-25	Stop the corresponding task(s) immediately and seek to reduce risk.

Table 3: Preliminary Risk Assessment

References

[1] M. Oakes, D. Bhowmik, and C. Abhayaratne, "Global motion compensated visual attention-based video watermarking," *Journal of Electronic Imaging*, vol. 25, no. 6, p. 061624, 2016.

[2] K. Fukuchi, K. Miyazato, A. Kimura, S. Takagi, and J. Yamato, "Saliency-based video segmentation with graph cuts and sequentially updated priors," in *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*. IEEE, 2009, pp. 638–641.

[3] N. U. M. Noor, H. A. Karim, N. A. M. Arif, M. H. L. Abdullah, and A. Sali, "Multi-view video plus depth representation with saliency depth video," in *Multimedia and Broadcasting (APMediaCast), 2015 Asia Pacific Conference on*. IEEE, 2015, pp. 1–6.

[4] J. Ohm, "Call for proposals on 3d video coding technology. iso," IEC JTC1/SC29/WG11, Doc, Tech. Rep., 2011.

[5] J. Wang, M. P. Da Silva, P. Le Callet, and V. Ricordel, "Computational model of stereoscopic 3d visual saliency," *IEEE Transactions on Image Processing*, vol. 22, no. 6, pp. 2151–2165, 2013.

[6] K.-h. Park and H. W. Park, "Region-of-interest coding based on set partitioning in hierarchical trees," *IEEE transactions on circuits and systems for video technology*, vol. 12, no. 2, pp. 106–113, 2002.

[7] V. Setlur, T. Lechner, M. Nienhaus, and B. Gooch, "Retargeting images and video for preserving information saliency," *IEEE Computer Graphics and Applications*, vol. 27, no. 5, 2007.

[8] K. Vu, K. A. Hua, and W. Tavanapong, "Image retrieval based on regions of interest," *IEEE Transactions on knowledge and data engineering*, vol. 15, no. 4, pp. 1045–1049, 2003.

[9] H. Liu and I. Heynderickx, "Visual attention in objective image quality assessment: Based on eye-tracking data," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 7, pp. 971–982, 2011.

[10] E. Niebur and C. Koch, "Control of selective visual attention: Modeling the" where" pathway," in *Advances in neural information processing systems*, 1996, pp. 802–808.

[11] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.

[12] Q. Jiang, F. Shao, G. Jiang, M. Yu, Z. Peng, and C. Yu, "A depth perception and visual comfort guided computational model for stereoscopic 3d visual saliency," *Signal Processing: Image Communication*, vol. 38, pp. 57–69, 2015.

[13] J. Hosseinkhani and C. Joslin, "Significance of bottom-up attributes in video saliency detection without cognitive bias," in *2018 IEEE 17th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)*. IEEE, 2018, pp. 606–613.

[14] Y. Zhang, F. Zhang, and L. Guo, "Salient object detection based on background model," in *2018 37th Chinese Control Conference (CCC)*. IEEE, 2018, pp. 9374–9378.