Unsupervised Saliency Detection in 3D video based on Multi-scale Segmentation and Refinement

Ping Zhang*, Pengyu Yan, Jiang Wu, Jingwen Liu, Fengcan Shen

***Abstract*—In this paper, we propose an unsupervised saliency objects extraction method for 3D video. The proposed framework consists of three main stages: (i) The input video frame is segmented into non-overlapping superpixels by combining both appearance and depth information at the input. Depth information can be used to improve the accuracy of segmentation in complex regions, *i.e.*, foreground object with similar appearance as background. A multi-scale segmentation scheme is also deployed by using different segmentation parameters to extract varying shapes of the foregrounds in each frame. (ii) The initial saliency score of each segmented superpixel in each scale is calculated via global contrast which is defined by appearance, depth, and motion cues from two consecutive frames. (iii) The initial saliency scores in each scale are refined by smoothing over a graph built by the spatial neighboring of all the superpixels in the frame. The final result is generated by fusing the saliency maps in all scales. The experiments on two widely-used datasets illustrate that our method outperforms state-of-the-art algorithms in terms of accuracy, robustness, and reliability.**

***Index Terms*—Saliency detection, Segmentation, Multi scale, Appearance, Motion, Depth, Graphical model**

I. Introduction

Human visual attention is most significant in human visual system, as it ensures us to detect the most important object in a complex scene.

In 1998, Itti [1] defined "saliency" as the area which is much different from the surrounding. This definition was broadly accepted by the research community. In [2], [3], the center-surround algorithm was enhanced using the KL distance of two feature histograms. Later, instead of using the part contrast, global contrast algorithm was designed to segment the saliency area [4]. Considering the feature dispersal effect, a "saliency filter" [5] was designed to calculate both the contrast and dispersal degree of color to determine the saliency area. In 2006. Harel et. al. [6] proposed a graphical model to represent the similarity between each part within the picture. Random Walk (RW) algorithm was adopted to give a score for each area. In 2013, Yang et. [7] used Manifold Ranking algorithm to distinguish the foreground and background areas.

With the development of the 3D measuring device like Kinect [8], depth information is naturally incorporated in the saliency detection model. Zhang et. [9] used appearance, depth, motion, illumination and direction to calculate bottom-up saliency in an image. In [10], depth is used to weigh the 2D saliency result. Based on that, Niu [11] constructed a depth weight curve by assuming that the comfort zone and popping out object are more salient than any other region or object. In [12], [13], instead using depth as a weight bias, the authors considered it as another feature and calculated the depth-only saliency map via global contrast. Finally, they combined the 2D saliency map with the depth saliency map and obtained 3D saliency result. Later, with the growing focus on video, motion and depth information is naturally utilized in saliency detection. In 2010, Zhang proposed the 3DV [9] method which fused the depth feature map with appearance and motion saliency map. In 2015, Lino [14] designed a saliency detection method which spatial, motion and depth saliency maps are obtained based on Itti's model, block matching and Difference of Gaussian (DoG) filter, separately. These results are fused to yield the final saliency map.

Since the success of convolutional neural networks (CNN) [15] in image classification during ImageNet2012 competition [16], deep learning was gradually utilized in saliency detection. Li et. [17] used CNN to extract multi-scale feature from image to obtain saliency result. In 2015, Zhao et. [18] calculated the saliency via high-level feature, which is obtained through CNN networks. Later, Xiang Wang et. [19] proposed Mask-RCNN model, which preserved a clearer edge compared with the former methods.

Although there are many proposed methods including deep-learning methods, they are limited in preserving a clear edges of the foreground object when considering for a wide range of 3D video. Besides, some methods need quantities of data and supervised training to achieve good performance.

In order to overcome the aforementioned challenge, as shown in Fig. 1, we proposed a multi-scale segmentation and refinement to make the detection method general and robust. In our paper, we upgraded SLIC algorithm [20] by adding the depth information into the clustering part. As a result, two regions which share a similar appearance but belong to different objects can be separated with well-preserved object edges. Multi-scale architecture is generated by setting different number of superpixels in each scale and then fusing the final result of each scale together in the last step. This architecture enables robustness and generality by combining the merit under each scale and fixing the problem of setting optimized superpixel number due to the continuously changed inputs. The initial saliency map are obtained via global contrast with appearance, depth, and motion information. Finally, the refinement based on graphical model is utilized to make the

This research is supported by the Science and Technology Planning Project of Sichuan Province, China (No. 2018GZ0166), National Natural Science Foundation of China (No. 61308102). (Corresponding author: Ping Zhang.)

The authors are with the School of Optoelectronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China (e-mail: pingzh@uestc.edu.cn, py.yan965@gmail.com, 1625523461@qq.com, 18702505013@163.com, robertshen@outlook.com)

saliency area smooth in interior and regular in shape.

II. Proposed Method

A. Segmentation with multi-scale

1) Upgraded SLIC Model: In this part, our aim is to decompose the input image into non-overlapping segments and preserve accurate object's shape and edge. While the traditional K-means algorithm in SLIC used appearance, *i.e.*, color, and distance to cluster pixels, we further add depth information with appropriate weight, in the proposed method. We set the depth feature distance as $D_d = \sqrt{(d_j - d_i)^2}$, where d_j and d_i represent the j^{th} and i^{th} superpixels' depth value, respectively. By adding the weight coefficient of depth feature distance as β, the complete feature distance evaluation coefficient becomes:

$$D' = (1-\beta)D_c + \beta D_d \tag{1}$$

$$D = \sqrt{(\frac{D'}{m})^2 + (\frac{D_S}{S})^2} \tag{2}$$

Here, D' is the weighted mean of color and depth feature distance, noted as D_c and D_d, respectively. D_S is the space Euclidean distance calculated by the coordinate of each pixel. m is a parameter related to the largest color distance in one image, with typical range of $[1, 40]$ and S is the expected number of superpixel. Throughout the experiment, β and m are set to 0.5 and 15, as they give the best segmentation performance.

2) Multi-scale Architecture: The number of superpixels is a parameter in image segmentation. Small number of superpixels reduces the algorithm's ability to preserve fine edges of an object, while large number of superpixels misses important corners. Furthermore, the number of optimal superpixels depends on the input images.

In order to increase the generality and robustness of the proposed algorithm, a Multi-scale approach is integrated with the SLIC algorithm. The first part of Fig. 1 shows the modified SLIC algorithm's output for a image with appearance and depth information and segmented to 200, 600 and 1000 superpixels. The feature value in each superpixel is the mean of all pixel's feature value in the superpixel area.

The remaining part of the model is calculated upon superpixel level and within each scale. At the last step, the results of all scales are fused using the algorithm described in Section II-D.

B. Initial saliency

According to visual perceptual research [1], contrast is the most important factor in visual saliency. In this part, two steps are proposed to calculate the initial saliency map via feature contrast.

First, we use CIE-Lab color space to represent the appearance of input image, since LAB color correlates better to perceptual of human eyes. Motion is extracted by the TV-L_1 Optical Flow [21], [22] and the two parameters of motion, v_x and v_y, can be obtained by solving the equation $I(x, y, t) = I(x + \mathrm{d}x, y + \mathrm{d}y, t + \mathrm{d}t)$ with the constraint of by $I_x v_x + I_y v_y + I_t = 0$. I is the luminance.

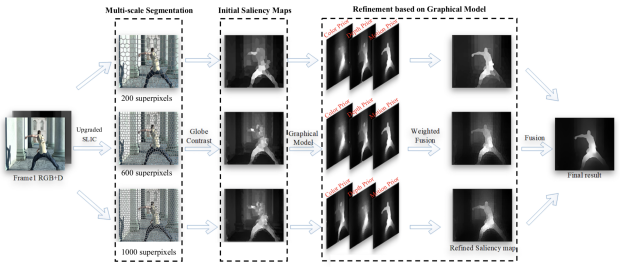


Fig. 1: Flowchart of the proposed method

Second, we use global contrast with color, depth, motion and spatial distance to generate the initial saliency maps. Since the images have been segmented by SLIC, feature contrast can be calculated on the superpixel level. We use arithmetic mean of color, motion and depth of all pixels in each superpixel area to represent the feature value of whole superpixel. The spatial distance is calculated among the barycenters of superpixels, and the feature distance is calculated by Euclidean distance:

$$d_C(R_j, R_i) = \sqrt{(L_{R_j} - L_{R_i})^2 + (a_{R_j} - a_{R_i})^2 + (b_{R_j} - b_{R_i})^2} \tag{3}$$

$$d_M(R_j, R_i) = \sqrt{(v_{x_{R_j}} - v_{x_{R_i}})^2 + (v_{y_{R_j}} - v_{y_{R_i}})^2} \tag{4}$$

$$d_D(R_j, R_i) = \sqrt{(D_{R_j} - D_{R_i})^2} \tag{5}$$

where R_i denotes the i^{th} superpixel, and L, a, b are the color values under CIE-Lab. d_C, d_M and d_D are the feature distance of color, motion and depth, respectively.

After the feature distance calculation, according to the visual mechanism of human where the closer the two superpixels are, the more influence they had for each other, we designed the weight coefficient based on spatial distance as follows:

$$\omega(R_j, R_i) = \exp(-\frac{d_S(R_j, R_i)}{\sigma}) \tag{6}$$

d_S represents the spatial distance among the barycenter of the superpixels. The value of the parameter σ is 0.6.

Using weight coefficient, we can obtain the score of each superpixel under each feature:

$$\begin{aligned} S_F(R_i) &= \sum_{R_j \in \Omega} s_F(R_j, R_i) \\ &= \sum_{R_j \in \Omega} \omega(R_j, R_i) \cdot d_F(R_j, R_i) \end{aligned} \tag{7}$$

where the notation "F" stands for either "C","M" or "D", *i.e.*, color, motion, and depth, correspondingly. s_F is "F" feature distance between two superpixels, which is weighted by spatial distance and $S_F(R_i)$ is the saliency score of R_i^{th} superpixel under "F" feature. Ω is the set of all the superpixels in one frame.

By normalizing all superpixels' score, the "F" feature saliency map is obtained. Then, we calculate the aggregation degree as the evaluation weight to fuse all feature saliency maps:

$$\mu_F = \frac{\sum_{R_i \in \Omega} \sqrt{(y_{R_i} - \overline{p_{y_F}})^2 + (x_{R_i} - \overline{p_{x_F}})^2} \cdot S_F(R_i)}{\sum_{R_i \in \Omega} S_F(R_i)} \tag{8}$$

$$S = \frac{\sum_{F\in\{C,M,D\}} 1/\mu_F \times S_F}{3} \tag{9}$$

Here, "F" means this equation is used to calculate value of "F" feature saliency map. (x_{R_i}, y_{R_i}) is the coordinate barycenter of the R_i^{th} superpixel, $(\overline{p_{x_F}}, \overline{p_{y_F}})$ is the barycenter of the saliency area calculated by saliency scores of all superpixels, and μ_F is the aggregation degree of each feature saliency map. $S_F(R_i)$ is the "F" feature saliency score of R_i^{th} superpixel. S is the initial saliency map.

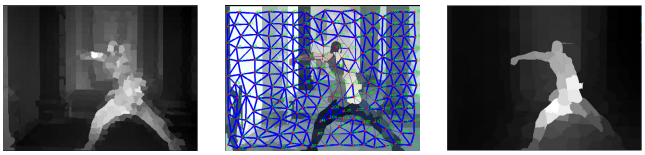
C. Saliency refinement based on Graphical Model

Initial saliency map obtained from global contrast has many artifacts at object's edge, as seen in Fig. 2(a). To regularize the saliency map and maintain the object's shape, a refinement method based on graphical model is proposed. A probability transition function [23] can reduce the large difference of saliency values between two near superpixels inside the saliency region and finally make the saliency area smoother and the edge between salient and non-salient area clearer.

First, upon initial saliency maps, we separately use motion, color and depth features to calculate the transition probability $P_{R_j,R_i,F}$:

$$P_{R_j,R_i,F} = \frac{W_{R_i,R_j,F}}{\sum_{R_j\in\Omega_{R_i}} W_{R_i,F}} \tag{10}$$

where $F \in \{C, M, D\}$, and $W_{R_i,R_j,F} = \exp(-\frac{d_F(R_j,R_i)}{\sigma})$. We observe that the more different the two superpixels feature is, the smaller the transition probability would be.



(a) Initial Saliency (b) Graphical Model (c) Optimized Saliency

Fig. 2: Optimization based on Graphical Model

Based on the transition probability, we draw the lines among the superpixels' barycenters and use the thickness of the lines to represent the value of transition probability. If the feature contrast of these two superpixels are small, the line would be thick, and vice versa (Fig. 2(b)). In Fig. 2(b), the rough outline of the objects could be presented by the thin lines, which reflect a big difference between the objects and the background. It is also the location where the line just cross the objects' edge.

Second, we use the transition probability as the weight to process the initial saliency maps:

$$U_F = \sum_{R_j,R_i\in\Omega} P_{R_j,R_i,F} \cdot s(R_j, R_i) \tag{11}$$

Here, $s(R_j, R_i)$ means the fused feature distance between two superpixels implied in the initial saliency maps and it can be calculated from eq. (7) and eq. (8) by $s(R_i, R_j) = \frac{1}{3}\sum_{F\in C,M,D} \frac{1}{\mu_F} s_F(R_j, R_i)$.

From this equation, we can gain the refined saliency map by each feature prior, for example U_M represents the refined saliency map by motion prior.

Third, after the normalization, the three refined saliency maps by three features priors would be fused by a simple way:

$$U = \frac{1}{3}\sum_{F\in\{C,M,D\}} U_F \tag{12}$$

The proposed fusion method yields refined saliency map with well-preserved object's shape and edge, as shown in Fig. 2(c).

Now, the refined saliency map under one scale is obtained. Similarly, refined saliency maps in the other two scales are calculated similarly.

D. Fusion of Multi-scale Saliency Maps

After section II-C, there are three refined saliency maps under three scales. We name each saliency map as U_n, *e.g.*, U_1 means the saliency map in the first scale.

$$U_A = \frac{1}{3}\sum_{n=1}^{3} U_n + \prod_{n=1}^{3} U_n \tag{13}$$

Considering the slight difference among saliency maps of all scales, we developed a fusing method by combining an average sum and multiplication. In the sum part, we average it to make the saliency value of each superpixel still between $[0, 1]$. In the multiplication part, we multiple saliency maps together and then the common saliency areas in all saliency maps would be enlarged, and the uncommon saliency area or none saliency area in all saliency maps would be greatly diminished (because it is the multiply between 0 and 1). In multiplication part, the saliency value of each superpixel is among $[0, 1]$ as well. Then we add these two parts together to improve the performance of the final refused result. The normalization: $U(R_i) = [U_A(R_i) - min(U_A)]/[max(U_A) - min(U_A)]$ is also used and U represents the final saliency result of the whole proposal.

III. Experiments and Evaluation

In this section, two steps are adapted to evaluate our experiment with two public 3D video sequence datasets, which are 3D-HEVC [24] and NAMA3DS1 [25]. The first step is to vindicate every part of our proposal is rational and necessary. The second step is to compare our method with other state-of-the-art methods.

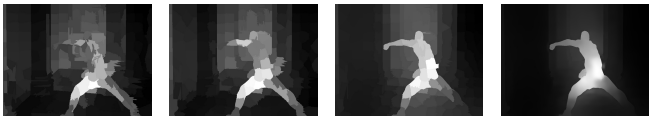


Fig. 3: the images from left to right are:initial saliency map without depth, initial saliency map with depth, refining based on graphical model, result of multi-scale approach

A. Vindication of Proposal

From Fig. 3, by comparing the result of each steps of our approach, we can observe the importance of each step in our proposal. As observed in the first two sub-images in Fig. 3, adding depth could help to improve the performance of the

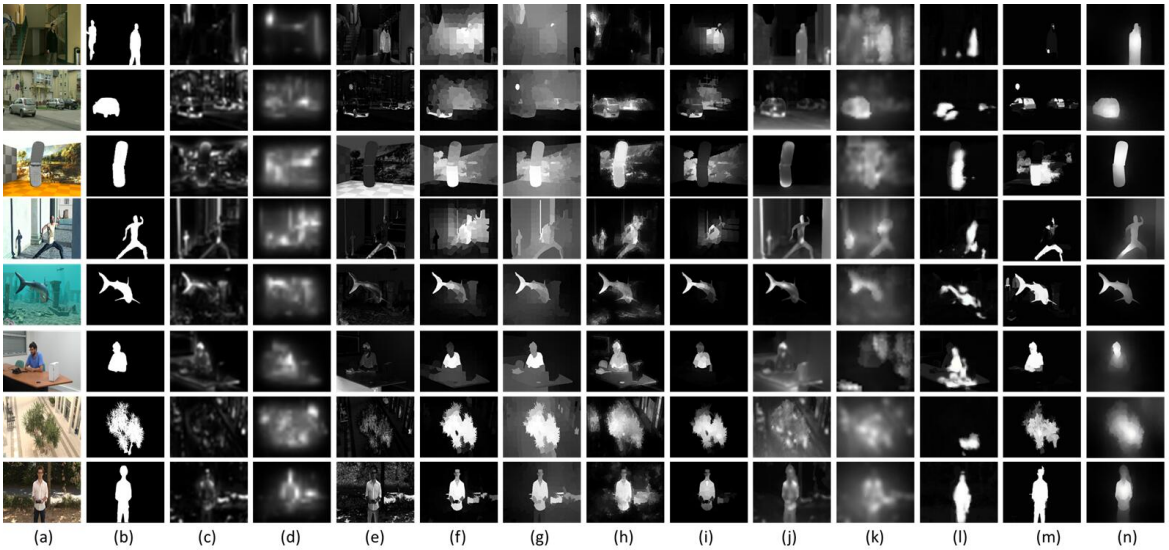


Fig. 4: Example of saliency results of differen methods. (a) Original frame. (b) Ground Truth. (c)-(m) Results of ITTI [1], GBVS [23], FT [26], MR [7], SPL [27], WANG2DV [28], RGBD [13], ZHANG3DV [9], LINO3DV [14], RFCN [29] and MDF [17]. (n) Our proposal

TABLE I: Measurement

Models	ITTI	GBVS	FT	MR	SPL	WANG2DV	RGBD	ZHANG3DV	LINO3DV	RFCN*	MDF*	**Proposed**
Precision	0.2549	0.3255	0.4123	0.4255	0.5536	0.4197	0.5416	0.4308	0.5507	0.4087	0.6938	**0.7020**
Recall	0.2992	0.4143	0.4023	0.5811	0.3726	0.6325	0.5378	0.4422	0.4804	0.3993	0.6533	**0.6977**
F-measure	0.2464	0.3171	0.3859	0.4142	0.4322	0.4231	0.4661	0.3931	0.4813	0.3968	0.6439	**0.6783**

segmentation and initial saliency maps. After refinement, the quality of the saliency maps improves significantly, which preserved the object edge and regularized the object shape. The multi-scale approach further improves the saliency result in the last subfigure as shown in Fig. 3.

TABLE II: Performance of two Fusion Ways

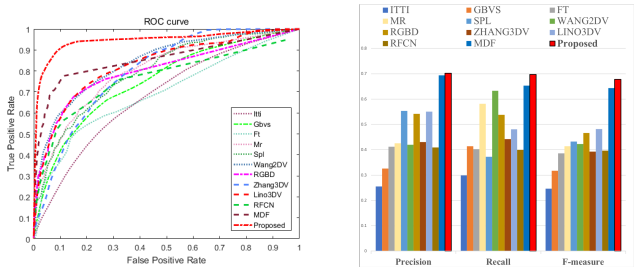
Fusion Way	Precision	Recall	F-measure
Only Average Sum	0.6717	0.7019	0.6683
Average Sum + Multiplication	**0.7020**	**0.6977**	**0.6783**

Meanwhile, as shown in Table II, the performance of the second fusion way is better than the first one which indicate that our proposed fusion method in Section II-D is reasonable and well-performed.

B. Comparison with other Methods

In the next experiments, we compare our proposal with eleven state-of-the-art salient detection methods: ITTI [1], GBVS [23], FT [26], MR [7], SPL [27], WANG2DV [28], RGBD [13], ZHANG3DV [9], LINO3DV [14] and other two deep-learning methods, MDF [17] and RFCN [29], in 2015 and 2016, respectively. In each methods, the parameters are set as default.

Results of seven sets are selected (shown in Fig. 4). From the results, we observe that the proposals except for ours would always fail in some scenery, and even the deep-learning methods, which shows in (l) and (m) column, yield poor results in 2^{nd} and 7^{th} rows. However, our proposal performs well at extracting the saliency object with a well-preserved edge and shape.



(a) ROC Curves (b) Evaluate Chart

Fig. 5: Evaluation

In addition, ROC curves, average Recall, F-measure and accuracy have been measured as well. Table I shows the Precision, Recall and F-measure among our proposal and other eleven methods (The methods with '*' are deep-learning methods). Fig. 5 summarizes the comparison between our method and other methods, in terms of ROC, Precision, Recall and F-measure. All these indexes shows the state-of-the-art performance of our proposal. Moreover, comparing with deep-learning methods, our proposed method is unsupervised.

IV. Conclusion

In this paper, we improved the SLIC algorithm and build an multi-scale architecture based on segmentation. Additionally, refining method based on graphical model is adopted to improve the object's shape and edge in the saliency maps. The experiments on 3D-HEVC and NAMA3DS1 datasets verify the excellent performance of our proposal. The saliency maps using the proposed method consist of objects with clear shape and edge, demonstrating the generality and robustness of the proposed algorithm.

REFERENCES

[1] 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, Nov 1998.

[2] D. Gao and N. Vasconcelos, "Bottom-up saliency is a discriminant process," in *2007 IEEE 11th International Conference on Computer Vision*, Oct 2007, pp. 1–6.

[3] D. Gao, V. Mahadevan, and N. Vasconcelos, "On the plausibility of the discriminant center-surround hypothesis for visual saliency," *Journal of Vision*, vol. 8, no. 7, pp. 1–18, 2008.

[4] M. M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, and S. M. Hu, "Global contrast based salient region detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 569–582, March 2015.

[5] F. Perazzi, P. Krhenbhl, Y. Pritch, and A. Hornung, "Saliency filters: Contrast based filtering for salient region detection," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, June 2012, pp. 733–740.

[6] B. Schlkopf, J. Platt, and T. Hofmann, *Graph-Based Visual Saliency*. MIT Press, 2007, pp. 545–552. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6287326

[7] C. Yang, L. Zhang, H. Lu, X. Ruan, and M. H. Yang, "Saliency detection via graph-based manifold ranking," in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, June 2013, pp. 3166–3173.

[8] J. Smisek, M. Jancosek, and T. Pajdla, "3d with kinect," in *IEEE International Conference on Computer Vision Workshops*, 2011, pp. 1154–1160.

[9] Y. Zhang, G. Jiang, M. Yu, and K. Chen, "Stereoscopic visual attention model for 3d video," in *International Conference on Advances in Multimedia Modeling*, 2010, pp. 314–324.

[10] C. Chamaret, S. Godeffroy, P. Lopez, and O. L. Meur, "Adaptive 3d rendering based on region-of-interest," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 7524, 2010.

[11] Y. Niu, Y. Geng, X. Li, and F. Liu, "Leveraging stereopsis for saliency analysis," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, June 2012, pp. 454–461.

[12] Y. Fang, J. Wang, M. Narwaria, P. L. Callet, and W. Lin, "Saliency detection for stereoscopic images," *IEEE Transactions on Image Processing*, vol. 23, no. 6, pp. 2625–2636, June 2014.

[13] H. Peng, B. Li, W. Xiong, W. Hu, and R. Ji, "Rgbd salient object detection: A benchmark and algorithms," *National Laboratory of Pattern Recognition*, vol. 8691, pp. 92–109, 2014.

[14] L. Ferreira, L. A. da Silva Cruz, and P. Assuncao, "A method to compute saliency regions in 3d video based on fusion of feature maps," in *2015 IEEE International Conference on Multimedia and Expo (ICME)*, June 2015, pp. 1–6.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *International Conference on Neural Information Processing Systems*, 2012, pp. 1097–1105.

[16] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, June 2009, pp. 248–255.

[17] G. Li and Y. Yu, "Visual saliency based on multiscale deep features," in *Computer Vision and Pattern Recognition*, 2015, pp. 5455–5463.

[18] R. Zhao, W. Ouyang, H. Li, and X. Wang, "Saliency detection by multi-context deep learning," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015, pp. 1265–1274.

[19] X. Wang, H. Ma, and X. Chen, "Salient object detection via fast r-cnn and low-level cues," in *2016 IEEE International Conference on Image Processing (ICIP)*, Sept 2016, pp. 1042–1046.

[20] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[21] C. Zach, T. Pock, and H. Bischof, "A duality based approach for realtime tv-l 1 optical flow," *Pattern Recognition*, pp. 214–223, 2007.

[22] J. S. Pérez, E. Meinhardt-Llopis, and G. Facciolo, "Tv-l1 optical flow estimation," *Image Processing On Line*, vol. 2013, pp. 137–150, 2013.

[23] B. Scholkopf, J. Platt, and T. Hofmann, *Graph-Based Visual Saliency*. MIT Press, 2007, pp. 545–552. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6287326

[24] K. Mller, H. Schwarz, D. Marpe, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, P. Merkle, F. H. Rhee, G. Tech, M. Winken, and T. Wiegand, "3d high-efficiency video coding for multi-view video and depth data," *IEEE Transactions on Image Processing*, vol. 22, no. 9, pp. 3366–3378, Sept 2013.

[25] M. Urvoy, M. Barkowsky, R. Cousseau, Y. Koudota, V. Ricorde, P. L. Callet, J. Gutirrez, and N. Garca, "Nama3ds1-cospad1: Subjective video quality assessment database on coding conditions introducing freely available high quality 3d stereoscopic sequences," in *2012 Fourth International Workshop on Quality of Multimedia Experience*, July 2012, pp. 109–114.

[26] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, June 2009, pp. 1597–1604.

[27] C. Yang, L. Zhang, and H. Lu, "Graph-regularized saliency detection with convex-hull-based center prior," *IEEE Signal Processing Letters*, vol. 20, no. 7, pp. 637–640, July 2013.

[28] W. Wang, J. Shen, and F. Porikli, "Saliency-aware geodesic video object segmentation," in *2015 IEEE Conference on Computer Vision and Pattern Recognition*, June 2015, pp. 3395–3402.

[29] L. Wang, L. Wang, H. Lu, P. Zhang, and X. Ruan, "Saliency detection with recurrent fully convolutional networks," in *European Conference on Computer Vision*. Springer, 2016, pp. 825–841.