

REFLECTANCE-BASED SURFACE SALIENCY

Gilles Pitard^{*,a} Gaëtan Le Goïc^b Alamin Mansouri^b Hugues Favrelière^c
Maurice Pillet^c Sony George^a Jon Yngve Hardeberg^a

^a The Norwegian Colour and Visual Computing Laboratory, Dept. of Computer Science, NTNU, Norway

^b Laboratoire LE2I, FRE CNRS 2005, Université de Bourgogne Franche-Comté, Auxerre, France

^c Laboratoire SYMME, EA 4144, Université Savoie Mont-Blanc, Annecy, France

ABSTRACT

In this paper, we propose an original methodology allowing the computation of the saliency maps for high dimensional RTI data (Reflectance Transformation Imaging). Unlike most of the classical methods, our approach aims at devising an intrinsic visual saliency of the surface, independent of the sensor (image) and the geometry of the scene (light-object-camera). From RTI data, we use the DMD (Discrete Modal Decomposition) technique for the angular reflectance reconstruction, which we extend by a new transformation on the modal basis enabling a rotation-invariant representation of reconstructed reflectances. This orientation-invariance of the resulting reflectance shapes fosters a robust estimation of saliency maps linked to the local visual appearance behaviour of surfaces on the scene. The proposed methodology has been tested and validated on real surfaces with controlled singularities, and the results demonstrated its efficiency since the estimated saliency maps show strong correlation with sensorial visual assessments.

Index Terms— Saliency, RTI, Surface appearance, DMD

1. INTRODUCTION

The visual appearance properties of surfaces are of primary importance in many fields such as in the quality inspection of manufacturing products in industry, or in the context of studying art and cultural heritage (CH) objects. Although the context, the imaging modalities or the purpose may vary, many common scientific issues regarding measuring and controlling the surface visual appearance remain the same, for instance, the automatic extraction of the most visually salient features of an inspected surface. One of the possible ways to achieve this objective is to measure and analyze locally (in each point/pixel) the way the incident light is re-emitted by the surface in all directions. Obviously, the Bidirectional Reflectance Distribution Function [1] (BRDF) is the most exhaustive quantity that describes this surface-light interaction. However its implementation is still complex, time-consuming and not affordable for the majority of applications where the visual inspection is applied. Beside, a recent and simpler approach for global appearance acquisition and modelling called Reflectance Transformation Imaging (RTI) has emerged. It consists in varying the illumination position, and acquiring an image for each position. Thus each point/pixel is de-

scribed by a high dimensional set of values rather than one value classically; each one corresponding to a specific illumination position. This set of values could be then interpolated or approximated by a method such as PTM (Polynomial Texture Mapping) [2, 3], HSH (Hemispherical Harmonics) [4] or recent one DMD (Discrete Modal Decomposition) [5] to construct the angular reflectance function for each pixel, thereby allowing to estimate the appearance of the inspected surface under novel light positions. This characterization of reflectance shapes is of high relevance for the description of the local visual appearance of surfaces, and is therefore capable of describing/capturing the singular elements on a surface. We therefore propose a methodology for the computation of saliency maps from RTI data, which is in addition rotation-invariant.

Although many reflectance measurements and models are reported in the literature [6], most of them rely on the assessment of either microgeometry or surface roughness to identify undesired local variations for which the light interaction induces changes in human perception [7, 8, 9]. In many cases however, salient features take their origin not only in local geometric variations, but also in many other physical causes affecting the appearance properties (gloss, sparkles, translucency or perceived color). The change in appearance properties is then mainly due to disruption on the spatial homogeneity of the local reflectance shapes, often at different scales. We build our methodology on the observation that rapid and local variations, such as specularities, shadows, or sparkles observed in reflectance shapes, has a strong correlation with the visual attraction (saliency). This proposal is consistent with the works of Itti *et al.* [10] who defined the saliency in terms of Bayesian surprise and suggest that attention/gaze is attracted by visually surprising signals. Our methodology extends the recent method DMD aiming at the reconstruction of the local angular reflectance shape [5]. We propose a new development on the DMD approach that makes the DMD basis invariant to the rotation. The similarity of the local estimated reflectance shapes is finally assessed by computing a multivariate analysis (section 2.4) of obtained rotation-invariant descriptors, and leads to significantly increasing the correlation between the estimated saliency maps and the human visual perception.

2. PROPOSED METHODOLOGY

As the high quality of the source images is essential, we first briefly recall the main requirements for RTI stereo-photometric data acquisitions. Then we detail the next three main steps of our methodology.

The authors would like to warmly thank the Regional Research Council (RFF-Innlandet, Norway) and the partners of the MeSurA project (Measuring Surface Appearance) for their support.

* Corresponding Author: gilles.pitard@ntnu.no

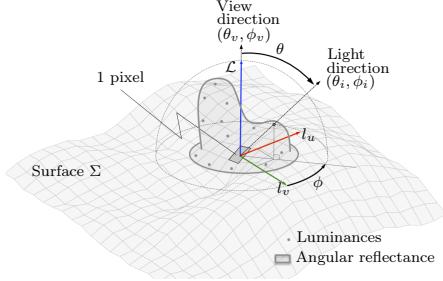


Fig. 1: Angular reflectance approximated from the measured luminances expressed in the coordinate system (l_u, l_v, \mathcal{L})

2.1. RTI data acquisition requirements

An RTI setup is used to acquire for each pixel of the scene the luminance values \mathcal{L} under varying light positions defined in the coordinate system (l_u, l_v, \mathcal{L}) as shown in Figure 1, where (l_u, l_v) are the components associated with light positions (LP) projected in the horizontal plane. The sensor is generally fixed at the centre of a dome [2], orthogonal to the inspected surface. In certain cases, the constraints intrinsic to the object or its environment may however restrain the deployment of the dome-based setup. In these cases, a manual acquisition protocol (H-RTI) can be implemented [11]. In order to be able to exploit the RTI data, it is of primary importance to control the intensity and the position/direction of the lighting sources to ensure the quality and subsequent modelling of angular reflectance.

2.2. Angular reflectance modelling

Modelling of the angular reflectance from RTI acquisitions aims at approximating the set of the discrete points representing the luminance values at each pixel for all light positions (hemispheric shaped) by a continuous surface (Fig. 2). This step is called surface parametrization for which, to our knowledge, only three methods exist in literature: PTM [2, 3], HSH [4] and DMD [5].

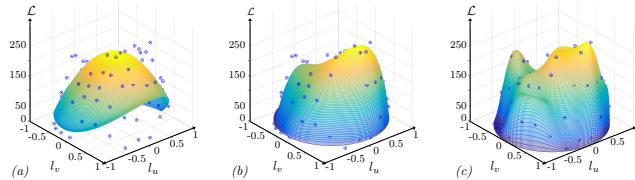


Fig. 2: Angular reflectance for the same pixel, approximated from measured luminances (blue dots) using respectively (a) PTM, (b) HSH and (c) DMD methods

The objective of the parametrization is to obtain the best fitting of the information hold by the discrete points while compacting it reliably. We chose to implement here the Discrete Modal Decomposition (DMD), which outperforms the others, especially in detecting and highlighting singularities on surfaces [5]. The DMD method is based on a projection on the Modal bases. The modal bases are composed of elementary forms that derive from a structural dynamic problem. We adapt these modal bases to the case of angular reflectance reconstruction by devising new modal shapes that we call Reflectance Modal Basis (RMB), solution of the structural dynamic

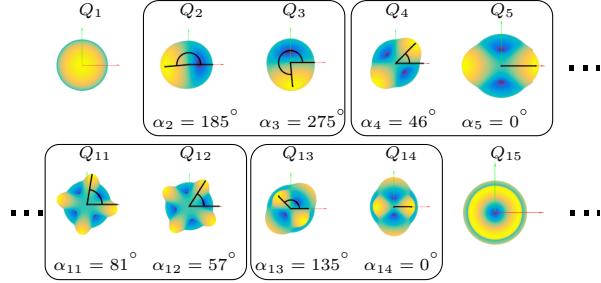


Fig. 3: Representation of the modal basis first modes

problem whose reference surface is a hemisphere. This set of elementary forms noted Q_i (with $i = 1..m$ where m is the number of modes) is used to reconstruct the reflectance shape from discrete values (luminance) obtained during the acquisition stage [12]. A representation of first modes of the RMB is given in Figure 3.

Therefore, the local reflectance associated to each point/pixel of the inspected surface is then described by a modal spectrum representing the contributions λ_i of each mode of the RMB.

2.3. Rotation-invariant representation

One interesting property of the modal shapes is that it can be separated in rotation-dependent and rotation invariant modes, named respectively *simple* and *congruent* modes. The shape of the *simple* modes presents a rotational symmetry which preserve their shape for any rotation around vertical axis (modes Q_1 or Q_{15} in Fig. 3). They are thereby originally rotation-invariants. Some pairs, namely the *congruent* modes, show the same shape but are oriented differently (for instance Q_{2-3} or Q_{4-5} in Fig. 3). As a consequence, in this original form, the modal spectrum, composed of λ_i coefficients will then vary when the orientation (rotation for example) of the object under the camera is modified. The same principle applies to, for example, a sample with a regular pattern oriented differently across its surface: two points with the same visual appearance behaviour but different orientation will be described by completely different modal spectrum. In order to use the modal information for the estimation of visual saliency maps, it is thus essential to transform the original modal shapes to make them rotation-invariant.

The proposed method to obtain a rotation-invariant representation is based on the separation of the phase and amplitude components. This separation is performed through a change applied on the congruent modes group presented below.

The amplitude and the phase-angle (resp. λ'_j et α'_j) can be derived from the expression of the linear combination of two congruent modes Q_i and Q_{i+1} and their modal coefficients λ_i and λ_{i+1} as shown as the Figure 4. The resulting amplitude λ'_j is obtained by computing the L2-norm: $\lambda'_j = (\lambda_i^2 + \lambda_{i+1}^2)^{1/2}$.

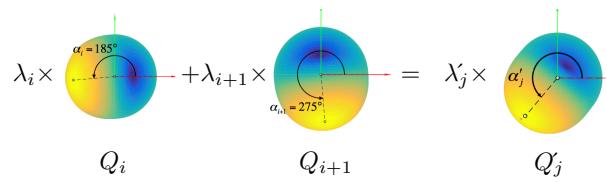


Fig. 4: Change of parametrization in the amplitude/phase-angle form

The L2-norm does not vary when the orientation of the reflectance is rotated around the vertical axis of (l_u, l_v, \mathcal{L}) . The vector of resulting amplitudes λ'_j is therefore a subset of the initial coefficients λ_i which characterize the shape of any reflectance shapes independently of its spatial orientation. In other words, the angular reflectances which are identical in shape but differently oriented, i.e. which can be brought into coincidence by rotation, present the same modal spectra in this rotation-invariant reflectance description. The phase-angle α'_j of the associated Q'_j mode can be determined by the following equation 1, where the phase-angle α_i of each mode is known *a priori* and saved with the RMB. This phase information is not used for visual saliency estimation methodology. However, it could be relevant for other applications for instance to assess the orientation of a textural pattern or to determine the appropriate light directions in order to find specific visual appearance features on surfaces.

$$\alpha'_j = \alpha_i + \arctan \frac{\lambda_{i+1} \sin(\alpha_{i+1} - \alpha_i)}{\lambda_i + \lambda_{i+1} \cos(\alpha_{i+1} - \alpha_i)} \quad (1)$$

An other interesting consequence of the proposed compaction is the reduction of dimensionality of reflectance representation: both the invariance and the compactness of representation give a significant advantage to, for example, automate visual inspection [13] or appearance classification tasks [14, 15].

2.4. Visual saliency estimation

In order to enable the identification of atypical (salient) behaviours in term of angular reflectance, we implement a multivariate statistical analysis on the local invariant modal spectra. The multivariate image is defined as a three-dimensional $n_1 \times n_2 \times m_B$ data matrix, where two spatial dimensions represent the $n_1 \times n_2$ image pixels and the third dimension represents the m_B modal rotation-invariant coefficients λ'_j . We can reshape the $n_1 \times n_2$ - pixel images of the amplitude series into column vectors with dimension $n = n_1 n_2$ and express the multivariate image data as two-dimensional $n \times m_B$ matrix $\boldsymbol{\lambda}' = [\lambda'_1, \lambda'_2, \dots, \lambda'_{m_B}]$. Then, we use the Mahalanobis distance D_{Mahal} [16] to assess the distance between the reflectance modal spectrum of each element (pixel) and the average modal spectrum of reflectance estimated on the whole image, that means from all vectors $\boldsymbol{\lambda}'$. The expression of D_{Mahal} is given by:

$$D_{\text{Mahal}} = \sqrt{(\boldsymbol{\lambda}' - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\lambda}' - \boldsymbol{\mu})} \quad (2)$$

where $\boldsymbol{\Sigma}$ and $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \dots, \mu_p)^T$ are respectively the covariance matrix and the average vector describing the mean reflectance shape across the surface. The visual saliency is finally computed from this distance estimation on each point of the inspected surface and plotted as saliency maps. Figure 5 summarizes the methodology we propose for saliency computation on high dimensional RTI data.

3. RESULTS AND DISCUSSION

The experimental results are presented on different challenging surfaces (metallic) exhibiting several visual appearance defects. The results are compared with both human observers assessment and a well known state-of-the art method for saliency maps computation. Since the state-of-the-art method was not initially adapted to RTI data and objects we studied, we introduced an adaptation on it. The samples acquisitions were acquired with a dome with 96 light sources evenly distributed spatially.

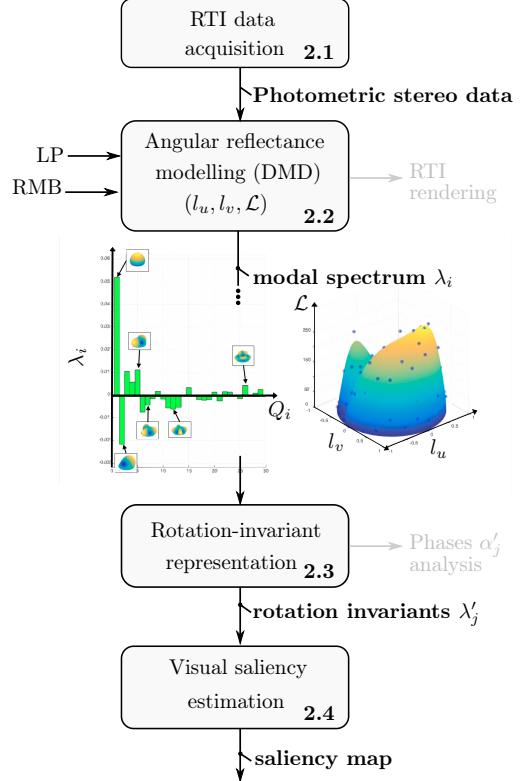


Fig. 5: Methodology flow chart

3.1. Datasets

The first dataset is associated to polished metallic surfaces (Fig. 6-top), on which nano-scratches were made with a Berkovish head. The scratches geometry is thereby controlled, and can be described by the parameters presented in Table 1. The visibility of the calibrated scratch on each sample may vary depending on its geometric parameters (the width between two peaks, the width at the surface and the depth) because these three characteristics are involved in the shape of specular lobes observed during the luminous reflection at their location on the surface. These samples were assessed through a sensorial analysis performed by trained quality controllers, as described in [17]. Particularly, the adaptation phenomenon [18] during this sensory evaluation was notably reduced by positioning the scratches at random locations on each sample (Fig. 6). A *visibility* ratio is thereby associated to each sample (Table 1). As other uncontrolled degradations are observed on the samples, these defects present a good challenge for our approach.

The second dataset is a set of two shiny metallic surfaces from the watch-making industry. The first sample is a steel bracelet link, mirror polished, with a convex shape (Fig. 7-left side). The second one is a watch dial surface which is finished with a sandblasting and a radial brushing (Fig. 7-right side).

3.2. Saliency results

The experimental saliency results using our methodology on these samples (datasets 1 and 2) are presented respectively in Figures 6 and 7. For the dataset 1, we also present, in a purpose of compar-

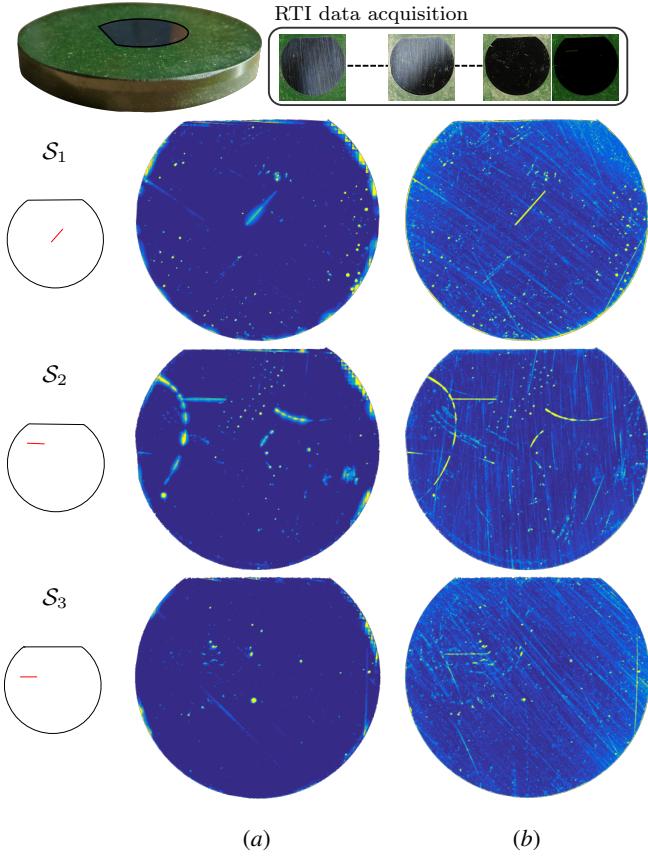


Fig. 6: Saliency maps S_{1-3} obtained, from RTI data, with (a) the Itti and Koch's saliency model and (b) our method

son, the experimental results with Itti and Koch's saliency approach [19, 20]. This approach is based on three visual features - luminance contrast, orientation and colour - with each feature computed over a multi-scale Gaussian pyramid. Itti's model was largely extended to the saliency estimation in a dynamic scene (video) by integrating motion estimators [21], applications on 3D meshes [22], or multispectral content [23]. As this model was optimized for a single image scene with macroscopic objects, a change is thus applied on the parameters related to the scale levels ($c \in \{1, 2, 3, 4\}$ and $\delta \in \{1, 2, 3\}$). These parameters were validated on synthetic rough surfaces [24, 25] in order to be sensitive to high frequency spatial information contained in the surface images. The Itti's method is applied on each input image of the RTI dataset in order to generate a saliency map associated with the different light directions, and then

Sample	$a (\mu m)$	$b (\mu m)$	$c (\mu m)$	Visibility (%)
S_1	1.25	0.75	58.32	64.6
S_2	0.98	0.61	43.57	14.5
S_3	0.68	0.44	26.1	5



Fig. 7: Saliency maps of metallic surfaces, with an input image of the associated RTI dataset

all of those maps are combined in a final saliency map (Fig. 6a). An effective separation of the scratches on the polished region is obtained by the two approaches. The efficiency of the method is confirmed by the high correlation with the sensorial visibility ratio (Table 1). With regards to Itti's -based results, the saliency maps are rather rotation-dependent since the small scratches are not detected. In addition, major limitations of Itti's method are related to its computational complexity and time costs of processing all images of RTI data structure: our method allows significant time savings (150 times faster) thereby satisfying time requirements in many inspection applications of real-world surfaces. Dataset 2 results (watch-making parts - Fig. 7), indicate that the proposed methodology has a significant robustness to non-flat surfaces (bracelet link) results, in absence or presence of directional pattern (watch dial). This confirms the relevance of the rotation invariance property introduced in this methodology. However, the performance of our proposed methodology could be affected in the cases of more complex freeform surfaces due to the effects of the various gradient distributions upon the corresponding visual appearance, which would require to take into account the local curvature of the object, that affects the inclination of the angular reflectance shapes.

4. CONCLUSION

This paper proposed a novel methodology for the detection of visually-salient features of real surfaces. The method uses RTI data acquisitions to reconstruct local angular reflectance on each point/pixel of the inspected surface. The reflectance reconstructions are achieved using the Discrete Modal Decomposition, which provides an accurate assessment of this information. We moreover introduced a change in the DMD parametrization to make this representation invariant to the orientation, and thus insensitive to the object rotation or the roughness pattern orientation on the surface. The saliency maps computed in this discriminative reflectance space provide a robust assessment of the visually-salient features of a surface in terms of accuracy, time effective and perception correlation.

Table 1: Results of the sensory analysis on S_{1-3} samples [17]

5. REFERENCES

- [1] F E Nicodemus, J C Richmond, J J Hsia, I W Ginsberg, and T Limperis, *Geometrical considerations and nomenclature for Reflectance*, Institute for Basic Standards, National Bureau of Standards, Washington, 1977.
- [2] T Malzbender, D Gelb, and H Wolters, “Polynomial texture maps,” *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, 2001.
- [3] Mark S Drew, Yacov Hel-Or, Tom Malzbender, and Nasim Ha-jari, “Robust estimation of surface properties and interpolation of shadow/specularity components,” *Image and Vision Computing*, vol. 30, no. 4-5, pp. 317–331, 2012.
- [4] Mingjing Zhang and Mark S Drew, “Efficient robust image interpolation and surface properties using polynomial texture mapping,” *EURASIP Journal on Image and Video Processing*, vol. 2014, no. 1, pp. 1–19, 2014.
- [5] G Pitard, G Le Goic, H Favreliere, S Samper, SF Desage, and M Pillet, “Discrete Modal Decomposition for surface appearance modelling and rendering,” *SPIE Optical Metrology*, vol. 9525, pp. 952523–952523–10, 2015.
- [6] D Guarnera, G C Guarnera, A Ghosh, C Denk, and M Glen-cross, “BRDF Representation and Acquisition,” *Computer Graphics Forum*, vol. 35, no. 2, pp. 625–650, 2016.
- [7] M L Smith, “The analysis of surface texture using photometric stereo acquisition and gradient space domain mapping,” *Image and Vision Computing*, vol. 17, no. 14, pp. 1009–1019, 1999.
- [8] Thomas Leung and Jitendra Malik, “Representing and Recognizing the Visual Appearance of Materials using Three-dimensional Textons,” *International Journal of Computer Vision*, vol. 43, no. 1, pp. 29–44, 2001.
- [9] Jiahua Wu and Mike J Chantler, “Combining gradient and albedo data for rotation invariant classification of 3D surface texture,” *ICCV 2003: 9th International Conference on Computer Vision*, pp. 848–855 vol.2, 2003.
- [10] L Itti and P F Baldi, “Bayesian surprise attracts human attention,” *Neural Information Processing Systems*, pp. 547–554, 2005.
- [11] G Earl and K Martinez, “Archaeological applications of polynomial texture mapping: analysis, conservation and representation,” *Journal of Archaeological Science*, vol. 37, no. 8, pp. 2040–2050, 2010.
- [12] G Pitard, *Surface appearance metrology and modeling for industrial quality inspection*, Ph.D. thesis, Université Grenoble Alpes, 2016.
- [13] Robin Gruna and Jurgen Beyerer, “Acquisition and evaluation of illumination series for unsupervised defect detection,” *IEEE International Instrumentation and Measurement Technology Conference*, 2011.
- [14] O Wang, P Gunawardane, S Scher, and J Davis, “Material classification using BRDF slices,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops)*, pp. 2805–2811, 2009.
- [15] Chao Liu and Jinwei Gu, “Discriminative Illumination: Per-Pixel Classification of Raw Materials Based on Optimal Projections of Spectral BRDF,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36, no. 1, pp. 86–98, 2014.
- [16] P C Mahalanobis, “On the generalized distance in statistics,” *Proceedings of the National Institute of Sciences (Calcutta)*, 1936.
- [17] T Puntous, S Pavan, D Delafosse, M Jourlin, and J Rech, “Ability of quality controllers to detect standard scratches on polished surfaces,” *Precision Engineering*, vol. 37, no. 4, pp. 924–928, 2013.
- [18] Jillian H Fecteau and Douglas P Munoz, “Exploring the consequences of the previous trial,” *Nature Reviews Neuroscience*, vol. 4, no. 6, pp. 435–443, 2003.
- [19] Laurent Itti and Christof Koch, “A saliency-based search mechanism for overt and covert shifts of visual attention,” *Vision research*, vol. 40, no. 10-12, pp. 1489–1506, 2000.
- [20] D Walther and C Koch, “Saliency Toolbox 2.3,” <http://www.saliencytoolbox.net>, 2006.
- [21] Laurent Itti, Nitin Dhavale, and Frederic Pighin, “Realistic avatar eye and head animation using a neurobiological model of visual attention,” *Optical Science and Technology, SPIE's 48th Annual Meeting*, vol. 5200, pp. 64–78, Jan. 2004.
- [22] Chang Ha Lee, Amitabh Varshney, David W Jacobs, Chang Ha Lee, Amitabh Varshney, and David W Jacobs, “Mesh saliency,” *ACM Transactions on Graphics (TOG)*, vol. 24, no. 3, pp. 659–666, July 2005.
- [23] Steven Le Moan, Alamin Mansouri, Jon Yngve Hardeberg, and Yvon Voisin, “Saliency for Spectral Image Analysis,” *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 6, no. 6, pp. 2472–2479, 2013.
- [24] A D F Clarke, P R Green, M J Chantler, and K Emrith, “Visual search for a target against a $1/f\beta$ continuous textured background,” *Vision research*, vol. 48, no. 21, pp. 2193–2203, 2008.
- [25] Alasdair D F Clarke, Mike J Chantler, and Patrick R Green, “Modeling visual search on a rough surface,” *Journal of Vision*, vol. 9(4), no. 11, pp. 1–12, 2009.