QUATERNION HARRIS FOR MULTISPECTRAL KEYPOINT DETECTION

Giorgos Sfikas[1,2]*, Dimosthenis Ioannidis*[2] *and Dimitrios Tzovaras*[2]

[1]Dept. of Computer Science & Engineering, University of Ioannina, Greece
[2] Centre for Research & Technology, Information Technologies Institute, Thessaloniki, Greece

ABSTRACT

We present a new keypoint detection method that generalizes Harris corners for multispectral images by considering the input as a quaternionic matrix. Standard keypoint detectors run on scalar-valued inputs, neglecting input multimodality and potentially missing highly distinctive features. The proposed detector uses information from all channel inputs by defining a quaternionic autocorrelation matrix that possesses quaternionic eigenvectors and real eigenvalues, for the computation of which channel cross-correlations are also taken into account. We have tested the proposed detector on a variety of multispectral images (color, near-infrared), where we have validated its usefulness.

***Index Terms*—** Quaternions, Keypoint detection, Multispectral images

1. INTRODUCTION

Keypoint or feature detection is a process that is used as a first step in a multitude of problems in computer vision and image processing [1]. It finds uses in various applications, including panorama stitching, object detection, image retrieval, photogrammetry and 3D reconstruction [2, 3]. Since the first keypoint detectors around the early 1980s and the introduction of major milestones such as the Harris and SIFT algorithms, today the trend on keypoint detection and description is centered around learning and deep learning-based methods [4, 5, 6]. Success of learning-based methods in creating keypoint detectors has not been as spectacular as in other fields in vision or even keypoint description [4], and state-of-the-art methods that are based on keypoint detection (e.g. 3D reconstruction) still make use of standard, "hand-crafted" methods [2, 7]. The recent state-of-the-art Key.Net detector makes use of both hand-crafted and learned filters [4].

While hand-crafted detectors are still relevant today, the vast majority of either older or more recent detectors relies on estimating a feature response metric based on first and second-order gradients of a single input channel (e.g. structure tensors / autocorrelation for Harris, Difference of Gaussians for SIFT, etc.). If the input comprises more than one channel, a significant amount of information, including channel cross-correlation, must simply be neglected. Multispectral inputs include color images as the simplest case, comprising 3 channels. A paradigm that is becoming increasingly more relevant is the availability of extra channels corresponding to non-visible band sensors such as Depth sensors, Near Infrared [8] or thermography imaging [9] among other cases [10]. Decreasing sensor hardware costs have contributed to this trend, with imaging inputs to digital image processing pipelines comprising up to 4 channels being much less rare than a decade ago.

With this work, we propose a keypoint detection scheme that can handle multispectral inputs by using a quaternion image representation and defining a corresponding quaternionic structure tensor and cornerness function, generalizing the classical Harris keypoint detection algorithm. Quaternions are mathematical entities that can be seen as 4-dimensional generalizations of complex numbers [11, 12, 13]. Hence, the imaged multispectral cues can be treated each as a single entity per pixel. The multimodal natural of the input, including cross-channel correlations can consequently be taken into account in a natural manner. In [8], a multi-band extension of the autocorrelation matrix had been proposed, that cannot take into account cross-channel correlations, as it is defined as the sum of autocorrelation over all bands separately. Furthermore, the eigenstructure of the proposed quaternionic autocorrelation matrix preserves information over all channels. Experimental results on color and multispectral images (color + infrared) validate the usefulness of the proposed detector.

In the remainder of the paper, we examine preliminaries on quaternionic analysis in section 2, define the proposed Quaternionic Harris (QuatHarris) detector in section 3, and discuss eigenstructure computation of the defined quaternionic autocorrelation in section 4. We showcase numerical and qualitative experiments in section 5 and close the paper with a short discussion in section 6.

2. PRELIMINARIES ON QUATERNIONS

Quaternions are mathematical objects that form a skew-field $\mathbb{H}$, *i.e.* quaternion addition and multiplication are defined

The work presented was carried out as part of the eDREAM project which is co-funded by the EUs Horizon 2020 framework programme for research and innovation under grant agreement No 774478.

with all the properties of a field, except that of multiplication commutativity. Quaternions $q \in \mathbb{H}$ share the basic form: $q = a+b\boldsymbol{i}+c\boldsymbol{j}+d\boldsymbol{k}$, where $a, b, c, d \in \mathbb{R}$ and $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ are independent imaginary units, and in general $pq \neq qp$ for $p, q \in \mathbb{H}$. Real and complex numbers can be regarded as quaternions with $b, c, d = 0$ or $c, d = 0$ respectively. By definition it holds that $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = \boldsymbol{ijk} = -1$, and consequently $\boldsymbol{ij} = -\boldsymbol{ji} = \boldsymbol{k}, \boldsymbol{jk} = -\boldsymbol{kj} = \boldsymbol{i}, \boldsymbol{ki} = -\boldsymbol{ik} = \boldsymbol{j}$. The length or magnitude of a quaternion is defined as $|q| = \sqrt{q\bar{q}} = \sqrt{\bar{q}q} = \sqrt{a^2+b^2+c^2+d^2}$, where $\bar{q}$ is the conjugate of q, defined as $\bar{q} = a - b\boldsymbol{i} - c\boldsymbol{j} - d\boldsymbol{k}$. The property $\mu^2 = -1$ holds if and only if μ is a unit pure (i.e. with zero real part) quaternion [12, 11]; hence, since there is an infinite number of unit pure quaternions, $\mu^2 = -1$ possesses an infinite number of solutions, in contrast to complex algebra.

A useful representation of quaternions is through the Caley-Dickson form [12, 14], where a quaternion may be represented in a unique way as a complex number with complex real and imaginary parts in particular $q = A + B\boldsymbol{j}$ with $A = a + b\boldsymbol{i}, B = c + d\boldsymbol{i}$. An analogous operation can be performed for quaternion matrices, which can be written as tuples of complex matrices [11, 15]. Indeed, for any quaternionic matrix A, there exist unique quaternionic matrices A_1, A_2 such that $A_1 + A_2\boldsymbol{j}$ [11]. Furthermore, the following function can be defined, mapping any $n \times n$ quaternionic matrix to a $2n \times 2n$ complex matrix:

$$\chi_A = \begin{bmatrix} A_1 & A_2 \\ -\bar{A}_2 & \bar{A}_1 \end{bmatrix} \tag{1}$$

Matrix χ_A is called the adjoint of A. The mapping χ to the adjoint bears interesting properties [11], from which follows that it can be used to compute eigenvalues and eigenvectors of quaternionic matrices. Note that eigenvalues are discerned between left and right ones for quaternionic matrices, as $Ax = \lambda x$ is a different problem than $Ax = x\lambda$ [11], while right eigenvalues are in general infinite in number.

3. PROPOSED DETECTOR

Similar to the motivation behind with the classical Harris keypoint detector (or the older Moraveć detector), we aim to find image points where the following quantity is maximized for small displacements Δx:

$$E(x) = \sum_{x_n \in N(x)} g(x_n)|I(x_n) - I(x_n + \Delta x)|^2 \tag{2}$$

where $N(x)$ represents a set that contains points in a neighbourhood around x and including itself, and $I(x)$ represents image intensity at point x. Intensity for each point x is considered quaternionic, i.e. $I(x) \in \mathbb{H} \forall x$, so up to 4 channels can be taken into account by assigning each channel to one of the real or imaginary components. If the number of channels is less than 4, the corresponding quaternionic component is set to zero for all image points. The weight function $g(x_n) \in \mathbb{R}^+$ is used to weigh summation terms that correspond to points x_n according to their distance to point x. By taking the first term of the quaternion-valued Taylor expansion [16] and $|x|^2 = x\bar{x}$ for $x \in \mathbb{H}$ [11], eq. 2 is written as

$$\begin{aligned} E(x) &= \sum_{x_n \in N(x)} g(x_n)\Delta x^T \nabla I(x_n)\overline{\Delta x^T \nabla I(x_n)} \\ &= \sum_{x_n \in N(x)} g(x_n)\Delta x^T \nabla I(x_n)\overline{\nabla I(x_n)^T}\Delta x \\ &= \Delta x^T [\sum_{x_n \in N(x)} g(x_n)\nabla I(x_n)\nabla I(x_n)^H]\ \Delta x \end{aligned} \tag{3}$$

where H denotes the conjugate transpose. From the above formula, the terms under summation result to the 2×2 quaternion-valued matrix:

$$A_q = \sum_{x_n \in N(x)} g(x_n)\nabla I(x_n)\nabla I(x_n)^H \tag{4}$$

It is straightforward that A_q is Hermitian. Consequently, the right eigenvalues of A_q are real [11], exactly two and are also left eigenvalues, since $\lambda x = x\lambda$ when $\lambda \in \mathbb{R}$ [1]. A cornerness measure $c(A_q)$ with a parameter k can then be defined in order to classify candidate points as keypoints as done in the standard Harris algorithm [3], with $c(A_q) = (\lambda_1\lambda_2) - \kappa(\lambda_1+\lambda_2)^2$. Interestingly, note that the off-diagonal elements of each rank-one component of A_q are of the form $cI_x\overline{I_y}$ or its conjugate. These elements are in turn a function of factors combining information from different channels, thus expressing channel cross-correlation in A_q.

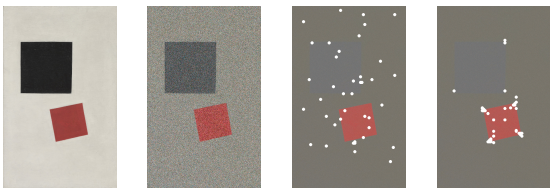

Fig. 1. From left to right: Original image, image degraded with LUN=70% and WGN $\sigma^2 = 10$, result of Harris detector on image degraded with LUN=99% and WGN $\sigma^2 = 1$, result of proposed QuatHarris detector on the same image.

4. EIGENSTRUCTURE OF QUATERNIONIC AUTOCORRELATION

Concerning computation of the eigenstructure of A_q, we first consider its diagonalisation as $U^H A_q U = \Lambda$, where Λ is a real diagonal matrix with the two eigenvalues in its diagonal and U a quaternion unitary matrix [11]. We can proceed by

[1] This is a consequence of theorem 5.4 and corollary 6.2 in [11].

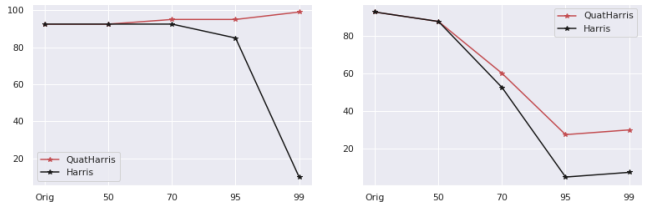

Fig. 2. Comparison of detector accuracy (higher values are better). Tests were run on degraded versions of the image in Fig. 1. The horizontal axis corresponds to increasing levels of LUN degradation. WGN levels are $\sigma^2 = 1$ and 10 on the left and right plots respectively.

taking the adjoint of both sides: $\chi_U^H \chi_{A_q} \chi_U = \chi_\Lambda$, where we have used $\chi_{AB} = \chi_A \chi_B$ and $\chi_{U^H} = \chi_U^H$ [11]. It now suffices to diagonalize the complex 4×4 matrix χ_{A_q} (also Hermitian [11]). With a similar argument as for the Quaternionic Singular Value Decomposition (QSVD) [15], eigenvalues will be obtained in two pairs of repeating values, and eigenvectors are found as $v_n = c_{n'} - \overline{d_{n'}}\boldsymbol{j}$ for $n = 1, 2$ and $n' = 2n - 1$, where $w_k = [\, c_k \quad d_k \,]^T$ is the k^{th} column of χ_U after indices are reordered by eigenvalue magnitude.

5. EXPERIMENTS

In this section, we present a set of experiments where we test the proposed quaternionic detector on various color and multispectral inputs. In all cases, $k = 0.04$ was used as the parameter of the cornerness function for all Harris variants, including the proposed one, and a multiscale-affine implementation was employed [2].

Proof of concept. In the first proof-of-concept experiment, we compare QuatHarris against the standard Harris detector. Unlike Harris, which must run over a single modality regardless of the number of input modalities, the proposed QuatHarris can exploit information from all modalities jointly. To this end, we ran the two detectors on degraded versions of the image shown in Figure 1 [2]. We have applied two types of degradations on the input: *Luminance Uniformity Noise (LUN)* and White Gaussian Noise (WGN). With what we call LUN noise in this paper, we suppress information on the luminance (L) channel of the input up to a certain amount specified by the noise parameter, expressed in terms of a percentage ($\alpha \in [0, 1]$). In particular, assuming an undegraded input $I(x, y)$, we define $I_{chroma}(x, y)$ as a version of I where the luminance channel is set to a fixed value for all pixels (50, i.e. average luminosity), hence only chrominance information is preserved. The image degraded with level α LUN is then defined as: $I_\alpha^{LUN}(x, y) = (1 - \alpha)I(x, y) + \alpha I_{chroma}(x, y)$. Hence, LUN intensity increases as α tends to 1. We have measured keypoint detection accuracy by counting the number of keypoint estimates that were located around square vertices (since these are the only corners of the image). By examining the qualitative result in Fig. 1, we can observe that the standard Harris detector fails to find meaningul keypoints, as it works on luminance information. Note that we could expect a similar result for any hand-crafted keypoint detector that is working on single-channel gradient information (e.g. SIFT, SURF, etc.). On the contrary, keypoints are concentrated around actual corners with the proposed detector. Numerical results over detections on differing degradation parameters can be examined in Figure 2, where QuatHarris is again shown to outperform Harris.

Visualization of quaternionic eigenvectors: We have seen that both eigenvalues of A_q are real, as it is by construction Hermitian (sec. 2); however, both of its eigenvectors are in general quaternionic. In Figure 3 we show visualizations of its eigenvectors. We take unit-norm eigenvectors $v = [v_1 v_2]^T$; note however that this does not mean that the quaternionic components $v_1, v_2 \in \mathbb{H}$ of the eigenvector are also unit-norm. Hence we visualize each eigenvector component as two images, one that contains its magnitude, and one that contains information about its angle. For this latter, we normalize each eigenvector component to unity and use its Euler angle representation to map it onto $\Re^3$ [17] and subsequently as a color image. Note that, as is the case with complex Hermitian matrices, quaternionic eigenvectors of Hermitian matrices are orthogonal to one another as well.

Image matching: We then use the proposed keypoint detector to find matches over tuples of frames of a color video. The video used for the test has been shot with a DJI M200 Unmanned Aerial Vehicle (UAV) on the premises of the Centre of Research and Technology (CERTH) in Thessaloniki, Greece (Fig. 4). A total of 23 tuples of frames (each at a resolution of 500×375 pixels) were given as input to a matching pipeline consisting of the following steps: Detection of K keypoints on each frame QuatHarris and description with SIFT, and matching them with RANSAC. As all input images are consecutive frames of the same time series, we can expect a high number of keypoint matches. As a check to test robustness of the QuatHarris cornerness measure, we have extracted only a low number of keypoints, starting from $K = 50$. In Table 1, we can view a numerical comparison between the proposed QuatHarris detector and the Multispectral Harris detector [8]. We show the cumulative total number of matches, i.e. over all 23 tuples, as well as the number of tuples where the one method or the other has scored more matches than the other. (Note that for a number of tuples a RANSAC match could not be established, hence the average is less than 3 in $K = 50$). A slight advantage for lower K values for the proposed QuatHarris method may be observed. This could be attributed to the fact that while both detectors take all channels into account, only the QuatHarris cornerness measure also uses cross-channel correlation information, resulting in slightly more distinctive keypoints. This advantage becomes

[2] Digitized version of the painting "Boy with Knapsack - Color Masses in the Fourth Dimension" by K.Malevich (1915).

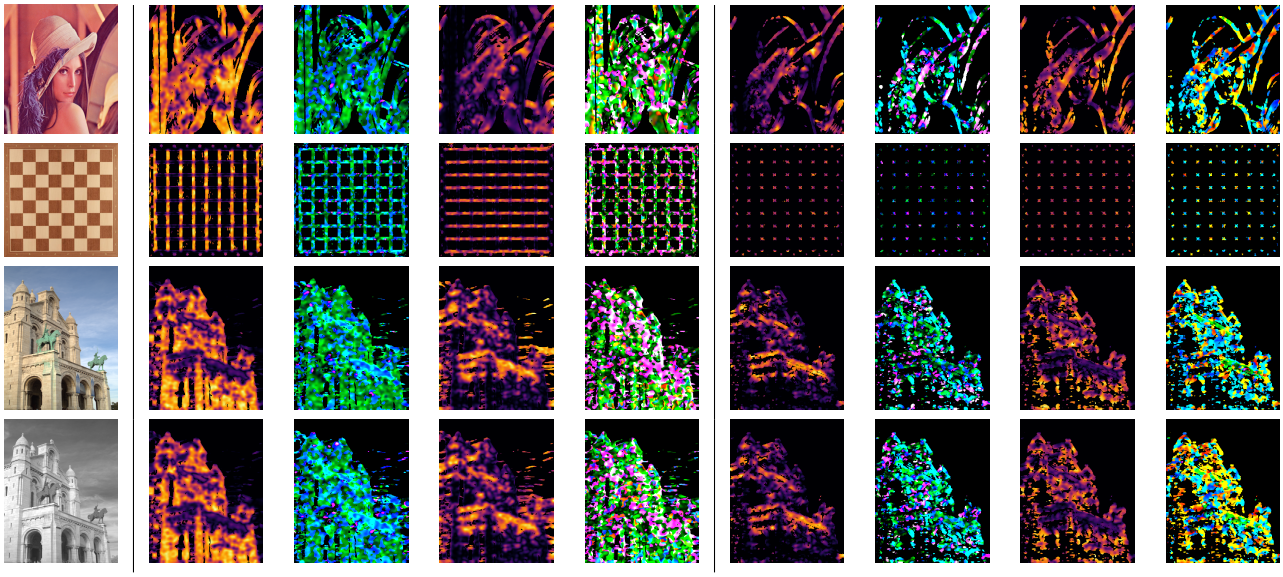

Fig. 3. Visualization of quaternionic autocorrelation eigenvectors. Each eigenvector is comprised of two quaternionic components. On each column, from left to right: Original image, magnitude/angle of the 1^{st} component of the major eigenvector, magnitude/angle of the 2^{nd} component of the major eigenvector, magnitude/angle of the 1^{st} component of the minor eigenvector, magnitude/angle of the 2^{nd} component of the minor eigenvector. In the last row, input is a 4-dimensional image comprised of the RGB component of the input image of the row above and the input shown on the current row as the NIR channel.

Fig. 4. Samples of the CERTH dataset.

less pronounced as K increases and RANSAC attains a larger pool of candidate matches.

Table 1. Comparison between QuatHarris and Multispectral Harris. The number of CERTH dataset tuples where one detector outperforms the other is reported.

Method / #kpoints	50	100	150	200	250
QuatHarris	**2**	**9**	**7**	6	7
Multispectral	1	4	3	**7**	**9**

Object detection: Finally, we have tested QuatHarris in the context of object detection. The test was run on a multispectral image and template of the database introduced in [18] (see Fig. 5). The image is comprised of 4 channels, namely the RGB channels plus a Near Infrared (NIR) channel. NIR is known to bear information that is complementary to the channels corresponding to the visible spectrum [18, 8], hence including it will in principle lead to better, more distinctive keypoints. We have extracted 500 keypoints for the template and target image, for which we have computed SIFT descriptors. While this figure may seem high, a correspondence can be established with QuatHarris keypoints, but not so with standard Harris keypoints. As Harris cannot take advantage of all available channels, its cornerness response is led to prioritize less distinctive keypoints.

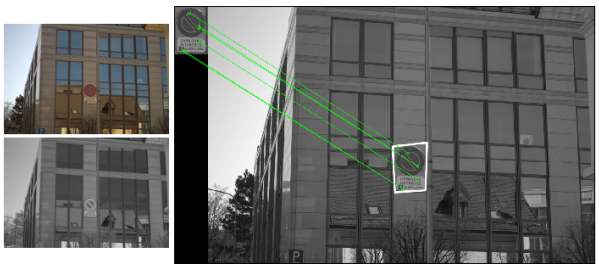

Fig. 5. Matching a template for object detection. Information from all four channels (RGB+NIR, shown at top left and bottom respectively) is taken into account.

6. CONCLUSION AND FUTURE WORK

We have presented an extension of the Harris detector that can take advantage of the content and cross-correlations of multiple-channel inputs, unlike standard hand-crafted detectors. As future work, we plan to work on coupling with a quaternionic analysis-based descriptor, examining uses with a learning-based scheme (e.g. HardNet [5]) or researching extensions to more complex algebras such as octonions [19]. The potential of using quaternionic eigenvectors as a basis for a better detector or an image cue will also be researched, as well as other application perspectives [20, 21, 22].

7. REFERENCES

[1] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, "A comparison of affine region detectors," *International journal of computer vision*, vol. 65, no. 1-2, pp. 43–72, 2005.

[2] E. Riba, D. Mishkin, D. Ponsa, E. Rublee, and G. Bradski, "Kornia: an open source differentiable computer vision library for pytorch," 2019.

[3] S.J.D. Prince, *Computer vision: models, learning, and inference*, Cambridge University Press, 2012.

[4] Axel Barroso Laguna, Edgar Riba, Daniel Ponsa, and Krystian Mikolajczyk, "Key. net: Keypoint detection by handcrafted and learned cnn filters," *arXiv preprint arXiv:1904.00889*, 2019.

[5] Anastasiia Mishchuk, Dmytro Mishkin, Filip Radenovic, and Jiri Matas, "Working hard to know your neighbor's margins: Local descriptor learning loss," in *Advances in neural information processing systems (NIPS)*, 2017, pp. 4826–4837.

[6] Jérôme Revaud, César De Souza, Martin Humenberger, and Philippe Weinzaepfel, "R2D2: Reliable and repeatable detector and descriptor," in *Advances in neural information processing systems (NIPS)*, 2019, pp. 12405–12415.

[7] Johannes L Schonberger, Hans Hardmeier, Torsten Sattler, and Marc Pollefeys, "Comparative evaluation of hand-crafted and learned local features," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1482–1491.

[8] D. Firmenichy, M. Brown, and S. Süsstrunk, "Multispectral interest points for RGB-NIR image registration," in *2011 18th IEEE International Conference on Image Processing*. IEEE, 2011, pp. 181–184.

[9] Giorgos Sfikas, João Patacas, Charalampos Psarros, Antigoni Noula, Dimosthenis Ioannidis, and Dimitrios Tzovaras, "A deep neural network-based method for the detection and accurate thermography statistics estimation of aerially surveyed structures," in *19th International Conference on Construction Applications of Virtual Reality*, 2019.

[10] Giorgos Sfikas, Christian Heinrich, Jihad Zallat, Christophoros Nikou, and Nikos Galatsanos, "Recovery of polarimetric stokes images by spatial mixture models," *JOSA A*, vol. 28, no. 3, pp. 465–474, 2011.

[11] Fuzhen Zhang, "Quaternions and matrices of quaternions," *Linear algebra and its applications*, vol. 251, pp. 21–57, 1997.

[12] Todd A Ell and Stephen J Sangwine, "Hypercomplex fourier transforms of color images," *IEEE Transactions on image processing*, vol. 16, no. 1, pp. 22–35, 2007.

[13] Lianghai Jin, Enmin Song, Lei Li, and Xiang Li, "A quaternion gradient operator for color image edge detection," in *2013 IEEE International Conference on Image Processing*. IEEE, 2013, pp. 3040–3044.

[14] Takehiko Mizoguchi and Isao Yamada, "Hypercomplex tensor completion with Cayley-Dickson Singular Value Decomposition," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 3979–3983.

[15] Nicolas Le Bihan and Jérôme Mars, "Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing," *Signal processing*, vol. 84, no. 7, pp. 1177–1199, 2004.

[16] Dong Cheng and Kit Ian Kou, "Properties of Quaternion Fourier transforms," *Preprint*, 2016.

[17] M. Boyle, "Quaternion: Add built-in support for quaternions to numpy," 2020.

[18] Matthew Brown and Sabine Süsstrunk, "Multi-spectral SIFT for scene category recognition," in *CVPR 2011*. IEEE, 2011, pp. 177–184.

[19] Hong-Yun Gao and Kin-Man Lam, "Salient object detection using octonion with bayesian inference," in *2014 IEEE international conference on image processing (ICIP)*. IEEE, 2014, pp. 3292–3296.

[20] Dimitrios Tzovaras, Nikitas Karagiannis, and Michael G Strintzis, "Robust image watermarking in the subband or discrete cosine transform domain," in *9th European Signal Processing Conference (EUSIPCO 1998)*. IEEE, 1998, pp. 1–4.

[21] Dimitrios Tzovaras, Ioannis Kompatsiaris, and Michael G Strintzis, "3D object articulation and motion estimation in model-based stereoscopic videoconference image sequence analysis and coding," *Signal Processing: Image Communication*, vol. 14, no. 10, pp. 817–840, 1999.

[22] Giorgos Sfikas, Christophoros Nikou, Nikolaos Galatsanos, and Christian Heinrich, "Majorization-minimization mixture model determination in image segmentation," in *CVPR 2011*. IEEE, 2011, pp. 2169–2176.