35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Impact of Image Preprocessing Methods on Polyp Localization in Colonoscopy Frames

Jorge Bernal[1], Javier Sánchez[1] and Fernando Vilariño[1]

Abstract— **In this paper we present our image preprocessing methods as a key part of our automatic polyp localization scheme. These methods are used to assess the impact of different endoluminal scene elements when characterizing polyps. More precisely we tackle the influence of specular highlights, blood vessels and black mask surrounding the scene. Experimental results prove that the appropriate handling of these elements leads to a great improvement in polyp localization results.**

I. INTRODUCTION

Colorectal cancer is the third most common cancer in incidence and the fourth most common cause of cancer death worldwide. Its survival rate decreases the later it is detected [1], hence the importance of colon screening techniques such as colonoscopy. Although colonoscopy is still the gold standard for colon screening, it has some drawbacks being polyp miss-rate (reported to be as high as 6% [2]) the most relevant problem. The work presented in this paper is enclosed into the field of intelligent systems for colonoscopy which aim at providing additional information to the colonoscopy procedure. More precisely we are focused on the development of automatic polyp localization methods, which still nowadays present several difficulties.

We present here the first study that takes into account the impact of different endoluminal scene elements in polyp localization results. We will address the influence of specular highlights, blood vessels and the black mask that surrounds the endoluminal scene.

Our automatic polyp localization method integrates valley information to locate the polyp. We must discern between valley information that comes from polyps and the one that is related to other elements in order to improve polyp localization results. The novelty of the work presented is the assessment of the impact that different elements of the endoluminal scene have on polyp localization results, as the three of them are also source of valley information which can affect the performance of our algorithms.

The structure of the paper is as follows: in Section II we introduce previous approaches on polyp characterization. We introduce our model of appearance for polyps in Section III. Our polyp localization method, in which the image preprocessing methods are enclosed, is presented in Section IV. In Section V we show our experimental setup along with image preprocessing and polyp localization results. Finally we finish this paper in Section VI with the main conclusions extracted along with proposals for future work.

[1]J. Bernal, J. Sánchez and F. Vilariño are from Computer Vision Center and Computer Science Department, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

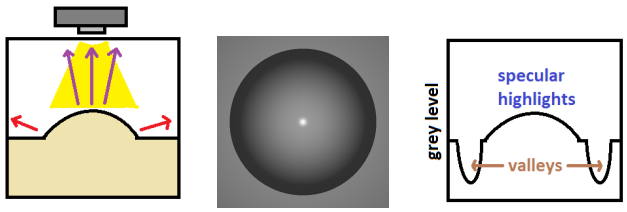
II. RELATED WORK

The description and characterization of endoluminal scene elements has been recently addressed in the literature. Although the majority of the available works are related to polyp characterization, there are also some works related to other elements such as the lumen [3] or specular highlights [4]. Other works devoted to either enhance the quality of colonoscopy frames or to discard low quality frames [3].

Related to polyp characterization, we can divide the existing bibliography [3] into three separate groups: 1) Shape; 2) Texture and 3) Color. Related to shape, we can also make a subdivision between two different types of methods: based on the curvature of the boundaries [5] or based on shape fitting [6]. There is a number of works on the field of texture description using specific texture descriptors such as wavelets *wavelet descriptors*, local binary patterns or co-ocurrence matrices [7]. The work of [8] presents MPEG-7 texture and color descriptors used in polyp characterization methods. Our previous work [9] departs from this specific approaches by building a general model of polyp appearance which takes into account both the processes of image acquisition and image generation. This model defines polyps delimited by boundaries corresponding to valleys in the intensity image as explained in the next section.

III. MODEL OF APPEARANCE FOR POLYPS

In order to define our model of appearance for polyps we use an a priori model about the polyp and a model of the illumination. In this case, for the sake of simplicity we consider polyps as semi-spherical shapes that protrude from the colon wall plane [9]. We also consider that polyp's surface is homogeneous and its reflectance can be approximated by the Phong's illumination model [10]. We can model the colonoscope by a pinhole camera and a punctual illumination source placed in the same position. As can be seen in Figure



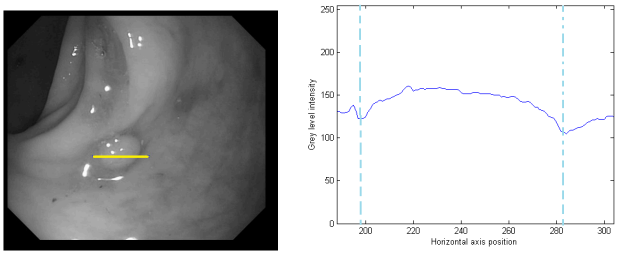
(a) (b) (c)

Fig. 1. Model of appearance and illumination of polyps: (a) Graphical representation of an illuminated prominent surface (a polyp); (b) Synthetic model rendering of a polyp (c) Corresponding grey level profile of both graphical and synthetical representations of the polyp .

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

1 the center of protruding objects, such as polyps, reflects the incident light back to the camera but lateral surfaces reflect light outside the camera.

Taking this into account, the characterization of the polyp is obtained through the shadings related to valleys in the intensity image as can be seen in Figure 2 where we intensity valleys constitute the boundary of the polyp. The detection of polyps is thus linked to the identification of the valleys that constitute their boundaries.



(a) (b)

Fig. 2. Presence of intensity valleys in polyp boundaries: (a) Original image; (b) Grey level profile under the yellow line drawn in (a).

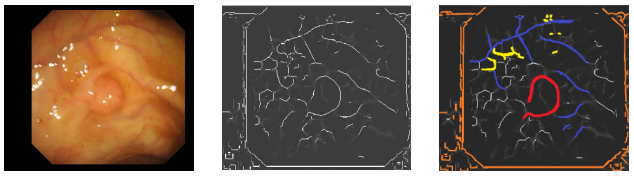
In order to obtain valley information we use the Multilocal Creaseness valley detector that was presented in [11] - other alternatives such as *Second Derivative of Gaussians* [12] could provide equivalent results -. In our case we opt to use the former since its output is more geometrical, leading to eliminate the response that non-desired structures may provide. Our valley detector is good at localizing the valleys in the image but it fails in terms of quantifying them, as its output is somewhat binary. In order to solve this the output of a valley detector (V) can be multiplied by the output of the morphological gradient (M_{grad}) to generate a Depth Of Valleys image $DoV = V \cdot M_{grad}$ [9]. DoV image will ideally present high values in pixels that constitute polyp boundary. We use this DoV image as the source of our polyp localization algorithm but, as it will be seen in the next section, we need to eliminate non-polyp valley information to make our algorithms perform robustly.

IV. Methodology

Our polyp localization processing scheme consists of three different stages: 1) Image preprocessing; 2) Depth of Valleys Accumulation (DOVA) energy maps and 3) Final polyp location from the maxima of DOVA energy map.

A. Image Preprocessing

Unfortunately, as can be seen in Figure 3, polyps are not the only source of valley information. We study four different sources of valley information in colonoscopy images: 1) Polyps; 2) Specular highlights; 3) Black mask and 4) Blood vessels. As our polyp localization method uses valley information to give its output, it is necessary to address the effect of the different non-polyp sources to ease later processing stages.

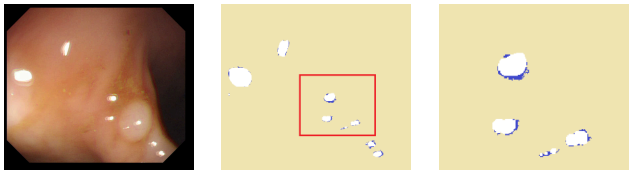


(a) (b) (c)

Fig. 3. Valley information sources: (a) Original image; (b) Valley image; (c) Manually-marked valley image. Marked valleys are from polyps (red), blood vessels (blue), specular highlights (yellow) and black mask (orange).

a) Specular highlights: The impact of specular highlights is twofold: we can have valley information within the specular highlight area, and we can also have this information around the specular highlight. We apply two different operations: 1) detection and 2) inpainting.

1) Detection: our method extends current state-of-the-art in colonoscopy videos [4] (Figure 4 (c)). In our case, we are concerned on those pixels that are suspected to be part of a specular highlight but they can not be easily identified. We assume that the intensity value inside the specular highlight is higher than its surroundings and pixels nearby to specular highlights will continue having higher intensity values, although smaller than inside the specularity. We find these pixels by calculating the difference between the original image and its median (not considering pixels already part of specular highlight) so we can obtain which pixels in the image have a intensity value marginally higher than its neighborhoods. Then, by means of a threshold value, we keep only those where the difference is higher.



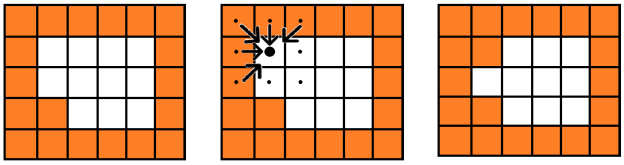
(a) (b) (c)

Fig. 4. Specular highlights detection: (a) Original image; (b) Extension of the detection obtained with [4]; (c) Zoom of the red square area in the detection mask. True positive pixels are painted in white and those TP pixels that we detect with our method that were not detected by [4] in blue.

2) Inpainting: The inpainting method consists of two different stages:

a) Diffusion: In this stage we diffuse values from the original image into pixels with no value which are under the detection mask M. We track the positions of the pixels under M and, for each of them we perform as follows: we obtain a 3×3 neighborhood around the pixel and change its original value by the mean value of the valid neighbors. Valid neighbours are those pixels which either do not belong to the original M mask or that have already being modified by the diffusion process. This process is repeated until every pixel under M has a new value. Once this happens, we repeat the process until the difference between the new and the previous value of pixels under M is smaller than a threshold values_{th}. A graphical example of the diffusion algorithm is shown in Figure 5. We can see that for the calculation of the diffused

value of the pixels under M, which are painted in white, we only use information from valid neighbors, painted in orange in the image.



(a) (b) (c)

Fig. 5. Example of the diffusion stage of the inpainting algorithm for one pixel: (a) Representation of the initial stage of the diffusion algorithm: pixels under M mask are painted in white whereas pixels outside M are painted in orange; (b) Calculation of the new value from valid neighbors, and (c) Image with new value obtained.

The complete diffusion algorithm is:

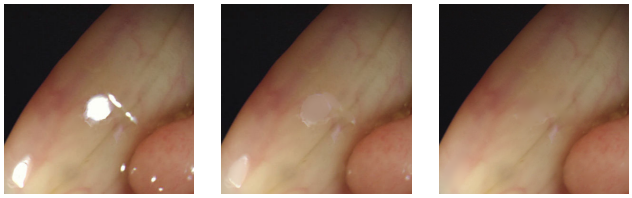
Algorithm 1: Inpainting diffusion algorithm

Data: Diffusion(I,FM,MC)
Arg:(I: original image, M: detection mask, MC: minimum change threshold)
Result: Diffused image(I_d)
Initialization of valid neighbors mask;
$VNM = \neg M$;
Calculation of diffused values for pixels in M;
repeat while the image is modified over s_{th}
  $stop = true$;
  forall the $\vec{x} \in I : M(\vec{x}) == 1$ **do**
    Definition of a neighborhood around a pixel;
    $Neigh = \{\vec{p} | \vec{p} \in Neighborhood(\vec{x}), VNM(\vec{p}) == 1\}$;
    if $\#Neigh > 0$ **then**
      Calculation of the diffused value;
      $nv = \frac{\sum_{\vec{p} \in Neigh} I_d(\vec{p})}{\#Neigh}$;
      Calculation of the stop flag;
      if $VNM(\vec{x}) == 1$ **then**
        if $|nv - I_d(\vec{x})| > s_{th}$ **then**
          $stop = false$;
      else
        $stop = false$;
      end
      Actualization of the diffused image value;
      $I_d(\vec{x}) = nv$;
    end
  end
until $stop == true$;

b) Obtention of the final inpainted image: To create the final inpainted image we also have to consider that if we do a direct substitution of the pixels under M there will still remain a clear frontier between pixels inside and outside the final image, as happens with the method explained in [4] (an example of this is shown in Figure 6 (b)). In order to solve this we create an extended mask which ponders the way we combine the original image I_o and the diffused I_d in the final inpainted I_{np} image. This extended mask M_1 is created by dilating the original M mask with a circular structural element and later convolving the result with a gaussian kernel. Once this mask is obtained the final inpainted image I_{np} is calculated as:

$$I_{np} = M_1 \cdot I_o + (1 - M_1) \cdot I_d \quad (1)$$

where $I_o(x, y)$ and $I_d(x, y)$ respectively correspond to the original image and the diffused image. In pixels under M mask, the intensity values are completely replaced by their corresponding values in the I_d. On the other hand, as we depart from the original M mask, the contribution of the original I_o values increases. An example of the final inpainted image can be seen in Figure 6 (c).



(a) (b) (c)

Fig. 6. Comparison of specular highlights inpainting results: (a) Original image; (b) Inpainted impage obtained by using method explained in [4]; (c) Inpainted image obtained with the method proposed in this paper.

b) Black mask: Colonoscopy video frames are natively acquired with a black mask (see Figure 3 (a)). The borders of the black mask do generate valley information, as can be seen in Figure 3 (c). In order to cope with this problem we have two alternatives, either to crop the image and analyze what is within the limits of the black mask, or to do an inpainting below the black mask. In our case we have opted for the second because by cropping we loose boundary's information, potentially showing polyp content.

c) Blood vessels: Blood vessels segmentation is a complicated task out of the scope of our current research, but it is possible to mitigate their impact in terms of valley information. In this case we are interested in finding a color channel that both enhances polyp boundary information while mitigating blood vessels. We have explored the use of several color spaces such as sRGB, HSV or CieLab and all the possible combinations (including channel subtraction) within a given color channel.

B. DOVA energy maps

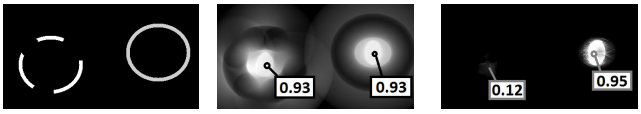
Once we have a cleaner valley image after the application of our image preprocessing methods, we are able to impose more restrictions in the final stage of our polyp localization method. We defined in our previous work [9] Sector Accumulation Depth of Valley Accumulation (SA-DOVA) energy maps integrating DoV information, which combined valley localization provided by a valley detector with a better quantization provided by morphological gradient. SA-DOVA was built by placing a series of radial sectors centred on each pixel and summing the maxima of DoV image under each sector. The rationale behind this approach was that

pixels inside the polyp should be surrounded by boundaries constituted by pixels with high value of the DoV image.

SA-DOVA is heavily affected by how the DoV image is in a way such if the source image is clean it works as suspected, but under the presence of abundant valley information related to non-polyp elements of the scene, its performance gets damaged. A simple but effective solution to eliminate noise and also benefit circularity is the use of median in the accumulation process instead of the sum of the maxima under each sector (See Eq(2)). The novel Median SA-DOVA (MSA-DOVA) is calculated as follows:

$$\begin{aligned} MaxL(\vec{x},\alpha) &= \max_{r}\{DoV(\vec{x}+r*(\cos(\alpha),\sin(\alpha)))\}, \\ Acc_{MSA}(\vec{x}) &= \operatorname*{Med}_{\alpha}(MaxL(\vec{x},\alpha)), \\ Acc_{SA}(\vec{x}) &= \sum_{\alpha}(MaxL(\vec{x},\alpha)), \end{aligned} \tag{2}$$

where $\alpha \in [0, 2\pi]$ and $r \in [R_{min}, R_{max}]$, R_{min} and R_{max} correspond respectively to the minimum and maximum radius of the sectors used in the accumulation process. We show in Figure 7 a qualitative comparison between the results obtained by SA-DOVA and MSA-DOVA. As we use the position of the maxima of the energy map to locate the polyp we will be interested in having a higher value associated to polyp boundaries than to other structures in the image. We can see in Figure 7 (b) that by using SA-DOVA there is no difference in terms of maxima of accumulation between a non-continuous structure composed by a few pixels with high DoV value (maxima value: 0.93) and a continuous structure composed by more pixels with smaller DoV value (maxima value: 0.93). By changing from sum-based to median-based accumulation we keep a similar maxima value under the continuous structure (0.95) but we almost eliminate accumulation inside the non-continuous one (maxima value: 0.12).



(a) (b) (c)

Fig. 7. Impact of non-continuous boundaries with high DoV value on the output of SA-DOVA energy maps: (a) Original image; (b) SA-DOVA energy map; (c) Median-DOVA energy map.

V. Experimental Results

In order to assess the performance of our methods we will use the, up to our knowledge, only available labeled database of colonoscopy videos, which was introduced in [9]. In this case, as in the original paper, we will only use a subset of 300 frames from the database, as some of the videos either content too much fecal content or do not have enough quality to be analyzed. In order to tune the parameters for the MSA-DOVA we used 30 images, different from the ones in the test database and we selected as final parameter values those which yielded better localization results. More precisely we set the minimum and maximum radii to 30 and 130 respectively and we also fixed to 180 the number of radial sectors used in the accumulation process.

We will perform two different experiments: a) Evaluation of the different image preprocessing methods, and b) assessment of the impact of image preprocessing in polyp localization results.

A. Image preprocessing results

For the case of **specular highlights detection** we will compare the performance of our method with two general state-of-the-art approaches [13], [14] and with the method we base our approach on [4]. In this case we use as a metric the Detection Rate (DR), which is defined as the percentage of specular highlights pixels that have been detected by each method. Results from Table I show that our contribution improves the state-of-the-art in specular highlight detection [4]. The errors in our approach are caused by failing on detecting some pixels close to the specular highlight but not part of it.

Method	Yang et al	Yoon et al	Arnold et al	Bernal et al
Detection rate %	53.04%	42.12%	81.44%	84.20%

TABLE I
Comparison of specular highlight detection methods.

For the case of **specular highlights inpainting** we will only compare our method with current state-of-the-art in colonoscopy [4] and in this case we will measure the ratio between the valley energy around the specular highlight before (E_0) and after the inpainting (E_{inp}). Both methods have the same M mask as input. As can be seen from Table II, by using our method we improve the mitigation of specular highlights-originated valleys.

Method	E_0	E_{inp}	$\%(E_{inp}/E_0)$
Arnold et al	1083.99	574.38	52.98%
Bernal et al	1083.99	445.84	41.13%

TABLE II
Comparison of specular highlights inpainting methods.

The objective of **blood vessels mitigation** experiment is to test, in terms the two low-level image processing algorithms used to generate the DoV image (valley detection and morphological gradient), if we can mitigate the effects of blood vessels in the image without losing polyp information. We selected and annotated a subset of 29 images with high blood vessel presence to test our mitigation method. We measure, for each input image, the relative difference (%) in energy under both vessels and polyp contour masks from the original value obtained from the grey scale image. We summarize the main results obtained in Table III. We can observe from the table that by using the B channel of the sRGB image we mitigate blood vessels' energy while enhancing polyp contours for both valley detection and morphological gradient.

Measure	Polyp	Vessel	Difference
Morph. gradient	9.21%	−0.44%	9.65%
Valley detection	20.33%	−6.61%	26.95%
Depth of valleys	1.87%	−0.02%	1.89%

TABLE III

BLOOD VESSELS MITIGATION RESULTS BY USING CHANNEL B OF THE SRGB IMAGE

B. Polyp localization results

We show a graph which presents the polyp localization results based on the application of image preprocessing methods in Figure 8:

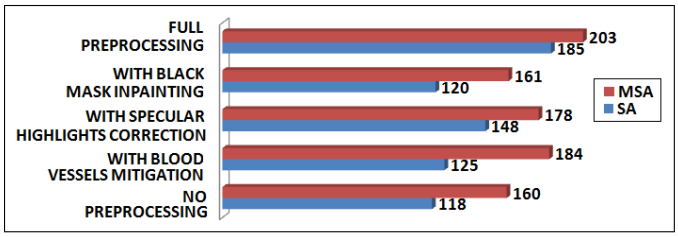
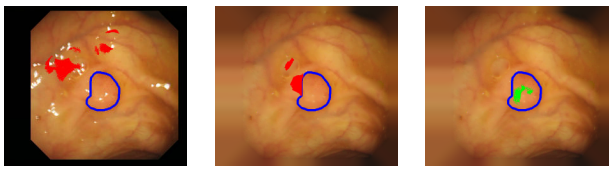


Fig. 8. Break down of the number of images with correct polyp localization (total of images: 300) according to the preprocessing applied to the original image and how the accumulation is performed.

There are several conclusions that can be extracted from this Figure:

1) The preprocessing method that has more impact on polyp localization results is blood vessels mitigation, followed by specular highlights correction.
2) By applying all the preprocessing methods to the original image and by changing the accumulation method we improve our polyp localization results in almost 90 images (30%) - see Figure 9)-.
3) The change of summing-based to median-based accumulation results on an improvement of polyp localization results applied on the greyscale image.



(a) (b) (c)

Fig. 9. Maxima of accumulation both inside (green) and outside (red) polyp contour (blue) for: (a) No preprocessed image and SA-DOVA; (b) Full preprocessed image and SA-DOVA; (c) Full preprocessed image and MSA-DOVA;

VI. CONCLUSIONS AND FUTURE WORK

Considering that valley information is the source of information of our polyp localization algorithm, in this paper we tackled the impact of several elements of the endoluminal scene (specular highlights, blood vessels and black mask) that also give valley response. For the case of specular highlights we introduced our detection and inpainting algorithms that improve general and specific approaches. Our preliminar study on blood vessels mitigation shows that the use of the B channel of the sRGB image leads to a decrease in the valley information related to blood vessels. Finally we applied the same inpainting method to the black mask superimposed to the endoluminal scene. Once preprocessing is done we apply a simple but effective improvement to a previous iteration of DOVA energy maps as the last step of the polyp localization algorithm. The experimental results show that all the three preprocessing methods have an impact on the overall performance on polyp localization methods although blood vessels mitigation and specular highlights correction are the techniques that lead to a better improvement. By means of the proposed preprocessing we improve polyp localization results in almost 30% of the images, which confirms its necessity.

The future work will involve the consideration of the rest of elements of the endoluminal scene such as the lumen or wrinkles and folds along with a development of a future scale-space implementation of the DOVA algorithm.

ACKNOWLEDGMENTS

This work was supported in part the Spanish Government through the founded projects "COLON-QA" ($TIN2009-10435$) and "FISIOLOGICA" ($TIN2012-33116$).

REFERENCES

[1] C. Eng, P. Lynch, and J. Skibber, "Colon cancer," in *60 Years of Survival Outcomes at The University of Texas MD Anderson Cancer Center*. Springer, 2013, pp. 77–84.
[2] B. Bressler, L. Paszat, Z. Chen, D. Rothwell, C. Vinden, and L. Rabeneck, "Rates of new or missed colorectal cancers after colonoscopy and their risk factors: A population-based analysis," *Gastroenterology*, vol. 132, no. 1, pp. 96–102, 2007.
[3] J. Bernal, F. Vilariño, and J. Sánchez, *Colonoscopy Book 1: Towards Intelligent Systems for Colonoscopy*. In-Tech, 2011.
[4] M. Arnold, A. Ghosh, S. Ameling, and G. Lacey, "Automatic segmentation and inpainting of specular highlights for endoscopic imaging," *Journal on Image and Video Processing*, vol. 2010, p. 9, 2010.
[5] H. Zhu, Y. Fan, and Z. Liang, "Improved Curvature Estimation for Shape Analysis in Computer-Aided Detection of Colonic Polyps," *Beijing, China*, p. 19, 2010.
[6] S. Hwang, J. Oh, and W. Tavanapong, "Polyp detection in colonoscopy video using elliptical shape feature," in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, vol. 2. IEEE, 2007, pp. II–465.
[7] S. Ameling, S. Wirth, D. Paulus, G. Lacey, and F. Vilariño, "Texture-based polyp detection in colonoscopy," *Bildverarbeitung für die Medizin 2009*, pp. 346–350, 2009.
[8] M. Coimbra and J. Cunha, "MPEG-7 visual descriptors and contributions for automated feature extraction in capsule endoscopy," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 16, no. 5, pp. 628–637, 2006.
[9] J. Bernal, J. Sánchez, and F. Vilariño, "Towards automatic polyp detection with a polyp appearance model," *Pattern Recognition*, 2012.
[10] B. T. Phong, "Illumination for computer generated pictures," *Communications of ACM*, vol. 18, no. 6, pp. 311–317, 1975.
[11] A. López and F. Lumbreras, "Evaluation of methods for ridge and valley detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, pp. 327–335, 1999.
[12] D. Gil *et al.*, "Structure-preserving smoothing of biomedical images," in *Computer Analysis of Images and Patterns*. Springer, 2009, pp. 427–434.
[13] Q. Yang, S. Wang, and N. Ahuja, "Real-time specular highlight removal using bilateral filtering," *Computer Vision–ECCV 2010*, pp. 87–100, 2010.
[14] K. Yoon and I. Kweon, "Correspondence search in the presence of specular highlights using specular-free two-band images," *Computer Vision–ACCV 2006*, pp. 761–770, 2006.