



# Efficient Enhancement of Stereo Endoscopic Images Based on Joint Wavelet Decomposition and Binocular Combination

Bilel Sdiri, Mounir Kaaniche, Faouzi Alaya Cheikh, Azeddine Beghdadi, Ole Jakob Elle

## ► To cite this version:

Bilel Sdiri, Mounir Kaaniche, Faouzi Alaya Cheikh, Azeddine Beghdadi, Ole Jakob Elle. Efficient Enhancement of Stereo Endoscopic Images Based on Joint Wavelet Decomposition and Binocular Combination. IEEE Transactions on Medical Imaging, 2019, 38 (1), pp.33-45. 10.1109/TMI.2018.2853808 . hal-04069232

**HAL Id: hal-04069232**

**<https://hal.science/hal-04069232>**

Submitted on 3 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Efficient enhancement of stereo endoscopic images based on joint wavelet decomposition and binocular combination

Bilel Sdiri, Mounir Kaaniche, *Member IEEE*, Faouzi Alaya Cheikh, *Senior Member IEEE*, Azeddine Beghdadi, *Senior Member IEEE*, and Ole Jakob Elle

**Abstract**—The success of minimally invasive interventions and the remarkable technological and medical progress have made endoscopic image enhancement a very active research field. Due to the intrinsic endoscopic domain characteristics and the surgical exercise, stereo endoscopic images may suffer from different degradations which affect its quality. Therefore, in order to provide the surgeons with a better visual feedback and improve the outcomes of possible subsequent processing steps, namely a 3D organ reconstruction/registration, it would be interesting to improve the stereo endoscopic image quality. To this end, we propose in this paper two joint enhancement methods which operate in the wavelet transform domain. More precisely, by resorting to a joint wavelet decomposition, the wavelet subbands of the right and left views are simultaneously processed to exploit the binocular vision properties. While the first proposed technique combines only the approximation subbands of both views, the second method combines all the wavelet subbands yielding an inter-view processing fully adapted to the local features of the stereo endoscopic images. Experimental results, carried out on various stereo endoscopic datasets, have demonstrated the efficiency of the proposed enhancement methods in terms of perceived visual image quality.

**Index Terms**—Endoscopic images, stereo pairs, wavelet transform, enhancement, joint processing, binocular vision.

## I. INTRODUCTION

### A. Medical Context

MINIMALLY invasive surgery (MIS) consists in using techniques that limit the size and number of cuts during the operation. Indeed, unlike conventional open surgery that relies on making large incisions to get a direct access to the surgical target, MIS is performed through tiny incisions not exceeding 1.5 cm by which the surgeons pass the surgical instruments and a long thin tube equipped with miniature camera, called an endoscope, to visualize the operating field. In the last decades, MIS has made remarkable medical and technological progress and became a popular diagnosis and treatment tool due to its benefits which include: decreasing significantly the patient trauma and morbidity, reducing the

length of hospital stay and recovery time, and yielding less scarring.

In a traditional laparoscopic/endoscopic surgery with a single two-dimensional (2D) visual system, different challenges are facing the surgeons, namely the restricted field of view and the loss of both tactile feedback and depth perception which are important for the navigation and to perform the surgical tasks. To solve the latter problem, three dimensional (3D) vision surgical systems [1], based on stereoscopic images, have been developed. Indeed, this data is a combination of left and right images taken from two different viewpoints. One main advantage of using stereo endoscopic images (SEI) is their ability to extract the 3D information related to the anatomical structure of the organs. This information is then used to perform accurate model-to-patient registration and improve surgery planning such as the resection surface in a *Hepatectomy*. Moreover, having good depth perception of the endoscopic scene allows the surgeons to distinguish easily the relative positions of the organs and surgical instruments, and so, improve the navigation conditions. Due to the intrinsic characteristics of the endoscopic environment including dynamic illumination conditions and moist tissues with high reflectance, SEI may suffer, however, from different degradations such as specular reflection, noise, low contrast and inhomogeneous illumination, which may affect some subsequent processing steps like segmentation, feature extraction/detection [2] and 3D reconstruction [3], [4], [5]. Therefore, it becomes necessary to enhance the quality of SEI.

### B. Related Works

Generally, endoscopic image quality can be improved by enhancing their contrast/sharpness or removing the surgical tasks artifacts (e.g. surgical smoke [6]). To this end, different techniques have been reported in the literature. More precisely, the enhancement of endoscopic images has been mainly studied in the case of *single* images (i.e one 2D view). In this context, enhancement methods based on retinex theory have been proposed [7], [8], [9]. While a single scale retinex approach is used in [7], [8], the method developed in [9] resorts to a multiscale retinex approach [10]. However, such methods present two main drawbacks. First, setting the retinex parameters is a difficult task. Moreover, they may lead to partial enhancement of the image. This problem can occur especially if the image illumination is not well

B. Sdiri, M. Kaaniche and A. Beghdadi are with L2TI-Institut Galilée, Université Paris 13, Sorbonne Paris Cité, 99 avenue J. B. Clément, Villeteuse, France, 93430. E-mail: bilel.sdiri@univ-paris13.fr, mounir.kaaniche@univ-paris13.fr, azeddine.beghdadi@univ-paris13.fr. F. A. Cheikh is with the Norwegian Colour and Visual Computing Lab, Norwegian University of Science and Technology; Teknologivgen 22, 2815 Gjøvik, Norway. E-mail: faouzi.cheikh@ntnu.no. O.J. Elle is with the The Intervention Centre, Oslo University Hospital; Sognsvannsveien 20, 0372 Oslo, Norway. E-mail: oelle@ous-hf.no

distributed and presents both over and under saturated regions. As a result, other medical image enhancement methods which adjust the coefficients using specific mapping functions, neural networks [11], [12], or filtering techniques [13], [14], have been designed. For instance, in [15], the luminance component is adjusted using an adaptive sigmoid function and new chrominance components are generated based on the texture information. In [16], the authors propose two simple enhancement methods and study their impact on feature tracking. The first technique is based on histogram equalization and the second one uses a morphological operator to enhance the local contrast. A similar work is proposed in [17]. The endoscopic images are first processed using a median filter to highlight the image details. Then, an histogram information correction is performed to adjust the dynamic range of the image.

It is worth pointing out that unlike the two-dimensional endoscopic images case, only *few* research works have been devoted for stereo endoscopic image enhancement [18], [19]. For instance, Hai *et al.* [18] propose to apply contrast limited adaptive histogram equalization (CLAHE) to the left and right views of the SEI. CLAHE is a contrast enhancement technique aiming at computing local histograms in different regions of the image to avoid some drawbacks such as noise amplification in homogeneous areas. In [19], an adaptive contrast enhancement method combining depth data and edge information is proposed. To avoid any over enhancement, the inter-view luminance differences are then processed based on a binocular just noticeable difference (BJND) model [20]. The latter measures the minimal distortion or noise in one of the stereo images that can trigger a noticeable visual difference when binocularly combined with the other view by the human visual system (HVS). The application of the BJND model for sharpness enhancement of stereo images has already been performed as well in [21]. In addition to the endoscopic context, other research works related to the enhancement of natural stereo images are found in the literature [21], [22], [23], [24]. Indeed, an adaptive sharpness enhancement algorithm is developed in [22]. The effect of depth perception is studied and a concept of just noticeable blur model is introduced. In [23], a contrast enhancement method is designed by combining local edges of both views and depth information. In [24], the left view is decomposed using a wavelet transform (WT) and the resulting subband coefficients are adjusted using a 2D contrast-measure-based enhancement technique in the wavelet domain [25]. The enhanced right image is then obtained by simply performing a compensation of the enhanced left view using the disparity map. Thus, such approach is not so efficient since the inter-view redundancies are only exploited to generate the enhanced right view whereas the left image is enhanced independently. It should be noted that natural stereo images enhancement techniques have been also reviewed since they may be applied to endoscopic data and used especially for comparison purpose as it will be performed in Section IV.

### C. Contributions

In this paper, we propose to further investigate the quality enhancement of SEI. The proposed methods operate in the

transform domain and rely on a joint wavelet decomposition and a cross-view processing to exploit both depth information and the inter-view redundancies of stereo images. More precisely, an efficient multiscale decomposition, based on Vector Lifting Scheme (VLS) [26], is first applied to both views. Note that, unlike conventional lifting scheme, the VLS is a joint wavelet decomposition that aims at exploiting the inter-view correlation to generate two compact multi-resolution representations of the left and right images. Then, the approximation and detail subbands of both views are separately adjusted using an adaptive parameterized mapping function as performed in conventional 2D enhancement techniques [27], [28]. After that, the resulting intermediate enhanced views, referred to as pre-enhanced views, are jointly processed to exploit the binocular vision features of the HVS, namely the binocular rivalry. To this end, inspired by HVS-based perceptual studies [29], [30], [31], a binocular combination of the left and right processed views is performed to generate the final enhanced images. Note that such binocular combination process aims at synthesizing a cyclopean image which is close to what is formed in the human visual cortex based on two separate left and right retinal images [29], [30]. To the best of our knowledge, this paper presents the first research work which exploits a cyclopean image model for the enhancement of stereo images. To perform this joint processing step, two stereo combination models based on local image features, including energy and contrast measures, are given. Finally, the inverse VLS-based decomposition is applied to the resulting wavelet representations to reconstruct the enhanced stereo endoscopic images.

The remainder of this paper is organized as follows. In Section II, the VLS-based joint decomposition applied to the left and right views is presented. The proposed stereo adaptive inter-view processing techniques are described in Section III. Finally, experimental results are given in Section IV and some conclusions are drawn in Section V.

## II. SYMMETRIC JOINT WAVELET DECOMPOSITION

Wavelets have been attracting much attention in various image processing applications due to their intrinsic properties: multiscale representation and good space-frequency localization. Among these applications, we mention super resolution [32], denoising [33], enhancement [34], and pattern/shape recognition [35]. For this reason, we propose to enhance SEI in the wavelet transform domain. To this respect, two wavelet representations of the left and right images are first generated. While classical (i.e 2D) wavelet transformations have already been employed in recent still and stereo image enhancement [24], [28], we resort here to a joint wavelet decomposition based on the concept of VLS [26]. This later has the advantage of allowing the exploitation of inter-view redundancies and correlation thanks to the estimated disparity maps.

By assuming that the disparity map is obtained by using the standard correlation-based disparity estimation (DE) technique, we focus on the description of the joint wavelet

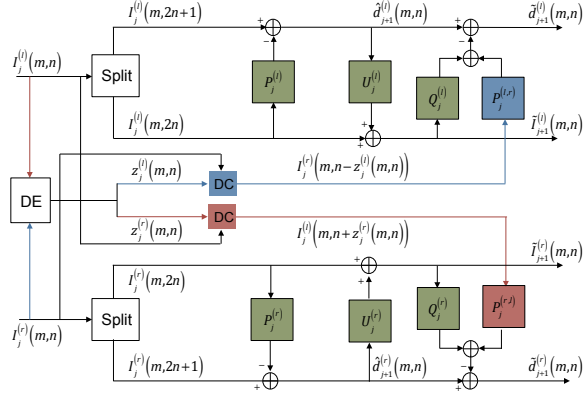


Fig. 1: VLS-based joint wavelet decomposition. Note that the green color is employed to show the intra-prediction and update filters used for each view. However, the red (resp. blue) color is employed to show the inter prediction step in the right (resp. left) view based on the information coming from the other left (resp. right) view.

decomposition. The analysis structure of this decomposition is shown in Fig. 1. For both left  $I_j^{(l)}$  and right  $I_j^{(r)}$  images at resolution level  $j$ , a lifting stage composed of Predict-Update-Predict (PUP) steps is applied to each line  $m$ . To describe these steps, we consider in the following the wavelet decomposition of the right image. For instance, an intermediate detail signal  $\hat{d}_{j+1}^{(r)}$  (at level  $j+1$ ) is obtained thanks to the first prediction step as follows:

$$\hat{d}_{j+1}^{(r)}(m, n) = I_j^{(r)}(m, 2n+1) - \sum_{k \in \mathcal{P}_j^{(r)}} p_{j,k}^{(r)} I_j^{(r)}(m, 2n-2k), \quad (1)$$

where  $\mathcal{P}_j^{(r)}$  and  $p_{j,k}^{(r)}$  denote respectively the predictor support and the weights of the first predictor operator. The following step consists in applying an update process to generate the approximation coefficients  $\tilde{I}_{j+1}^{(r)}$ :

$$\tilde{I}_{j+1}^{(r)}(m, n) = I_j^{(r)}(m, 2n) + \sum_{k \in \mathcal{U}_j^{(r)}} u_{j,k}^{(r)} \hat{d}_{j+1}^{(r)}(m, n-k), \quad (2)$$

where  $\mathcal{U}_j^{(r)}$  and  $u_{j,k}^{(r)}$  designate respectively the support and the weights of the update filter. In the last step, the final detail coefficients  $\tilde{d}_{j+1}^{(r)}$  are obtained using the second hybrid prediction step:

$$\begin{aligned} \tilde{d}_{j+1}^{(r)}(m, n) = & \hat{d}_{j+1}^{(r)}(m, n) - \left( \sum_{k \in \mathcal{Q}_j^{(r)}} q_{j,k}^{(r)} \tilde{I}_{j+1}^{(r)}(m, n-k) \right. \\ & \left. + \sum_{k \in \mathcal{P}_j^{(r,l)}} p_{j,k}^{(r,l)} I_j^{(l)}(m, 2n+1+z_j^{(r)}(m, 2n+1-k)) \right), \quad (3) \end{aligned}$$

where  $q_{j,k}^{(r)}$  (resp.  $p_{j,k}^{(r,l)}$ ) and  $\mathcal{Q}_j^{(r)}$  (resp.  $\mathcal{P}_j^{(r,l)}$ ) represent the weights and support of the second intra (resp. inter)-image predictor, and  $z_j^{(r)}$  is generated by down-sampling and dividing by  $2^j$  the input estimated disparity map  $z^{(r)}$  associated to the right view:

$$z_j^{(r)}(m, n) = \frac{1}{2^j} z^{(r)}(2^j m, 2^j n). \quad (4)$$

In this paper, this decomposition is carried out by taking the following supports and weights for the intra prediction and update filters:  $\mathcal{P}_j^{(r)} = \{-1, 0\}$ ,  $\mathcal{U}_j^{(r)} = \{0, 1\}$ ,  $p_{j,-1}^{(r)} = p_{j,0}^{(r)} = \frac{1}{2}$  and  $u_{j,0}^{(r)} = u_{j,1}^{(r)} = \frac{1}{4}$ , which are similar to those used in the standard 5/3 lifting scheme (LS) [36]. For the hybrid prediction step, it is performed with the spatial supports  $\mathcal{Q}_j^{(r)} = \{-1, 0\}$  and  $\mathcal{P}_j^{(r,l)} = \{-1, 0, 1\}$ . The associated weights  $q_{j,k}^{(r)}$  and  $p_{j,k}^{(r,l)}$  are optimized by minimizing the variance of the detail signal  $\tilde{d}_{j+1}^{(r)}$ .

Once this decomposition strategy is performed for each line of the right view, a similar P-U-P structure is applied to each line of the left view, which will result in an approximation subband and a detail one for each image. By repeating the same decomposition along the columns of the resulting subbands, one approximation subband  $I_{j+1}^{(v)}$  and three detail subbands oriented horizontally  $I_{j+1}^{(HH,v)}$ , vertically  $I_{j+1}^{(LH,v)}$  and diagonally  $I_{j+1}^{(HH,v)}$ , are generated for each view  $v \in \{r, l\}$ . Finally, two multi-resolution representations of both views can be produced by iterating the same decomposition strategy on their approximation subbands.

We should note that the main feature of this joint wavelet decomposition compared to a standard LS-based WT concerns the prediction stage. For instance, while only intra prediction and update filters are used in conventional LS, the VLS is designed by adding an hybrid prediction step which aims at exploiting simultaneously the intra and inter-view redundancies. Indeed, in order to generate the final detail coefficients of the right image, our decomposition uses the corresponding left view pixels  $(I_j^{(l)}(m, 2n+1+z_j^{(r)}(m, 2n+1-k)))_{k \in \mathcal{P}_j^{(r,l)}}$  (using the disparity compensation (DC) process), in addition to the surrounding pixels of the right view sample to be predicted  $I_j^{(r)}(m, 2n+1)$ .

Moreover, compared to VLS-based decomposition which was firstly developed for stereo image coding purpose [26], the current analysis structure, shown in Fig. 1, presents one main difference. Indeed, in stereo image compression applications, and in order to guarantee their decoding, *only* one view is encoded with such joint decomposition and the second view (known as reference image) is encoded in intra mode (i.e. independently of the other view) using a classical LS. However, since such decoding constraint is not required in the context of stereo endoscopic image enhancement, the joint wavelet decomposition developed in this work is applied in a symmetric way to both left and right images to exploit efficiently the intra and inter-view redundancies, as shown in Fig. 1 and explained in its caption.

Finally, it is important to note that such symmetric VLS decomposition has been recently investigated in [37] to design a statistical-based no reference stereo image quality assessment method. However, this joint decomposition has not been previously exploited for enhancement. While a dense and smooth disparity map is required and used in [37] for relevant statistical depth feature extraction purpose, the current developed joint wavelet decomposition can be simply performed using a standard block-matching technique to capture the matching pixels and enhance images. Once the stereo images

are transformed and enhanced, more accurate disparity maps that preserve edges should be estimated from the new data for the 3D reconstruction step as performed in [5].

### III. PROPOSED JOINT ENHANCEMENT TECHNIQUES

#### A. Background and Motivation

As mentioned in Section II, wavelets have been found to be an efficient tool for image enhancement. In this context, the main idea consists in modifying the approximation and/or detail coefficients by applying a given transformation [24], [27], [28]. For instance, in [27], [28], and for a given image  $I$  with approximation subband  $I_j$  and detail subbands  $I_j^{(o)}$  where  $o \in \{HL, LH, HH\}$ , the authors propose first to apply the following transformation to the approximation coefficients:

$$\hat{I}_j(m, n) = \kappa_j(m, n) \cdot I_j(m, n) \quad (5)$$

where  $\hat{I}_j$  is the enhanced approximation subband,  $\kappa_j(m, n) = f(I_j(m, n))$  is a scale factor and  $f(\cdot)$  is a mapping function which aims to adjust local image intensity.

Then, based on the fact that the detail coefficients contain generally small number of significant coefficients, which represent the edge structure, and large number of non-significant coefficients which can be considered as noise, the different details subbands  $I_j^{(o)}$  are transformed as follows:

$$\hat{I}_j^{(o)}(m, n) = \kappa_j(m, n) \cdot \kappa_j^{(o)}(m, n) \cdot \lambda_j^{(o)}(m, n) \cdot I_j^{(o)}(m, n) \quad (6)$$

where  $\hat{I}_j^{(o)}$  is the enhanced detail subband,  $\kappa_j^{(o)}$  is a locality factor obtained from detail coefficients using the mapping function  $f$  (i.e.  $\kappa_j^{(o)} = f(I_j^{(o)}(m, n))$ ), and  $\lambda_j^{(o)}$  is a shrinkage factor. According to [27], [28], it should be noted that the term  $\kappa_j$  is re-used based on the fact that the energy of the detail coefficients is proportional to the intensity level associated to the approximation coefficients. The term  $\kappa_j^{(o)}$  contributes to the preservation of edge information, and the shrinkage factor  $\lambda_j^{(o)}$  is used for noise removal. The latter (i.e.  $\lambda_j^{(o)}$ ) can be computed by employing the technique presented in [38].

As it has been already shown in [28], the above wavelet-based processing techniques have been found to be very efficient for image enhancement. For this reason, we propose in this paper to enhance the wavelet coefficients of the stereo endoscopic images by first resorting to similar processing techniques. Since this enhancement technique has been designed for single (i.e. mono-view) images, a straightforward solution will consist in applying these processing strategies separately on each view. However, such enhancement method is not so efficient since it neglects the redundancies existing between the left and right views and intrinsic characteristics of stereo images, namely depth data. Therefore, we design an additional joint processing stage that further exploits the binocular vision features and depth information.

To this respect, two techniques are proposed in this paper. In the first one, referred to as Adaptive Inter-View Processing (AIVP), an energy-based combination model is only applied on the approximation subbands of both views while the detail subbands are kept unchanged using the intra subband mapping

as performed in [28]. In the second one, referred to as Fully Adaptive Inter-View Processing (FAIVP), we propose to add a contrast-based model to further combine the detail subbands of both views. As it will be shown later, one main advantage of the second method is the design of a local adaptive weighting strategy. The flowcharts of the proposed enhancement methods are shown in Figs. 2 and 3, and they will be described in the following.

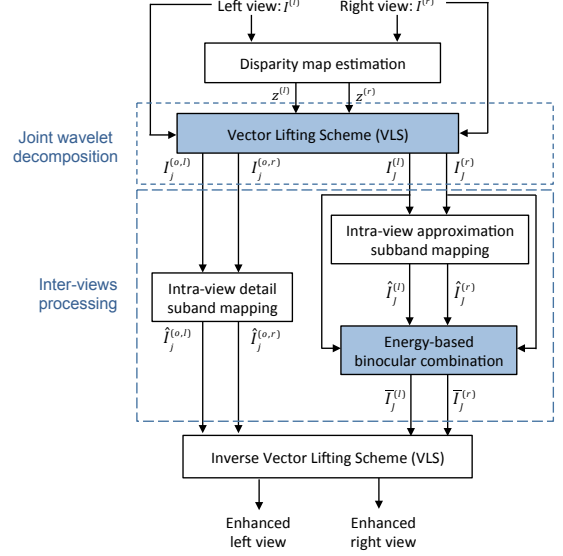


Fig. 2: The proposed adaptive inter-view processing technique.

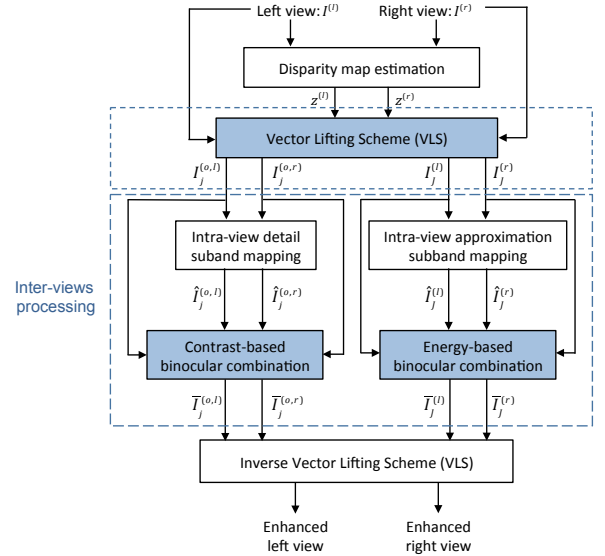


Fig. 3: The proposed fully adaptive inter-view processing technique.

#### B. Adaptive Inter-View Processing Technique

Given stereo endoscopic images, a disparity map is first estimated and the joint wavelet decomposition is performed on left and right views over  $J$  resolution levels. Thus, we obtain one approximation subband at the coarsest scale  $I_j^{(v)}$

and  $3J$  detail subbands  $\left(I_j^{(o,v)}\right)_{o \in \{HL, LH, HH\}, j \in \{1, \dots, J\}}$  for each view  $v \in \{l, r\}$ .

Then, wavelet-based mapping techniques, similar to those given in Section III-A, are applied to the generated approximation and detail subbands yielding the intermediate enhanced wavelet subbands of both views. The latter, called also pre-enhanced wavelet subbands, are expressed as follows:

$$\begin{aligned} \forall v \in \{l, r\}, \quad \forall o \in \{HL, LH, HH\}, \\ \begin{cases} \hat{I}_j^{(v)}(m, n) = \kappa_j^{(v)}(m, n) \cdot I_j^{(v)}(m, n) \\ \hat{I}_j^{(o,v)}(m, n) = \kappa_j^{(v)}(m, n) \cdot \kappa_j^{(o,v)}(m, n) \cdot \lambda_j^{(o,v)}(m, n) \cdot I_j^{(o,v)}(m, n) \end{cases} \end{aligned} \quad (7)$$

where  $\kappa_j^{(v)}$ ,  $\kappa_j^{(o,v)}$ , and  $\lambda_j^{(o,v)}$  are weighting terms similar to those used in Eqs. (5) and (6). Finally, the pre-enhanced left and right images are combined by exploiting the binocular vision property to generate the final enhanced stereo images. Indeed, since the approximation coefficients correspond to a good approximation of the original image at a coarsest scale, we propose in our first approach to combine *only* the approximation subbands of both views while keeping unchanged the other pre-enhanced detail subbands. More precisely, the new enhanced approximation subbands of the right image  $\bar{I}_J^{(r)}$  and the left one  $\bar{I}_J^{(l)}$  are given by:

$$\begin{cases} \bar{I}_J^{(r)}(m, n) = w_J^{(r)}(m, n) \cdot \hat{I}_J^{(r)}(m, n) + w_J^{(l)}(m, n) \cdot \hat{I}_J^{(l)}(m, n + z_J^{(r)}) \\ \bar{I}_J^{(l)}(m, n) = w_J^{(l)}(m, n) \cdot \hat{I}_J^{(l)}(m, n) + w_J^{(r)}(m, n) \cdot \hat{I}_J^{(r)}(m, n - z_J^{(l)}) \end{cases} \quad (8)$$

where  $w_J^{(r)}$  and  $w_J^{(l)}$  are the weighting coefficients of the corresponding right and left approximation subbands. Note that the dependence on  $(m, n)$  of  $(z_J^{(r)}, z_J^{(l)})$  has been dropped to simplify the notation.

One simple way to define these weights consists in taking an average combination of both views (i.e  $w_J^{(r)} = w_J^{(l)} = \frac{1}{2}$ ). However, this choice is not efficient since it is not correlated with the binocular perception, especially in a rivalry situation when the two views have different qualities. For this reason, we propose to resort to a local adaptive weighting combination model that is correlated with the HVS binocular vision features. Indeed, according to Levelt's study [31] and others investigations on binocular vision/rivalry [39], the linear weighting model given by Eq. (8) is equivalent to that used for simulating a *cyclopean* image, which is formed within the observer's visual cortex when two retinal views are captured by the left and right eyes [29], [40]. It is important to note that the use of the cyclopean image model has already been investigated for the quality assessment of stereo images [41], [42]. However, to the best of our knowledge, this paper presents the first research work which exploits such model for the enhancement of stereo images.

To set the stereo weights in a way simulating the visual stimulus strength, we used the gain control theory model [40],

which is one of the latest successful theories for modeling the cyclopean perception. Since our approach operates in the wavelet transform domain, we have employed an energy weighted summation model to synthesize the final enhanced approximation subbands  $\bar{I}_J^{(r)}$  and  $\bar{I}_J^{(l)}$ . Therefore, the weighting terms  $w_J^{(r)}$  and  $w_J^{(l)}$  used in Eq. (8), are computed as follows:

$$\begin{aligned} \forall v \in \{l, r\}, \quad \forall (m, n) \in \Omega_J, \\ w_J^{(v)}(m, n) = \frac{E_J^{(v)}}{E^{(l)} + E^{(r)}}, \end{aligned} \quad (9)$$

where  $\Omega_J$  represents the coordinates set of the approximation coefficients  $I_J^{(v)}$ , and  $E_J^{(v)}$  and  $E^{(v)}$  are respectively the energy of the approximation and all the wavelet subbands given by:

$$\begin{aligned} \forall v \in \{l, r\}, \\ E_J^{(v)} = \sum_{(m, n) \in \Omega_J} |I_J^{(v)}(m, n)|^2 \\ E^{(v)} = E_J^{(v)} + \sum_{j=1}^J \sum_{o \in \{HL, LH, HH\}} \sum_{(m, n) \in \Omega_j} |I_j^{(o,v)}(m, n)|^2 \end{aligned} \quad (10)$$

with  $\Omega_j$  represents the coordinates set of the detail subband coefficients  $I_j^{(o,v)}$ . Thus, increased energy of either (right and left) stimulus reduces the contribution of the other view when there is binocular rivalry. The steps of the proposed enhancement method are summarized in Fig. 2.

### C. Fully Adaptive Inter-View Processing Technique

As mentioned in the previous section, the joint processing of the left and right images has been only performed on the approximation subbands of both views while their detail subbands have been separately enhanced using Eq. (6). In the second approach, we propose to add also an efficient joint processing technique of the detail subbands yielding a fully adaptive method. The block diagram of this method is shown in Fig. 3. More precisely, similarly to the combination model used with the approximation subbands, the new enhanced detail subbands of the right image  $\bar{I}_j^{(o,r)}$  and the left one  $\bar{I}_j^{(o,l)}$  are obtained as follows:

$$\begin{aligned} \forall o \in \{HL, LH, HH\}, \quad \forall j \in \{1, \dots, J\}, \\ \begin{cases} \bar{I}_j^{(o,r)}(m, n) = w_j^{(o,r)}(m, n) \cdot \hat{I}_j^{(o,r)}(m, n) + w_j^{(o,l)}(m, n) \cdot \hat{I}_j^{(o,l)}(m, n + z_j^{(r)}) \\ \bar{I}_j^{(o,l)}(m, n) = w_j^{(o,l)}(m, n) \cdot \hat{I}_j^{(o,l)}(m, n) + w_j^{(o,r)}(m, n) \cdot \hat{I}_j^{(o,r)}(m, n - z_j^{(l)}) \end{cases} \end{aligned} \quad (11)$$

where  $w_j^{(o,r)}$  and  $w_j^{(o,l)}$  are the weights used with the right and left detail subbands.

Unlike the weights used with the approximation subbands, we propose to set those associated to the detail subbands by using a perceptual measure of the image. More precisely, based on the observation that human contrast sensitivity is highly dependent on spatial frequency [43], we adopt a local contrast

weighted summation model to synthesize the final enhanced details subbands  $\bar{I}_j^{(o,r)}$  and  $\bar{I}_j^{(o,l)}$ . Thus the resulting weighting terms are computed as follows:

$$\forall o \in \{HL, LH, HH\}, \quad \forall j \in \{1, \dots, J\},$$

$$\begin{cases} w_j^{(o,r)}(m, n) = \frac{C_j^{(o,r)}(m, n)}{C_j^{(o,r)}(m, n) + C_j^{(o,l)}(m, n + z_j^{(l)})} \\ w_j^{(o,l)}(m, n) = \frac{C_j^{(o,l)}(m, n)}{C_j^{(o,l)}(m, n) + C_j^{(o,r)}(m, n - z_j^{(r)})} \end{cases} \quad (12)$$

where  $C_j^{(o,v)}$  is the contrast measure in the wavelet transform domain [44], [25]. While different definitions of contrast have been proposed in the literature, we retain the local band-limited contrast measure, developed by Peli [43], which has the advantage of assigning a contrast value to each pixel of each frequency band of the image. The latter is given by:

$$C_j^{(o,v)}(m, n) = \frac{I_j^{(o)}(m, n)}{I_j^{(v)}(\lceil \frac{m}{2^{J-j}} \rceil, \lceil \frac{n}{2^{J-j}} \rceil)} \quad (13)$$

with  $\lceil \cdot \rceil$  is the rounding up operator. Note that this operator is applied to the coordinates of the approximation coefficients since the size of  $I_j^{(v)}$  is equal to that of  $I_j^{(o,v)}$  divided by  $(2^{J-j} \times 2^{J-j})$ .

It is worth pointing out that, contrary to the previous energy-based weighting model where a similar weight is applied to all the approximation coefficients as shown in Eq. (9), the contrast-based weighting model uses a local measure adapted to each pixel of the stereo image.

#### IV. EXPERIMENTAL RESULTS

##### A. Experimental setup and datasets

The experimental tests are carried out on 4 various datasets containing a total number of 120 SEI. The main characteristics of these datasets are summarized in Table I and some samples are illustrated in Fig. 4. Indeed, the first dataset is composed of 20 stereo pairs extracted from five stereo *ex-vivo* video sequences of a pig liver. The choice of the pig can be explained by the fact that its liver has similar structures and tissue textures to the human liver. The acquisition is performed in an operating room in the *The Intervention Center (IVS)* at *Oslo University Hospital* with a stereo laparoscope used in the laparoscopic surgical routine work by IVS surgeons. The liver was captured in illumination conditions simulating the dark closed endoscopic environment and with clinical conditions that are similar to a real laparoscopic liver resection surgery in terms of camera settings. The laparoscope was moved according to different angles and zooming scales, simulating the navigation of the surgeons and leading to different image contents with some imperfections such as blur, dark regions and specular reflections. The 3 other datasets contain a total number of 100 SEI, taken from different *in-vivo* sequences of the *Hamlyn Centre* laparoscopic/endoscopic dataset<sup>1</sup> [45], [46]. The SEI are categorized into three main classes: Vessels, Liver and Heart. Therefore, using different organs with various tissue structures (fine details, homogeneous/texture regions)

and acquisition settings (scale change, rotation, motion of the laparoscope and/or patient) allow us to increase the diversity of our experimental data.

##### B. Comparison Methods

In order to evaluate the performance of the proposed joint enhancement approaches, we compare them to the following state-of-the-art methods developed by: Selka *et al.* [16], Attar *et al.* [17], Hai *et al.* [18], Sdiri *et al.* [19], Subedar and Karam [22], Hachicha *et al.* [23], and Cho *et al.* [28]. The main ideas behind these enhancement techniques have been addressed in Section I. For instance, let us recall that, in [16], [17], the methods are developed for single (i.e one view) endoscopic images, and so, they have been separately applied to the left and right images in this work. The methods designed in [18], [19] (rep. [22], [23]) correspond to stereo endoscopic (resp. natural) images. The wavelet-based approach developed in [28] is for 2D images, and so, it has been applied to each view of the SEI. It should be noted here that our first (resp. second) approach based on VLS followed by an adaptive (resp. fully adaptive) inter-view processing are performed using two resolution levels (i.e  $J = 2$ ), and will be designated in the following by VLS-AIVP and VLS-FAIVP, respectively.

##### C. Objective Quality Assessment Metrics

To assess the contrast enhancement methods, the proposed and baseline comparison ones are evaluated using metrics based on local and global image properties such as edginess information, image intensity and local contrast. It is worth pointing out that the stereoscopic image quality assessment for contrast enhancement is not well investigated in the literature. Indeed, only few works have been developed relying mainly on subjective evaluations which are generally expensive and time consuming [47], [48]. Therefore, due to the lack of stereo contrast enhancement evaluation (CEE) metrics, the following conventional 2D-CEE measures are used: Region Contrast (RC) [25], Edge Content (EC) [49], Absolute Measure of Enhancement (AME) [50], Second Derivative like Measure of Enhancement (SDME) [51], the Image Enhancement Metric (IEM) [52], Visual Information Fidelity (VIF) [53] and Mean Brightness (MB) [28]. While MB, AME and SDME measure the image intensity adjustment, RC, EC and IEM evaluate the local contrast and indicate the ability of detecting local image features which is a key step for different subsequent processing tasks like 3D reconstruction and tracking. MB is an indicator about the average image intensity level that shows whether the image is too bright or too dark. A good MB should be close to the mid-intensity range defined as the mean of the maximum and minimum intensity values. AME combines the *Weber's* contrast law and *Michelson's* contrast law modulation to simulate the HVS appreciation of the image contrast enhancement. The final score is estimated based on an average value of the relation between the sum and spread of luminance intensities within each bloc. In the same way, the SDME incorporates the center pixel within each block in the computation of the visibility operator to reduce the sensitivity to noise. Both RC and EC metrics are based on local image

<sup>1</sup>Hamlyn.doc.ic.ac.uk/vision/



intensity variations computed either by Sobel operator or any other gradient-based filter depending on the level of edges to be detected and the sensitivity to noise. Based on the idea that variations in contrast/sharpness reflect the changes between a given pixel and its neighbors, the IEM consists in comparing the absolute intra-bloc differences between the initial and the enhanced image. Inspired by HVS features, the VIF predicts the perceived quality of contrast enhanced images using an image information measure that quantifies an estimation of the information extracted by the brain when observing an image. Compared with other image quality assessment metrics, VIF is reported to be consistent and correlated with the subjective evaluation of the HVS [53], [54]. Instead of using the spatial domain implementation of the VIF, we used the wavelet-based version which has been found to yield better performance [53]. In case the enhancement processing does not add or amplify noise, the obtained VIF score is larger than 1, which implies a superior perceptible quality compared to the original image. Note that these metrics are applied on the left and right enhanced views, and their average values correspond to the final objectives scores. The mathematical expressions of all these quality measures detailed above can be found in [54]. Higher scores of RC, EC, IEM and VIF, and lower scores of AME and SDME, show good performance.

#### D. Results

The performance of the different enhancement methods for a sample SEI as well as the average scores on all images are given in Tables II-V and Tables VI-IX, for the *IVS* and *Hamlyn Centre* datasets respectively. For each dataset, the proposed methods are compared separately in two different tables with 2D and 3D enhancement techniques. Note that the first line of each table, denoted by "Original", indicate the initial scores computed on the input data before the enhancement step. Moreover, to easily identify the efficient enhancement methods, the objective quality scores of the two best methods have been highlighted in bold. From these tables, it can be firstly noticed that the endoscopic image enhancement methods [16], [17] as well as the 3D enhancement approach [22] often outperform the other state-of-the-art methods. Moreover, our first proposed approach "VLS-AIVP" relying on the stereo approximation subbands processing leads to better results, especially for the *IVS* liver dataset which is characterized by large homogeneous liver structures and more drastic dynamic illumination conditions. Finally, thanks to the fully adaptive inter-view processing technique "VLS-FAIVP", a significant gain is achieved for both datasets images in terms of image intensity, contrast and edginess information. The gain is more significant for the *Hamlyn* dataset SEI in terms of sharpness and edge information (RC and EC metrics) since these images contain much more structure details with regular/micro edges representing fine veins and vessels. Indeed, the stereo high-frequency subbands processing of the proposed "VLS-FAIVP" technique enhances the images details and local contrast, which exhibits the subtle tissues structures and fine vessels.

In addition to this quantitative evaluation, the visual quality of the processed images shows the efficiency of the proposed

methods. Figs. 5 and 6 illustrate the original and enhanced right views for two SEI taken from *IVS* and *Hamlyn Centre* datasets, respectively. It can be observed that the 2D endoscopic image enhancement methods [16], [17] result in an over-enhancement illustrated by both over brightened and darkened regions despite better exposure of the edges in some other areas. Furthermore, the 3D endoscopic enhancement methods [18], [19] yield better results without efficiently enhancing the small image details. By using our approach "VLS-FAIVP", the obtained images show better image luminance and contrast while improving and emphasizing the local image details such as the liver edges of the *IVS* dataset and the fine textures and structures of the *Hamlyn* dataset.

Moreover, in order to better show the benefits of the different blocks related to the proposed approaches, we will focus now on the following comparisons. First, our methods will be also evaluated by modifying the wavelet decomposition step. Indeed, instead of using a joint wavelet transform based on VLS, a standard 2D 5/3 Lifting Scheme (LS) is used by applying it separately to each view of the stereo pairs. These methods are denoted by "LS-AIVP" and "LS-FAIVP". Note that the results obtained with the wavelet-based intra-view processing technique [28] allow us to show the interest of exploiting both the depth and binocular stereo information by resorting to an inter-view processing. Tables X and XI provide the performance of these latter methods. Thus, it can be seen that our approaches outperform significantly the method developed by Cho *et al.* [28] which confirms the benefits of the proposed inter-view processing techniques compared to a simple intra-view processing. Further improvements are also achieved using the joint wavelet decomposition (VLS) compared to a conventional 2D LS-based one. Figs. 7 and 8 illustrate the enhanced images obtained by our approaches "VLS-AIVP" and "VLS-FAIVP" as well as that of Cho *et al.* [28]. Compared to the latter, it can be noticed that the "VLS-AIVP" leads to better visual image quality in terms of local contrast and image intensity. Indeed, the dark regions are reduced and the liver structures and edges are sharper and clearer. The images are further enhanced by resorting to an inter-view processing technique fully adapted to the local characteristics and activity of the stereo endoscopic images.

Finally, we propose to evaluate the impact of the proposed method VLS-FAIVP on two typical subsequent processing steps. The first one concerns the feature matching which is crucial in many robotic-assisted MIS applications [55] and aims to find the image similarities between the two views. The second one is the 3D organ surface reconstruction step. To this end, we have used the Hierarchical Multi-Affine (HMA) Feature-Matching Toolbox [55] and the dense surface reconstruction algorithm developed recently for stereo endoscopic images [5]. The results, obtained with the original SEI and the enhanced versions using different methods, are shown in Fig. 9. More precisely, the original and enhanced right views are displayed in the first column of Fig. 9, while the corresponding feature matching results between the left and right views as well as their 3D reconstructed organs are illustrated in the second and third columns, respectively. It can be firstly observed that using an enhancement step improves clearly the matching process



TABLE I: Datasets description. <sup>1</sup> Taken from *Hamlyn Centre* dataset [45], <sup>2</sup> Taken from the dataset of *The Intervention Center (IVS)* at *Oslo University Hospital*.

Dataset	Type	Number of SEI	Size	Acquisition settings	Tissue features
Vessels <sup>1</sup>	<i>in-vivo</i>	10	640 × 480	Scale variation (laparoscope navigation)	Fine details, edges, and micro edges
		10	640 × 480	Orientation variation (laparoscope rotation)	
Liver <sup>1</sup>	<i>in-vivo</i>	20	360 × 288	Static laparoscope and liver deformation due to respiration	Medium homogeneous region, few edges
		20	640 × 480		
Heart Phantom <sup>1</sup>	<i>in-vivo</i>	20	360 × 288	Heart phantom deformation due to cardiac motion with different angles	Wide homogeneous regions with edges
		20	360 × 288		
Liver <sup>2</sup>	<i>ex-vivo</i>	20	1080 × 1728	Static liver and orientatin/scale variation due to laparoscope rotation/navigation	

TABLE II: Performance comparison of the proposed methods with 2D enhancement techniques for SEI-1 of the *IVS* dataset.

Method/Measure	MB	AME	SDME	EC	RC	IEM	VIF
Original	122.87	60.43	134.11	65.49	5.97	-	1
VLS-AIVP	<b>134.44</b>	<b>50.35</b>	<b>105.87</b>	<b>91.69</b>	<b>33.10</b>	<b>2.52</b>	<b>1.31</b>
VLS-FAIVP	<b>128.54</b>	<b>48.14</b>	<b>98.40</b>	<b>110.33</b>	<b>46.27</b>	<b>3.03</b>	<b>1.52</b>
Cho <i>et al.</i> [28]	165.42	61.26	116.82	77.63	7.89	1.26	1.11
Selka <i>et al.</i> [16]	98.87	51.94	118.11	81.27	5.58	1.25	1.24
Attar <i>et al.</i> [17]	122.42	56.84	137.97	88.30	12.39	1.27	1.17

TABLE III: Average scores of the proposed methods compared with 2D enhancement techniques for the *IVS* dataset images.

Method/Measure	MB	AME	SDME	EC	RC	IEM	VIF
VLS-AIVP	<b>121.70</b>	<b>40.33</b>	<b>91.13</b>	<b>81.19</b>	<b>20.41</b>	<b>1.97</b>	<b>1.39</b>
VLS-FAIVP	<b>123.73</b>	<b>38.59</b>	<b>88.60</b>	<b>96.53</b>	<b>23.81</b>	<b>2.19</b>	<b>1.47</b>
Cho <i>et al.</i> [28]	108.23	48.42	112.95	52.17	5.51	0.89	1.14
Selka <i>et al.</i> [16]	78.37	50.41	92.46	64.35	4.74	1.11	1.33
Attar <i>et al.</i> [17]	102.02	47.71	108.87	70.24	10.05	1.13	1.26

TABLE IV: Performance comparison of the proposed methods with stereo enhancement techniques for SEI-1 of the *IVS* dataset.

Method/Measure	MB	AME	SDME	EC	RC	IEM	VIF
Original	122.87	60.43	134.11	65.49	5.97	-	1
VLS-AIVP	134.44	<b>50.35</b>	<b>105.87</b>	<b>91.69</b>	<b>33.10</b>	<b>2.52</b>	<b>1.31</b>
VLS-FAIVP	<b>128.54</b>	<b>48.14</b>	<b>98.40</b>	<b>110.33</b>	<b>46.27</b>	<b>3.03</b>	<b>1.52</b>
Hai <i>et al.</i> [18]	115.11	54.04	118.86	80.58	8.47	1.29	1.17
Sdiri <i>et al.</i> [19]	122.86	53.78	120.20	86.97	11.00	1.52	1.19
Subedar and Karam [22]	<b>123.10</b>	50.79	118.18	88.33	25.31	2.01	1.16
Hachicha <i>et al.</i> [23]	122.31	56.66	122.71	79.06	7.74	1.26	1.14

TABLE V: Average scores of the proposed methods compared with stereo enhancement techniques for the *IVS* dataset.

Method/Measure	MB	AME	SDME	EC	RC	IEM	VIF
VLS-AIVP	<b>123.39</b>	<b>43.89</b>	<b>94.60</b>	<b>84.95</b>	<b>24.13</b>	<b>2.10</b>	<b>1.34</b>
VLS-FAIVP	<b>132.47</b>	<b>36.96</b>	<b>85.97</b>	<b>98.16</b>	<b>28.49</b>	<b>2.66</b>	<b>1.68</b>
Hai <i>et al.</i> [18]	109.98	55.49	105.81	67.15	17.43	1.71	1.19
Sdiri <i>et al.</i> [19]	101.80	57.66	106.68	66.75	16.40	1.67	1.16
Subedar and Karam [22]	114.82	51.13	101.92	70.64	19.10	1.97	1.22
Hachicha <i>et al.</i> [23]	97.18	63.71	110.04	60.28	14.10	1.51	1.11

TABLE VI: Performance comparison of the proposed methods with 2D enhancement techniques for SEI-1 of the *Hamlyn Centre* dataset.

Method/Measure	MB	AME	SDME	EC	RC	IEM	VIF
Original	100.56	57.22	137.83	64.95	42.75	-	1
VLS-AIVP	<b>123.08</b>	<b>50.35</b>	<b>101.54</b>	<b>99.98</b>	<b>92.20</b>	<b>1.98</b>	<b>1.19</b>
VLS-FAIVP	<b>130.54</b>	<b>48.14</b>	<b>96.04</b>	<b>119.25</b>	<b>126.27</b>	<b>2.65</b>	<b>1.28</b>
Cho <i>et al.</i> [28]	150.58	59.90	112.38	84.62	67.65	1.44	1.09
Selka <i>et al.</i> [16]	98.41	54.04	107.12	70.63	40.01	1.03	1.04
Attar <i>et al.</i> [17]	117.58	50.41	147.67	88.77	45.07	1.05	1.06

TABLE VII: Average scores of the proposed methods compared with 2D enhancement methods for the *Hamlyn Centre* dataset.

Method/Measure	MB	AME	SDME	EC	RC	IEM	VIF
VLS-AIVP	126.71	<b>42.78</b>	<b>99.40</b>	99.31	<b>66.36</b>	<b>1.87</b>	<b>1.32</b>
VLS-FAIVP	<b>128.35</b>	<b>40.44</b>	<b>93.22</b>	<b>119.81</b>	<b>87.03</b>	<b>2.59</b>	<b>1.48</b>
Cho <i>et al.</i> [28]	130.41	60.20	116.20	71.94	29.83	0.88	1.09
Selka <i>et al.</i> [16]	93.53	45.32	122.59	84.56	22.53	1.02	1.07
Attar <i>et al.</i> [17]	<b>127.54</b>	47.01	135.32	<b>110.99</b>	45.46	1.21	1.13

TABLE VIII: Performance of the proposed and conventional stereo enhancement techniques for SEI-1 of the *Hamlyn* dataset.

Method/Measure	MB	AME	SDME	EC	RC	IEM	VIF
Original	100.56	57.22	137.83	64.95	42.75	-	1
VLS-AIVP	<b>123.08</b>	<b>50.35</b>	<b>101.54</b>	<b>99.98</b>	<b>92.20</b>	<b>1.98</b>	<b>1.19</b>
VLS-FAIVP	<b>130.54</b>	<b>48.14</b>	<b>96.04</b>	<b>119.25</b>	<b>126.27</b>	<b>2.65</b>	<b>1.28</b>
Hai <i>et al.</i> [18]	103.10	53.63	127.02	74.64	42.71	1.19	1.09
Sdiri <i>et al.</i> [19]	100.47	53.26	129.76	82.20	68.12	1.45	1.11
Subedar and Karam [22]	100.66	52.18	110.50	77.05	64.66	1.73	1.13
Hachicha <i>et al.</i> [23]	100.53	56.91	136.91	65.96	43.23	1.02	1.03

TABLE IX: Average scores of the proposed methods compared with stereo enhancement methods for the *Hamlyn* dataset.

Method/Measure	MB	AME	SDME	EC	RC	IEM	VIF
VLS-AIVP	<b>134.92</b>	<b>36.41</b>	<b>86.77</b>	<b>149.32</b>	<b>68.83</b>	<b>2.04</b>	<b>1.41</b>
VLS-FAIVP	<b>122.66</b>	<b>32.82</b>	<b>80.04</b>	<b>162.15</b>	<b>76.98</b>	<b>2.65</b>	<b>1.54</b>
Hai <i>et al.</i> [18]	86.39	46.65	95.47	123.39	62.05	1.72	1.26
Sdiri <i>et al.</i> [19]	84.82	48.75	99.11	122.61	59.72	1.66	1.25
Subedar and Karam [22]	89.24	39.83	92.94	131.18	66.12	1.89	1.29
Hachicha <i>et al.</i> [23]	78.81	50.96	102.70	120.35	44.25	1.51	1.18

TABLE X: Average scores of LS and VLS-based enhancement methods for the *IVS* dataset.

Method/Measure	MB	AME	SDME	EC	RC	IEM	VIF
Cho <i>et al.</i> [28]	172.83	69.32	119.44	73.39	16.77	1.38	1.13
LS-AIVP	144.65	52.71	107.28	79.66	20.68	1.81	1.26
VLS-AIVP	123.39	43.89	94.60	84.95	24.13	2.10	1.34
LS-FAIVP	137.12	40.50	92.35	89.59	26.23	2.44	1.52
VLS-FAIVP	<b>132.47</b>	<b>36.96</b>	<b>85.97</b>	<b>98.16</b>	<b>28.49</b>	<b>2.66</b>	<b>1.68</b>

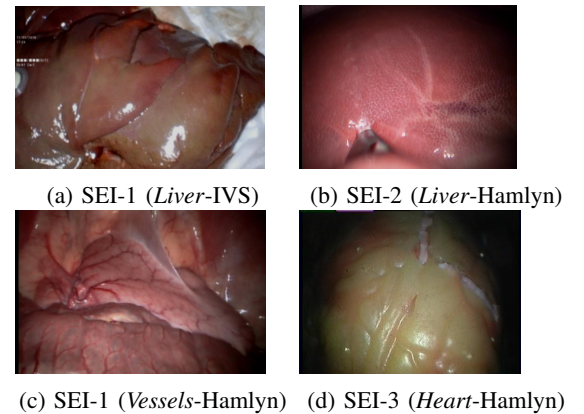
TABLE XI: Average scores of LS and VLS-based enhancement methods for the *Hamlyn Centre* dataset.

Method/Measure	MB	AME	SDME	EC	RC	IEM	VIF
Cho <i>et al.</i> [28]	159.76	63.62	108.51	97.70	54.86	1.66	1.17
LS-AIVP	147.51	47.14	95.26	134.85	62.12	1.90	1.28
VLS-AIVP	134.92	36.41	86.77	149.32	68.83	2.04	1.41
LS-FAIVP	136.26	35.62	84.48	155.36	73.84	2.36	1.44
VLS-FAIVP	<b>122.66</b>	<b>32.82</b>	<b>80.04</b>	<b>162.15</b>	<b>76.98</b>	<b>2.65</b>	<b>1.54</b>

by increasing the number of matched features. Moreover, the stereo pairs enhanced with our proposed method result in more matched features that are well distributed around almost the entire organ surface especially in edges and textures. In terms of 3D organ reconstruction, our method yields larger reconstructed organ surfaces that present better details with sharper edges and textures, while reducing the staircasing effect which appears on the right side of the liver test image. These improvements are expected since having dense well distributed inter-view feature matching should result in a better reconstructed 3D shape.

Before concluding the paper, we evaluate the proposed enhancement methods in terms of execution time. Indeed, using a computer with an Intel Core i7 processor (2.3 GHz), a matlab implementation and a test image with the highest spatial resolution in our dataset ( $1080 \times 1728$ ), the FAIVP (resp. AIVP) technique takes 0.46 (resp. 0.31) seconds while the VLS decomposition takes 2.5 seconds. Note that this execution time can be further reduced since these two steps could be

efficiently implemented on parallel computing architectures.

Fig. 4: Sample right views taken from the *IVS* and *Hamlyn Centre* datasets.

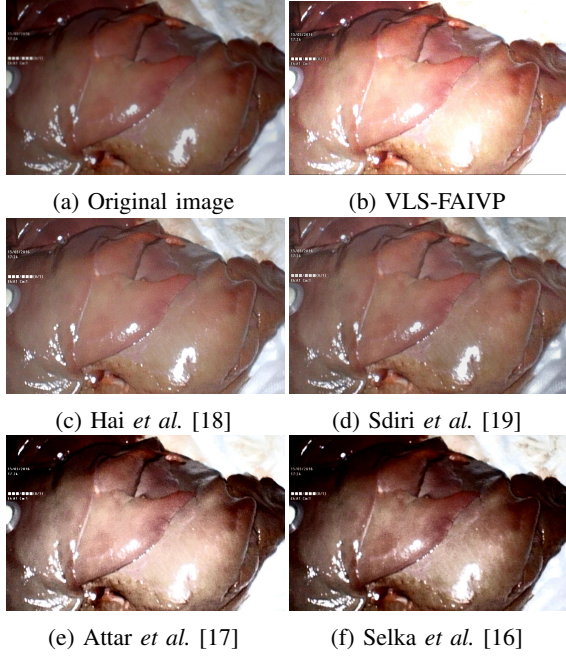


Fig. 5: The enhanced right image for the SEI-1 of the *IVS* dataset using different methods.

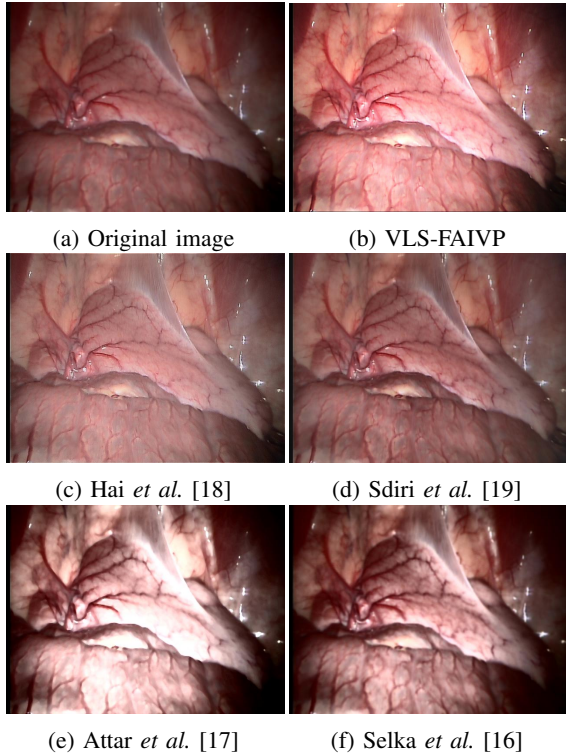


Fig. 6: The enhanced right image for the SEI-1 of the *Hamlyn Centre* dataset using different methods.

## V. DISCUSSIONS AND CONCLUSION

In this paper, we have focused on the enhancement of stereo endoscopic images. For this purpose, two efficient methods have been designed in the wavelet transform domain. More precisely, a joint stereo wavelet decomposition is used to

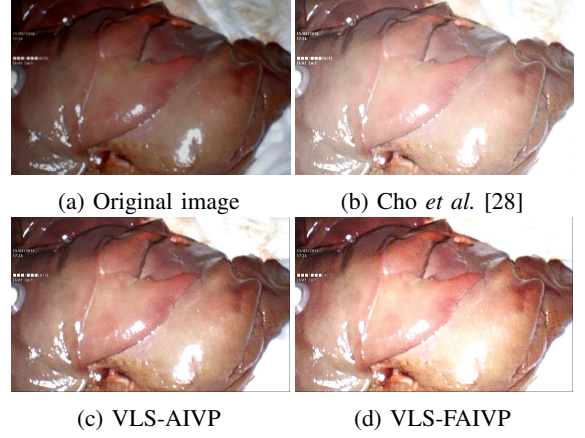


Fig. 7: The enhanced right images for the SEI-1 of *IVS* dataset using the two proposed inter-view processing techniques and the intra-view processing one [28].

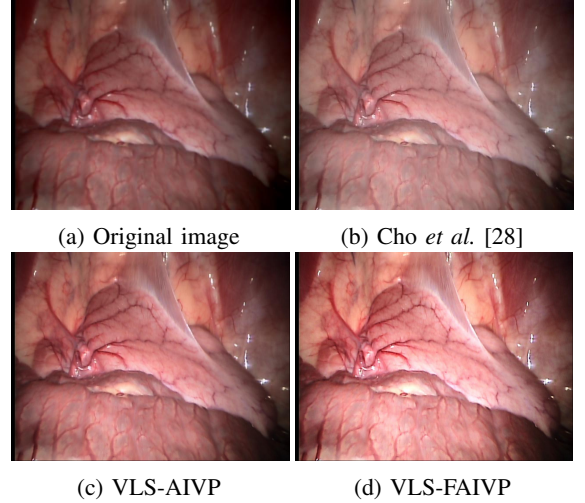


Fig. 8: The enhanced right images for the SEI-1 of *Hamlyn Centre* dataset using the two proposed inter-view processing techniques and the intra-view processing one [28].

generate the wavelet subbands of the left and right views, and two inter-view processing techniques are proposed to exploit the depth information and binocular vision properties. The obtained experimental results confirm the efficiency of the proposed techniques compared to the state-of-the-art methods. It should be noted that there are two main novelties in this work. The first one aims to show the interest of the VLS-based joint decomposition compared to a standard lifting scheme-based decomposition for enhancing the left and right endoscopic images. The second one is the design of a local contrast-based binocular combination model fully adapted to the image contents.

While the proposed methods are more performant than state-of-the-art methods, some possible limitations should be mentioned. First, the proposed enhancement methods seem to increase in some cases the specular reflection components as shown in the results obtained with the liver *IVS* dataset. However, to overcome this drawback, one can firstly apply



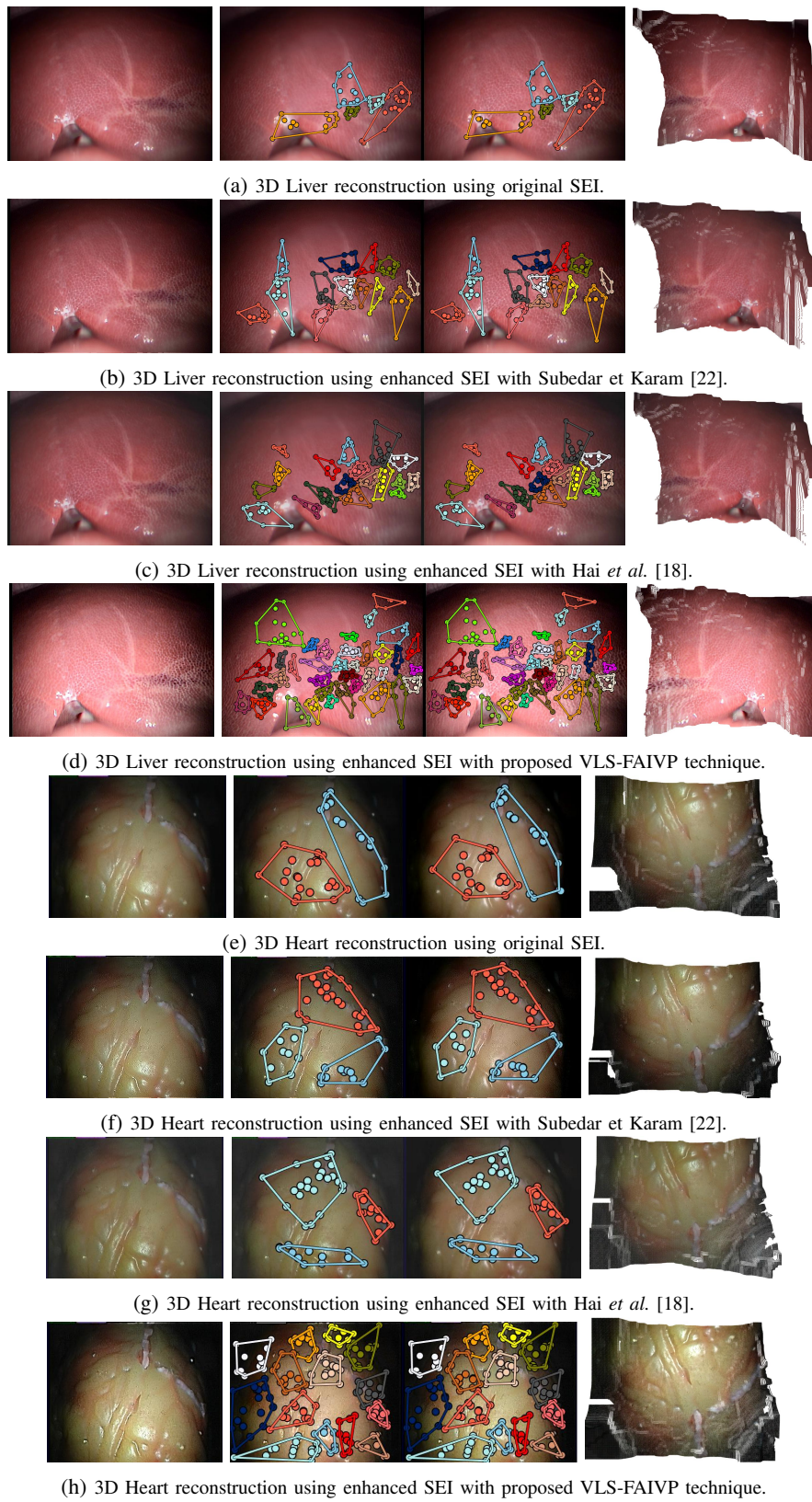


Fig. 9: Enhancement evaluation in terms of feature matching and 3D organ reconstruction. The first column illustrates the original and enhanced right views of the input stereo images, the second column displays the corresponding feature matching results using HMA toolbox [55], and the third column shows the 3D organ reconstruction results obtained using [5].

an efficient pre-processing step to remove the specularities before performing the enhancement step. Another solution is to perform a selective image quality enhancement by using a specular reflection detection algorithm [56]. Moreover, the efficiency of the proposed inter-views processing techniques may be affected in the case where the stereo endoscopic images present large occlusions.

In future work, we aim at resorting to other combination models involving perceptual weights and taking into account the occlusion effect. Moreover, the inter-subband correlations within each view could be further exploited in the joint enhancement process. Finally, the use of other mapping functions to adjust the wavelet coefficients could be investigated.

#### ACKNOWLEDGMENT

The authors would like to acknowledge Congcong Wang for her valuable feedback regarding the 3D reconstruction results. This research work has been supported by HyPerCept project, funded by both Université Paris 13 and the Research Council of Norway through project no. 221073 (*HyPerCept - Color and Quality in higher dimensions*), and performed between the Laboratory of Information Processing and Transmission (L2TI) in Université Paris 13, the Norwegian Colour and Visual Computing Laboratory (ColourLab) in the Norwegian University of Science and Technology, and the Intervention Centre at Oslo University Hospital.

#### REFERENCES

- [1] G. H. Ballantyne and F. Moll, "The Da Vinci telerobotic surgical system: the virtual operative field and telepresence surgery," *Surgical Clinics of North America*, vol. 6, no. 83, pp. 1293–1304, 2003.
- [2] B. Lin, Y. Sun, J. E. Sanchez, and X. Qian, "Efficient vessel feature detection for endoscopic image analysis," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 4, pp. 1141–1150, 2014.
- [3] L. Maier-Hein, A. Groch, A. Bartoli, S. Bodenstedt, G. Boissonnat, P.-L. Chang, N. Clancy, D. S. Elson, S. Haase, E. Heim, J. Hornegger et al., "Comparative validation of single-shot optical techniques for laparoscopic 3-D surface reconstruction," *IEEE Transactions on Medical Imaging*, vol. 33, no. 10, pp. 1913–1930, 2014.
- [4] R. Palomar, F. A. Cheikh, B. Edwin, A. Beghdadi, and O. J. Elle, "Surface reconstruction for planning and navigation of liver resections," *Computerized Medical Imaging and Graphics*, vol. 53, pp. 30–42, 2016.
- [5] C. Wang, F. A. Cheikh, M. Kaaniche, and O. J. Elle, "Liver surface reconstruction for image guided surgery," in *Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 10576. International Society for Optics and Photonics, 2018, p. 105762H.
- [6] X. Luo, A. McLeod, S. Pautler, C. Schlachta, and T. Peters, "Vision-based surgical field defogging," *IEEE Transactions on Medical Imaging*, vol. 36, no. 10, pp. 2021–2030, 2017.
- [7] H. Okuhata, H. Nakamura, S. Hara, H. Tsutsui, and T. Onoye, "Application of the real-time retinex image enhancement for endoscopic images," in *International Conference of the Engineering in Medicine and Biology Society*. IEEE, 2013, pp. 3407–3410.
- [8] F. Deeba, S. K. Mohammed, F. M. Bui, and K. A. Wahid, "Unsupervised abnormality detection using saliency and retinex based color enhancement," in *International Conference of the Engineering in Medicine and Biology Society*. IEEE, 2016, pp. 3871–3874.
- [9] J. C.-H. Wu, G.-S. Lin, H.-T. Hsu, Y.-P. Tan, K.-C. Liu, and W.-N. Lie, "Quality enhancement based on retinex and pseudo-hdr synthesis algorithms for endoscopic images," in *Visual Communications and Image Processing*. IEEE, 2013, pp. 1–5.
- [10] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image processing*, vol. 6, no. 7, pp. 965–976, 1997.
- [11] O. Oktay, E. Ferrante, K. Kamnitsas, M. Heinrich, W. Bai, J. Caballero, R. Guerrero, S. Cook, A. de Marvao, D. O'Regan, B. Kainz, B. Glocker, and D. Rueckert, "Anatomically constrained neural networks (ACNN): Application to cardiac image enhancement and segmentation," *IEEE Transactions on Medical Imaging*, 2017.
- [12] C. F. Baumgartner, O. Oktay, and D. Rueckert, "Fully convolutional networks in medical imaging: Applications to image enhancement and recognition," in *Deep Learning and Convolutional Neural Networks for Medical Image Computing*. Springer, 2017, pp. 159–179.
- [13] H. Chen, A. Li, L. Kaufman, and J. Hale, "A fast filtering algorithm for image enhancement," *IEEE Transactions on medical imaging*, vol. 13, no. 3, pp. 557–564, 1994.
- [14] A. Makeev and S. J. Glick, "Low-dose contrast-enhanced breast ct using spectral shaping filters: An experimental study," *IEEE transactions on medical imaging*, vol. 36, no. 12, pp. 2417–2423, 2017.
- [15] M. S. Imtiaz and K. A. Wahid, "Image enhancement and space-variant color reproduction method for endoscopic images using adaptive sigmoid function," in *International Conference of the Engineering in Medicine and Biology Society*. IEEE, 2014, pp. 3905–3908.
- [16] F. Selka, S. A. Nicolau, V. Agnus, A. Bessaid, J. Marescaux, and L. Soler, "Evaluation of endoscopic image enhancement for feature tracking: a new validation framework," in *Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions*. Springer, 2013, pp. 75–85.
- [17] A. Attar, X. Xie, C. Zhang, Z. Wang, and S. Yue, "Wireless micro-ball endoscopic image enhancement using histogram information," in *International Conference of the Engineering in Medicine and Biology Society*. IEEE, 2014, pp. 3337–3340.
- [18] Y. Hai, L. Li, and J. Gu, "Image enhancement based on contrast limited adaptive histogram equalization for 3D images of stereoscopic endoscopy," in *International Conference on Information and Automation*. IEEE, 2015, pp. 668–672.
- [19] B. Sdiri, A. Beghdadi, F. A. Cheikh, M. Pedersen, and O. J. Elle, "An adaptive contrast enhancement method for stereo endoscopic images combining binocular just noticeable difference model and depth information," *Electronic Imaging*, vol. 2016, no. 13, pp. 1–7, 2016.
- [20] Y. Zhao, Z. Chen, C. Zhu, Y.-P. Tan, and L. Yu, "Binocular just-noticeable-difference model for stereoscopic images," *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 19–22, 2011.
- [21] S.-W. Jung, J.-Y. Jeong, and S.-J. Ko, "Sharpness enhancement of stereo images using binocular just-noticeable difference," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 1191–1199, 2012.
- [22] M. M. Subedar and L. J. Karam, "Increased depth perception with sharpness enhancement for stereo video," in *Electronic Imaging*. International Society for Optics and Photonics, 2010, pp. 75 241B–75 241B.
- [23] W. Hachicha, A. Beghdadi, and F. A. Cheikh, "Combining depth information and local edge detection for stereo image enhancement," in *European Signal Processing Conference*. IEEE, 2012, pp. 250–254.
- [24] Y. Qiu and J. Tang, "An enhancement technique for stereoscopic images using salient features and wavelet transform," in *Proc. SPIE 9120, Mobile Multimedia/Image Processing, Security, and Applications*, 2014.
- [25] J. Tang, X. Liu, and Q. Sun, "A direct image contrast enhancement algorithm in the wavelet domain for screening mammograms," *Journal of Selected Topics in Signal Processing*, vol. 3, no. 1, pp. 74–80, 2009.
- [26] M. Kaaniche, A. Benazza-Benyahia, B. Pesquet-Popescu, and J.-C. Pesquet, "Vector lifting schemes for stereo image coding," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2463–2475, 2009.
- [27] D. Cho and T. D. Bui, "Image contrast enhancement in compressed wavelet domain," in *International Conference on Image Processing*. IEEE, 2011, pp. 3421–3424.
- [28] D. Cho and T. Bui, "Fast image enhancement in compressed wavelet domain," *Signal Processing*, vol. 98, pp. 295–307, 2014.
- [29] B. Julesz, *Foundations of cyclopean perception*. Chicago: University of Chicago Press, 1971.
- [30] A. Maalouf and M.-C. Larabi, "Cyclop: A stereo color image quality assessment metric," in *International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2011, pp. 1161–1164.
- [31] W. J. Levelt, *On binocular rivalry*. Van Gorcum Assen, 1965.
- [32] H. Demirel and G. Anbarjafari, "Image resolution enhancement by using discrete and stationary wavelet decomposition," *IEEE Transactions on Image Processing*, vol. 20, no. 5, pp. 1458–1460, 2011.
- [33] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on image processing*, vol. 9, no. 9, pp. 1532–1546, 2000.
- [34] A. Laine, J. Fan, and W. Yang, "Wavelets for contrast enhancement of digital mammography," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, no. 5, pp. 536–550, 1995.

- [35] Y. Y. Tang, *Wavelet theory approach to pattern recognition*. World Scientific, 2009.
- [36] W. Sweldens *et al.*, “The lifting scheme: A new philosophy in biorthogonal wavelet constructions,” *Wavelet Applications in Signal and Image Processing*, vol. 3, pp. 68–79, 1995.
- [37] W. Hachicha, M. Kaaniche, A. Beghdadi, and F. A. Cheikh, “No-reference stereo image quality assessment based on joint wavelet decomposition and statistical models,” *Signal Processing: Image Communication*, vol. 54, pp. 107–117, 2017.
- [38] D. Cho and T. D. Bui, “Multivariate statistical modeling for image denoising using wavelet transforms,” *Signal Processing: Image Communication*, vol. 20, no. 1, pp. 77–89, 2005.
- [39] R. Blake, Y. Yang, and H. R. Wilson, “On the coexistence of stereopsis and binocular rivalry,” *Elsevier, Vision Research*, vol. 31, no. 7-8, pp. 1191–1203, 1991.
- [40] J. Ding and G. Sperling, “A gain-control theory of binocular combination,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 4, pp. 1141–1146, 2006.
- [41] M.-J. Chen, C.-C. Su, D.-K. Kwon, L. K. Cormack, and A. C. Bovik, “Full-reference quality assessment of stereopairs accounting for rivalry,” *Signal Processing: Image Communication*, vol. 28, no. 9, pp. 1143–1155, 2013.
- [42] Y.-H. Lin and J.-L. Wu, “Quality assessment of stereoscopic 3d image compression by binocular integration behaviors,” *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1527–1542, 2014.
- [43] E. Peli, “Contrast in complex images,” *Journal of the Optical Society of America A*, vol. 7, no. 10, pp. 2032–2040, 1990.
- [44] J. Tang, Q. Sun, and K. Agyepong, “An image enhancement algorithm based on a contrast measure in the wavelet domain for screening mammograms,” in *International Conference on Image Processing*, vol. 5. IEEE, 2007, pp. 29–32.
- [45] P. Mountney, D. Stoyanov, and G.-Z. Yang, “Three-dimensional tissue deformation recovery and tracking: Introducing techniques based on laparoscopic or endoscopic images,” *Signal Processing Magazine*, vol. 27, no. 4, pp. 14–24, 2010.
- [46] D. Stoyanov, M. V. Scarzanella, P. Pratt, and G.-Z. Yang, “Real-time stereo reconstruction in robotically assisted minimally invasive surgery,” in *Medical Image Computing and Computer Assisted Interventions*, 2010, pp. 275–282.
- [47] F. Battisti, M. Carli, and A. Neri, “Subjective evaluation of 3D video enhancement algorithm,” in *European Signal Processing Conference*. IEEE, 2014, pp. 250–254.
- [48] N. Kawabata and M. Miyao, “Multi-view 3D CG image quality assessment for contrast enhancement based on S-CIELAB color space,” *IEICE Transactions on Information and Systems*, vol. 100, no. 7, pp. 1448–1462, 2017.
- [49] A. Saleem, A. Beghdadi, and B. Boashash, “Image fusion-based contrast enhancement,” *Journal on Image and Video Processing*, vol. 2012, no. 1, p. 10, 2012.
- [50] S. S. Agaian, B. Silver, and K. A. Panetta, “Transform coefficient histogram-based image enhancement algorithms using contrast entropy,” *IEEE Transactions on Image Processing*, vol. 16, no. 3, pp. 741–758, 2007.
- [51] K. Panetta, C. Gao, and S. Agaian, “No reference color image contrast and quality measures,” *IEEE Transactions on Consumer Electronics*, vol. 59, no. 3, pp. 643–651, 2013.
- [52] V. Jaya and R. Gopikakumari, “IEM: a new image enhancement metric for contrast and sharpness measurements,” *International Journal of Computer Applications*, vol. 79, no. 9, 2013.
- [53] H. R. Sheikh and A. C. Bovik, “Image information and visual quality,” *IEEE Transactions on image processing*, vol. 15, no. 2, pp. 430–444, 2006.
- [54] M. A. Qureshi, A. Beghdadi, and M. Deriche, “Towards the design of a consistent image contrast enhancement evaluation measure,” *Signal Processing: Image Communication*, vol. 58, pp. 212–227, 2017.
- [55] G. A. Puerto-Souza and G.-L. Mariottini, “A fast and accurate feature-matching algorithm for minimally-invasive endoscopic images,” *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1201–1214, 2013.
- [56] R. T. Tan and K. Ikeuchi, “Separating reflection components of textured surfaces using a single image,” in *Digitally Archiving Cultural Objects*. Springer, 2008, pp. 353–384.