# Morphological Difference of Closings Operator for No-Reference Quality Evaluation of DIBR-Synthesized Images

Dragana Sandic-Stankovic, Dragan Kukolj, Patrick Le Callet

# Morphological Difference of Closings Operator for No-Reference Quality Evaluation of DIBR-Synthesized Images

Dragana D. Sandiĺ-Stankoviĺ
Belgrade, Serbia
draganasandic68@gmail.com

Dragan D. Kukolj
Department of Computer Engineering,
Faculty of Technical Sciences,
University of Novi Sad
Novi Sad, Serbia
dragan.kukolj@rt-rk.uns.ac.rs

Patrick Le Callet
CNRS, LS2N, UMR 6004
Nantes Université, Ecole Centrale Nantes
F-44000 Nantes, France
patrick.lecallet@univ-nantes.fr

***Abstract*—Images synthesized using Depth-Image-Based Rendering (DIBR) techniques are characterized by complex structural distortion. Multi-resolution multi-scale sparse image representation generated using morphological Difference of Closings operator (DoC) is used to efficiently capture structure-related distortion of synthesized images in the no-reference DoC-GRNN image quality assessment model. Non-linear morphological Difference of Closings operator (DoC) with an array of line-shaped structuring elements of increasing length is used to extract perceptually important details of object structure at different scales and resolutions. The sparsity of DoC band is calculated as scalar feature. The extracted features are mapped to the quality score by general regression neural network (GRNN). We have explored the influence of the direction of an array of line-shaped structuring elements on the model's performances. The DoC-GRNN model shows high agreement with perceptual quality scores, comparable to the state-of-the-art metrics, when evaluated on the stereoscopic DIBR-synthesized images of MCL-3D dataset.**

***Keywords— No-reference image quality assessment, structural image distortion, morphological multi-resolution multi-scale image representation, sparse representation, granulometry, difference of closings operator***

## I. INTRODUCTION

View synthesis using depth-image-based rendering (DIBR) algorithms is important for immersive imaging technologies such as multi-view video and free-viewpoint video with application in various fields: medicine, education, remote surveillance and entertainment. Since only limited number of camera is used to capture different views, new video sequences have to be synthesized at the viewpoints, where captured video sequences are missing from captured color video and depth maps using DIBR algorithm. Since an area around object edges which are invisible from the reference view become visible from the position of the synthesized view, black holes appear in the synthesized images [1]. Although inpainting and interpolation techniques are used for hole filling, a texture distortion may appear. Geometric distortions occur locally at object contours. Distortions in the synthesized images may also appear due to distortions of color images and depth maps, which may occur during the process of acquisition, compression and transmission.

The image quality assessment (IQA) of the synthesized views is of great importance for the development of 3D video applications and for the development of the view synthesis algorithms. In multi-view video systems, reference views are not available for all viewpoints. Therefore, referenceless quality metrics are highly desirable.

It is believed that the human visual system (HVS) is complex nonlinear system which separate images into a set of "spatial frequency" channels as a first encoding of visual information. Sparse image representations that capture the salient structures in line with the human perception are reported to be a processing strategy of the nervous system [2]. In order to mimic the hierarchical structure of HVS, multi-resolution and multi-scale image representations have been used in image quality assessment models. Multi-resolution and multi-scale image representations have also been used in image quality assessment models designed for DIBR-synthesized images: morphological pyramid of Laplacian type [3-5], Gaussian type pyramid [6], multi-resolution representation generated using only down-sampling [7], [8], morphological wavelets [9] and linear wavelets [10], [11]. The sparsity of morphological pyramid and wavelet bands have been used in the image quality model based on general regression neural network [12].

The scale-space framework developed to handle the multi-scale nature of image data has been used in IQA models designed for DIBR-synthesized images. The quality of the whole view synthesis process was estimated based on the array of low-pass smoothed images from Gaussian scale space [13]. The array of band-pass detail images generated by the Difference of Gaussian operator (DoG) of Gaussian scale space, has been used to capture low-order structure (edges) in IQA models [14], [15]. Morphological Difference of Closings (DoC) operator has been used to extract higher-order structures (fine grained details, texture) of high curvature, which corresponds to the distortion of shapes to which the HVS is highly sensitive in the combined DoC-DoG-GRNN model [16]. DoC operator nonlinearly removes redundancies and extracts local structure and contrast to which HVS is highly sensitive. An array of line-shaped structuring elements of increasing sizes has been used by morphological close filter to create a set of simplified signals at different scales. In this work, we explore the influence of the direction of an array of the line-shaped structuring elements used in the generation of DoC bands on the DoC-GRNN modelôs performances. DoC-GRNN model can be successfully used to evaluate complex distortions of

synthesized images which appear due to the rendering process and the compromised quality of color images and depth maps.

In the next section, the proposed model is described in more details. The experimental setup and the performances of the proposed model are presented in Section three.

## II. IMAGE QUALITY MODEL

The framework of the DoC-GRNN model is shown on Fig. 1. In the first stage, multi-resolution and multi-scale image representation using morphological Difference of Closings (DoC) operator is generated to mimic the hierarchical property of human visual system. DoC operator nonlinearly removes redundancies and extracts local structure and contrast to which HVS is highly sensitive. DoC operator extracts structures (fine grained details, texture) of high curvature, which corresponds to the distortion of shapes to which the HVS is highly sensitive. In the second stage, Hoyer index [17] is used to measure the sparsity of DoC band. The extracted features and subjective scores are used to train the general regression neural network (GRNN) which is particularly advantageous with sparse data in real-time environment [18]. In the third stage, the trained network GRNN is used to predict the quality score.

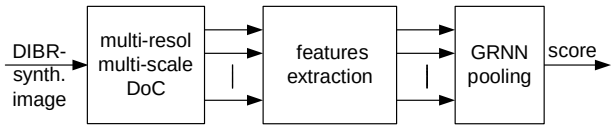


Fig. 1. The framework of the DoC-GRNN model

First, we review the morphological operators used in the calculation of DoC bands. Basic morphological operators, dilation, $f \oplus SE$, and erosion, [illegible], are calculated using operators maximum and minimum on the area of a grayscale image *f,* defined by the flat structuring element *SE* [19]:

$$(f \oplus SE)(x,y) = \max_{u,v \in SE} (f(x-u, y-v)) \quad (1)$$

[illegible] $= \min_{u,v \in SE} (f(x+u, y+v))$ (2)

The morphological filter closing, $f \bullet SE$, which locally modifies the geometric signal features, iteratively applies dilation and erosion (3). Closing filter produces a smooth signal approximation with details removed. Visually, the closing filter smooths contours of image dark regions, tends to fuse narrow breaks, eliminates small holes and fills gaps in the contours. Morphological closing filter is useful for visual perception since it removes valleys from the signal, which correspond to the critical points of contour.

$$f \bullet SE = (f \oplus SE) \ominus SE \quad (3)$$

### A. Multi-resolution multi-scale image decomposition using morphological DoC operator

Image decomposition in the first stage of the framework can be explained using granulometry, a tool for the computation of size distribution in images [20]. At each resolution level *i*, *i=1, ... M*, a series of image approximations, $s_{i,j}$, *j=1, ..., N* with less and less details is created by applying a morphological filter closing to the input image, $s_{i,0}$, using an array of linear structuring elements with increasing length, $SE_j$ (4). The increasing sequence of closings, $\{s_{i,j}\}, j = 1,...N$, constitutes granulometry by closing. A fast algorithm can be used for efficient computation of linear granulometries [21].

$$s_{i,j} = s_{i,0} \bullet SE_j, \quad j = 1,...,N,\ i = 1,...,M \quad (4)$$

The length of the structuring element determines the scale *j*. The structuring element at scale *j*, $SE_j$ is created using the dilation of the structuring element at the first scale, $SE_1$, with itself $j-1$ times (5).

$$SE_j = SE_1 \underbrace{\oplus SE_1 \ldots \oplus SE_1}_{j-1\ times},\ 2 \le j \le N \quad (5)$$

We have explored the morphological filtering using an array of linear structuring elements with increasing length in three directions: 0, 45, 90 degrees. Structuring elements at scale 1, $SE_1$, in three directions are shown on Fig. 2.

Fig. 2. The structuring element of length 2 at scale 1, $SE_1$, in three directions: 0, 45, 90.

At the first resolution level, the input image $s_{1,0}$ is the image under test *f*. To create the initial image for the next resolution level, $s_{i+1,0}$, the initial image from the current resolution level, $s_{i,0}$, is filtered using morphological filter closing with the structuring element $SE_R$ and downsampled by a factor of two ( $\sigma^{\downarrow}$ ) by rows and by columns (6). In this work, we've used square 2x2 as $SE_R$.

$$s_{1,0} = f$$

$$s_{i+1,0} = \sigma^{\downarrow}(s_{i,0} \bullet SE_R)\ \ i = 1,...,M \quad (6)$$

The Difference of Closing (DoC) operator has been defined as the difference between morphologically filtered images of adjacent scales using morphological filter closing (7) [16]. The generation of non-negative integer valued DoC band is shown on Fig. 3.

$$\text{DoC:} \quad d_{i,j} = s_{i,j} - s_{i,j-1}\ \ , j = 1,...,N \quad (7)$$

The DoC bands, $d_{i,j}$, at resolution levels *i=1, ... M*, and scales *j=1, ..., N* constitute multi-resolution and multi-scale image representation (MR-MS-DoC) used in the DoC-GRNN model. Since the morphological operators involve only integers and only max, min and addition, the calculation

of DoC-based image representation has low computational complexity.

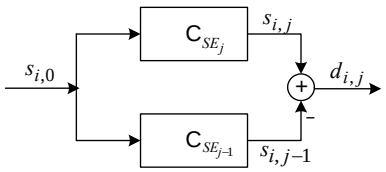


Fig. 3. DoC band is created as the difference between closings of nearby scales; $s_{i,0}$ is the input image at resolution level *i*, $s_{i,j}$ is the low-pass copy and $d_{i,j}$ is the detail image (DoC band) at resolution level *i* and scale *j*.

The DoC bands at resolution level 4 and scale 5, created using line-shaped structuring elements in three directions 0, 45, 90 degrees from the image Kendo synthesized from the content with JPEG2000 distortion are shown on Fig. 4. The oriented lines extracted in DoC bands correspond to perceptually important building blocks of object structures.

Fig. 4. Image Kendo synthesized from the content with JPEG2000 distortion (top); DoC bands at resolution level 4 and scale 5 created using line-shaped structuring elements in three directions: horizontal (bottom left), 45 degrees (bottom middle), vertical (bottom right).

## III. Experimental Results

The proposed metric is evaluated using publicly available dataset of stereoscopic DIBR-synthesized images MCL-3D [22]. The MCL-3D dataset contains 684 synthesized stereoscopic image pairs synthesized from nine image-plus-depth sources and associated mean opinion score (MOS) values. One third of images are of resolution 1024×728 and two thirds are of resolution 1920×1080. The majority of images (648 stereo images) are synthesized using both left and right views distorted by one of six distortion types: Gaussian blur, additive white noise, down-sampling blur, JPEG and JPEG-2000 compression and transmission error. Distortions have been symmetrically applied to the left and right viewpoints to either the color images or the depth maps or both the color images and the depth maps at four distortion levels before stereoscopic image rendering using the View Synthesis Reference Software [23]. The minority of the images (36 stereo images) have been synthesized from the undistorted color images and depth maps of single view using four DIBR synthesis algorithms. They contain only the distortions caused by imperfect rendering.

In order to evaluate the performances of the proposed metric and other metrics, three evaluation criteria were used: Root Mean Squared Error (RMSE) to compute the prediction error, Pearson Linear Correlation Coefficient (PLCC) to compute prediction accuracy and Spearman Rank-order Correlation Coefficient (SROCC) for prediction monotonicity.

In the evaluation of the DoC-GRNN model, the k-fold cross validation strategy is used to split the dataset with D images to k=5 disjoint test subsets of similar size, each with D/k images (20% images of the dataset), and to select the train subsets with D-D/k images (80% images of the dataset) such that there is not overlap between the test and the train subsets. At each iteration of the k-fold cross validation strategy, GRNN is used to predict the scores of the test subset. The median value of the performances through cross-validation iterations and through 1000 repetitions is calculated as the final modelôs performances.

We have explored modelôs performances using three arrays of line-shaped structuring elements in three directions: 0, 45, 90 degrees, in the multi-resolution and multi-scale DoC-based decomposition. When the array of horizontal structuring elements (0 degrees) is used, the best performances are achieved using M=7 resolution levels and N=5 scales in the MR-MS-DoC decomposition and 36 features are extracted. In the case when the array of structuring elements with direction 45 degrees is used, the best performances are achieved using M=7 resolution levels and N=6 scales, leading to 43 features. When the array of vertical structuring elements is used, the best performances are achieved using M=4 resolution levels and N=6 scales leading to 25 features. The only parameter of GRNN is the spread, that represents the width of radial basis function of GRNN. The PLCC of the DoC-GRNN scores using three arrays of structuring elements in different directions, 0, 45, 90, with fluctuation of the GRNN spread parameter is shown on Fig. 5.

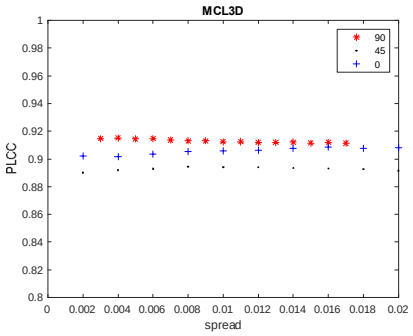


Fig. 5. PLCC of the DoC-GRNN scores for different directions of an array of line-shaped structuring elements used in the DoC-based decomposition.

TABLE I. Performance of the DoC-GRNN Model using Three Arrays of Line-shaped Structuring Elements in Three Directions

| direction (degrees) | num. features | spread | PLCC | SROCC | RMSE |
|---|---|---|---|---|---|
| 0 | 36 | 0.016 | 0.9084 | 0.9092 | 1.1222 |
| 45 | 43 | 0.008 | 0.8943 | 0.8966 | 1.2055 |
| **90** | **25** | 0.004 | **0.9147** | 0.9091 | **1.0750** |

The best performances of DoC-GRNN model using three arrays of structuring elements in three directions are shown in Table I. The results show that the performance dependency from the direction of line-shaped structuring elements is small. The lowest number of features, 25, is

needed when the array of vertical structuring elements is used and the performances are slightly better in that case. The performances are slightly worse when the array of structuring elements with orientation 45 degrees is used.

The performances of the proposed model are compared to the no-reference DoG-GRNN metric proposed for DIBR-synthesized images [14] and to 2D IQA metrics for natural images which use reference: PSNR, SSIM, MS-SSIM, IW-SSIM, Table II. The DoC-GRNN metric shows reliable agreement with human scores, PLCC=0.914, SROCC=0.909. High performances are achieved mostly thanks to the capacity of morphological series to capture higher order properties of spatial random processes. DoG-GRNN model, using 31 features calculated from DoG-based image representation in 5 resolution levels and 6 scales, achieves similar performances as the DoC-GRNN model. Among the metrics proposed for natural images, IW-SSIM achieves the highest performances in the quality evaluation of MCL-3D dataset.

TABLE II. PERFORMANCES OF THE DOC-GRNN METRIC AND OTHER METRICS

| Metric | | PLCC | SROCC | RMSE |
|---|---|---|---|---|
| NR | **DoC-GRNN** | **0.9147** | **0.9091** | **1.0750** |
| | DoG-GRNN | 0.9041 | 0.8988 | 1.1397 |
| FR | PSNR | 0.7779 | 0.7900 | 1.6397 |
| | SSIM | 0.7539 | 0.7741 | 1.7144 |
| | MS-SSIM | 0.8386 | 0.8518 | 1.4213 |
| | IW-SSIM | 0.8694 | 0.8878 | 1.2894 |

Several more reference-based metrics designed for the evaluation of natural images can be successfully used to assess the quality of images from MCL-3D dataset [16]. Reference-based metrics, designed to predict the quality of DIBR-synthesized images show limited performances in the evaluation of MCL-3D dataset [16]. No-reference machine learning-based techniques proposed for the quality evaluation of DIBR-synthesized images achieve the best performances in the evaluation of MCL-3D dataset [16].

## IV. CONCLUSION

No-reference training-based DoC-GRNN model uses sparsity of Difference of Closings (DoC) bands as features. DoC bands which contain perceptually important structure-related information at different scales and resolutions are generated by morphological filtering using an array of line-shaped structuring elements. We have analysed the performances of DoC-GRNN model using an arrays of line-shaped structuring elements in three directions: horizontal, 45 degrees, vertical. The performances are slightly higher using vertical line-shaped structuring elements and the feature number is the lowest in that case. The DoC-GRNN model achieves high performance in the evaluation of DIBR-synthesized images with rendering distortions and the distortion which appear during acquisition, compression and transmission of color images and depth maps. The code for feature extraction is available at https://sites.google.com/site/draganasandicstankovic/code/doc.

## REFERENCES


[1] E. Bosc, R Pepion, P. Le Callet, M. Koppel, P.Ndjiki-Nya, M.Pressigout, L. Morin, ñTowards a New Quality Metric for 3-D Synthesized View Assessment,ò IEEE Journal on Selected Topics in Signal Processing,vol. 5, no. 7, pp. 1332-1343, Nov. 2011.

[2] B. Olshausen, D. Field, ñSparse coding of sensory inputs,ò Current Opinion in Neurobiology, vol. 14, no. 4, pp. 481-487, Aug. 2004.

[3] D. SandiĹ-StankoviĹ, D. Kukolj, and P. Le Callet, ñDIBR synthesized image quality assessment based on morphological multiscale approach,ò EURASIP J. Image and Video Process. 2017: 4, 2016.

[4] D. SandiĹ-StankoviĹ, D. Kukolj, P. Le Callet, ñDIBR synthesized image quality assessment based on morphological pyramids,ò 3DTV-CON Immers. inter. 3D media exp. over networks, Lisbon, July 2015.

[5] D. SandiĹ-StankoviĹ, D Kukolj, P Le Callet, òMulti-scale synthesized view assessment based on morphological pyramid,ò Journal of Electrical Engineering, Vol. 67 (1), pp. 1ï9, 2016.

[6] Y. Zhou, L. Yang, L. Li, K. Gu, L Tang, ñReduced-reference quality assessment of DIBR-synthesized images based on multi-scale edge intensity similarity,ò Multim. Tools and Appl., pp.1-20, Dec. 2017.

[7] K. Gu, J.F. Qiao, P. Le Callet, Z. Xia, W. Lin, ñUsing multiscale analysis for blind quality assessment of DIBR-synthesized images,ò *IEEE Int. Conf on Image Processing*, 2017.

[8] K. Gu, J Qiao, S Lee, H Liu, W. Lin, P. Le Callet, ñMultiscale Natural Scene Statistical Analysis for No-Reference Quality Evaluation of DIBR-Synthesized Views,ò IEEE Trans. Broadcast, 2019.

[9] D. SandiĹ-StankoviĹ, D. Kukolj, P. Le Callet, ñDIBR synthesized image quality assessment based on morphological wavelets,ò Int. Workshop on Quality of Multimedia Experience QoMEX, Costa Navarino, Greece, May 2015.

[10] E. Bosc, F. Battisti, M. Carli and P.L.Callet, ñA wavelet-based image quality metric for the assessment of 3D synthesized views,ò Proc. SPIE 8648, Stereoscopic Displays and Applications, March 2013.

[11] G. Wang, Z. Wang, K. Gu, L. Li, Z. Xia, L.Wu, ñBlind Quality Metric of DIBR-Synthesized Images in the Discrete Wavelet Trasform Domain,ò IEEE Trans Image Process., Oct. 2019.

[12] D. Bokan, G VelikiĹ, D Kukolj, D SandiĹ-StankoviĹ, ñBlind DIBR-synthesized image quality assessment based on sparsity features in morphological multiscale domain,ò Int. Conf. on Quality of Multimedia Experience (QoMEX), 2017.

[13] Y Zhou, L. Li, S. Ling, P Le Callet, ñQauality assessment for view synthesis using low-level and mid-level structural representation,ò Signal Processing: Image Communication, pp.309-321, March 2019.

[14] D. SandiĹ-StankoviĹ, D. Bokan, D. Kukolj, ñBlind DIBR-synthesized Image Quality Assessment using multi-scale DoG and GRNN,ò Symposium on Neural Networks and Applications, Nov. 2018.

[15] Y. Zhou, L. Li, S. Wang, J. Wu, Y. Fang, and X. Gao, ñNo-Reference Quality Assessment for View Synthesis Using DoG-based Edge Statistics and Texture Naturalness,ò IEEE Trans Image Process., vol. 28, issue 9, pp. 4566-4579, Apr. 2019.

[16] D. SandiĹ-StankoviĹ, D. Kukolj, P. Le Callet, "Quality Assessment of DIBR-synthesized views based on Sparsity of Difference of Closings and Difference of Gaussians", IEEE Transactions on Image Processing, vol. 31, pp. 1161-1175, January 2022.

[17] P. Hoyer, ñNon-negative Matrix Factorization with Sparseness Constraints,ò Jour. Machine Learn. Research 5, pp. 1457-1469, 2004.

[18] D. F. Specht, ñA general regression neural network,ò IEEE Trans. Neural Networks, vol. 2, no. 6, pp. 568ï576, Nov. 1991.

[19] R. Haralick, S. Sternberg, X. Zhuang, ñImage Analysis using Mathematical Morphology,ò IEEE Trans. On Pattern Analysis and Machine Intelligence, vol.9, no. 4, pp. 532-550, July 1987.

[20] P. Maragos, ñPattern Spectrum and Multiscale Shape Representation,ò IEEE Trans. Pattern and Machine Intelligence, vol. 11, no.7, July 1989.

[21] L. Vincent, ñFast grayscale granulometry algorithms,ò Int. Symposium on Mathematical Morphology, pp. 265-272, Sept 1994.

[22] R. Song, H. Ko, C C J Kuo, ñMCL-3D: A database for stereoscopic image quality assessment using 2D-image-plus-depth source,òJour. of Inform. Science Engineer., vol.31, no. 5, pp. 1593-1611, Sept. 2015.

[23] M Tanimoto, T.Fujii, and K.Suzuki, ñView synthesis algorithm in view synthesis reference software 3.5 (VSRS3.5),ò (2009)