

RESEARCH

Open Access



DIBR-synthesized image quality assessment based on morphological multi-scale approach

Dragana Sandić-Stanković^{1*}, Dragan Kukulj² and Patrick Le Callet³

Abstract

The depth-image-based rendering (DIBR) algorithms used for 3D video applications introduce new types of artifacts mostly located around the disoccluded regions. As the DIBR algorithms involve geometric transformations, most of them introduce non-uniform geometric distortions affecting the edge coherency in the synthesized images. Such distortions are not handled efficiently by the common image quality assessment metrics which are primarily designed for other types of distortions. In order to better deal with specific geometric distortions in the DIBR-synthesized images, we propose a full-reference metric based on multi-scale image decomposition applying morphological filters. Using non-linear morphological filters in multi-scale image decomposition, important geometric information such as edges is maintained across different resolution levels. Edge distortion between the multi-scale representation subbands of the reference image and the DIBR-synthesized image is measured precisely using mean squared error. In this way, areas around edges that are prone to synthesis artifacts are emphasized in the metric score. Two versions of morphological multiscale metric have been explored: (a) Morphological Pyramid Peak Signal-to-Noise Ratio metric (MP-PSNR) based on morphological pyramid decomposition, and (b) Morphological Wavelet Peak Signal-to-Noise Ratio metric (MW-PSNR) based on morphological wavelet decomposition. The performances of the proposed metrics have been tested using two databases which contain DIBR-synthesized images: the IRCCyN/IVC DIBR image database and MCL-3D stereoscopic image database. Proposed metrics achieve significantly higher correlation with human judgment compared to the state-of-the-art image quality metrics and compared to the tested metric dedicated to synthesis-related artifacts. The proposed metrics are computationally efficient given that the morphological operators involve only integer numbers and simple computations like min, max, and sum as well as simple calculation of MSE. MP-PSNR has slightly better performances than MW-PSNR. It has very good agreement with human judgment, Pearson's 0.894, Spearman 0.77 when it is tested on the MCL-3D stereoscopic image database. We have demonstrated that PSNR has particularly good agreement with human judgment when it is calculated between images at higher scales of morphological multi-scale representations. Consequently, simplified and in essence reduced versions of multi-scale metrics are proposed, taking into account only detailed images at higher decomposition scales. The reduced version of MP-PSNR has very good agreement with human judgment, Pearson's 0.904, Spearman 0.863 using IRCCyN/IVC DIBR image database.

Keywords: DIBR-synthesized image quality assessment, Multi-scale IQA metric using morphological operations, Geometric distortions, Morphological pyramid, Morphological wavelets

* Correspondence: dragana.sandic@iritel.com

¹Institute for Telecommunications and Electronics IRITEL, Belgrade, Serbia
Full list of author information is available at the end of the article

1 Introduction

The advanced 3D video (3DV) systems are mostly based on multi-view video plus depth (MVD) format [1] as the recommended 3D video format adopted by the moving picture experts group (MPEG). In the 3DV system, smaller number of captured views is transmitted and greater number of views is generated at the receiver side from the transmitted texture views and their associated depth maps using depth-image-based rendering (DIBR) technology. DIBR techniques can be used to generate views for different 3D video applications: free viewpoint television, 3DTV, 3D technology based entertainment products, and 3D medical applications. The perceptual quality of the synthesized view is considered as the most significant evaluation criterion for the whole 3D video processing system. Reliable quality assessment metric for synthesized views is of a great importance for the 3D video technology development. The use of subjective tests is expensive, time consuming, cumbersome, and practically no feasible in systems where real-time quality score of an image or video sequence is needed. Objective metrics are intended to predict human judgment. The reliability of objective metrics is based on their correlation to subjective assessment results.

The evaluation of DIBR system depends on the application. The main difference between free viewpoint video (FVV) and 3DTV is the stereopsis phenomenon (fusion of left and right views in human visual system) existing in 3DTV. FVV does not have to be used in 3D context. It can be applied in 2D context. In this paper, the quality assessment of still images from MVD video sequences in both 2D and 3D contexts as a first step of 3D quality assessment is concerned. The evaluation of still images is important scenario in the case when the user switches the video in pause mode [2].

For the comparison of DIBR algorithms, virtual views synthesized from the uncompressed data which contain only synthesis artifact need to be evaluated. When encoding either depth data or color sequences before performing the synthesis, compression-related artifacts are combined with synthesis artifact. In this paper, the distortions introduced only by view synthesis algorithms are evaluated using the IRCCyN/IVC DIBR image dataset [3, 4] and part of the MCL-3D image dataset [5, 6].

DIBR algorithms introduce new types of artifacts mostly located around disoccluded regions [2]. They are not scattered in the entire image such as 2D video compression distortions. As DIBR algorithms involve geometric transformations, most of them introduce mainly geometric distortions affecting edges coherency in the synthesized images. These artifacts are consequently challenging for standard quality metrics, usually tuned for other types of distortions. In order to better deal

with specific geometric distortions in DIBR-synthesized images, we propose multi-scale image quality assessment metric based on morphological filters in multi-resolution image decomposition. Due to multi-scale character of primate visual system [7], the introduction of multi-resolution image decomposition in the image quality assessment contributes to the improvement of metric performances relative to single-resolution method. Introduced non-linear morphological filters in multi-resolution image decomposition maintain important geometric information such as edges on their true positions, neither drifted nor blurred, across different resolution levels [8]. Edge distortion between appropriate subbands of the multi-scale representations of the reference image and the DIBR-synthesized image is precisely measured pixel-by-pixel using mean squared error (MSE). In this way, areas around edges that are prone to synthesis artifacts are emphasized in the metric score. Mean squared errors of subbands are combined into multi-scale mean squared error, which is transformed into multi-scale peak signal-to-noise ratio measure. More precisely, two types of morphological multi-scale decompositions for the multi-scale image quality assessment (IQA) have been explored: morphological bandpass pyramid decomposition in the Morphological Pyramid Peak Signal-to-Noise Ratio measure (MP-PSNR) and morphological wavelet decomposition in the Morphological Wavelet Peak Signal-to-Noise Ratio measure (MW-PSNR). Morphological bandpass pyramid decomposition can be interpreted as a structural image decomposition tending to enhance image features such as edges which are segregated by scale at the various pyramid levels [9]. Using non-linear morphological wavelet decomposition, geometric structures such as edges are better preserved in the lower resolution images compared to the case when the linear wavelets are used in the decomposition [10]. Both separable and true non-separable morphological wavelet decompositions using the lifting scheme have been investigated.

Both measures, MP-PSNR and MW-PSNR, are highly correlated with the judgment of human observers, much better than standard IQA metrics and much better than their linear counterparts. They have better performances than tested metric dedicated to synthesis-related artifacts also. Since the morphological operators involve only integers and only max, min, and addition in their computation, as well as simple calculation of MSE, the proposed morphological multi-scale metrics are of low computational complexity.

Moreover, it is experimentally shown that PSNR has very good agreement with human judgment when it is calculated for the subbands at higher morphological decomposition scales. We propose the reduced versions of morphological multi-scale measures, reduced MP-PSNR, and reduced MW-PSNR, using only detail images from higher decomposition scales. The performances of the

reduced versions of the morphological multi-scale measures are improved comparing to their full versions.

In the next section, the distortion of the DIBR-synthesized view is shortly described. Previous work on the quality assessment of the DIBR-synthesized views and multi-scale image quality assessment is also shortly reviewed in Section 2. In Section 3, we describe two versions of the proposed multi-scale metric, based on two types of multi-resolution decomposition schemes, morphological pyramid, and morphological wavelets. Description of the distortion computation stage and pooling stage of the proposed multi-scale measures is given also in Section 3. The performances of MP-PSNR and MW-PSNR and discussion of results are presented in Section 4, while the conclusion is given in Section 5.

2 Related works

2.1 Distortion in the DIBR-synthesized view

The synthesis process changes the pixels position in the synthesized image and induces new types of distortion in DIBR-synthesized views. View synthesis noise mainly appears along object edges. Typical DIBR artifacts include object shifting, geometric distortions, edge displacements or misalignments, boundary blur, and flickering. Incorrect depth map induces object shifting in the synthesized image. Object shifting artifact or ghost artifact manifests as slight translation or resize of an image regions due to depth map errors. A large number of tiny geometric distortions are caused by the depth inaccuracy and the numerical rounding operation of pixel positions. Geometric distortions appear in the synthesized images because the pixels are projected to wrong positions. Blurry regions appear due to inpainting method used to fill the disoccluded areas. Incorrect rendering of textured areas appears when inpainting method fails in filling complex textured areas. When the objects move, the distortion around edges is more noticeable. The view synthesis distortion flickering locates on the edge of the foreground object which has a movement. Flickering can be observed as significant and high-frequency alternated variation between different luminance levels [11]. The temporal flicker distortion is the most significant difference between the traditional 2D video and the synthesized video. Some of the typical artifacts due to DIBR synthesis are shown on Fig. 1.

2.2 Quality assessment of DIBR-synthesized view

The evaluation of DIBR views synthesized from uncompressed data using standard image quality metrics has been discussed in literature for still images from FVV in 2D context [3] using IRCCyN/IVC DIBR image database. It has been demonstrated that 2D quality metrics originally designed to address image compression distortions are very far to be effective to assess the visual quality of synthesized views.



Fig. 1 Typical artifacts due to DIBR synthesis. Original images are in the left column and synthesized images are in the right column

Full-reference objective image quality assessment metrics, VSQA [12], and 3DswIM [13], have been proposed to improve the performances obtained by standard quality metrics in the evaluation of the DIBR-synthesized images. Both metrics are dedicated to synthesis-related artifacts without compression-related artifacts and both metrics are tested using IRCCyN/IVC DIBR images dataset. VSQA [12] metric dedicated to view synthesis quality assessment is aimed to handle areas where disparity estimation may fail. It uses three visibility maps which characterize complexity in terms of textures, diversity of gradient orientations, and presence of high contrast. SSIM-based VSQA metric achieves the gain of 17.8 % over SSIM in correlation with subjective measurements. 3DswIM [13], relies on a comparison of statistical features of wavelet subbands of the original and DIBR-synthesized images. Only horizontal detail subbands from the first level of Haar wavelet decomposition are used for the degradation measurement. A registration step is included before the comparison to ensure shifting-

resilience property. A skin detection step weights the final quality score in order to penalize distorted blocks containing skin-pixels based on the assumption that a human observer is most sensitive to impairments affecting human subjects. It was reported that 3DswIM metric outperforms the conventional 2D metrics and tested DIBR-synthesized views dedicated metrics.

Edge-based structural distortion indicator addressing the distortion related to DIBR systems is proposed in [14]. The method relies on the analysis of edges in the synthesized view. The proposed method does not assess the image quality, but it is able to detect the structural distortion. Since it does not take the color consistency into account, the method remains a tool for assessing the structural consistency of an image.

Vision-based quality measures for 3D DIBR-based video, both full-reference FR-3VQM [15], and no-reference NR-3VQM [16] are proposed to evaluate the quality of stereoscopic 3D video generated by DIBR. Both measures are a combination of three measures: temporal outliers, temporal inconsistencies, and spatial outliers, using ideal depth. Ideal depth is derived for both no-reference and for full-reference metric for distortion-free rendered video. 3VQM metrics show better performances than PSNR and SSIM using a database of DIBR-generated video sequences.

Quality metric proposed in [17] is designed for the evaluation of synthesized images which contain artifacts introduced by the rendering process due to depth map errors. It consists of two parts. One part is the calculation of the conventional 2D metric after the consistent object shifts. After shift compensation, the 2D QA model matches the subjective quality score better. The other part is the calculation of the structural score by the Hausdorff distance. The Hausdorff distance identifies the degree of the inconsistent object shift or ghost-type artifact at object boundaries. The proposed metric shows better performances than traditional IQA metrics in the evaluation of synthesized stereo images from MVD video sequences.

SIQE metric [18] proposed to estimate the quality of DIBR-synthesized images compares the statistical characteristics of the synthesized and the original views estimated using the divisive normalization transform. In the evaluation of compressed MVD video sequences, it achieves high correlation with widely used image and video quality metrics.

A full-reference video quality assessment of synthesized view with texture/depth compression presented in [11] focuses on the temporal flicker distortion due to depth compression distortion and the view synthesis process. It is based on two quality features which are extracted from both spatial and temporal domains of the synthesized sequence. The first feature focuses on

capturing the temporal flicker distortion and the second feature is used to measure the change of the spatio-temporal activity in the synthesized sequence due to blurring and blockiness distortion caused by texture compression. The performances of the proposed metric evaluated on the synthesized video quality database SIAT [11] are better than the performances of the commonly used image/video quality assessment methods.

2.3 Multi-scale image quality assessment

As in most other areas of image processing and analysis, multi-resolution methods have improved performances relative to single-resolution methods also for the image quality assessment. Pyramids and wavelets are among the most common tools for constructing multi-resolution signal decomposition schemes used in image processing and computer vision. Both redundant image pyramid representation and non-redundant image wavelet representations have been explored for multi-scale image quality assessment metrics.

Multi-scale structural similarity measure, MS-SSIM [19] is based on linear low-pass pyramid decomposition. Multi-scale image quality measures using information content weighted pooling, IW-SSIM, and IW-PSNR [20], use Laplacian pyramid decomposition [21]. CW-SSIM [22] simultaneously insensitive to luminance and contrast changes and small geometric distortions of image is based on multi-orientation steerable pyramid decomposition using multi-scale bandpass-oriented filters.

It has been shown that the local contrast in different resolutions can be easily represented in terms of Haar wavelet transform coefficients and computational models of visual mechanisms were incorporated into a quality measurement system [23]. Experiments have shown that Haar filters have good ability to simulate the human visual system (HVS) and the proposed metric is successful in measuring compressed image artifacts.

Error-based image quality metric using Haar wavelet decomposition has been proposed in [24]. It has been reported that Haar wavelet provided more accurate quality scores than other wavelet bases. PSNR has been calculated between the edge maps calculated from detail subbands as well as between approximation subbands of the original and the distorted images. These two PSNR have been linearly combined to the overall quality score. The proposed metric predicts quality scores more accurately than the conventional PSNR and can be used efficiently in real-time applications.

Reduced-reference image quality assessment based on multi-scale geometric analysis (MGA) to mimic multi-channel structure of HVS, contrast sensitivity function to re-weights MGA coefficients to mimic nonlinearities in HVS and the just noticeable difference threshold to remove visually insensitive MGA coefficients has been

presented in [25]. The quality of the distorted image was measured by comparing the normalized histograms of the distorted and the reference images. MGA was utilized to decompose images by a series of transforms including wavelet, curvelet, bandelet, contourlet, wavelet-based contourlet, hybrid wavelets, and directional filter banks. MGA can capture the characteristics of image, e.g., lines, curves, contour of object. IQA based on MGA and IQ metric using Haar wavelet decomposition [24] have been evaluated on the database which contains compressed, white noisy, Gaussian-blurred, and fast-fading Rayleigh channel noisy images.

3 Proposed morphological multi-scale metric

Multi-scale image quality assessment (IQA) framework can be described as three-stage process. In the first stage, both the reference and the distorted images are decomposed into a set of lower resolution images using multi-resolution decomposition. In the second stage, image quality/distortion maps are evaluated for all subbands at all scales. In the third stage, a pooling is employed to convert each map into a quality score, and these scores are combined into the final multi-scale image quality measure score.

The key stage of the multi-scale image quality assessment may be how to represent images effectively and efficiently, so it is necessary to investigate various kinds of transforms. Most of the current multi-scale IQA metrics use linear filters in the multi-resolution decomposition. In this paper, we propose to use non-linear morphological operators in the multi-scale decompositions in the first stage of multi-scale IQA framework, Fig. 2, in order to better deal with specific geometric distortions in DIBR-synthesized images. Introduced non-linear morphological filters used in the multi-scale image decomposition maintain important geometric information such as edges on their true positions, across different resolution levels [8]. More precisely, we investigate two types of morphological multi-scale decompositions in the first stage of multi-scale IQA framework: morphological bandpass pyramid decomposition in MP-PSNR and morphological wavelet decomposition in MW-PSNR. In the second stage of the multi-scale IQA framework, Fig. 2, we propose to calculate squared error maps between the appropriate images of the multi-scale representations of the two images, the

reference image and the DIBR-synthesized image, in order to measure precisely, pixel-by-pixel, the edge distortion. In this way, the areas around edges that are prone to synthesis artifacts are emphasized in the metric score. In the third stage of IQA multi-scale framework, MSE is calculated from each squared error map. MSE of all multi-scale representation images are combined into multi-scale mean squared error, which is transformed into morphological multi-scale peak signal-to-noise ratio measure.

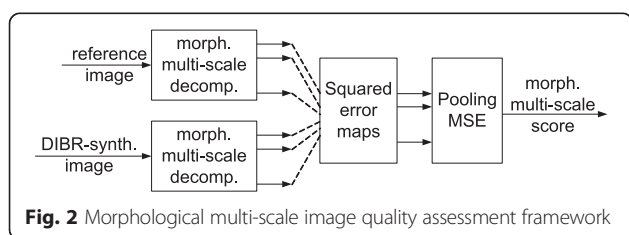
3.1 Morphological multi-scale image decomposition

The importance of analyzing images at many scales arises from the nature of images themselves [26]. Scenes contain objects of many sizes and these objects contain features of many sizes. Objects can be at various distances from the viewer. Any analysis procedure that is applied only at a single-scale may miss information at other scales. The solution is to carry out analysis at all scales simultaneously. Psychophysics and physiological experiments have shown that multi-scale transforms seem to appear in the visual cortex of mammals [27].

A multi-scale representation is completely specified by the transformation from a finer scale to a coarser scale. In linear scale-spaces the operator for changing scale is a convolution by a Gaussian kernel. After the convolution with Gaussian kernel the images are uniformly blurred, also the regions of particular interest like the edges [28]. This is a drawback as the edges often correspond to the physical boundaries of objects. The edge and contour information may be the most important of an image's structure for human to capture the scene. To overcome this issue, non-linear multi-resolution signal decomposition schemes based on morphological operators have been proposed to maintain edges through scales [8].

In morphological image processing, geometric properties such as size and shape are emphasized rather than the frequency properties of signals. Mathematical morphology [29, 30] is a set-theoretic method for image analysis which provides a quantitative description of geometric structure of an image. It considers images as sets which permits geometry-oriented transformations of the images. The structuring element offers flexibility because it can be designed in different shapes and sizes according to the purpose. Morphological filters are non-linear signal transformations that locally modify geometric signal features.

In the first stage of morphological multi-scale IQA framework, we have explored two types of multi-scale image decomposition using morphological pyramid and morphological wavelets.



3.1.1 Multi-scale image decomposition using morphological pyramid

The image pyramid offers a flexible, convenient multi-resolution format that matches the multiple scales found in the visual scenes and mirrors the multiple scales of processing in the human visual system [26]. Pyramid representations have much in common with the way people see the world, i.e., primate visual systems achieve a multi-scale character [7].

In this paper, we propose to use morphological bandpass pyramid (MBP) decomposition in the first stage of morphological multi-scale IQA framework. Morphological bandpass pyramid is generated using the Laplacian type pyramid decomposition scheme [21], but instead of linear filters, morphological filters are used. We propose to use morphological operator erosion (E) for low-pass filtering in analysis step and morphological operator dilation (D) for interpolation filtering in synthesis step leading to the morphological bandpass pyramid decomposition erosion/dilation (MBP ED) introduced in [31] and reviewed in [32]. One level of the proposed MBP ED pyramid is shown on Fig. 3.

In the MBP ED scheme, Fig. 3, a lower resolution image s_{j+1} is obtained by applying morphological operator erosion on the previous pyramid level image s_j and downsampling the eroded image by factor 2 on both image dimensions (σ^{\downarrow}) (1). We've used the square structuring element of size $(2r+1) \times (2r+1)$, $r=1, \dots, 6$ for erosion.

$$s_{jE}(m, n) = \min \{ s_j(m+k, n+l), \quad | -r \leq k, l \leq r \} \quad (1)$$

$$s_{j+1} = \sigma^{\downarrow}(s_{jE})$$

The erosion as the analysis operator removes fine details smaller than the structuring element. A detail image is derived by subtracting from each level an interpolated version of the next coarser level. The image s_{j+1} of the next pyramid level is upsampled by factor 2 on both dimensions (σ^{\uparrow}) leading to the image

s_{jU} . Morphological operator dilation is applied on the upsampled image s_{jU} to produce expanded image \hat{s}_j . The detail image d_j is obtained as the difference of the pyramid image s_j and expanded image from the next pyramid level \hat{s}_j :

$$s_{jU} = \sigma^{\uparrow}(s_{j+1})$$

$$\hat{s}_j(m, n) = \max \{ s_{jU}(m-k, n-l), \quad | -r \leq k, l \leq r \} \quad (2)$$

$$d_j = s_j - \hat{s}_j$$

Using square structuring element, morphological reduce and expand filtering can be implemented more efficiently separably by rows and columns using the structuring elements of size $1 \times (2r+1)$ for rows and $(2r+1) \times 1$ for columns.

Morphological bandpass pyramid with M decomposition levels consists of detail (error) images of decreasing size d_j , $j=0, \dots, M-1$ and the coarse lowest resolution image s_M [9]. MBP ED pyramid generated using SE of size 7×7 of the synthesized frame from the video sequence Newspaper is shown on Fig. 4.

MBP ED pyramid based on adjunction satisfies the property that the detail signal is always non-negative. At any scale change, maximum luminance at the coarser scale is always lower than the maximum luminance at the finer scale, the minimum is always higher. Morphological bandpass pyramid decomposition can be interpreted as a structural image decomposition tending to enhance image features such as edges which are segregated by scale at the various pyramid levels [9]. Enhanced features are segregated by size: fine details are prominent in the lower level images while progressively coarser features are prominent in the higher level images. MBP ED pyramid using structuring element of size 2×2 is morphological Haar pyramid [31]. MBP satisfies pyramid condition [31] which states that synthesis of a signal followed by analysis returns the original signal, meaning that no information is lost by these two consecutive steps and the original image can be perfectly reconstructed from the pyramid representation. Perfect reconstruction, while not mandatory for image quality assessment is a valuable property for a representation in early vision not because a visual system needs to literally reconstruct the image from its representation but rather because it guarantees that no information has been lost, ie that if two images are different then their representations are different also [7]. There is neurophysiological evidence that the human visual system uses a similar kind of decomposition [33]. There is inherent congruence between the morphological pyramid decomposition scheme and human visual perception [9].

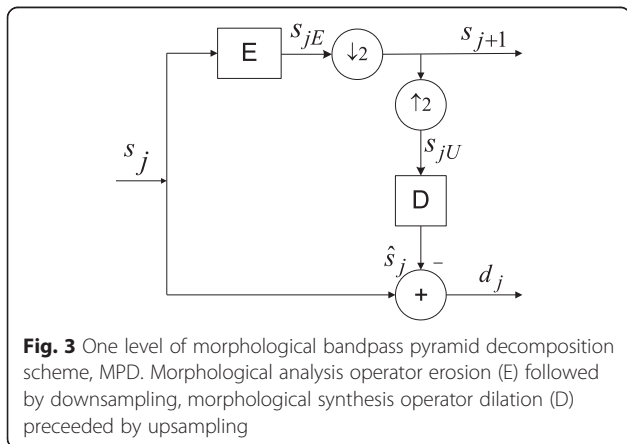


Fig. 3 One level of morphological bandpass pyramid decomposition scheme, MPD. Morphological analysis operator erosion (E) followed by downsampling, morphological synthesis operator dilation (D) preceded by upsampling

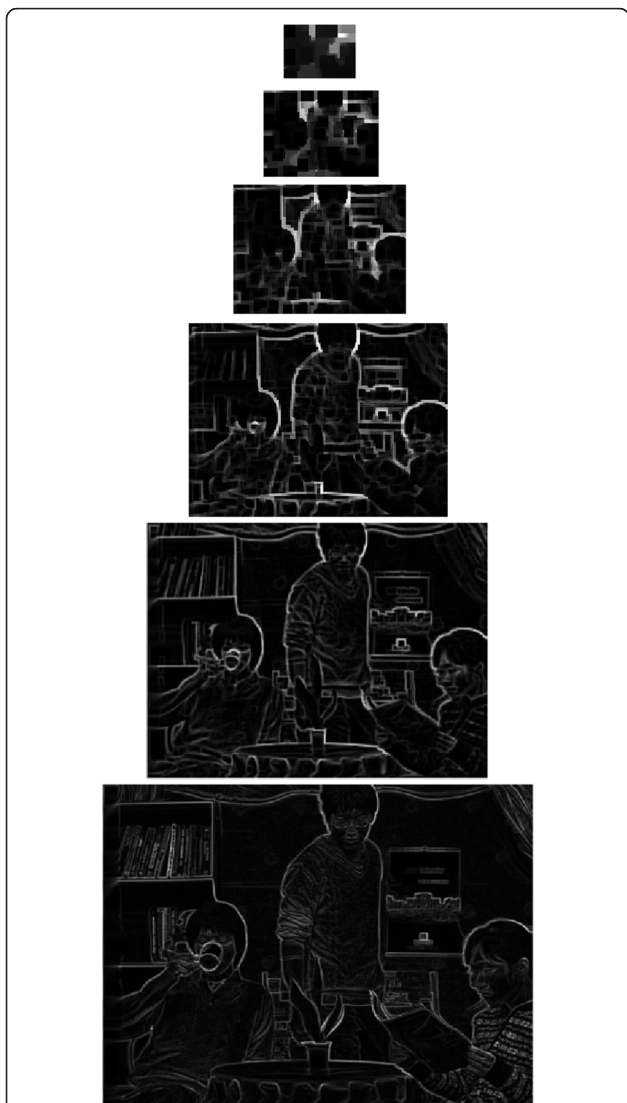


Fig. 4 Morphological bandpass pyramid representation of the synthesized frame from the video sequence Newspaper. Squared structuring element of size 7×7 is used for morphological reduce filtering in MBP ED

3.1.2 Multi-scale image decomposition using morphological wavelets

Most current image quality assessment methods based on discrete wavelet transform use linear wavelet kernels [23, 24, 34]. In this paper, we propose to use morphological wavelet decomposition in order to better preserve geometric structures such as edges in the lower resolution images. The morphological wavelet transforms introduced in [10] and reviewed in [32] are non-linear wavelet transforms that use min and max operators. Due to non-linear nature of the morphological operators, important geometric information such as edges are well preserved across different resolution levels. A general and flexible approach for the construction of non-linear

morphological wavelets in the spatial domain is provided by the lifting scheme using morphological lifting operators in prediction (P) step and update (U) step [35], Fig. 5. We have explored both separable and true non-separable morphological wavelet decompositions using the lifting scheme.

Separable 2D discrete wavelet transform (DWT) is implemented by cascading two 1D DWT along the vertical and horizontal directions [36] producing three detail subbands and approximation signal. Separable wavelet decompositions using 1D morphological Haar wavelet (minHaar) and 1D morphological wavelet using min-lifting scheme (minLift) [10, 37] are explored. Their linear counterparts, Haar wavelet and biorthogonal wavelet of Cohen-Daubechies-Feauveau (cdf (2,2)) [38], are also tested for comparison.

Non-separable sampling opens a possibility of having schemes better adapted to the human visual system [39]. Non-separable 2D morphological wavelet decomposition on a quincunx lattice using the min-lifting scheme (minLiftQ) [40] is also explored. Non-separable wavelet decomposition with linear wavelet of Cohen-Daubechies-Feauveau on a quincunx lattice (cdf(2,2)Q) [41] is implemented for comparison.

- 1D Morphological Haar min wavelet transformation (minHaar)

One of the simplest example of non-linear morphological wavelets is the morphological Haar wavelet (minHaar) [10]. It is very similar structure to the linear Haar but it uses non-linear morphological operator erosion (by taking the minimum over two samples) in the update step of the lifting scheme [32, 37]. An illustration of one step of the wavelet transform with minHaar wavelet using the lifting scheme is shown on Fig. 6. Initially, the signal x (the first row in Fig. 6) is splitted to the even samples array (white nodes) and odd samples array (black nodes). The detail signal d (middle row in Fig. 6) is calculated as the difference of the odd array and the even array (3). The lower resolution signal s (bottom row in Fig. 6) is calculated from the even array and detail signal (4).

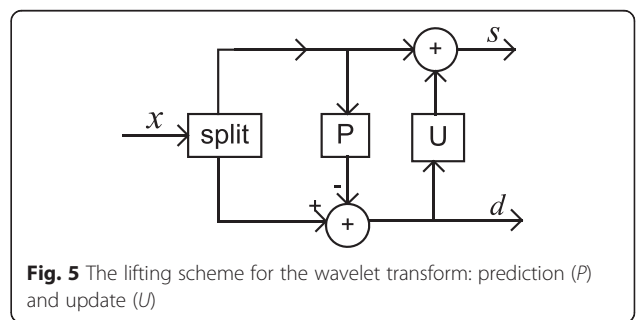
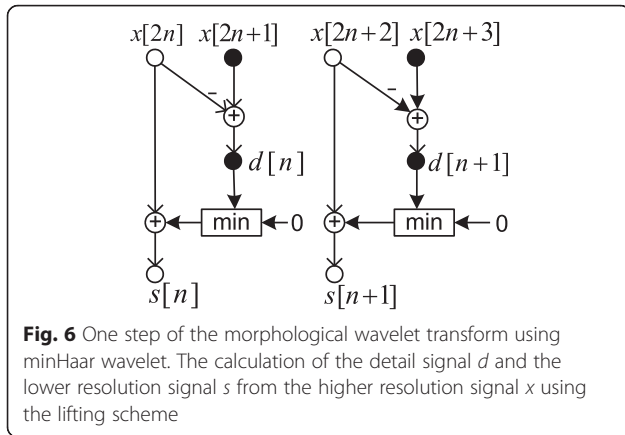


Fig. 5 The lifting scheme for the wavelet transform: prediction (P) and update (U)



$$d[n] = x[2n + 1] - x[2n] \tag{3}$$

$$s[n] = x[2n] + \min(0, d[n]) \tag{4}$$

The morphological Haar wavelet decomposition scheme may do a better job in preserving edges as compared to linear case [10]. The morphological Haar wavelet has some specific invariance properties. Besides of being translation invariant in the spatial domain, it is also gray-shift invariant and gray-multiplication invariant [37].

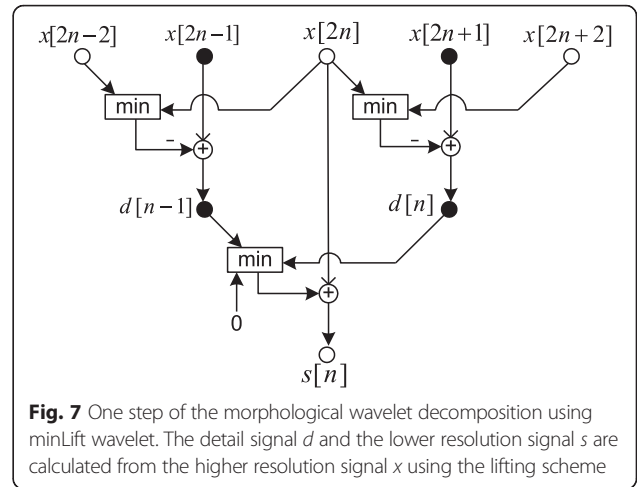
- *1D Morphological wavelet transformation using min-lifting scheme (minLift)*

Min-lifting scheme [10] is constructed using two non-linear lifting steps: non-linear prediction and non-linear update, both using operator erosion (by taking the minimum over two/three samples). After splitting the signal x to an odd samples array (black nodes in the first row of Fig. 7) and an even samples array (white nodes in the first row of Fig. 7), each sample of the detail signal d (second row on Fig. 7) is calculated according to (5). The update step is chosen in such a way that local minimum of the input signal is mapped to scaled signal and a sample of the approximation signal s (third row on Fig. 7) is calculated according to (6).

$$d[n] = x[2n + 1] - \min(x[2n], x[2n + 2]) \tag{5}$$

$$s[n] = x[2n] + \min(0, d[n-1], d[n]) \tag{6}$$

Morphological wavelet decomposition using minLift wavelet is both gray-shift invariant and gray-multiplication invariant [37]. Min-lifting scheme has the nice property that it preserves local minima of a signal, respectively, over several scales. It does not generate any new local minima. The detail signal is almost zero at areas of smooth gray level variation and sharp gray level variations are mapped to positive detail signal values (white). As an illustration of the wavelet decomposition using morphological minLift



wavelet, the oriented wavelet subbands from the first decomposition level which contain vertical, horizontal, and corner details are shown on Fig. 8 for the synthesized frame from the video sequence Newspaper.

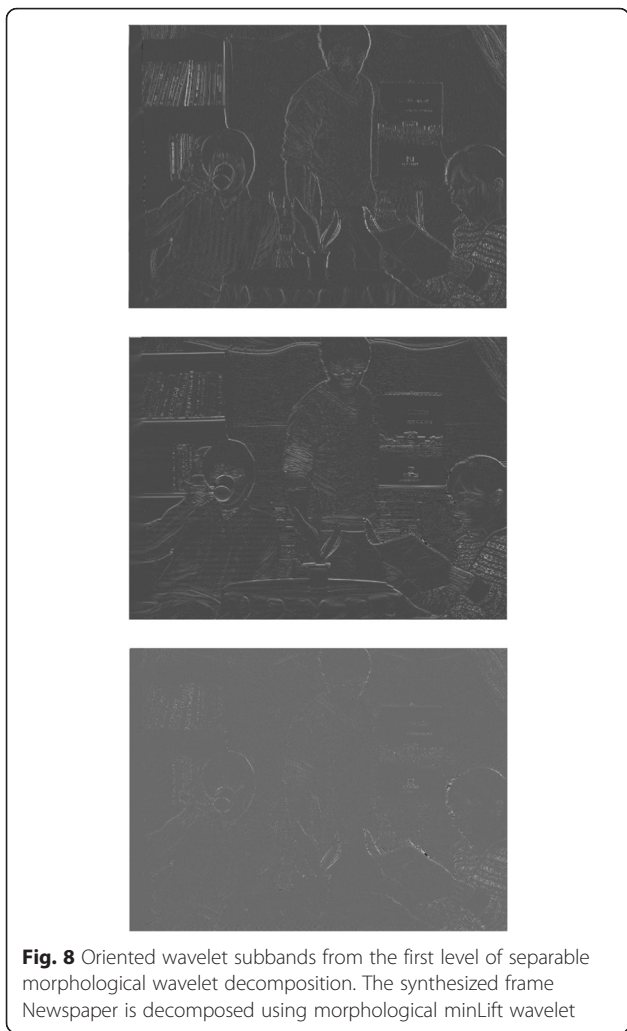
- *Non-separable morphological wavelet transformation with quincunx sampling using min-lifting scheme (minLiftQ)*

Two-dimensional non-separable morphological wavelet decomposition on a quincunx lattice using the min-lifting scheme minLiftQ [40] is analog to separable morphological wavelet decomposition using minLift wavelet. Non-separable 2D wavelet transform on a quincunx lattice using the lifting scheme is performed through odd and even steps alternately, producing a detail subband at each step and an approximation image which is decomposed further. Each step, odd and even, is implemented using the lifting scheme which consists of three parts: splitting, prediction and update. In the odd step, the image pixels are splitted in two subsets, both on quincunx lattice, Fig. 9 upper row, one subset with white pixels, x and the other subset with black pixels, y . The pixel of the error signal d is calculated using the minimum of the four nearest pixels in the horizontal and vertical directions (7), Fig. 9 bottom row left, and the lower resolution signal s is updated from the four nearest detail signal pixels (8), Fig. 9 bottom row right.

$$d = y - \min(x, x_1, x_2, x_3) \tag{7}$$

$$s = x + \min(d, d_1, d_2, d_3, 0) \tag{8}$$

In the even step, the signal on the quincunx lattice is separated on two subsets, both on Cartesian lattice, one subset with white pixels x and the other subset with gray pixels y , Fig. 10 upper row. The pixel of the error signal d is calculated from the four nearest pixels on diagonal

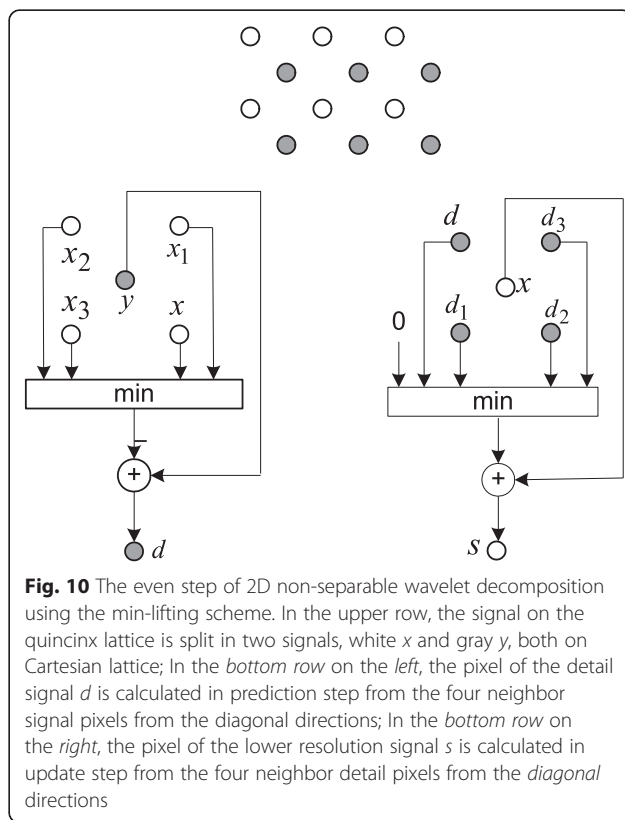
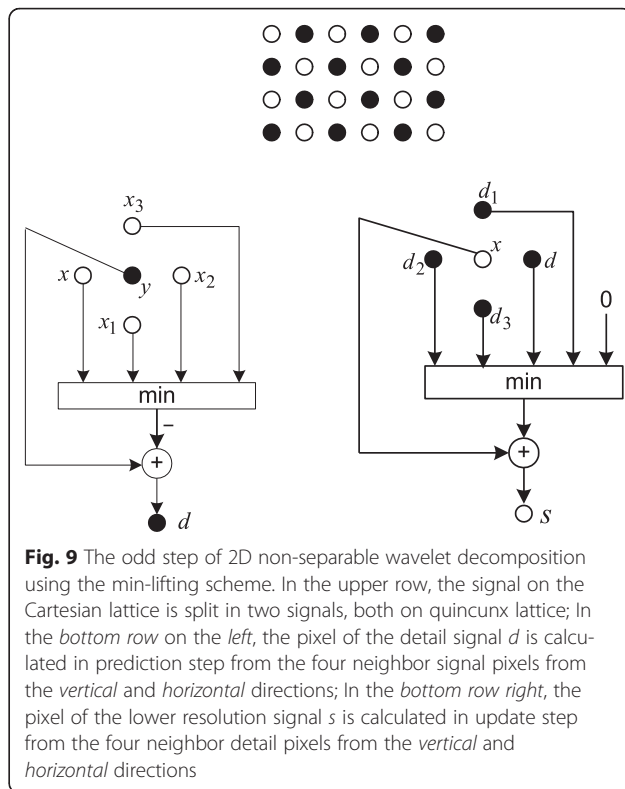


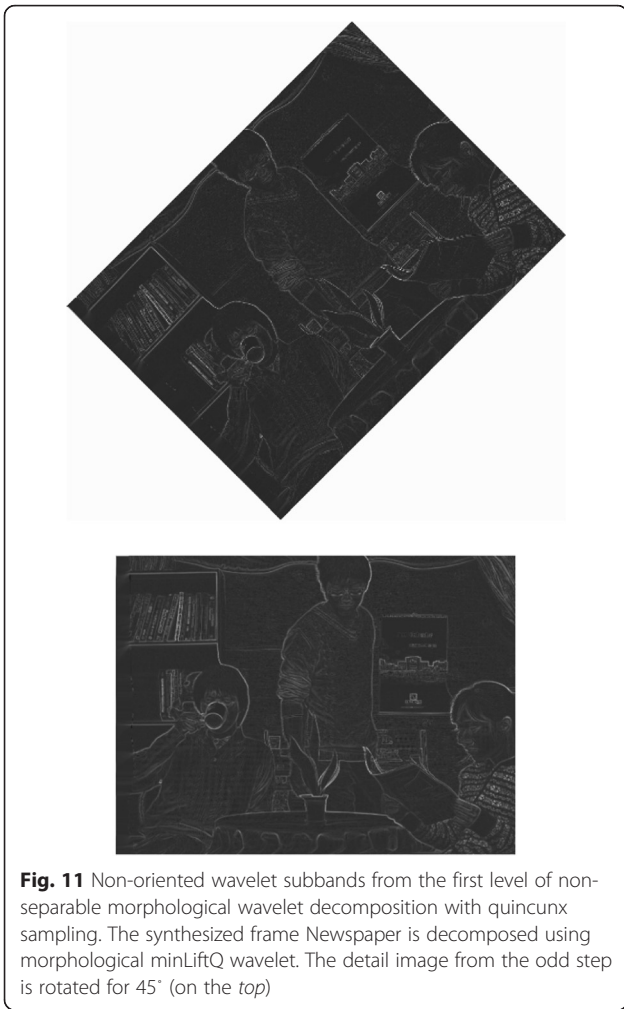
directions (7), Fig. 10 bottom row left, and the lower resolution signal s is updated from four nearest detail signal pixels on diagonal directions (8), Fig. 10 bottom row right.

Owing to the symmetry in the quincunx grid, the non-separable transform is insensitive to edge directions and image orientation. Non-oriented wavelet subbands from the first level of non-separable wavelet decomposition with quincunx sampling using morphological minLiftQ wavelet of the synthesized frame from the video sequence Newspaper are shown on Fig. 11. The detail image from the odd step is rotated 45° before display. The detail images are almost zero at areas of smooth gray level variation. Sharp gray level variations are mapped to positive (white) detail image values.

3.2 Distortion computation and pooling stage

Mean squared error (MSE) and peak signal-to-noise ratio (PSNR) are the most widely used objective image distortion/quality metrics. They are probably the simplest





way to quantify the similarity between two images. The mean squared error remains the standard criterion for the assessment of signal quality and fidelity. It has many attractive features: simplicity, parameter free, memoryless [42]. The MSE is an excellent metric in the context of optimization. Moreover, competing algorithms have most often been compared using MSE/PSNR [42]. It is shown that MSE has poor performances in some cases (contrast

stretch, mean luminance shift, contamination by additive white Gaussian noise, impulsive noise distortion, JPEG compression, blur, spatial scaling, spatial shift, rotation) when it is used as a single-scale metric on the full resolution images in the base band [42, 43].

In this paper, we propose to use MSE for distortion measurement between pyramid images in MP-PSNR and between wavelet subbands in MW-PSNR. In the second stage of the multi-scale IQA framework we use squared error maps between the morphological multi-scale representations of the two images: the reference image and the DIBR-synthesized image. Squared error maps calculated pixel-by-pixel show wrong displacement of the object edges induced by DIBR process through different scales of multi-scale representations. From the squared error maps, mean squared errors are calculated and combined into the multi-scale mean squared error which is transformed into multi-scale peak signal-to-noise ratio in the third stage of the multi-scale IQA framework.

3.2.1 The calculation of MP-PSNR

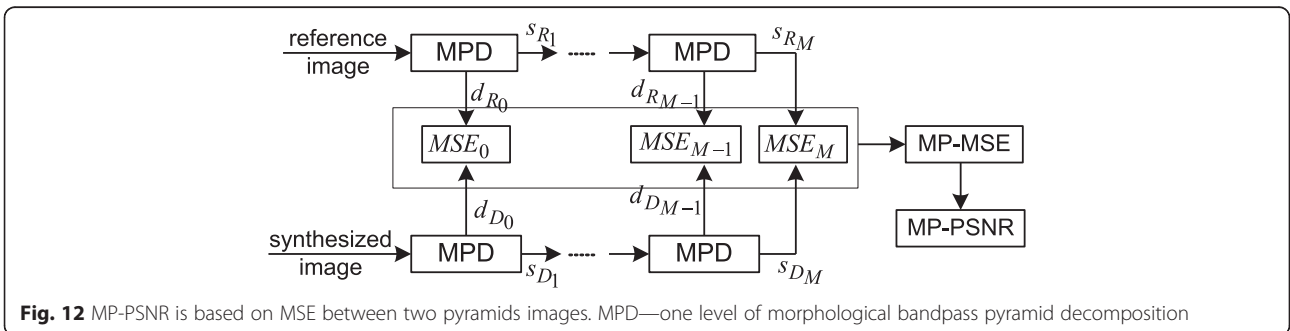
When the morphological pyramid decomposition is used in the first stage of morphological multi-scale IQA framework, Fig. 12, multi-scale pyramid mean squared error MP_MSE is calculated as weighted product of MSE_j values at all pyramid levels (9).

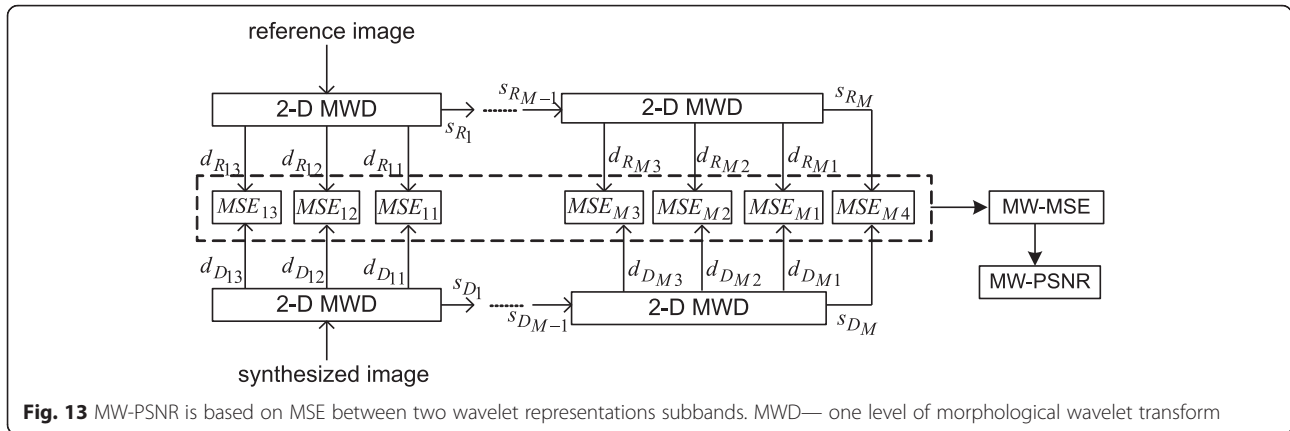
$$MP_MSE = \prod_{j=0}^M [MSE_j]^{\beta_j} \tag{9}$$

where equal value weights $\beta_j = \frac{1}{M+1}$ are used, M is the number of decomposition levels and $M + 1$ is the number of pyramid images. Finally, MP_MSE is transformed into Morphological Pyramid Peak Signal-to-Noise Ratio MP_PSNR (10).

$$MP_PSNR = 10 \cdot \log_{10} \left(\frac{R^2}{MP_MSE} \right) \tag{10}$$

where R is the maximum dynamic range of the image.





3.2.2 The calculation of MW-PSNR

When the morphological wavelet decomposition is used in the first stage of morphological multi-scale IQA framework, multi-scale wavelet mean squared error (MW-MSE) is calculated as weighted sum of MSE_{ji} values for all subbands at all scales of the two wavelet representations as final pooling (11).

$$MW_MSE = MSE_{M,D+1} \cdot \beta_{M,D+1} + \sum_{j=1}^M \sum_{i=1}^D MSE_{ji} \cdot \beta_{ji} \quad (11)$$

where equal value weights $\beta_{ji} = \frac{1}{M \cdot D + 1}$ are used. M is the number of decomposition levels, D is the number of detail subbands at one decomposition level. In the case of separable wavelet transforms, $D = 3$, Fig. 13, while for the non-separable wavelet decomposition, $D = 2$, MSE_{ji} is the mean value of the squared error map of the subband i at decomposition level j .

Finally, multi-scale metric Morphological Wavelet Peak Signal-to-Noise Ratio, MW-PSNR, is calculated as:

$$MW_PSNR = 10 \cdot \log_{10} \left(\frac{R^2}{MW_MSE} \right) \quad (12)$$

4 Results

In this section, experimental setup for the validation of proposed morphological multi-scale measures is described. The performances of two versions of the proposed morphological multi-scale metric, the Morphological Pyramid Peak Signal-to-Noise Ratio measure, MP-PSNR, and the Morphological Wavelet Peak Signal-to-Noise Ratio measure, MW-PSNR, are presented and discussed. Moreover, the PSNR performances by multi-scale decomposition subbands are analyzed. It is shown experimentally that PSNR has very good agreement with human judgment when it is calculated for the images at higher morphological decomposition scales. Therefore, we propose the reduced versions

of the morphological multi-scale measures, reduced MP-PSNR, and reduced MW-PSNR, using only detail images from higher decomposition scales. The performances of the reduced morphological multi-scale measures are presented also.

Since the morphological operators used in morphological multi-resolution decomposition schemes involve only integers and only max, min, and addition in their computation the calculation of morphological multi-resolution decompositions have low computational complexity. The calculation of MSE is of low computational complexity also. Therefore, the calculation of both measures, MP-PSNR and MW-PSNR, is not computationally demanding.

4.1 Experimental setup

To compare the performances of the image quality measures the following evaluation metrics are used: root mean squared error between the subjective and objective scores (RMSE), Pearson's correlation coefficient with non-linear mapping between the subjective scores and objective measures (PCC) and Spearman's rank order correlation coefficient (SCC). The calculation of DMOS from given MOS and non-linear mapping between the subjective scores and objective measures are done according to test plan for evaluation of video quality models for use with high definition TV content by VQEG HDTV group [44].

The performances of the metrics MP-PSNR and MW-PSNR are evaluated using two publicly available databases which contain DIBR-synthesized images: the IRCCyN/IVC DIBR image database [3, 4] and part of the MCL-3D stereoscopic image database [5, 6].

4.1.1 The IRCCyN/IVC DIBR image quality database

The IRCCyN/IVC DIBR image quality database contains frames from three multi-view video sequences: Book arrival (1024×768 , 16 cameras with 6.5 cm spacing), Lovebird1 (1024×768 , 12 cameras with 3.5 cm spacing) and

Newspaper (1024×768 , 9 cameras with 5 cm spacing). The selected contents are representative and used by MPEG also. For each sequence four virtual views are generated on the positions corresponding to those positions obtained by the real cameras using seven depth-image-based rendering algorithms, named A1-A7 [45–50]. One key frame from each synthesized sequence is randomly chosen for the database. For these key frames subjective assessment in form of mean opinion scores (MOS) is provided. The difference mean opinion scores (DMOS) is calculated as the difference between the reference frame's MOS and the synthesized frame's MOS. In the algorithm A1 [45], the depth-image is pre-processed by a low-pass filter. Borders are cropped and then the image is interpolated to reach its original size. The algorithm A2 is based on A1 except that the borders are not cropped but inpainted by the method described in [46]. The algorithm A3 [47] use inpainting method [46] to fill in the missing parts in the virtual image which introduces blur in the disoccluded area. This algorithm was adopted as the reference software for MPEG standardization experiments in 3D Video group. The algorithm A4 performs hole-filling method aided by depth information [48]. The algorithm A5 uses a patch-based texture synthesis as the hole-filling method [49]. The algorithm A6 uses depth temporal information to improve synthesis in the disoccluded areas [50]. The frames generated by algorithm A7 contain unfilled holes. Due to very noticeable object shifting artifacts in the frames generated by algorithm A1, these frames are excluded from the tests. The focus remains on images synthesized using A2–A7 DIBR algorithms and without registration procedure for alignment of the synthesized and the original frames. The results presented in Sections 4.2–4.4 for the IRCCyN/IVC DIBR database are based on the mixed statistics of the DIBR algorithms A2-A7.

4.1.2 The MCL-3D stereoscopic image quality database

The part of the stereoscopic image quality database MCL-3D which contains 36 stereopairs generated using four DIBR algorithms and associated mean opinion score (MOS) values is used for testing. These stereoscopic image pairs are rendered from nine image-plus-depth sources: Baloons, Kendo and Lovebird1 of resolution 1024×728 and Shark, Microworld, Poznan street, Poznan Hall2, Gt_fly, Undo_dancer of resolution 1920×1088 .

For each source, three views are used for the calculation of the metric score, Fig. 14. Original textures (T1, T2, T3) and their associated depth maps (D1, D2, D3) are obtained by selecting key frames from each of nine multi-view test sequences associated with depth maps. From the middle view (T2, D2), using one of the four DIBR algorithms, the stereoscopic image pair (SL, SR) is generated. The textures from the outer views, (T1, T3)

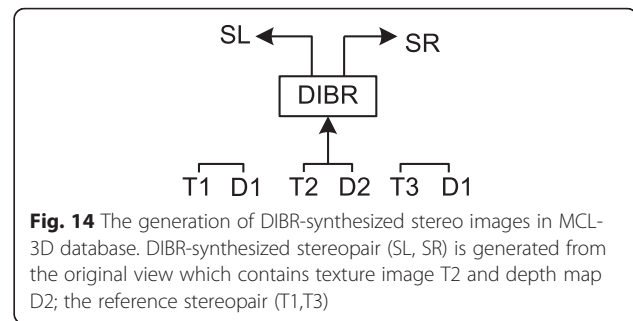


Fig. 14 The generation of DIBR-synthesized stereo images in MCL-3D database. DIBR-synthesized stereopair (SL, SR) is generated from the original view which contains texture image T2 and depth map D2; the reference stereopair (T1,T3)

are used as the reference stereo pair. We have calculated IQA metric score between the DIBR-synthesized stereopair (SL, SR) and the reference stereopair (T1,T3). The score for the stereo pair is calculated as the average of the left and right image scores.

In the generation of the MCL-3D database, four DIBR algorithms are used: DIBR with filtering, A1 [45], DIBR with inpainting, A2 [46], DIBR without hole-filling, A7 and DIBR with hierarchical hole-filling (HHF), A8 [51]. HHF uses pyramid-like approach to estimate the hole pixels from lower resolution estimates of the 3D wrapped image yielding to the virtual images that are free of any geometric distortions. Adding the depth adaptive preprocessing step before applying the hierarchical hole-filling, the edges and texture around the disoccluded areas can be sharpened and enhanced. The results presented in sections 4.2 – 4.4 for the MCL-3D database are based on the mixed statistics of four DIBR algorithms A1, A2, A7, and A8. The original image Shark and the left images from the stereopairs synthesized using four DIBR algorithms (A1, A2, A7, A8) are shown on Fig. 15 from top to bottom and from left to right.

4.2 Analysis of MP-PSNR performances

In this section, the performances of the Morphological Pyramid Peak Signal-to-Noise Ratio measure, MP-PSNR, are analyzed. Morphological bandpass pyramid decomposition using morphological operator erosion for low-pass filtering in analysis step and morphological operator dilation for interpolation filtering in synthesis step (MBP ED) is applied on the reference image and the DIBR-synthesized image. The influence of different size and shape of structuring element used in morphological operations and different number of decomposition levels in MBP ED pyramid decompositions on MP-PSNR performances are explored. For comparison with linear case, MP-PSNR performances are calculated using Laplacian pyramid decomposition with linear filters. In addition, PSNR performances calculated between two pyramids' images on different pyramid scales are investigated. The reduced version of MP-PSNR using only lower resolution images from higher pyramid scales is proposed and its performances are analyzed.

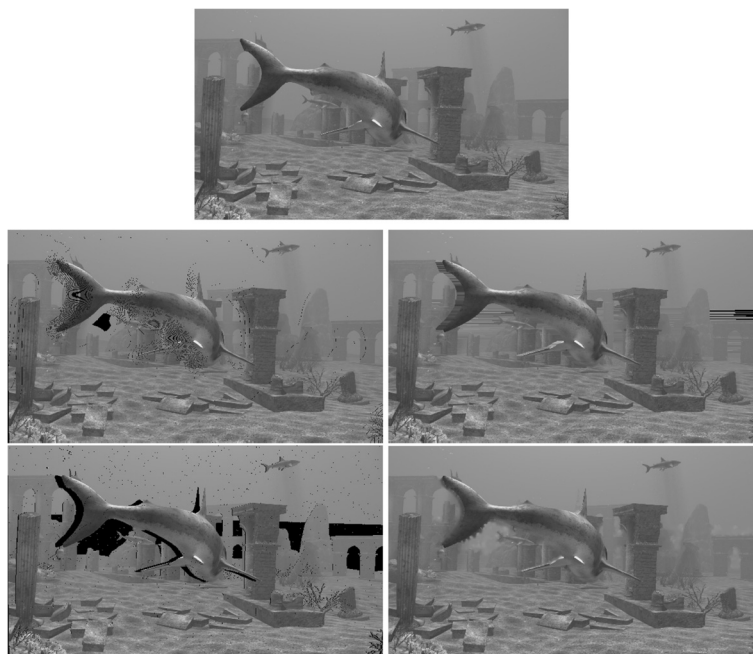


Fig. 15 The image Shark: original and DIBR-synthesized. Original image, *left image* of the stereoscopic pair synthesized using DIBR algorithms: A1, A2, A7, and A8, from *top to bottom, from left to right*

The shape and the size of the structuring element (SE) used in morphological filtering determine which geometrical features are preserved in the filtered image especially the direction of object’s enlargement or shrinking. Using square structuring element the objects are enlarged or shrunked equally in all directions. Squared-shaped structuring element is suitable to detect straight lines while round SE is suitable to detect circular features. The MP-PSNR performances using different shapes of structuring element (square, round, rhomb and cross type structuring element, Fig. 16) for morphological filtering in analysis step are evaluated. Better performances of MP-PSNR are achieved with square or round type SE than by rhomb or cross type SE. The results are similar with square and round type structuring element, but the computational complexity is significantly lower when the square structuring element is used. Namely, in that case separable pyramid decomposition by rows and columns with down-sampling after each step can be easily implemented.

In the images from the two chosen databases, straight lines are dominant and squared-shaped structuring element is chosen.

Moreover, the impact of structuring element size used in morphological operations and the number of decomposition levels in MBP ED pyramid decompositions on MP-PSNR performances is investigated. MP-PSNR performances are calculated using MBP ED pyramid decomposition with different number of decomposition levels (1–7 for IRCCyN/IVC DIBR database and 1–8 for MCL-3D database) and with square structuring elements of different sizes from 2×2 to 13×13 . More features are removed from the image at each decomposition level as larger structuring element is used. The number of decomposition levels for the best MP-PSNR performances depends on the size of structuring element.

The performances of MP-PSNR using SE of different sizes and the best number of decomposition levels for that size of SE are shown in the upper part of Table 1. For the IRCCyN/IVC DIBR database, the MP-PSNR performances show improvement with enlargement of the structuring element. The MP-PSNR performances are noticeable better for SE of size 5×5 and higher. Matlab implementation of MP-PSNR is available online [52].

In the case of MCL-3D database, the operation sum is used in the calculation of MP-MSE (9) as better performances of MP-PSNR are achieved. For the

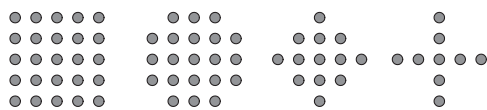


Fig. 16 Structuring elements of size 5×5 in different shapes. From left to right: square, round, rhomb, cross

Table 1 Performances of the full and the reduced versions of MP-PSNR

SE	IRCCyN/IVC DIBR				MCL-3D			
	Levels	RMSE	PCC	SCC	Levels	RMSE	PCC	SCC
full MP-PSNR								
2 × 2	6	0.4101	0.8019	0.7083	8	1.3364	0.8735	0.7446
3 × 3	5	0.3996	0.8131	0.7101	5	1.3506	0.8706	0.7228
5 × 5	5	0.3561	0.8549	0.7759	4	1.3014	0.8805	0.7373
7 × 7	5	0.3264	0.8796	0.8050	4	1.2600	0.8885	0.7566
9 × 9	5	0.3263	0.8798	0.8015	3	1.2713	0.8863	0.7601
11 × 11	4	0.3165	0.8874	0.8175	3	1.2560	0.8892	0.7691
13 × 13	4	0.3221	0.8830	0.8021	3	1.2277	0.8945	0.7700
Reduced MP-PSNR								
2 × 2	4–6	0.3660	0.8459	0.7775	4–9	1.3033	0.8801	0.7701
3 × 3	3–5	0.3252	0.8806	0.8185	4–6	1.3392	0.8730	0.7551
5 × 5	3–5	0.2936	0.9039	0.8634	4–5	1.2954	0.8817	0.7820
7 × 7	3–5	0.2931	0.9042	0.8573	2–5	1.2565	0.8891	0.7656
9 × 9	2–4	0.2997	0.8996	0.8614	2–4	1.2759	0.8855	0.7535
11 × 11	2–4	0.2922	0.9048	0.8684	2–4	1.2599	0.8885	0.7869
13 × 13	2–4	0.2920	0.9050	0.8684	2–4	1.2325	0.8936	0.7821

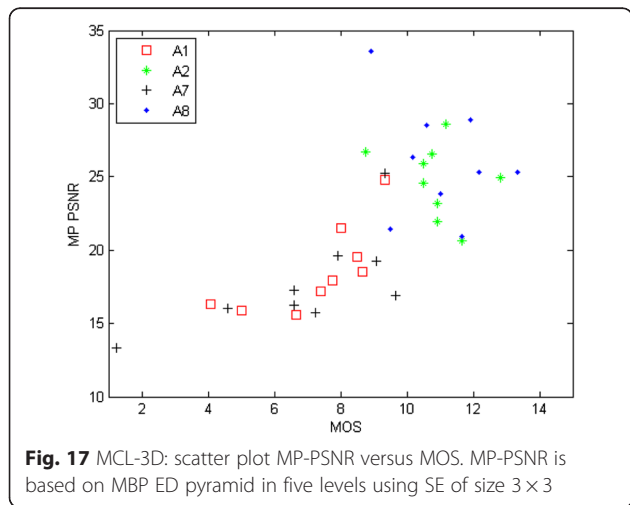
MCL-3D database, there is just a slight improvement of MP-PSNR performances with the enlargement of the structuring element. Scatter plot of MP-PSNR using SE of size 3 × 3 versus MOS for MCL-3D database is shown in Fig. 17. Each point represents one stereopair from the database.

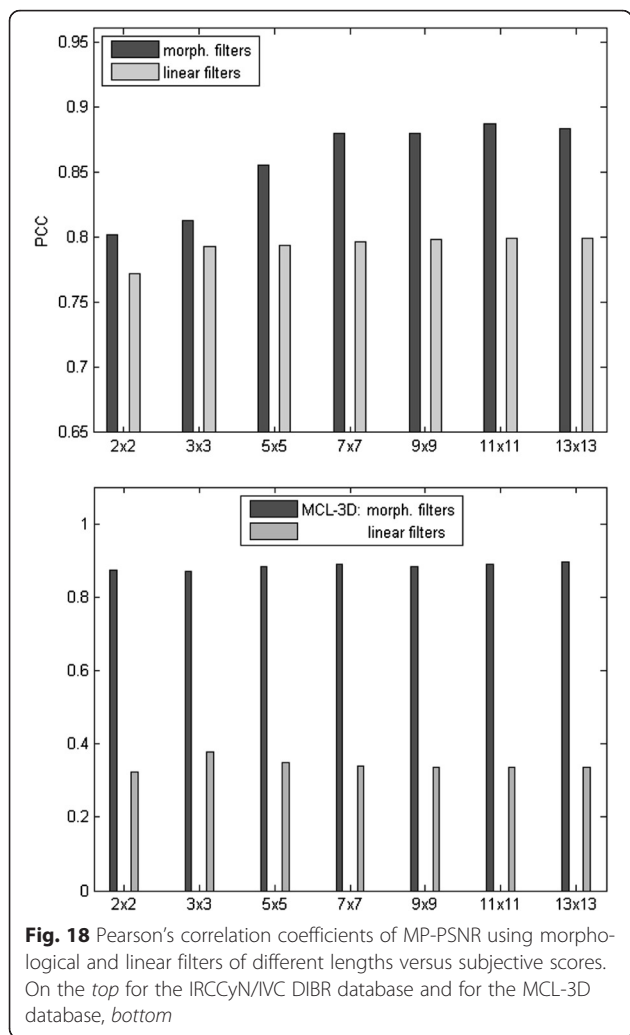
For the comparison with linear case, the image decomposition is performed using Laplacian pyramid with linear filters. Simple and efficient binomial filters [53] as approximation of a Gaussian filters are used. Binomial filters' coefficients are from Pascal's triangle, normalized with

their sum. Two-dimensional filter is implemented as cascade of one-dimensional filters. The MP-PSNR performances using pyramid decompositions with linear filters are similar for all filter lengths. For the IRCCyN/IVC DIBR database, Pearson's correlation varies from 0.771 for the linear filter of length 2 to 0.799 for the linear filter of length 13. For the MCL-3D database, Pearson's correlation varies from 0.322 for the linear filter of length 2 to 0.377 for the linear filter of length 3. Pearson's correlation coefficients of MP-PSNR versus DMOS for different filter lengths used in linear pyramid decomposition and for different sizes of SE used in morphological pyramid decomposition are shown on Fig. 18, left for the IRCCyN/IVC DIBR database and right for the MCL-3D database. The results on Fig. 18 are based on the mixed statistics of the DIBR algorithms A2–A7 for the IRCCyN/IVC DIBR database and A1, A2, A7, A8 for the MCL-3D database. MP-PSNR using pyramid decomposition with morphological filters has much better performances than MP-PSNR using pyramid decomposition with linear filters.

- Analysis of PSNR performances by pyramid images

It is shown in [54] that better performances of IQA metrics PSNR and SSIM are achieved when these metrics are calculated for the lower resolution images after low-pass filtering and downsampling than for the full resolution images. The downsampling scale depends on the image size and the viewing distance. We have investigated PSNR





performances for the detail images of the morphological bandpass pyramid at different pyramid scales. The reference image and the DIBR-synthesized image are decomposed into a set of lower resolution pyramid images using morphological bandpass erosion/dilation pyramid decomposition. At each pyramid scale, PSNR is calculated between the detail images of the two pyramids, the reference image pyramid and the DIBR-synthesized image pyramid.

For the IRCCyN/IVC DIBR database, Pearson's correlation coefficients of PSNR versus DMOS for pyramid images by pyramid scales using structuring elements of different sizes are shown on Fig. 19.

The smallest PCC is for the first pyramid scale (d_0) for all sizes of SE. Higher value PCC is for the middle and high scales. For the morphological pyramid decomposition using SE of size 2×2 and 3×3 , the highest PCC is at scale 5 (d_4). For the SE of size 5×5 , the best PSNR performances are obtained at pyramid scale 4 (d_3). For the pyramid decomposition with larger SE, the best

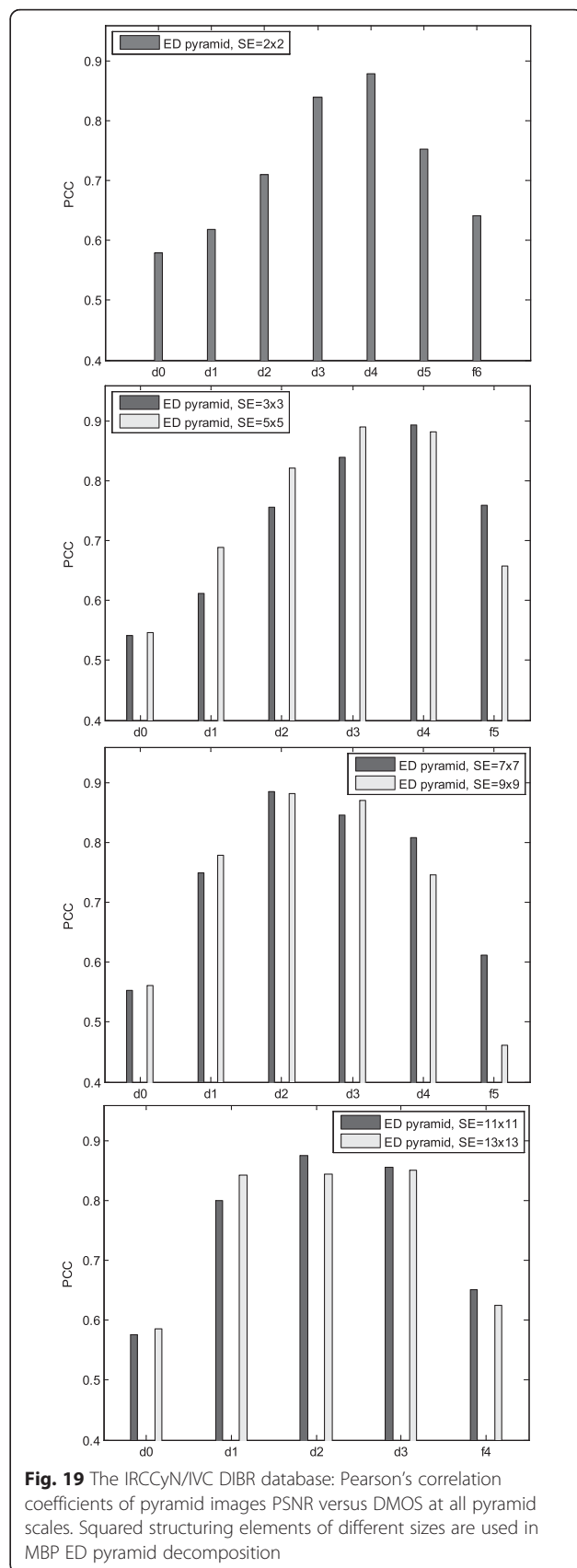


Table 2 The best performances of PSNR by pyramid scale for structuring element (SE) of different sizes

SE	IRCCyN/IVC DIBR				MCL-3D			
	Image	RMSE	PCC	SCC	Image	RMSE	PCC	SCC
2 × 2	d4	0.3270	0.8792	0.8147	d8	1.2364	0.8929	0.7877
3 × 3	d4	0.3076	0.8939	0.8671	s5	1.3454	0.8717	0.8124
5 × 5	d3	0.3130	0.8899	0.8656	s4	1.3100	0.8788	0.8145
7 × 7	d2	0.3180	0.8862	0.8485	s4	1.2750	0.8856	0.8209
9 × 9	d2	0.3239	0.8816	0.8697	s3	1.2948	0.8818	0.8100
11 × 11	d2	0.3307	0.8763	0.8513	s3	1.2804	0.8846	0.8165
13 × 13	d3	0.3597	0.8517	0.7859	s3	1.2643	0.8877	0.8245
-	f0	0.4525	0.7519	0.6766	f0	2.6090	0.3113	0.2630

PSNR performances are obtained at scale 3 for detail images d_2 . Also, PSNR performances at middle and higher pyramid scales are much better than the PSNR performances for the case when the PSNR is calculated between the original and the DIBR-synthesized images without decomposition, in the base band. The best PSNR performances by pyramid images for different sizes of SE used in morphological pyramid decomposition are shown in Table 2. For the morphological pyramid decomposition using SE of size 3×3 , the best PSNR performances are achieved for the detail image at pyramid level 5, Pearson

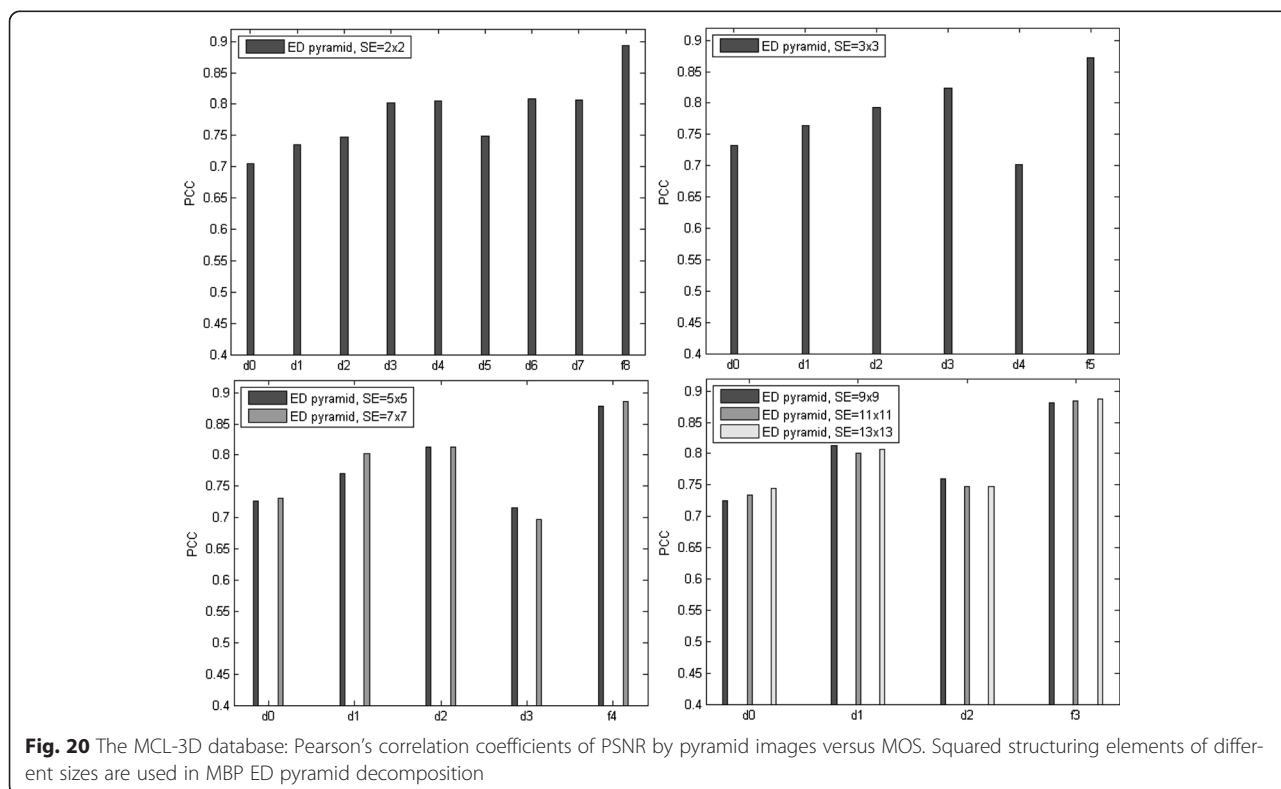
correlation coefficient 0.89 and Spearman correlation coefficient 0.867.

For the MCL-3D database, Pearson’s correlation coefficients of PSNR versus MOS for pyramid images at all pyramid scales using structuring elements of different sizes are shown on Fig. 20. For this database, smaller differences between PCC for pyramid images at different scales exist. The smallest PCC is at the first scale (detail images d_0) and the highest PCC is for the approximation images at the highest scale. The best pyramid image PSNR performances for different sizes of SE used in morphological pyramid decomposition are shown in Table 2.

For both databases, it is shown that PSNR shows very good agreement with human quality judgments when it is calculated at higher scales of MBP ED pyramid, much better than for the full resolution images in the base band. Matlab implementation of PSNR by morphological pyramid images is available online [52].

- The performances of the reduced version of MP-PSNR

Based on the results of PSNR performances calculated separately by pyramid scales, we propose reduced version of MP-PSNR using only pyramid images with higher PCC values of PSNR towards subjective scores. Reduced



version of MP_MSE is calculated as the weighted sum of the used subbands' MSE (9).

For the IRCCyN/IVCDIBR database, the reduced version of MP-PSNR is calculated using only three detail images with higher PCC values of PSNR towards DMOS. The performances of the reduced versions of MP-PSNR using equal value weights are presented in the bottom left part of Table 1. Reduced version of MP-PSNR has better performances than its full version: from 1.74 % when the MBP ED pyramid decomposition with SE of size 11×11 is used to 6.75 % when the MBP ED pyramid decomposition with SE of size 3×3 is used. HVS visually integrates an image edges in a coarse-to-fine-scale (global-to-local) fashion [34]. Visual cortex cells integrate activity across spatial frequency in an effort to enhance the representation of edges. Because the edges are visually integrated in a coarse-to-fine-scale order, the visual fidelity of an image can be maintained by preserving coarse scales at the expense of fine scales. Reduced version of MP-PSNR is computationally more efficient than its full version as the MSE is only calculated for lower resolution pyramid images. The reliable and fast evaluation is obtained with reduced version MP-PSNR using MBP ED pyramid with SE of size 5×5 (Pearson's 90.39 %, Spearman 86.3 %). Scatter plot of nonlinearly mapped reduced MP-PSNR versus subjective DMOS for that case is shown in Fig. 21. Each point represents one frame from the database. Matlab implementation of reduced version of MP-PSNR is available online [52].

For the MCL-3D database, the reduced version of MP-PSNR is calculated without detail images from the first three pyramid scales when the SE of size less than 7×7 is used. When the SE of size 7×7 and bigger is used, only the pyramid image from the first scale is omitted in the calculation of the reduced version of MP-PSNR. The performances of the reduced versions of MP-PSNR

using equal value weights are presented in the bottom right part of Table 1. Only marginal improvement is achieved using reduced version of MP-PSNR for MCL-3D database.

4.3 Analysis of MW-PSNR performances

In this section, the performances of the Morphological Wavelet Peak Signal-to-Noise Ratio measure, MW-PSNR, are analyzed. MW-PSNR uses morphological wavelet decomposition of the reference and the DIBR-synthesized images. Both separable morphological wavelet decompositions using morphological Haar min wavelet (minHaar) and min-lifting wavelet (minLift) and non-separable morphological wavelet decomposition with quincunx sampling using min-lifting wavelet (minLiftQ) are investigated. Separable morphological wavelet decompositions are computationally less expensive than non-separable wavelet decompositions. Also, they are less expensive than morphological pyramid decompositions for the same filter length. The influence of different number of wavelet decomposition levels on MW-PSNR performances are explored. For the comparison with linear wavelet decompositions, MW-PSNR performances are calculated using separable linear wavelet decompositions using Haar wavelet (Haar) and Cohen-Daubechies-Feauveau wavelet cdf(2,2) and non-separable linear wavelet decomposition with quincunx sampling using cdf(2,2)Q. PSNR performances calculated by wavelet subbands through decomposition scales are investigated. The reduced version of MW-PSNR using only wavelet subbands with better PSNR performances is analyzed.

The number of decomposition levels has been varied between 1 and 8 and the configurations with the best MW-PSNR performances have been chosen. The best MW-PSNR performances have been achieved using separable wavelet transformations in $M=7$ levels producing 22 subbands. Using non-separable wavelet transformation with quincunx sampling for the IRCCyN/IVC DIBR database, the best MW-PSNR performances have been achieved also with $M=7$ levels producing 15 subbands. For the MCL-3D database the best MW-PSNR performances using non-separable wavelet transformation have been achieved with $M=4$ levels producing nine subbands. Equal value weights are used in the calculation of MW-MSE (11). Matlab implementation of MW-PSNR is available online [55].

The performances of MW-PSNR for different wavelet transformations are presented in the upper part of Table 3. The performances of MW-PSNR using morphological wavelet transforms are better than the performances of MW-PSNR using linear wavelet transforms. The best MW-PSNR performances have been obtained

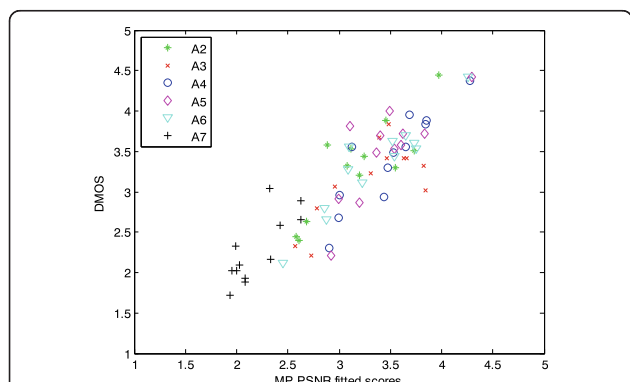


Fig. 21 Fitted scores of reduced MP-PSNR versus DMOS for the IRCCyN/IVC DIBR database. Reduced MP-PSNR is based on pyramid detail images from scales 3–5 of MBP ED pyramid using SE = 5×5

Table 3 Performances of the full and reduced versions of MW-PSNR

	IRCCyN/IVC DIBR			MCL-3D		
	RMSE	PCC	SCC	RMSE	PCC	SCC
Full MW-PSNR						
minHaar	0.3565	0.8545	0.7750	1.3529	0.8702	0.7076
Haar	0.4435	0.7632	0.6491	2.3611	0.5103	0.4760
minLift	0.4017	0.8108	0.6816	1.3882	0.8627	0.7029
cdf(2,2)	0.5009	0.6836	0.5450	2.3954	0.4887	0.4583
minLiftQ	0.3922	0.8206	0.7382	1.6299	0.8047	0.6463
cdf(2,2)Q	0.4756	0.7210	0.5779	2.5184	0.3982	0.3629
Reduced MW-PSNR						
minHaar	0.3188	0.8855	0.8298	1.3131	0.8782	0.7686
Haar	0.3935	0.8194	0.7695	1.9152	0.7165	0.6938
minLift	0.3878	0.8251	0.6990	1.3873	0.8629	0.7011
cdf(2,2)	0.4735	0.7239	0.5958	1.7729	0.7635	0.7352
minLiftQ	0.3599	0.8514	0.7641	1.6029	0.8119	0.6410
cdf(2,2)Q	0.4508	0.7541	0.6126	2.0357	0.6710	0.7040

using separable wavelet decomposition with morphological Haar wavelet which is of the lowest computational complexity: for the IRCCyN/IVC DIBR database, Pearson 0.85, Spearman 0.77 and for the MCL-3D database, Pearson 0.87, Spearman 0.70. Scatter plot of MW-PSNR using separable wavelet decomposition with morphological Haar wavelet versus MOS for MCL-3D database is shown on Fig. 22.

- Analysis of PSNR performances by wavelet subbands

We have investigated PSNR performances by wavelet subbands at different wavelet decomposition scales. The reference image and the DIBR-synthesized image are decomposed into a sets of lower resolution

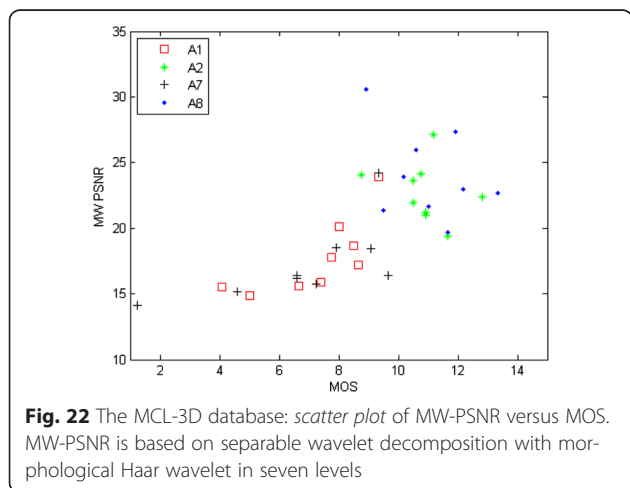


Fig. 22 The MCL-3D database: scatter plot of MW-PSNR versus MOS. MW-PSNR is based on separable wavelet decomposition with morphological Haar wavelet in seven levels

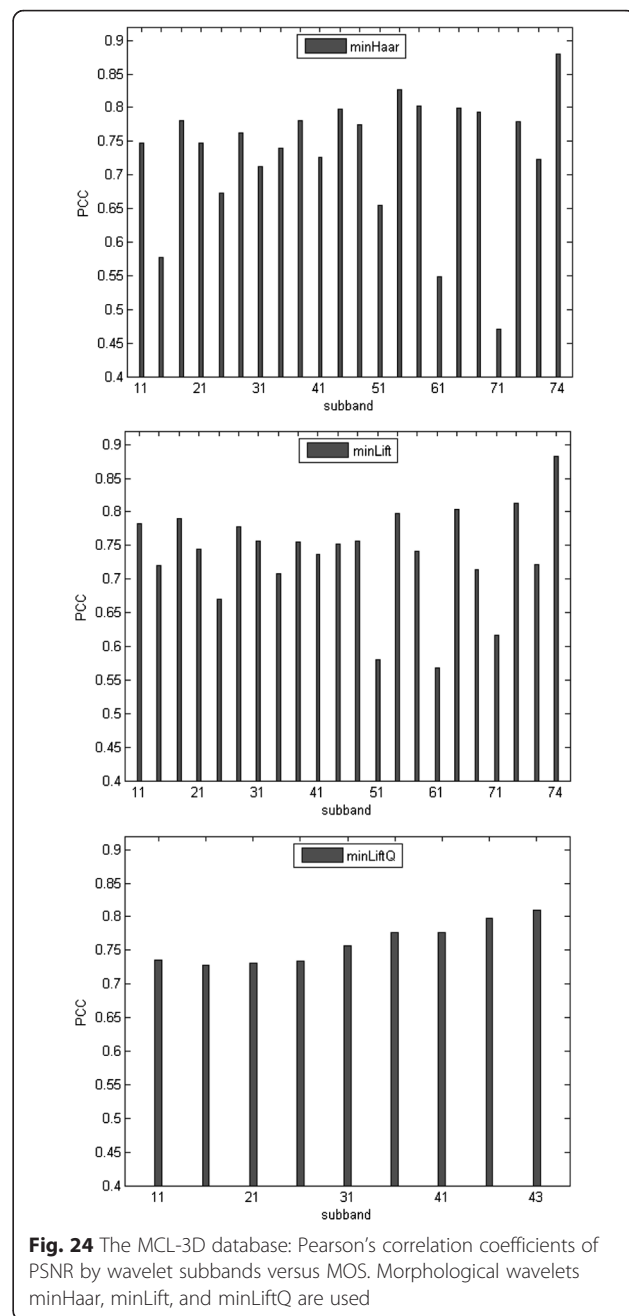
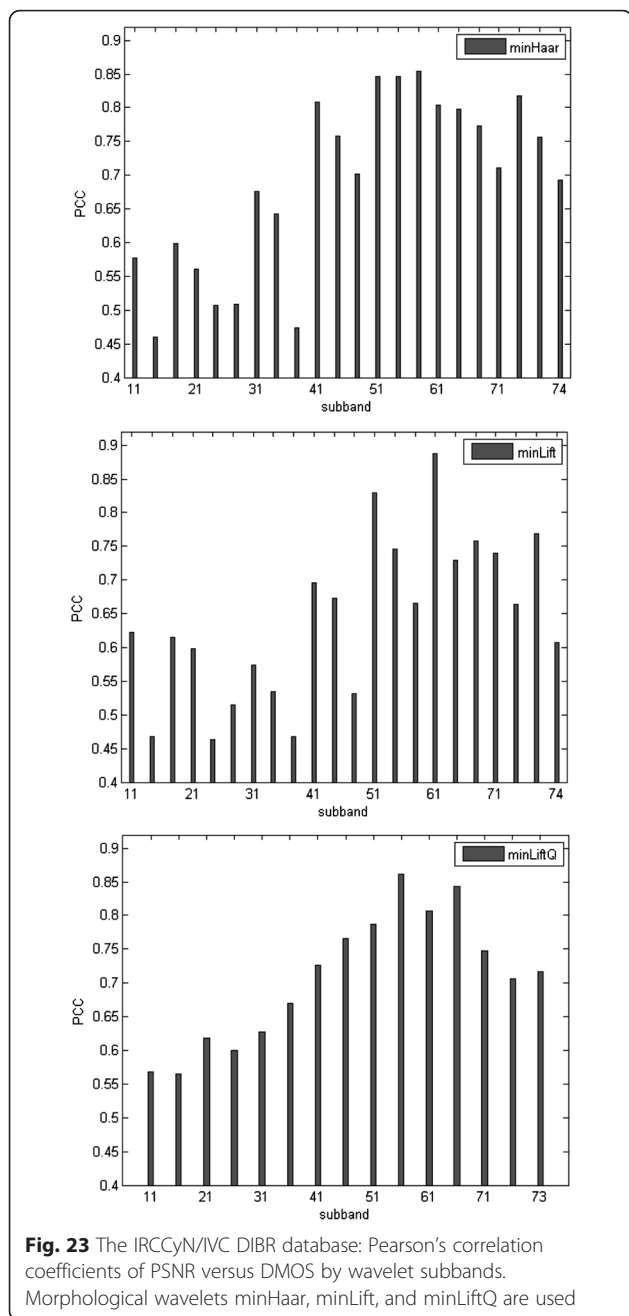
subbands using morphological wavelet decomposition. At each decomposition scale, for each wavelet subband, PSNR is calculated between the subbands of the two wavelet representations, the reference image wavelet representation and the DIBR-synthesized image wavelet representation. Pearson’s correlation coefficient (PCC) of PSNR to subjective scores is calculated for each subband for three types of morphological wavelets: minHaar, minLift and minLiftQ. Matlab implementation of PSNR by morphological wavelet subbands is available online [55].

For the IRCCyN/IVC DIBR database, Fig. 23, Pearson’s correlation coefficients calculated for wavelet subbands on decomposition levels 4–7 are higher than Pearson’s correlation coefficients calculated for wavelet subbands on decomposition levels 1–3. For the MCL-3D database, smaller differences by wavelet subbands between Pearson’s correlation coefficients can be noticed, Fig. 24.

Moreover, the best PSNR performances by wavelet subbands for each wavelet decomposition are shown in Table 4. For instance, for the IRCCyN/IVC DIBR database for the separable wavelet decomposition using morphological minLift wavelet, the best PSNR performances are obtained for subband on the scale 6 with vertical details (d61), PCC 0.887 and SCC 0.828. Also, for all tested wavelets, the PSNR of the wavelet subband with the highest PCC show much better performances than PSNR calculated between the reference image and the DIBR-synthesized image without decomposition in the base band.

- Analysis of the reduced version MW-PSNR performances

Based on the PSNR performances by subbands for the IRCCyN/IVC DIBR database given in Fig. 23, it can be concluded that the PSNR performances of wavelet subbands at decomposition levels 4–7 are much better than the subband PSNR performances on levels 1–3. Therefore, we propose reduced version of MW-PSNR using only these higher level subbands. Reduced versions of MW_MSE is calculated as weighted sum of the used subbands’ MSE. For the separable wavelet decomposition, the reduced version of MW-PSNR is calculated using only 11 subbands from levels 4–7 with indices 41–72. For the non-separable wavelet decomposition with quincunx sampling, reduced version of MW-PSNR is calculated using 6 subbands from decomposition levels 4–7 with indices 42–71. Matlab implementation of the reduced version of MW-PSNR is available online [55]. The performances of the reduced MW-PSNR are presented in the bottom left part of Table 3. It is shown



that for each wavelet type, the performances of the reduced version MW-PSNR are better than the performances of the full version MW-PSNR: 3.1 % for minHaar, 1.43 % for minLift and 3.08 % for minLiftQ. The best reduced version MW-PSNR performances are obtained using separable wavelet decomposition with morphological minHaar wavelet, Pearson's 88.5 %, Spearman 82.98 %. Scatter plot of nonlinearly mapped reduced MW-PSNR versus subjective DMOS for that case is shown in Fig. 25.

For the MCL-3D database, only marginal improvement is achieved using reduced version of MW-PSNR, Table 3 bottom right.

4.4 Summary of the results

The performances of the selected proposed metrics, the commonly used 2D image quality assessment metrics and the metric dedicated to synthesis-related artifacts, 3DswIM [13], are presented in Table 5. The considered commonly used 2D metrics are: PSNR, universal quality index UQI [56], structural similarity index SSIM [57],

Table 4 The best performances of PSNR by wavelet subbands for each wavelet

Decomp.	Wavelet	Subb.	IRCCyN/IVC DIBR			MCL-3D			
			RMSE	PCC	SCC	Subb.	RMSE	PCC	SCC
Separable	minHaar	d53	0.3576	0.8535	0.7831	s7	1.3084	0.8791	0.8239
	Haar	d61	0.3691	0.8431	0.7939	d13	1.6986	0.7856	0.7361
	minLift	d61	0.3167	0.8872	0.8281	s7	1.2877	0.8832	0.8107
	cdf(2,2)	d61	0.3558	0.8551	0.7671	d13	1.6301	0.8047	0.7593
Non-separable	minLiftQ	d52	0.3478	0.8621	0.7777	s4	1.6119	0.8095	0.7225
	cdf(2,2)Q	d52	0.4279	0.7818	0.6493	d1	1.7613	0.7671	0.7323
-	-	Base band	0.4525	0.7519	0.6766	Base band	2.6090	0.3113	0.2630

multi-scale structural similarity MS-SSIM [19], information weighted IW-PSNR [20], and IW-SSIM [20]. Single-scale structural similarity SSIM [57] is calculated between the original and the synthesized images using the given matlab code [58]. 3DswIM [13] is calculated using the given matlab p-code [59]. Selected versions of the proposed metrics using morphological pyramid decompositions presented in Table 5 are: PSNR calculated on scale 5 of the MBP ED pyramid representations using SE of size 3×3 ; reduced version of MP-PSNR using SE of size 5×5 in pyramid MBP ED decomposition; full versions of MP-PSNR using SE of size 5×5 . The selected proposed metrics using morphological wavelet decompositions shown in Table 5 are: PSNR calculated on scale 6 between wavelet subbands with vertical details of the two wavelet representations using minLift wavelet for the IRCCyN/IVC DIBR database and PSNR calculated on scale 7 between approximation wavelet subbands using minLift wavelet for the MCL-3D database; reduced and full versions of MW-PSNR using minHaar wavelet.

The performances of the proposed metrics are much better than the performances of the commonly used 2D metrics and better than the performances of the metric dedicated to synthesis-related artifacts, 3DswIM. The Pearson’s correlation coefficients of the selected commonly used 2D metrics, the metric dedicated to synthesis-related artifacts, 3DswIM, and the reduced versions of MP-PSNR and of MW-PSNR are shown on Fig. 26.

5 Conclusions

Most of the depth-image-based rendering (DIBR) techniques produce images which contain non-uniform geometric distortions affecting the edge coherency. This type of distortions are challenging for common image quality assessment (IQA) metrics. We propose full-reference metric based on multi-scale decomposition using morphological filters in order to better deal with specific geometric distortions in the DIBR-synthesized images. Introduced non-linear morphological filters in multi-resolution image decomposition maintain important geometric information such as edges across different resolution scales. The proposed metric is dedicated to artifact detection in DIBR-synthesized images by measuring the edge distortion between the multi-scale representations of the reference image and the DIBR-synthesized image using MSE. We have explored two versions of morphological multi-scale metric, Morphological Pyramid Peak Signal-to-Noise Ratio measure, MP-PSNR, based on morphological pyramid decomposition and Morphological Wavelet Peak Signal-to-Noise Ratio measure, MW-PSNR, based on morphological wavelet decomposition. The proposed metrics are evaluated using two databases which contain images synthesized by DIBR algorithms: IRCCyN/IVC DIBR image database and MCL-3D stereoscopic image database. Both metric versions demonstrate high improvement of performances over standard IQA metrics and over tested metric dedicated to synthesis-related artifacts. Also, they have much better performances than their linear counterparts for the evaluation of DIBR-synthesized

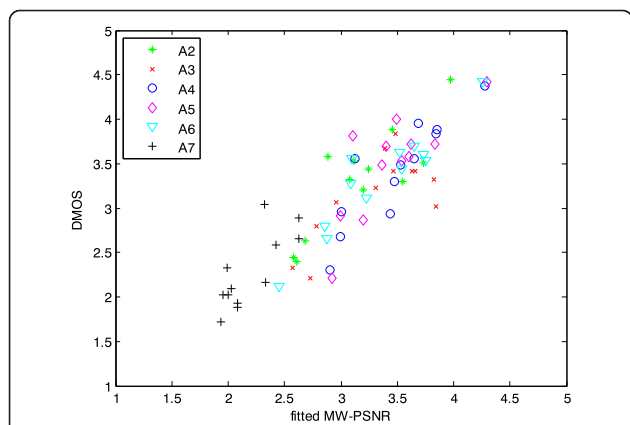
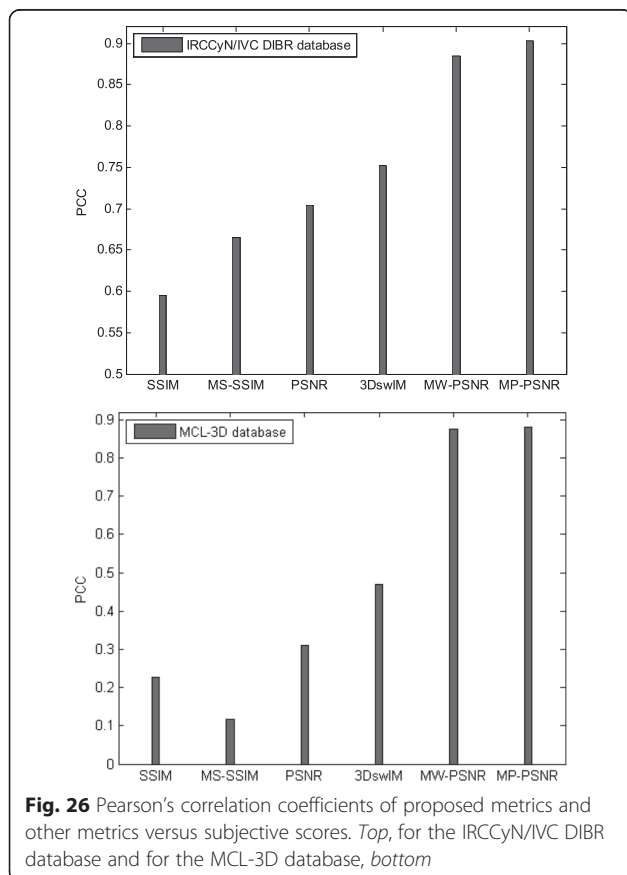


Fig. 25 The IRCCyN/IVC DIBR database: Fitted scores of reduced MW-PSNR versus MOS. Reduced version of MW-PSNR is based on wavelet subbands from decomposition levels 4–7; morphological wavelet decomposition using minHaar wavelet in seven levels is used

Table 5 Performances of the selected proposed metrics and other metrics

Metric	IRCCyN/IVC DIBR			MCL-3D		
	RMSE	PCC	SCC	RMSE	PCC	SCC
Commonly used metric						
PSNR	0.4525	0.7519	0.6766	2.6090	0.3113	0.2630
IW-PSNR [20]	0.5267	0.6411	0.5320	2.5961	0.3253	0.1726
UQI [56]	0.5199	0.6529	0.5708	2.2491	0.5735	0.3121
SSIM [57]	0.5513	0.5956	0.4424	2.6730	0.2283	0.0821
MS-SSIM [19]	0.5127	0.6649	0.5188	2.7267	0.1168	0.0301
IW-SSIM [20]	0.5350	0.6265	0.4856	2.4872	0.4235	0.0546
Dedicated to DIBR synth. images						
3DswIM [13]	0.4868	0.7049	0.6396	2.4232	0.4701	0.2559
Proposed metric						
PSNR, SE = 3 × 3	0.3076	0.8939	0.8671	1.3454	0.8717	0.8124
Reduced MP-PSNR SE = 5 × 5	0.2936	0.9039	0.8634	1.2954	0.8817	0.7820
Full MP-PSNR SE = 5 × 5	0.3561	0.8549	0.7759	1.3014	0.8805	0.7373
PSNR, minLift wavelet	0.3167	0.8872	0.8281	1.2877	0.8832	0.8107
Reduced MW-PSNR, minHaar	0.3188	0.8855	0.8298	1.3131	0.8782	0.7686
Full MW-PSNR, minHaar	0.3565	0.8545	0.7750	1.3529	0.8702	0.7076



images. MP-PSNR has slightly better performances than MW-PSNR. For the MCL-3D database, MP-PSNR achieves Pearson 0.888 and Spearman 0.756 using MBP ED pyramid decomposition with square structuring element of size 7×7 in 4 levels. For the same database, MW-PSNR achieves Pearson 0.87 and Spearman 0.707 using separable wavelet decomposition with morphological Haar wavelet in 7 levels.

It is shown that PSNR has particularly good agreement with human judgment when it is calculated between the appropriate detail images at higher decomposition scales of the two morphological multi-scale image representations. For IRCCyN/IVC DIBR images database, PSNR calculated on scale 5 of the MBP ED pyramid image representations using structuring element of size 3×3 has very good performances, Pearson's 0.89 and Spearman 0.86. For MCL-3D database, PSNR calculated on scale 4 of the MBP ED pyramid image representations using square structuring element of size 7×7 achieves Pearson's 0.88, Spearman 0.82. For IRCCyN/IVC DIBR images database, it has been shown that reduced versions of multi-scale metrics, reduced MP-PSNR and reduced MW-PSNR, can be used for the assessment of DIBR-synthesized frames with high reliability. Reduced version of MP-PSNR using morphological pyramid decomposition MBP ED with square structuring element of size 5×5 achieves the improvement 15.2 % of correlation over PSNR (Pearson's 0.904, Spearman 0.863) and reduced version of MW-PSNR using morphological wavelet decomposition with minHaar wavelet gains the improvement of

13.3 % of correlation over PSNR (Pearson's 0.885, Spearman 0.829).

Since the morphological operators involve only integers and only min, max, and addition in their computation, as well as simple calculation of MSE, the multi-scale metrics MP-PSNR and MW-PSNR are computationally efficient procedures. They provide reliable DIBR-synthesized image quality assessment even without any parameter optimization and precise registration procedure.

Abbreviations

cdf(2,2), linear, biorthogonal (2,2) of Cohen-Daubechies-Feauveau wavelet; cdf(2,2)Q, non-separable linear (2,2) of Cohen-Daubechies-Feauveau wavelet with quincunx sampling; DIBR, Depth-Image-Based Rendering; DMOS, Difference Mean Opinion Score; FVV, Free Viewpoint Video; Haar, Haar wavelet transformation; IQA, Image Quality Assessment; MBP ED, morphological bandpass pyramid erosion/ dilation; minHaar, morphological Haar min wavelet transformation; minLift, morphological min-lifting wavelet transformation; minLiftQ, non-separable morphological min-lifting wavelet transformation with quincunx sampling; MOS, Mean Opinion Score; MP-PSNR, Morphological Pyramid Peak Signal-to-Noise Ratio metric; MSE, Mean Squared Error; MW-PSNR, Morphological Wavelet Peak Signal-to-Noise ratio metric; PCC, Pearson's Correlation Coefficient; PSNR, Peak Signal-to-Noise ratio; SE, structuring element for morphological operations

Acknowledgements

This work was partially supported by COST Action IC1105-3D ConTourNet, the Ministry of Education, Science and Technological Development of the Republic of Serbia under Grant TR-32034 and by the Secretary of Science and Technology Development of the Province of Vojvodina under Grant 114-451-813/2015-03.

Authors' contributions

DSS proposed the framework of this work, carried out the whole experiments, and drafted the manuscript. DK supervised the whole work, offered useful suggestions, and helped to modify the manuscript. PLC participated in the discussion of this work and helped to polish the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Institute for Telecommunications and Electronics IRITEL, Belgrade, Serbia.

²Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia.

³Ecole polytechnique de l'Université de Nantes, IRCCyN Lab, Nantes, France.

Received: 20 August 2015 Accepted: 30 June 2016

Published online: 26 July 2016

References

1. K Mueller, P Merkle, T Wiegand, 3D video representation using depth maps. *Proc. IEEE* **99**(4), 643–656 (2011)
2. E Bosc, P Le Callet, L Morin, M Pressigout, Visual quality assessment of synthesized views in the context of 3DTV, in *3D-TV system with depth-image-based rendering*, ed. by C Zhu, Y Zhao, L Yu, M Tanimoto (Springer, New York, 2013), pp. 439–473
3. E Bosc, R Pepion, P Le Callet, M Koppel, P Ndjiki-Nya, M Pressigout, L Morin, Towards a new quality metric for 3-d synthesized view assessment. *IEEE Journal on Selected Topics in Signal Processing* **5**(7), 1332–1343 (2011)
4. IRCCyN/IVC DIBR image quality database. ftp://ftp.ivc.polytech.univ-nantes.fr/IRCCyN_IVC_DIBR_Images
5. R Song, H Ko, CCJ Kuo, *MCL-3D: a database for stereoscopic image quality assessment using 2D-image-plus-depth source*, 2014. <http://arxiv.org/abs/1405.1403>
6. MCL-3D stereoscopic image quality database. <http://mcl.usc.edu/mcl-3d-database>
7. E. Adelson, E. Simoncelli, W. Freeman, Pyramids and multiscale representations. *Proc. European Conf. on Visual Perception*, Paris (1990)
8. P Maragos, R Schafer, Morphological systems for multidimensional signal processing. *Proc. IEEE* **78**(4), 690–710 (1990)
9. A Toet, A morphological pyramidal image decomposition. *Pattern Recogn. Lett.* **9**(4), 255–261 (1989)
10. H Heijmans, J Goutsias, Multiresolution signal decomposition schemes-Part II: morphological wavelets. *IEEE Trans. Image Process.* **9**(11), 1897–1913 (2000)
11. X Liu, Y Zhang, S Hu, S Kwong, CCJ Kuo, Q Peng, Subjective and objective video quality assessment of 3D synthesized views with texture/depth compression distortion. *IEEE Trans. Image Process.* **24**(12), 4847–4861 (2015)
12. P. Conze, P. Robert, L. Morin, Objective view synthesis quality assessment. *Proc. SPIE 8288, Stereoscopic Displays and Applications XXIII* (2012)
13. F Battisti, E Bosc, M Carli, P Le Callet, S Perugia, Objective image quality assessment of 3D synthesized views. *Elsevier Signal Processing: Image Communication.* **30**(1), 78–88 (2015)
14. E Bosc, P Le Callet, L Morin, M Pressigout, *An edge-based structural distortion indicator for the quality assessment of 3D synthesized views, Picture Coding Symposium*, 2012, pp. 249–252
15. M. Solh, G. AlRegib, J.M. Bauza, 3VQM: A 3D video quality measure, 3VQM: a vision-based quality measure for DIBR-based 3D videos, *IEEE Int. Conf. on Multimedia and Expo (ICME)* (2011)
16. M. Solh, G. AlRegib, J.M. Bauza, A no reference quality measure for DIBR based 3D videos, *IEEE Int. Conf. on Multimedia and Expo (ICME)* (2011)
17. CT Tsai, HM Hang, *Quality assessment of 3D synthesized views with depth map distortion, visual communications and image processing (VCIP)*, 2013
18. M.S. Farid, M. Lucenteforte, M. Grangetto, Objective quality metric for 3d virtual views, *IEEE Int. Conf. on Image Processing (ICIP)* (2015)
19. Z. Wang, E. Simoncelli, A.C. Bovik, Multi-scale structural similarity for image quality assessment. *Asilomar Conference on Signals, Systems and Computers* (2003)
20. Z Wang, Q Li, Information content weighting for perceptual image quality assessment. *IEEE Trans. On Image Processing* **20**(5), 1185–1198 (2011)
21. PJ Burt, EH Adelson, The Laplacian pyramid as a compact image code. *IEEE Trans. on Communications* **31**(4), 532–540 (1983)
22. Z. Wang, E. Simoncelli, Translation insensitive image similarity in complex wavelet domain. *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal processing*, 573–576 (2005)
23. Y.K. Lai, C.C. Jay, Kuo, Image quality measurement using the Haar wavelet. *Proc. SPIE 3169, Wavelet Applications in Signal and Image Processing V*, 127 (1997)
24. S Rezazadeh, S Coulombe, A novel wavelet domain error-based image quality metric with enhanced perceptual performance. *Int. J. Comput. Electrical Eng.* **4**(3), 390–395 (2012)
25. X Gao, W Lu, D Tao, X Li, Image quality assessment based on multiscale geometric analysis. *IEEE Trans. Image Process.* **18**(7), 1409–1423 (2009)
26. E. Adelson, C. Anderson, J. Bergen, P. Burt, J. Ogden, Pyramid methods in image processing. *RCA Engineer* (1984)
27. S Mallat, Wavelets for a vision. *Proc. IEEE* **84**(4), 604–614 (1996)
28. F Meyer, P Maragos, Nonlinear scale-space representation with morphological levelings. *J. Vis. Commun. Image Represent.* **11**, 245–265 (2000)
29. G Matheron, *Random sets and integral geometry* (Wiley, New York, 1975)
30. J Serra, Introduction to mathematical morphology. *J. on Comput. Vision, Graph. Image Process.* **35**(3), 283–305 (1986)
31. J Goutsias, H Heijmans, Nonlinear multiresolution signal decomposition schemes—Part I: morphological pyramids. *IEEE Trans. Image Process.* **9**(11), 1862–1876 (2000)
32. D. Sandić-Stanković, Multiresolution decomposition using morphological filters for 3D volume image decorrelation. *European Signal Processing Conf. EUSIPCO, Barcelona* (2011)
33. H. Heijmans, J. Goutsias, Some thoughts on morphological pyramids and wavelets. *European Signal Processing Conf. EUSIPCO, Rodos* (1998)
34. D Chandler, S Hemami, VSNR: a wavelet-based visual signal-to-noise ratio for natural images. *IEEE Trans. Image Process.* **16**(9), 2284–2298 (2007)
35. H. Heijmans, J. Goutsias, Constructing morphological wavelets with the lifting scheme. *Int. Conf. on Pattern Recognition and Information Processing, Belarus*, 65–72 (1999)
36. S Mallat, Multifrequency channel decompositions of images and wavelet models. *IEEE Trans. on Acoustics, Speech and Signal Process.* **37**(12), 2091–2110 (1989)

37. H Heijmans, J Goutsias, *Multiresolution signal decomposition schemes Part2: morphological wavelets*. Tech. Rep. PNA-R9905 (CWI, Amsterdam, The Netherlands, 1999)
38. I Daubechies, W Sweldens, Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl.* **4**(3), 247–269 (1998)
39. J Kovacevic, M Vetterli, Nonseparable two- and three-dimensional wavelets. *IEEE Trans. on Signal Processing* **43**(5), 1269–1273 (1995)
40. H. Heijmans, J. Goutsias, Morphological pyramids and wavelets based on the quincunx lattice. in *Mathematical morphology and its applications to image and signal processing*, ed. by J Goutsias, L Vincent, D Bloomberg, (Springer US, 2000), 273–281
41. G. Uytterhoeven, A. Bultheel, The red-black wavelet transform. *Proc. of IEEE Benelux Signal Processing Symposium* (1997)
42. Z Wang, A Bovik, Mean squared error: love it or leave it. *IEEE Signal Process. Mag.* **26**(1), 98–117 (2009)
43. Z. Wang, A. Bovik, L. Lu, Why is image quality assessment so difficult. *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ASSP)*, 4, 3313–3316, Orlando FL, US (2002)
44. VQEG HDTV Group, *Test plan for evaluation of video quality models for use with high definition tv content*, 2009
45. C. Fehn, Depth image based rendering (DIBR), compression and transmission for a new approach on 3D-TV. *Proc. SPIE, Stereoscopic Displays and Applications XV*, 5291, 93–104, San Jose, CA (2004)
46. A Telea, An image inpainting technique based on the fast matching method. *J. Graph, GPU and Game Tools* **9**(1), 23–34 (2004)
47. Y Mori, N Fukushima, T Yendo, T Fujii, M Tanimoto, View generation with 3D warping using depth information for FTV. *Signal Process. Image Commun.* **24**(1–2), 65–72 (2009)
48. K Muller, A Smolic, K Dix, P Merkle, P Kauff, T Wiegand, View synthesis for advanced 3D video systems. *EURASIP Journal on Image and Video Processing* **2008**, 438148 (2008)
49. P. Ndjiki-Nya, P. Koppel, M. Doshkov, H. Lakshman, P. Merkle, K. Muller, T. Wiegand, Depth image based rendering with advanced texture synthesis. *IEEE Int. Conf. on Multimedia&Expo*, 424–429, Suntec City (2010)
50. M. Koppel, P. Ndjiki-Nya, M. Doshkov, H. Lakshman, P. Merkle, K. Muller, T. Wiegand, Temporally consistent handling of disocclusions with texture synthesis for depth-image-based rendering. *IEEE Int. Conf. on Image Processing*, 1809–1812, Hong Kong (2010)
51. M Solh, G AlRegib, *Depth adaptive hierarchical hole filling for DIBR-based 3D videos*, *Proceedings of SPIE*, 8290, 829004 (Burlingame, CA, US, 2012)
52. MP-PSNR matlab p-code. <https://sites.google.com/site/draganasandicstankovic/code/mp-psnr>
53. M Aubury, W Luk, Binomial filters. *Journal of VLSI Signal Processing for Signal, Image and Video Technology* **12**(1), 35–50 (1995)
54. K Gu, M Liu, G Zhai, X Yang, W Zhang, Quality assessment considering viewing distance and image resolution. *IEEE Trans. On Broadcasting* **61**(3), 520–531 (2015)
55. MW-PSNR matlab p-code. <https://sites.google.com/site/draganasandicstankovic/code/mw-psnr>
56. Z Wang, AC Bovik, A universal image quality index. *IEEE Signal Processing Letters* **9**(3), 81–84 (2002)
57. Z Wang, AC Bovik, HR Sheikh, E Simoncelli, Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)
58. SSIM matlab code. https://ece.uwaterloo.ca/~z70wang/research/ssim/ssim_index.m
59. 3DSwIM matlab p-code. <http://www.comlab.uniroma3.it/3DSwIM.html>

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
