Comparison of modern nonlinear multichannel filtering techniques using recent full-reference image quality assessment methods

Krzysztof Okarma*

Chair of Signal Processing and Multimedia Engineering, Szczecin University of Technology, 26. Kwietnia 10, 71-126 Szczecin, Poland

Abstract

In the paper the quality analysis of some modern nonlinear color image filtering methods is presented. Traditionally, many image filtering algorithms are analyzed using classical image quality assessment metrics, mainly based on the Mean Square Error (MSE). However, they are all poorly correlated with subjective evaluation of images performed by observers.

Due to necessity of better image quality estimation, some other methods have been recently proposed. They are especially useful for development of new lossy image compression algorithms, as well as evaluation of images obtained after applying some image processing algorithms e.g. filtering methods.

Most of image quality algorithms are based on the comparison of similarity between two images: the original (reference) one and the second one which is processed e.g. contaminated by noise, filtered or lossily compressed. Such a group of full-reference methods is actually the only existing universal solution for automatic image quality assessment. There are also some "blind" (no-reference) algorithms but they are “specialized” for some kinds of distortions e.g. blocky effects in the JPEG compressed images. The last years’ state-of-the-art full-reference metrics are Structural Similarity (SSIM) and M-SVD based on the Singular Value Decomposition of two images' respective blocks.

Another important aspect of color image quality assessment is the way the color information is utilized in the quality metric. The authors of two analyzed metrics generally do not consider the effects of using color information at all or limit the usage of their metrics to luminance information in YUV color model only so in this article the solutions based on RGB and CIE LAB models are compared.

In the paper the results of quality assessment using the SSIM and M-SVD methods obtained for some modern median-based filters and Distance-Directional Filter for color images are presented with comparison to those obtained using classical metrics as the verification of their usefulness.

* e-mail adress: okarma@ps.pl
1. Introduction

Median filtering is one of the most popular nonlinear techniques of noise elimination in digital images, especially useful for impulse noise removal. One of the most important features related to such an algorithm for grayscale images is the fact that, unlike in linear low-pass filters where the resulting pixel’s luminance is calculated as the weighted average of luminance of some pixels from the neighbourhood (within the filter mask), the median filter always chooses one of the luminance levels existing in the pixel’s neighbourhood as the result. As the consequence of such choice, there is a guarantee that no new luminance levels are introduced into the image which is especially important for images with limited palette.

Median filtering algorithms for color images also should prevent the situation when new colors, treated as combinations of values for three channels, are introduced into the image. For that reason, independent filtering of each color channel (red, green and blue typically) should not be used. Instead of that some vector based median filtering algorithms have been proposed but their performance is typically compared to existing solutions using conventional image quality metrics [1], such as Peak Signal to Noise Ratio (PSNR) or Mean Absolute Error (MAE). For color images there is a possibility of using Normalized Color Difference (NCD) metric defined in the CIE LAB color space but in many publications their authors use only classical quality measures which are poorly correlated with the way humans perceive images.

However, in recent years some new image quality assessment methods have been proposed, much better correlated with Human Visual System, so there is a need for verification of usefulness of some multichannel filtering methods from the point of view based on some modern image quality metrics.

2. Multichannel median filtering algorithms

The simplest approach for color median filtering is independent filtering of each color channel (marginal ordering) which can lead to many artifacts in the resulting image and introduce new colors into the image. Regardless of the disagreement with one of the most important aspects of median filtration, this effect can lead to poor results of quality assessment because of potential discolorations caused by ignoring the correlation among color channels. The only reasonable usage can be related to the situations when channels are strongly decorrelated (e.g. so called “television” color spaces such as YUV and YIQ).

The proper methodology of color image filtration is using vector-based filters which utilize data related to all channels. Assuming the most popular RGB color space for each pixel values from each channel are stored in the three-element vector and the ordering criterion of vectors corresponding to the pixels within
the current mask is dependent on the particular filter. The simplest approach is just the choice of specified channel as the main one and the values from other channels are analyzed only if two or more pixels have the same value of the main channel. Such approach is known as conditional ordering.

In reduced (aggregate) ordering techniques, the ordering criterion is calculated on the basis of all elements of the vector. The main difference between grayscale and such kind of vector processing is the choice of the output element. In the grayscale median filter (and in the simplest conditional ordering filters) the middle element of the sorted vector is chosen as the result but in vector processing pixels are usually ordered by the criterion related to e.g. aggregated distance to other elements in the current mask (window) so the lowest ranked element is chosen as the result. In these algorithms pixels significantly changing from their neighborhood (“outliers”) are usually high ranked and have small chances to be chosen.

Two major groups of reduced ordering filters are Vector Median Filters (VMF) and Vector Directional Filters (VDF). In the first approach the pixel with the minimum aggregated distance to the others within the sliding window is chosen as the result and typical distance measure is Minkowski metric [2]. One of the extensions of such filter is the combination with the linear low-pass filter (Arithmetic Mean Filter – AMF) known as Extended Vector Median Filter [3], useful for elimination of non-impulse noise e.g. Gaussian one. The other well known combination of AMF and VMF is $\alpha$-trimmed VMF where $\alpha$ elements with the lowest rank in VMF are averaged by AMF [4]. However, such approach can be used for mixed noise and does not guarantee that no new colors will be introduced into the resulting image. Similarly to grayscale weighted median filters such approach is also present in the field of color image processing. The most popular filter is known as Central Weighted Vector Median Filter [5].

The simplest filter from the VDF class is Basic Vector Directional Filter (BVDF) [6] eliminating the pixels with atypical directions in the specified color space (not necessarily RGB). The angular distance can be calculated using the following formula:

$$A_k = \sum_{j=1}^{N} \cos^{-1} \left( \frac{x_k \cdot x_j}{\|x_k\| \|x_j\|} \right), \quad k = 1..N,$$

where $x_j$ and $x_k$ are the three-element vectors in the specified color space and $N$ denotes the number of pixels inside the processing window. As the result, the pixel with the minimum angular distance $A_k$ is chosen. The modification of such algorithm known as Generalized Vector Directional Filter (GVDF) utilizes the cascading structure where the number of BVDF outputs with the minimum angular distance is filtered using e.g. distance norm in the second stage.
A combination of median and directional approaches described above is the Distance Directional Filter (DDF) which uses the weighted average of both magnitude and direction defined as [7]

\[
D_k = \left[ \sum_{j=1}^{N} \cos^{-1} \left( \frac{x_k \cdot x_j}{\|x_k\| \|x_j\|} \right) \right]^p \cdot \left[ \sum_{j=1}^{N} d(x_k, x_j) \right]^{1-p},
\]

where \(d(x_k, x_j)\) is calculated using the Minkowski metric (usually Euclidean distance) as in VMF.

Another interesting idea is using HSV color space with conditional ordering as proposed in paper [8]. Such filter, denoted as VMED, is based on the ordering primarily using V component in the ascending order, and then (only for equal values of V) in the descending ordering by S and finally, in the case of equal values of V and S, a hue component is used for sorting in the ascending order. As the result, the middle element is chosen similarly to the grayscale filter.

The filters described above have a number of further modifications, especially towards adaptive algorithms, but generally, the results presented by their authors are based usually only on the classical image quality assessment methods and NCD metric. In this paper, the results obtained by some of the discussed filters are verified using the modern image quality metrics in the CIE LAB color space as well as using the RGB model.

### 3. Modern image quality assessment methods

There are two general approaches for image quality assessment – objective and subjective methods. Only the first can be successfully applied in computer systems and for optimization tasks because of the necessity of time consuming human interaction in the methods based on subjective evaluation. A typical approach for subjective quality assessment is using Mean Opinion Score (MOS) as the metric calculated by analysis of the quality scores given by a number of observers. Certainly, such approach can be helpful for verification of existing algorithms but during research with the required direct usage of quality metric, it should be treated as rather a useless solution.

Objective image quality assessment methods are based on the calculation of preferably single scalar value corresponding to the overall quality of the image. More sophisticated algorithms allow the creation of quality maps or may give the opportunity to calculate a vector measure consisting of some independent metrics related to some specific distortions. The only limitation of such approach is the necessity of componing the distorted image with the original one which should be known (full-reference methods). However, existing blind (no-reference) or reduced-reference image quality measures are usually sensitive only to specific types of distortions (e.g. block effects on JPEG compression)
and their universality is far from the requirements of image processing society. Currently only full-reference objective image quality assessment methods, especially scalar ones, are treated as universal solution and can be utilized in many areas of digital image processing e.g. development of new filtration algorithms, lossy compression methods or video transmission techniques.

The classical measures of that type are based on Mean Square Error (MSE), e.g. Peak Signal to Noise Ratio (PSNR) or similar metrics like Mean Absolute Error (MAE). Most of them, unfortunately poorly correlated with Human Visual System, are briefly described in [1].

The first modern image quality metric is the Universal Image Quality Index [9] proposed by Wang and Bovik in 2002, further extended into Structural Similarity [10] defined as:

\[
SSIM = \frac{(2 \cdot \bar{X} \cdot \bar{X'} + C_1) \cdot (2 \cdot \sigma_{X'X} + C_2)}{(\bar{X}^2 + \bar{X'}^2 + C_1) \cdot (\sigma_X^2 + \sigma_{X'}^2 + C_2)},
\]

where \(X\) and \(X'\) stand for the original and distorted images.

The constants \(C_1\) and \(C_2\) are the most relevant extension of the Universal Image Quality Index preventing the stability of obtained results (possible division by zero especially for large flat and dark regions) and should be chosen as e.g. \(C_1 = (0.01 \times 255)^2\) and \(C_2 = (0.03 \times 255)^2\) as suggested by the authors of the paper [10] because they should not introduce any significant changes of the results. \(\bar{X}\) and \(\bar{X'}\) are the average values and \(\sigma\) is the standard deviation in the original and distorted image blocks respectively.

Using the SSIM index with the sliding window approach, it is possible to create quality map of the image. The default window is 11×11 pixels Gaussian one. The mean value of SSIM index from the whole quality map is treated as the overall image quality metric, sensitive to the three common types of distortions: loss of correlation, luminance distortion and loss of contrast.

Another interesting approach to objective image quality assessment has been presented in the paper [11] and is based on Singular Value Decomposition (SVD) calculated for 8×8 pixels blocks of original and distorted images. As the comparison of such values for each block, the following value is calculated:

\[
D_k = \sqrt{\sum_{k=1}^{n} \left( s_k - s'_k \right)^2 },
\]

where \(s_k\) and \(s'_k\) stand for the singular values for original and distorted image blocks.

The overall quality metric is calculated using the following formula:

\[
M_{SVD} = \frac{1}{K} \sum_{i=0}^{K} |D_i - D_{mid}|,
\]
where $K$ is the total number of blocks in the whole image and $D_{\text{mid}}$ denotes the middle element of the sorted vector $D$ (in fact, the median value).

Most image quality metrics are based on the assumption that evaluated images are grayscale ones. The same situation takes place for the SSIM and M-SVD measures. Their authors argue that color information is not important in image quality assessment at all or using simple conversion of color images from RGB to YUV color space in order to use only the luminance channel ($Y$) for quality evaluation. However, such approach is rather useless for color image processing, including multichannel median filtering techniques.

On the other hand, there is only one worth noticing quality metric defined particularly for evaluation of color images utilizing the CIE LAB color model (in some applications CIE LUV model is also used), known as Normalized Color Difference (NCD) defined as:

$$
NCD = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{(L_{i,j} - L'_{i,j})^2 + (a_{i,j} - a'_{i,j})^2 + (b_{i,j} - b'_{i,j})^2}}{\sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{L_{i,j}^2 + a_{i,j}^2 + b_{i,j}^2}},
$$

(6)

where $L,a,b$ denote the CIE LAB channels of the original image and $L',a',b'$ stand for the same for the distorted one. The conversion from RGB to CIE XYZ and then to CIE LAB color space is based on nonlinear functions depending also on the white point [12].

![Fig. 1. Contaminated versions of the example image used in experiments](image-url)
Although the NCD measure is based on the perceptually uniform character of CIE LAB color space, it is sensitive only to the color preservation, so its usefulness for quality assessment of noised and filtered images is limited.

As the method potentially useful for verification of some nonlinear filtering algorithms, the usage of SSIM and M-SVD measures in CIE LAB color space (also compared with the application of RGB) instead of often used traditional metrics such as MAE or PSNR is investigated. Test results have been obtained for several commonly used color test images contaminated by achromatic and color impulse noise with various strengths (interpreted as the number of contaminated pixels) filtered using some of the most popular color median filters.

4. Discussion of results

In the experiments several typical images have been used. Chosen multichannel filters have been applied for images without any noise and contaminated by achromatic and RGB impulse noise as well as by mixed achromatic (Gaussian and impulse) noise. The results obtained for intentionally chosen example image known as “Mandrill” or “Baboon” characterized by a large number of details are presented in Tables 1-7 (PSNR values are expressed in dB). The filters denoted as C-R, C-G and C-B are based on conditional ordering using specified channel as the main criterion. Gaussian noise has been added to the images assuming the mean value equal to zero and standard deviation equal to 50. Original image and its contaminated versions are presented in Fig. 1.

<table>
<thead>
<tr>
<th>Filter</th>
<th>MAE</th>
<th>PSNR</th>
<th>NCD</th>
<th>SSIM</th>
<th>M-SVD</th>
<th>SSIM (Lab)</th>
<th>M-SVD (Lab)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C – R</td>
<td>6.9877</td>
<td>-17.6152</td>
<td>0.0943</td>
<td>0.6108</td>
<td>29.6569</td>
<td>0.7353</td>
<td>9.0632</td>
</tr>
<tr>
<td>C – G</td>
<td>6.4056</td>
<td>-17.3188</td>
<td>0.0904</td>
<td>0.6557</td>
<td>32.8279</td>
<td>0.7537</td>
<td>9.6792</td>
</tr>
<tr>
<td>C – B</td>
<td>7.2064</td>
<td>-17.6951</td>
<td>0.0947</td>
<td>0.6012</td>
<td>26.2041</td>
<td>0.7350</td>
<td>8.5289</td>
</tr>
<tr>
<td>VMF</td>
<td>6.8161</td>
<td>-17.4284</td>
<td>0.0922</td>
<td>0.6194</td>
<td>29.6225</td>
<td>0.7413</td>
<td>9.2844</td>
</tr>
<tr>
<td>BVDF</td>
<td>18.8656</td>
<td>-20.8363</td>
<td>0.1130</td>
<td>0.4189</td>
<td>77.5342</td>
<td>0.6504</td>
<td>14.2793</td>
</tr>
<tr>
<td>DDF</td>
<td>6.8993</td>
<td>-17.4667</td>
<td>0.0922</td>
<td>0.6193</td>
<td>29.7504</td>
<td>0.7412</td>
<td>9.2671</td>
</tr>
<tr>
<td>VMED</td>
<td>6.6257</td>
<td>-17.4530</td>
<td>0.0925</td>
<td>0.6241</td>
<td>32.3511</td>
<td>0.7430</td>
<td>9.5595</td>
</tr>
</tbody>
</table>

Comparing the results obtained for different amounts of noise as well as for noiseless images, it can be stated that in most cases analysis of the results obtained using all the metrics lead to similar conclusions. Modern metrics (SSIM and M-SVD) calculated using CIE LAB color space are less sensitive in
### Table 2. Results obtained for the example test image with 10% impulse achromatic noise

<table>
<thead>
<tr>
<th>Filter</th>
<th>MAE</th>
<th>PSNR</th>
<th>NCD</th>
<th>SSIM</th>
<th>M-SVD</th>
<th>SSIM (Lab)</th>
<th>M-SVD (Lab)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>7.7770</td>
<td>-11.9591</td>
<td>0.0876</td>
<td>0.3340</td>
<td>69.4411</td>
<td>0.6428</td>
<td>17.3784</td>
</tr>
<tr>
<td>C – R</td>
<td>7.3132</td>
<td>-17.7100</td>
<td>0.0965</td>
<td>0.5937</td>
<td>30.9356</td>
<td>0.7284</td>
<td>9.2892</td>
</tr>
<tr>
<td>C – G</td>
<td>6.7979</td>
<td>-17.4682</td>
<td>0.0930</td>
<td>0.6341</td>
<td>33.5432</td>
<td>0.7453</td>
<td>9.8042</td>
</tr>
<tr>
<td>C – B</td>
<td>7.5312</td>
<td>-17.7880</td>
<td>0.0971</td>
<td>0.5844</td>
<td>28.1679</td>
<td>0.7277</td>
<td>8.8355</td>
</tr>
<tr>
<td>VMF</td>
<td>7.3413</td>
<td>-17.6131</td>
<td>0.0960</td>
<td>0.5869</td>
<td>29.9536</td>
<td>0.7275</td>
<td>9.3334</td>
</tr>
<tr>
<td>BVDF</td>
<td>21.8037</td>
<td>-20.4464</td>
<td>0.1468</td>
<td>0.1484</td>
<td>88.4815</td>
<td>0.4104</td>
<td>20.6454</td>
</tr>
<tr>
<td>DDF</td>
<td>14.5138</td>
<td>-18.5467</td>
<td>0.1547</td>
<td>0.2295</td>
<td>65.7433</td>
<td>0.4618</td>
<td>15.8062</td>
</tr>
<tr>
<td>VMED</td>
<td>6.9466</td>
<td>-17.5792</td>
<td>0.0944</td>
<td>0.6054</td>
<td>31.7789</td>
<td>0.7357</td>
<td>9.5320</td>
</tr>
</tbody>
</table>

### Table 3. Results obtained for the example test image with 10% impulse RGB noise (for each channel)

<table>
<thead>
<tr>
<th>Filter</th>
<th>MAE</th>
<th>PSNR</th>
<th>NCD</th>
<th>SSIM</th>
<th>M-SVD</th>
<th>SSIM (Lab)</th>
<th>M-SVD (Lab)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>5.8435</td>
<td>-10.6609</td>
<td>0.1879</td>
<td>0.3902</td>
<td>67.3842</td>
<td>0.3108</td>
<td>21.2670</td>
</tr>
<tr>
<td>C – R</td>
<td>10.6556</td>
<td>-18.6023</td>
<td>0.1648</td>
<td>0.3403</td>
<td>39.5144</td>
<td>0.4855</td>
<td>12.6873</td>
</tr>
<tr>
<td>C – G</td>
<td>9.3773</td>
<td>-18.4239</td>
<td>0.1399</td>
<td>0.4198</td>
<td>36.1377</td>
<td>0.5942</td>
<td>10.5547</td>
</tr>
<tr>
<td>C – B</td>
<td>12.1197</td>
<td>-18.9603</td>
<td>0.1622</td>
<td>0.3056</td>
<td>42.6909</td>
<td>0.4827</td>
<td>12.5533</td>
</tr>
<tr>
<td>VMF</td>
<td>7.2466</td>
<td>-17.6203</td>
<td>0.0974</td>
<td>0.5806</td>
<td>28.2981</td>
<td>0.7248</td>
<td>8.9625</td>
</tr>
<tr>
<td>BVDF</td>
<td>38.0191</td>
<td>-21.5425</td>
<td>0.4477</td>
<td>0.0785</td>
<td>139.5848</td>
<td>0.1300</td>
<td>44.4058</td>
</tr>
<tr>
<td>DDF</td>
<td>7.4165</td>
<td>-17.6903</td>
<td>0.0979</td>
<td>0.5785</td>
<td>28.4798</td>
<td>0.7226</td>
<td>8.9219</td>
</tr>
<tr>
<td>VMED</td>
<td>9.7919</td>
<td>-18.2037</td>
<td>0.1346</td>
<td>0.4218</td>
<td>34.3941</td>
<td>0.5632</td>
<td>10.3038</td>
</tr>
</tbody>
</table>

### Table 4. Results obtained for the example test image with 20% impulse achromatic noise

<table>
<thead>
<tr>
<th>Filter</th>
<th>MAE</th>
<th>PSNR</th>
<th>NCD</th>
<th>SSIM</th>
<th>M-SVD</th>
<th>SSIM (Lab)</th>
<th>M-SVD (Lab)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>21.3214</td>
<td>-16.3188</td>
<td>0.2312</td>
<td>0.1361</td>
<td>89.6137</td>
<td>0.5176</td>
<td>25.4763</td>
</tr>
<tr>
<td>C – R</td>
<td>8.4112</td>
<td>-18.0207</td>
<td>0.1038</td>
<td>0.5319</td>
<td>34.8514</td>
<td>0.6994</td>
<td>10.1082</td>
</tr>
<tr>
<td>C – G</td>
<td>8.0740</td>
<td>-17.8866</td>
<td>0.1011</td>
<td>0.5582</td>
<td>36.3815</td>
<td>0.7104</td>
<td>10.3616</td>
</tr>
<tr>
<td>C – B</td>
<td>8.6323</td>
<td>-18.0962</td>
<td>0.1045</td>
<td>0.5230</td>
<td>33.5931</td>
<td>0.6980</td>
<td>9.8294</td>
</tr>
<tr>
<td>VMF</td>
<td>9.3896</td>
<td>-18.0596</td>
<td>0.1137</td>
<td>0.4508</td>
<td>40.9190</td>
<td>0.6460</td>
<td>12.3321</td>
</tr>
<tr>
<td>BVDF</td>
<td>29.8866</td>
<td>-20.3305</td>
<td>0.1812</td>
<td>0.0564</td>
<td>97.0186</td>
<td>0.3257</td>
<td>26.2931</td>
</tr>
<tr>
<td>DDF</td>
<td>27.3535</td>
<td>-19.7821</td>
<td>0.2770</td>
<td>0.0659</td>
<td>86.4945</td>
<td>0.3348</td>
<td>24.2526</td>
</tr>
<tr>
<td>VMED</td>
<td>8.0185</td>
<td>-17.9305</td>
<td>0.1009</td>
<td>0.5398</td>
<td>32.7516</td>
<td>0.7051</td>
<td>9.8514</td>
</tr>
</tbody>
</table>

comparison to their application using RGB model. Relatively good results obtained using conditional ordering approach in the RGB color space for images without noise or with achromatic noise are caused by the specific character of the presented image with many small details (often very dark and bright).
Table 5. Results obtained for the example test image with 20% impulse RGB noise (for each channel)

<table>
<thead>
<tr>
<th>Filter</th>
<th>MAE</th>
<th>PSNR</th>
<th>NCD</th>
<th>SSIM</th>
<th>M-SVD</th>
<th>SSIM (Lab)</th>
<th>M-SVD (Lab)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>12.3485</td>
<td>-13.9330</td>
<td>0.3712</td>
<td>0.2218</td>
<td>79.6784</td>
<td>0.1704</td>
<td>24.0205</td>
</tr>
<tr>
<td>C – R</td>
<td>14.3979</td>
<td>-19.2073</td>
<td>0.2296</td>
<td>0.2248</td>
<td>58.0095</td>
<td>0.3441</td>
<td>19.1273</td>
</tr>
<tr>
<td>C – G</td>
<td>12.2771</td>
<td>-19.0931</td>
<td>0.1886</td>
<td>0.3064</td>
<td>50.6389</td>
<td>0.4823</td>
<td>15.4824</td>
</tr>
<tr>
<td>C – B</td>
<td>16.5274</td>
<td>-19.5947</td>
<td>0.2226</td>
<td>0.2044</td>
<td>59.5240</td>
<td>0.3552</td>
<td>18.5451</td>
</tr>
<tr>
<td>VMF</td>
<td>8.1136</td>
<td>-17.9324</td>
<td>0.1104</td>
<td>0.5070</td>
<td>26.9605</td>
<td>0.6760</td>
<td>8.7358</td>
</tr>
<tr>
<td>BVDF</td>
<td>48.6378</td>
<td>-21.9680</td>
<td>0.5231</td>
<td>0.0549</td>
<td>176.0226</td>
<td>0.0972</td>
<td>53.0522</td>
</tr>
<tr>
<td>DDF</td>
<td>8.9997</td>
<td>-18.1850</td>
<td>0.1168</td>
<td>0.4820</td>
<td>28.9840</td>
<td>0.6477</td>
<td>9.5710</td>
</tr>
<tr>
<td>VMED</td>
<td>14.2807</td>
<td>-19.1306</td>
<td>0.2043</td>
<td>0.2666</td>
<td>51.7398</td>
<td>0.3855</td>
<td>17.1105</td>
</tr>
</tbody>
</table>

Table 6. Results obtained for the example test image with 20% mixed achromatic noise

<table>
<thead>
<tr>
<th>Filter</th>
<th>MAE</th>
<th>PSNR</th>
<th>NCD</th>
<th>SSIM</th>
<th>M-SVD</th>
<th>SSIM (Lab)</th>
<th>M-SVD (Lab)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>7.7609</td>
<td>-11.9685</td>
<td>0.1064</td>
<td>0.3397</td>
<td>67.0485</td>
<td>0.6261</td>
<td>17.0522</td>
</tr>
<tr>
<td>C – R</td>
<td>5.4904</td>
<td>-16.4362</td>
<td>0.1005</td>
<td>0.5788</td>
<td>27.5555</td>
<td>0.7164</td>
<td>9.0956</td>
</tr>
<tr>
<td>C – G</td>
<td>4.9341</td>
<td>-16.0538</td>
<td>0.0954</td>
<td>0.6358</td>
<td>26.8894</td>
<td>0.7450</td>
<td>9.3280</td>
</tr>
<tr>
<td>C – B</td>
<td>5.6764</td>
<td>-16.5136</td>
<td>0.0996</td>
<td>0.5797</td>
<td>24.8318</td>
<td>0.7249</td>
<td>8.3613</td>
</tr>
<tr>
<td>VMF</td>
<td>5.5039</td>
<td>-16.3319</td>
<td>0.0983</td>
<td>0.5883</td>
<td>25.7287</td>
<td>0.7272</td>
<td>8.8714</td>
</tr>
<tr>
<td>BVDF</td>
<td>51.6325</td>
<td>-21.7006</td>
<td>0.1819</td>
<td>0.0743</td>
<td>158.2048</td>
<td>0.3298</td>
<td>36.1246</td>
</tr>
<tr>
<td>DDF</td>
<td>7.7420</td>
<td>-16.7831</td>
<td>0.1155</td>
<td>0.4391</td>
<td>37.9591</td>
<td>0.6130</td>
<td>12.2670</td>
</tr>
<tr>
<td>VMED</td>
<td>5.0889</td>
<td>-16.2010</td>
<td>0.0968</td>
<td>0.6061</td>
<td>25.8238</td>
<td>0.7345</td>
<td>9.0021</td>
</tr>
</tbody>
</table>

Table 7. Results obtained for the example test image with 20% mixed RGB noise (10% Gaussian and 10% impulse for each channel)

<table>
<thead>
<tr>
<th>Filter</th>
<th>MAE</th>
<th>PSNR</th>
<th>NCD</th>
<th>SSIM</th>
<th>M-SVD</th>
<th>SSIM (Lab)</th>
<th>M-SVD (Lab)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>5.6781</td>
<td>-10.7103</td>
<td>0.1958</td>
<td>0.4007</td>
<td>63.3730</td>
<td>0.3288</td>
<td>19.7451</td>
</tr>
<tr>
<td>C – R</td>
<td>8.6036</td>
<td>-17.5477</td>
<td>0.1628</td>
<td>0.3469</td>
<td>39.4056</td>
<td>0.4957</td>
<td>12.1941</td>
</tr>
<tr>
<td>C – G</td>
<td>7.3027</td>
<td>-17.3180</td>
<td>0.1385</td>
<td>0.4274</td>
<td>34.5436</td>
<td>0.6026</td>
<td>10.0461</td>
</tr>
<tr>
<td>C – B</td>
<td>10.0033</td>
<td>-18.0395</td>
<td>0.1604</td>
<td>0.3125</td>
<td>41.3757</td>
<td>0.4928</td>
<td>11.9684</td>
</tr>
<tr>
<td>VMF</td>
<td>5.4083</td>
<td>-16.3313</td>
<td>0.0998</td>
<td>0.5798</td>
<td>25.1091</td>
<td>0.7237</td>
<td>8.6641</td>
</tr>
<tr>
<td>BVDF</td>
<td>34.9835</td>
<td>-21.2612</td>
<td>0.4320</td>
<td>0.0833</td>
<td>133.7732</td>
<td>0.1400</td>
<td>42.3179</td>
</tr>
<tr>
<td>DDF</td>
<td>5.5405</td>
<td>-16.4094</td>
<td>0.1000</td>
<td>0.5793</td>
<td>24.9382</td>
<td>0.7232</td>
<td>8.5630</td>
</tr>
<tr>
<td>VMED</td>
<td>7.9058</td>
<td>-17.1966</td>
<td>0.1376</td>
<td>0.4170</td>
<td>34.9377</td>
<td>0.5621</td>
<td>10.1461</td>
</tr>
</tbody>
</table>

Usage of CIE LAB color space causes the necessity of modification of typical ranges for each channel, because dynamic range of L channel is from 0 to 100, and for A and B chrominances the range is from – 120 to 120. The assumption
of 256 levels for each RGB channel cannot be made for the CIE LAB model so it affects the obtained results of SSIM and M-SVD metrics in that color space. It is clearly visible for the M-SVD measure where much better quality can be expected analyzing the results obtained for the CIE LAB model in comparison to the assessment of the same images in RGB color space. Nevertheless, one should avoid the direct comparison of results obtained for different color models. The presented results show that there is a serious danger of incorrect conclusions in such a case.

Analyzing the obtained results it can be observed that application of CIE LAB color space for modern image quality metrics such as SSIM and M-SVD does not lead to better correlation with the human perception of images. As the illustration of that problem, the images presented in Fig. 2 can be investigated, especially for the SSIM measure as more attractive because of its much lower computational cost.

The result of filtration using the conditional – red filter presented in Fig. 2a is noticeably worse quality of the image while SSIM calculated in LAB color space increases from 0.3288 to 0.4957 (SSIM values in RGB are 0.4007 and 0.3469 respectively as shown in Table 7). The images presented in Fig. 2b illustrate the effect of using DDF filter for 20% achromatic mixed noise, the result of filtration is perceptually better but SSIM for CIE LAB gives almost identical results (0.6130 after filtration and 0.6261 for noisy image).

Fig. 2c illustrates the fragments of original image and that obtained using the conditional-blue filter. The resulting image has noticeably worse quality but the results of NCD and SSIM for CIE LAB as well as PSNR and M-SVD do not illustrate this effect adequately. It can be treated also as the example of M-SVD measure imperfection since its values indicate much better quality after filtration. In that case, only MAE and SSIM calculated on the RGB model illustrate the loss of quality properly, but one should notice that Mean Absolute Error can often produce unexpected results for other types of distortions.

Conclusions

The presented results can be treated both as the verification of usefulness of some filters and the differences between them as well as the confirmation of suitability of some modern image quality assessment methods for evaluation of multichannel nonlinear filters. The methods presented in the paper can be a useful tool for development of color image processing algorithms. It is worth noticing that SSIM is generally more convenient metric than M-SVD because of its lower computational complexity regardless of the fact that calculations are performed using the sliding window approach, not for the blocks.
Application of modern image quality metrics for comparison of multichannel median filtering algorithms usually leads to similar conclusions to the usage of classical methods but prevents some errors caused by poor correlation of traditional measures. Taking into account their correlation with the subjective evaluation, modern metrics can be treated as an effective alternative for traditional measures, especially for applications where several classical metrics should be used for reliable comparison of the obtained effects, such as in multichannel nonlinear image filtering methods. In those cases the usage of single modern quality measure should be enough for reliable comparison. Nevertheless, the application of some other color spaces for image quality...
assessment, also for evaluation of some adaptive multichannel median filters, can be an interesting field of further research.

References