Full and Reduced Reference Image Quality Assessment of Panoramic View using Novel Hybrid Image Stitching Method

Omkar S. Vaidya, Sanjay T. Gandhe

Abstract: Image Stitching is becoming more popular in field of computer vision because of rapid development of efficient algorithms that replaces the high cost wider lens cameras and commercial image stitching tools. The existing methods used global geometric transformation in registration stage and hence suffered from object deformation, parallax error, ghosting effect and motion blur in output result. In this paper, newly developed Hybrid Warping of weighted linearized homography matrix and similarity transform matrix is implemented over standard image stitching database. The visual quality of stitched image using proposed method has been examined in terms of performance metrics of Full Reference Image Quality Assessment (FRIQA) such as Root Mean Square Error (RMSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM) and Reduced Reference Image Quality Assessment (RRIQA). Also, the performance analysis of proposed method is compared against existing image stitching methods in terms of field of view and stitching time. This analysis has ascertained the outperformance of Novel Hybrid Image Stitching method.

Keywords: Full Reference and Reduced Reference Image Quality Assessment, Homography, Hybrid Warping, Image Stitching, Panoramic View.

I. INTRODUCTION

Image Stitching has remarkable inclination in computer vision and computer graphics because of its multifaceted applications in transportation, medical imaging, surveillance and multimedia technology etc. Most of the smartphone image stitching applications has limited control settings and require human interventions while lining up images to obtain limited field of view of panoramic image. Image Stitching is a method to align small multiple sequential images having considerable overlap among them and produce seamless panoramic image of broader field of view over common reference frame [1]. Image stitching has three main stages viz. image acquisition, image registration and image blending [2], [3]. The perfect image registration is key success to better image alignment between pair of images to be stitched.

The proposed method used hybrid combinations of linearized homography matrix and similarity transform matrix into single resultant homography matrix which then used to warp the input images for accurate image alignment. Further, this matrix has been filtered using Homography Screening technique before the warping process. The proposed method is tested on As-Projective-As-Possible Image Stitching Database [4] and compared with existing image stitching methods viz. As-Projective-A-Possible (APAP) [5], Shape Preserving Half Projective (SPHP) [6] and Elastic Local Alignment (ELA) [7]. The perceptual quality of stitched image is examined with the help of objective image quality assessment approach like FRIQA and RRIQA. When original pristine reference image is available against the stitched image whose quality is to be assessed, then such approach is FRIQA [8] while RRIQA possesses partial information regarding original reference image apart from stitched image [9].

The outline of the paper is as follows: Section II highlights detailed literature review of image stitching methods which is relevant to our proposed work. Section III explores the structured dataflow of proposed method. Section IV elaborates the objective quality assessment of stitched image. The performance analysis is evaluated and comparative results are shown in Section V. In the end, the paper is concluded in Section VI.

II. RELATED WORKS

The 2-D global transformation models are not sufficient to achieve accurate image alignment. The mapping and resampling of pixels is performed digitally to alleviate the local deformation while registering pair of image. The local warping models have more degrees of freedom than global transformation models [10].

Junhong Gao et.al. introduced Dual Homography Warping (DHW) which uses two homographies like ground plane (Hg) and distant plane (Hp) in [11]. As both homographies are of projective, the stitched image suffers from excessive extrapolation and severe stretching. Feng Liu and Michael Gleicher implemented 3-D video stabilization by using Content Preserve Warping (CPW) in [12] which not only resizes the image but it also preserves important content of image.
Igarashi et.al. introduced As Rigid As Possible (ARAP), a two-step algorithm to manipulate the object shape in [13]. This is an interactive method in which several points of object are transformed either by using rotation or warping to desired location.

Wen-Yan Lin et.al. used global affine transformation to primarily align images [14]. The affine parameter is averaged to get smooth affine stitching field. The affinity preserves the parallel lines in non-overlapping regions. The smoothly varying stitching field interpolates accurately and flexibly in overlapping regions to obtain better alignment. But performance of this method is degraded when there are no correspondence points (i.e. non-overlapping region) and produces distorted results. Fan Zhang and Feng Liu adopted optimal local homography for rough image alignment followed by content preserving warping technique in [15].

Julio Zaragoza et.al. presented APAP warp [5] used 2-D projective warp which aligns source and target image followed by Moving Direct Linear Transform (MDLT) to estimate local warp. The output of APAP warp gives better alignment than DHW and reduced extrapolation than SVA.

Chang et.al. proposed SPHP warp [6] which mitigate the excessive extrapolation in stitched image. But, this algorithm completely fails when structure is full of lines. Chung-Ching Lin et.al. proposed stitching method based on least intersection angle between projection plane with scene to achieve better alignment in [16]. Jing Li et.al. proposed the warping method that combined mesh based model with direct deformation scheme in [7]. The warping model successfully eliminates the parallax error and maintains alignment accuracy.

Nan Li et.al. developed Quasi homography warp that overcomes the problem of projective and perspective distortion by using local consistent scale linearization and global consistent slope preservation respectively [17].

In most of the existing image stitching based research, the local projective warp is used to improve image alignment accuracy. But due to perspective nature of warp, severe extrapolation is observed especially in non-overlapping region of stitched image. This deteriorates the output result in terms of area and shape distortion.

The importance of selecting quality homography has been ignored by various image stitching methods up till. This poses major challenges such as structural misalignment, ghosting effect, motion blur and parallax error in output result. Therefore, Development of Novel Hybrid Image Stitching method can be a one stop solution to address above problems.

### III. NOVEL HYBRID IMAGE STITCHING METHOD

Thorough insight into the research work has been attempted by means of a detailed dataflow of proposed method as shown in Fig. 1.

The pair of input images having a certain overlap of a particular view is assumed as Source Image (I) and Target Image (I'). To reduce the complexity and improve the accuracy of preceding algorithms, the original input image is converted to standardized form. It includes conversion of colour image into grayscale image, changing size of image to lower dimensions and conversion of single precision floating points. All of them help to avoid memory to go out of range. Features are extracted from each image using SIFT algorithm. For the purpose of image registration, rough feature matching is carried out using Sum of Squared Euclidean Distance (SSED) with specified threshold value. The matched points are first normalized because the normalization step nullifies the effect of coordinate change by selecting canonical coordinate frame [18].

![Fig. 1 Flowchart of Proposed Method](image-url)
Hence to rectify the non-linear distortion, the local homography is linearized by incorporating partial derivatives. The global similarity transformation is computed from normalized inlier points of source and target image. The resultant homography is calculated using Hybrid combination of Similarity Transform and Linearized local warp. After this, the homography is filtered using Homography Screening algorithm then warping operation on both homographies is performed which is known as Hybrid Warping [19]. In addition to this, one more approach is adopted to repair any impaired structure in stitched image by using edge strength computation.

The compositing surface over which both source and target images are projected is mapped with horizontal and vertical lines. Thus, the rectilinear projection is a qualifying approach among various types used for composting the surface. After image composition, the task is to conceal the visible seams, any intensity difference during image registration and ghosting effect in the overlapping part. An efficient, speedy and computationally least complex Feathering blending technique is chosen to rectify post composting problems. The stitched image is obtained once the blending operation is performed and then the Angular and Linear Field of View (AFOV and LFOV) of stitched image is computed. In the end, to evaluate the quality of stitched image, Full Reference Image Quality Assessment (FRIQA) and Reduced Reference Image Quality Assessment (RRIQA) models are explored.

IV. OBJECTIVE QUALITY ASSESSMENT OF STITCHED IMAGE

The perceptual quality of stitched image is examined with the help of objective image quality assessment approach.

A. Full Reference Image Quality Assessment (FRIQA)

If original pristine reference image is available against the stitched image whose quality is to be assessed, then such approach is referred to as Full Reference Image Quality Assessment (FRIQA) [8]. Performance metrics of FRIQA are Root Mean Square Error (RMSE), Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM). These matrices provide equations which are used to quantify error and help to characterize the image quality [20].

Root Mean Square Error gives an average value of error in the stitched image as compared with its corresponding original image. Consider \( f(x, y) \) is original input image while \( f(x, y) \) is output stitched image. RMSE is obtained as given in eq. (1).

\[
RMSE = \sqrt{\frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [f(x, y) - f(x, y)]^2}
\]  

Where, \( M \) and \( N \) are dimensions of image.

For an eight bit image (\( b=8 \)), Peak Signal-to-Noise Ratio (PSNR) is expressed in eq. (2).

\[
PSNR = 20 \log_{10} \left( \frac{2^{b-1}}{\sqrt{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [f(x, y) - f(x, y)]^2}} \right)
\]  

Against noise, the PSNR is measured with respect to peak signal power which is a constant value and SNR is measured with respect to actual signal power which is a fluctuating signal. Therefore, PSNR value is more meaningful when quantifiably comparing with different image stitching methods.

Zhou Wang et. al. developed the quality assessment method named Structural Similarity Index (SSIM) which uses the knowledge of Human Visual System (HVS) in [21]. Structural Similarity Index is a performance metric that compares local patterns of normalized pixel intensities for luminance and contrast. SSIM estimates the shift in luminance, contrast distortion and structure changes in stitched image with reference to original image.

The Structural Similarity Index (SSIM) between original signal \( x \) and stitched signal \( y \) is represented in eq. (3).

\[
SSIM(x,y) = l(x,y) \cdot c(x,y) \cdot s(x,y)
\]  

Where, \( \alpha, \beta \) and \( \gamma \) are non-zero adjustable parameters which then adjust relative importance of luminance, contrast and structure respectively.

The performance metric of FRIQA is widely used because it is mathematically convenient and easy to compute if reference image is available but it does not match well with the perceived visual quality of reference image. Also, the FRIQA does not estimate the type of distortion in image. To overcome this issue, Reduced Reference Image Quality Assessment (RRIQA) is proposed.

B. Reduced Reference Image Quality Assessment (RRIQA)

Wang and Simoncelli developed Reduced Reference Image Quality Assessment model based on natural scene statistics model in wavelet transform domain in [9]. Reduced Reference approach possesses partial information regarding original reference image apart from stitched image whose quality is to be assessed. It provides distortion score between original image and stitched image. Therefore lower the value of RRIQA score, better the quality of stitched image. If the standard pristine resulting image using any image stitching method is not available, then authors in [22] proposed Blind Image Quality Assessment.

V. PERFORMANCE ANALYSIS

The computing platform used for image stitching methods is Corei3 Intel Processor with 4 GB RAM, 2 GHz processor speed and Windows 10 Operating System. The experimentation is performed in MATLAB 2017a and results obtained over As-Projective-As Possible Image Stitching Database [4].

A. Performance Analysis of FRIQA

To validate the structural misalignment error, spatial patches are extracted from overlapping region of stitched image and it is compared with that of original image.

To quantify the accuracy of image alignment in stitched image, RMSE is calculated between original image and stitched image. Computation of RMSE value over different image stitching methods for each dataset is given in Table I.
Table I: Comparison of Root Mean Square Error (RMSE) Value

<table>
<thead>
<tr>
<th>Image Dataset</th>
<th>APAP</th>
<th>SPHP</th>
<th>ELA</th>
<th>Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temple</td>
<td>15.95</td>
<td>20.22</td>
<td>14.27</td>
<td>10.89</td>
</tr>
<tr>
<td>Railtracks</td>
<td>58.64</td>
<td>57.33</td>
<td>57.55</td>
<td>46.27</td>
</tr>
<tr>
<td>Rooftops</td>
<td>27.52</td>
<td>24.26</td>
<td>29.34</td>
<td>20.73</td>
</tr>
<tr>
<td>Garden</td>
<td>41.03</td>
<td>30.08</td>
<td>32.98</td>
<td>21.59</td>
</tr>
<tr>
<td>Apartment</td>
<td>24.34</td>
<td>17.73</td>
<td>22.73</td>
<td>14.21</td>
</tr>
<tr>
<td>Chess Girl</td>
<td>20.23</td>
<td>26.41</td>
<td>25.37</td>
<td>20.11</td>
</tr>
<tr>
<td>Couch</td>
<td>23.09</td>
<td>18.16</td>
<td>21.61</td>
<td>16.49</td>
</tr>
</tbody>
</table>

Table I shows that our proposed method gives lower value of RMSE as compared with other image stitching method in each dataset. This is the evidence of improvement in image alignment being achieved using our proposed method. The RMSE value of Railtracks dataset is higher as compared to other datasets. Railtracks is the most challenging dataset which has dense and complicated structure of railway routes and the slightest change in translation produces erroneous result.

To assess the quality of stitched image in terms of information loss, the Peak Signal-to-Noise Ratio (PSNR) is computed between original image and stitched image. Computation of PSNR value over different image stitching methods for each dataset is given in Table II.

Table II: Comparison of Peak Signal-to-Noise Ratio (PSNR) Value in dB

<table>
<thead>
<tr>
<th>Image Dataset</th>
<th>APAP</th>
<th>SPHP</th>
<th>ELA</th>
<th>Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temple</td>
<td>24.07</td>
<td>22.01</td>
<td>25.04</td>
<td>27.38</td>
</tr>
<tr>
<td>Railtracks</td>
<td>12.76</td>
<td>12.96</td>
<td>12.92</td>
<td>14.82</td>
</tr>
<tr>
<td>Rooftops</td>
<td>19.33</td>
<td>20.43</td>
<td>18.78</td>
<td>21.79</td>
</tr>
<tr>
<td>Garden</td>
<td>15.86</td>
<td>18.56</td>
<td>17.76</td>
<td>21.44</td>
</tr>
<tr>
<td>Apartment</td>
<td>20.4</td>
<td>23.15</td>
<td>20.99</td>
<td>25.07</td>
</tr>
<tr>
<td>Chess Girl</td>
<td>22</td>
<td>19.69</td>
<td>20.04</td>
<td>22.06</td>
</tr>
<tr>
<td>Couch</td>
<td>20.85</td>
<td>22.94</td>
<td>21.43</td>
<td>23.78</td>
</tr>
</tbody>
</table>

From Table II, it is clear that for each dataset, our proposed method gives higher value of PSNR.

To assess the quality of stitched image in terms of degradation caused due to noise or blur, Structural Similarity (SSIM) Index is computed between original image and stitched image. Computation of SSIM value over different image stitching methods for each dataset is given in Table III.

Table III: Comparison of Structural Similarity (SSIM) Index Value

<table>
<thead>
<tr>
<th>Image Dataset</th>
<th>APAP</th>
<th>SPHP</th>
<th>ELA</th>
<th>Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temple</td>
<td>0.72</td>
<td>0.60</td>
<td>0.67</td>
<td>0.75</td>
</tr>
<tr>
<td>Railtracks</td>
<td>0.53</td>
<td>0.49</td>
<td>0.52</td>
<td>0.61</td>
</tr>
<tr>
<td>Rooftops</td>
<td>0.72</td>
<td>0.77</td>
<td>0.71</td>
<td>0.80</td>
</tr>
<tr>
<td>Garden</td>
<td>0.51</td>
<td>0.6613</td>
<td>0.6655</td>
<td>0.80</td>
</tr>
<tr>
<td>Apartment</td>
<td>0.65</td>
<td>0.75</td>
<td>0.63</td>
<td>0.81</td>
</tr>
<tr>
<td>Chess Girl</td>
<td>0.91</td>
<td>0.8856</td>
<td>0.8845</td>
<td>0.9132</td>
</tr>
<tr>
<td>Couch</td>
<td>0.76</td>
<td>0.82</td>
<td>0.77</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Table III shows that Structural Similarity index for each dataset in our proposed method is higher than other image stitching methods. This indicates that our proposed method gives better quality of stitched image against undesired degradation due to employment of Homography Screening before hybrid Warping.
Fig. 2 Geometry of overlapping images \[1\]

\[
\theta = 2\tan^{-1}\left(\frac{\text{Sensor width (mm)}}{2 \times \text{focal length (mm)}}\right) \tag{4} 
\]

The angle of rotation between successive images is denoted by \(\alpha\) and is computed from equation (5) as given below:

\[
\frac{l}{L} = \frac{1}{2} \left(\frac{\tan\left(\alpha - \frac{\theta}{2}\right)}{\tan\left(\frac{\theta}{2}\right)}\right) \tag{5} 
\]

The LFOV of stitched image is computed from its horizontal width but AFOV is computed from equation (6) as given below:

\[
\theta_{\text{stitched image}} = 2\tan^{-1}\left(\frac{L_{\text{stitched image}}}{D}\right) \tag{6} 
\]

Comparison of AFOV and LFOV among image stitching methods for each dataset is given in Table V and Table VI respectively.

**Table- IV: Comparison of LFOV of stitched image**

<table>
<thead>
<tr>
<th>Image Dataset</th>
<th>LFOV of input image (mm)</th>
<th>LFOV of stitched image (mm)</th>
<th>SPHP</th>
<th>ELA</th>
<th>Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temple</td>
<td>258</td>
<td>250</td>
<td>323</td>
<td>279</td>
<td></td>
</tr>
<tr>
<td>Railtracks</td>
<td>135</td>
<td>248</td>
<td>278</td>
<td>392</td>
<td></td>
</tr>
<tr>
<td>Rooftops</td>
<td>113</td>
<td>244</td>
<td>118</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>Garden</td>
<td>74</td>
<td>252</td>
<td>257</td>
<td>358</td>
<td></td>
</tr>
<tr>
<td>Apartment</td>
<td>151</td>
<td>253</td>
<td>275</td>
<td>405</td>
<td></td>
</tr>
<tr>
<td>Chess Girl</td>
<td>97</td>
<td>205</td>
<td>171</td>
<td>328</td>
<td></td>
</tr>
<tr>
<td>Couch</td>
<td>97</td>
<td>204</td>
<td>169</td>
<td>321</td>
<td></td>
</tr>
</tbody>
</table>

D. Performance Analysis of Stitching Processing Time

Stitching processing time is the aggregation of time required by feature detection and matching, outlier removal, warping and blending process. The overall stitching processing time each dataset of our proposed method is tabulated and compared with APAP, SPHP and ELA image stitching method as shown in Table VII.

**Table- VIII: Comparison of Stitching Processing Time (sec)**

<table>
<thead>
<tr>
<th>Image Dataset</th>
<th>Overall Stitching Processing Time (sec)</th>
<th>APAP</th>
<th>SPHP</th>
<th>ELA</th>
<th>Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temple</td>
<td>8.02</td>
<td>10.61</td>
<td>4.68</td>
<td>5.19</td>
<td></td>
</tr>
<tr>
<td>Railtracks</td>
<td>12.74</td>
<td>13.00</td>
<td>10.02</td>
<td>8.94</td>
<td></td>
</tr>
<tr>
<td>Rooftops</td>
<td>3.9</td>
<td>3.77</td>
<td>1.98</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>Garden</td>
<td>16.7</td>
<td>12.51</td>
<td>9.82</td>
<td>12.28</td>
<td></td>
</tr>
<tr>
<td>Apartment</td>
<td>16.58</td>
<td>14.67</td>
<td>11.49</td>
<td>13.06</td>
<td></td>
</tr>
<tr>
<td>Chess Girl</td>
<td>14.73</td>
<td>10.31</td>
<td>10.39</td>
<td>10.79</td>
<td></td>
</tr>
<tr>
<td>Couch</td>
<td>15.99</td>
<td>11.11</td>
<td>10.17</td>
<td>11.71</td>
<td></td>
</tr>
</tbody>
</table>

Table VII shows that due to MATLAB Executable acceleration function, the warping processing time gets reduced in our proposed method. Computation time depends on size of image and number of features extracted from each image. Therefore, the computation time is less for Rooftops dataset and computation time recorded for Apartment, Garden, and Couch and Chess Girl dataset is more than 10 sec.

E. Qualitative comparison of stitched image on Temple image dataset

Consider the pair of Temple dataset of image size 487 × 730 as shown in Fig. 3.
The final stitched image obtained by use of different image stitching methods on Temple dataset is shown in Fig. 4 through Fig. 7.

From Fig. 4 and Fig. 5, structural misalignments are clearly noticed especially around region of ground plane, top of temple roof and person standing on the road. These objects are highlighted by red box.

In case of APAP and ELA method, the stitched image produces severe extrapolation. The unnatural tilting of source image in SPHP method loses perceptivity of panoramic image. The temple object and building structures in Fig. 4 – Fig. 6 are not parallel with each other and it is highlighted by red color circle.

In Fig. 6, the object around the person standing on road suffers from motion blur. The stitched image obtained by our proposed method alleviates all the above issues and is shown in Fig. 7.

VI. CONCLUSION

Full Reference Image Quality Assessment (FRIQA) and Reduced Reference Image Quality Assessment (RRIQA) have been evaluated between reference input image and local spatial patch of stitched image. The result analysis shows that Hybrid warping of source and target image in such a way that image alignment in overlapping region has been matched perfectly. Reduction in stitching processing time using our proposed method has been achieved as compared with SPHP and APAP image stitching method and is comparable with ELA image stitching method. The wider angular and linear field of view has been computed in our proposed method as compared to other image stitching method. The developed image stitching model would benefit the medical, transportation, multimedia, photogrammetry and surveillance fields due to wider field of view of panoramic image which is free from motion parallax, ghosting artifact, visible seams and structure deformation.

ACKNOWLEDGMENT

This research work is financially supported by Savitribai Phule Pune University (SPPU), India through Assistance by SPPU for Project-based Innovative Research (ASPIRE) under Research Mentorship Program Scheme (18TEC001416) which is supervised by Internal Quality Assurance Cell (IQAC), SPPU, Pune, India.
REFERENCES


4. As-Projective-As-Possible Image Stitching Dataset
   The School of Computer Science, The University of Adelaide, Australia. [Available online at]: https://cs.adelaide.edu.au/˜tchinh/app/#/Datasets


AUTHORS PROFILE

Omkar S. Vaidya received B.E. degree in 2009, M.E. degree in 2013 and currently pursuing Ph.D. He is an Assistant Professor in E&C Dept. of Sandip Institute of Technology & Research Centre, Nashik. He is a life member of IEEE, ISTE, IEI, ISCA, IAENG and IRED. He has published 25 papers in International Journals and presented 24 papers in National/International Conferences. He has received Best Paper Award twice in conferences. His research interests are Image Processing and Computer Vision.

Dr. Sanjay T. Gandhe is Research Supervisor and Principal of Sandip Institute of Technology & Research Centre, Nashik. He completed B.E., M.E. and Ph.D. (VNIT, Nagpur) in 1990, 1996 and 2008 respectively. He has published 33 research papers in International Journals and presented 41 papers in National/International Conferences. He is fellow member of IETE, IEI, Senior Member of IEEE and life member of ISTE. He has fetched several Research Grants from Governmental Apex Bodies. He is honored as Best Principal Award and Best Teachers in Professional Educational Category. He was Reviewer of IEEE Transaction on Dependable and Secure Computing in 2011-12 and IEEE Access in 2019-2020. He served as Session Chair in various National and International Conferences. His area of expertise is Image Processing, Robotics and Machine Vision.