Response to Reviewer 1 Comments

1. Summary

Thank you very much for taking the time to review this manuscript. Please find the detailed responses below and the corrections highlighted in the re-submitted files.

2. Questions for General Evaluation	**Reviewer's Evaluation**	**Response and Revisions**
Does the introduction provide sufficient background and include all relevant references?	Yes	The revisions have been made based on reviewer's comments.
Are all the cited references relevant to the research?	Yes	The revisions have been made based on reviewer's comments.
Is the research design appropriate?	Yes	The revisions have been made based on reviewer's comments.
Are the methods adequately described?	Must be improved	The revisions have been made based on reviewer's comments.
Are the results clearly presented?	Must be improved	The revisions have been made based on reviewer's comments.
Are the conclusions supported by the results?	Can be improved	The revisions have been made based on reviewer's comments.

3. Point-by-point response to Comments and Suggestions for Authors

Comments 1: Figure 1: Text is not in English (it was correct in version 1).

Response 1: Thank you very much for your correction. We have corrected this issue. You can view the revised Figure 1 in the revised manuscript.

Comments 2: Terminology is not consistent. For example the following terms all appear to refer to the same thing: contour, edge, curve, spindle (but this is not the only example). Pick one word or phrase for each concept and use it throughout.

Response 2: Thank you for your suggestion. Based on your suggestions, we have unified definitions of the operation object terms in sections 3.2 and 3.3. Local edge object: In the process of extracting local edge contour features, we name the object of the entire operation as a local edge object, and continuously update its two attribute requirements after each operation. Local edge contour: used to indicate the contour structure of local edge objects during operations. For example, in the contour merging process, we believe that the term "local edge contour" is more appropriate. In the process of "structural separation", "edge" is not only a commonly used description operation object in closed operations, but also a detection result of the canny algorithm. "Curve" is also the local edge object in the separation process, now we change it to "local edge object". At the same time, we have unified the

terms used such as threshold. You can find the revised terminology in L246、L249-L253、L255、L256、L262、L275、L289、L290、L297、L308、L314、L318、L328-L329、L335、L361、L389-L390、L407、L422 and L433 of the revised manuscript.

Comments 3:249-256: Don't need to explain how to establish a "minimum bounding box" as this is a well-established concept.

Response 3: Thanks for your advice. Based on your suggestion, we have removed the explanation of minimum enclosing rectangle and used a noun instead. You can find the modifications in L254 of the revised manuscript.

Comments 4: 263:"the preset distance threshold" - I think "a" was meant rather than "the", to indicate this parameter is being introduced here rather than referring to something established earlier. It may seem like a minor grammatical quibble but the difference is important. This sort of issue is common throughout the text making it difficult to read accurately.

Response 4: Thank you for your suggestion. Based on your suggestion, we have corrected the syntax error. You can find the revised content in L257, L262, L289, and L297 of the revised manuscript.

Comments 5: 277: What are the requirements for "the required regular shape contours"? These are not stated.

Response 5: Thank you for your suggestion. We explained the regular shape outline required in this article at the beginning of Section 3.2. That is to say, the local edge objects that conform to the quadratic curve are separated, and the separated local edge objects conform to the straight line structure or the curve structure that approximates the parabola. Here we explain it in detail. And in the parts mentioned later, we no longer use the word "regular shape", but use specific descriptions. You can find the modifications in L242-L245 and L276-L279 of the revised manuscript.

Comments 6: 284-286: Peak detection has a deep prior literature, and this is perhaps the most naive possible implementation. Figure 5 even shows what should probably be considered spurious peaks due to very small changes in edge direction. A more sophisticated peak detection algorithm would fix this.

Response 6: Thank you for your suggestion. We strongly agree with your perspective. In this paper, we employed the classical peak detection method to identify local maxima in the data, which may lead to the detection of false peaks. Therefore, it has certain limitations. Due to our limited understanding of the peak detection research direction and the constraints on modification time, we are currently retaining the current method. Moving forward, we will delve into the latest literature and methods related to peak detection to enhance the capability of the edge detection module in our approach. Thank you for your understanding and thank you again for your suggestions.

Comments 7: 287: What's a "spindle" here? This word is never used again in the manuscript.

Response 7: Thank you for your opinion. In the detection results of the canny method, each local edge object has a main axis attribute, including the direction and length of the edge. We set a length threshold according to the length attribute in the main axis attribute to filter smaller outline structures. You can find "spindle length threshold" in L289 of the revised manuscript.

Comments 8: 293: It looks like a sentence was cut off around "the local edge Similar structures exist" (after "edge").

Response 8: Thank you for correcting our shortcomings. We reorganize the language of this description. This sentence mainly describes the cause of the problem that the separation operation produces similar edge structures. cause: The new local edge objects separated in the X and Y

directions have similar structures. You can find the revised content in L294-L295 of the revised manuscript.

Comments 9: 311: Is this threshold parameter the same one mentioned on line 296, or a new one?

Response 9: Thank you for your suggestion. These two parameter thresholds are consistent. In the paper, we designate a unique name (edge distance threshold) to represent this threshold parameter and specify the parameter value in the experimental parameter description. You can find these two thresholds at L297 and L314 of the revised manuscript.

Comments 10: 318-319: Don’t need to explain what a polynomial is.

Response 10: Thank you for your suggestion. We have deleted the introduction of polynomial formulas and principles and use terminology to refer to them in the text. You can find the corresponding content in L321-L322 of the revised manuscript.

Comments 11: 319: What polynomial order n was selected, and why?

Response 11: Thank you for your suggestion. Based on your suggestion, we explicitly indicate in the paper the order of the polynomial used. Because the edge structures we obtained conform to quadratic curves, we chose to use a binomial fitting model to meet this requirement. You can find the revised content in L321-L322 of the revised manuscript.

Comments 12: 324-325: Don’t need to define Residual Sum of Squares.

Response 12: Thank you for your suggestion. Based on your comments, we have deleted the formula and principle introduction of the residual sum of squares, which are mentioned using terminology in the text. You can find the revised content in L322-L325 of the revised manuscript.

Comments 13: Figure 6(c): I believe Section 3.2.2 should reference panel C as an example of the results of the contour merging operation.

Response 13: Thank you for your suggestion. Based on your suggestion, we have included Figure 7(c) as an example of the result of the contour merging operation in Section 3.2.2. After modification, Figure 6 has been changed to Figure 7. You can find the added content in L327-L328 of the revised manuscript.

Comments 14: 336: As a homography only maps a plane to a plane, how robust is this step in the face of non-coplanar structures in the image? Addressing this issue was one of the strengths of Multi-GS, so why is that algorithm not being leveraged here?

Response 14: Thank you for your suggestion. First, we explain the issues regarding the robustness of Multi-GS in this process: The image dataset we spliced is an image containing multiple planar structures. The article shows that the global homography matrix used only achieves a rough alignment of the overlapping areas of the images. Before calculating the global homography, we use the more robust Multi-GS algorithm to filter out outliers and quickly generate hypothesis models in multi-structure data, thereby processing large-scale or complex data more efficiently. Secondly, there is no mention of the Multi-GS algorithm issue regarding this process: The Multi-GS algorithm is an outlier filtering operation before global homography calculation. Only the calculation of global homography is mentioned in this paragraph. Based on your suggestion, we have added the expression used by the Multi-GS algorithm in this snippet. You can find the added content in L340-L341 of the revised manuscript.

Comments 15: 338: Which "matching feature pairs" are used to calculate the homography?

Response 15: Thank you for your suggestion. Based on your suggestion, we describe in detail the feature point pairs used to calculate the global homography matrix in this article. We calculate the

global homography matrix using matching feature point pairs extracted by SIFT and point-line consistency. You can find the revised content in L337-L340 of the revised manuscript.

Comments 16: 347: What are mask_refer and mask_target? They were never defined.

Response 16: Thank you for your suggestion. The previous article stated that mask_refer and mask_target represent the binary image of the reference image and the target image, in order to facilitate the calculation formula of the overlapping area mask. Based on your suggestion, this paragraph has been changed to a text description and the symbols have been deleted. You can find the modifications in L352-L359 of the revised manuscript.

Comments 17: 349: Don't need Equation 3-8 for this, the text description is enough.

Response 17: Thank you for your suggestion. Based on your suggestion, this paragraph has been changed to a text description and the symbols have been deleted. You can find the modifications in L352-L359 of the revised manuscript.

Comments 18: 356-373: Is edge contour filtering also applied to the target image? If not, why?

Response 18: Thank you for your suggestion. Both the reference image and the target image require local edge contour filtering operations, and only the reference image is used as an example of filtering in this article. The target image needs to undergo the same local edge contour filtering operation. You can find the modifications in L386-L387 of the revised manuscript.

Comments 19: 363: L(e_i) is defined but never referenced again. Why was it defined?

Response 19: Thank you for your suggestion. Based on your suggestion, we have removed the definition of the principal axis length of the local edge object. In order to facilitate the understanding of dense edge filtering operations, local edge objects are defined. L(e_i) is defined in the article, but the definition L(e_i) is not used in the follow-up. We have deleted the definition L(e_i). You can find the modifications in L365-L366 of the revised manuscript.

Comments 20: 364: should reference formula 3-4, not 3-7.

Response 20: Thanks for your correction. We have corrected the use of formula . You can find the modifications in L370 of the revised manuscript.

Comments 21: 366: Another cut off sentence "considered to be The two edge" (after "be").

Response 21: Thank you for correcting our shortcomings. We rephrase this sentence. This section mainly determines the relationship between two local edge objects through the distance matrix by setting the minimum edge distance threshold. The specific modifications are as follows: "A minimum edge distance threshold T is defined to assess the distance relationship between two local edge objects. If the distance value is greater than T, it is considered that the two local edge objects do not affect each other, and the corresponding position of the distance matrix is set to 0. Otherwise, the two local edge objects are considered to interfere with each other, and the corresponding position of the distance matrix is set to 1."

Comments 22: 370: In the phrase "delete the relationship between the remaining edge objects in the edge object set and the edge object", is this referring to a row in a distance matrix? It would be clearer to describe this whole section as operating on a distance matrix.

Response 22: Thank you for your suggestion. When we designed the algorithm, we used a matrix for filtering operations. Based on your suggestion, we describe this part in conjunction with the distance matrix. The sentence "delete the relationship between the remaining edge objects in the edge object set and the edge object" explains: In the distance matrix, rows and columns represent sequential local edge objects, and the corresponding position represents the distance value of two local edge objects.

We set an "edge distance threshold T." Positions with distances greater than T are set to 0, indicating no interference, while positions with distances less than T are set to 1, indicating mutual interference. To delete the relationship between two local edge objects is to set the corresponding position of the distance matrix to 0. You can find the modifications in L372-L376 of the revised manuscript.

Comments 23: 403-404: What are "partially matched edge contours?" The purpose of this paragraph is unclear without that definition.

Response 23: Thanks for your correction. Based on your suggestion, we are more specific about "partially matched edge contours". The first is to use the vertices of the enclosing rectangular box to calculate two edge contour features that are relatively close overall. Secondly, there are also data at the intersection positions of some local edge contours that match, so this section is about calculating whether the two local edge contour features match using the data at the intersection positions. We have restated this passage. You can find the revised content in L423-L425 of the revised manuscript.

Comments 24: 413: What does “their” refer to? Edge contours?

Response 24: Thank you for your opinion. We describe this referent in more detail. The phrase "their" has been replaced with "local edge contour features." You can find the revised content in L432-L433 of the revised manuscript.

Comments 25: 414-419: This paragraph does not provide enough details for this step. What are the "structure type" classes used? There is another cut off sentence on line 417. What is done with the results of the "same structure type" determination and "similarity of the two edges" in order to affect the contour edge matching between image pairs?

Response 25: Thank you for your suggestion. Based on your suggestion, we added a specific description of the other two factors for local edge contour feature matching. This section is to judge the similarity of local edge contour features through the shape of the local edge contour and the slope of the connecting line, and observe whether these two conditions are met. If it is not satisfied, it is considered a mismatch. In addressing the punctuation within the text, we emphasize the modification of sentence structures at the paragraph level. You can find the modifications in L433-L448 of the revised manuscript.

Comments 26: It would be great to see more image examples with curved contours. The top panel of Figure 7 does have a few curves, but perhaps the other panels could be replaced with examples that show curved contours rather than all straight lines. Otherwise the improvement over Jia et al. is not obvious.

Response 26: Thank you for your suggestion. Based on your suggestions, we have provided additional examples of images with curved outlines.You can see Figure 8 in the revised manuscript.

Comments 27: Section 3.4 has been cut down dramatically but there is still no mention of Multi-GS. The previous version of the text appeared to be a rough explanation of Multi-GS, but most of that text has been removed in this version. Is this pre-alignment step directly using the Multi-GS method, or something different but related that the authors derived themselves? If it’s just Multi-GS then say so, along with a clearer explanation of how the fitting problem is posed in terms of the point-line and edge contour features. If, however, this is a new method then it requires a much more thorough explanation. Furthermore, the authors should explain why they are not just using Multi-GS which seems well-suited to this exact problem.

Response 27: Thank you for your suggestion. First, in this section, the Multi-GS method is used directly. The function of Multi-GS is to filter mismatches on the extracted feature point pairs to obtain

filtered interior points. Compared with the RANSAC algorithm, Multi-GS can quickly generate hypotheses and efficiently process complex data. Based on your suggestion, a clear term, "Multi-GS", has been added to the text, along with literature citations. The process of selecting the best hypothesized model is described in detail in the paper. Content explanation: For each hypothetical model, the distance value of the local edge contour feature matching pair is calculated through fitting. This value also represents the alignment effect of local edge contour feature pairs. The distance sums of all local edge contour feature matching pairs are arranged in descending order to obtain the overall alignment result value. At the same time, the root mean square error of the distance values of all local edge contour feature matching pairs under a single hypothesis model is calculated, and the results are sorted. The best hypothesis model is obtained through the above two sorted lists. You can find the revised content in L454 and L470-L473 of the revised manuscript.

Comments 28: Having read Sections 3.2 and 3.3 in more detail, I have found five different threshold parameters mentioned: contour distance (263), spindle length (287), edge density (296), polynomial fit (328), and edge close distance threshold T (364). However line 529 only mentions two threshold parameters when describing the experimental setup. What about the others?

Response 28: Thank you very much for your opinion. Based on your suggestions, we have added the threshold parameters used. Details as follows: the point distance threshold is 3, the principal axis length threshold is 50, the edge similarity threshold is 15, the fitting threshold is 4, and the minimum edge distance threshold is 80. You can find the modifications in L562-L564 of the revised manuscript.

Comments 29: The data shown in Tables 2 and 3 and Figure 9 do not support the claim that the contour constraints provide a higher quality image alignment. The constraints=Yes values are not even consistently lower (better) than the constraints=No values.

Response 29: Thank you very much for your advice. The previous Figure 9 has changed to Figure 10. In response to your question, after discussion, we believe: What this comparison experiment highlights is the alignment effect of local edge contour features under feature constraints. In addition to the alignment effect of point features and line features, the alignment strength of the image overlapping areas is also reflected through the overall alignment of local edge contour features. Although the calculated values in Table 2, Table 3 and Figure 10 with constraints are not always lower than those calculated in the unconstrained case, overall, the fitting results of the local edge contour features under the constrained case are better than those without the constrained case. In the case of local edge contour feature constraints, the local edge feature alignment effect is better. Based on your suggestions, we will consider the experimental results more comprehensively later. Thank you again for your suggestions.

Comments 30: Why are the SSIM results only reported on a subset of the datasets that RMSE was reported on?

Response 30: Thanks for pointing out the shortcomings. According to your suggestion, we used the same dataset as RMSE to calculate the SSIM value. Thank you again for your suggestion.

Comments 31: Figures 12 & 13: Yellow and pink lines and points are misaligned. It looks like the yellow points are on the pink line, and the yellow line has no points. Yellow-on-white is very hard to see and a darker color should be used, anyway.

Response 31: Thanks for your correction. We've redrawn the picture. For clearer display, the yellow color has been removed and the value of Ours is represented in red instead.

Comments 32: The English was understandable overall, but Sections 3.2-3.4 need more attention.

The text has frequent passages written in the imperative ("Merge the data...") whereas most of the text uses the preferred passive voice ("The data were merged..."). All instances of imperative voice should be corrected.

Response 32: Thank you for your suggestion. Based on your suggestion, we carefully checked the paragraphs where imperative sentences appeared and changed the imperative voice to passive voice.