Answers to the comments from referee #2

We would like to sincerely thank the reviewer for their valuable suggestions and comments. In the following sections we have tried to address all of them and have made appropriate changes to the manuscript itself.

Responses to the comments made by the reviewer are prefixed with the term **Authors** in a new line. References to specific lines are with regards to the unrevised version of the manuscript.

Line 90: "This reference is incomplete."

Authors: We acknowledge the remark of the reviewer and have added three more papers of recent date which use IMU/gyro/control action data for the stabilisation of UAV video, or UAV itself:

1. *Auysakul et al., 2018* – which describes a hybrid KLT + IMU data approach,
2. *Hanning et al., 2011* – which describes approach for IMU data for stabilisation of videos with rolling shutter effects, and
3. *Stegagno et al., 2014* – which (while not aimed specifically at videos) describes a method of UAV stabilisation using IMU, visual, and control action information.

Line 215: "By "structural" are you referring to shape, size, and orientation of features? Please clarify."

Authors: We agree with the reviewer that this term could be explained in more depth. When comparing two images (of the same size) shape, size and orientation will all have impact on the SSIM score. However, other effects such as image noise will have far lower impact on the SSIM score than with metrics such as MSE. We have added a brief explanation to the existing sentence: "… *as it implicitly relies on* <u>*the information on shape, size and orientation of features – structural information.*</u>"

Figure 3: "Scaling, not scalling."

Authors: We thank the reviewer for pointing out this mistake. Text in the figure was changed accordingly.

Tables 2-4: "It would be helpful to display these vectors on a background image to help orient the reader."

Authors: Presented vectors could indeed benefit from background image, and this is something that we discussed during manuscript preparation. However, readability issues quickly arose where vectors of lower intensities could not be perceived correctly with their directions. For those reasons we opted to omit the background. However, for the revised version we have prepared supplementary material – unaggregated velocity fields, which would allow for deeper insights into the results of our work.

Equation 3: "What about an intercept term of the form + c at the end of this equation?"

Authors: The intercept term in the polynomial expression was deliberately omitted by the authors in order to retain as much physical meaning as possible. Keeping the intercept term would indicate non-zero displacement even for RMSD=0. Furthermore, the difference in R^2 values between the two relationships (with and without the intercept term) have been found to be negligent (up to order of 0.001). Intercept term values have been found between 0.1 and 0.2 for all verification points. However, the reviewer's point is valid, and it is worth explaining in the text why the intercept term was not used. We have added the following sentence to the paragraph after Fig. 10: "*The intercept term in the polynomial relationship in Eq. (3) was intentionally omitted in order to retain some physical meaning in the relationship and prevent Eq. (3) from indicating small displacements even when RMSD=0.*"

Figures 11-14: "Are the distributions summarized in these boxed plots over the frames in the sequence (i.e., over time)? Please clarify."

Authors: We thank the reviewer for pointing out this issue. Indeed, a small clarification may be required regarding these results. Boxplots presented in Figs. 11-14 contain information about all estimated displacements of verification points, either relative to their position in the first frame of the sequence (Figs. 11 and 13), or relative to their position in the previous frame (Figs. 12 and 14). Total number of points summarised by one column in the boxplot is equal to the number of frames in the sequence minus one. Lines 557 and 579 (in the previous version of the manuscript) have now been changed to read:

> "*Figures 11 and 13 <u>summarize all displacements of validation points across the entire frame sequence</u> relative to their location in the initial frame ...*"

> "*Figures 12 and 14 <u>summarise all displacements of validation points across the frame sequence</u> relative to their position in the previous frame*."

We hope that this small correction will clarify the origin of the values presented in Figs. 11-14.

Line 652: "Missing a number here."

Authors: The value in question was indeed accidentally omitted. This is rectified in the revised version of the manuscript.

Line 758: "Should this be table 6? I don't see a table 7."

Authors: We thank the reviewer for pointing this out. The text should indeed point towards the Table 6, and this was rectified in the revised version of the manuscript.

Other minor linguistic corrections have been made as per reviewer's suggestions.