Effectiveness of Multi-bispectral Passive Infrared Range Finding

James Utley
utleyj@bu.edu

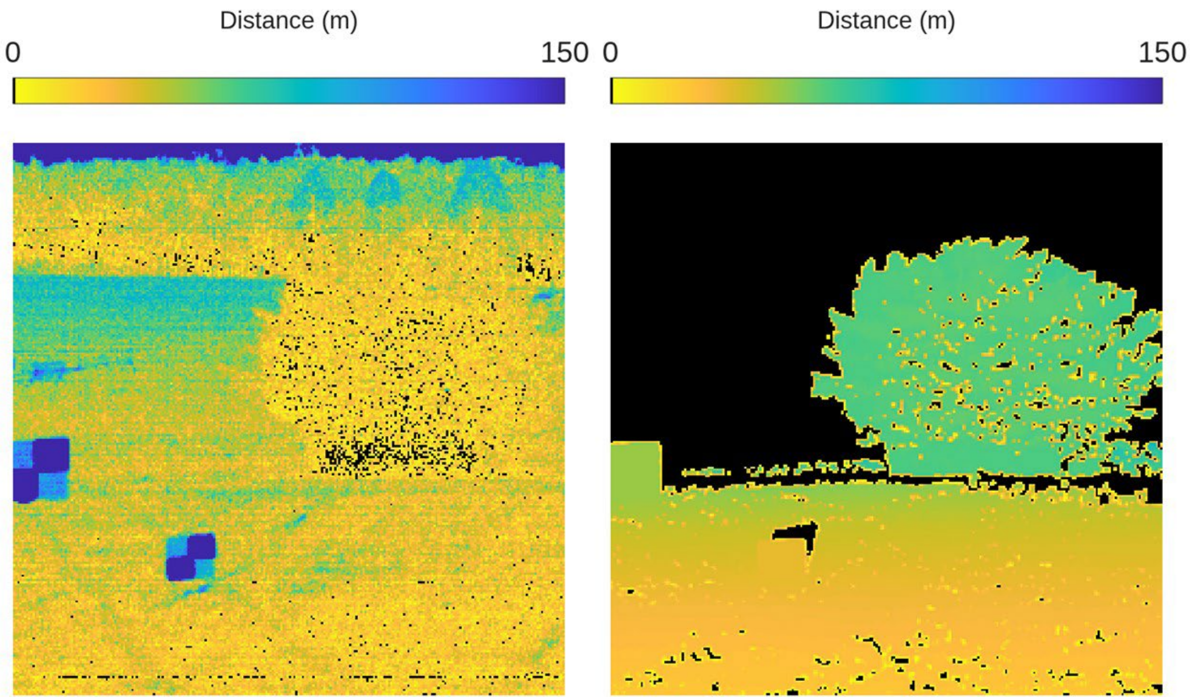

Figure 1 - Single Bispectral Estimation Pair Index 7 (left) compared to LiDAR ground truth (right)

1. Introduction

Infrared imaging and ranging is an established field of research with many military and civilian applications. Currently, passive absorption-based ranging is used for settings with high temperature contrast, while settings with low temperature contrasts are largely ignored. This is due to the fact that key assumptions used when ranging high temperature difference objects are not applicable for those with small temperature differences. The potential uses for passive ranging similar temperature objects are many, namely ones where conserving electricity is important. But the effectiveness and reliability of this technique in real-time systems has not yet been shown.

Recent work has investigated the usefulness and reliability of using passive range finding of similar temperature objects using two methods: bispectral and hyperspectral range finding. Hyperspectral imaging produced good but slow estimates and bispectral imaging produced fast but inaccurate estimates. Therefore, this research did not show that passive infrared ranging was effective for use in real-time systems.

Following in the footsteps of this research, this project investigates three methods of combining multiple bispectral estimations: Type A combinations (pixel-by-pixel mean), Type B combinations (pixel-by-pixel median), and Type C combinations (pixel-by-pixel error minimization). The goal of this project is to show the potential of using combinations of only bispectral estimations to give a reliable estimate.

2. Related Work

This project is based on two papers, sources [1] and [2], which investigate aspects of passive infrared range finding for similar temperature objects.

The greatest intellectual inspiration for this project comes from source [2] which investigated the use of bispectral versus hyperspectral readings for passive range finding. Using a high-powered HgMeTe infrared sensor, Gallastegi et al. ultimately found that hyperspectral readings were more accurate, but their compute time makes the method intractable for realtime systems. Furthermore, the team found that bispectral estimation had a lower compute time but returned a low-confidence distance estimation. With these results, the team discontinued further stages of this project, since its deployment in real-time autonomous vehicles and systems seemed unlikely.

Despite this setback, source [1] returns to a similar question of passive ranging, this time utilizing quadspectral analysis. Primarily using earlier data from [2], the team sought to further mitigate effects of atmospheric downwelling in an attempt to increase the reliability of the distance estimation while reducing compute time. Unfortunately, like the earlier paper, this research did not reach the conclusion that the deployment of passive LWIR in real time systems would be possible.

What this project seeks to do is reanalyze the data used by Gallastegi et al. in both sources [1] and [2] to investigate the validity of using multiple combinations of bispectral estimations to reach a meaningful conclusion with low compute time.

3. Method

3.1. Dataset

The data for this project comes from one scene imaged at 256x256 pixels. The ground truth image referred to throughout comes from LiDAR data of the same scene, showing midground and foreground distances, with 23550 pixels out of a possible 65536 pixels containing ground truth information. The remaining pixels are blacked out and represent NaN values.

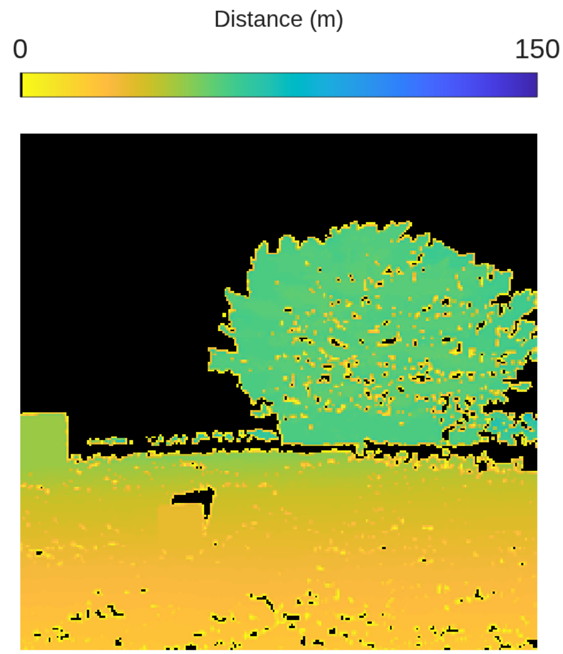

Figure 2 – LiDAR Ground Truth

This data is the same data used in sources [1] and [2] by Gallastegi et al. and comes to this project by Wentao Shangguan and Dr. Vivek Goyal who are members of the research team responsible for sources [1] and [2]. A special thanks to them for providing this data.

3.2. Root Mean Square Error (RMSE)

This project uses Root Mean Square Error (RMSE) to assess the correctness of the estimations. A mask is applied over all NaN values in the given estimation combination and in the LiDAR ground truth, such that only pixels with valid values in both the given estimation and the ground truth are compared. This choice is to ensure that NaN values are not compared to real values and keeps errors from producing in MATLAB.

A pixel-by-pixel difference is then taken between the estimation and the ground truth image. This difference is then elementwise squared, summed, and a mean operator is applied to collapse the array to a scalar. The square root is then taken of this scalar which gives the RMSE value for the given estimation. Pseudocode for this process is provided below:

```
Combo_givn=all_combos(:,:,i);
valid_mask = ~isnan(combo_givn) & ~isnan(lidar);

difference = combo_givn(valid_mask)-
    lidar(valid_mask);

rmse=sqrt(mean(difference.^2));
```

Furthermore, this project utilizes thirteen pairs of wavelengths defined by Gallastegi et al. to produce thirteen distinct bispectral estimations. From these thirteen pairs, this project enumerated the 8178 possible combinations ($2^{13} - 13 - 1$) in MATLAB. Each combination contains two to thirteen bispectral estimations. The combination methods are defined in 3.3 and 3.4. The thirteen single bispectral estimations are used as control values to compare the performance of the combinations, and appear throughout the project as 'control' values.

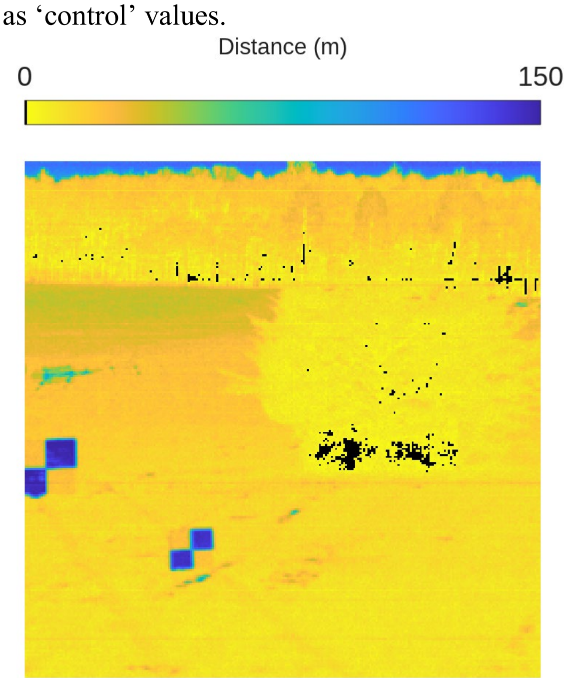

Figure 3 - Lowest RMSE Bispectral Estimation, Pair Index 3.

3.3. Type A Combinations - Mean

Type A combinations are combinations of estimations using an elementwise mean operator. For a given combination of the original thirteen bispectral estimations, a Type A combination adds all of the 256x256 matrices together and then it performs elementwise division by the number of original estimations summed.

The mean operator was chosen primarily for its simplicity: it is trivial to find the mean arrays of the same size in a vectorized coding language like MATLAB. But after initial data returned curious results, it was determined that the mean operator was not itself enough to prove the effectiveness of the combination, and therefore another type of combination was required.

3.4. Type B Combinations – Median

Type B combinations are combinations of estimations using an elementwise median operator. For a given combination of the original thirteen bispectral estimations, a Type B combination considers a single pixel from all of the estimations being combined and finds the median value of each pixel. (For combinations of size two, a Type B combination and a Type A combination produce the same output.)

The median operator was selected because it is less sensitive to outliers skewing data than the mean operator. Therefore, it was hypothesized that the median operator would provide cleaner results that were more accurate and thereby have lower RMSE values than the mean operator.

Other methods of combination were considered, namely weighted least squares (WLS) and generalized least squares (GLS). This project seeks to find an estimation simpler than hyperspectral estimation, which, as used in sources [1] and [2], resembles a regularized least squares formulation. Therefore, to avoid similar faults with computational time, WLS and GLS were not pursued in favor of median combinations.

3.5. Type C Combination – Minimization

Type C Combinations are elementwise pixel-by-pixel combinations using the minimization operator and absolute error. Each pixel of the thirteen original bispectral estimations is measured against the pixel of the ground truth image using absolute error, ignoring NaN values. This absolute error calculation is the following: $|\text{estimate} - \text{truth}|$. Absolute error was chosen because the direction of the difference (positive or negative) did not matter, prioritizing magnitude over direction.

After the combinations were enumerated, each final pixel was determined by selecting the original pixel with the minimum absolute error.

This project hypothesizes that Type C combinations will achieve lower RMSE but higher computational complexity than Type A or Type B.

4. Description of the Theory

Before we discuss the methods used for combining the bispectral measurements, it is important first to understand how bispectral estimations are calculated. To explain bispectral estimations, this project will be borrowing much of the explanation of the forward model described in source [2], which is the same forward model this project uses.

Put simply, bispectral estimations can be solved by using measurements from two different wavelengths of light (in this case infrared light) to find the radiance of the emission, and then using the Beer-Lambert Law to estimate the distance from the emissivity.

The first step is to find the radiance of the emission. For this we assume that each object observed emits thermal radiation and travels through a standard Earth atmospheric composition. To calculate the radiance, take two measurements at two close wavelengths. It is important that the two wavelengths selected are close in wavelength to one another. The reason for this comes from the Fisher Information Analysis explained in source [2]. In short, signals from radiating objects attenuate to different degrees the further they are from the sensor. This attenuation is the reason one can use infrared information to estimate the distance to an observed object: because one knows how the signal should attenuate by the time it reaches the sensor for different distances, then one can use the observed attenuation to assess the distance. From source [2]'s derivation of the Fisher Information Analysis, we know that the ideal attenuation to measure this distance is approximately 4.3 dB on a logarithmic scale or $\frac{1}{e}$ on a linear scale. Therefore, the ideal wavelength λ is one that satisfies $\alpha(\lambda) = \alpha^*(d) = \frac{4.3}{d} = \frac{1}{(d)(e)}$. Furthermore, the smoothness assumption for bispectral estimation assumes that the emissivity of one wavelength multiplied by the blackbody radiation of the atmosphere for that wavelength is roughly equal to the emissivity of a second wavelength times the blackbody radiation of the atmosphere for that wavelength.

$$\epsilon(\lambda_1)B(\lambda_1; T_{air}) \approx \epsilon(\lambda_2)B(\lambda_2; T_{air})$$

This assumption is critical to bispectral estimation, and it only holds when the wavelengths are close to one another. The further the wavelengths are from one another, the more this smoothness approximation fails and the less accurate the estimation becomes. Therefore, particular wavelengths carry more information about particular distances than others and two wavelengths must be selected that are close to one another to allow for the smoothness assumption to hold.

Next, from each measurement, subtract the effects of downwelling (atmospheric blackbody radiation) for the known air temperature to get a corrected reading for the object's radiance alone. It is important to note that the solving of air temperature can be done either through a dedicated process described in source

[2], or, as here, through the readings from a local weatherstation and can therefore be treated as a given for the purposes of this project. It also bears note that while reflected downwelling can lead to systematic biases in measurements, the derivation above and in sources [1] and [2] assumes negligible reflective downwelling. To get the ratio of downwelling-corrected radiances for both wavelengths, L, perform elementwise division between the result of the first measurement and the second measurement. Using MATLAB style vectorized pseudocode, this process looks like this:

$$\mathrm{L} = (\mathrm{m}_1 - \mathrm{B}(\lambda_1, \mathrm{T}_{\mathrm{air}}))./(\mathrm{m}_2 - \mathrm{B}(\lambda_2, \mathrm{T}_{\mathrm{air}}))$$

With L calculated, one can rearrange the Beer-Lambert Law to get an estimate for the distance. The Beer Lambert Law can be written as:

$$\mathrm{A} = \epsilon \mathrm{d} \mathrm{C}$$

Where d is the distance, C is the concentration, and ϵ is the molar absorptivity. We can therefore rearrange the Beer-Lambert Law to estimate distance for a known absorbance, concentration, and molar absorptivity to be:

$$\mathrm{d} = \frac{\mathrm{A}}{\epsilon \mathrm{C}}$$

Since we know the concentration and molar absorptivity of the atmosphere, we can use substitutions found in source [2] to rewrite the Beer-Lambert Law to be the following:

$$\hat{\mathrm{d}} = -10 \log_{10} 10(\mathrm{L})./(\mathrm{att}_1 - \mathrm{att}_2)$$

Which represents the logarithm result of the emissivity elementwise divided by the attenuation of the signal over the two wavelengths.

This project uses thirteen predetermined pairs of wavelengths determined from water vapor absorption bands, as shown in Figure 4.

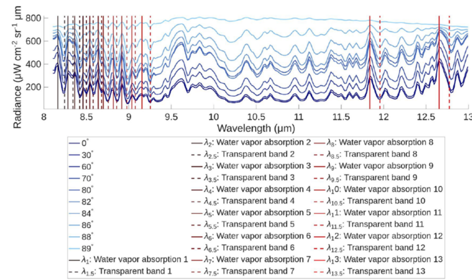

Figure 4 – 13 Predetermined Pairs for Bispectral Estimation

Now that we have an understanding of bispectral estimation, let us discuss the methods for combining these wavelength pairs. While Section 3 of this project discusses the methods of combining the bispectral estimations, the precise mathematical methods of combination are outlined here.

5. Results

5.1. Type A

Type A combinations produced 2016 estimations with an RMSE value lower than the control value. Figure 5 shows the top 49 lowest RMSE values for Type A combinations. Notice how most of the lowest RMSE values belong to combinations of length 6 and 7. This result is not ideal, since larger combination lengths should correlate with a lower RMSE value. However, as is seen in Figure 6, this is not the case.

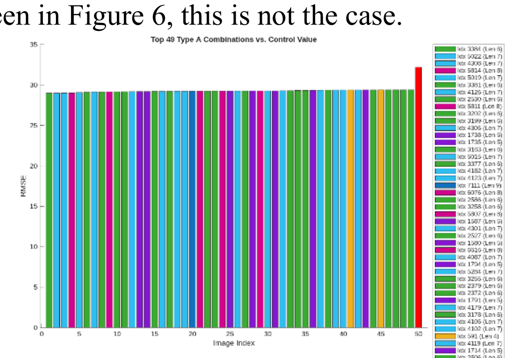

Figure 5 – Type A Lowest 49 RMSE values, Lowest on the Left

Figure 6 shows the lowest RMSE values by combination length for Type A combinations. There is a slight decrease in the lowest RMSE value as combination lengths increase from 2 to 6, and then begin to increase from combinations of length 8 to 13. As in Figure 5, Figure 6 shows combinations of length 6 and 7 to have the lowest RMSE score, which is not the ideal outcome. Between these two figures, a relationship is evident between longer combinations and higher RMSE scores.

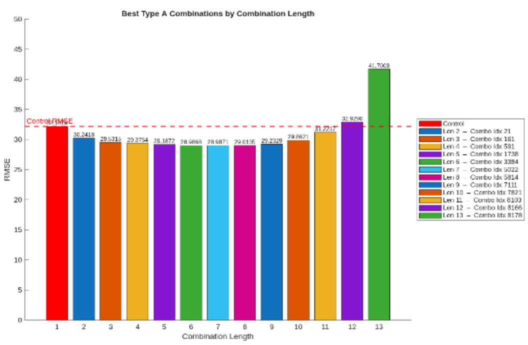

Figure 6 – Type A Lowest RMSE by Combination Length

This relationship is underscored in Figure 7 which shows the average RMSE value for Type A combinations by combination length. As the combination length increases beyond 5, the average RMSE value increases. This result implies that Type A combinations are susceptible to outliers, an expected result from prior knowledge of the mean operator as outlined in Section 3.3.

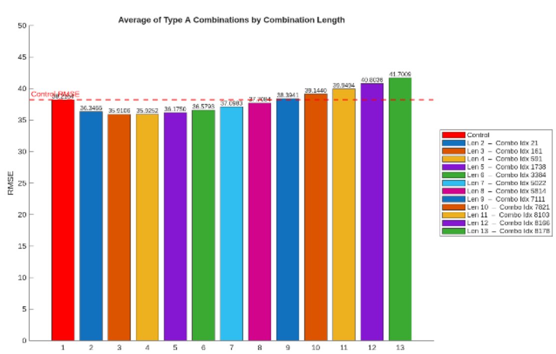

Figure 7 – Type A Average RMSE by Combination Length

Finally, Figure 8 shows the distribution of data points by combination length, as well as the mean average (the black line) and the minimum (the red line) of each combination length. Note the smoothness of the distribution, how there are few large noticeable gaps in the data, which is different from Figure 12 below.

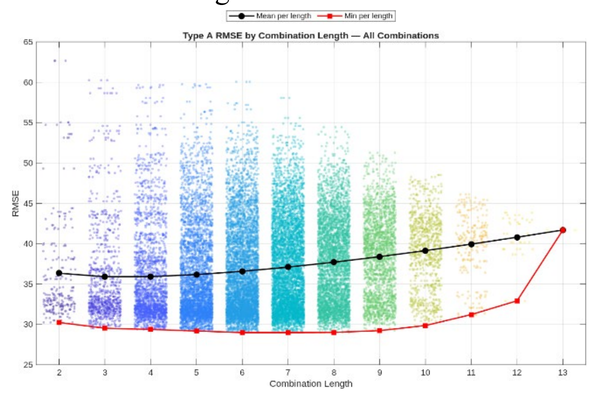

Figure 8 - Type A Jittered Scatter Plot

5.2. Type B

Type B combinations produced 4225 combinations with an RMSE value lower than the control value. This is more than double the viable estimations produced by Type A combinations. The increase in viable combinations shows that Type B combinations are better for combining bispectral estimations since they produce more results than Type A combinations. This assertion is furthered by Figure 9 which shows the 49 lowest RMSE scores of Type B combinations.

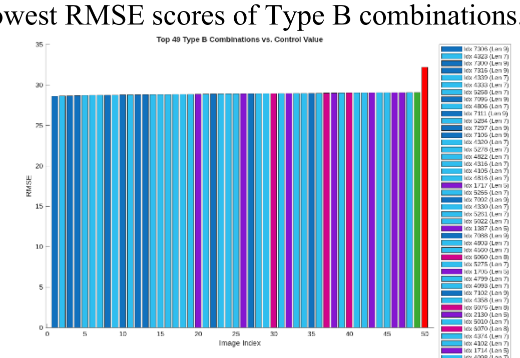

Figure 9 – Type B Lowest 49 RMSE Combinations, lowest on the left

The lowest RMSE scores shown in Figure 9 come from combination sizes 8 to 11, an improvement from Type A combinations.

However, the combination of length 13 is not the lowest RMSE score. Further analysis on the distribution, shown in Figure 12, shows that there exist more combinations for medium sized lengths (i.e. lengths of 5 to 8) than exist at the largest combination sizes (i.e. lengths of 11 to 13). This sparsity of information in the distribution is expected by the nature of the n-choose-k operator used to enumerate all of the possible combinations.

The median operator was chosen for Type B combinations because it is less sensitive to outliers than Type A. The increase in accuracy gained from less sensitivity to outliers can be seen in Figure 10 and Figure 11. In Figure 10, the lowest RMSE scores trend downwards from combinations of length 3 to 11. Similarly in Figure 11, the average RMSE value by combination length decreases until combinations of 9 or more where they increase. This behavior is improved from their respective Type A metrics.

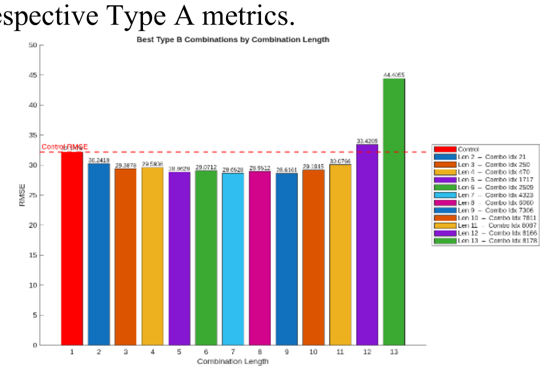

Figure 10 -Type B Lowest RMSE by Combination Length

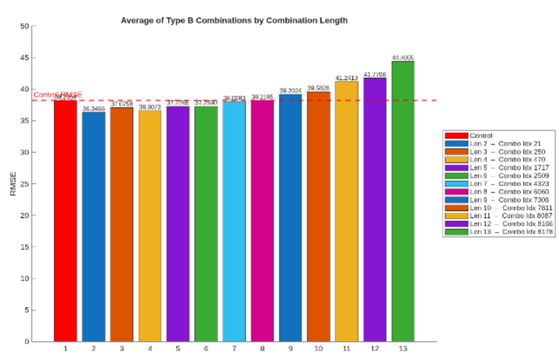

Figure 11 – Type B Average RMSE by Combination Length

However, like Type A combinations, Type B combination RMSE values are not monotonically decreasing, with increasing RMSE values at higher combination sizes. This result was not expected and required additional analysis to understand why it was happening.

Furthermore, a key difference between Figure 12 and Figure 8 is the discontinuities in the Figure 12 (i.e. the blank spaces in each bar). This is because the median operator used in Type B combinations calculates based on position, not on value, so the introduction of an estimator can skew the combination to create the blank spaces seen in Figure 12.

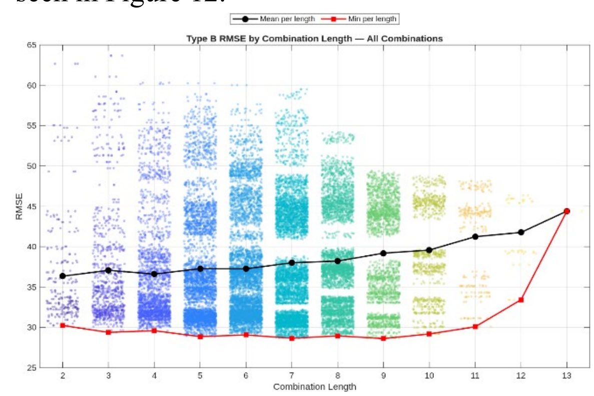

Figure 12 - Type B Jittered Scatter Plot

5.2.1. Top-k estimator analysis

In the development of the project, the question of if Type B combinations would be improved by more careful selection and combination of the original bispectral estimations. To answer this, this project uses a top-k analysis of a particular combination length to assess performance and whether outlier data influenced Type B combinations in a significant way.

For the top-k analysis, this project selected five subsets of the original thirteen, ranging from the top seven to the top eleven, each comparing metrics for combinations of length five. Figure 13 shows all thirteen combined on the left, and then proceeding left to right with the top seven to the top eleven.

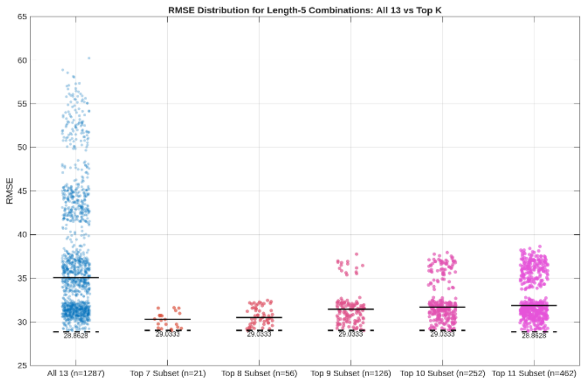

Figure 13 - Top 7-11 Original Estimators vs. All 13 Estimators for Length-5 Combinations

There is little change in the minimum RMSE measured across all six top-k (shown with a dotted line), but there is a great change to the average value recorded value (shown with a full line). From Figure 13, it can be seen that outliers from some original estimators skew the averages of estimation combinations, but the minimum is largely unaffected by outliers.

5.3. Type C Combination – An Upper Bound

Type C combinations produce 8151 combinations with an RMSE lower than the control value. This improvement from Type A and Type B combinations is reflected in Figure 14 and Figure 15 which both show a monotonic decrease in RMSE values as combination sizes grow. This monotonic decrease is further supported in the jittered scatter plot Figure 16 which shows that the distributions also generally decrease with combination length. Furthermore, Figure 17 shows that the lowest RMSE values are sizes 9 to 13, unlike Types A and B which showed best RMSE values around size 7 to 9.

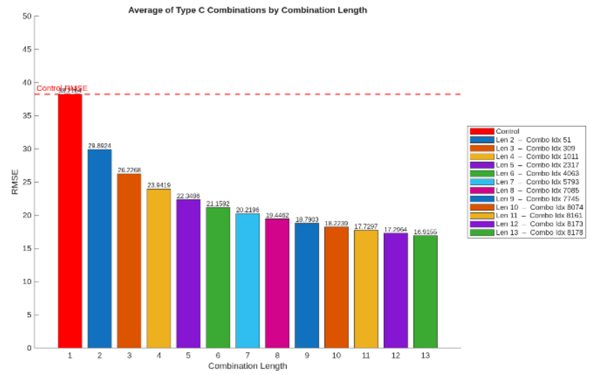

Figure 14 - Type C Average RMSE by Combination Length

By RMSE, Type C performs the best of the three combination methods assessed in this project. As the combination length gets longer, the RMSE gets smaller, without the increase at larger combination sizes seen in Types A and B.

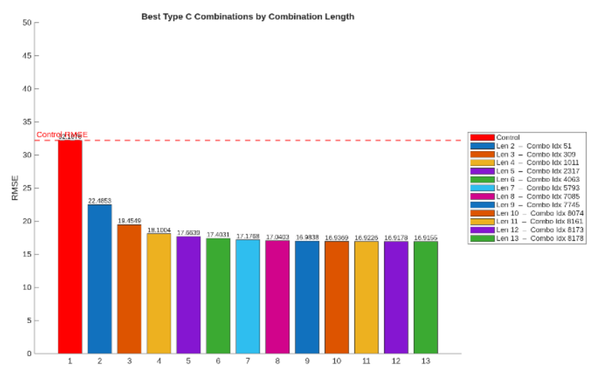

Figure 15 - Type C Lowest RMSE by Combination Length

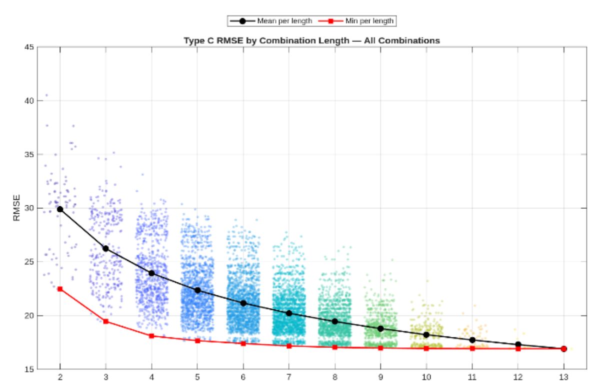

Figure 16 - Type C Distribution; Jittered Scatter Plot

The difference in the performance of Type C versus Type A or B is not surprising when considering the types of operators each type uses. Type A and B operators rely on finding averages among a set of values. Type C, however, uses the minimization operator to minimize the error between individual pixels and the ground truth image. This also minimizes the overall error of the estimation and drives the RMSE down.

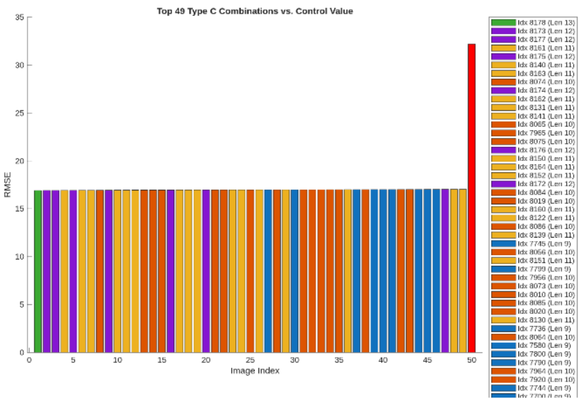

Figure 17 - Type C Lowest 49 RMSE Values, Lowest on the Left

6. Conclusion

Of the three combination types, Type C combinations performed the best overall, achieving both lower average and minimum RMSE by combination length than both Type A and Type B. Type B also performs well at finding a combination that provides a lower minimum RMSE than in Type A, but it does so through many calculations that would slow overall analysis in a real-time system.

This project has shown that bispectral estimations can be combined to give estimations that are closer to the ground truth than by the single bispectral estimations alone. To do this effectively, however, the process by finding the lowest estimator needs optimization. Figure 13 emphasizes this problem by showing that the number of combinations to find the minimum in a length 5 combination. These calculations may be able to be optimized by deep learning networks but would require more dedicated research into optimizing the process for real-time deployment.

Initially, this project sought to find differences between Type A and Type B combinations of bispectral combinations. The idea of minimization and introducing Type C combinations did not happen until late in the project's development. Because of this and for the sake of time, several avenues of analysis were not pursued, namely computational time analysis for the three types of combinations and code optimization. Future directions for this project include optimizing the computational time for the combination operations, and finding the amount of estimations needed to have a high probability of finding a good minimum RMSE.

7. Bibliography

1. Dorken Gallastegi, U., Shangguan, W., Choudhary, V., Agarwal, A., Rueda-Chacón, H., Stevens, M. J., & Goyal, V. K. (2026). Ozone cues mitigate reflected downwelling radiance in LWIR absorption-based ranging. arXiv:2603.00273. https://arxiv.org/abs/2603.00273
2. Dorken Gallastegi, U., Rueda-Chacon, H., Stevens, M. J., & Goyal, V. K. (2023). Absorption-based, passive range imaging from hyperspectral thermal measurements. arXiv:2308.05818. https://arxiv.org/abs/2308.05818
3. Hawks MR, Vincent RA, Martin J, Perram GP. Short-range demonstrations of monocular passive ranging using O2 (X3Σg- → b1Σg+) absorption spectra. Appl Spectrosc. 2013 May;67(5):513-9. doi: 10.1366/12-06673. PMID: 23643040.
4. Li JH, Wang ZB, Wang Zhi. [Passive ranging of infrared target using oxygen A-band and Elsasser model]. Guang Pu Xue Yu Guang Pu Fen Xi. 2014 Sep;34(9):2582-6. Chinese. PMID: 25532368.
5. Yu H, Liu B, Yan Z, Zhang Y. Passive ranging using a filter-based non-imaging

method based on oxygen absorption. Appl Opt. 2017 Oct 1;56(28):7803-7807. doi: 10.1364/AO.56.007803. PMID: 29047763.

Table of Figures