A Variable Level IR/RGB Fusion Methodology

Greg Search
Stanford University
EE367

Abstract

The goal of infrared (IR) and red, green, blue (RGB) image fusion is to provide the viewer with more information than each component can provide alone. In this paper, we assess a collection of fusion techniques that span from low to high complexity. We present a new fusion method (Level 3) that can drastically decrease computational cost without much degradation in picture quality. Level 1 fusion uses a simple average between the IR and RGB image. In Level 2, we convert an RGB image to YCbCr space and replace the luminance with the IR intensity. Levels 3-5 create a laplacian pyramid (LAP) to divide the image into a base layer and detailed layers. Level 3 combines base layers using a simplified version of the visual saliency map (VSM). It then combines detail layers using the Max Absolute Rule (MAR). In Level 4, we implement a full VSM to fuse the base layers. Finally, in level 5 we use VSM on base layers and apply a weighted least squares (WLS) on the fusion layers. The intent of the WLS is to obtain salient color features lost in MAR. Experimental results demonstrate that our new method (Level 3) can decrease computational cost dramatically while retaining a high level of quality. We recommend Level 3 for readers who desire near state-of-the-art performance but would like to avoid heavy computation. Applications of this method include detecting pedestrians from vehicles, determining crop health and facial recognition.

1. Introduction

Image fusion provides more information than can be obtained through the sum of each individual image. Common examples include: depth mapping, auto-focus, and multi-spectral imaging. IR/RGB fusion is superior to other fusion methods for two main reasons. First, IR and RGB capture completely different modalities and thus can provide a greater variety of information to the viewer. For example, if the subject is obscured by fog, shadows or darkness, RGB is no longer a viable option. [12] Second, IR and RGB are inherent in all materials, as opposed to MRI which requires magnetic materials.[3]

1.1. IR vs RGB Image Capture

The difference between these two types of capture methods comes from the wavelength of light that is captured in each case. Instead of the 400-700nm wavelength light captured by RGB, IR cameras are sensitive to light between approximately 10,000nm and 14,000nm. IR sensors often only collect intensity of IR in one wavelength, which leads to a monochrome image. Yet they can be colored using a technique called density slicing which samples the IR frequency at multiple different ranges of wavelength that correspond to colors.[8] Infrared sensors, and thus IR cameras, are often much more expensive than their RGB counterparts in large part because there is less energy contained in IR waves than visible light rays.

The benefits of RGB are well known. In the context of fusion, features can often be easier to detect at certain wavelengths of light. The human eye is most sensitive to colors in the middle of the visible light spectrum (green). However, it can be shown that increasing the level of red can allow edges to be detected more effectively [5].

1.2. IR and RGB Fusion

There are many different algorithms for image fusion: PCA based image fusion, Wavelet transform image fusion, and Pair-wise spatial frequency matching. [3] In all cases, the goal is to identify the most salient features, delete unnecessary or redundant data and fuse the result. This may include capturing certain spatial frequencies, gradients or intensities. RGB fusion with IR adds extra considerations since the images are captured in very different ways and provide a much different set of benefits to the viewer.

The first obvious difference between IR and RGB image processing is that IR is only grayscale (in most cases). Thus, it has $\frac{1}{3}$ the amount of storage and processing needs as the RGB image for rudimentary operations. However, there are many methods to reduce the computational cost of RGB images and video. Our eyes are much more sensitive to an undersampled image intensity than an undersampled image color. As a result, there are many compression algorithms based on converting RGB images to either luminance, chrominance (YCbCr) space or hue, saturation value

(HSV) space. [1]. The image can be compressed by undersampling in the color space.

In the last decades, multi-scale decomposition has become very popular in the fusion community. [7] The idea is to break the image down into separate components that can be processed in different ways (different sampling rates, different filter threshold, etc.) and then recombined. Perhaps one of the most popular forms of this is the creation of the laplacian pyramid (LAP). This is a method that applies an increasing gaussian blur on a picture. At each interval, it collects image and thus creates a ”stack” that captures the image at different frequency levels.

Figure 1. Laplace Pyramid (LAP)

A method of detecting the ”important” aspects of an image is by constructing a visual saliency map (VSM). This compares the image pixel by pixel in order to detect pixels that have the greatest difference overall. Note that VSM are only recommended for images at low spatial frequencies. If used on high frequency images, a significant amount of noise can be introduced to the fused image. For high frequencies, the Max Absolute Rule (MAR) is commonly used. This is a step function that measures whether the IR or RGB image has higher intensity. Normally a smoothing function is applied to the MAR output to avoid aliasing. [7]

2. Related Work

The two main objectives of this paper are to decrease computation (in low level fusion) and increase fidelity (in high level fusion). This section presents fusion algorithms that attempt to achieve both objectives.

2.1. Low Cost Fusion

The first and simplest fusion method is a simple averaging of the IR and RGB image. This surprisingly can be relatively successful despite the extremely low computational cost. Zhou et al describes a simple weighting algorithm based on the intensity of the pixel [11]. For each pixel, the IR weight is determined by its intensity, which demonstrates the amount of information it’s adding to the pixel. The same follows for the RGB weight. This method cannot be used with an image that provides glare.

Averaging can also be done by images that are not exactly aligned through a process called image registration. As B. Zitova explains, there are many registration methods that can be accomplished without a need for high computation. [13] In addition to providing low cost fusion, registration allows for the processing of images that aren’t aligned and thus enables dynamic image capture.

There are methods that utilize the color basis in order to decrease computational cost. One paper uses the color transfer technique based on the linear YCBCR space. The method directly uses the grayscale fused image and the difference signals of the input images to construct the source YCBCR components, then uses the statistical color transfer technique to form a color fused image that takes the target image’s color characteristics. Two different strategies, are designed to fulfill different user needs. [9]

2.2. High Fidelity Fusion

There are situations where the goal of the processing is to create the highest fidelity image possible. A very common way to achieve this is through the use of a neural network. There are countless algorithms. The work by by Li et al in January 2019 gives a sense for the level of complexity and precision that can be introduced in these methods. [6] The method uses IR and RGB pairs as the input to the deep learning neural network. Within the network, a convolution layer encodes the data in order to be fused and then decoded in order to be viewed. Neural networks can be extremely powerful once created, however they require a large amount of training data. This paper does not discuss neural networks since the audience are those seeking to decrease computational complexity.

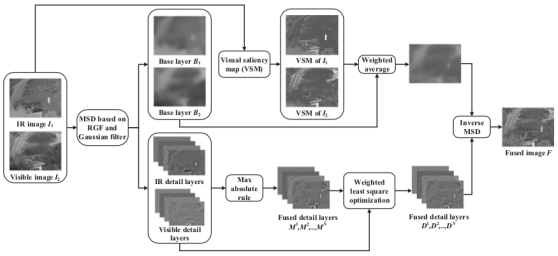

Figure 2. MAR optimized with WLS

Figure 2 visually describes the work done by J Ma et al. It serves as a template for our methodology as a state-of-the-art RGB IR fusion technique. After creating the LAP, the bottom (lowest frequency) layer is labelled the ”base layer” and the others are ”detail layers.” The base layer is fused using a VSM in order to determine the pixel that has the most salient features. The detail layers are first combined in a way that creates the MAR. The MAR takes the pixel from either IR or RGB based on which has the highest intensity. This is applied at each of the detail layers. Before the image is constructed, a weighted least squares optimization is applied. Each of the detailed layers is attempting to match the MAR, while subject to a penalty imposed for the

relative intensity of the pixel. In other words, pixels are only chosen that are salient in the image overall, but not salient compared to their neighbors. Overall, this gives a much less noisy image than the MAR itself. [7]

3. Method

The mathematics for each level of fusion are presented. The image used to demonstrate were obtained through a registered database from Oklahoma state. [2]

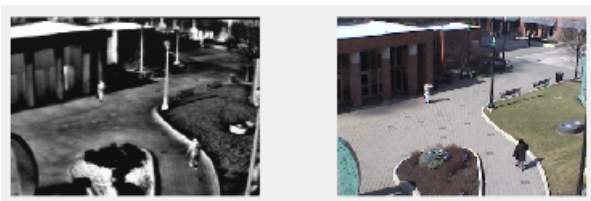

Figure 3. Static Images IR(left), RGB(right)

3.1. Averaging (Level 1)

The first and most simple means of fusion is a pure average. At every pixel, the RGB and IR intensity are added together and divided by 2. Note that since there are three color dimensions in the RGB image, we averaged the IR image for each individual color channel.

$$pix_{Fred} = 0.5 * (pix_{RGBred} + pix_{IR})$$

$$pix_{Fgreen} = 0.5 * (pix_{RGBgreen} + pix_{IR})$$

$$pix_{Fblue} = 0.5 * (pix_{RGBblue} + pix_{IR})$$

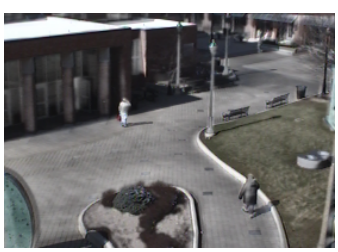

Figure 4. Level 1 Fusion

3.2. IR as Luminance (Level 2)

Level 2 fusion involves transforming the RGB image into the YCbCr domain. This technique was motivated by the work by G Li et al. They used IR images to construct the YCbCr space through linear functions.[9] In Level 2, the grayscale IR intensity is substituted directly for the luminance in the YCbCr picture.

Figure 5. RGB in YCbCr Space

Figure 6. Y (left), IR(right)

1. Convert RGB image into YCbCr Space
2. Isolate the luminance (Y) and compare to IR
3. Substitute luminance for IR values
4. Convert YCbCr image back to RGB

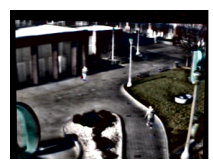

Figure 7. Level 2 Fusion

It is now possible to glean much more data from the IR image. However there was some increased complexity in the calculation. Clearly, there is much more that can be done to make the image more aesthetically pleasing to the viewer.

3.3. MAR Variation (Level 3)

At this level, a new low-cost version of the MAR is implemented. The purpose of our alteration is to achieve most of the value added by the MAR with much less computation. This loosely follows the process in Figure 2.

1. Construct LAP

   Apply an increasing gaussian blur to both IR and RGB. We used a spatial standard deviation of $\sigma = 2$ for the first blur. At each subsequent layer, we multiply the previous σ by 2. We found this achieves the maximum amount of desired blurring by the 5th layer. Note that the first layer, layer 1, is just the original IR/RGB image with no blurring induced. u^j is the image at layer j.

$$u^2 = Gaussian(u^1, \sigma = 2)$$

$$u^5 = Gaussian(u^4, \sigma = 2^4)$$

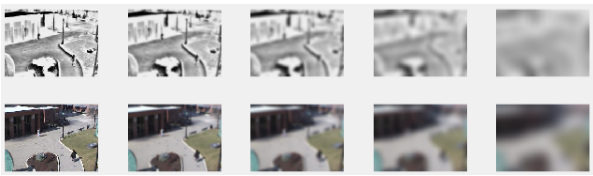

Figure 8. Laplacian Pyramid

2. Separate Base and Detail Layers

   The IR and RGB pictures that experienced the greatest degree of gaussian blur are called the "base layer." B_1 refers to the IR base and B_2 to the RGB base. The de-

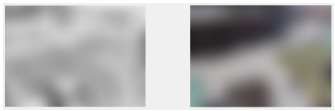

Figure 9. Base Layers

   tail layers are constructed by taking the difference between each layer and the next layer. The detailed layers for IR will be denoted d_1^j while detail layers for RGB will be d_2^j, where j refers to the layer. Thus, the IR detail layers will be $d_1^1, d_1^2, ..., d_1^N$ and the RGB detail layers will be $d_2^1, d_2^2, ..., d_2^N$. N designates the number of detail layers, which is 1 less than the original stack. We call detail layer 1 the layer with the least degree of gaussian blur applied. In the figure below, j increases from 1 to 4 left to right. Thus far, our method follows along with the MAR WLS method introduced by J. Ma etal.

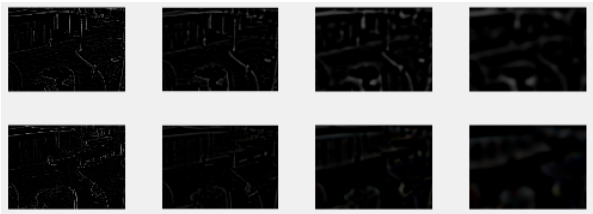

Figure 10. Detail Layers

3. Fusion of the Base Layer

   Now, the simplified version of the VSM is introduced. The intent is to determine the saliency of each pixel compared to the image at large. In this simplification, the salience is the difference between the pixel and the mean pixel value for the image. This is done at each layer. V_1 and V_2 correspond to IR and RGB respectively.

$$\bar{B}_1 = \frac{\sum_p B_1(p)}{R_1}$$

$$\bar{B}_2 = \frac{\sum_p B_2(p)}{3R_2}$$

   Where R_1 is the resolution of the IR image and R_2 is the resolution of each color of the RGB image.

$$V_1(p) = |p - \bar{B}_1|$$

$$V_2(p) = |p - \bar{B}_2|$$

   The pixels with the highest difference from the mean will have a higher intensity and thus a higher weight moving forward.

Figure 11. IR V_1 (left), RGB V_2 (right)

   Overall, the RGB image has higher saliency using this method. However, note that the pattern is very different. There are details in the bottom left of the image that will be taken from the IR image and not from the RGB image. We use the weights generated by V_1 and V_1 to fuse the base layers.

$$W_b = 0.5 + \frac{V_1 - V_2}{2}$$

$$B_F = W_b B_1 + (1 - W_b) B_2$$

   Where W_b is the weight of the IR base layer B_1.

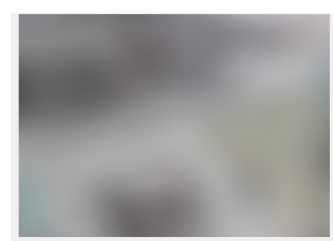

Figure 12. Level 3 Base Fusion

4. Fusion of Detail Layers

   Similar to the base layer, we would like to pick, for each pixel, whether the IR or RGB image has the more salient presentation. However, because of the higher freqency content in the detail layers than the base layers, it is not recommended to use the absolute difference, as in step 3. Instead, the IR and RGB image is compared at each pixel.

$$W_d^j = \begin{cases} 1, & |d_1^j| > |d_2^j| \\ 0, & otherwise \end{cases}$$

Where W_d^j is the weights of the detail layer j. A gaussian blur is then applied, with $\sigma = 2$ to smooth out the noise. Again, the layers increase from 1 to 4 left to right.

Figure 13. Wd

Now, the MAR is generated by fusing the detail layers together. Since the W_d was created based on portions of the image where IR dominates, W_d must now be attached to weight of the IR image d_1.

$$M^j = W_d^j d_1^j + (1 - W_d^j) d_2^j$$

Where M^j is our MAR image at each layer j. MAR represents fusion at each detail layer. To construct the full image, the detailed layers and base layers are fused together.

$$I_F = B_F + \sum_j M^j$$

Figure 14. Level 3 Fusion

The RGB and IR image fusion has now become useful to the viewer. There is enough color contained in the image to portray proper understanding of the visible light in the scene. From the IR image, the heat map is now visible.

3.4. VSM and MAR (Level 4)

For this level, the visual saliency map (VSM) will be introduced in the base layer fusion. The intent is to compare each pixel to each other pixel to generate our V_1 and V_2 for the IR and RGB base respectively. Thus:

$$V(p) = |I_p - I_1| + |I_p - I_2| + ... + |I_p - I_P|$$

Where p is our pixel of interest and P is the total number of pixels. Note that this must be done in two different dimensions for B_1 and three different dimensions for B_2. Because the RGB base must be computed for red green and blue color channels, we normalize the V_2 by a factor of 3 before comparing it to V_1. Below is a comparison of the VSM of the IR base V_1 compared to the comparison to the average used in Level 3.

Figure 15. Saliency: B_1 V_1 (left), Average (right)

The saliency map is of higher fidelity now. The two maps are similar. However, the higher fidelity VSM now captures direct comparisons between the IR pixels and the RGB pixels. From here we use the same equations from the Level 3 base fusion:

$$W_b = 0.5 + \frac{V_1 - V_2}{2}$$

$$B_F = W_b B_1 + (1 - W_b) B_2$$

This new VSM improves the averaging fusion rule from level 3 by taking into account the VSM. If $V_1 = V_2$ at some locations, the weight W_b would degrade to the common average weight. If $V_1 > V_2$, W_b would be larger than 0.5, and thus B_F would fuse more information from the base layer B_1 . Otherwise, when V_1 is relatively smaller than V_2 ; B_F would fuse more information from the base layer B_2 .[7] Here is our new base layer fusion compared to that of level 3.

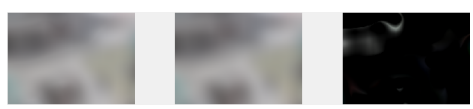

Figure 16. Base Fusion: Level 4 (left), Level 3 (center), Diff (right)

The difference between the two base layers is subtle. We magnified the difference by a factor of 50 to make it clearer in Figure 16. The VSM detects some pixels that are salient between neighbors but not salient in terms of the image as a whole. This creates a better blending for the viewer.

3.5. Level 5: VSM, MAR and WLS

At the highest level of fidelity, we combine all the techniques used up to Level 4 and add a weighted least squares (WLS) optimization for the detail layers. This is the full procedure proposed by J Ma etal. Similar to the VSM, the

intent is to detect pixels that are not only salient compared to the picture overall, but salient compared to their neighbors. We'll be picking up from step 4 in Level 3.

1. Create LAP

   The LAP is created using the same gaussian blur with $\sigma = 2$. Full description is given in step 1 of Level 3.

2. Collect Base and Detail Layer for IR and RGB

   The Base is identified as the image that has the greatest gaussian blur applied to it (lowest frequency data). If needed, reference step 2 in Level 3.

3. Fuse Base Layer using VSM

   Reference Level 4 for VSM method for fusion of base layers.

4. Construct MAR solution

   This will be done in the same manner we accomplished step 4 in Level 3. We should now have a VSM fused base layer B_F as well as a MAR image M^j for each of the j detail layers.

5. Construct our Spatial-Varying Weight

   The goal is to penalize pixels of high intensity in high intensity neighborhoods as well as remove penalty on low intensity pixel in low intensity neighborhoods. In order to do this, a matrix A^j is constructed for each detail layer in the following manner:

$$A_p^j = (|\sum_{q\epsilon\omega_p} d_1^j| + \epsilon)^{-1}$$

   Where epsilon is an extremely small value intended to avoid null returns, we used $\epsilon = 0.00001$. Our ω is a window around the pixel of interest. In other words, our A measures the strength of the neighborhood of the pixel. A large size of window would blur the fused image and increase the computational cost. On the other hand, a much smaller size of window cannot remove the influence of the noise and irrelevant IR details. Generally, the size 7 x 7 would be a good choice to produce satisfactory fusion results. [7] We used a window of size 7x7.

6. WLS Optimization

   Now, the building blocks for the WLS are ready.

$$\sum_p ((D_p^j - M_p^j)^2 + \lambda A_p^j (D_p^j - d_2^j)_p)^2$$

   Where D_p is the optimal detail layer. We can see that an increased λ value will increase the penalty imposed by the distance from the A matrix. In our method, we used a λ value of $\lambda = 0.01$. However, a λ within the range of [0.005, 0.02] should provide good results, depending on the image.[7] A higher λ value should be used in cases where there is already high contrast in the image. Let's write our equation in matrix form.

$$(D^j - M^j)^T (D^j - M^j) + \lambda (D^j - d_2^j)^T A^j (D^j - d_2^j)$$

   The optimal D^j will be found by solving the equation:

$$[2I + \lambda(A^j + (A^j)^T]D^j = 2M^j + \lambda(A^j + (A^j)^T)d_2^j$$

   Note that A^j is a diagonal matrix, thus it simplifies to:

$$(1 + \lambda A^j)D^j = M^j + \lambda A^j d_2^j$$

   .

   And thus we solve for our D^j in the following manner:

$$D^j = (1 + \lambda A^j)^{-1}(M^j + \lambda A^j d_2^j$$

   Once the optimal detail layer is obtained, the layers are combined in the following fashion:

$$I_5 = B_{VSM} + D^1 + D^2 + ... + D^J$$

   Where J is the maximum number of detail layers.

Figure 17. MAR (left), MAR WLS (center), Diff (right)

Note the difference the WLS makes in the image. It is now easier to see color data. The far right displays the color information that is not maintained with the MAR method alone. The WLS gathers all useful data from IR and is still able to create some color information.

4. Analysis, Evaluation and Comparison

The results will be analyzed in terms of qualitative and quantitative assessments of the computational complexity and the fidelity of the image. Computational cost is quantified by a simple tic toc function in MATLAB which measures the time for each computation. Fidelity is quantified using the standard deviation measurement, as recommended by a survey done by B Meher et al. Meher claims that it is difficult to quantify fidelity in a fused image for a variety of reasons, and thus qualitative assessment should be considered more important than qualitative in most cases.[4] Comparison is done through all 5 levels.

5. Results

In order to see the effect on multiple images, a new image is introduced. The image used for demonstration in section 3 is now Image 1. The new image is Image 2, below:

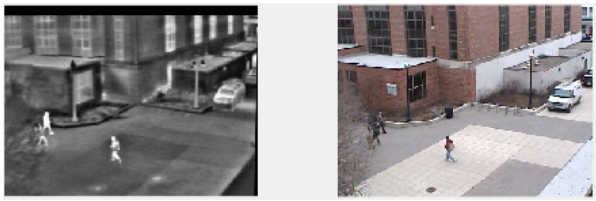

Figure 18. Image 2: IR (left), RGB (right)

5.1. Computational Cost

Below is the computational time for each level for each image. This was measured agnostic of any initialization of variables and image loading.

Table 1. Computational Cost Comparison (s)

Image	Level 1	Level 2	Level 3	Level 4	Level 5
1	0.0026	0.0086	1.5	35	35
2	0.0047	0.45	2.2	71	72

Level 4 and Level 5 fusion requires much more processing than Levels 1,2 and 3. One consideration is that these computations were performed on a standard Macbook in MATLAB. However, consider that instead of a single image, many of image processing involves gigabytes of data. Although the magnitude of time it will take to accomplish the computation may be exaggerated, the trend is still valid. Another consideration is the scale of the computational time based on image size. Image 1 is 201 x 281 (56280 pixels) and Image 2 is 240 x 380 (91200 pixels). The increase in the computational time is roughly proportional to the increase in the number of pixels between image 1 and 2. For applications with incredibly high resolution, it's recommended to use lower level fusion. There are many applications where little computational complexity is needed. One examples is using algorithms on mobile phones. A recent study attempts to use mobile phones to perform tongue diagnosis. But this requires simple, low cost algorithms in order to be performed on mobile phones. [10]

A qualitative assessment is now provided. All methods were coded from scratch and thus it's possible to speak to the level of challenge commensurate with the different levels. Understand that the most difficult portion of the experiment was finding registered images in IR and RGB. Image registration is an alternative, but it was not used. Level 1 required essentially zero effort. As long as one has access to a coding platform that can load images and knows how to do simple addition, this should be no issue. Level 2 required a small level of knowledge of YCbCr space. MATLAB provided easy conversions, but some programming platforms may not. Level 3 began the creation of the LAP as well as the base and detail layers, which was tedious but simple. Although Level 4 had a much higher computation time, coding V_1 in Level 4 was nearly the same as V_1 in Level 3. By contrast, the WLS optimization in Level 5 required some understanding of linear algebra and statistics. In summary, Level 3 to Level 4 caused the greatest increase in computation time, yet Level 4 to Level 5 increased the degree of qualitative complexity of operations the most.

5.2. Fidelity Comparison

The assumption inherent in the creation of Levels 1-5 was that the fidelity of the image would increase at each level. Images are first evaluated from a quantitative perspective. There are two main types of quantitative means of assessing images. The first is if we know the reference image. For this, the most common measure is the PSNR. [4] PSNR is used to compare each of the Levels 1-4 to Level 5 which is considered the reference image. PSNR is calculated as:

$$PSNR = 10log10(\frac{max(I_5)^2}{MSE})$$

Where

$$MSE = \frac{\sum_p (I_i - I_5)^2}{MN}$$

Where I_i is the fused image at level i. And the image has a resolution of MxN is the resolution of the images (dimensions).

Table 2. PSNR compared to Level 5

Image	Level 1	Level 2	Level 3	Level 4
1	15	13	44	43
2	16	14	51	49

Overall, a higher PSNR was obtained using image 2 than in image 1 over each level. The difference was not drastic, less than 10%. Most likely the cause is that image 2 has more piece-wise constant edges than image 1. Thus, the optimization problem can more easily identify edges and high salience in the pixels. Image 2 also has a higher average pixel intensity, which could be increasing the noise detected by the PSNR. In general, the PSNR increases as level increases, as expected. However, we have not yet verified that our Level 5 is the highest quality image.

The second type of quantitative analysis is much more unlikely. It is when we do not have a reference image. In

this case, the most common means of determining fidelity of a fused image is by using its standard deviation. [4] Ideally, a good image will have a high standard deviation. We use the following equation for the standard deviation:

$$SD = \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}(I_{ij} - \bar{\mu})^2}$$

Where $\bar{\mu}$ is the average pixel value. Below we show the standard deviation of each image at each level.

Table 3. Standard Deviation

Image	Level 1	Level 2	Level 3	Level 4	Level 5
1	0.0192	0.0710	0.0386	0.0383	0.0380
2	0.0209	0.0335	0.0307	0.0305	0.0305

As expected, the standard deviation increase in general as the Level increases. The main exception is Level 2, which has the highest standard deviation in both images. Understand that standard deviation is only one measure of quality. This will be reassessed during the qualitative portion. Excluding Level 2, we see no significance between the standard deviation of Levels 3, 4 and 5. This agrees with our hypothesis that Level 3 is viable, quality fusion solution.

Now a qualitative assessment of each level of the images is provided. For all images, image 1 is on the first row and image 2 is on the second. First are low level fusions: Level 1 (left) and Level 2 (right). Next are high level fusions: Level 3 (left), Level 4 (center) and Level 5 (right).

Figure 19. Low Level Fusion Comparison

Level 1 is not viable for most applications since the higher RGB intensity weights IR too lowly. At Level 2, a noticeable use of the IR image can first be seen. However, the image has incredibly high contrast and not appealing to the eye qualitatively.

Most likely, Levels 3, 4 and 5 provide a more accurate representation of what would add the most value. There are small differences within these layers. Level 4 highlights salient low frequency information more effectively than Level 3. For example, the heat created on the side of the wall in the top left of the picture is clearer in Level 4 than Level 3. This is due to the addition of the VSM in the base layer fusion. Comparing Level 4 and 5, the image is more colorful. The qualitative assessment concludes that Level 5 is the highest fidelity image.

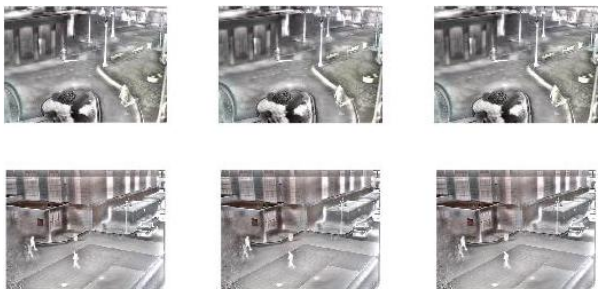

Figure 20. High Level Fusion Comparison

6. Discussion

Computational cost and fidelity of all 5 Levels were analyzed quantitatively and qualitatively. Increasing level increased cost. The main contributor is the VSM introduced in Level 4. Qualitative complexity in the model increases significantly from Level 4 to Level 5. Fidelity increases with level. The exception was Level 2. However, qualitative assessments will be considered more heavily than quantitative.[4] Optimal level is based on a variety of factors including application priorities, processing capability and image number/size. In most cases, it is recommended for the target audience to use Level 3 fusion.

Future research could provide concrete decision points based on specific applications for switching Levels. Further research could be done to determine how neural networks compare in terms of computational cost.

It's important to discuss limitations in this paper. This was performed by a single writer on a standard Macbook using MATLAB functionality. There may be built in function on other coding platforms that make these computations easier. There are absolutely more capable processors as well. This experiment only had a sample size of two images. More images would have made trends easier to identify.

As processing speeds increase, fusion is becoming more available to the individuals. The imaging mechanisms of infrared and visible sensors are different; thus, we should design niche-targeting saliency detection methods for infrared and visible images separately. [4] Thus, it's important that we adapt these niche algorithms at an institutional level as well as an individual level.

Acknowledgement

Thanks to Chris Metzler for his mentorship on this paper.

References

[1] Information technology – digital compression and coding of continuous-tone still images: Jpeg file interchange format (jfif). https://www.itu.int/rec/T-REC-T.871. Accessed: 2020-03-09.

[2] Osu color and thermal database. http://vcipl-okstate.org/pbvs/bench/Data/03/download.html. Accessed: 2020-03-09.

[3] B. D. A.P. James. Medical image fusion: A survey of state of the art. *Information Fusion*, 19(866):4–19, 2019.

[4] S. A. B Meher. A survey of region based image fusion methods. *ELSEVIER*, 2018.

[5] A. Gudipalli, S. Mandava, and P. Saravanan. Hybrid colour infrared image edge detection using rgb-ycbcr image fusion. 8, 07 2019.

[6] X. W. H Li. Densefuse: A fusion approach to infrared and visible images. *Image and Vision Computing*, 2019.

[7] B. W. H. Z. J. Ma, Z. Zhou a. Infrared and visible image fusion based on visual saliency map and weighted least square optimization. *Elsvier*.

[8] J. X. J.A. Richards. Remote sensing digital image analysis: An introduction. *Birkhäuser*, 4(866):102–104, 2019.

[9] G. Li, S. Xu, and X. Zhao. Fast color-transfer-based image fusion method for merging infrared and visible images. In J. J. Braun, editor, *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2010*, volume 7710, pages 219 – 230. International Society for Optics and Photonics, SPIE, 2010.

[10] K. T. L. M. H. Tania. Computational complexity of image processing algorithms for an intelligent mobile enabled tongue diagnosis scheme. *SKIMA*, pages 29–36, 2019.

[11] Y. Z. . M. Omar. Fusion for infrared and visible acquisitions, international journal of optomechatronic. pages 41–53, 2009.

[12] K. Rodhouse. A comparison of near-infrared and visible imaging for surveillance applications. *Masters Thesis*, 2012.

[13] B. Zitová and J. Flusser. *Image and Vision Computing*, 10 2003.