

Feedback to Reviewer's Comments – Second Revision

1)The reviewer has thankfully pointed out that we have significantly improved the paper in the revised version and that we have critically discussed the limitations of each reference. But he also added that we didn't mention how to solve these limitations. Actually, this is not quite accurate.

It is well known in research community that the literature review section should furnish old works, explore the related ideas and concepts, and discuss their possible limitations. Those limitations might be common to a bunch of papers. Eventually, the new proposed paper should address a solution to those limitations (taking into consideration the nature of the problem). And this is what we have done and considered in our new work.

We have addressed three major limitations: 1) the computational complexity, 2) effect of time lapse, and 3) enhancement of recognition rate. The computational complexity effects the time consumed and consequently the system speed. In page 4, lines [140 – 141] we addressed this issue of computational complexity. The same issue was noted in the same page - lines [153 – 154]. Moreover, the issue of time lapse can be studied by the virtue of having a database with images taken at different times. Most of the available works have used databases that do not address this issue, and we have stated this matter in our paper. For example, in page 3 lines (98 – 104): “Bhowmik et al. [14] introduced the role of different IR spectrums, their applications, and some interesting critical observations. In their experimental work, they fused both thermal and visible images to enhance the recognition rate, as it is expected that the fusion process improves the overall performance of the system. They have tested their method on IRIS and Terravic databases. In both databases, the images were taken in one time session. In other words, the effect of time lapse has not been taken into consideration.” As this paragraph reveals, we have mentioned the time lapselimitation clearly. With respect to performance issue, we have explained that some techniques did not show attractive performance. For example, Ref [15] page 3 – line 109, and Ref [18] page 3 – lines [121 – 122].

Now, we explain how we have mentioned the way to solve the limitations stated above. First, we have clearly stated in page 2, lines (49 – 56): “Decision level fusion finds a combination of multiple statistical patterns to produce an integrated result that is enhanced in terms of information content for pattern recognition and classification. Local representations offer robustness against variability due to the changes in localized regions of the objects. The features used in the local feature analysis methods are less sensitive to illumination changes, easier for estimating the rotations, and have less computational burden. Hence, the proposed system employs the advantages and the characteristics of: thermal images, component-based approach, the statistical features, and fusion.” As it is obvious from the paragraph, the new proposed technique solves these limitations by exploiting the characteristics of:

- i. thermal images,
- ii. component approach,
- iii. statistical features, and
- iv. fusion.

Then, we further add:

- a) Computational complexity: this issue is studied by examining the amount of time consumed and the related speed, and we have shown this in Section 5, page 11 – lines (364 – 367): “An important advantage of our system is its speed and efficiency. Table 4 shows the time consumed by the system to complete the histograms of the different components. As can be seen from Tables 3 and 4, the system performance demonstrates its effectiveness when both the recognition rate and the time consumed are considered.” Table 4 shown in page 20 demonstrates an example of the time consumed (in second) needed to calculate the histogram for each component for different images of the same subject. The noticeably small amount of time taken by the system is quite obvious from this table. This reflects that the proposed system is not computationally greedy.
- b) Time lapse: we have indicated in Section 4, pages (7 – 8) – lines (261 – 264): “We have built a database of 570 images for 19 different subjects taken at different sessions. The database consists of male and female coming from various ethnic backgrounds. Acquisitions were held at different times and most subjects participated in multiple sessions across a number of different weeks.” So, when our method works effectively on a database characterized by this “time lapse” challenge then it shows that this method can solve an important limitation that was not previously solved or addressed.
- c) Enhancement of recognition rate: our technique shows very promising results with a recorded success rate of 96.4%. We have shown this clearly in the comparison made with other methods shown in Table 3 - page 19.

On the other hand, the reviewer has said that the proposed method is published in several papers. The reviewer should note that the feature vector that we are proposing doesn’t consist only of histogram calculations. Rather, the first and second moments of both the global image and the components (parts) of the image are also included. Page 5, lines (204 – 210) states: “In this paper, we propose a feature vector that consists of the following different feature parameters: first moment, second moment, and the thermal image histogram. The calculation of the first, second moments and the histogram distribution is implemented at the component level, beside the whole face image. The final decision results depend on the *local* mean, variance and histogram beside the computation of the *global* mean, variance and histogram over the entire thermal face image. These local parameters are more powerful in preserving the discriminant features in a neighborhood of a pixel in an image.” To the best of our knowledge, this new combination of tools and techniques is novel and is presented for the first time. These lines

further demonstrate our strategy in tackling the limitations of other works. Also, we emphasize on the effect of fusion and its role on the final results. This is also been mentioned in page 6, lines (239 – 249): “Decision level fusion combines the results to yield a final fused decision based on a joint assertion of multiple single source results (or decisions) to achieve an improved classification. Prior knowledge and domain specific information can also be incorporated. Fusion techniques exploit synergistic integration of the information obtained from different data sources or from multiple pattern classifiers to produce a more informative form than the original, and hence to improve the overall classification accuracy [24]. Fusion methods include AND fusion, OR fusion, majority voting, and ranked list combination. In our experiments, we have implemented the AND fusion, as it has proven its superiority among the other fusion methods [24].” Again, this is a newly added dimension to our method that enhances the overall performance.

2)OK.

3)Actually we have clearly addressed the description of the database in Section 4, pages (7 – 8) – lines (261 – 264): “We have built a database of 570 images for 19 different subjects taken at different sessions. The database consists of male and female coming from various ethnic backgrounds. Acquisitions were held at different times and most subjects participated in multiple sessions across a number of different weeks.” In this since, we have clearly identified the following variations that our database considers: the time lapse (since the images were taken at different times), gender variation (since we have images for male and female), and the skin color effect or the genetic effect (since we have images for people from different ethnic backgrounds, i.e.; from Middle East, India, and Europe). We have also mentioned the type of camera that has been used in image capturing. Further, we have given complete details about each one of the three sets we designed (the number of images used in every set and the way these images have been chosen).

4)The reviewer has repeatedly mentioned that “*the paper suffers the problem of “lack of novelty”*”. Well, we need to clarify an important point here. The idea of histogram by itself is not new, however; the intelligent combination of the different components (where the histogram being one of them) is novel. We repeat here what we have stated earlier in page 2, lines (46 – 56) that: “In this paper, we present a new technique for face recognition that exploits the local characteristics of the thermal image. The “*whole*” face image is divided into components of different sizes. The statistical features of these components, beside the “*whole*” image, are combined together using fusion methods. Decision level fusion finds a combination of multiple statistical patterns to produce an integrated result that is enhanced in terms of information content for pattern recognition and classification. Local representations offer robustness against variability due to the changes in localized regions of the objects. The features used in the local feature analysis methods are less sensitive to illumination changes, easier for estimating the rotations, and have less computational burden. Hence, the proposed

system employs the advantages and the characteristics of: thermal images, component-based approach, the statistical features, and fusion.”

Further, the reviewer may have noticed that the feature vector doesn’t consist of only histogram calculations. Rather, the first and second moments of both the global image and the components (parts) of the image are also included. In page 5, lines (204 – 210) we state: “In this paper, we propose a feature vector that consists of the following different feature parameters: first moment, second moment, and the thermal image histogram. The calculation of the first, second moments and the histogram distribution is implemented at the component level, beside the whole face image. The final decision results depend on the *local* mean, variance and histogram beside the computation of the *global* mean, variance and histogram over the entire thermal face image. These local parameters are more powerful in preserving the discriminant features in a neighborhood of a pixel in an image.”

Moreover, we have mentioned in page 6, lines (239 – 249): “Decision level fusion combines the results to yield a final fused decision based on a joint assertion of multiple single source results (or decisions) to achieve an improved classification. Prior knowledge and domain specific information can also be incorporated. Fusion techniques exploit synergistic integration of the information obtained from different data sources or from multiple pattern classifiers to produce a more informative form than the original, and hence to improve the overall classification accuracy [24]. Fusion methods include AND fusion, OR fusion, majority voting, and ranked list combination. In our experiments, we have implemented the AND fusion, as it has proven its superiority among the other fusion methods [24].” We would like further to emphasize that histogram calculations are made at different levels. It is NOT only based on the global image, rather, we have clearly stated that it is also implemented at the local level (for different component sizes). We refer the reviewer again to Figure 1 where we clearly demonstrate this.

It should be clear that when we implement the histogram on local components, we are considering only the local characteristics of the component, and not the global one. In other words, we are considering the effect of the neighborhood region (neighborhood temperature) not the whole face temperature. So, the obtained features are for the local parts (which are discriminant) and can contribute differently to the final system decision. This is clearly shown in Figure 14 – page 18.

To summarize it, it is the effect of the combination of the different tools and techniques that offers multiple advantages: high recognition rate, low computational complexity, efficient process time, beside the issue of illumination variation that is considered inherently by the nature of IR image theory.

In this regard, we would like to point out that PCA or LDA, for example, are not new methods. Nevertheless, researchers have exhaustively utilized them where these two techniques can be found as major elements in many new proposed techniques that address the various issues

offace recognition domain. It is the new combination of these famous techniques with other methods, which validates the originality and novelty of the proposed method. In other words, it is the whole concept or system that we finally look at, that reveals its novelty, not only one part of it. And this can be easily verified in the various scholars' literature.

5) We agree that we should have a general protocol to set a common standard that enables researchers to compare between the different techniques. However, this field is relatively new, and has not been tackled yet so aggressively by researchers, especially when compared to visible light domain. Hence, there is no common ground for comparison. At this stage, we aimed at proving our concept. The initial experimental results are encouraging and motivate us to further explore new directions in the field of thermal image recognition. As the reviewer may agree, it is infeasible to implement each of the existing methods on the new database in ONE week time (one week time is the period given for us to send our feedback to reviewer's comments). So, due to this time constraint we will postpone this task to our future work.

Another important issue is, it is true that we should aim at increasing and enhancing the success rate, but it is also important to enhance this rate in an *efficient* way. In this work, we are not only considering the success rate, but also the consumed time. It is important to have a system that performs superiorly, but it is also important that this superior performance does not tolerate the system speed. This was the idea of this paper.