EG UK Computer Graphics & Visual Computing (2016)
Cagatay Turkay and Tao R. Wan (Editors)

Exploring Face Recognition under Complex Lighting Conditions with HDR Imaging

E. O. Ige, K. Debattista, R. Mukherjee and A. Chalmers

Visualisation Group, WMG, University of Warwick,
Coventry, United Kingdom

Abstract

Applying image processing applications under complex or harsh lighting conditions can be a difficult challenge. In particular, face recognition can be prone to such limitations due to the uncontrolled nature of the applications to which it is applied. One of the conventional ways used to resolve this concern is by capturing images under controlled light or pre-processing the affected images, which can change the perception of the resultant images. One of the primary issues with this is the lack of information present in the original images due to over-exposed and under-exposed pixels. High Dynamic Range (HDR) imaging offers an alternative due to its capability of handling natural lighting. This paper explores the use HDR imaging for face recognition. A training and testing set of HDR images under different harsh lighting conditions was created. Traditional low dynamic range methods were compared with using the full range and applying HDR methods to a traditional face recognition method. Results demonstrate that adapting HDR captured images for use with traditional face recognition methods via a tone mapping provides sufficient improvement and enables traditional algorithms to cope well with harsh lighting scenarios.

Keywords: face recognition, dynamic range, high dynamic range image, tone mapping.

1. Introduction

Face recognition (FR) systems have found application in a number of areas, particularly, in security and surveillance and have achieved success under controlled lighting conditions. However, FR performance tends to drop in more realistic and uncontrolled lit scenes, thus making most of the conventional approaches inapplicable without additional tasks. Low performance under uncontrolled lighting conditions can be due to expression, pose, gender, image resolution, illumination etc. [LHK05, BJ11]. Amongst these, illumination is considered one of the fundamental characteristics capable of reducing performance. Thus, images of the same person can appear different to computer algorithms under a changing light [BJ11] as shown in Figure 1. However, HDR techniques [BADC11] present an alternative to the traditional low (or standard) dynamic range (LDR) imaging methods. While LDR imaging suffers from over or under-exposed pixels when handling scenarios with a wide dynamic range, HDR can capture, process and present the entire range of natural lighting.

This work explores whether HDR can benefit FR systems. In order to demonstrate the benefit, a new, modest, HDR database of faces has been created under three complex lighting scenarios. Using a traditional FR method for LDR data, FR rates were compared across a number of conditions. The main conditions considered in this work are LDR input for FR, LDR input pre-processed

Figure 1: *Sample images of the same person under different light.*

with traditional methods to reduce illumination issues, and finally HDR input converted to LDR using a traditional HDR optimal exposure and tone mapping method. Results demonstrate that traditional methods struggle to recognise faces under the complex lighting conditions presented here. However, the use of straightforward HDR techniques, by just capturing in HDR and tone mapping present successful performance. Crucially, this approach does not require the development of new FR systems but can be used on already functioning methods without modification.

The rest of this paper is organised as the following. Section 2 discusses the related work. Section 3 describes the overall methodology of the work and the steps taken to achieve the results. In

DOI: 10.2312/cgvc.20161296

Section 4, we describe the adopted face recognition approach and presented the chosen illumination pre-processing method. Section 5 provides the experimental results and discussion. Conclusion is presented in Section 6.

2. Related Work

FR under harsh lighting conditions has received some attention in the research domain. The traditional approaches for dealing with this issue can be broadly described under normalisation based and feature based methods. In [TT10], the appearance based approach, part of the feature based methods, was used to define training examples under different lighting conditions and directly used without undergoing lighting pre-processing and further used to learn possible illumination variations. These are then generalised to the variations present in the images. Also an approach involving modelling the effect of illumination on human faces was implemented in [LHK05], they showed the existence of configurations of single source directions of light to be effective for FR and report that a linear subspace spanned by the corresponding images can be approximated to illumination cone, thus performing better with FR under wide range of difficult lighting conditions.

With the normalised based approach, images are reduced to a canonical form in a way to suppress the lighting variations [TT10]. For example in a face, naturally occurring incoming lighting distributions have information predominantly as low spatial frequency and soft edges, such that the high frequency information in the image are mainly signal (inherent facial appearance). The Multiscale Retinex (MSR) method [JRW97] of Jobson was used to cancel the low frequency information, the images were smoothed and used to divide the smooth version of itself. In a similar way, the Self Quotient Image (SQI) method [GB03] was used with a different local filter. This was improved [CYZ*06] with the Logarithmic Total variation (LTV) smoothing. Recently, a comparative analysis was conducted [SKM04] on the above and related methods, they concluded that the normalised based approach as been effective but limited in there ability to handle spatially non-uniform lighting variations, which is the main focus in this paper.

The popular approach involving the extraction of insensitive illumination feature sets directly in a presented image, such as geometric features, image derivative, edge maps, Local Binary Patterns (LBP), local autocorrelation filter, Gabor wavelet, etc. [AHP06, PYL08, ZSCG07], perform great with raw gray images, but are limited in resistance to the harsh lighting variations present in the real world scenes. Accordingly, the LBP features have proved to be effectively invariant to monotonic global gray-level transformation [TT10], yet it degrades in performance easily under harsh changes of lighting directions and shadows, and similar performance drop also applies to other features discussed earlier.

Pereira, et al. [PMPP14] tested the automatic face recognition using sparse feature representations with tone mapped operators applied on 20 HDR faces from five selected individuals, they presented a preliminary result in which the logarithmic tone mapping operator performed best. In this paper, our work is conducted under controlled conditions with lighting environments clearly identified and quantified and with more participants. We carefully detailed procedures taken for our image capturing exercise, image dynamic range is indicated as shown in Section 3.2. Furthermore, we generated more datasets for zeroth exposure and optimal exposure respectively from the original HDR database created, which contains 498 HDR images.

3. Methodology

This section describes the overall motivation of this work and the steps taken to achieve the results. Furthermore, the creation of the HDR dataset and its capture are described.

3.1. Motivation

For FR to be used "in anger" it must be capable of coping under the harsh lighting conditions exhibited in the natural world. Unfortunately, traditional imaging methods perform poorly in such environments, particularly when detailed information for an application such as FR are required. HDR data provides the ability to handle content beyond the traditional 8 bits per colour channel attributed to LDR imaging, by storing floating point values per colour channel per pixel. However, a broad number of applications have been designed based on LDR content. The motivation behind this work is to establish whether HDR can benefit FR with the minimum possible intrusion. In this work the use of capturing with HDR content is proposed and the straightforward HDR method of tone mapping used to attempt to maintain as much data as possible from the original captured scenario that can be used by a traditional FR system. In order to study this proposal, three harsh lighting scenarios are contemplated one with a bright light source behind the face, one to the side producing harsh shadows across the face and a similar one at an angle producing shadows at an alternative angle.

3.2. HDR capture

In this work, we capture high quality LDR images under three different categories of harsh lighting conditions via the process of bracketed exposures to create an HDR image. The captured images were visually inspected for quality. The poor quality images in cases where participants blink eyes, out of focus or appear blur were all removed. The final HDR images were cropped at the facial region and resized to $150x150$ pixels.

In order to obtain the HDR image, seven LDR images were bracketed within the range of $R \in [-8, +8]$ Ev with an interval of 2.42 Ev (exposure compensation value) between each LDR exposure. Since the exposure compensation values were changed, the sensor sensitivity/film speed (ISO setting) of the camera was fixed to 1600 and the base (0^{th}) exposure was fixed to a relative aperture of $f/5.6$. The shutter-speed was allowed to vary to obtain the required exposure. The capture details are given as below:

- Camera body: Canon EOS 5D Mark III
- Lens: Canon EF 24-105 f/4.0
- ISO: 1600
- Aperture: $f/5.6$ (base exposure)
- Bracketing range: -8 to + 8 Ev
- Number of exposures: 7
- Bracketing interval: 2.42 Ev

Figure 2: *Schematic diagram of the illumination setup.*

- Capture format: .cr2 (14 bits/pixel/channel)

The illumination setup was created with the help of an ARRI studio lamp [Gro] with a peak luminance value of $\approx 2000\ \text{cd}/\text{m}^2$ where the lamp was placed at three different positions relative to the subject i.e. a) to the right, b) behind and c) overhead. Such a setup ensured that the artificially induced harsh-lighting conditions with prominent illumination edges on the subject's face cannot be captured with a single exposure and significant details would be lost in either the low or high luminance regions of the image. The camera settings were unchanged throughout the capture session. Figure 2 illustrates the illumination setup. The HDR dataset is composed of 21 participants with ages ranging between 23-50 years comprising of 17 males and 4 females. The average dynamic range of the images is 15 ($\log_2$). The LDR images were visually inspected and out of focus/blurry (due to motion) images were discarded. Subsequently, the candidate images from each of the sessions were merged into HDR images, cropped at the facial region and resized to a resolution of 150×150 pixels. The resultant HDR images were stored in the .exr format [FKH] for further processing.

3.3. HDR tone-mapping

Face identification/recognition and feature detection algorithms such as Speeded Up Robust Features (SURF) and Scale-invariant Feature Transform (SIFT) typically operate on LDR images wherein the luminance pixel values accepted by the algorithms are in the normalised range of $\text{RGB}_{(x,y,z)} \in [0,1]$. In order to take advantage of traditional methods the captured HDR data is converted to the detection algorithm suitable format while keeping the overall appearance/tone of the scene as similar as possible to the reference HDR. Such a task is usually accomplished by either choosing one of the exposures (typically the base 0^{th} or optimal exposure [DBRS*15]) or by using a tone-mapping operator (TMO) to map the HDR to its corresponding LDR representation. Analogous to the display driven data characteristics, the tone-mapped image is gamma corrected and passed to the feature detection algorithm. However, there are two major issues. Significant amount of details are unavailable in the base exposure and there is a plethora of TMOs which perform equally well for various requirements which presents a challenge of definitively choosing one. Since the goal in this work was maximal tone preservation, the Display Adaptive TMO proposed by Mantiuk et al. [MDK08] was chosen since it has been classified as a Scene Reproduction Operator (SRO) [EWMU13] and has performed well in previous evaluations [MBDC15, UMM*10]. The results of another, more straightforward TMO a logarithmic TMO as used by Perriera et al. [PMPP14] in their HDR study for facial recognition is also presented as a comparison.

The goal of the SRO is to preserve the appearance of the original HDR scene including contrast, sharpness and colours by adjusting the image with the pre-notion of the ambient illumination and capabilities of the target display. The authors demonstrate that this can be defined as a non-linear optimisation problem which when simplified by reducing the degrees of freedom results in the introduction of a mapping technique with adjustable parameters. The SRO employs a piecewise linear tone-curve to map the HDR luminance to its corresponding LDR luminance. Given a particular display characteristics, the TMO produces the least distorted image in terms of visible contrast distortions (measured in Just Noticeable Difference (JND) steps) which when weighed by an HVS model accounts for luminance masking, spatial contrast sensitivity and contrast masking. Moreover, the SRO also uses chroma preservation techniques introduced by Schlick [Sch95] to preserve accurate chroma information. This SRO was chosen over many other candidates because it reproduces the reference HDR scene with minimal visible distortions and, as mentioned above, has performed well in previous evaluations. Figure 3 shows results of tone mapping on images in the captured database.

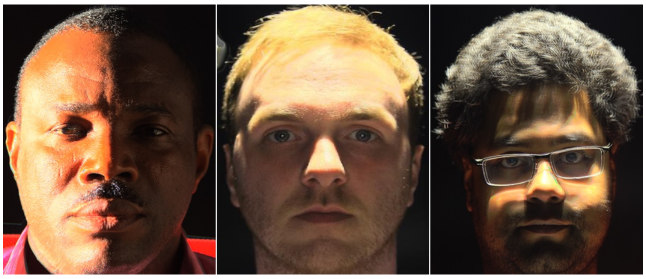

Figure 3: *Sample tone mapped faces from our HDR dataset captured under different light.*

4. Face Recognition

This section describes the adopted FR approach. Furthermore, the chosen method for illumination pre-processing for LDR is presented.

4.1. SURF using BOF Technique

The SURF algorithm using BOF techniques is discussed in this section as the next step following the image processing stage.

SURF

SURF (Speeded-Up Robust Features) is an invariant detector and descriptor using image interest points [BETVG08]. The most important SURF property is the ability to repeat interest points, which goes to define how reliable and robust the detector is in finding the same interest points under different lighting conditions, as in our case. Thus, the interest points are detected at distinctive image locations, such as corners, blobs and T-junctions.

Given a single channel image, I_0 with pixel intensity I_0 $(x;y)$ at a given point $X = (x;y)$, if the central difference method is applied, the first-order difference achieve no response when applied to unchanging signals. To check for any changes in intensity, the second derivatives of I_0(x; y) is expressed as:

$$\frac{\delta^2 I_0(x,y)}{\delta x^2} = I_0(x+1,y) - 2I_0(x,y) + I_0(x-1,y) \tag{1}$$

If the case that I_0 corresponds to another image I_u is considered, with an unknown lighting condition, and if we assume that the diagonal-offset models the two images (I_0 and I_u) as related by a linear transformation determined by a scalar constant α and an offset β, the pixel intensity can simply be modelled such that $I_u(x;y)$ of the image I_u at the same point $X = (x;y)$ is given as:

$$I_u(x,y) = \alpha I_0(x,y) + \beta \tag{2}$$

From equation 2, the second derivative of $I_u(x,y)$ with respect to x, can be expressed as:

$$\frac{\delta^2 I_u(x,y)}{\delta x^2} = \frac{\alpha\delta^2 I_0(x,y)}{\delta x^2} \tag{3}$$

Similarly, the same principle can be applied to the second derivatives in y and XY, where the offset term β is cancelled out in the computation of the derivatives, without effect on the final result. But when the illumination is varied with a scalar α, there will be a proportional variation in the second derivatives with the scalar.

Typically, localising feature with SURF algorithm involves - interpolations, discarding low-contrast key points below a given threshold, (e.g. threshold can be fixed at 0.03) and eliminating edges to increase key point stability. The challenge is, if a detector suffer from varying response to changing light, a feature (keypoint) for instance in a bright image region may not be detected in the corresponding image with lower lighting levels. Therefore, SURF detector response $R_u(x;y)$ of a given pixel $I_u(x;y)$ is given by the determinant of the Hessian matrix:

$$R_u(x,y) = \frac{\delta^2 I_u(x,y)}{\delta x^2}\frac{\delta^2 I_u(x,y)}{\delta y^2} - (\frac{\delta^2 I_u(x,y)}{\delta xy})^2 \tag{4}$$

Therefore, in order to compute SURF response, we substitute 3 into 4, so that the filter response R_u can be represented as $R_o(x;y)$ of the I_o and summarised into 5

$$R_u(x,y) = \alpha\frac{\delta^2 I_0(x,y)}{\delta x^2}\alpha\frac{\delta^2 I_0(x,y)}{\delta y^2} - (\alpha\frac{\delta^2 I_0(x,y)}{\delta xy})^2 = \alpha^2 R_0 \tag{5}$$

In 5, because of the degree of α^2, even with small variation in scene light, there can be significant variations in the size of the detector response.

BOF

A Bag Of Feature (BOF) [FFP05] algorithm is used in the study to build feature dictionary. Like most modern approaches to category-level object detection, this algorithm attempts to use intensity features as input in both training and testing tasks. In this case, each image is represented as a collection of detected patches over a facial patch of 8x8 grid used to generate 500 visual BOF, where images are characterized by their illumination-invariant regions, along with their SURF descriptors. Quantisation is carried out through clustering to construct the image's signature formed by the centers of the clusters and their relative sizes into a more manageable size. The centroids is then used to provide an encoding method for counting feature frequency in each image, where each image membership is used to build a histogram of length k where the $i'th$ value is the frequency of the $i'th$ dictionary feature. Multi-class SVM is used to train on the objects categories and used for image classification.

4.2. Face Classification

SVMs (Support Vector Machines) are a widely used supervised learning algorithm for data analysis and pattern recognition [ZTC14]. SVM implementation is adopted for this problem. A RBF (Radial Basis Function) kernel was used to map the inputs to a higher dimensional space. This has been successfully used previously for face recognition [GJ10]. For instance, given a pair of faces, the SVM is trained to determine whether faces belong to the same or different subjects. The SVM was trained using the encoded histogram represented in the form of feature vector from the 21 subjects of the HDR-tone mapped dataset using 80% randomly generated pairs and testing with remaining 20% randomly generated pairs. As previously mentioned, we implement three induced harsh lighting conditions in our dataset to create three different representations, see Figure 2. Based on this, we decided to implement disjoint lighting conditions for training and testing. This means, we presented different input to the SVM for training and different input for testing.

4.3. Normalised Discrete Cosine Transform (NDCT)

One of the methods explored in the results section employs the use of pre-processing of the LDR image to remove aspects of the harsh lighting. The chosen method for pre-processing adopted for this work is normalised discrete cosine transform (NDCT). The authors in [IDC16] compared four popular pre-processing methods for solving illumination problems used for image enhancement for emotion recognition, which shares similar characteristics to FR. It was found that the NDCT approach performed well above the other methods for the pre-processing methods. Also in [GNV11], the authors report lowest error rate with DCT even with down sampling coefficients.

The NDCT transform image representation as a sum of sinusoids of varying magnitudes and frequencies. Most salient information exists in low frequency coefficients. Unfortunately sometimes, the region where illumination varies are present in coefficients with low frequencies. Typically, illumination variations are reduced by setting the low-frequency DCT coefficients in logarithmic domain to zero [FN09]. In frequency domain, it is widely believed that illumination changes mainly in the low-frequency band, and research [GNV11] has shown that illumination changes slowly in the facial region. This means that, in order to obtain robust facial features under harsh light, the recovery of reflectance characteristics is important. Therefore, for illumination compensation, setting the DCT coefficient to zero is equivalent to subtracting the DCT basis image product and the corresponding coefficients from the original image. Setting the n low frequency DCT coefficients to zero gives:

$$F^{'}(x,y) = \sum_{\mu=0}^{M-1}\sum_{\nu=0}^{N-1} E(\mu,\nu) - \sum_{i=1}^{n} E(\mu_i,\nu_i)\, i = F(x,y) - \sum_{i=1}^{n} E(\mu_i - \nu_i) \tag{6}$$

where $E(\mu,\nu) = \alpha(\mu)\alpha(\nu)C(\mu,\nu)Cos[\frac{\pi(2x+1)\mu}{2M}]cos[\frac{\pi(2y+1)\nu}{2N}]$

And the illumination component term can be regarded as $\sum_{i=1}^{n} E(\mu_i,\nu_i)$.

5. Results and Discussion

This section presents a series of results. The different experimental conditions are meant to validate the performance of HDR on the FR algorithm used in this paper. To demonstrate this, results are presented across the three different induced lighting representations (back, left and overhead) both individually and as a whole. For these results, the accuracy of the algorithm in learning a set of faces from training images and then correctly recognising the same people from a test set of different images is evaluated. For these tests, both the training and testing sets contain the same people. To achieve this, tests are carried out based on 80% training and the rest on testing. This setup was repeated five times and the recognition rates averaged over the five trials. Due to limitations in the data set size a procedure known as cross-validation was adopted. This separates the data set ($N = 21$) into two parts leaving one part to be unknown, such that the prediction accuracy obtained from the "unknown" set more precisely reflects the performance on the classification of an independent data set.

To prevent data over-fitting the k-fold cross-validation was used [ZTC14], whereby the training set is divided into k subsets of equal size. Thus, one subset is tested using the classifier trained on the remaining $k-1$ subsets. Where each instance of the whole training set is predicted once, consequently the cross-validation accuracy is the percentage of data that is correctly classified. Our classification problem falls into the category of the mutually exclusive (multi-value classification). Wherefore, a decision on one class leaves all options open for the other classes. Thus, in a sense the classes are suppose to be independent of one another, but the classes are rarely statistically independent. So multi-class SVM classifiers are learnt and applied on each training set. Finally, the decisions of all classifiers is set as the recognition rate.

Initially the proposed FR method is validated against a traditional database. The Caltech 101 database [FFFP07], the de facto validation dataset for object recognition, containing images with large variations in light, pose and expression, was chosen for this test. In the comparative evaluation report of several recognition algorithms on Caltech 101 dataset conducted by [ZBMM06], the authors reported 65%, although their study was not limited to faces only. In this test a performance of 86% was achieved for the method presented in section 4; this result was limited to testing only images that were clearly faces.

5.1. Overall FR performance

FR on the HDR dataset is tested across five methods - $0th$ exposure (naive), $0th$ exposure with NDCT preprocessing (NDCT), optimal exposure (optimal), logarithmic tone mapping (Lg_TMO) and Display Adaptive tone mapping (DA_TMO) with the type based lighting conditions - back light, left light and overhead light. The naive approach is based on the set consisting of the $0th$ exposure from the seven exposures captured within bracketing range (-8 to $+8$). Since the naive approach is unable to capture the full scene lighting for the HDR scenarios, NDCT [IDC16] described in section 4.3, was adopted and use as a pre-processing technique. This method would represent the LDR approach of dealing with harsh lighting. The optimal method is an HDR to LDR method that modifies the

HDR data into an LDR by selecting the largest contiguous area in luminance space to fit into an LDR. The algorithm used here is based on Debattista et al.'s method [DBRS*15]. The resulting outputs from this method are not dissimilar to what a camera sensor uses to attempt to fit the wider dynamic range of a scene that is being captured. Lg_TMO is a logarithmic tone mapper that is considered relatively straightforward compared to other tone mappers; it is added here to provide a comparisons with results in the related work of Perriera et al. [PMPP14]. DA_TMO was based on the DA_TMO as discussed in section 3.3 and Lg_TMO for comparing our result with the work in [OPAHC*14]; this represents the state of the art of HDR to LDR methods. Table 1 presents results for the four methods. The performance increases as expected naive (82%), NDCT (84%), Opt_exp (87%), Lg_TMO (87.7%) and DA_TMO (93%). Naive under-performs as expected and NDCT does not improve much. Optimal is better but the TMOs are best. Lg_TMO does not perform as well as DA_TMO as expected. DA_TMO does very well with an overall of 93% indicating that a robust TMO may be sufficient for general FR performance for scenarios with harsh lighting.

5.2. Comparison of FR performance with disjoint training and testing set

In this section, the performance when the datasets is separated across the three lighting conditions is presented.

Generally, with HDR imaging [MBDC15] there is the advantage of capturing wide range of available scene light. But when images are presented in a scene where the area of interest is away from the camera lens, the performance of such HDR imaging would be low. Therefore, we observe that using images captured under the left light for training or testing lead to a drop in performance as shown in Table 2, but slightly higher in Lg_TMO and AD_TMO. Tables 3, 4, 5, 6 and 7 show the confusion matrices for the recognition. The confusion matrix is used to display statistics for assessing supervised classification accuracy, all correct guesses are located in the diagonal of the table with degree of mis-classification among classes (errors) shown on the outside. As reported, DA_TMO performs best and is not producing many false positives.

So far, we have used the FER accuracy parameter to discuss the correctness of our classifier's performance. To further strengthen our discussion, we explore other clues to give a further understanding of where the classifier is going wrong. Since we adopted the multi-class classifier, for the purpose of generalisation, we decided to compute the success rates of the presented confusion matrices using the *precision* and *recall* metrics from one label versus all other labels. Precision, gives all the predicted labels (for example, the class Bck_light). Similarly, Recall, gives all instances that should have a label Bck_light, meaning how many of these were correctly captured? [SL09].

To compute precision and recall, the following shall be defined:

- Our confusion matrix tables assumes three possible output labels: Bck_light, Lft_light and Ovh_light.
- The diagonals of the matrices contains the *True Positives (TP)* for each label.
- The sum of a column would be the total number of instances that should have label X_light
- The sum of a row would be total number of instances predicted as a particular label X_light
- Given the above, the precision of a label X_light is computed as: $precision = TP\_X\_light/(TotalPredicted\_X\_light)$
- The recall of a label X_light is computed as: $recall = TP\_X\_light/(TotalLabel\_X\_light)$

We notice that all the TP (accuracy) in Tables 3 - 7 computed come out to be recall metric. Based on this, we decided to ignore the computation for recall. To compute precision, we take all rows as the emotional labels being predicted and all columns as the predicted emotional labels and used the expression given below:
$precision = TP\_X\_light/(TP\_X\_light + FP\_X\_light)$. The result of our computation is presented in Table 8.

Table 8 focuses on providing further understanding of the classifier's ability in predicting the labels correctly. We recorded high precision with DA_TMO in all 3 lights, followed by optimal. These are expected, as was also reported in the confusion matrices above.

Table 1: *Face Recognition rates with naive, NDCT, Opt_exp and TMO datasets (%).*

Data instance	Recognition rate
Naive	82
NDCT	84
Opt_exp	87
Lg_TMO	88
DA_TMO	93

Table 2: *Recognition rate base on individual lighting conditions. BL (back_light), LL (left_light), Ovh (overhead_light) (%).*

	Bck_light	Lft_light	Ovh_light
Naive	75	82	90
NDCT	76	86	89
Opt_exp	78	91	91
Lg_TMO	89	88	86
DA_TMO	91	93	95

Table 3: *FR based on Naive (0^{th}exposure). Recognition rate 82%.*

Data Instance	Bck_light	Lft_light	Ovh_light
Bck_light	**75**	13	12
Lft_light	7	**82**	11
Ovh_light	4	6	**90**

6. Conclusions

This paper has presented the use of HDR in face recognition. Results demonstrate that HDR is beneficial for harsh lighting conditions. Results with the modest database presented show that TMOs

Table 4: *FR based on NDCT. Recognition rate 84%.*

Data Instance	Bck_light	Lft_light	Ovh_light
Bck_light	**76**	6	18
Lft_light	10	**86**	4
Ovh_light	9	2	**89**

Table 5: *FR based on optimal exposure. Recognition rate 87%.*

Data Instance	Bck_light	Lft_light	Ovh_light
Bck_light	**78**	10	12
Lft_light	1	**91**	8
Ovh_light	3	6	**91**

Table 6: *FR based on Lg_TMO. Recognition rate 87.7%.*

Data Instance	Bck_light	Lft_light	Ovh_light
Bck_light	**89**	6	5
Lft_light	8	**88**	4
Ovh_light	8	6	**86**

Table 7: *FR based on DA_TMO. Recognition rate 93%.*

Data Instance	Bck_light	Lft_light	Ovh_light
Bck_light	**91**	0	9
Lft_light	7	**93**	0
Ovh_light	1	4	**95**

Table 8: *Summary of Precision across the 3 lights and 5 datasets (%).*

	Bck_light	**Lft_light**	**Ovh_light**
Naive	87	81	80
NDCT	80	92	80
Opt_exp	95	85	82
Lg_TMO	85	88	91
DA_TMO	92	96	91

can provide a good method of using HDR data, therefore making it possible to make use of HDR on legacy FR systems without requiring complex changes to adapt the methods to HDR.

The dataset chosen however is relatively small and further work is required to consolidate these results. Further work will also investigate facial recognition under complex dynamic lighting as the challenges for quickly changing illumination are even broader and would require the use of more sophisticated HDR methods.

Acknowledgement

Thanks to all colleagues who contributed to this work and the anonymous reviewers for their suggestions. This work is supported by research funding from the Petroleum Technology Development Fund, Nigeria and Debattista is partially supported by a Royal Society Industrial Fellowship.

References

[AHP06] AHONEN T., HADID A., PIETIKAINEN M.: Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 28*, 12 (2006), 2037–2041. 2

[BADC11] BANTERLE F., ARTUSI A., DEBATTISTA K., CHALMERS A.: *Advanced high dynamic range imaging: theory and practice*. CRC Press, 2011. 1

[BETVG08] BAY H., ESS A., TUYTELAARS T., VAN GOOL L.: Speeded-up robust features (surf). *Computer vision and image understanding 110*, 3 (2008), 346–359. 4

[BJ11] BASRI R., JACOBS D.: Illumination modeling for face recognition. In *Handbook of Face Recognition*. Springer, 2011, pp. 169–195. 1

[CYZ*06] CHEN T., YIN W., ZHOU X. S., COMANICIU D., HUANG T. S.: Total variation models for variable lighting face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 28*, 9 (2006), 1519–1524. 2

[DBRS*15] DEBATTISTA K., BASHFORD-ROGERS T., SELMANOVIĆ E., MUKHERJEE R., CHALMERS A.: Optimal exposure compression for high dynamic range content. *The Visual Computer 31*, 6-8 (2015), 1089–1099. 3, 6

[EWMU13] EILERTSEN G., WANAT R., MANTIUK R. K., UNGER J.: Evaluation of tone mapping operators for hdr-video. In *Computer Graphics Forum* (2013), vol. 32, Wiley Online Library, pp. 275–284. 3

[FFFP07] FEI-FEI L., FERGUS R., PERONA P.: Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. *Computer Vision and Image Understanding 106*, 1 (2007), 59–70. 5

[FFP05] FEI-FEI L., PERONA P.: A bayesian hierarchical model for learning natural scene categories. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* (2005), vol. 2, IEEE, pp. 524–531. 4

[FKH] FLORIAN KAINZ ROD BOGART P. S., HILLMAN P.: *Technical Introduction to OpenEXR*. Industrial Light & Magic, Weta Digital. URL: http://openexr.com/TechnicalIntroduction.pdf. 3

[FN09] FRANCO A., NANNI L.: Fusion of classifiers for illumination robust face recognition. *Expert Systems with Applications 36*, 5 (2009), 8946–8954. 5

[GB03] GROSS R., BRAJOVIC V.: An image preprocessing algorithm for illumination invariant face recognition. In *Audio-and Video-Based Biometric Person Authentication* (2003), Springer, pp. 10–18. 2

[GJ10] GOPALAN R., JACOBS D.: Comparing and combining lighting insensitive approaches for face recognition. *Computer Vision and Image Understanding 114*, 1 (2010), 135–145. 5

[GNV11] GOEL T., NEHRA V., VISHWAKARMA V. P.: Comparative analysis of various illumination normalization techniques for face recognition. *International Journal of Computer Applications 28*, 9 (2011). 5

[Gro] GROUP A.: Arri studio lights. https://www.arri.com/lighting/. 3

[IDC16] IGE E. O., DEBATTISTA K., CHALMERS A.: Towards hdr based facial expression recognition under complex lighting. In *Proceedings of the 33rd ACM international conference on Computer Graphics* (2016), ACM, pp. 131–140. 5

[JRW97] JOBSON D. J., RAHMAN Z.-U., WOODELL G. A.: A multi-scale retinex for bridging the gap between color images and the human observation of scenes. *Image Processing, IEEE Transactions on 6*, 7 (1997), 965–976. 2

[LHK05] LEE K.-C., HO J., KRIEGMAN D. J.: Acquiring linear subspaces for face recognition under variable lighting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 27*, 5 (2005), 684–698. 1, 2

[MBDC15] MELO M., BESSA M., DEBATTISTA K., CHALMERS A.: Evaluation of tone-mapping operators for hdr video under different ambient luminance levels. *Computer Graphics Forum 34*, 8 (2015), 38–49. URL: http://dx.doi.org/10.1111/cgf.12606, doi:10.1111/cgf.12606. 3, 6

[MDK08] MANTIUK R., DALY S., KEROFSKY L.: Display adaptive tone mapping. In *ACM SIGGRAPH 2008 Papers* (New York, NY, USA, 2008), SIGGRAPH '08, ACM, pp. 68:1–68:10. URL: http://doi.acm.org/10.1145/1399504.1360667, doi:10.1145/1399504.1360667. 3

[OPAHC*14] OUSSALAH M., PROFESSOR ALI HESSAMI D., CHERMAK L., AOUF N., RICHARDSON M.: Hdr imaging for feature tracking in challenging visibility scenes. *Kybernetes 43*, 8 (2014), 1129–1149. 6

[PMPP14] PEREIRA M., MORENO J.-C., PROENÇA H., PINHEIRO A. M.: Automatic face recognition in hdr imaging. In *SPIE Photonics Europe* (2014), International Society for Optics and Photonics, pp. 913804–913804. 2, 3, 6

[PYL08] PANG Y., YUAN Y., LI X.: Gabor-based region covariance matrices for face recognition. *IEEE Transactions on Circuits and Systems for Video Technology 18*, 7 (2008), 989–993. 2

[Sch95] SCHLICK C.: Quantization techniques for visualization of high dynamic range pictures. In *Photorealistic Rendering Techniques*. Springer, 1995, pp. 7–20. 4

[SKM04] SHORT J., KITTLER J., MESSER K.: A comparison of photometric normalisation algorithms for face verification. In *null* (2004), IEEE, p. 254. 2

[SL09] SOKOLOVA M., LAPALME G.: A systematic analysis of performance measures for classification tasks. *Information Processing & Management 45*, 4 (2009), 427–437. 6

[TT10] TAN X., TRIGGS B.: Enhanced local texture feature sets for face recognition under difficult lighting conditions. *Image Processing, IEEE Transactions on 19*, 6 (2010), 1635–1650. 2

[UMM*10] URBANO C., MAGALHÃČES L., MOURA J., BESSA M., MARCOS A., CHALMERS A.: Tone mapping operators on small screen devices: An evaluation study. *Computer Graphics Forum 29*, 8 (2010), 2469–2478. URL: http://dx.doi.org/10.1111/j.1467-8659.2010.01758.x, doi:10.1111/j.1467-8659.2010.01758.x. 3

[ZBMM06] ZHANG H., BERG A. C., MAIRE M., MALIK J.: Svm-knn: Discriminative nearest neighbor classification for visual category recognition. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)* (2006), vol. 2, IEEE, pp. 2126–2136. 5

[ZSCG07] ZHANG B., SHAN S., CHEN X., GAO W.: Histogram of gabor phase patterns (hgpp): a novel object representation approach for face recognition. *Image Processing, IEEE Transactions on 16*, 1 (2007), 57–68. 2

[ZTC14] ZHANG L., TJONDRONEGORO D., CHANDRAN V.: Facial expression recognition experiments with data from television broadcasts and the world wide web. *Image and Vision Computing 32*, 2 (2014), 107–119. 5