

Inter-image outliers and their application to image classification

Alex Pappachen James^{a,1,*}, Sima Dimitrijev^{a,2}

^a*Griffith School of Engineering, Griffith University, QLD, 4111, Australia*

Abstract

Image variability that is impossible or difficult to restore by intra-image processing, such as the variability caused by occlusions, significantly reduces the performance of image-recognition methods. To address this issue, we propose that the pixels associated with large distances obtained by inter-image pixel-by-pixels comparisons should be considered as inter-image outliers and should be removed from the similarity calculation used for the image classification. When this method is combined with the template-matching method for image recognition, it leads to state-of-the-art recognition performance: 91% with AR database that includes occluded face images, 90% with PUT database that includes pose variations of face images and 100% with EYale B database that includes images with large illumination variation.

Keywords:

Outliers, Image recognition, Template matching

1. Introduction

An outlier in data is broadly defined as an observation or a set of observations that appears to be inconsistent with the remaining data. In data and pattern analysis, large outliers influence the statistics of the analysis and can lead to inconclusive and/or wrong results. There are many methods for detection of outliers, including methods based on statistical data distributions, prior knowledge of the nature of distributions, expected number of outliers, and the nature of expected outliers [1, 2]. Majority of these methods are only applicable to univariate data or to a specific type of data distribution. However, practical applications that involve images are multidimensional and do not always follow a specific distribution. In such situations, detection of outliers becomes difficult and the conventional methods fail to work.

In the past, outliers were used in image analysis to detect changes within images [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13]. However, the idea of outliers has not been explored directly for applications involving image comparison and classification. In this paper, we introduce a new approach for outlier detection that is based on pixel-wise distances obtained by inter-image pixel comparisons. In this approach, no assumptions about the nature of the underlying data is needed. This detection of inter-image outliers is combined with the template-matching method for image recognition and applied to practical examples of face recognition under difficult recognition conditions that include heavy face occlusions and pose variations.

2. Background

2.1. Spatial outliers

The idea of outlier detection has been explored across various fields and applications [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17] and a number of these approaches are applicable to image analysis and recognition

*Corresponding author

Email addresses: a.james@griffith.edu.au (Alex Pappachen James), s.dimitrijev@griffith.edu.au (Sima Dimitrijev)

¹A. P. James is a Research Fellow with Queensland Micro- and Nanotechnology Centre, Griffith University. Tel: +61 07 37358008

²S. Dimitrijev is a Professor with the Griffith School of Engineering, Griffith University.

[14, 3, 4, 5, 6, 7, 8, 9, 10]. Outliers in images are often termed as *spatial outliers* and can be seen to represent neighborhoods of pixels in an image that are significantly different from the rest of the image. Based on this idea, the following is a definition of a spatial outlier [6]: “A spatial outlier is a spatially referenced object whose non-spatial attribute values are significantly different from those of other spatially referenced objects in its spatial neighborhood.”

Spatial outlier detection methods are either based on distances [14, 15, 16, 17], to define spatial neighborhoods, or on graph connectivity [3, 4, 5, 6, 7, 8, 9, 10]. A comprehensive study on distance based methods is presented in [14, 15]. In this study, it is proposed that for a normally distributed random variable T with mean μ_d and standard deviation σ_d , $t \in T$ is an outlier if and only if $\frac{t-\mu_d}{\sigma_d} \geq 3$ or $\frac{t-\mu_d}{\sigma_d} \leq -3$. In the case of Poisson distribution with parameter $\mu = 3.0$, t is an outlier if and only if $t \geq 8$. Using these criteria, the outliers within an image are detected by indexing, nearest neighbor search or through a cell-based approach [14].

Graph based spatial outliers [8, 9, 10] use the statistic $S(x) = [f(x) - E_{y \in N(x)}(f(y))]$ for a data record x with attribute $f(x)$, where $N(x)$ is the fixed cardinality set of neighbors of x and $E_{y \in N(x)}(f(y))$ is the average attribute value for the neighbors of x . In other words, the statistic $S(x)$ denotes the difference between the attribute value of each data object x and the average attribute value of x ’s neighbors. When $f(x)$ is normally distributed, $S(x)$ is also normally distributed. If μ_s and σ_s are the mean and the standard deviation of S , then the outlier $t \in S$ can be defined as $|\frac{t-\mu_s}{\sigma_s}| > 2$. As can be seen in the literature [8, 9], majority of methods define the candidate outliers using an approach based on Chebyshev inequality.

Outliers detected by methods based on local image properties are called local outliers. The method in [7] describes a measure to detect local spatial outliers and shows its relative importance in comparison with global outliers. Detection of local outliers can reveal outliers specific to the local regions within images and is more likely to provide useful outliers than global ones.

The existing methods and applications of outlier detection can be classified as within image detection of spatial outliers. As distinct from this, this paper is focused on the idea of inter-image detection of outliers and its application to image classification.

2.2. Image recognition using templates

Template matching is one of the most immediate approaches applied in multi-class image recognition [18]. The template based methods can be widely classified into rigid templates [19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31] and deformable templates[32, 33]. Rigid templates use templates that undergo simple transformations such as rotation, scaling, and translation. They rely on the proper localization of images and similarity between the images. The similarity between the images is calculated using the sum of differences [25], the correlation coefficients [26], haar transform coefficients [27], Walsh-Hadamard transform coefficients [28], and the invariant moments [29, 30, 31]. With rigid templates, similarity measures tend to be sensitive on improper localization and show severe drop in recognition performance with deformations in object shape. On the other hand, deformable templates have the capacity to deal with shape deformations [32, 33]. The major methods include active contours [34, 35, 36] and analytical and prototype based parametric models [37, 38, 39].

Image recognition using template based methods suffers in performance as a direct result of natural variability in images. Natural variability in images that are used for recognition of faces, objects etc. can be considered as a result of pixel intensity variations that occur due to variations in illumination, view point, and occlusions [40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59]. We can define an image x_l of an object that has no natural variability as the gallery or reference image and an unknown instance of the object with any arbitrary degree of natural variability as test image x_t . In a recognition problem, test image is required to be compared with many gallery images, and ranking the degree of similarity among all comparisons results in the identification of the class of the test image with respect to the known set of gallery images. The presence of natural variability introduces two problems: (1) reduction in intra-class similarity and (2) increase in inter-class similarity. To achieve class discrimination, two objectives that we need to achieve are (1) to increase similarity between intra-class images and (2) to reduce similarity between inter-class images. To achieve these two objectives of class discrimination, we apply the idea of inter-image outlier detection and explore its role in enhancing the performance of the often disregarded rigid template matching method to a new competitive level.

3. Application of outlier detection in image classification

A new method for detecting distance based outliers in application to image classification is presented in this section. In this type of comparisons, pixel to pixel differences are used to calculate a distance vector. When two intra-class images taken under same photographic conditions are compared, the distances will be zero (the highest possible similarity). For an inter-class comparison, two images under the same photographic conditions show non-zero distances (low similarity).

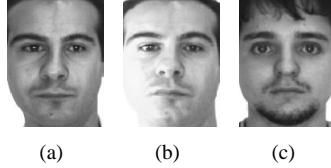


Figure 1: Examples of three images from AR database [60, 61] showing inter-class and intra-class images: (a) neutral image of a person A, (b) neutral image with changed illumination conditions of person A, and (c) neutral image of person B. The comparison of gallery image A to that of B is an example of inter-class comparison, whereas the comparison of test image A with gallery image A is an example of intra-class comparison.

Examples of inter-class and intra-class face images are shown in Fig. 1(a), 1(b) and 1(c) with the corresponding inter-class and intra-class scatter plots shown in Fig. 2(a) and 2(b). The central lines in this type of scatter plots show the reference points where the distances between two pixels are zero. The distances of the pixels from the center line can be used as measures for detection of the outliers. The pixels that fall beyond a predefined distance threshold (θ) can be defined as outliers. In this paper, the threshold value is fixed at $\theta = 2\sigma$, where σ is the overall standard deviation of the inter-class pixels (Fig. 2(a)). It can be noted that the threshold value can also be determined empirically such that differences from the inter-class comparisons become always greater than the intra-class comparisons under the same photographic conditions. However, as can be seen from Fig. 2(b), when the photographic conditions change (e.g. due to illumination), there can be large differences between the images. The test image that is used in this illustration contains increased amount of light intensity variation and an increased intra-class variability (densities depart from the center line). These large differences, if detected as outliers, would be incorrect because these differences represent false differences that are caused by the change in illumination. Since identity information of a pixel in an image should not be influenced by illumination, there should be a method to remove the effect of illumination before any comparison is made.

Intensity offsets that occur across regions of pixels depend on the brightness and strength of the source of light in the environment. To reduce the large differences that are due to comparisons of illuminated pixels and normal pixels, normalization of the pixels can be performed by using the local image information. One of the first techniques employed to normalize data in statistics is to normalize the elements in a vector by adjusting its mean to zero and its standard deviation to one. Since the intensity offsets introduced by illumination are nonuniform globally and approximately uniform at local regions, the key idea is to remove such offsets by applying local normalization in images.

Definition 1. Local Gaussian normalized image y formed from a raw intensity image x with a size of $U \times V$ pixels with local standard deviation σ_g and local mean \bar{x}_g of the image pixels encompassed by a window w consisting of $M \times N$ pixels is defined as $y(i, j) = \frac{x(i, j) - \bar{x}_g(i, j)}{\sigma_g(i, j)}$

In Def. 1, the local standard deviation $\sigma_g(i, j)$ is calculated by the following equation:

$$\sigma_g(i, j) = \sqrt{\frac{1}{MN} \sum_{z=-a}^a \sum_{t=-b}^b [x(i+z, j+t) - \bar{x}_g(i, j)]^2} \quad (1)$$

where $a = (M - 1)/2$ and $b = (N - 1)/2$. The value of local mean $\bar{x}(i, j)$ is calculated by the following equation:

$$\bar{x}_g(i, j) = \frac{1}{MN} \sum_{s=-a}^a \sum_{t=-b}^b x(i+s, j+t) \quad (2)$$

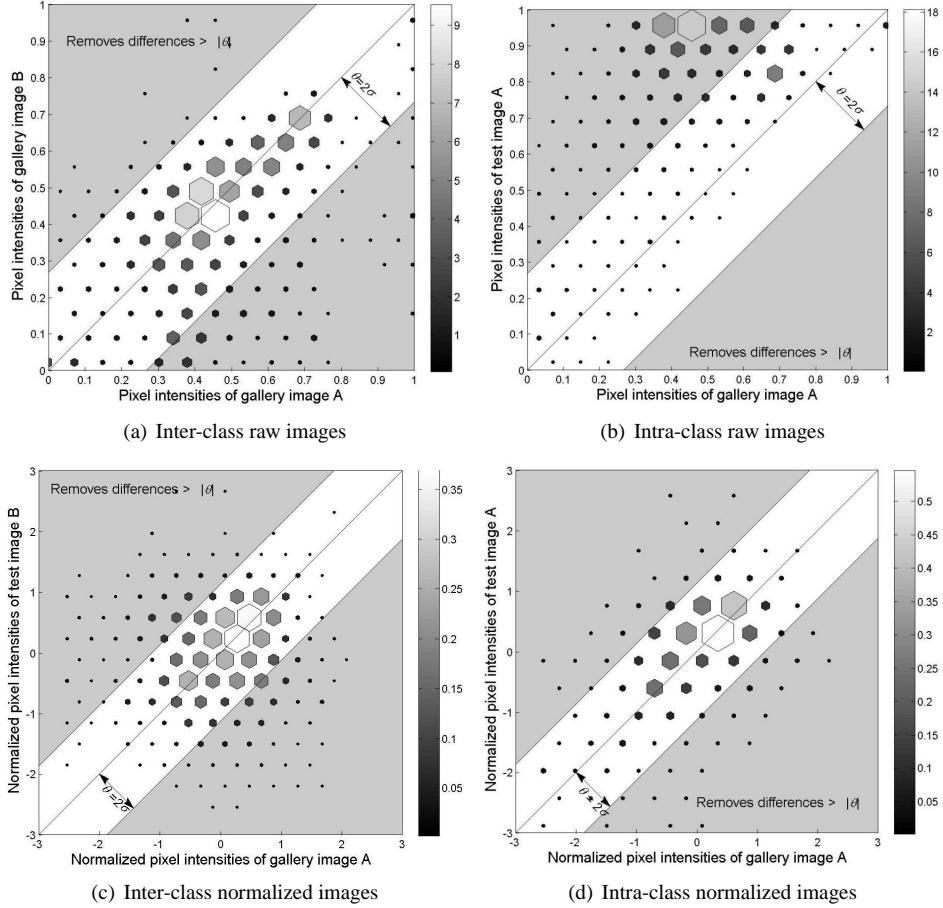


Figure 2: Scatter plots illustrating graphically the issues in inter-class and intra-class comparisons using raw images shown in Fig. 1. The gray shaded regions in the graphs illustrate the outliers selected based on threshold θ . The vertical gray colored bars associated with each plot show densities as described in [62]. Inter-class comparisons using raw images are shown in (a), while (b) illustrates a problem with intra-class comparisons using raw images having illumination that result in false differences being detected as incorrect outliers. The scatter plots for inter-class comparison using normalised images are shown in (c), while (d) shows the role of normalisation in correctly detecting outliers by reducing the differences due to illumination changes for intra-class image comparisons.

To illustrate the effect of normalisation on images, the images shown in Fig. 1 are normalized and drawn as the scatter plots shown in Fig. 2(c) and 2(d). The idea of normalization is to bring the pixel groups closer to the center line in the scatter plot. As an another illustrative example, Fig. 3(a) and 3(b) show the important role that normalisation plays in compensating large variations in illumination. It can be seen that the proposed normalization helps to bring out features that are otherwise hardly visible by human eye because of difficult illumination conditions.

The distance of the pixels from the center line can be used as a measure to detect the outliers. The distance between gallery image pixel $y_g(i, j)$ and test image pixel $y_t(i, j)$ can be calculated using simple trigonometry as shown in Fig. 4. Since any point in the line following SO would mean that the gallery image pixel $y_g(i, j)$ and test image pixel $y_t(i, j)$ to take same values, we have $\left| \frac{y_g(i, j)}{y_t(i, j)} \right| = 1$ or $|y_g(i, j) - y_t(i, j)| = 0$. However, the point D will have $\left| \frac{y_g(i, j)}{y_t(i, j)} \right| > 1$ or $|y_g(i, j) - y_t(i, j)| > 0$. The distance between the points S and D is denoted as $\delta' (= |y_g(i, j) - y_t(i, j)|)$, while that between O and E is denoted as δ . Since the slope of the line SO is always 1, $\angle OSD = 45^\circ$, and as the line OD is perpendicular to the x-axis, $\angle ODS = 90^\circ$ and since the acute angles of a right triangle are complementary, $\angle SOD = 45^\circ$. Further, as the lines SO and DE is parallel to each other, $\angle EOS = 90^\circ$, forcing $\angle SOD = \angle DOE = \angle ODE = 45^\circ$. This would mean, $\triangle OED$ to be isosceles, and hence length of OE to be equal to length of DE, represented as δ . Applying



(a) Raw images



(b) Normalised images

Figure 3: Example of raw and normalised test images. The raw images shown in (a) contain features in a human face that are hard to observe by human eye. The application of the proposed normalisation as shown in (b) brings out those features that are otherwise not visible to the human eye.

Pythagoras theorem on $\triangle OED$, $\delta'^2 = \delta^2 + \delta^2$ or $\delta = \frac{\delta'}{\sqrt{2}}$ and can be represented as:

$$\delta(i, j) = \frac{|y_g(i, j) - y_t(i, j)|}{\sqrt{2}} \quad (3)$$

The differences between the individual pixels becomes high when there are occlusions in the test image. The differences are also large for the comparisons between inter-class images. In the ideal case one would prefer to remove the differences due to occlusions without removing any difference due to inter-class comparison to achieve maximum class discrimination. It can be noted that in essence the removal of difference as a result of occlusions would also remove the difference from inter-class comparisons. This essentially means we are faced with the need to have an optimal criterion for difference removal such that the removal of difference from occlusions outweighs the removal of inter-class differences.

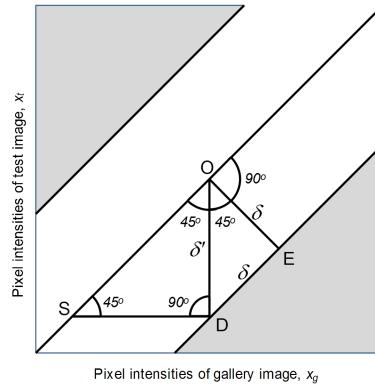


Figure 4: Graphical illustration for the calculation of the distance δ using simple trigonometric relationships. The value of δ is proportional to σ , where σ is the net standard deviation of inter-class differences between the gallery images.

Definition 2. The distance $\delta(i, j)$ between the Gaussian normalized gallery pixel $y_g(i, j)$ and test pixel $y_t(i, j)$ can be used to determine the outlier $t_t \in y_t$ if and only if $\frac{\delta(i, j)}{\sigma_{avg}} > 2$, where the outlier detection threshold is $\theta = 2\sigma_{avg}$ and σ_{avg} is the average standard deviation among all inter-class gallery distances.

The criterion from Def. 2 corresponds to removal of 5% of the inter-class differences. It should be noted that the percentage of removed intra-class differences is much smaller when natural variability is limited, and this percentage of removed intra-class differences can increase to a larger value when there are increased levels of natural variability.

A graphical illustration showing the removal of large differences in a Gaussian normalized images using this threshold θ is shown in Fig. 2c and 2d. The remaining differences that fall within the boundaries of θ , as shown in Fig. 2c and 2d, can be represented as:

$$\delta'(i, j) = \begin{cases} \delta(i, j) & \text{if } \delta(i, j) \leq \theta \\ \phi & \text{if } \delta(i, j) > \theta \end{cases} \quad (4)$$

where formally ϕ represents the *empty* element as all $\delta(i, j)$'s greater than θ are removed. The main idea behind this approach is based on the assumption that in comparison with ordinary inter-class differences, the occlusions will have an increased number of large differences that fall outside 2σ region. By removing such outliers, we can ensure that the differences from occlusions do not interfere with the overall similarity calculation for a comparison between two images. Figure 5 shows the regions where outliers are detected using the distance based threshold approach presented in this paper using two examples of intra-class image comparisons. It can be seen from Fig. 5(b) and 5(d) that since the threshold does not relate to a specific nature of occlusion, any large difference that occur due to a genuine mismatch (that come from variations in illumination, facial expressions and facial pose) between the pixels is also removed.

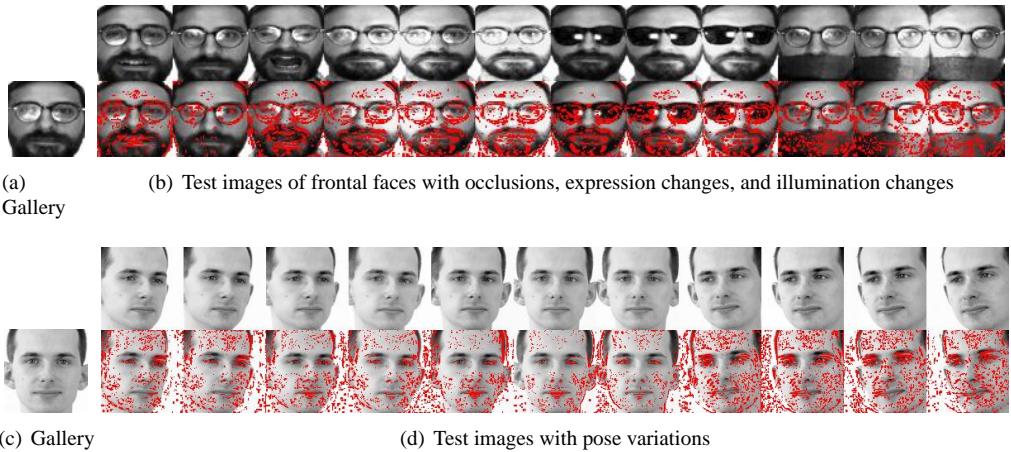


Figure 5: Detected inter-image outliers are marked in red color in two examples of intra-class image comparisons. In the first example, the gallery image (a) is used for intra-class inter-image comparison with test images shown in (b) to visualise the detected outliers with illumination changes, facial expression changes and occlusions. Similarly, in the second example, the gallery image (c) of a person in a frontal pose is compared with the test images shown in (d) to visualise the detected outliers in test images with variations in pose.

4. Image recognition algorithm

The presented method of removal of large differences relies on pixel-to-pixel inter-image comparisons. Each pixel in a test image is compared with the corresponding pixels from the gallery images along the same spatial coordinate in the images. This is very similar to a template comparison where a major concern is proper localization of features [18].

4.1. Image Alignment by Feature Localization

Automatic object or face detection and alignment is a difficult problem when natural variability in images is high. In any method that is based on pixel-by-pixel comparisons, it is essential that the features of the compared images are well aligned. For example, irrespective of the face detection method employed, natural variability can cause pixel-level misalignments. To compensate for the localization errors that occur after an automatic or manual alignment, we apply either test or gallery image shifts with respect to a set of registration points in the feature matrices. For example, the localization of face images can be achieved by detecting the location of eye coordinates. An error in localization means the eye coordinates are shifted. A scale error means that the eye coordinates are shifted toward

each other or away from each other. A rotation error causes shifts of the two eye coordinates in opposite vertical directions. We perturb the reference eye coordinates by applying such shifts and re-localize the face images using the shifted eye coordinates. After each perturbation, the feature matrix of the image is cropped to size of $U \times V$. In total, we apply $r \times c$ shifts for each eye coordinate, which generates $p = 2rc$ synthetic images labeled as x^p . For example, for $r = c = 3$, there will be 18 shifted feature matrices generated from a single image x for each registration point used in the process of perturbation.

4.2. Algorithm summary

The following are the steps in the proposed image recognition algorithm:

1. Generate x_t^p synthetic test images from a test image x_t .
2. Normalize the test x_t^p and gallery x_g images using *Def. 1*.
3. Calculate the pixel-wise distance $\delta^k(i, j)$ between test $y_t^p(i, j)$ and gallery images $y_g(i, j)$ using Eq. (3).
4. Remove the outliers by applying a single-sided threshold θ to the difference vector $\delta^p(i, j)$ to obtain $\delta'^p(i, j)$ using Eq. (4).
5. Calculate the mean distance $\bar{\delta}'^p$ by taking mean of the difference vector with the removed outliers, $\delta'^p(i, j)$.
6. Repeat steps 1 to 5 for all the comparisons between a test image and the gallery images.
7. The mean distances $\bar{\delta}'^p$ are ranked and the minimum value selected for determining the best match.

5. Parameter selection

The parameters used in the presented algorithm are: (1) the size of the normalization filter window $M \times N$ pixels, (2) the inter-image outlier threshold, θ , and (3) the perturbation parameters $r \times c$ pixels. To determine the values of these heuristics, 50 neutral face images are selected from the AR database for forming the gallery set and 50 neutral face images from another session are selected as training set. Figure 6(a) and 6(b) shows example gallery and test images used in this analysis. The gallery and test image size was kept at 60×60 pixels. A recognition experiment with this training set will be an easy task as it only includes limited effects of natural variability. Further, for parameter selection, it is unfair to use original images from the database having a range of natural variability as this knowledge of exact test conditions may have a bias on recognition performance of the testing methods. Due to this reason, a dataset is created by adding different type of synthetic variabilities to the training images. This dataset is grouped based on 3 different types of variability, test set A having different types of occlusions, test set B having different illumination changes, and test set C having variable blurring effects. Using the training images and overlaying the 6 occlusions as shown in Fig. 6(c), the test set A shown in Fig. 6(d) is created. Here occlusions are fixed to a size of 30×60 pixels and occludes 50% of test images. Figure 6(e) shows the test set B created again from the training image by applying intensity offsets of $\pm 0.1, \pm 0.2$ and ± 0.3 from the training image intensity values that are bounded within $[0, 1]$. Figure 6(f) shows the test set C created from the training image by applying a circular averaging filter within the square matrix of sides 3, 5, 7, 9 and 11. Test set C images tries to simulate the situation when the resolution on gallery and test images are drastically very different or when the test image camera is out of focus introducing blurring. Since the variations introduced in this test data sets is high to the extent that it maybe unrecoverable using any known techniques, it is expected that the performance of this small dataset against any image classification method to never reach 100%. This generated synthetic training data will be used in our experiments to optimise the parameters of the proposed method.

5.1. Optimization of normalization window and threshold

Using all the images in the synthetic dataset described in Fig. 6, recognition experiments are performed to optimize the value of $M \times N$. To reduce the number of parameters, we keep the size of the length and width of the images to be same, $M = N$. In a methodological point of view, if M and N needs to be determined separately because of unequal width and height of test image, the same method of independent determination of M can be extended to determine N in such two dimensional scenarios.

Figure 7 shows the impact of normalization window size and threshold on the the recognition performance of the presented method. To avoid any bias as a result of localization error compensation, the values of perturbation

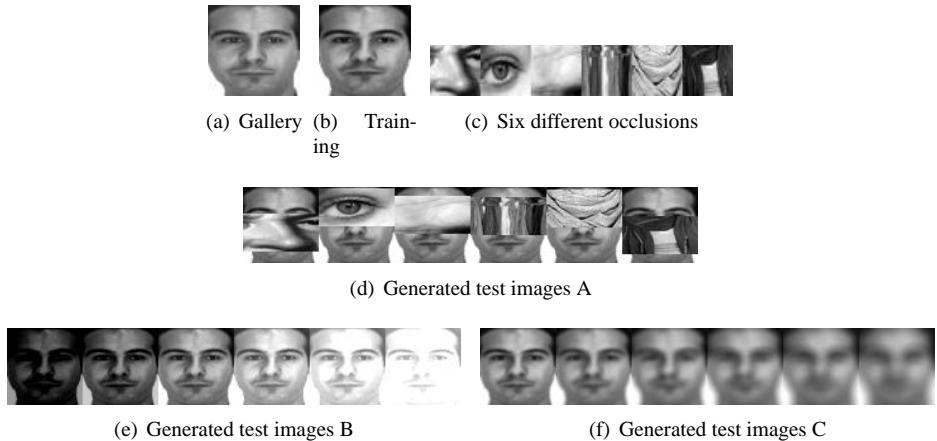


Figure 6: To make the comparison between gallery image (a) and the training image (b) a difficult recognition task, a set of synthetic test images is generated. Using a set of predefined non monotonic occlusions shown in (c), a set of images is generated as shown in (d). The six different types of occlusions from images of nose, eye, hand, towel, scarf, and shirt are used to generate the test images by overlaying occlusions that cover 50% of the training image size. Further, by adding intensity offsets to the training images, another set of test images is generated as shown in (e), and finally, by applying circular averaging filter on the training image, a set of test images shown in (f) is generated to account for the effects of low resolution images.

parameters are set at zero ($r = c = 0$). Using Fig. 7(a), it is determined that the normalization window size is robust within the range of 5×5 to 15×15 pixels. It was also observed from Fig. 7(a) that a low threshold of σ considerably reduces the recognition rate, while around 2σ the recognition accuracies peaked independently of the size of the normalization window. Figure 7(b) shows the impact of threshold variation on recognition performance of the presented method. It is observed from the Fig. 7(b) that the recognition performance of the method is robust near the threshold range from 1.5σ to 2σ . Further, Fig. 7(b) shows a clear indication that the detection of outliers improves recognition performance even with application of a wide range of threshold values. Figure 8, shows a generalisation on stability of normalisation window with respect to the size of the image. The shaded region shows the range of normalization window sizes as a percentage of image size where the proposed method yields stable recognition performance. Although, this wide range of stable normalization window are shown for an optimal inter-image outlier threshold of 2σ , this approach to determine the normalization window size can be extended to include any values of threshold.

5.2. Optimization of perturbation parameters

A perfectly aligned gallery and test images will be devoid of localization errors and does not require the need for perturbations. Since such pixel level alignments are impossible to achieve even with meticulous calculations, an exact optimal value of r and c can vary from one database to another. Figure 9 shows the role of perturbation in correcting the localization errors as indicated by the increase in recognition performance. It is observed that for the dataset described in Fig. 6, any value of the perturbation parameter greater than 3×3 gives considerably better results than the one without perturbation. Although, the performance improvement can vary from a database to another, any correction in localization errors corrects for large number of pixel mismatches and results in better recognition performance.

6. Experimental verification

The recognition of human-face images under natural variability and limited training data represents a difficult recognition problem. Majority of the 2D face image recognition algorithms use machine learning that involves the use of several training instances to create gallery features or models [42, 43, 44, 45, 63]. However, when the number of training samples is limited, such as in single sample mug-shot databases, a straightforward use of the machine learning schemes is not useful. Since this is an interesting and practical problem, we include a study on image to

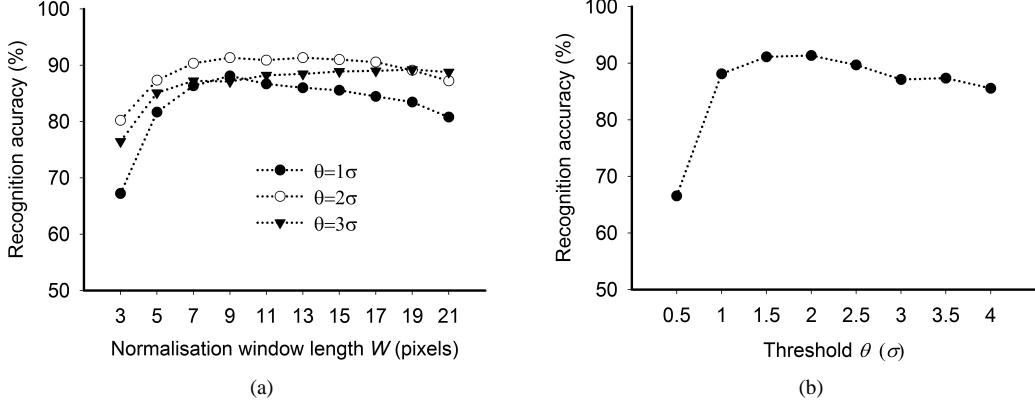


Figure 7: Dependence of recognition performance on (a) normalization window size and (b) threshold of the presented method. Using an image with size of 60×60 pixels it is shown in (a) that the normalization window is stable in a wide range from 5×5 pixels to 15×15 pixels within 1% change in the top recognition accuracy. Graph (b) shows that the maximum recognition accuracies are obtained in the range from $\theta = 1.5\sigma$ to $\theta = 2\sigma$ for set normalisation window size of 9×9 pixels. The values of r and c are kept at zero in these experiments.

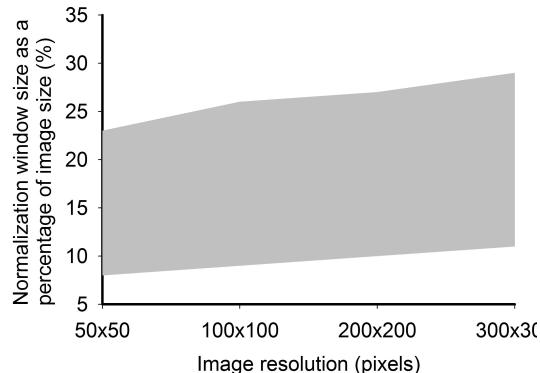


Figure 8: Graphical illustration showing the stable range of values for the normalization window size as a percentage of image resolution that fall within 1% change in the top recognition accuracy. The values of r and c are kept at zero in these experiments and the threshold θ is set to 2σ .

image matching accuracy that do not require extensive machine learning and will enable us to verify the baseline use of the outlier-detection method in image recognition.

AR [60, 61], Extended Yale B (Eyale B) [65] and PUT [64] face databases are used in our study because they include different numbers of test images under a wide range of natural conditions, including heavy occlusions, facial expressions, illumination variations, and pose variations (see Fig. 10, Fig. 11 and Fig. 12). AR database is used to test the performance of the proposed method against the frontal face image recognition using single gallery image per person, PUT database is used to test the performance of the proposed method against large changes in face pose, while Eyale B database is used to test the performance of the proposed method against large changes in illumination conditions.

The images in the AR database are localized by the location of the eye coordinates and cropped to a size of 160×120 pixels. The color information is disregarded by converting it to the corresponding gray levels. For the experiments, we randomly select images of 100 persons from a pool of images from 126 persons. From the 26 images taken over two sessions representing 13 conditions for each person, we select one neutral image from a session as gallery image and remaining images as test images (see Fig. 10(a) and 10(b)).

Eyale B database [65] is a benchmark database used to test the performance of a face recognition method against large illumination changes. The Yale B database consists of 5850 facial images of 10 persons, which has 585 images

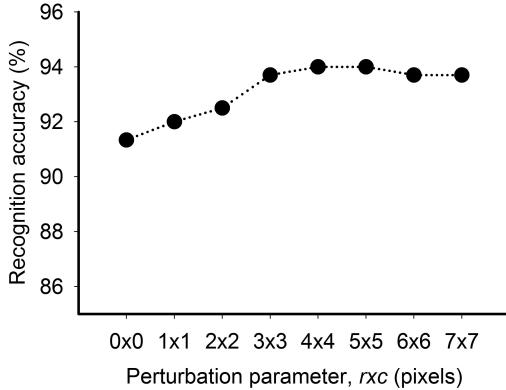


Figure 9: Dependence of the recognition accuracy on the values of perturbation parameters r and c . The outlier threshold is set at 2σ and normalization window in kept at a lower bound of 9×9 pixels.

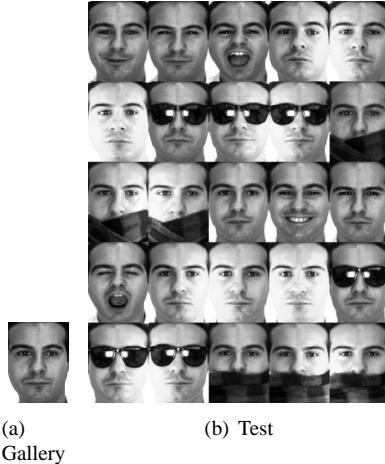


Figure 10: Example images from AR database. Face image with neural facial expression and frontal pose shown in (a) is used as the gallery image. The test images shown in (b) correspond to different conditions including variations in facial expressions, illumination, and occlusions .

per person having nine poses under 65 illumination conditions with an image resolution of 480×640 pixels. Images with 45 out of 65 conditions were usually reported in previous works [66]. The angle of the light source (12° , 25° , 50° , and 77°) is used to divide the database into four subsets. These divided subsets 1, 2, 3, and 4 consists of 70, 120, 120, and 140 images per pose, respectively. The remaining images with larger variations than that of subsets 1-4 are denoted as subset 5. Subset 5 is not often used to report results, however it contains the largest variations in illumination in comparison with subset 1. Example set of images from each of these subsets are shown in Fig 11. To report the performance of proposed method across all the subsets against illumination only variations, the pose variations are kept fixed. In our experiments, 1 image of each person per pose from subset 1 are selected for gallery and the remaining 65 images are used for testing. In other words, we consider all available images in the database representing different illumination conditions for testing. The total numbers of gallery and test images are 90 (10 persons, 9 poses, 1 images) and 5760 (10 persons, 9 poses, 64 images), respectively.

The images from PUT database are also aligned and cropped based on the ground truth eye coordinates to a size of 360×360 pixels. The database contains 9971 images of 100 people. The images are taken in partially controlled illumination conditions over a uniform background. The database has a face appearance variations in head pose. Figures 11(a) and 11(b) show sample images from the galley and test set from the PUT database. The gallery set consists of 2200 images (22 images of each person). For each person, this set contains the frontal view from the

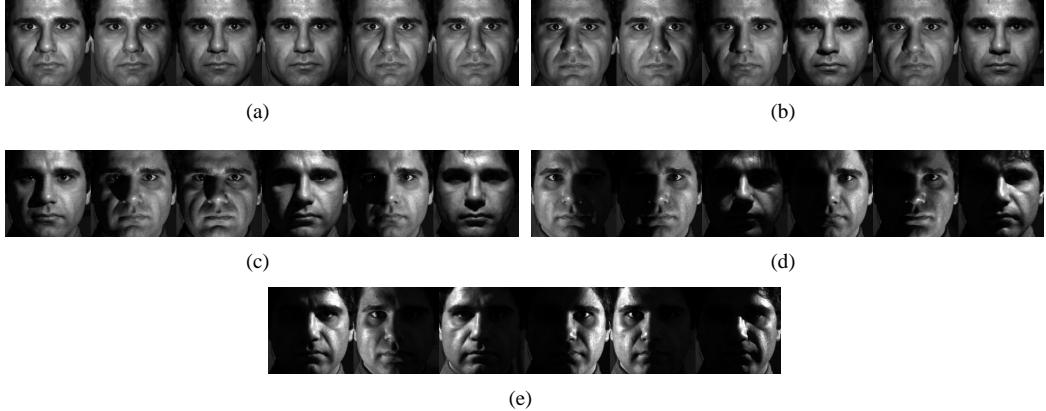


Figure 11: Example images from EYale B database. The images shown in (a), (b), (c), (d), and (e) have increasing degree of illumination variation and are selected from subsets 1, 2, 3, 4, and 5 respectively.

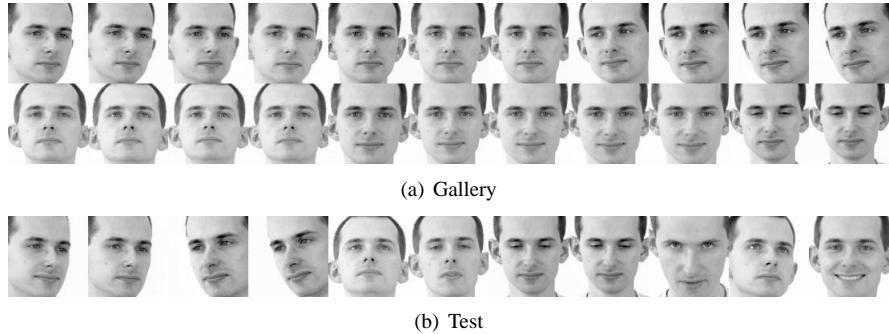


Figure 12: Example images from PUT database. Face images with controlled changes in pose shown in (a) are used to form the gallery set. The test images shown in (b) provide combinations of controlled and uncontrolled variations in pose and facial expressions.

first sequence, 5 images with gradually increasing turn to the left from the central position, 5 images with gradually increasing turn to the right from the central position, frontal view from the second sequence, 5 images with gradually increasing tilt over the central position, 5 images with gradually increasing tilt below the central position [64]. The test set contains the following images of each person: 2 images with head turned more to the left than in any image in the gallery, 2 images with head turned more to the right than in any image in the gallery, 2 images with head raised more than in any image in the gallery, 2 images with head lowered more than in any image in the gallery, 3 frontal face images with unconstrained pose and facial expression [64].

For all the results presented in this paper, the size of the normalization filter is approximated to a lower bound of $U/10 \times V/10$ pixels (determined from Fig. 8), where $U \times V$ is the size of the image, that falls in the stable region of filter window. The values of r and c are set at 5. The threshold θ is kept at a constant value of 2σ .

6.1. Recognition experiments

The overall face-recognition results for the AR, EYale B and PUT databases are shown in Table 1. In Table 1, we also compare the recognition performance of the presented method in the absence and presence of the outlier detection and removal. It should be emphasised that the results reported in Table 1 include the subsets 1-5 in the EYale B database (as distinct from this most conventional approaches use only subsets 1-4, and contained limited variations in illumination). By detecting and removing outliers for similarity calculation, the false similarities from mismatched regions are removed, resulting in improved accuracies. Table 2 shows the recognition performance of the presented method against different conditions in AR database. The high recognition accuracies across wide range of testing conditions shows the robustness of the presented method against natural variability in images.

Table 3 shows the comparison of the presented method with that of conventional and immediate template matching methods such as correlation and distance based nearest-neighbour (N-N) methods. To maintain a fair comparison the presented normalisation scheme used in this paper is applied to both N-N methods. The conventional methods such as correlation or distance based N-N methods do not detect outliers, which results in false similarity calculations and reduced recognition accuracy. Table 4 shows the comparison of the presented method with other reported results from major benchmark recognition methods [67, 68, 69, 70, 71, 72, 73]. To maintain a fair comparison, the test conditions for the presented method were kept the same to that of the reported results in the literature. It can be seen that the presented method shows increased recognition performance in comparison with the reported benchmark methods across different testing conditions and databases.

Table 1: Performance improvement by detecting outliers

Database	Recognition accuracy (%)	
	With detection of outliers	Without detection of outlier
AR	91	80
PUT	90	79
EYALE	100	98

Table 2: Recognition results for the AR database

Condition ^a	Recognition accuracy (%)
Session 1	
Expression	99
Illumination	97
Eye occlusion	100
Eye occlusion and illumination	95
Mouth occlusion	97
Mouth occlusion and illumination	93
Session 2	
Neutral	100
Expression	86
Illumination	85
Eye occlusion	90
Eye occlusion and illumination	77
Mouth occlusion	89
Mouth occlusion and illumination	78
Overall accuracy	91

^a The neutral image of each person in session 1 is used as the gallery image. The remaining 25 images with various photographic and facial conditions are used as test images.

It can be seen from the recognition accuracies that the removal of outliers helps to reduce the effect of natural variability in the images. Occlusions can affect any part of an image and result in information loss that cannot be recovered causing it to be a serious problem in face recognition [68, 42, 43, 44, 45]. However, the removal of outliers

Table 3: Comparison of presented method with other immediate template matching methods (the image normalization proposed in this paper is applied in all cases)

Databases	Recognition accuracy (%)		
	Presented method	Correlation method	N-N method
AR	91	76	80
PUT	90	78	79
EYALE	100	97	98

Table 4: Performance comparison with benchmark methods using various testing conditions

Database	Benchmark Method	Recognition accuracy (%)		Notes
		Benchmark method	Presented method	
AR	SVM+PCA[71]	92.7	99.6	300 face images (12 per individual), with changes in facial expression and illumination
	SVM+ICA[71]	94.0	99.6	
	LEM[72]	96.0	100.0	
	PCA[69]	32.0	91.0	126 frontal faces
EYALE	Eigenface [69]	51.2	100.0	100 frontal faces gallery, 2500 test images Images from subset 1 were used for gallery and subsets 2-4 were used for testing.
	LDA[70]	64.4	100.0	
	Fisherface [74]	66.6	100.0	
	KSLDA[73]	81.4	100.0	
	LCE[66]	98.5	100.0	
PUT	LDA[70]	71.3	90.1	22 gallery images and 11 test images per person, for a total of 100 persons.
	PCA[69]	64.1	90.1	
	ICA[75]	65.2	90.1	

in the occluded areas ensures that only useful information is used in the process of similarity calculation.

Illumination, facial expressions, and other random variabilities such as pose that occur in face images also cause information loss that can affect the performance of face recognition systems. In terms of the presented method, however, these effects can be viewed as different forms of occlusions. This explains the demonstrated robustness of the outlier detection across these very different conditions (Table 1).

6.2. Inter-image variability and outlier statistics

Inter-image variabilities can result from the occurrence of natural variabilities in images such as occlusion, pose and illumination variations. To set a numerical benchmark on the performance of the presented method against occlusion, we select 100 images with neutral frontal face images each from AR and PUT across two sessions. The first 200 images selected from a single session form the gallery and next 200 images selected from a different session forms the test. It is expected that such a selection makes the class discrimination easy, however the addition of occlusions on such images makes the classification problem extremely difficult. Occlusions shown in Fig. 6(c) are applied to the test images to generate images with synthetic occlusion with various sizes. Occlusions with the same size is applied several times at different spatial locations to generate over 100 occluded images. Figure 13 shows the dependence of percentage of detected outliers on the percentage of occlusion with respect to the image size. It is expected that any increase in occlusion size would result in a larger number of large inter-pixel distances and hence in

increased number of detected outliers. With an increase in size of occlusion from 1% to 49%, the percentage number of outliers detected increases from $9.2 \pm 4.2\%$ to $23.5 \pm 4.5\%$.

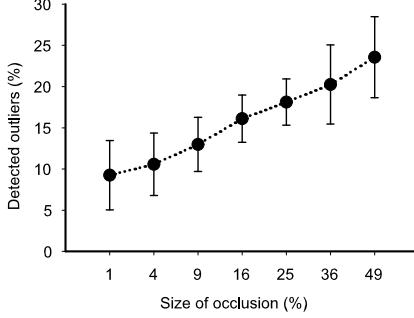


Figure 13: The percentage of detected outliers depends linearly on the occlusion size.

Figure 14 shows the percentage of total number of outliers with respect to the total number of pixels in the image for images in various subsets of EYale B database. It is observed that an increased level of illumination change from subsets 1 to 5 (as shown in Fig. 12), increases the percentage number of detected outliers. EYale B database contains images with large variation in illumination making it suitable to study the statistics of outliers under illumination changes. As shown in Fig. 3(a) and 3(b) normalisation helps to bring out details in image that are otherwise not clearly visible to human eye, it is expected that by such normalisation scheme when illumination changes in an image is fixed, the percentage of outliers in an image that gets detected would be less than that could occur with images having same number of pixels affected with variabilities such as occlusions. Accordingly, the results in Fig. 14 also shows that increased natural variability can result in increased loss of information that is unrecoverable using normalization and results in more number of detected outliers.

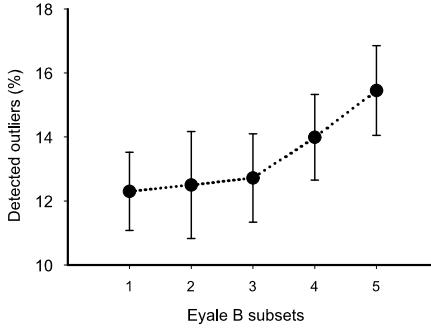


Figure 14: The percentage of detected outliers increases with increased levels of illumination change (subset 1 correspond to the lowest illumination change while subset 5 corresponds to the highest illumination change).

PUT database contains face pose variations in test set with reference to the face pose contained in the gallery images. Figure 15 is a graphical illustration of the percentage number of inter-image outliers detected with reference to pose variations in the test images. A single gallery image having a frontal face pose is used in this experiment. It can be seen that an increased variation in pose with respect to the frontal pose results in increased number of outliers. The asymmetry in the range of values across pose conditions shown in Fig. 15 merely indicates the nature of images in the database, for example a broader range implies a larger variation in the pose within a pose condition. Further, when multiple gallery images in Fig. 11(a) are tested against images in Fig. 11(b), the total percentage of outliers detected for 2 test images with head turned more to the left than in any image in the gallery, 2 test images with head turned more to the right than in any image in the gallery, 2 test images with head raised more than in any image in the gallery, 2 test images with head lowered more than in any image in the gallery and 3 frontal face test images with unconstrained

pose and facial expression, are $15.8 \pm 1.4\%$, $21.3 \pm 2.7\%$, $15.1 \pm 4.7\%$, $11.1 \pm 7.4\%$ and $15.8 \pm 1.4\%$, respectively. This shows an overall decrease in the number of outliers when using multiple gallery images than single gallery image per person; $22.7 \pm 6.1\%$ using single gallery image per person and $15.8 \pm 3.2\%$ with multiple gallery image per person when tested with the same set of test images. With either single or multiple gallery images per person, the presented results reiterate the idea that increased variation in pose between the gallery and test images introduces increased levels of outlier numbers. It may be noted that compensation schemes such as multiple number of gallery images can always result in increased recognition performance as there will be an increase in outlier numbers resulting from inter-class image comparisons and decrease in outlier numbers resulting from intra-class image comparisons.

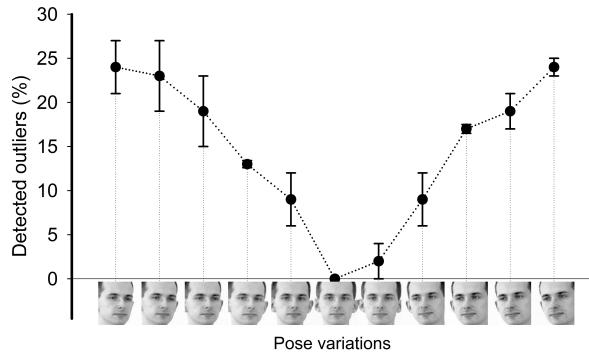


Figure 15: The outliers detected with images from PUT database having varying poses from left to right and right to left directions show an increasing trend with increased levels of pose change.

7. Conclusion

This paper presents a simple method to detect local outliers in images by inter-image comparisons. The important role of normalization for proper detection of the inter-image outliers is also addressed. When combined with the image-recognition method based on rigid template matching and applied to face recognition problems that include heavy occlusions, large illumination changes, and pose variations, state-of-the-art recognition performance is achieved. The fact that such accuracy rates can be achieved across a very wide range of different conditions with this simple method demonstrates that the idea of outlier detection and removal is of fundamental importance for a robust image matching and recognition. This could prove very useful in a broad range of applications, from multi-dimensional clustering and classification tasks to video processing applications such as activity monitoring, motion detection, fraud detection, biometric scanners, and object searching and verification.

Acknowledgements

Funding support by Qs Semiconductor is gratefully acknowledged. The authors wish to acknowledge the anonymous reviewers whose constructive peer-review comments has helped to improve the paper.

References

- [1] V. Barnett, T. Lewis, *Outliers in Statistical Data*, John Wiley and Sons, 1994.
- [2] D. Hawkins, *Identification of outliers*, Chapman and Hall, 1980.
- [3] C.-T. Lu, D. Chen, Y. Kou, Algorithms for spatial outlier detection, in: *Third IEEE International Conference on Data Mining*, p. 597.
- [4] V. P. Jeneja, V. Atluri, Spatial outlier detection in heterogeneous neighborhoods, *Intelligent Data Analysis* 13 (2009) 85–107.
- [5] Z.-Q. Wang, S.-K. Wang, T. Hong, X.-H. Wan, A spatial outlier detection algorithm based multi-attributive correlation, in: *Machine Learning and Cybernetics*, 2004. Proceedings of 2004 International Conference on, pp. 1727–1732.
- [6] S. Shekhar, C.-T. Lu, P. Zhang, S. Shekhar, C. T. Lu, P. Zhang, A unified approach to spatial outliers detection, *GeoInformatica* 7 (2003) 139–166.
- [7] S. Chawla, P. Sun, Slom a new measure for local spatial outliers, *Knowledge and Information Systems* 9 (2006) 412–429.

- [8] Y. Kou, C. T. Lu, R. F. DosSantos, Spatial outlier detection: A graph-based approach, in: Proceedings of the 19th International Conference on Tools with Artificial Intelligence.
- [9] S. Shekhar, C.-T. Lu, P. Zhang, Detecting graph-based spatial outliers: algorithms and applications (a summary of results), in: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 371–376.
- [10] M. Agyemang, K. Barker, R. Alhajj, A comprehensive survey of numeric and symbolic outlier mining techniques, *Intelligent Data Analysis* 10 (2006) 521–538.
- [11] M. E. abd H-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, in: E. Simondis, J. Han, U. Fayyad (eds) Proc. KDD, AAAI Press, Menlo Park CA, 1996, pp. 226–231.
- [12] R. Ng, J. Han, Efficient and effective clustering methods for spatial data mining, in: Proceedings of VLDB, Morgan Kaufmann, San francisco, 1994, pp. 144–155.
- [13] T. Zhang, R. Ramakrishnan, M. Livay, Brich: an efficient data clustering method for very large databases, in: Proceedings of ACM SIGMOD, ACM New York, 1996, pp. 103–114.
- [14] E. M. Knorr, R. T. Ng, V. Tucakov, Distance-based outliers: algorithms and applications, *The VLDB Journal The International Journal on Very Large Data Bases* 8 (2000) 237–253.
- [15] E. M. Knorr, R. T. Ng, Algorithms for mining distance-based outliers in large datasets, in: Proceedings of the 24rd International Conference on Very Large Data Bases, pp. 392–403.
- [16] M. Agyemang, K. Barker, R. Alhajj, A comprehensive survey of numeric and symbolic outlier mining techniques, *Intelligent Data Analysis* 10 (2006) 521–538.
- [17] F. Angiulli, F. Fassetti, Dolphin: An efficient algorithm for mining distance-based outliers in very large datasets, *ACM Transactions on Knowledge Discovery from Data (TKDD)* 3 (2009) 1–57.
- [18] R. Brunelli, *Template Matching Techniques in Computer Vision: Theory and Practice*, Wiley, 2009.
- [19] R. J. Baron, Mechanisms of human facial recognition, *International journal of man machine studies* 15 (1981) 137–178.
- [20] D. H. Ballard, Generalizing the hough transform to detect arbitary shapes, *Pattern recognition* 13 (1981) 111–122.
- [21] R. O. Duda, P. E. Hart, Use of hough transforms to detect lines and curves in picture, *ACM communications* 15 (1972) 11–15.
- [22] R. T. Chin, C. R. Dyer, Model-based recognition in robot vision, *ACM compt. surveys* 18 (1986) 67–108.
- [23] W. E. L. Grimson, Object recognition by computer: the role ofgeometric constraints, MIT press, Cambridge, MA, 1990.
- [24] P. E. Anuta, Digital image registration of multi-spectral video imagery, *Soc. Photo-Opt. Instrum. Engrs. J.* 7 (1969) 168–175.
- [25] D. I. Barnea, H. F. Silverman, A class of algorithms for fast digital image registration, *IEEE Transactions on computers C-21* (1972) 179–186.
- [26] M. Svedlow, C. D. McGillem, P. E. Anuta, Experimental examination of similaity measures and processing methods used for image registration, in: Symposium on machine processing remotely sensed data.
- [27] D. V. S. Chandra, W. W. M. J. S. Boland, H. S. Ranganth, Feature matching multiple image registration using haar coefficients, in: SOUTHCON, pp. 549–552.
- [28] H. Schutte, S. Frydrychowicz, J. Schnroder, Scene matching with translation invariant transforms, in: 5th international joint conference on pattern recognition, pp. 195–198.
- [29] R. Y. Wong, E. L. Hall, Scene matching with invariant moments, *Compt. graphics image processing* 8 (1978) 16–24.
- [30] M.-K. Hu, Visual pattern recognition by momont invariants, *Trans. Inform. Theory* (1962) 179–187.
- [31] A. Goshtasby, Template matching in rotated images, *IEEE Transactions on pattern analysis and machine intelligence* 7 (1985) 338–344.
- [32] B. Widrow, The rubber mask technique, parts i and ii, *Pattern recognition* 5 (1973) 175–211.
- [33] M. Fischler, R. Elschlager, The representation and matching of pictorial structures, *IEEE Trans. Computers* 22 (1973) 67–92.
- [34] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active counter models, *International journal on computer vision* 1 (1988) 321–331.
- [35] R. Malladi, J. Sethain, B. Vemuri, Shape modelling with front propagation: A level set approach, *IEEE Transaction on pattern analysis and machine intelligence* 17 (1995) 158–175.
- [36] L. Cohen, Note on active counter models and ballons, *CVGIP: Image understanding* 53 (1991) 211–218.
- [37] L. H. Staib, J. S. Duncan, Boundary finding with parametrically deformable models, *IEEE Transaction on pattern analysis and machine intelligence* 14 (1992) 1061–1075.
- [38] A. L. Yuille, P. W. Hallinan, D. S. Cohen, Feature extraction from faces using deformable templates, *International journal of computer vision* 8 (1992) 133–144.
- [39] Y. Zhong, Object matching using deformable templates, Ph.D. thesis, Michigan State University, 1997.
- [40] L. C. Jain, U. Halici, I. Hayashi, S. Lee, S. Tsutsui (Eds.), *Intelligent Biometric Techniques in Fingerprint and Face Recognition*, Crc Press International Series on Computational Intelligence., CRC, 1999.
- [41] G. Guo, *Face, Expression, and Iris Recognition Using Learning-Based Approaches*, VDM Verlag, 2008.
- [42] K. Delac, M. Grgic, *Face Recognition*, I-Tech Education and Publishing, Vienna, Austria, 2007.
- [43] H. Wechsler, *Reliable Face Recognition Methods*, Springer Berlin / Heidelberg, 2006.
- [44] W. Zhao, R. Chellappa, *Face Processing : Advanced Modeling and Methods*, ACADEMIC PRESS, 2005.
- [45] S. Z. Li, A. K. Jain, *Handbook of Face Recognition*, Springer Berlin / Heidelberg, 2005.
- [46] M. Lades, J. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. Wurtz, W. Konen, Distortion invariant object recognition in the dynamic link architecture, *IEEE Trans. on Computers* 42 (1993) 300311.
- [47] H. Chen, P. Belhumeur, D. Jacobs, In search of illumination invariants, *IEEE Proc. Computer Vision and Pattern Recognition* 1 (2000) 254261.
- [48] D. Jacobs, P. Belhumeur, R. Basri, Comparing images under variable illumination, *IEEE Proc. Computer Vision and Pattern Recognition* (1998) 610617.
- [49] P. Belhumeur, J. Hespanha, D. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, *IEEE Trans. on Pattern Analysis and Machine Intelligence* 19 (1997) 711720.
- [50] Y. Adini, Y. Moses, S. Ullman, Face recognition: The problem of compensating for changes in illumination direction, *IEEE Trans. on Pattern Analysis and Machine Intelligence* 19 (1997) 721732.

- [51] H. Chen, P. Belhumeur, D. Jacobs, In search of illumination invariants, *IEEE Proc. Computer Vision and Pattern Recognition* (2000) 254261.
- [52] Y. Moses, S. Ullman, Limitations of non model-based recognition schemes, in: *European Conference on Computer Vision*, p. 820828.
- [53] A. M. Martinez, Recognizing imprecisely localized, partially occluded and expression variant faces from a single sample per class, *IEEE Trans. on Pattern Analysis and Machine Intelligence* 24 (2002).
- [54] D. Beymer, T. Poggio, Face recognition from a single example view, in: *Proceedings of 5th International Conference on Computer Vision*, Boston, MA, USA, pp. 500–507.
- [55] V. Blanz, T. Vetter, Face recognition based on fitting a 3d morphable model, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25 (2003) 1063–1074.
- [56] D. Jiang, Y. Hu, S. Yan, L. Zhang, H. Zhang, W. Gao, Efficient 3d reconstruction for face recognition, *Pattern Recognition* 38 (2005) 787–798.
- [57] H. Ling, S. Soatto, N. Ramanathan, D. Jacobs, A study of face recognition as people age, in: *IEEE International Conference on Computer Vision (ICCV)*.
- [58] N. Ramanathan, R. Chellappa, Face verification across age progression, in: *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, volume 2, p. 462469.
- [59] N. Ramanathan, R. Chellappa, Modeling age progression in young faces, in: *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, volume 1, p. 387394.
- [60] A. M. Martinez, R. Benavente, The ar face database, *CVC Technical Report* 24, 1998.
- [61] A. M. Martinez, R. Benavente, Ar face database, 2000. Available from <http://rvl.ewn.ecn.purdue.edu/RVL/database.htm>.
- [62] W. L. Martinez, A. R. Martinez, *Exploratory Data Analysis with MATLAB*, Chapman and Hall, 2004.
- [63] P. W. Hallinan, G. Gordon, A. L. Yuille, P. Giblin, D. Mumford, *Two- and Three-Dimensional Patterns of the Face*, AK Peters,Ltd, 1999.
- [64] A. Kasinski, A. Florek, A. Schmidt, The put face database, *Image Processing and Communications* 13 (2008) 59–64.
- [65] A. Georghiades, P. Belhumeur, D. Kriegman, From few to many: Illumination cone models for face recognition under variable lighting and pose, *IEEE Trans. Pattern Anal. Mach. Intelligence* 23 (2001) 643–660.
- [66] W.-C. Kao, M.-C. Hsua, Y.-Y. Yang, Local contrast enhancement and adaptive feature extraction for illumination-invariant face recognition, *Pattern Recognition* 43 (2010) 1736–1747.
- [67] A. S. Tolba, A. El-Baz, A. El-Harby, Face recognition: A literature review, *International Journal of Signal Processing* 2 (2006) 88–103.
- [68] W. Zhao, R. Chellappa, A. Rosenfeld, P. J. Phillips, Face recognition: A literature survey, *ACM Computing Surveys* (2003) 399–458.
- [69] M. Turk, A. Pentland, Eigenfacesforrecognition, *Journal of Cognitive Neuroscience* 3 (1991) 71–86.
- [70] H. Yu, J. Yan, A direct lda algorithm for high-dimensional data with application to face recognition, *Pattern Recognition* 34 (2001) 2067–2070.
- [71] O. Deniz, M. Castrillon, M. Hernandez, Face recognition using independent component analysis and support vector machines, *Pattern Recognition Letters* 24 (2003) 2153–2157.
- [72] Y. Gao, K. Leung, Face recognition using line edge map, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (2002).
- [73] J. Huang, P. Yuen, W. Chen, J. Lai, Choosing parameters of kernel subspace lda for recognition of face images under pose and illumination variations, *IEEE Trans. Syst. Man Cybern. B* 37 (2007) 847–862.
- [74] P. Belhumeur, J. Hespanha, D. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, *IEEE Transactions on Pattern Analysis and Machine Inteligence* 9 (1997) 711–720.
- [75] J. R. M. M. S. Bartlett, T. J. Sejnowski, Face recognition by independent component analysis, *IEEE Transactions on neural network* 13 (2002) 1450–1464.
- [76] G. B. Huang, M. Ramesh, T. Berg, E. Learned-Miller, Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments, Technical Report 07-49, University of Massachusetts, Amherst, 2007.
- [77] Y. Taigman, L. Wolf, T. Hassner, Multiple one-shots for utilizing class label information, in: *The British Machine Vision Conference (BMVC)*.