# A New Distance Measure for Face Recognition System

**Conference Paper** · September 2009

DOI: 10.1109/ICIG.2009.97 · Source: DBLP



2 authors:


Aditya Nigam
Indian Institute of Technology Mandi



Phalguni Gupta
National Institute of Techical Teachers Training and Research, Kolkata

2009 Fifth International Conference on Image and Graphics

# A New Distance Measure for Face Recognition System

Aditya Nigam
Department of Computer Science and Engineering
Indian Institute of Technology
Kanpur, INDIA 208016
Email: naditya@cse.iitk.ac.in

Phalguni Gupta
Department of Computer Science and Engineering
Indian Institute of Technology
Kanpur, INDIA 208016
Email: pg@cse.iitk.ac.in

***Abstract*—This paper proposes a new powerful distance measure called Normalized Unmatched Points (NUP). This measure can be used in a face recognition system to discriminate facial images. It works by counting the number of unmatched pixels between query and database images. A face recognition system has been proposed which makes use of this proposed distance measure for taking the decision on matching. This system has been tested on four publicly available databases, viz. ORL, YALE, BERN and CALTECH databases. Experimental results show that the proposed measure achieves recognition rates more than 98.66% for the first five likely matched faces. It is observed that the NUP distance measure performs better than other existing similar variants on these databases.**

## I. INTRODUCTION

Face recognition can be applied in identifying criminals, automatic video surveillance and many other security systems. A good algorithm for face recognition should be able to tolerate small amount of variations in the environmental factors like facial poses, illumination, image backgrounds, facial expressions, human ageing. Also it should not take must time and space. The work done in the field of face recognition is surveyed well in [13], [5]. There exist many different methods to recognize human faces such as PCA [16], EBGM [18], Neural Networks [10] and Support Vector Machines (SVM) [8]. All of the above mentioned methods work on gray scale images and hence do not provide too much tolerance to varying conditions.

Initially the Hausdorff Distance (*HD*) and Partial Hausdorff Distance (*PHD*) have been proposed by Huttenlocher and Rucklidge *et al* [9] for the purpose of image comparison. Images are treated by these measures as set of edge points. The property which makes these distance measures suitable for face recognition is their robustness for small amount of local non-rigid distortions.

*HD* and *PHD* measures are used successfully for object localization by Rucklidge [12]. Dubuisson [6] has modified *HD* measure while Takacs [15] has modified *PHD* to *M2HD* measure. *M2HD* measure has introduced a "neighborhood" and penalizes those points that have got matched outside it. In [7] some of the important feature points of the face like nose, mouth, eyes are given more importance in the proposed *SWHD* and *SW2HD* measures. Eigen faces are also used as the weighing function to improve *SW2HD* in [11].

Most of the vital three-dimensional facial information is not used while working with *HD* measure and all its variants because they are defined on edge maps. They cannot perform well beyond a certain level because drastic changes may be caused due to pose and expression variances in edge images. Two new measures $H_g$ and $H_{pg}$ which can work on gray quantized images directly are proposed by Vivek and Sudha [17]. Quantized images are obtained by considering $n$ most significant bits. These measures search for a correspondence between sets of pixels having the same quantized value.

In this paper, a new measure, called Normalized Unmatched Points ($NUP$), is proposed for comparison of gray facial images which has tolerance to varying environmental conditions. This measure has been found to perform much better than similar existing measures on the benchmark face databases.

The paper is organised as follows. Section II discusses about the $HD$ measure. In Section III the proposed distance measure has been proposed. Experimental results have been analysed in Section IV. Last section concludes the paper.

## II. HAUSDORFF DISTANCE MEASURE

Hausdorff distance (*HD*) measure can be seen as the dissimilarity between two sets of points. It can also be applied on edge maps to compare shapes. This measure can be calculated without explicit pairing up of the points from the two sets because it considers proximity instead of exact superposition. Suppose, $A = \{a_1, a_2, a_3, a_4..a_m\}$ and $B = \{b_1, b_2, b_3, b_4..b_n\}$ are two sets of points. The undirected *HD* [9] between $A$ and $B$ is given by:

$$HD(A, B) = max(hd(A, B), hd(B, A)) = HD(B, A) \quad (1)$$

In the above equation $hd(A, B)$ is directed Hausdorff distance defined by:

$$hd(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (2)$$

where $\|.\|$ is the vector's usual second norm.

Occlusion and missing part may pose some serious troubles to be $HD$ measure. Hence, to overcome this Partial Hausdorff Distance (*PHD*) has been proposed. Directed version of *PHD* is defined as:

$$phd(A, B) = K^{th} \max_{a \in A} \min_{b \in B} \|a - b\| \quad (3)$$

978-0-7695-3883-9/09 $26.00 © 2009 IEEE
DOI 10.1109/ICIG.2009.97

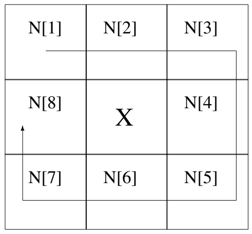



**Fig. 1:** 8-neighborhood

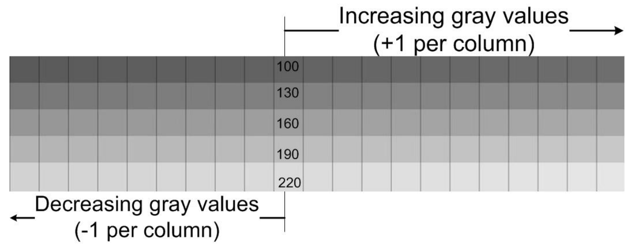



**Fig. 2:** Gray-value spectrum.

while undirected version remains to be the same. Non-linearity of $max$ and $min$ functions really makes these measures susceptible to noise.

## III. Proposed Measure

Gray values are very sensitive to the environmental conditions. Hence, it is very hard for a measure to acquire useful information in varying environment from the gray scale images.

A transformation is proposed by Sudha and Wong [14] by which gray-scale images can be converted into transformed images to preserve the intensity distribution. One can make use of it to get some robustness against illumination variation and local non-rigid distortions. One can observe that a pixel's relative gray value with respect to its 8-neighborhood pixels can be more stable than its own gray value.

Define a vector $V$ as

$$V[i] = \begin{cases} -1 & 0 \leq N[i] < X \\ 0 & N[i] = X \\ 1 & X < N[i] \leq 255 \end{cases} \tag{4}$$

where X is the gray value of pixel $A_{j,k}$ and $N[i], i = 1, 2, ...8$ are the gray values of $8$ neighboring pixels of $3 \times 3$ sub-image with center pixel $A_{j,k}$ (as shown in Fig. 1). Note that each pixel will now have an $8$ element vector $V$ associated to it containing the sign of first-order derivative with respect to its 8-neighborhood. This transformation ensures that if the gray value of a pixel in an image is slightly changed in its pose, the corresponding vector of pixel does not get affected. It is clear that this property holds good when the gray values of $8$ neighbors are not too close to each other. As there is not too much variation among the $8$ neighbors in facial images, the above property does not hold true for most of the pixels. Hence, if the vector $V$ is defined as in Eq. (4), it may not convey sufficient information.

Gray values of pixels within the $8$ neighborhood with respect to the center pixel of $3 \times 3$ sub-image are observed to be similar. Also, gray levels are hardly distinguishable within a range of $\pm 5$ units (as shown in Fig. 2).

### A. Transformation

To provide tolerance to illumination variation and local non-rigid distortions in facial image, a transformation termed as $gt$-transformation has been proposed. We can define $8$ element vector $V$, for any pixel having the gray value $X$, as:

$$V[i] = \begin{cases} 0 & 0 \leq N[i] < (X - gt) \\ 1 & (X - gt) \leq N[i] \leq (X + gt) \\ 2 & (X + gt) < N[i] \leq 255 \end{cases} \tag{5}$$

In the above equation, $gt$ is defined to be the gray value tolerance. The value of $gt$ totally depends on the type of images in the database. It has been observed that choosing the gray value tolerance parameter $gt$ more than 5 may affect the performance of the system adversely. Since every element of $V \in \{0, 1, 2\}$, hence any pixel $A_{j,k}$ can now be seen as an $8$-digit number in base $3$. The transformed value of the pixel $A_{j,k}$ is basically $\sum_{i=0}^{7} 3^i \cdot V[i+1]$ which always lies between $\{0, 3^8 - 1\}$. As compared to the original gray value of the pixel, this transformed value is found to be more stable because it is obtained from the gray values of $8$ neighbors of a pixel. High $gt$ values are not suggested in case of some directional lighting or heavy illumination variations because it may create some considerably dark patches in the image within which if the vector $V$ is calculated with big $gt$ value it may loose its meaning (as also suggested by Eq. (5)).

For $A$ and $B$, two $gt$-transformed images $NUP(A, B)$ can be defined as:

$$NUP(A, B) = \| \langle \frac{N^U_{AB}}{N_A}, \frac{N^U_{BA}}{N_B} \rangle \|_p \tag{6}$$

where $N^U_{AB}$, $N^U_{BA}$, $N_A$ and $N_B$ are the number of unmatched points in $A$ with respect to $B$, that in $B$ with respect to $A$, number of points in $A$ and that in $B$ respectively and $\|.\|_p$ is the $p^{th}$ norm.

The number of unmatched points in image $A$ with respect to image $B$ is defined as:

$$N^U_{AB} = \sum_{a \in A} (1 - Match(a, B)) \tag{7}$$

where $Match(a, B)$ returns 1 if there exists a pixel $b \in N^a_B$ such that both $a$ and $b$ have the same transformed value; otherwise it returns 0.

For any pixel *a* its neighborhood in image *B* (*i.e.* $N^a_B$) is defined as:

$$N^a_B = \{b \in B \mid \|a - b\| \leq d\sqrt{2}\} \tag{8}$$

where $d\sqrt{2}$ is the radius of circumcircle of $(2d+1) \times (2d+1)$ square neighborhood centered at pixel $a$.

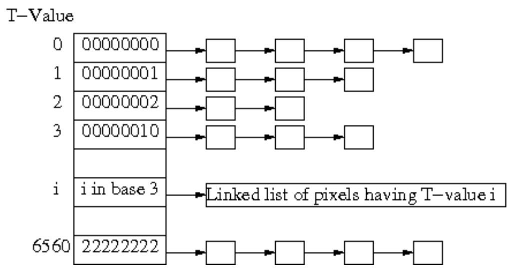



**Fig. 3:** Data structure: LIST

### B. *Computation of* $\boldsymbol{NUP}$ *measure*

Gray scale images $P$ and $Q$ are converted into transformed images $A$ and $B$ using $gt$-transformation. To compute $NUP(P, Q)$ one has to check for all pixels $a \in A$, whether there exist some pixel $b \in B$ such that both $a$ and $b$ have the same transformed value and $b \in N_B^a$. This can be done in an efficient manner by using an array of pointers to linked list (as shown in Fig. 3). The array of pointers $LIST[B, 0:6560]$ [14] has been created from the transformed image $B$ having 6561 elements. In this data structure every element $LIST[B, i],\ i = 0, 1, 2...6560$ points a linked list that contains all those pixels of $B$ which are having the transformed value $i$.

Let the transformed value of a pixel $a$ be $x$. Using the $LIST$ data structure, Algorithm 1 can be used to search a node $b$ such that $b \in N_B^a$, in all nodes (*i.e.* pixels) of the linked list pointed by $LIST[B, x]$. If such a node exists, it returns 1 else it returns 0.

**Algorithm 1** $Match(a, B)$

**Require:** A pixel $a$ and $gt$-transformed image $B$.
**Ensure:** If $a$ got $Matched$ then return 1, else return 0.

1: $tval\_a \leftarrow gt$-transformed value of pixel $a$;
2: Search $LIST[B, tval\_a]$, for a point $P \in N_B^a$;
3: **if** no point found in step 2 **then**
4:    **return** 0;
5: **else**
6:    **return** 1;
7: **end if**

Algorithm 2 can be used to calculate number of unmatched points between two $(r-2) \times (c-2)$ sized transformed images $A$ and $B$ with the help of Algorithm 1.

Algorithm 3, first computes directed normalized unmatched points between $A$ and $B$ (*i.e.* $N_{AB}^U$) and that between $B$ and $A$ (*i.e.* $N_{BA}^U$) and finally obtains their $p^{th}$ norm to determine the distance between the two images.

### C. *Running Time Analysis*

Initially gray scale images are converted into $gt$-Transformed images only once and then saved. For this, full $r \times c$ image is scanned once and hence its time complexity is $O(rc)$.

**Algorithm 2** $Compare(A, B)$

**Require:** Two $gt$-transformed images $A$ and $B$.
**Ensure:** Return $N_{AB}^U$.

Construct $LIST$ (array of pointers to linked list) for $B$;
$unmatched \leftarrow 0$;
**for** $i = 0$ to $(r-3)$ **do**
   **for** $j = 0$ to $(c-3)$ **do**
      **if** $Match(A_{ij}, B)$ is 0 **then**
         $unmatched \leftarrow unmatched + 1$;
      **end if**
   **end for**
**end for**
**return** $unmatched$;

**Algorithm 3** $NUP(A, B)$

**Require:** $gt$-transformed images $A$ and $B$.
**Ensure:** Return $NUP(A, B)$.

Load $gt$-transformed images $A$ and $B$ from the Disk;
$N_{AB}^U \leftarrow Compare(A, B)$;
$N_{BA}^U \leftarrow Compare(B, A)$;
$NUP(A, B) \leftarrow \sqrt[p]{\left(\frac{N_{AB}^U}{N_a}\right)^p + \left(\frac{N_{BA}^U}{N_b}\right)^p}$;
**return** $NUP(A, B)$;

A linked list of pixels have to be searched linearly for the execution of Algorithm 1. Therefore the time taken by this algorithm depends on the length of the list. Assuming that the length of the largest linked list is $k$. *NUP* measure is obtained by calling Algorithm 2 for $2rc$ times. Therefore time required to compute *NUP* measure is $O(krc)$.

## IV. Experimental Results

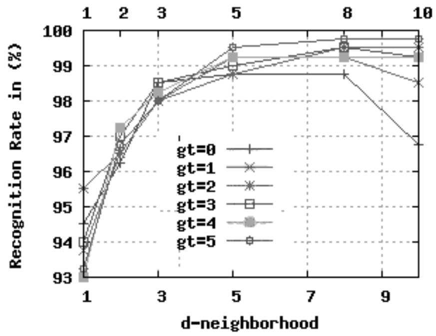
### A. $\boldsymbol{NUP}$ *measure applied to face recognition*

Haar cascades are used to extract face images. Then these images are resized to the ORL standard size $92 \times 112$ pixels. Images after transformation are saved as color images (in TIFF format) of size $90 \times 110$.

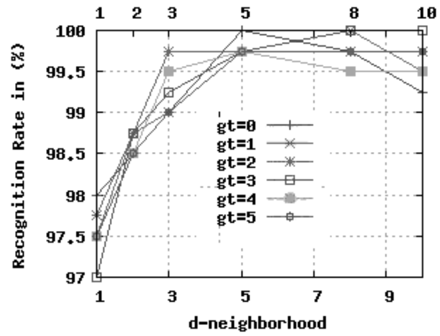
Each image in the testing set stored in the database is matched with all other images excluding itself. Finally, top $T$ best matches are reported. Parameter $T$ can range from 1 to (total number of poses per subject - 1). For the experiments, $T$ is set to 1 and 5. Any test image is announced to be $matched$ if there exists a image of that subject in top $T$ best matches. Recognition rate is defined as:

$$\text{Recognition rate} = \frac{\text{Number of matched images}}{\text{Total number of images}} \tag{9}$$

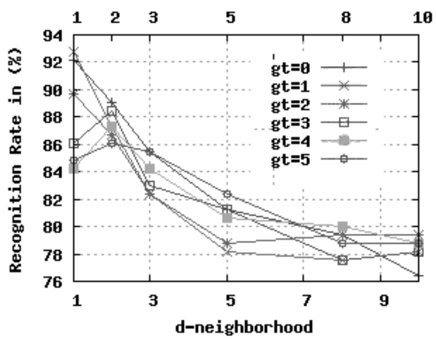
which is generally used to analyze the performance of any measure. This $NUP$ measure can tolerate slight pose variation. One can store templates of faces in different poses at the time of registration for handling wide pose variations. Some of the standard benchmark facial image databases such as ORL [3], YALE [4], BERN [1] and CALTECH [2] are used for the evaluation of $NUP$ measure. $NUP$ has shown



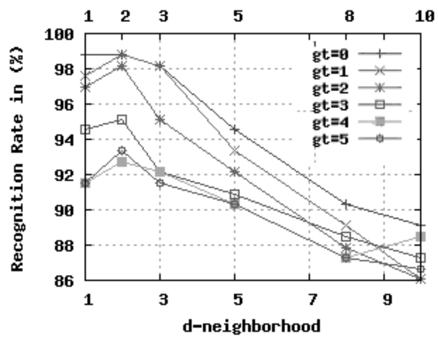

(a) ORL, Considering $top$ 1 best matched



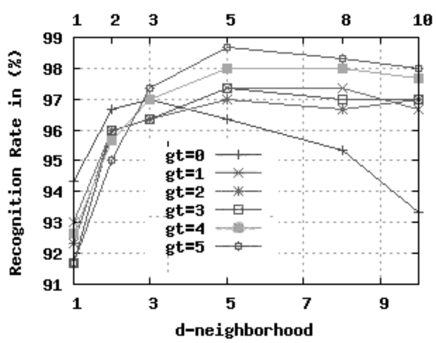

(b) ORL, Considering $top$ 5 best matched



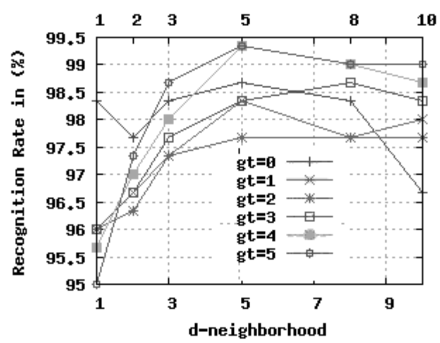

(c) YALE, Considering $top$ 1 best matched



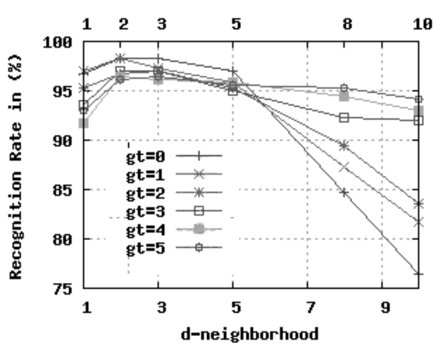

(d) YALE, Considering $top$ 5 best matched



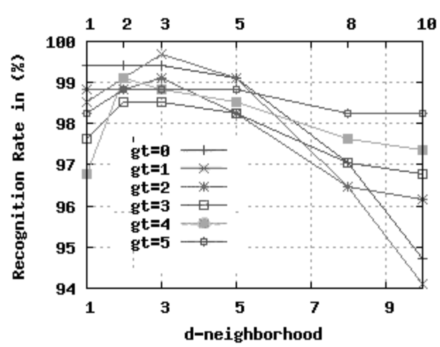

(e) BERN, Considering $top$ 1 best matched




(f) BERN, Considering $top$ 5 best matched




(g) CALTECH, Considering $top$ 1 best matched




(h) CALTECH, Considering $top$ 5 best matched

**Fig. 4:** Results of $NUP$ based Recognition with Varying $gt$ and $d$ on Various Databases

**TABLE I:** Databases vs Best Parameters

| Db | TOP 1 | | | TOP 5 | | |
|---|---|---|---|---|---|---|
| | gt | d | RR % | gt | d | RR % |
| **ORL** | 5 | 8 | 99.75 | 5 | 8 | 100 |
| **YALE** | 1 | 1 | 92.75 | 1 | 2 | 98.78 |
| **BERN** | 5 | 5 | 98.66 | 5 | 5 | 99.33 |
| **CALTECH** | 1 | 2 | 98.23 | 1 | 3 | 99.70 |

**TABLE II:** Comparative Analysis for Top Best Match

| Distance Measure | Recognition rate (%) ORL | YALE |
|---|---|---|
| PCA | 63.00 | 50.00 |
| HD | 46.00 | 66.00 |
| PHD | 72.08 ($f = 0.85$) | 84.00 ($f = 0.70$) |
| M2HD | 75.00 | 80.00 |
| SWHD | 82.00 | 82.00 |
| SW2HD | 88.00 | 83.00 |
| SEWHD | 88.00 | 85.00 |
| SEW2HD | 91.00 | 89.00 |
| $H_{pg}$ | 91.25 | 83.30 ($f = 0.55$) |
| **NUP** | **99.75** ($gt = 5,\ d = 8$) | **92.72** ($gt = 1, d = 1$) |

good recognition rates if there is small amount of variation in lighting conditions, poses and expressions.

### B. Parameterized analysis of *NUP*

*NUP* measure is parameterized primarily by parameters, viz. gray value tolerance $gt$ varying within range $[0, 5]$ and neighborhood parameter $d$ varying within range $[1, 10]$ and order of the norm $p$.

This Measure requires to give weights to $\left[\frac{N^U_{AB}}{N_a}\right]$ and $\left[\frac{N^U_{BA}}{N_b}\right]$ in the order of their importance (*i.e.* higher weight have to be given to the bigger one). Higher $p$ value starts to favor the bigger element and with $p = \infty$, we end up choosing the bigger one. In this experiment, the last value of $p$ is found to be 20.

Lower $gt$ values perform well on the databases where there is illumination varying such as YALE and CALTECH databases. In the absence of directional light and slight illumination variation as in ORL and BERN databases, higher $gt$ values can provide better discrimination (as shown in Table I and Fig. 4).

Bigger neighborhood can yield good performance if the databases have pose and expression variations as in ORL and BERN databases. Also both of these databases are having images that are not normalized with respect to any facial features. YALE and CALTECH databases contain fairly normalized images without too much pose and expression variations; hence smaller neighborhood is expected (as shown in Table I and Fig. 4).

### C. Comparative Analysis

Table II and Table III show the superiority in terms of discriminative power of $NUP$ measure. Experimental results have shown that $NUP$ measure has a better discriminative power and can achieve a higher recognition rate than $HD$ [9],

**TABLE III:** Results on BERN Database for Top Match

| Test Faces | Recognition rate (%) PHD ($f = 0.85$) | LEM | $H_{pg}$ | ***NUP*** ($gt = 5, d = 5$) |
|---|---|---|---|---|
| Looks right/left | 74.17 | 74.17 | 95.83 | **99.00** |
| Looks up | 43.33 | 70.00 | 90.00 | **99.00** |
| Looks down | 61.66 | 70.00 | 68.33 | **98.00** |
| Average | 58.75 | 72.09 | 87.50 | **98.66** |

$PHD$ [9], $MHD$ [6], $M2HD$ [15], $SWHD$ [7], $SW2HD$ [7], $SEWHD$ [11], $SEW2HD$ [11], $H_g$ [17] and $H_{pg}$ [17].

## V. Conclusion

This paper has proposed Normalized Unmatched Points ($NUP$) measure to compare gray facial images. All the previous existing Hausdorff distance based measure works on edge images but the face recognition approach based on $NUP$ is different from all of them. It actually works on transformed images that are obtained from gray images rather than edge images enabling $NUP$ measure to achieve the appearance based comparison of faces. An algorithm by which $NUP$ measure can be computed efficiently is also presented in the paper.

The values of $gt$ and $d$ which are the two parameters guiding the $NUP$ measure are chosen taking into account the nature of the images in the face database. $NUP$ distance measure has shown tolerance to slight variation in expressions, illumination and poses. High recognition rates ($98.00\%$ and above) are achieved while experimenting on ORL, YALE, BERN and CALTECH benchmark face databases with the $NUP$ distance measure. In uniform illumination and constrained environment, $NUP$ measure can also be used for scene segmentation in videos, video surveillance, face detection, face authentication and so on.

## References


[1] Bern database. ftp://ftp.iam.unibe.ch/pub/images/faceimages/.

[2] Caltech database. http://www.vision.caltech.edu/html-files/archive.html.

[3] Orl database. http://www.uk.research.att.com/facedatabase.html.

[4] Yale database. http://cvc.yale.edu/projects/yalefaces/yalefaces.html.

[5] R. Chellappa, C. L. Wilson, and S. Sirohey. Human and machine recognition of faces: A survey. *Proceedings of the IEEE*, 83(5):705–740, May 1995.

[6] M. Dubuisson and A. K. Jain. A modified hausdorff distance for object matching. *Proceedings of the 12th IAPR International Conference*, 1(1):566–568, 1994.

[7] B. Guo, K.-M. Lam, K.-H. Lin, and W.-C. Siu. Human face recognition based on spatially weighted hausdorff distance. *Pattern Recognition Letters*, 24(1-3):499–507, 2003.

[8] G. Guo, S. Z. Li, and K. L. Chan. Face recognition by support vector machines. In *FG*, pages 196–201, 2000.

[9] D. P. Huttenlocher, G. A. Klanderman, and W. Rucklidge. Comparing images using the hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(9):850–863, 1993.

[10] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back. Face recognition: A convolutional neural network approach. *IEEE Transactions on Neural Networks*, 8(1):98–113, 1998.

[11] K.-H. Lin, K.-M. Lam, and W.-C. Siu. Spatially eigen-weighted hausdorff distances for human face recognition. *Pattern Recognition*, 36(8):1827–1834, 2003.

[12] W. Rucklidge. Locating objects using the hausdorff distance. In *ICCV*, pages 457–464, 1995.

[13] A. Samal and P. A. Iyengar. Automatic recognition and analysis of human faces and facial expressions: a survey. *Pattern Recognition*, 25(1):65–77, 1992.
[14] N. Sudha and Y. Wong. Hausdorff distance for iris recognition. *Proceeding of 22nd IEEE International Symposium Intelligent Control*, 1(1):614–619, 2007.
[15] B. Takács. Comparing face images using the modified Hausdorff distance. *Pattern Recognition*, 31(12):1873–1881, Dec. 1998.
[16] M. Turk and A. P. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.
[17] E. P. Vivek and N. Sudha. Gray hausdorff distance measure for comparing face images. *IEEE Transactions on Information Forensics and Security*, 1(3):342–349, 2006.
[18] L. Wiskott, J.-M. Fellous, N. Krüger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):775–779, July 1997.