

# SVDD-Based Illumination Compensation for Face Recognition

Sang-Woong Lee<sup>1</sup> and Seong-Whan Lee<sup>2,\*</sup>

<sup>1</sup> The Robotics Institute, Carnegie Mellon University,  
5000 Forbes Ave., Pittsburgh, PA 15213, USA  
rhiephil@cs.cmu.edu

<sup>2</sup> Center for Artificial Vision Research, Korea University,  
Anam-dong, Seongbuk-ku, Seoul 136-713, Korea  
swlee@image.korea.ac.kr

**Abstract.** Illumination change is one of most important and difficult problems which prevent from applying face recognition to real applications. For solving this, we propose a method to compensate for different illumination conditions based on SVDD(Support Vector Data Description). In the proposed method, we first consider the SVDD training for the data belonging to the facial images under various illuminations, and model the data region for each illumination as the ball resulting from the SVDD training. Next, we compensate for illumination changes using feature vector projection onto the decision boundary of the SVDD ball. Finally, we obtain the pre-image under the identical illumination with input image. By repeated for each person, we can recognize a person with facial images under same illumination. We also perform the face recognition in order to verify the efficacy of proposed method.

**Keywords:** Illumination compensation, face reconstruction, noise, support vector data description, face recognition.

## 1 Introduction

With various biometric methods such as finger print, iris, palm, gait and so forth, face recognition technologies have been one of most interesting fields up to recently. While finger print systems are generally used in our life, any face recognition system has not made an attraction to our daily life. Even though several products using face analysis for entertainments, it seems that it takes long times to use face recognition in real life.

The delay of a common product is caused by such actual problems as illumination changes, facial pose variation, facial expression changes, facial aging, different ethnic groups, etc. In the last decade, many researchers have brought up these problem and solution. Recently, Gross and Cohn surveyed the current state of the art in face recognition and made several experiments with variations of these problem and facial databases. They experimentally found that pose and illumination

---

\* Corresponding author.

changes are the most affective to the recognition rate [1]. Zhao et al. also identified two key problems with the illumination and pose problem in their survey paper [2]. Tan et al. mentioned the problem from a single image per person in training step and introduced holistic, local, and hybrid method for solving this problem [3]. They have paid much attention to illumination and pose changes.

We focus on solving an illumination problem based on holistic methods. In the holistic methods, while these keep all essential information of facial images, these should consider the variation of illumination using small number of training data. If trained faces have quite a different illumination with a gallery image, recognition rate decreases drastically. In earlier day, Pentland et al. proposed view-based face recognition. They prepared various training data in each change [12]. However, view-based method can not have all training data.

In order to compensating for unknown illumination changes, many approaches have tried to synthesize a new face or remove illumination effect. Gross and Brajovic proposed an image processing approach compensating illumination change. They tried to obtain facial images under consistent illumination condition based on weber's law [4]. Xudong and Lam proposed a local normalization of uneven illumination [5]. Georghiadis et al. proposed illumination cone models, which can render an arbitrary illumination using several bases [6]. Zhou et al. proposed an illumination-invariant face recognition using generalized photometric stereo [7]. This also carried out the iteration steps, which takes 2-3 seconds. Song et al. also used photometric stereo method and compensated for the difference between different illumination conditions by linear combination of facial images under different lighting [8]. Both of two approaches combines linear combination of models with Lambertian reflectance and have the limit of linear analysis. Zhang and Samaras proposed harmonic image exemplars, which assume that there is no pose variation [9].

We aim to also synthesize a facial image of each trained person under the same illumination condition as one of input face. In order to accomplish this goal, we extend the concept of SVDD to illumination problem. In section 2, we address a SVDD-based training with general concepts. We explain our compensation approach using vector projection and pre-image synthesis in section 3 and make some experiments to verify our idea in section 4. Finally we make a conclusion on our intuition and experimental proof.

## 2 Training Diverse Illumination Condition

The SVDD method trains data belonging normal class and defines the region including these data. With trained region, it is possible to detect abnormal data and outlier. This region is called as the ball,  $B$  with the center  $a \in \mathfrak{R}^d$  and the radius  $R$ , and the training data set  $D$  consisting of facial images,  $x_i \in \mathfrak{R}^d$ ,  $i = 1, \dots, N$  under diverse illumination. SVDD method used to find a ball which can meet two constraints concurrently. The ball should be as small as possible and should contain as many training data as possible. The trained ball can includes ideally most illumination conditions of one person.

Since it is generally said that illumination change lies on the domain of non-linear problem, balls can wrap very limited class of subsets in SVDD method, which deals with linear problem. To represent more complex decision regions in  $\mathfrak{R}^d$ , one can use the so-called feature map  $\phi: \mathfrak{R}^d \rightarrow F$  and hyper-balls,  $B_F$  with the center,  $a_F$  and the radius,  $R_F$  defined on the feature space  $F$ . We use the concepts of the slack variable  $\xi_i$ , the trade-off constant  $C$ , and kernel function  $K$  for optimal  $B_F$  which contains reasonably large portion of the (transformed) training face data set  $D_F = \{\phi(x_i) | i = 1, \dots, N\} \subset F$ . Finally, we can obtain the solution by the following QP:

$$\begin{aligned} \min_{\alpha} \quad & \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j K(x_i, x_j) - \sum_{i=1}^N \alpha_i K(x_i, x_i) \\ \text{s. t.} \quad & \sum_{i=1}^N \alpha_i = 1, \quad \alpha_i \in [0, C], \quad \forall i \end{aligned} \quad (1)$$

The criterion for the normality can be summarized as follows:

$$\begin{aligned} f_F(x) &\triangleq R_F^2 - \|\phi(x) - a_F\|^2 \\ &= R_F^2 - 1 + 2 \sum_{i=1}^N \alpha_i K(x_i, x) \\ &\quad - \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j K(x_i, x_j) \\ &\geq 0. \end{aligned} \quad (2)$$

This explanation is similarly referred from the paper of Park et al., where they use this SVDD concept for de-noising. They considered noised numeric images as abnormal data and clear numeric images as normal data, respectively in feature space. The projection of the vector of abnormal data is transformed to image space. The pre-image of projected vector results in de-noising one [10]. Lee et al. applied this knowledge to low resolution face recognition. They trained facial images of high resolution. A facial vector is moved to feature space, projected on the surface of the ball, and returned in image space [11].

### 3 Illumination Compensation

Briefly introducing our approaches, we train several facial images of a person in different illumination as normal class on off-line stage. In Fig. 1, training data set was composed of facial images of same person with various illuminations. As on-line stages, when a facial image is given, this is transformed in feature space and projected onto surface of trained ball,  $B_F$ . With this projected vector, we imitate illumination condition of arbitrary facial images (as abnormal) of another person.

The support is expressed by a reasonably small ball containing a reasonably large portion of face feature vectors,  $\phi(x_i)$ , of different illumination in a feature space. When an abnormal facial image is given as a test input,  $x$ , of different person from training class, the network resulting from the SVDD is supposed to recognize the fact that the  $x$  does not belong to the normal class; apparently all images in the ball has the same identity. If input image is used from another person under unknown illumination, it may be out of the ball because it has

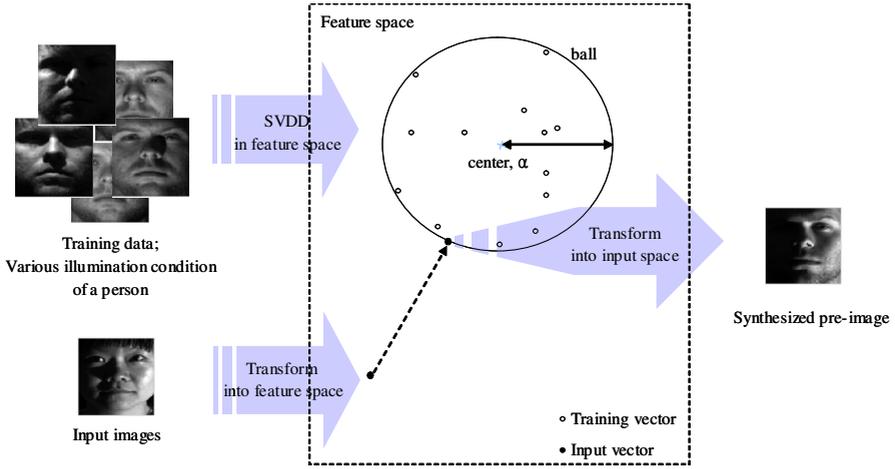


Fig. 1. Basic idea in example for illumination compensation

different shape. When input vector,  $x$  is transformed to input feature vector,  $\phi(x)$ , our intuition is the followings with regard to the vector projection.

- Pre-image of projected input vector,  $\phi^{-1}(P\phi(x))$  resembles the facial shape of the trained person
- The illumination condition of  $\phi^{-1}(P\phi(x))$  is similar with one of input images,  $x$ .

For vector projection and pre-image synthesis, we adopt Park and Kwok’s method [10]. It is quite easy to let the feature vector,  $\phi(x_i)$ , of the abnormal image move toward the center,  $a_F$ , of the SVDD ball,  $B_F$ , until it reaches the decision boundary, so that it can be tailored enough to be considered as normal<sup>1</sup>. Of course, since the movement starts from the feature,  $\phi(x_i)$ , there are plenty of reasons to believe that the tailored feature,  $P\phi(x_i)$ , still contains essential information about illumination condition. In the concrete, when the decision function  $f_F$  of Eq. (2) yields a nonnegative value for  $x$ , the test input is accepted normal as it is. In the otherwise case, the test input  $x$  is considered to be abnormal. For illumination compensation, we move the feature vector  $\phi(x)$  toward the center  $a_F$  up to the point where it touches the ball  $B_F$ . Thus, the outcome of this movement is the following:

$$P\phi(x) = a_F + \frac{R_F}{\|\phi(x) - a_F\|}(\phi(x) - a_F). \quad (3)$$

As a result of the projection, we have the obvious result  $\|P\phi(x) - a_F\| = R_F$ . Also, note that with  $\lambda \triangleq R_F/\|\phi(x) - a_F\|$ , the Eq. (3) can be further simplified into

<sup>1</sup> In this paper, we use Gaussian kernel where projected feature vector is located on the surface of the hyper-ball.

$$P\phi(x) = \lambda\phi(x) + (1 - \lambda)a_F, \quad (4)$$

where  $\lambda$  can be computed from

$$\lambda^2 = \frac{R_F^2}{\|\phi(x) - a_F\|^2} = \frac{R_F^2}{(1 - 2 \sum_i \alpha_i K(x_i, x) + \sum_i \sum_j \alpha_i \alpha_j K(x_i, x_j))}. \quad (5)$$

We try to find the pre-image of the  $P\phi(x)$ . If the inverse map  $\phi^{-1} : F \rightarrow \mathfrak{R}^d$  is well-defined and available, this final step attempting to get the pre-image via  $\hat{x} = \phi^{-1}(P\phi(x))$  will be trivial. However, the exact pre-image typically does not exist. Thus, we need to seek an approximate solution instead. For this, we follow the strategy, which uses a simple relationship between feature-space distance and input-space distance together with the MDS (multi-dimensional scaling).

In this paper, we insist that  $P\phi(x_i)$  resembles  $\phi(x_i)$  of trained person under the illumination of  $\phi(x)$ . We also insist that it is possible to perform better recognition with  $P\phi(x_i)$  than with  $\phi(x_i)$ . That is to say, we can find the illumination condition in training dataset, which means that we can match the new gallery image under unknown illumination and we can approximate the illumination parameter from the training dataset. For verifying our assumption, we perform some experiments in Section 4 with popular face database, which includes various illumination conditions.

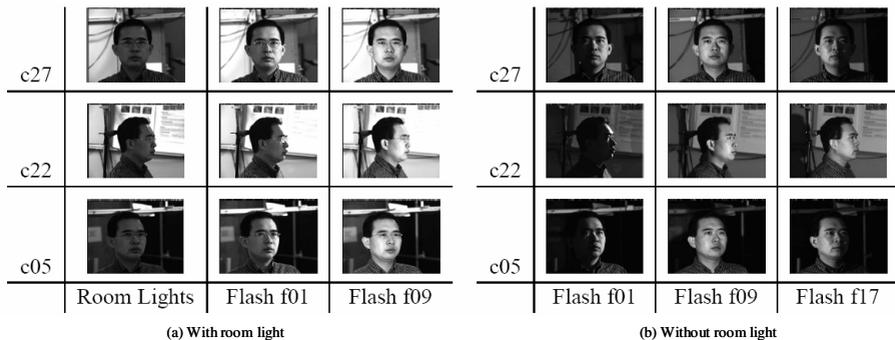
## 4 Experimental Results and Analysis

### 4.1 Facial Databases

For making experiments, we used the CMU PIE DB[13]. It contains 41,368 images of 68 subjects under 13 different poses, 43 different illumination conditions, and with 4 different expressions. 43 illumination variations are composed of 21 different flash directions with room light, 21 different flash directions without room light, and only room light without flash. Among this, we selected 1,428 facial images of  $C27$  pose under 21 illumination changes without room light (like 3 images of first row in Fig. 2 (b)). Then we converted them to gray scale images and normalized as same size.

### 4.2 Pre-Image Synthesis

If a facial image under arbitrary illumination condition is given, we synthesize a new facial image, which is under same illumination with input image by proposed method. Fig. 3 shows the 12 examples of experiments. Input image (left column) is randomly selected in PIE database. It is used for synthesizing a new facial image of another person, who is also randomly selected. As mentioned in previous sections, we carried out several processes and obtained the experimental result (middle column). The original image (right column) is shown by selecting the facial image of ground truth, which we have already known. We aim that



**Fig. 2.** Examples of facial images in CMU PIE Database [13]

the result is exactly same with the ground truth image. As shown in Fig. 3, two images almost resemble each other. However, in detailed view, there is little difference of shadow and reflectance. This may be caused by sleek surface. We could not consider this property. In order to compare two images, we measured the difference of pixel intensity. In other view, our SVDD-based method is called as non-linear combination of trained data using kernel trick. In this view, it is efficient way to compare linear combination method proposed by Moon et al.[8]. Table 1 shows the comparison with this method. As a result, the mean error of our proposed method is lower than one of theirs.

### 4.3 Performance in Face Recognition

In addition, our method is also useful in view-based face recognition. When recognizing a facial image, each training face is matched with input image. If two images were captured under different illumination condition, the recognition rate would be decreased. In case of this, we can repeat our approach for the illumination compensation of each training data to the direction of the illumination in input image.

In this strategy, we performed similar experiments with previous one. Firstly, we selected an input face under unknown illumination condition. Secondly, we synthesized each face by repeating all people in training set. Finally we recognized

**Table 1.** The results of face recognition with comparison

	Average of face recognition rate (%)	Average of image difference (grayscale)
Previous Method [8]	89	10.76
Proposed Method	94	6.73



Fig. 3. Example images of synthesized results with CMU PIE database

an input person by matching it with all synthesized faces. We can identify a person by selecting the index, which has the minimum error. In this experiment, we perform this process 100 times. One hundred of arbitrary faces were randomly selected and we recognized them repeatedly. The result of this experiment is also compared with linear combination method [8] and is shown in Table 1. The

experiment resulted in improvement of 5% recognition rate. As a result, our proposed method experimentally showed better performance, which is enough to draw the attraction in real application.

## 5 Conclusions

In this paper, we proposed a new method of synthesizing facial exemplars based on support vector data description and applied our approaches to illumination problem, which enables to recognize a face under arbitrary lighting condition. Our method is especially efficient in solving the limitation of view-based approaches. We first considered the SVDD problem for the data belonging to the trained facial images under various illuminations, and modelled the data region for each illumination as the ball resulting from the SVDD problem. Next, we impersonate an arbitrary illumination condition using feature vector projection onto the decision boundary of the SVDD ball. Finally, we combined facial images which are obtained from the pre-image of the projection, and built a facial image under compensated illumination. The applicability of the proposed method was illustrated via some experiments dealing with faces in popular facial databases. The experimental results showed that the facial images synthesized using the proposed method were quite similar to the original facial images, thus making it possible to recognize a person under various illuminations.

## Acknowledgments

We would like to thank Prof. Takeo Kanade for his kind help and this work was supported by the Korea Research Foundation Grant funded by the Korean Government(MOEHRD)(KRF-2006-214-D00144).

## References

1. Gross, R., Shi, J., Cohn, J.: Quo Vadis Face Recognition? In: Third Workshop on Empirical Evaluation Methods in Computer Vision (2001)
2. Zhao, W., Chellappa, R., Rosenfeld, A., Phillips, P.J.: Face Recognition: A Literature Survey. *ACM Computing Surveys*, 99–458 (2003)
3. Tan, X., Chen, S., Zhou, Z.-H., Zhang, F.: Face recognition from a single image per person: A survey. *Pattern Recognition* 39(9), 1725–1745 (2006)
4. Gross, R., Brajovic, V.: An Image Preprocessing Algorithm for Illumination Invariant Face Recognition. In: Kittler, J., Nixon, M.S. (eds.) AVBPA 2003. LNCS, vol. 2688, pp. 10–18. Springer, Heidelberg (2003)
5. Xie, X., Lam, K.-M.: An efficient method for face recognition under varying illumination. *IEEE Int. Symp. on Circuits and Systems* 4, 3841–3844 (2005)
6. Georghiadis, A., Belhumeur, P., Kriegman, D.: From Few to Many: Illumination Cone Models for Face Recognition under Variable lighting and Pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23(6), 643–660 (2001)

7. Zhou, S.K., Aggarwal, G., Chellappa, R., Jacobs, D.W.: Appearance Characterization of Linear Lambertian Objects, Generalized Photometric Stereo, and Illumination-Invariant Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29(2), 230–245 (2007)
8. Moon, S.-H., Lee, S.-W., Lee, S.-W.: Illumination Invariant Face Recognition Using Linear Combination of Face Exemplars. In: Kanade, T., Jain, A., Ratha, N.K. (eds.) AVBPA 2005. LNCS, vol. 3546, pp. 112–121. Springer, Heidelberg (2005)
9. Zhang, L., Samaras, D.: Face Recognition Under Variable Lighting using Harmonic Image Exemplars. *IEEE Conference Computer Vision and Pattern Recognition* 1, 19–25 (2003)
10. Park, J., Kang, D., Kim, J., Kwok, J.T., Tsang, I.W.: Pattern de-noising based on support vector data description. In: *Proceedings of International Joint Conference on Neural Networks*, Montreal, Canada, July 31-August 4, pp. 949–953 (2005)
11. Lee, S.-W., Park, J., Lee, S.-W.: Low Resolution Face Recognition Based on Support Vector Data Description. *Pattern Recognition* 39(9), 1809–1812 (2006)
12. Pentland, A., Moghaddam, B., Starner, T.: View-Based and Modular Eigenspaces for Face Recognition. In: *IEEE Conf. Computer Vision and Pattern Recognition*, pp. 84–91. IEEE Computer Society Press, Los Alamitos (1994)
13. Sim, T., Baker, S., Bsat, M.: The CMU Pose, Illumination, and Expression (PIE) Database of Human Faces. In: *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, Washington, USA, pp. 53–58 (2002)