Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Online Evolutionary Context-Aware Classifier Ensemble Framework For Object Recognition

Zhan Yu
Computer and Information Engineering
Inha University
Incheon, Korea
yshmzhan@126.com

Mi Young Nam
Computer and Information Engineering
Inha University
Incheon, Korea
rera@im.inha.ac.kr

Phill Kyu Rhee
Computer and Information Engineering
Inha University
Incheon, Korea
pkrhee@inha.ac.kr

***Abstract*—In this paper, we propose an online evolutionary Context-Aware classifier ensemble framework for object recognition systems which are adaptive to various environments. The starting point utilizes a context recognizer, a context knowledge base and a classifier ensemble generator to provide optimal solutions for classifier ensemble. Even though our framework is quite general and could be applied to various classification tasks, here we focus on the cases of face recognition. The proposed framework uses an unsupervised learning method to carry out context modeling tasks for various environments. The data for classifier ensemble are assigned to corresponding contexts based on supervised learning and an evolutionary algorithm processes all the information to generate context knowledge for online adaptation. Experimental comparisons with systems based on conventional face recognition algorithms upon four extended benchmark data sets, E-FERET, E-Yale, E-INHA, and our own database showed that the system based on our framework was able to operate in dynamic environments with stable performance which others could not achieve.

***Keywords*—Context-Aware, online classifier ensemble, evolutionary computing, object recognition,**

I. Introduction

Context-Aware computing has been investigated for years by researchers [1] with the ultimate goal of providing highly robust systems that can be applied to meet the user's needs in various fields such as visualization [2], distributed systems [3], mobile devices [4], [5] and image coding [6]. In order to make the decision of the system even more accurate, literature claim that by combining the outputs of a team of classifiers, the accuracy of the decision can be superior to the best one in the team [7]. However, selecting a suitable combination method has proven to be difficult [7]. Despite numerous efforts spent in this area, methods developed so far have been based on intuitive techniques or aimed at special cases [8-14]. In this study, we propose a way of performing the classifier ensemble relying on the context information. We separate our framework into two modes, a learning mode and a operation mode, in accordance with their functionalities. As for the learning mode, a context recognizer and a classifier ensemble generator are trained with unsupervised and supervised methods respectively. An evolutionary processor optimizes the resulting context knowledge. For the operation mode, context and input data are gathered separately. The former is used to generate classifier ensemble scheme according to the context knowledge accumulated during learning, while the latter is the input data of the classifier ensemble. In addition, an online learning scheme is established by means of collecting the output of the framework during the operation mode and repeating the process of the learning mode in a short period.

The text is organized as follows. Section II introduces the basic structure of the whole framework. Section III presents the details of the learning and operation modes. In Section IV, the experiment set up and results are included and Section V offers our conclusion.

II. Basic Framework Structure

The basic idea under the framework is to automatically modify the classifier ensemble scheme using the information derived from context data. As shown in Figure 1, a Context Recognizer (CR) is used for recognizing the context data. A Classifier Ensemble Generator (CEG) is employed for generating the classifier ensemble scheme and classifying the problem according to the context recognition result from the CR. A CR Training Module and CEG Training Module are used for training CR and CEG during the learning mode respectively. The Control Module (CM) assembles and dispatches the context information and corresponding classifier ensemble information. An Evolutionary Processor (EP) and Context Knowledge Base (CKB) are used for optimizing and saving the context knowledge. Finally, a Result Evaluator (RE) decides whether the classification result is incorrect and Data Accumulator (DA) saves the data in case of further training.

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

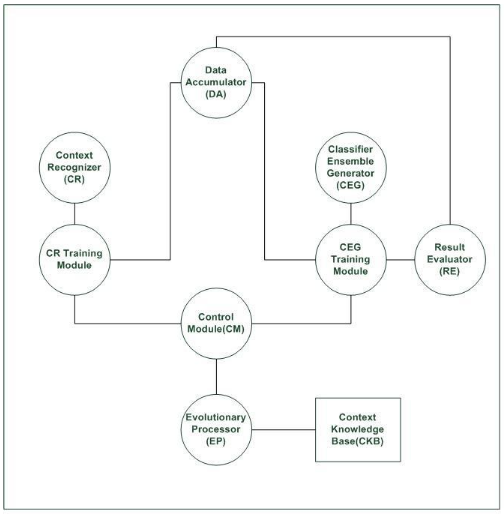

Figure 1. Basic Structure of online evolutionary Context-Aware classifier ensemble framework.

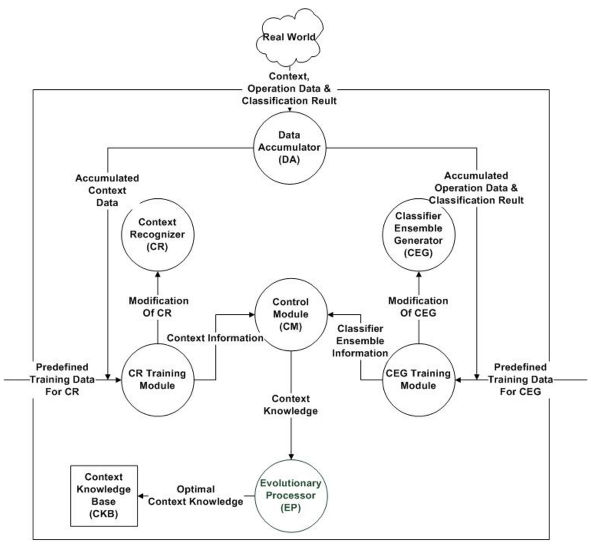

Figure 2. Flowchart of learning mode of the proposed framework.

III. Learning Mode and Operation Mode

A. Learning Mode

In this subsection we describe in detail the learning mode of our proposed framework (see Figure 2 for an illustration). We begin by considering CR Training Module, which takes predefined or accumulated context training data D as the input. The parameters of probabilities of all contexts are represented by parameter vector $\theta = (\theta_1, \theta_2, ..., \theta_c)^T$. The *mixture density* [15] for the data could be given by

$$p(D \mid \theta) = \sum_{i=1}^{c} p(D \mid \omega_i, \theta_i) P(\omega_i), \qquad (1)$$

where ω_i represents a state of nature to obtain D, $p(D \mid \omega_i, \theta_i)$ represents the component densities, and $P(\omega_i)$ represents the mixing parameters. Since various sources for context information and clues could be obtained [5], we assume here that we know the complete probability structure for the contexts and the probability density function $p(D \mid \theta)$ is identifiable. The procedure of unsupervised context learning is given by Figure 3.

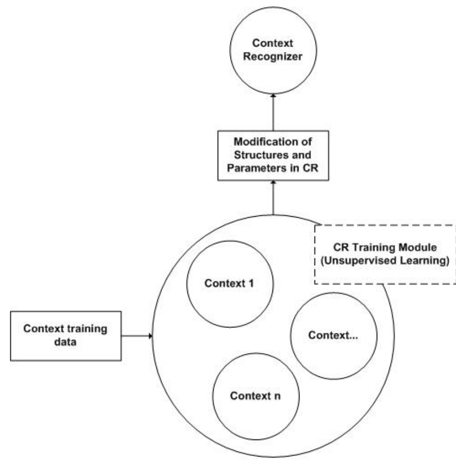

Figure 3. Context Learning Procedure

Another important issue for the CR Training Module is that the number of contexts is unknown most of the time. Because we want to use the context information to optimize the classifier ensemble, we need to make sure the context information itself is collected correctly. If the number of contexts is known, many approaches for a predefined number

Validity Checking

Given

- c : the number of context clusters
- E_t : validity criterion(e.g. sum-of-square-error)
- ε : criterion threshold
- n : number of samples

Do $c = c + 1$

- Calculate E_c
- Get the gradient descent g of E_c and E_{c-1}

Until $g < \varepsilon$ or $c = n$

Return c.

Figure 4. Validity checking procedure

of clusters such as Maximum-Likelihood, k-Means Clustering, and Unsupervised Bayesian Learning could be used. However, in most cases, we can't predict how many different kinds of contexts we are facing. Therefore, we first give certain c as the number of context clusters and then perform a validity checking procedure. The common validity checking procedure is shown in Figure 4. After we achieve a satisfying c as the number of contexts, many approaches as mentioned above could take c as the predefined number to determine θ.

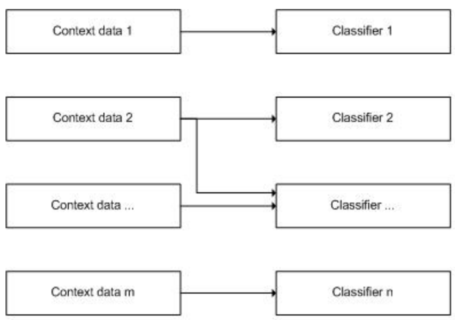

Figure 5. Training Classifiers

The detail of CEG training module is as follows: Let $C = \{C_1, C_2, ..., C_n\}$ be a set of classifiers in Classifier Ensemble Generator. As shown in Figure 5, we train each classifier C_i with data under different context conditions (one context can be assign to more than one classifier). After all classifiers have been trained to satisfying levels, we follow a supervised learning procedure to optimize the classifier ensemble regarding to the given context information. Let $\omega = \{\omega_1, \omega_2, ..., \omega_d\}$ and $D' = \{D'_1, D'_2, ..., D'_e\}$ be the training data. Each D'_i has the predefined class label $\omega_{Di} \in \omega$, $D'_i \in D'$, $i \in [1, e]$. Given context information CI_l, Classifiers in C take $v \in D'$ as their input and assign it to a class label in ω, i.e., $D' \rightarrow \omega$, $j \in [1, n]$. Suppose the estimate of classifier C_j gets a output ω_i with input D'_k is $P(\omega_i \mid C_j, D'_k, CI_l)$, we have

$$\sum_{i=0}^{d} P(\omega_i \mid C_j, D'_k, CI_l) = 1 \ . \qquad (2)$$

Therefore we can acquire this *decision profile* [16]

$$DP(v) = \begin{bmatrix} P(\omega_1 \mid C_1, v, CI_l) & P(\omega_2 \mid C_1, v, CI_l) & ... & P(\omega_d \mid C_1, v, CI_l) \\ P(\omega_1 \mid C_2, v, CI_l) & P(\omega_2 \mid C_2, v, CI_l) & ... & P(\omega_d \mid C_2, v, CI_l) \\ ... & ... & ... & ... \\ P(\omega_1 \mid C_n, v, CI_l) & P(\omega_2 \mid C_n, v, CI_l) & ... & P(\omega_d \mid C_n, v, CI_l) \end{bmatrix}, \quad (3)$$

where each row of the decision profile is one output of a classifier and each column is the probability of obtaining a class label. In [7, 13, 14, 17], many classifier ensemble methods have been proposed such as Rotation Forest, Boosting, Trainable Combiners and Fuzzy Integral. With methodologies carefully selected, the classifier ensemble results could be satisfying.

The task for the Evolutionary Processor (EP) is basically a global optimization problem [17], which is designing algorithms that can find the optimal combinations of contexts and classifier ensembles. The problem can be represented mathematically by:

Find one of global minima $x^* \in G$ for which:

$$x^* = \min(\{ f(x) \mid x \in G\}) \ , \qquad (4)$$

where

$$G \in \{u, v : u \in contexts, v \in classifier\ ensembles\} \ . \qquad (5)$$

The searching procedure adopts ideas from the evolution under the natural selection and performs massive parallel search through G of possible solutions.

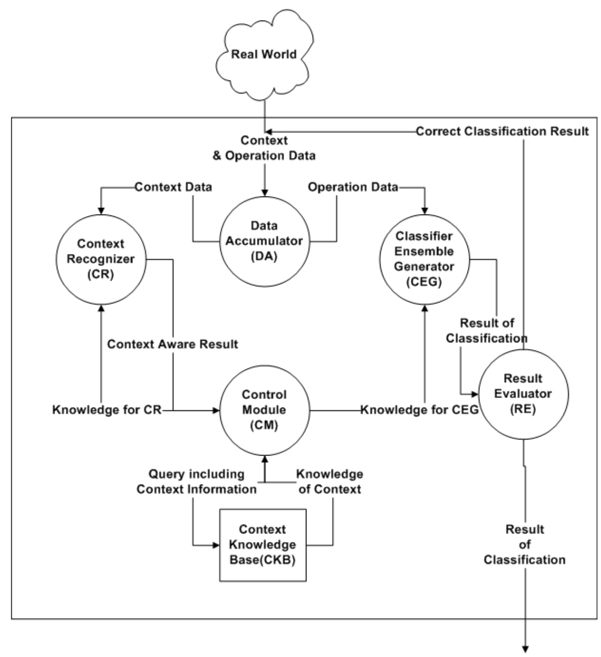

Figure 6. Flowchart of operation mode of the proposed framework.

B. *Operation Mode*

Figure 6 shows the operation mode of our proposed system. The data obtained from the real world could be subcategorized into context data D_c and operation data D_o. Then the context information will be achieved in CR by

$$c = CR(D_c) \ . \qquad (6)$$

After the Control Module (CM) generates the query q from c, q is sent to Context Knowledge Base (CKB) to acquire the corresponding knowledge for constructing classifier ensemble in Classifier Ensemble Generator (CEG), and the generated classifier ensemble will acquire the classification result from the operation data. We have

$$r = CEG(o,i), \tag{7}$$

where o is the operation data CEG, i is the information for generating appropriate classifier ensemble and r is the classification result. Result Evaluator (RE) judges whether r should be saved in Data Accumulator (DA) for further online training.

IV. EXPERIMENTS

We applied our framework on a face recognition system due to its wide range of applications and high level of difficulty [18-20]. Our system was tested using an AMD Turion 64*2 mobile technology laptop with 2GB of RAM. The algorithms were implemented in Microsoft Visual Studio 2005. Good Results achieved under different kinds of environments verified the robustness and adaptiveness of the framework. The experiments are discussed as follows.

A. Unsupervised learning for contexts

Since the number of the context categories is unknown to the system, based on our framework, we embed the validity concept inside clustering method to create a novel online context learning algorithm.

To solve the validating problem, we advanced the hypothesis that there are exactly n clusters present and $E(n)$, the error, is calculated by:

$$E(n) = \sum_{i=1}^{n} \sum_{x \in D_i} \| x - m_i \|^2 , \tag{8}$$

where we partition the set of samples D, which has the size of s, into n subsets in order to minimize $E(n)$. According to [15], for large s, it can be shown that the distribtuion of $E(n)$ is approximately normal with mean $s(d-\frac{2}{\pi})\Sigma$, and variance $2s(d-\frac{8}{\pi^2})\Sigma^2$, where d is the dimension of x and Σ is the covariance matrix of x in D. We initialize n to 1. If we have

$$\frac{E(2)}{E(1)} < 1 - \frac{2}{\pi d} - a\sqrt{\frac{2(1-8/\pi^2 d)}{nd}} , \tag{9}$$

where a is decided by

$$p = \int_a^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\mu^2/2} du , \tag{10}$$

we could reject the hypothesis at the significance of p. This can easily be applied to any n by doing the same test. If the hypothesis is rejected, we advance a new hypothesis that there are exactly $n+1$ clusters present, and do the validaty test again. The clustering algorithm for the learning problem with unknown clusters is given by Figure 7.

Clustering Algorithm For Unknown Clusters

Given

- η : learning rate
- θ : threshold
- C_i : context
- x : context data

Do accept new x

- $j = \arg\min_{j'} \| x - m_{Cj'} \|$
- $if \| x - m_{Cj} \| < \theta, then\ m_{Cj} = m_{Cj} + \eta x$
- *else add new x to m*
- Normalize m

Until no more patterns

Return.

Figure 7. Clustering Algorithm For Unknown number of Clusters

The context learning algorithm was tested on the data set of 40000 face images under 400 different kinds of contexts caused by changing of illumination level, illumination direction as well as background. The clustering correctness R which is calculated according to (8) should be

$$R = -E(n), \tag{11}$$

where n is the number of contexts. The test was performed by starting with data under ten kinds of contexts and gradually adding data under more kinds of contexts into the input of the algorithms. Here we use a set of three algorithms, k-means defined for 200 clusters, unsupervised Bayesian learning defined for 200 clusters, and our proposed online clustering algorithm. As shown in Figure 8, we can see that when the number of contexts of the data varies, the correctness of our algorithm remains high, while using conventional clustering algorithms for predefined number of contexts, satisfying results could only be achieved when the number of contexts in the given data is in the neighborhood of the predefined number.

TABLE I. FACE RECOGNITION SYSTEM PERFORMANCE COMPARISON

	Neural Network	Nonfuzzy Boosting Combination	Linear Oracle Ensemble	Our System
E-Yale I	74.41%	80.11%	81.26%	85.74%
E-Yale II	73.22%	81.23%	82.74%	85.61%
E-Yale III	75.10%	79.50%	80.41%	85.70%
E-FERET I	71.58%	80.17%	81.33%	85.88%
E-FERET II	79.43%	82.67%	79.15%	85.74%
E-FERET III	70.81%	80.14%	80.10%	85.89%
E-INHA I	76.09%	81.31%	82.74%	85.26%
E-INHA II	72.75%	81.09%	82.06%	86.10%
E-INHA III	69.38%	80.20%	81.95%	85.87%
IT I	73.71%	79.46%	79.13%	87.34%
IT II	70.65%	82.77%	81.22%	86.98%
IT III	72.82%	81.94%	80.70%	87.15%

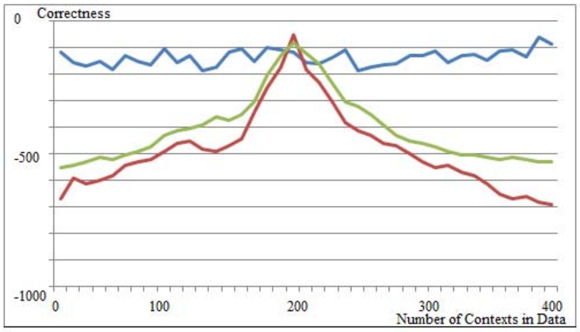

Figure 8. Clustering correctness comparison among our algorithm (blue), k-means defined for 200 clusters (red), and unsupervised Bayesian learning defined for 200 clusters (green).

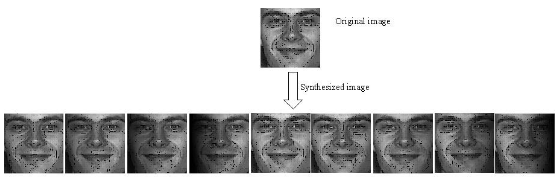

Figure 9. Artificial illumination procedure.

Figure 10. Examples of our real data set.

B. *Online Context-Aware Classifier ensemble result*

We artificially illuminated FERET, Yale and Inha data sets by 300 different kinds of illumination conditions (see Figure 9 for details) and generated E-Yale, E-FERET, and E-INHA data sets (E represents Extended). These extended data sets were further divided into nine data sets, each containing 10000 face images under 100 different kinds of illuminations. Three real data sets captured from our lab, each including 1000 face images under 10 different kinds of illumination conditions, were also added to the experiment (Figure 10). Four face recognition systems were trained with appropriate face images and tested with these data sets. As illustrated in Table I, our proposed system was found to yield the highest recognition rate and the most stable performance.

For the purpose of testing the robustness of our system with online learning functionalities, we created another similar

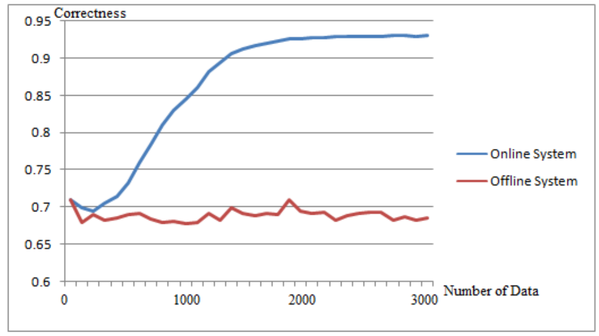

Figure 11. Online and Offline System Performance comparison.

system based on the same algorithms. The only difference was that the second system (offline system) didn't have a Result Evaluator (RE). The two systems were trained by the same gallery data sets and tested on new probe data sets with different illumination conditions from the training data. Figure 11 shows that the online system could gradually adapt itself to the new data with different illumination conditions while the system without online learning abilities kept giving a poor performance.

V. CONCLUSION

In this paper, we describe the idea of an online evolutionary Context-Aware classifier ensemble framework for object recognition systems. This framework exhibits not only a degree of accuracy, but also robust regarding the context changing results on our system.

The strength of our framework comes from embedding classifier ensemble into a Context-Aware Framework. The encouraging results of our face recognition system suggest that this framework could be used in other object recognition fields.

ACKNOWLEDGMENT

The authors would like to thank the reviews for their comments.

REFERENCES

[1] Gu, T., Pung, H.K., and D.Q. "Toward an OSGi-Based Infrastructure for Context-Aware Application," Pervasive Computing, IEEE, Vol. 3, Issue 4, pp. 66-74, 2004.

[2] Novotny, M., Hauser, H. "Outlier-preserving Focus+Context Visualization in Parallel Coordinates," IEEE Trans. on Visualization and Computer Graphics, Vol. 12, Issue 5, pp. 893-900, 2006.

[3] De Rocha, R.C.A., "Middleware: Context Management in Heterogeneous, Evolving Ubiquitous Environments," Distributed Systems Online, IEEE, Vol. 7, Issue 4, pp. 1-1, April 2006.

[4] Yap, Ghim-Eng; Tan, Ah-Hwee; Pang, and Hwee-Hwa, " Discovering and Exploiting Causal Dependencies for Robust Mobile Context-Aware Recommenders," IEEE Trans. on Knowledge and Data Engineering, Vol. 19, Issue 7, pp. 977-992, July 2007.

[5] Mantyjarvi, J., Huuskonen, P., and Himberg, J., "Collaborative Context Determination to Support Mobile Terminal Applications," Wireless Communications, IEEE, Vol. 9, Issue 5, pp. 39-45, Oct. 2002.

[6] Jianhua Chen, “Context Modeling Based on Context Quantization with Application in Wavelet Image Coding”, IEEE Trans. on Image Processing, Vol. 13, Issue 1, pp. 26-32, Jan. 2004.

[7] Kuncheva, L.I., ““Fuzzy” Versus “Nonfuzzy” in Combining Classifiers Designed by Boosting”, IEEE Trans. on Fuzzy Systems, Vol. 11, Issue 6, pp. 729-741, Dec. 2003.

[8] L. Breiman, “Combination predictors,” in *Combining Artificial Neural Nets*, A.J.C. Sharkey, Ed. New York: Springer-Verlag, 1999, pp.31-50.

[9] Kuncheva, L.I., “A theoretical study on expert fusion strategies,” Manuscript, 2000

[10] Kittler, J., Hatef, M., Duin, R.P.W., and Matas, J, “On Combining Classifiers,” IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 20, Issue 3, pp. 226-239, March 1998.

[11] Magimai-Doss, M., Hakkani-Tur, D., Cetin, O., Shriberg, E., Fung, J., and Mirghafori, N., “Entropy Based Classifier Combination for Sentence Segmentation,” IEEE International Conference on Acoustics, Speech and Signal Proccsing, 2007, Vol. 4, pp. VI 189-192, April 2007.

[12] Byungki Byun, Chengyuan Ma, and Chin-Hui Lee, “An experimental study on discriminative concept classifier combination for TRECVID high-level feature extraction,” IEEE International Conference on Image Processing, 2008, pp. 2532-2535, Oct. 2008.

[13] Rodriguez, J.J., Kuncheva, L.I., and Alonso, C.J., “Rotation Forest: A New Classifier Ensemble Method,” IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 28, Issue 10, pp. 1619-1630, Oct. 2006.

[14] Kuncheva, L.I., Rodriguez, J.J., “Classifier Ensembles with a Random Linear Oracle,” IEEE Trans. on Kownledge and Data Engineering, Vol. 19, Issue 4, pp. 500-508, April 2007.

[15] Richard O. Duda, Peter E. Hart, and David G. Stork, “Pattern Classification,” Stork;John Wiley & Sons, Inc., 2000.

[16] L. I. Kuncheva, J. C. Bezdek, and R. P. W. Duin, “Decision templates for multiple classifier fusion: An experiment comparison,” Pattern Recogn., Vol. 34, no. 2, pp. 299-314, 2001.

[17] A. Torn, and A. Zilinskas, “Global optimization,” New York: Springer-Verlag 1989.

[18] Chen, W., Meng Joo Er, and Shiqian Wu, “Illumination Compensation and Normalization for Robust Face Recognition Using Discrete Cosine Transform in Logarithm Domain,” IEEE Trans. on Systems, Man, and Cybemetics, Part B, vol 36, Issue 2, pp. 458-466, April 2006.

[19] R. Chellappa, C.L. Wilson, and S. Sirohey, “Human and machine recognition of faces: a survey,” Proc. IEEE, Vol. 83, no. 5, pp. 705-740, May 1995.

[20] Adini, Y., Moses, Y., and Ullman, S., “Face Recognition: The Problem of Compensating for Changes in Illumination Direction,” IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, Issue 7, pp. 721-732, July 1997.