

A Web-Based Face Recognition System Using Mobile Agent Technology

¹Md. Geaur Rahman, ²Somlal Das, ²A.R.S. Ahmed Siddique
and ²Md. Khademul Islam Molla

¹Department of Computer Science and Mathematics,
Bangladesh Agricultural University, Mymensingh, Bangladesh

²Department of Computer Science and Engineering,
University of Rajshahi, Rajshahi, Bangladesh

Abstract: This study presents an application of mobile agent technology to solve the problem of web-based face recognition. Faces represent complex, multidimensional, meaningful visual stimuli and developing a computational model for face recognition is difficult. The current face recognition models, which suffer from slow performance and platform dependence. The proposed system is a network-oriented system involving not only face detection but also facial feature extraction and graph matching schemes. For high performance in terms of reduced processing time and better flexibility, we introduce an innovative four-layer structural model and a three-dimension operational model. In addition, the proposed system integrates the computing advantages of mobile agents with several improved face recognition algorithms to enhance system robustness. Preliminary experimental results demonstrate the advantages and potentialities of the approach for face recognition on the web.

Key words: Mobile agent, face recognition, feature extraction, remote graph-matching test, Bangladesh

INTRODUCTION

Web-based face recognition systems have received more and more attention in recent years. Face Recognition (FR) is not a simple problem since a new image of a face seen in the recognition phase is usually different from the images previously seen by the system in the learning phase. There are several sources for these variations between images of the same face. The image depends on viewing conditions, device characteristics and environment. This includes viewing position (which determines the orientation, location and size of the face in the image), imaging quality (which influences the resolution, blurring and noise in the picture) and the light source (which influences the reflection). In addition to this the face is a dynamic object and it changes according to expressions, mood, age, hairstyle, glasses, etc. (Rahman *et al.*, 2003).

A Web-Based Face Recognition System (WBFRS), which is flexible and efficient should be able to solve these problems. Early attempts to model FR system generally uses a geometric coding in which measurements of the relations between features (as eyes, the nose, the mouth, the chin, etc.) were coded and used for recognition purpose. Aslandogan and Yu (2000)

developed a web-based image search agent named Diogenes which takes advantage of the text/HTML structure of web pages as well as the visual analysis of the images for personal image search and identification. Diogenes works well with web pages that contain a facial image accompanied by a body of text that contains textural identification of the image.

However in this system, the intelligent agent does not get real mobility when it traverses web pages. It can not migrate from host to host in a heterogeneous Internet environment. This disadvantage makes it difficult to incorporate it with legacy systems when the system has to access image databases in different formats over the Internet.

In this study, we propose a WBFRS using mobile agent technology. It is a network-oriented system involving not only face detection but also facial feature extraction and graph matching schemes. The mobile agent displays its full computing advantages in all system sub-layers.

MATERIALS AND METHODS

Mobile agent paradigm: Comparing with earlier paradigm such as process migration or remote evaluation in

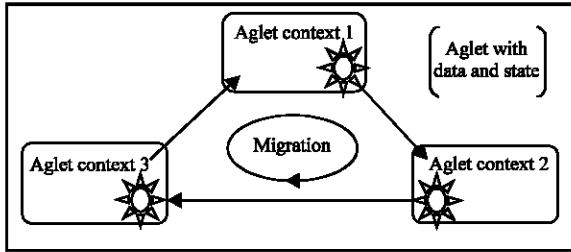


Fig. 1: Basic migration paradigm of an aglet

distributed computing, mobile agent model is becoming popular for network-centric programming. Traditional client/server paradigm relies on handshake mechanism to communicate over a network. The client requests information, while the server responds. Each request/response has to be a complete round trip on the network. The emerging mobile-agent paradigm has redefined the way Internet-based applications work. As an autonomy software entity with pre-defined functionality and certain intelligence, mobile agent is capable of migrating autonomously from one host machine to another making its request to the server directly and performing tasks on behalf of its master.

The Aglet system developed by IBM is chosen as the implementation example in the proposed system. Although, it is not a full-fledged platform until now, it has received the most press coverage and shows promises as a functional technology that fits very well into the JAVA world. Several phrases can be summarized to characterize an aglet: lightweight objects migration built with persistent support, event driven and so on. Central to the aglet architecture is the context which is the server environment for aglet execution. When an aglet has finished its work in a context, its state and data will be serialized to a stream of bytes and exported to the new context through Agent Transfer Protocol. In a reverse process, the state of the aglet can be reconstructed from the stream of bytes and become active at the new context (Lange and Oshima, 1998). The basic migration paradigm of aglet can be shown in Fig. 1.

System framework: To achieve high performance and better robustness, we propose a multi-agent system. The system framework can be viewed in two ways: one is the structural model and the other is the operational model. Fig. 2 shows the four-layer system structural model:

- The point-to-point (P2P) input/output layer is a P2P communication channel between the input/output device and the application server
- The line-like central controller layer controls all inner recognition schemes as well as implementing an intelligent detection strategy in a line-like fashion

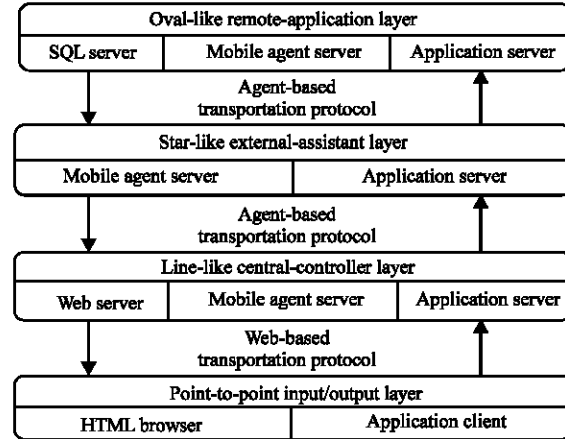


Fig. 2: Four layer system structural model

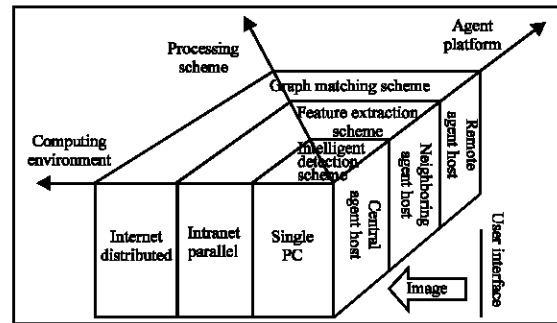


Fig. 3: Three-dimensional system operational model

- The star-like external assistant layer includes a group of external worker agents to help implementing the feature extraction scheme
- The oval-like remote application layer is the outer application part of the system which is used for connecting remote legacy databases

From an operational point of view, we propose a three-dimensional model consisting of three blocks, as shown in Fig. 3:

- A central agent host blocks implementing an intelligent detection scheme on a single PC
- A neighboring agent host blocks implementing a feature extraction scheme in a parallel-computing environment
- A remote agent host blocks implementing a remote graph-matching scheme in a distributed computing environment

The point-to-point input layer: To make accessing the system as easy as possible, a point-to-point input/output

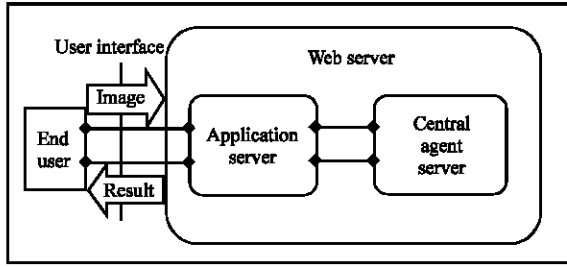


Fig. 4: The structure of the point-to-point input/output layer

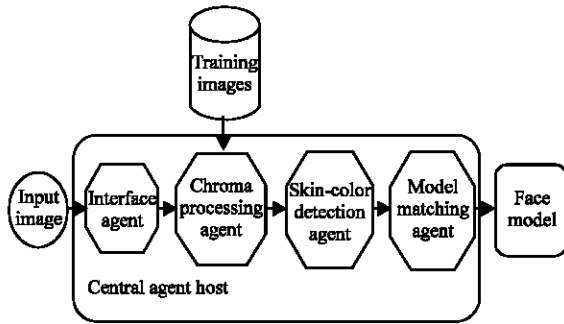


Fig. 5: The structure of the line-like central controller layer

layer is established explicitly connecting the input/output device and the application server. All recognition schemes are hidden from the end user. A user interface is used to accept the input image and return recognition results. The structure of the point-to-point input/output layer is shown in Fig. 4.

The line-like central controller layer: The central controller layer is the main part of the whole system connecting all sub-systems. A line-like intelligent detection scheme is implemented in this layer. Four agents reside at the central controller host: the interface agent, the chroma processing agent, the skin-color detection agent and the model matching agent. The structure of the line-like central controller layer is shown in Fig. 5.

IA and Chroma Processing Agent (CPA): The Interface Agent (IA) performs the interaction between the input and CPA. It takes the images as its input, then processes those images and finally sends those to the CPA. The chroma chart is the key element for automatic skin color region extraction. For example in the research of (Cai *et al.*, 2000), information about skin color (embedded in a chroma chart) has been used to find likely image regions where faces may exist. The knowledge about facial patterns (distribution of non-skin sub-regions) will be used to determine instances of face within such

regions. However in their research, the non-uniformity in the distribution of skin-colors weakens the minority skin colors because the statistical results are smoothed directly using Gauss filters. To solve this problem, we encapsulate a novel algorithm in the proposed CPA and all the skin-colors are taken as the same once the statistical value for one chroma exceeds the selected threshold.

Skin-Color Detection Agent (SCDA): Mathematical morphology analyses images based on concepts from algebra and geometry, such as set theory, translation, convexity and so on. Due to the pioneering research of Matheron (1975) and Serra (1982), it has been successfully used in contour retrieval, segmentation, shape analysis, etc. In the proposed system, we introduce mathematical morphology into the SCDA in order to suppress all kinds of noises and improve the reliability of the detecting result. The SCDA uses morphological operators to divide different convex objects, removing regions that are too small and recovering regions' sizes, while keeping the same topological structure.

Model Matching Agent (MMA): The majority of face detection methods are based on the model matching scheme, which combines prior knowledge about the face with lower-level processing results. In the proposed MMA, we introduce a model face and a matching function for model matching with emphasis on the topological characteristics which reduces the computation cost and improves matching efficiency.

The matching function includes two parts: one for topological constraints, the other for the geometric constraints of each facial feature. For each topological or geometrical attribute, the expected value and variance can be determined through experiments. Through changing the size of the Gauss filter, this method can detect faces of different sizes in an image avoiding noises in lower-level processing and forming a complete description of the face. The similarity item is defined in Eq. 1 and Fig. 6, where x is the difference between the measured and expected values of each topological attribute and σ is the reasonable fluctuation range. The matching function is defined as the weighting sum of all these similarity items.

$$f(x) = \begin{cases} 1 - \left(\frac{x}{\sigma}\right)^2, & |x| < \sigma \\ 0, & |x| \geq \sigma \end{cases} \quad (1)$$

The star-like external-assistant layer: To deal with the complexity of a facial feature's contour, we introduce the Multi-GVF (Gradient Vector Flow)-snakes paradigm in this layer. This approach uses open snakes to represent smooth contours and the cross points of two snakes for

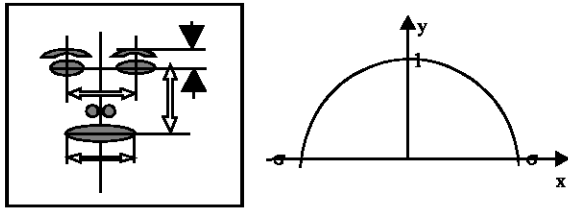


Fig. 6: Model face and similarity item

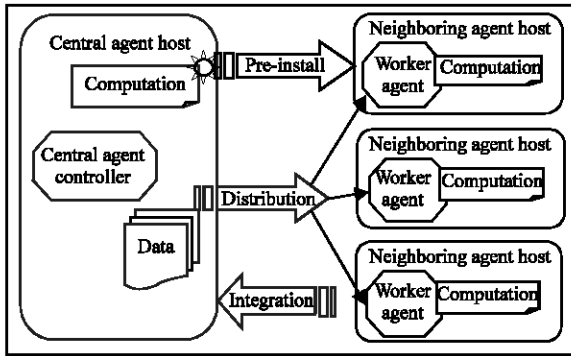


Fig. 7: The structure of the star-like external assistant layer

sharp corners. It performs better than the single snake model in accuracy but its long execution time is a penalty all the facial features (mouth, nose, eyes and so on) have to be extracted separately and sequentially.

To address the mentioned problem, we employ a group of external computation agents and adopt the Divide and Conquer strategy in this layer. The central agent controller dissects large amount of computation data into several operational portions and dispatches these smaller computation units to the neighboring worker agents for execution in parallel. A global result will be computed after all the sub-results are sent back to the central host for assembly. The system structure of the star-like external assistant layer is shown in Fig. 7.

The GVF was introduced to overcome the limitations of the traditional snake model in initialization for convergence to concave boundaries (Xu and Prince, 1998). It strongly improves the convergence of the classic snake towards the desired solution and always has a large capture range barring interference from other objects. However, it proves some lack of performance in the presence of corners because of the diffusion effects. Thus a single closed snake can extract a contour robustly only when the features are smooth. In the proposed star-like layer, we introduce the Multi-GVF-snakes paradigm and integrate it with the mobile agent's worker model to address these limitations and achieve high efficiency.

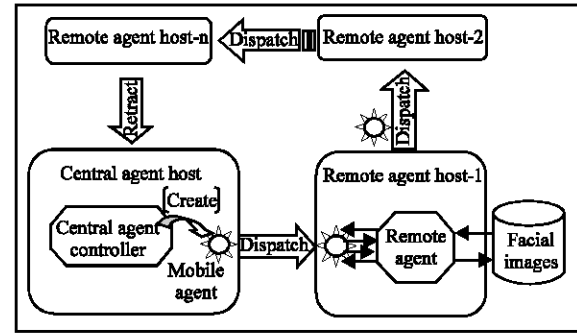


Fig. 8: The structure of the oval-like remote application layer

The oval-like remote application layer: Remote graph-matching scheme has been implemented in this layer. Facing the fact that large image databases with different formats might be located in different places over the network, it makes more sense if the mobile agent moves to the remote data source for searching and matching, rather than transferring large volumes of data over the network for processing. In this layer, we explicitly create a matching agent initializing it with matching algorithms and dispatching it to the Internet. Upon reaching a new host, the matching agent interacts with remote agents and communicates with the backend databases for searching and matching.

A two-step-matching scheme will be performed: geometric-based coarse matching and dynamic-link architecture based accurate matching. After the matching agent has achieved its pre-defined goal, it will migrate to the next host until returning home with the results. Some advantages of this scheme can be summarized as reduced design work, better bandwidth usage and broader searching range. The structure of the oval-like remote application layer is shown in Fig. 8.

Geometric-based coarse matching: Geometric, feature based face recognition is among the earlier algorithms proposed (Kanade, 1973). As the image database becomes larger however, it turned out to be impossible to perform accurate recognition by simply using this scheme. However, we embed this algorithm into our proposed matching agent as an effective pre-filtering for the second-step accurate matching.

In this scheme, the overall geometrical configuration of the face features is described with a vector of numerical data representing the position and size of the main facial features, e.g., eyes, nose and mouth and supplemented by the shape of the face outline (Kaya and Kobayashi, 1972). Nearest Neighbor (NN) classifier can be used for

performing general, non-parametric classification. Its performance is a function of the number of classes to be discriminated (people to be recognized) and of the number of examples per class.

Dynamiclink architecture-based accurate matching: The Dynamic Link Architecture (DLA) is initially proposed to solve certain conceptual problems of conventional artificial neural networks. The power of DLA can best be demonstrated by applying it to a complex problem like position and distortion invariant object recognition (Buhmann *et al.*, 1990; Laddes *et al.*, 1993). Elastic matching of an object graph to an image graph is a pattern classification strategy, which clearly accounts for local distortions.

The cost function, which combines the topological and feature terms into a measure of distance between the object domain and the image domain is used to evaluate the quality of a match. Gabor wavelet is chosen as a suitable data format for object recognition, which proved to be invariant with respect to background, translation, distortion and size.

RESULTS AND DISCUSSION

System implementation: To implement and evaluate the system performance, we apply the IBM Aglet as the implementation example. Three sets of tests are conducted: Intelligent Detection Test (IDT) in the line-like layer, Feature Extraction Test (FET) in the star-like layer and Remote Graph-matching Test (RGT) in the oval-like layer.

As explained below, the IDT Test aims at evaluating the algorithm's robustness in the proposed detection agents under different shooting conditions. Then the FET test will calculate the speedup efficiency of the agent based Multi-GVF-snakes model under different worker patterns. Finally the RGT test will examine the correctness and effectiveness of remote graph matching scheme in a distributed environment.

Intelligent detection test: In order to obtain a stable chroma chart with fuzzy character, >2000 face images with different shooting conditions and different types of skin are chosen as the training samples for the CPA.

More than 200 face images were used for the IDT test. The experimental results showed in Fig. 9. Table 1 shows that the improved algorithms in the proposed detection agents are robust to the unevenness of lighting, multiple faces, moderate tilt of faces and partial sheltering out.

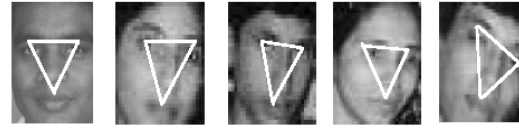
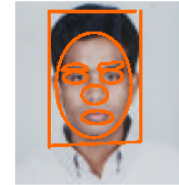


Fig. 9: Intelligent detection test



(a) Test image



(b) Final feature extraction

Fig. 10: Feature extraction test

Table 1: Intelligent detection test

Different shooting conditions	Unevenness of lighting	Multiple faces	Moderate tilt of faces	Partial sheltering
Correct detection rate	98%	95%	97%	85%

Feature extraction test: To examine the speedup efficiency of the proposed agent-based Multi-GVF-snakes model, a group of external assistant agents is employed. The central agent controller is responsible for dissecting the task and assembling the final result. The processing result is shown in Fig. 10. According to the number of available worker agents, different strategies will be adopted to partition and distribute sub-tasks.

Five/four-worker pattern: If all facial features-chin, left eye, right eye, mouth and nose-forms a set $X = \{c, l, r, m, n\}$, the five-worker strategy simply dissects the whole task into five parts and distributes sub-tasks to five worker agents. Actually, it takes no longer if $\{l, r\}$ is processed on one host because $\text{Host}(l + r) < \text{Host}(m)$ with the same CPU speed. The total processing time of the five or four-worker pattern is:

$$\text{Max}\{\text{Host}_1(c), \text{Host}_2(l + r), \text{Host}_3(m), \text{Host}_4(n)\} + \text{latencies}$$

Where latencies include network transmission and data packaging and unpacking.

Three-worker pattern: A three-worker pattern may be adopted due to the fact that $\text{Host}(l + r + n) \approx \text{Host}(m)$. Thus if $\{l, r, n\}$ is processed on one host, the total execution time is:

$$\text{Max}\{\text{Host}_1(c), \text{Host}_2(l + r + n), \text{Host}_3(m)\} + \text{latencies}$$

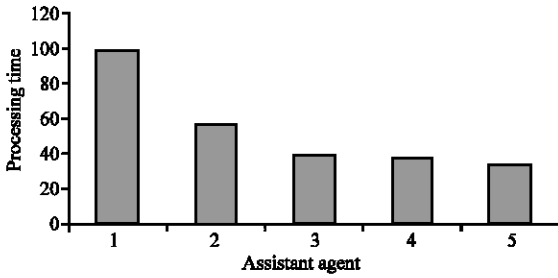


Fig. 11: The speedup ratio of different worker patterns

Table 2: Viewing perspectives test

Viewing perspective	From vertical axis (Degrees)				From horizontal axis (Degrees)			
	-20	-10	+10	+20	-20	-10	+10	+20
Correct classification (%)	84	86	86	83	85	87	87	84

Table 3: Occlusion and distortion test

Pattern occlusion and distortion	Partial occlusion	Various expressions	Wearing glasses
Correct classification	77%	81%	82%

Two-worker pattern: A two-worker pattern is the minimum requirement for this system structure. A simple algorithm is introduced to decide how to schedule tasks between two hosts. The task subset A ($A \subset X$, $A \neq \Phi$) that needs to be handled in host 1 must satisfy the formula:

$$\text{Min}_{A \subset X, A \neq \Phi} \left\{ \sum_{x \in A} \text{Host}_1(x) - \sum_{x \in A} \text{Host}_2(x) \right\}$$

and the total processing time is:

$$\text{Max}_{A \subset X, A \neq \Phi} \left\{ \sum_{x \in A} \text{Host}_1(x), \sum_{x \in A} \text{Host}_2(x) \right\} + \text{latencies}$$

The speedup ratio of different worker patterns is shown in Fig. 11.

Remote graph-matching test: To examine the correctness and effectiveness of the remote graph matching scheme in the oval-like layer, two sets of tests are conducted: the viewing perspective test and the pattern occlusion and distortion test.

Viewing perspectives test: In this test, about 100 test patterns with different viewing perspectives are examined. Sample patterns and recognition results are presented in Table 2. The overall correct recognition rate is about 85%.

Pattern occluded test: In this test, another 100 test patterns with partial occlusion, various expressions or wearing accessories are examined.

Sample patterns and recognition results are presented in Table 3. It should be explained that the classification rate of the partial occlusion test is heavily dependent on which portion and how much of the facial image is hidden.

CONCLUSION

This study presents a web-based face recognition system using mobile agent technology. In contrast to current face recognition models, which suffer from slow performance and platform dependence, a four-layer system structural model and a three-dimension system operational model are introduced to achieve high performance and better flexibility.

The proposed system model with several improved algorithms has been tested through experiments, demonstrating its feasibility and effectiveness.

Coupled with other supporting schemes, the system is potentially useful in a wide range of Internet-based face recognition services.

REFERENCES

- Aslandogan, Y.A. and C.T. Yu, 2000. Diogenes: A web search agent for content based indexing of personal images. Proceedings of ACM SIGIR, (SIGIR'00), Athens, Greece, pp: 1-19.
- Buhmann, J., M. Lades and C. von der Malsburg, 1990. Size and distortion invariant object recognition by hierarchical graph matching. Proceedings of International Conference on Neural Networks, June 21-17, San Diego, CA., pp: 411-416.
- Cai, J., A. Goshtasby and C. Yu, 2000. Detecting human faces in color images. Image Vision Comput., 18: 63-75.
- Kanade, T., 1973. Picture processing system by computer complex and recognition of human faces. Technical Report, Kyoto University, Department of Information Science, pp: 1-148. http://www-old.ri.cmu.edu/pub_files/pub3/kanade_takeo_1973_1/kanade_takeo_1973_1.pdf.
- Kaya, Y. and K. Kobayashi, 1972. A Basic Study on Human Face Recognition. Academic Press, New York, pp: 265-289.

- Laddes, M., J.C. Vorbruggen, J. Buchman, J. Lange, C.V. Malsburg, R.P. Wurtz and W. Konen, 1993. Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans. Comput.*, 42: 300-311.
- Lange, D.B. and M. Oshima, 1998. *Programming and Deploying Java Mobile Agents with Aglets*. Addison Wesley, New York.
- Matheron, G., 1975. *Random Sets and Integral Geometry*. 1st Edn., Wiley, New York.
- Rahman, M.G., M.A. Hossain, D. Das and M.K.I. Molla, 2003. Gender identification system from facial image using artificial neural network. *Proceedings of 6th International Conference on Computer and Information Technology*, Dec. 19-21, Jahangirnagar University, Dhaka, pp: 200-205.
- Serra, J., 1982. *Image Analysis and Mathematical Morphology*. Academic Press, New York.
- Xu, C. and J.L. Prince, 1998. Snakes, shapes and gradient vector flow. *IEEE Trans. Image Process.*, 7: 359-369.