Is High Resolution Representation More Effective for Content Based Image Classification?

Liang Chen
Computer Science Program, University of Northern British Columbia,
Prince George, B.C., Canada V2N 4Z9
E-mail: lchen@ieee.org
Naoyuki Tokuda
SunFlare Company, Shinjuku Hirose Bldg, Yotsuya 4-7, Shinjuku-ku,
Tokyo, Japan 160-0004
Email: tokuda_n@sunflare.co.jp

***Abstract*— This paper shows by a mathematical model that, for image classification / recognition purposes, high resolution pictures have lower recognition rate than relatively lower resolution pictures. The analysis is based on the matching approach by a simple neural network, but we believe that the conclusion remains valid even when the classification process involves complicated matching schemes such as principal component analysis and Gabor transforms.**

***Index Terms*— content based image classification, face recognition, image processing, stability, vision and scene understanding, perceptual reasoning, similarity measures**

I. Introduction

Current technique has made 5 million to 7 million pixel cameras affordable to non-professional users. It is beyond the authors knowledge to know if it is the best choice for professional photographers to take pictures with highest resolution digital cameras; it is known that human subjects can recognize human faces even from very coarse and low quality images, from distances, and under degraded viewing conditions [1], [2], [3], [4]. It is believed that low resolution images may be processed in a more general or holistic manner [5]. If computing technique can well simulate human subjects, there is no reason to believe that, for face recognition and general content based image retrieval purposes, extremely high resolution representations have any advantage. While taking further into consideration the computation and storage costs, relatively low resolution representations should be the best choice.

Recent experiments of appearance-based matching approaches in human face recognition seem to have shown that high resolution pictures have actually lower recognition rate than relatively lower resolution pictures[6], [7]. The purpose of this paper is to develop a mathematical model to show why the recognition / classification performances for relatively low resolution pictures are better that those for high resolution pictures. We will show that the matching algorithm work best on intermediate resolution representations.

While there are many appearance-based matching approaches that have been developed for content based pattern classifications which take an object / image as a vector of global features, we consider a simple model where a neural network is used to determine the classification of an object / image.

We suppose the high resolution images are of size $l \times w$ hereafter. A neural network used for such a classification/recognition problem is usually a multi-layer neural network with $l \times w$ inputs, some (more or less) hidden neurons, and one or several outputs. We have no interest in discussing the internal structure of a neural network in this paper; thus we represent a neural network as figure 1. We expect that, although it may finally generate some errors, the output(s) of the neural network with any input array $\mathbf{x}^{l\times w}$ should agree with, or output directly, the label of the object that the input intensity array represents.

While the choices of high/low resolution representations of an image usually come from the quality or setting of digital cameras, low resolution representation of an image can be obtained from a high resolution image by firstly partitioning the image into equal sized windows and then merging the pixel in each window into one pixel of which the intensity is the averaged value of all the pixel in the window. For convenience of discussion, we assume the low resolution images are obtained from high resolution images by above calculation. Under this assumption, there is a one-to-one correspondence between a pixel in the low resolution representation of an image and a window of pixel in the original high resolution representation image. The neural network for the classification of low resolution images will be the same as that for high resolution images, except that the input array is smaller.

We will study the stabilities of the neural network approach for the classification of images with high resolution representations and low resolution representations.

The rest of the paper is constructed as follows: We introduce the basic terms used in Section II. The main theorems regarding stability are given in Section III. The result is fully provided in Section IV, with experimental verification of the theory. The discussion and further work are given in Section V.

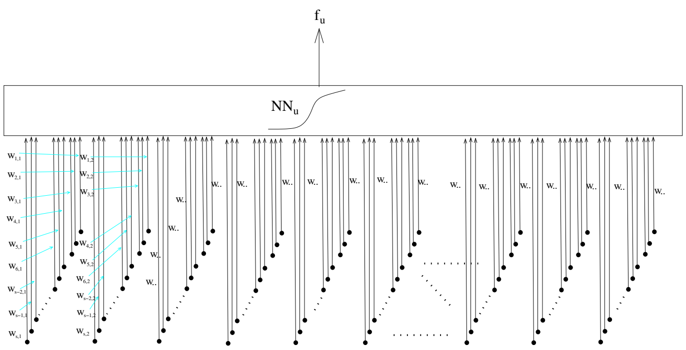

Fig. 1. Neural Network for Image Classification

II. Basic Models and Assumption

A. Bi-label classification problem description

We assume the images are bitmap (black-white) image, where the value of each pixel can be either "1" or "0", representing "black" or "white", and suppose the class label set $L = \{''+'', ''-''\}$.

1. For the convenience of discussion, we only consider a bi-label bitmap image classification problem. But note that the following derivation can also be generalized to M-label classification problems for $M > 2$. Each of the original images is represented by an $l \times w = N$ (l and w being very large integers) array of pixel. The value of a cell, is either *+1* or *-1*, which represents *black* or *white* [1]. We are to classify an image into one of the two clusters *Black* and *White*, by counting the $+1$-pixels and -1-pixels.
2. Let α and β denote the percentages of $+1$-pixels and -1-pixels in the absence of noise; we have $\alpha + \beta = 1$. In the absence of noise, the image will be classified as *black* image, so that we have $1 > \alpha > \beta > 0$.
As we suppose we are using high resolution images, i.e., l and w are very large integers, we could regard α (and β) as the probability of the pixel taking the value $+1$ (and -1) respectively when we arbitrarily choose one pixel in the image.
3. The low resolution representation of the original image is obtained as follows: the image area $l \times w$ is divided into equal shaped windows of size $r_l \times r_w$ (r_l and r_w being positive integers), where l and w are divisible by r_l and r_w independently, then the pixel in each window are merged into one pixel with value calculated by using a simple majority principle on the pixel in the window.
4. The classification of an image, either high resolution or low resolution, is implemented over the entire image by applying simple majority principle after counting the numbers of $+1$-pixels and -1-pixels.

We only discuss the simplified neural network as shown in figure 1, where the block NN$_u$ denotes a single output neuron with the sign function as its activity function. Hence the output of the neural network on input array $\mathbf{x}^{l\times w}$ is:

$$f_u(\mathbf{x}^{l\times w}) = \begin{cases} +1, & \text{if } \sum_{1\le i\le l, 1\le j\le w} w_{i,j}x_{i,j} \ge 0, \\ -1, & \text{if } \sum_{1\le i\le l, 1\le j\le w} w_{i,j}x_{i,j} < 0; \end{cases}$$

where $x_{i,j}$ ($1 \le i \le l$, $1 \le j \le w$) is an element of $\mathbf{x}^{l\times w}$.

B. Training Data

We assume that we have a very large and perfect data set for training purposes so that it covers all the possible situations. Ideally, there are $2^{l\times w}$ samples, and the label of each sample follows the following equation:

$$L = \begin{cases} ''+'', & \text{if } \sum_{i,j} x_{i,j} \ge 0, \\ ''-'', & \text{if } \sum_{i,j} x_{i,j} < 0; \end{cases} \tag{1}$$

where $x_{i,j}$ is an element of array $\mathbf{x}^{l\times w}$. Notice that we should always suppose that we do not know whether the training data follows Equation 1. We denote this sample set $\mathcal{U}$.
We shall use all the date in $\mathcal{U}$ for the training of the Neural Network. A sample of the training date should be in the form of $(\mathbf{x}^{l\times w}; *)$; where $*$ can be either $+1$, or -1, representing the label for $\mathbf{x}^{l\times w}$ is "+", or "-" respectively.

As we have a large and perfect data set described above, we can prove the following Lemma:

Lemma 1:
1. The neural network can be trained perfectly so that the accuracy on the training data set $\mathcal{U}$ is 100%.
2. After the neural network has been trained perfectly, all the weights of the connections $w_{i,j} > 0$.
3. Suppose, after the neural network has been trained perfectly, the sum of $\lfloor l \times w/2 \rfloor$ largest weights of the connections is S_{w_1}, and the sum of all $\lceil l \times w/2 \rceil$ smallest weights of the connections is S_{w_2}, then $S_{w_2} - S_{w_1} \ge 0$.

Proof:
1. This conclusion can be obtained by setting all $w_{i,j}$ to 1.
2. Suppose there is an (i_{-1}, j_{-1}) such that $w_{i_{-1},j_{-1}} \le 0$. We can construct an $\mathbf{x}^{l\times w}$ so that $\lceil l \times w/2 \rceil$ elements, including $x_{i_{-1},j_{-1}}$, take the value "+1", while the remaining elements take the values "-1". We should have $f_u(\mathbf{x}^{l\times w}) = +1$ according to equation (1), because $\sum_{1\le i\le l, 1\le j\le w} x_{i,j} \ge 0$. We can then obtain a new instance $\mathbf{x'}^{l\times w}$ from $\mathbf{x}^{l\times w}$ by letting $x'_{i_{-1},j_{-1}} = -1$, and $x'_{i,j} = x_{i,j}$ for any $(i,j) \ne (i_{-1}, j_{-1})$. $\mathbf{x'}^{l\times w}$ should be in $\mathcal{U}$ and the correct label should be "-". However, $f_u(\mathbf{x'}^{l\times w}) = +1$, as $\sum_{1\le i\le l, 1\le j\le w} w_{i,j}x'_{i,j} = \sum_{1\le i\le l, 1\le j\le w} w_{i,j}x_{i,j} - 2 \times w_{i_{-1},j_{-1}} > 0$. This causes a contradiction.
3. By letting all the pixel with connections in S_{w_2} take the values "+1" and the others take the values "-1", we should have $S_{w_2} - S_{w_1} \ge 0$ as we expect the output should be "+1". ■

[1]The intensity value of each pixel is either $+1$ or -1 for bitmap images, but there is no problem if we assume the values are either $+1$ or -1.

Lemma 1 simply implies that the weights of connections are very close, if not identical, especially when $l \times w$ is a large number. Thus, without losing generality, we assume $w_{i,j} = 1$ for all (i, j) after the neural network is trained.[2]

C. Basic Model

1. The noise is defined as a change of environment that enforces a change of pixel values; when subjected to noise, the values of some of the pixel will undergo a change from $+1$ to -1, some from -1 to $+1$, and others may remain unchanged. The noise that influence pixel to change from $+1$ to -1 (or -1 to $+1$) is called anti-$black$-noise (or anti-$white$-noise) respectively. A pixel which undergoes a change from $+1$ to -1 (-1 to $+1$) is called an anti-$black$-noise-contaminated pixel (anti-$white$-noise-contaminated cell) respectively.

2. Two types of noise are considered; concentrated noise which influences the values within a concentrated block(s) of pixels, and random noise that is distributed uniformly over the whole image randomly.

3. A set of anti-"black"-random noise (or anti-"white"-random noise) is dispersed uniformly over the image, producing a uniform chance of converting the values of pixel from $+1$ to -1 (or -1 to $+1$). The result of random noise thus could be regarded as a change in the probability for a pixel taking the value $+1$ from α to a new value (implying a change in the probability for a pixel taking the value -1 from β to a new value).

4. A set of anti-"black"-concentrated noise (or anti-"white"-concentrated noise) is defined as the union of non-overlapped rectangle blocks of size $n_l \times n_w$, on each of which all the pixels taking the value $+1$ (or -1) will be changed to taking -1 (or $+1$). The corresponding union of these rectangle areas is called a noise concentrated area, and $n_l \times n_w$ is called the size of noise blocks.

5. In accordance with the above two types of noise, the anti-"black"-noise-contaminated pixel (anti-"white"-noise-contaminated cells) comprise the two types depending on the noise type, namely the anti-"black"-random-noise-contaminated pixel (anti-"white"-random-noise-contaminated cells) and anti-"black"-concentrated-noise-contaminated pixel (anti-"white"-concentrated-noise-contaminated cells). [3]

6. We call a pixel in the low representation concentrated-noise-polluted *if and only if* the conjunction set of the corresponding window in its original high resolution representation and the noise-concentrated area is not empty.

7. Because we are interested in computing the lower bounds of the classification stability throughout this paper, we consider only the anti-"black" noise in the analysis. Thus when we refer to noise, concentrated noise, random noise, or contaminated pixels hereinafter, anti-"black" noise, anti-"black"-concentrated noise, anti-"black"-random noise, anti-"black" noise contaminated pixels are implied.

8. $\aleph_c$, and $\aleph_w$ denote the number of concentrated-noise contaminated pixels, and the number of random-noise contaminated pixels separately.

D. Assumption

We always assume that the high resolution image is large and the amount of noise is large so that both the total number of noise contaminated pixels and the size of concentrated noise contaminated area are large.

Basic Assumption *In the absence of concentrated noise, the percentage of "black" (or "white")-pixels in the low resolution representation image among all the pixels, or among a set of many arbitrarily chosen pixels, is equivalent to the chance that there are more pixel taking value $+1$ than those -1 (or more pixel taking value -1 than those $+1$) in a window in original high resolution image where the pixels merged into one pixel of the low resolution representation.*

This assumption implies:

Lemma 2: In the absence of both random and concentrated noise, among a set of a large number of arbitrarily chosen pixel in the low representation image, the proportion of "black" and "white" pixels P_{+1} and P_{-1} can be computed by:

$$P_{+1} = \sum_{y=0}^{\lfloor \frac{r_l r_w}{2} \rfloor} \binom{r_l r_w}{y} \beta^y \alpha^{r_l r_w - y},$$

$$P_{-1} = \sum_{y=0}^{\lfloor \frac{r_l r_w}{2} \rfloor} \binom{r_l r_w}{y} \alpha^y \beta^{r_l r_w - y}.$$

Lemma 3: In the presence of only random noise, in a set of a large number of arbitrarily chosen pixels in the low resolution image, P'_{+1} and P'_{-1} which denote the proportion of $+1$-pixels and -1-pixels can be computed by:

$$P'_{+1} = \sum_{y=0}^{\lfloor \frac{r_l r_w}{2} \rfloor} \binom{r_l r_w}{y} \beta'^y \alpha'^{r_l r_w - y},$$

$$P'_{-1} = \sum_{y=0}^{\lfloor \frac{r_l r_w}{2} \rfloor} \binom{r_l r_w}{y} \alpha'^y \beta'^{r_l r_w - y}.$$

where, α' and β' are the percentage of the pixels in the original high resolution image taking value $+1$ and taking value -1, in the presence of random noise.

The fact that $P_{+1} > P_{-1}$ when $\alpha > \beta$, simply says in noise free environment, both of high resolution and low resolution images will be classified into the same cluster.

We should regard the set of concentrated noise polluted pixel in low resolution image as a set of "arbitrarily" chosen pixel as in the above basic assumption, because concentrated noise also have to be "random", although they occur in the form of concentrated blocks. This is to say, the above lemmas should be valid for the set of concentrated noise polluted pixel in the low resolution image.

[2]Carefully checking the results in our theorems in the later sections, we can know that it actually does *not matter* if $w_{i,j} = 1$ is not true for all i and j, as long as the conclusion of item 3 in Lemma 1 is valid.

[3]Notice that, when both of random noise and concentrated-noise coexist, some noise-contaminated votes may belong to both of these two types, as it will be explained in the proof of theorem 1. In a noise-free environment, the rate of $+1$-pixels in a large subset under random noise will not change from that of the entire image. This implies that a percentage of random noise contaminated pixel in the concentrated area should not change from that of the whole image area. This is useful in computing an overlap of random noise contaminated pixel and concentrated noise contaminated pixel. It is understandable as the noise-concentrated area can be considered as reasonably large.

III. Main Theorems

A. The Stabilities for High Resolution and Low Resolution Representations

Theorem 1: The classification result of the high resolution image will be retained iff

$$\aleph_c + \aleph_w \times \frac{N - \aleph_c/\alpha}{N} < \frac{\alpha - \beta}{2} \times N.$$

This theorem can be proved by noticing the following two facts:
1. Among $\aleph_w$ anti-"black"-random-noise-contaminated votes, $\frac{\aleph_c/\alpha}{N} \times \aleph_w$ votes come from the anti-"black"-Noise-Concentrated area.
2. The classification of high resolution image is able to preserve the original decision, if and only if the number of overall anti-"black"-concentrated-noise-contaminated pixels is less then $\frac{\alpha-\beta}{2} \times N$.

B. The Stability of Low Resolution Image

Theorem 2: The original classification result will be retained if:

$$\aleph_c < \frac{\frac{n_l n_w}{r_l r_w}}{(\left\lceil \frac{n_l - 1}{r_l} \right\rceil + 1)(\left\lceil \frac{n_w - 1}{r_w} \right\rceil + 1)} \cdot \frac{P_{+1}(\aleph_w) - P_{-1}(\aleph_w)}{1 + P_{+1}(\aleph_w) - P_{-1}(\aleph_w)} \cdot \alpha \cdot N,$$

and

$$\aleph_w < (\alpha - \beta)/2 \times N,$$

where $P_{+1}(\aleph_w)$ and $P_{-1}(\aleph_w)$ which denote the percentages of $+1$-pixels and -1-pixels under random noise, and can be calculated as

$$P_{+1}(\aleph_w) = \sum_{y=0}^{\lfloor \frac{r_l r_w}{2} \rfloor} \binom{r_l r_w}{y} (\beta + \aleph_w/N)^y (\alpha - \aleph_w/N)^{r_l r_w - y},$$

$$P_{-1}(\aleph_w) = \sum_{y=0}^{\lfloor \frac{r_l r_w}{2} \rfloor} \binom{r_l r_w}{y} (\alpha - \aleph_w/N)^y (\beta + \aleph_w/N)^{r_l r_w - y}.$$

Proof: If $\aleph_w < (\alpha - \beta)/2 \times N$, it is easy to see that $P_{+1}(\aleph_w) > P_{-1}(\aleph_w)$, and thus the classification result will be preserved. Suppose that the rate of concentrated noise polluted pixels is X $(0 \le X \le 1)$. $P_{+1}(\aleph_w)$ and $P_{-1}(\aleph_w)$ portions of these concentrated noise polluted pixels are originally $+1$- and -1-pixels respectively in the presence of random noise only. So when $X < \frac{P_{+1}(\aleph_w) - P_{-1}(\aleph_w)}{1 + P_{+1}(\aleph_w) - P_{-1}(\aleph_w)}$, the number of $+1$-pixels should be still larger than the number of -1-pixels, even if we regard all concentrated noise polluted pixel as -1-pixels. Thus, the classification result of the low resolution image should be able to preserve the original classification result. As a noise block of size $n_l \times n_w$ can be partitioned at most into $(\left\lceil \frac{n_l - 1}{r_l} \right\rceil + 1)(\left\lceil \frac{n_w - 1}{r_w} \right\rceil + 1)$ different windows in high resolution representation, we have: $S_r/S_c \le (\lceil (n_l - 1)/r_l \rceil + 1)(\lceil (n_w - 1)/r_w \rceil + 1) \cdot r_l r_w/(n_l n_w)$, where $S_r = X \cdot N$ is the total size of concentrated noise polluted windows, S_c denotes the total size of the concentrated noise area. To meet the requirement on X of $X < \frac{P_{+1}(\aleph_w) - P_{-1}(\aleph_w)}{1 + P_{+1}(\aleph_w) - P_{-1}(\aleph_w)}$, we only need the following inequality:

$$|S_c| < \frac{\frac{n_l n_w}{r_l r_w}}{(\left\lceil \frac{n_l - 1}{r_l} \right\rceil + 1)(\left\lceil \frac{n_w - 1}{r_w} \right\rceil + 1)} \cdot \frac{P_{+1}(\aleph_w) - P_{-1}(\aleph_w)}{1 + P_{+1}(\aleph_w) - P_{-1}(\aleph_w)} \cdot N.$$

Substituting the relation of $\aleph_c = |S_c| \times \alpha$, the conclusion of Theorem 2 follows.

C. Averaging the Stability Boundary

In the theorem above, the ceiling operations are used to develop a sufficient condition of stability that constitutes a worst possible condition whereby each of the noise blocks pollutes a maximum number of pixels in low resolution representation. It is most doubtable that the worst situation for each noise block happens at the same time. Some appropriate averaging will be introduced here by shifting the partitions with noise distribution fixed.

By taking an average, we can have the following Theorem showing the improved result:

Theorem 3: In average, the classification result of low resolution image will retain the original classification result if:

$$\aleph_c < \frac{n_l n_w}{(r_l + n_l - 1)(r_w + n_w - 1)} \cdot \frac{P_{+1}(\aleph_w) - P_{-1}(\aleph_w)}{1 + P_{+1}(\aleph_w) - P_{-1}(\aleph_w)} \cdot \alpha \cdot N$$

and

$$\aleph_w < \frac{1}{2}(\alpha - \beta)N$$

where $P_{+1}(\aleph_w)$ and $P_{-1}(\aleph_w)$ are calculated as they are in Theorem 2.

IV. Conclusions, Experiment

A. Conclusions

Fig. 2 illustrates the number of noise-contaminated pixels in the original image that the classification on the low resolution representation and the high resolution representation can accommodate before the original classification result, "black", is reversed. We are using averaged estimation in the figure. Nearly equilibrium cases of $\alpha - \beta = 0.05$ are treated in the figure.

Based on this picture, we could see that, as the resolution decreases, the number of contaminated pixel it can accommodate increases continuously up to a certain point beyond which this starts to decrease. We claim that the classification of low resolution image is always more stable than the high resolution image. Thus, we have the conclusions here:
1. The low resolution is always more stable than the high resolution representation.
2. The lower the resolution representation is, the number of noise we can accommodate without changing the original classification results keeps increasing up to a certain limit but across this point, the number of the noise we can accommodate starts to decrease.

Conjecture We believe that the above conclusion related to high resolution representation and low resolution representation still remain valid even when the decision making process involves complicated matching schemes such as features extraction by PCA method.

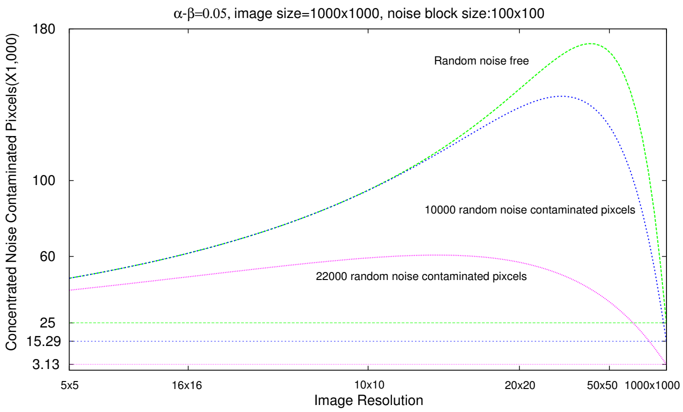

Fig. 2. Numbers of noise-contaminated pixels that the classification system can accommodate on different resolution representations

B. Experiments

The experiment was done on human face classification, and the algorithms were implemented with Matlab 6. We were to determine if any new photo belongs to one of a set of known people.

1) Data Set: The training data consist of two sets of photos. Set 1, the collection of photos of known people, comprises the pictures of 50 people, each has 4 different gray-scale pictures showing different facial expressions: blank, smiling, angry and surprised. Set 2, used as samples of unknown people, comprises 400 facial pictures, which we downloaded from Internet. We cropped each picture in Set 1 and Set 2 into a face of size 45×45 pixels where the line between two eyes has fixed length and is parallel to the horizontal line, as shown in figure 3, and stored them in raw format. We used all these photos as the training data of high resolution images.

The test date were obtained as follows: We used another set of 50 pictures of the known people in Set 1, and downloaded 50 other facial pictures of unknown people from Internet; then added noise into these 100 facial pictures by using Photoshop 6.0; and finally cropped them into faces of size 45×45 satisfying the above requirement on the line between two eyes. Several typical training faces and a test face is shown in figure 3.

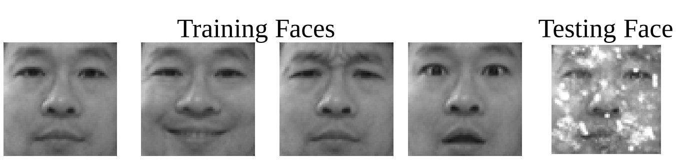

Fig. 3. Examples of Training and Testing Faces

2) Detailed Methods and Results: The pixel values of each image are normalized so that the mean is null before the images are fed into the input layers of neural networks. We implemented a neural network with three layers, 45×45 inputs, 90 hidden units and 2 outputs. The two outputs represent respectively the two different categories, known people and unknown people; each is used for predicating whether the photo to be classified belongs to that category. The output of the network is interpreted by believing the output neuron with the highest score. That is, when the output neuron representing "known people" is greater (smaller, respectively) than the neuron representing "unknown people", the input image is classified as "known" ("unknown", respectively).

The low resolution images and the neural networks for low resolution images were constructed as follows: we use photoshop 6.0 to reduce each image into $r\times r$ resolution image (r=40, 35, 30, 25, 20, 15, 10, separately). We implemented a 3-layer neural network for each low resolution level.

We used the hyperbolic tangent sigmoid transfer function (tansig) as the activation function for each neuron of the hidden layers of the neural networks. The log-sigmoid transfer function (logsig) was used as the activation function for all the output neurons of the neural networks.[4]

The neural networks were trained by the back-propagation algorithm with adaptive learning rate (traingda). During a training process, an "early stopping" technique was used to improve the generalization and avoid over-fitting. If the error did not decrease in 10 consecutive epochs, the training of a neural network was terminated to avoid over-fitting.

The trained neural networks were used on the test set. We have reduces the sizes of original images using PhotoShop6.0 into different sizes so as to compare the performances of the neural networks for different resolution images. The numbers of correctly classified pictures among the 100 test pictures for global matching and each regional matching neural networks are shown in Table I.

TABLE I
CLASSIFICATION RESULTS ON HUMAN FACES WITH DIFFERENT RESOLUTIONS

45×45	40×40	35×35	30×30	25×25	20×20	15×15	10×10
75	77	76	79	84	88	83	73

Note: The first raw represents pictures in different resolutions, the second the numbers of correctly classified pictures.

V. Discussion and Open Problems

The paper shows by a mathematical model that the classification by neural networks for relatively low resolution representations has better performance than that for high resolution representations. The model and discussion are based on the based approach. While it may be argued that, with a proper preprocessing stage, high resolution representations can be easily be converted to low resolution representations; our model does not consider such a preprocessing and requires the entire image be fed into the neural network directly. When such a preprocessing stage is to be considered, our model shows, for classification purpose, properly reducing the resolutions of images could save storage space but without reducing the classification performance.

The model we set up is based on bitmap images, although our experiment shows that the result remains valid

[4]The log-sigmoid transfer function was picked for output neurons because its output range (0 to 1) is perfect for learning to output boolean values. We used the hyperbolic tangent sigmoid function for hidden neurons as it reached best performance in our experiment.

for grayscale images, a proper model for grayscale and color images seems to be interesting to be developed.

REFERENCES

[1] P. Sinha, "Identifying perceptually significant features for recognizing faces," in *Proceedings of SPIE,Human Vision and Electronic Imaging VII*, vol. 4662. SPIE, 2002, pp. 12–21.
[2] A. M. Burton, P. Miller, V. Bruce, P. Hancock, and Z. Henderson, "Human and automatic face recognition: a comparison across image formats," *Vision Research*, vol. 41, pp. 3185–3195, 2001.
[3] V. Bruce, Z. Henderson, K. Greenwood, P. Hancock, A. Burton, and P. Miller, "Verification of face identities from images captured on video," *Journal of Experimental Psychology: Applied*, vol. 5, pp. 339–360, 1999.
[4] A. M. Burton, S. Wilson, M. Cowan, and V. Bruce, "Face recognition in poor-quality video: Evidence from security surveilance," *Psychological Science*, vol. 10, no. 3, pp. 243–248, May 1999.
[5] M. Pavlou, "Face kernel extraction from local features," Ph.D. dissertation, School of Electrical Engineering and Electronics, University of Manchester, 2005.
[6] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, "Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft knn ensemble," *IEEE Transactions on Neural Networks*, 2005, in press.
[7] M. A. Anjum, M. Y. Javed, and A. Basit, "A new approach to face recognition using dual dimension reduction," *International Journal of Signal Processing*, vol. 2, no. 1, pp. 1–6, 2005.