PR 1598
pp: 1--8 (col.fig.: Nil)
PROD. TYPE: TYP
ED: IS
PAGN: Vidya – SCAN: Mangala

ARTICLE IN PRESS

PERGAMON

Pattern Recognition ▮▮▮ (▮▮▮) ▮▮▮–▮▮▮

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Subspace morphing theory for appearance based object identification

Zhongfei (Mark) Zhang[a,*], Rohini K. Srihari[b]

[a] Computer Science Department, Watson School, State University of New York, Binghamton, NY 13902, USA
[b] Center of Excellence for Document Analysis and Recognition (CEDAR), State University of New York, Buffalo, NY 14228, USA

Received 2 January 2001; accepted 30 October 2001

Abstract

Object identification techniques have wide applications ranging from industry, business, military, law enforcement, to people's daily life. This research is motivated to develop a new theory for appearance based object identification with its applications in different areas. Although many successful techniques have been proposed in certain specific applications, object identification, in general, still remains as a difficult and challenging problem. In appearance based approaches, almost all the proposed methods are based on a fundamental assumption, i.e., all the images (both in the model base and to be queried) are in the same dimensions, so that the feature vectors are all in the same feature space; if images are provided with different dimensions, a normalization in scale to a pre-determined image space must be conducted. In this research, a theory for appearance based object identification called subspace morphing is developed, which allows scale-invariant identification of images of objects, and therefore, does not require normalization. Theoretical analysis and experimental evaluation show that in the situation where images are provided in different dimensions, which is common in many applications, subspace morphing theory is superior to the existing, normalization-based techniques in performance. © 2002 Published by Elsevier Science Ltd on behalf of Pattern Recognition Society.

Keywords: Appearance based object identification; Subspace morphing; SV vector; Projections; Morphed vectors; Essential SV vectors; Collection matrix; Essential collection matrix; Identification capability; Identification precision

UNCORRECTED PROOF

1. Introduction

Object identification techniques have wide applications ranging from industry, business, military, law enforcement, to people's daily life. This research is motivated to develop a new theory for appearance based object identification with its applications in different areas.

Although many successful techniques have been proposed in certain specific applications, object identification, in general, still remains as a difficult and challenging problem. In appearance based approaches, typical methods include eigenfaces and the related subspace analysis (e.g. [1–4]), singular value decomposition (e.g. [5]), Gabor wavelet features (e.g. [6]), neural networks based learning (e.g. [7,8]), evolutionary pursuit based learning (noticeably [9,10]), on active appearance models (noticeably [11,12]), and shape and appearance-based probabilistic models (e.g. [13–15]) as well as the combinations of several existing techniques in object identification (e.g., matrix distances proposed by Cox et al. [16]). All these methods are based on a fundamental assumption, i.e., all the images (both in the model base and to be queried) are in the same dimensions, so that the feature vectors are all in the same feature space; if images are provided with different dimensions, a normalization in scale to a pre-determined space must be conducted.

In this paper, a theory for appearance based object identification called *subspace morphing* is developed, which allows scale-invariant identification of images of objects, and therefore, does not require normalization in scale.

* Corresponding author. Tel.: +1-607-777-2935; fax: +1-607-777-4729.

E-mail addresses: zhongfei@cs.binghamton.edu (Z. Zhang), rohini@cedar.buffalo.edu (R.K. Srihari).

0031-3203/02/$22.00 © 2002 Published by Elsevier Science Ltd on behalf of Pattern Recognition Society.
PII: S0031-3203(01)00209-6

Theoretical analysis and experimental evaluation show that in the situation where images are provided in different dimensions, which is typical and common in many applications, subspace morphing theory is superior to the existing, normalization-based techniques in performance.

This paper is organized as follows. After this section, the subspace morphing theory is introduced in Section 2. The following section is dedicated to comparing this theory with the existing, normalization based techniques in terms of computation complexity, identification capability, and identification precision. Then the experimental evaluations of this theory and the normalization based techniques are conducted based on real image data. Finally, a conclusion is presented.

2. Subspace morphing theory

In the following text, R is used to denote a real space; thus, R^n is an n dimensional vector space, and $R^{m\times n}$ is an $m\times n$ dimensional matrix space (or image space). Also a boldface symbol is used to denote a vector or matrix.

The subspace morphing theory is based on a set of properties of the theory of Singular Value Decomposition [17].

Theorem 1. *Let $A\in R^{m\times n}$. $\exists$ column-orthonormal matrix $\boldsymbol{U}\in R^{m\times n}$, row-orthonormal matrix $\boldsymbol{V}\in R^{n\times n}$, and diagonal matrix $\boldsymbol{\Lambda}\in R^{n\times n}$, such that*

$$\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Lambda}V^{\mathrm{T}}, \tag{1}$$

where the diagonal values of $\boldsymbol{\Lambda}$ are in sorted order from the largest to the smallest.

Proof. See Ref. [17].

Definition 1. Let $A\in R^{m\times n}$. Eq. (1) defines a *singular value decomposition* (SVD) of the matrix $\boldsymbol{A}$, where the diagonal values of Λ are called *singular values* (*SVs*). The vector formed by the sorted diagonal values are called *SV vectors*.

Lemma 1. *Let $\boldsymbol{A}\in R^{m\times n}$. Let $\boldsymbol{P}\in R^{m\times m}$, and $\boldsymbol{Q}\in R^{n\times n}$ be orthonormal matrices. Then matrices $\boldsymbol{P}A$, AQ and matrix A share the same set of singular values.*

Proof. W.L.O.G. assume $m>n$. Then, by SVD, Eq. (1) holds for $\boldsymbol{A}$.

This is followed by

$$(\boldsymbol{P}A)^{\mathrm{T}}(\boldsymbol{P}A)=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{V}\Lambda^2\boldsymbol{V}^{\mathrm{T}}, \tag{2}$$

$$(AQ)^{\mathrm{T}}(AQ)=\boldsymbol{Q}^{\mathrm{T}}(A^{\mathrm{T}}A)\boldsymbol{Q}=\boldsymbol{W}\Lambda^2\boldsymbol{W}^{\mathrm{T}}, \tag{3}$$

where $\boldsymbol{W}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{V}$ is an orthonormal matrix, as the product of two orthonormal matrices is another orthonormal matrix [18]. This has proved the lemma. □

Lemma 2. *Let $A\in R^{m\times n}$. Let $\boldsymbol{B}\in R^{m\times n}$, and $\boldsymbol{B}$ is obtained by the transform of interchanging any two rows (or columns) of $\boldsymbol{A}$. Then $\boldsymbol{A}$ and $\boldsymbol{B}$ share the same set of singular values.*

Proof. Define the following matrix:

$$\boldsymbol{I}_{ij}=\boldsymbol{I}-(\boldsymbol{E}_i-\boldsymbol{E}_j)(\boldsymbol{E}_i-\boldsymbol{E}_j)^{\mathrm{T}}, \tag{4}$$

where $\boldsymbol{I}$ is the unit matrix, and $\boldsymbol{E}_i$ and $\boldsymbol{E}_j$ are the ith and the jth column vectors of the unit matrix, respectively. Hence, the transform of interchanging two rows of $\boldsymbol{A}$ is equivalent to the product of $\boldsymbol{I}_{ij}\boldsymbol{A}$, and the transform of interchanging two columns of $\boldsymbol{A}$ is equivalent to the product of $\boldsymbol{A}\boldsymbol{I}_{ij}$. It is clear that the matrix $\boldsymbol{I}_{ij}$ is an orthonormal matrix. Consequently, the proof follows by Lemma 1.

Theorem 2. *Let $\boldsymbol{A},\boldsymbol{B}\in R^{m\times n}$. By SVD, $\exists$ SV vectors $\boldsymbol{\Lambda}$ and $\boldsymbol{\Sigma}$ for $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively. Thus, the Euclidean norm of the difference between the vectors $\boldsymbol{\Lambda}$ and $\boldsymbol{\Sigma}$ is bounded by the Frobenius norm of the difference between the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, i.e.*

$$\sqrt{\sum_{i=1}^{n}(\lambda_i-\sigma_i)^2}\leqslant\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(a_{ij}-b_{ij})^2}, \tag{5}$$

where λ_i,σ_i are the ith elements of Λ and Σ, respectively, and a_{ij} and b_{ij} are the elements of $\boldsymbol{A}$ and $\boldsymbol{B}$ at the position of (i,j), respectively.

Proof. See [17].

Lemma 3. *Let $A\in R^{m\times n}$. Let $\boldsymbol{B}\in R^{p\times q}$, such that $\boldsymbol{B}$ is obtained by scaling $\boldsymbol{A}$ in row by k_r, and then in column by k_c, where k_r,k_c are integers, and $k_r,k_c\geqslant 1$, i.e., $p=k_rm$, and $q=k_cn$. Thus, $\boldsymbol{B}$ has the same rank as $\boldsymbol{A}$ does, and the singular values of $\boldsymbol{B}$ are $\sqrt{k_rk_c}$ times the corresponding singular values of $\boldsymbol{A}$.*

Proof. The fact that $\boldsymbol{B}$ has the same rank as $\boldsymbol{A}$ does is obvious, since duplication of any row or column in a matrix does not affect the rank of the matrix [18].

By SVD following the notations of Eq. (1):

$$\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Sigma}V^{\mathrm{T}}. \tag{6}$$

In the following, the symbol $\equiv$ is used to denote "equal by definition", and the symbol $\rightarrow$ is used to denote a transform.

To show the relationship of their corresponding singular values between $\boldsymbol{A}$ and $\boldsymbol{B}$, let

$$\boldsymbol{A}\equiv\begin{pmatrix}\boldsymbol{a}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{a}_m^{\mathrm{T}}\end{pmatrix},$$

where $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_m$ are the row vectors of $\boldsymbol{A}$.

UNCORRECTED PROOF

The following series of transforms are conducted:

$$A = \begin{pmatrix} a_1^{\mathrm{T}} \\ \vdots \\ a_m^{\mathrm{T}} \end{pmatrix} \rightarrow \begin{pmatrix} \left.\begin{matrix} a_1^{\mathrm{T}} \\ \vdots \\ a_1^{\mathrm{T}} \end{matrix}\right\} k_r \text{ times} \\ \vdots \\ \left.\begin{matrix} a_m^{\mathrm{T}} \\ \vdots \\ a_m^{\mathrm{T}} \end{matrix}\right\} k_r \text{ times} \end{pmatrix} \equiv A_r \rightarrow \begin{pmatrix} \left.\begin{matrix} A \\ \vdots \\ A \end{matrix}\right\} k_r \text{ times} \end{pmatrix} \equiv A_b.$$

The first transform duplicates each row of A by k_r times to obtain A_r, and the second transform "clusters" the row vectors of $k_r A$ by interchanging the positions of the row vectors to obtain A_b. Hence,

$$A_b^{\mathrm{T}} A_b = (A^{\mathrm{T}} \ldots A^{\mathrm{T}}) \begin{pmatrix} A \\ \vdots \\ A \end{pmatrix} = \sum_{i=1}^{k_r} A^{\mathrm{T}} A = k_r A^{\mathrm{T}} A = k_r V \Sigma^2 V^{\mathrm{T}} = V(\sqrt{k_r}\Sigma)^2 V^{\mathrm{T}}.$$

This means that the singular values of A_b are $\sqrt{k_r}$ times those of A. By Lemma 2, A_b and A_r have the same singular values. Therefore, the singular values of A_r are $\sqrt{k_r}$ times those of A

Similarly, the transform on A_r by duplicating each columns of $A_r k_c$ times to obtain a matrix A_c, which is B, and then by "clustering" the columns of $k_c A_r$ through interchanging the positions of the columns to obtain a matrix A_d. Then,

$$A_d A_d^{\mathrm{T}} = (A_r \ldots A_r) \begin{pmatrix} A_r^{\mathrm{T}} \\ \vdots \\ A_r^{\mathrm{T}} \end{pmatrix} = k_c A_r A_r^{\mathrm{T}}.$$

This means that the singular values of A_d are $\sqrt{k_c}$ times those of the corresponding singular values of A_r, which have been shown to be $\sqrt{k_r}$ times the corresponding singular values of A. Again, by Lemma 2, A_d and A_c have the same singular values. Therefore, the singular values of A_c are $\sqrt{k_r k_c}$ times those of A. This concludes the proof of this Lemma. □

Based on the above lemmas and Theorem 2, the foundation of the subspace morphing theory and its algorithm is based on the following theorem:

Theorem 3. *Let $A_1 \in R^{m_1 \times n_1}$, $A_2 \in R^{m_2 \times n_2}$. Λ and Σ are SV vectors of A_1 and A_2, respectively. For any positive integer n, the following holds*:

$$||\sqrt{m_2 n_2}\Lambda - \sqrt{m_1 n_1}\Sigma||_n \leqslant ||B_1 - B_2||_F, \tag{7}$$

where B_1 and B_2 are both in $R^{m_1 m_2 \times n_1 n_2}$, and are obtained by zero-hold interpolation [19] *from A_1* and *A_2, respectively. $||\cdot||_F$ stands for matrix Frobenius norm, and $||\cdot||_n$ the norm of an n dimensional vector.*

Proof. First scale up matrices A_1 and A_2, respectively, to B_1 and B_2, such that $B_1, B_2 \in R^{m_1 m_2 \times n_1 n_2}$. By Lemma 3, the singular value vectors of B_1 and B_2 are:

$$\sqrt{m_2 m_2}\lambda_i, \quad i = 1, \ldots, n_1$$

and

$$\sqrt{m_1 n_1}\sigma_i, \quad i = 1, \ldots, n_2.$$

By Theorem 2, Eq. (7) holds, and the proof follows.

This foundation theorem may be explained conceptually as follows. Given two images with different dimensions, they can always be mapped onto a common image space. This theorem establishes a relationship between the SV vectors extracted from these two images and the SV vectors that would be extracted after mapping them onto that common space. Thus, using this theorem, the SV vectors from the hypothetical common space can be immediately "extracted" based on the SV vectors obtained from the current images without actually mapping them onto that common space.

Based on this foundation theorem, the subspace morphing theory is introduced below.

Definition 2. Let v be a vector in R^n. Let R^m be an arbitrary space. Define vector v' as a *projection* of v from R^n to R^m as follows. If $m > n$, the dimensionality of v' is extended to m by filling values 0 to the extended dimensions of v; if $m < n$, the dimensionality of v' is truncated to m; if $m = n$, $v' = v$.

Definition 3. Given two arbitrary spaces $R^{m \times n}$ and $R^{p \times q}$, W.L.O.G. assume $m \geqslant n$, $p \geqslant q$. Let $A \in R^{m \times n}$, v be its *SV* vector, and u' be the projection of v from R^n to R^q. Vector v' is defined as the *morphed SV vector* of v from R^n to R^q

UNCORRECTED PROOF

as follows:

$$\mathbf{v}' = \frac{\sqrt{pq}}{\sqrt{mn}}, \quad \mathbf{u}' = \sqrt{pq}\mathbf{u}, \tag{8}$$

where $\mathbf{u} = (1/\sqrt{mn})\mathbf{u}'$ is called the *essential SV vector* of $\mathbf{v}$.

This means that $\mathbf{v}'$ is the "equivalent" SV vector of a matrix in $R^{p\times q}$ mapped from $\mathbf{A}$ in $R^{m\times n}$. Note that the values of an essential SV vector always stay the same regardless of whatever space the corresponding vector is morphed to.

Definition 4. Let $\mathbf{A} \in R^{m_1\times n_1}$, $\mathbf{B} \in R^{m_2\times n_2}$, W.L.O.G. assume $m_1 \geqslant n_1$, $m_2 \geqslant n_2$. Let $\mathbf{v}_A$ and $\mathbf{v}_B$ be their SV vectors, respectively. Given any $R^{p\times q}$, W.L.O.G. assume $p \geqslant q$. The *addition* of $\mathbf{v}_A$ and $\mathbf{v}_B$ in R^q is defined as

$$(\mathbf{v}_A \oplus \mathbf{v}_B)_q = (\mathbf{v}'_A + \mathbf{v}'_B) = \sqrt{pq}(\mathbf{u}_A + \mathbf{u}_B), \tag{9}$$

where $\mathbf{u}_A$ and $\mathbf{u}_B$, are the essential SV vectors of $\mathbf{v}_A$ and $\mathbf{v}_B$, and $\mathbf{v}'_A$ and $\mathbf{v}'_B$ are the morphed SV vectors of $\mathbf{v}_A$ and $\mathbf{v}_B$, respectively.

Similarly, subtraction, dot product, and cross product of two SV vectors with different dimensionalities may be defined. Note that all these algebraic operations between two SV vectors are defined as to first morphing them to a "common" space before the actual operation may be conducted; this "common" space may be any arbitrary space; this space is to be determined dynamically during the identification process.

Definition 5. Given a set of SV vectors: $\mathbf{v}_i \in R^{n_i}$, respectively, $i = 1, \ldots, k$, let $\mathbf{v}'_i$ be the morphed vector of $\mathbf{v}_i$ from R^{n_i} to a "common" space R^q, respectively. The matrix $\mathbf{V}$ formed by all $\mathbf{v}'_i$ is called the *collection matrix* of $\mathbf{v}_i$ in R^q, i.e.

$$\mathbf{V} = (\mathbf{v}'_1, \ldots, \mathbf{v}'_k) = \sqrt{pq}\,\mathbf{U}, \tag{10}$$

where $\mathbf{U}$ is called the *essential collection matrix* of all the vectors $\mathbf{v}_i$ $(i = 1, \ldots, k)$. Let $\mathbf{u}_i$ be the essential SV vector of $\mathbf{v}_i$. Then, $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_k)$.

Note that the essential collection matrix is independent of the "common" space that the given SV vectors are morphed to, which is to be determined dynamically during the identification process.

Based on the above theory, the algorithm (**SM**) for object identification is given below. Assume that there are N object classes in a model base, and each class has M_i sample images collected $(i = 1, \ldots, N)$. Note that sample images in each class may be in different dimensions. Consequently, the SV vectors of these samples are in different dimensionalities, and so is the SV vector of the query image. The idea of **SM** is to use nearest linear combination *in a morphed subspace* dynamically, as opposed to in a fixed, predetermined space, to find the best match for a query image. Specifically, assume that each sample image in a model base is represented by an essential SV vector $\mathbf{u}_{ij}$ (here i indexes the class number, and j indexes the sample number in the object class). Then a linear combination $\mathbf{U}_i\boldsymbol{\alpha}_i$ of all the essential SV vectors in class i is used to represent the "cluster region" of this class in a morphed feature space, where $\mathbf{U}_i$ is the essential collection matrix of the SV vectors in class i, and $\boldsymbol{\alpha}_i$ is a coefficient vector in R^{M_i}. Assume that $\mathbf{w}$ is the essential SV vector of the query image. Then, the following constraint holds for a specific $\boldsymbol{\alpha}_i$, if the query image is in class i:

$$\mathbf{w} = \mathbf{U}_i\boldsymbol{\alpha}_i. \tag{11}$$

Eq. (11) shows that a linear solution exists for the object identification problem *independent* of the hypothetical "common" image space $R^{p\times q}$ and the morphed subspace R^q. This property enables vector space comparison in different dimensionalities independent of the hypothetical "common" space. Eq. (11) is typically overdetermined, and a solution to $\boldsymbol{\alpha}_i$ may be obtained through a standard least square minimization.

3. Subspace morphing vs. normalization

This section is dedicated to comparing the performance of the subspace morphing theory with those of the normalization based techniques, and then to show that the former is superior to the latter in performance. In particular, the performance analysis is conducted in the following three aspects:

Computation complexity: It is clear that the process of normalization in scale *per se* requires extra computation. *In addition to* this extra computation, in the case of scaling up, redundancy is introduced, as data interpolation does not bring in new information. Even worse, *additional* computation is required in feature extraction. [1] On the other hand, in the case of scaling down, information is lost, although computation is saved in feature extraction. However, this saving in computation is at the cost of losing precious information. In the subspace morphing theory, since no normalization in scale is necessary, information is used "as is", and no extra cost in additional computation is required.

Identification capability: Identification capability is defined as the number of object classes an algorithm can identify before it starts to be confused. Regardless of what features are used, to simplify the analysis, assume that each class has the same number of samples, say N_c. If the feature space is fixed, say R^n, the upper bound of the number of object classes these techniques are able to identify is $C^n_{N_c}$. In the subspace morphing theory, on the other hand, since the feature space R^n is not fixed, and it can be morphed to an arbitrary space dynamically in the process of identification, this upper bound is open, and in general is much higher than the fixed upper bound, resulting in much larger

[1] The complexity of SVD is $O(mn^2)$ for a matrix in $m \times n$ [17].

UNCORRECTED PROOF

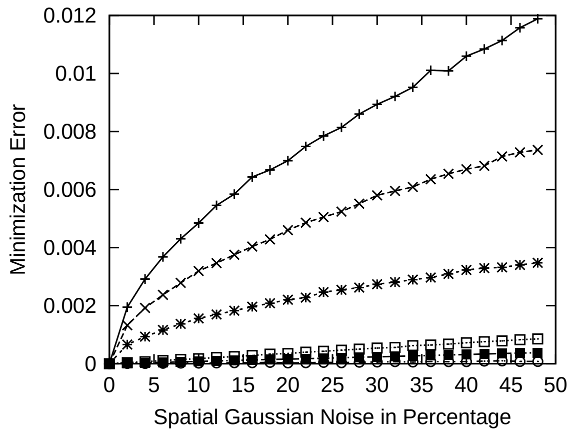

Fig. 1. Identification precision analysis based on noise corrupted intensity appearance. Simulated Gaussion noise is added into real query images at both spatial level (from 0% to 50%) and intensity level (5%, 10%, and 15%); the performances are computed by averaging over 1000 runs. The legends: $+, \times, *$—performances of the normalization based method after 15%, 10%, and 5% Gaussian noises are added into the intensity space, respectively; □, •, ○—performances of SM after 15%, 10%, and 5% Gaussian noises are added into the intensity space, respectively.

identification capability than that of the normalization based techniques.

Identification precision: Identification precision refers to how much appearance change is allowed before misclassification happens. Clearly, if normalization is conducted to scale up an image, the identification precision stays the same, as data interpolation does not bring in new information. However, if normalization is to scale down an image, the identification precision degrades significantly. This is shown below in two types of appearance change: noise corruption of original images and pose change of the same object. In both cases, images in the model base are both in half scale of the query images in row and column dimensions. In order to assure a fair comparison, a nearest neighbor based identification technique using SV features in the pre-defined model image space is implemented as a representative technique of the conventional, normalization based approaches.

In order to show the appearance change due to the change of intensity values, Fig. 1 gives the performances of the conventional technique and **SM**, when simulated Gaussian noise is added into real query images in both spatial level and intensity level. The degradation is measured as the standard minimization error in both techniques, and the results are obtained after averaging 1000 runs. Clearly, **SM** exhibits more stable values under the same levels of Gaussian noise than the conventional technique does.

In order to show the appearance change due to the change of object pose, Fig. 2 gives the performances of the both techniques when a query image of an asymmetric industrial part deviates a model image of the same object in pose from

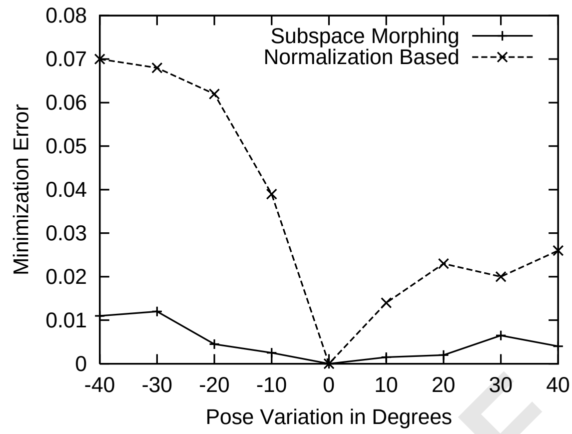

Fig. 2. Identification precision analysis based on pose change. Query images of nine different poses of the same object are input into both the normalization based method and SM for performance comparison.

-40° to 40° with 10° in each step. [2] Again, the performance degradation of the conventional technique is obviously more significant than that of **SM**.

In summary, subspace morphing theory performs significantly better than the conventional, normalization based methods if images are provided in different dimensions. This conclusion is also supported by experimental evaluations based on real data shown in Section 4.

4. Experiments

SM is implemented and has been used to identify different types of objects, including human faces, flowers, coffee mugs, space shuttles, toys, etc. Due to the appearance based nature, this method may be used to identify both 2D and 3D objects, both rigid and deformable objects. In this section, in order to show a performance comparison between this method and the conventional, normalization based methods, a human face database is used as a benchmark evaluation.

Note that the significance of the subspace morphing theory is that a better performance is achieved by taking advantage of the otherwise lost information in the normalization process when images are provided in different dimensions. In order to conduct a fair comparison through experiments, it would be ideal that a common, large scale testbed (such as FERET [20]) is available to test both the subspace morphing based technique and the conventional, normalization-based techniques. Regretfully, all the publicly existing face databases consist of normalized images, which nullifies the essential advantage of the subspace morphing theory. Based on this consideration, a special face database with images of different dimensions is constructed

[2] These images are obtained from University of Plymouth, UK.

UNCORRECTED PROOF

for this purpose. This database consists of two disjoint parts. A model base consists of 112 face images of 25 individuals, and a testing base consists of 226 face images of the same 25 individuals. These images were contributed voluntarily by people, or obtained through the Internet. In the process of data collection, special attention was paid to reflect significant variations in scale and appearances. The distribution of the number of images for each individual in the databases is not uniform, depending on the availability of the collected images from these individuals. Images in the model base are kept as all frontal or near-frontal images, while images in the testing base are of relatively large variations of appearances and scales to test the scalability and robustness of the evaluated algorithms.

Table 1
Face identification results[a]

Methods	SM	NE	LE	NS	LS
NS	208	92	104	102	56
NF	18	134	122	124	170
IR	92%	41%	46%	45%	25%

[a] Number of images in the model base: 112. Number of images in the testing base: 226. Number of subjects in both bases: 25. NS: Number of testing images succeeded. NF: Number of testing images failed. IR: Identification rate.

In order to show a comparison study between **SM** and the conventional techniques in this evaluation, four classic face identification techniques are implemented: nearest neighbor based eigenface (**NE**) [2], linear combination based eigenface (**LE**) [21], nearest neighbor based SVD (**NS**) [5], and linear combination based SVD (**LS**) by combining [5] and

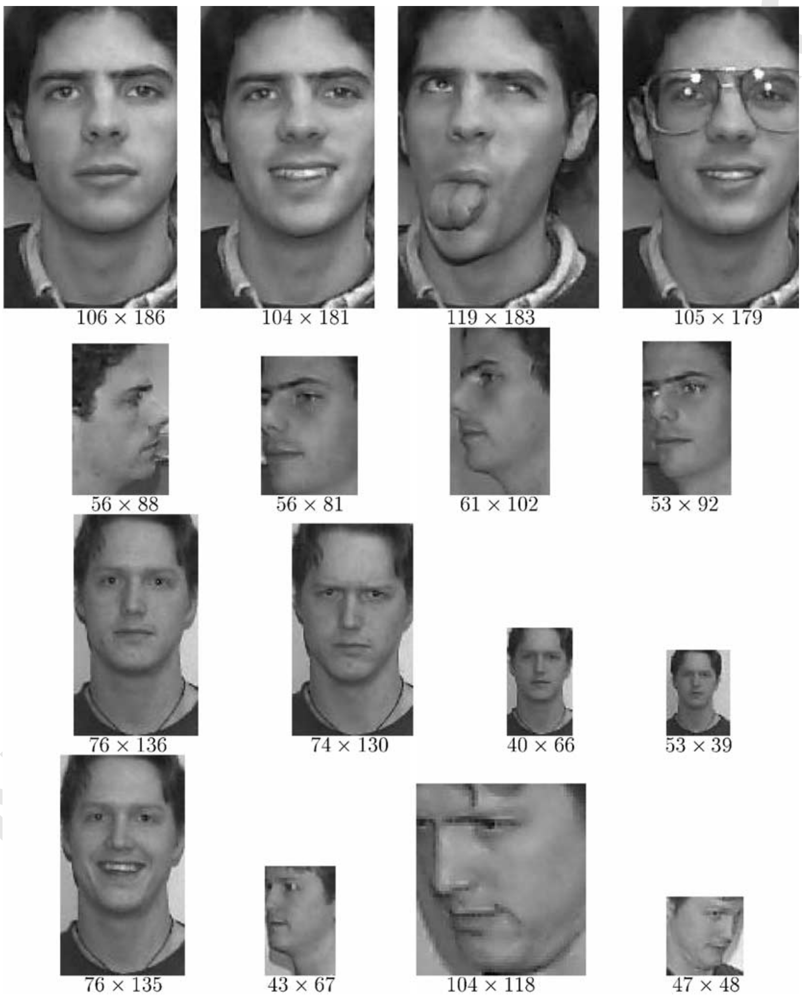

Fig. 3. Face examples in the experiments. The first and the third rows are the face images in the model database, and the second and the fourth rows are the images from the testing database; the dimensions of each images are indicated.

[21]. As all the conventional techniques require normalization in scale, all the 338 face images in the model and testing bases are scaled to a common dimension space of 120×80 to generate a separate version of the database for the sake of running these four classic algorithms. Table 1 shows the identification results of **SM** and the four conventional techniques. The performance of **SM** is noticeably better than those of the four conventional identification techniques under this evaluation, even with large variations of appearances between the model images and the testing images. Fig. 3 shows several examples of the recognition results in which **SM** succeeded and all the four classic techniques failed. Note the large variations of the appearances between the model images and the testing images. Also note that all the images either from the model base or from the testing base are in different dimensions.

For those testing images that **SM** fails to identify, they were all those with significant pose variations with respect to the images in the model base. In fact, some of the examples shown in Fig. 3 are in this "borderline" category (i.e., with significant pose variations): some of them were correctly identified by **SM** such as those shown in Fig. 3 while others failed to be correctly identified by **SM**. In other words, those testing images with significant pose variations shown in Fig. 3 are "good" cases, and there were other similar images in the same category as bad cases. In particular, for those individuals with fewer images in the model base, their testing images with significant pose variations tended to fail to be correctly identified more often than those with more images in the model base, as the numbers of the collected images in the model base for each individuals are not equal due to the uneven availabilities in the collection of the images. However, noise of the images in the "borderline" category were correctly identified by any of the four classic algorithms. In general, this evaluation indicates that **SM** appears to have substantially larger tolerance towards scale and appearance variations (especially in "internal deformation" such as expressions, with or without glasses, different backgrounds, different hair styles, different clothes, etc.) than the classic techniques do. More specific conclusion in comparison of **SM** with the classic techniques requires further larger scale evaluations.

The evaluation result for **LS** is a bit counter-intuitive. It is believed that in general the nearest linear combination based techniques should perform better than the nearest neighbor based techniques [21]. This may require further investigation in the future.

5. Conclusion and future work

A new theory for appearance based object identification called subspace morphing is presented in this paper. It differs from most existing identification techniques in the sense that it does not require normalization in image scale, and consequently is scale invariant. Furthermore, since this theory uses images "as is", when images are provided in different dimensions, which is a typical and common situation in many applications, it is shown that the "extra" information in the difference of the image scales contributes to a better performance of this technique in object identification. This conclusion is supported by both theoretical analysis and experimental evaluation.

The future work will be focused on determining how far the subspace morphing theory can go in object identification with large variations between testing and model images, especially with significant variations in pose, either through theoretic analyses or through experimental evaluations. This will allow further extensive and more "accurate" applications of this theory in solving for general object identification problems.

References

[1] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cog. Neurosci. 3 (1) (1991).

[2] A. Pentland, B. Moghaddam, T. Starner, View-based and modular eigenspaces for face recognition, in: Proceedings of the CVPR, 1994.

[3] B.J. Frey, A. Colmenarez, T.S. Huang, Mixtures of local linear subspaces for face recognition, in: Proceedings of the CVPR, IEEE Press, New York, 1998.

[4] D.B. Graham, N.M. Allinson, Norm^2-based face recognition, in: Proceedings of the CVPR, IEEE Computer Society Press, New York, 1999.

[5] Z. Hong, Algebraic feature extraction of image for recognition, Pattern Recognition 24 (1991).

[6] B.S. Manjunath, R. Chellappa, C.V.D. Malsburg, A feature based approach to face recognition, in: Proceedings of the CVPR, 1992.

[7] R. Brunelli, T. Poggio, Face recognition: features versus templates, IEEE Trans. Pattern Anal. Mach. Intell. 15 (1993).

[8] S. Lawrence, C.L. Giles, A.C. Tsoi, Convolutional neural networks for face recognition, in: Proceedings of the CVPR, 1996.

[9] C. Liu, H. Wechsler, Evolutionary pursuit and its application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 22 (6) (2000) 570–582.

[10] C. Liu, H. Wechsler, Learning the face space—representation and recognition, in: Proceedings of the ICPR, IAPR, 2000.

[11] G.J. Edwards, C.J. Taylor, T.F. Cootes, Improving identification performance by integrating evidence from sequences, in: Proceedings of the CVPR, IEEE Computer Society Press, New York, 1999.

[12] N. Costen, T.F. Cootes, G.J. Edwards, C.J. Taylor, Simultaneous extraction of functional face subspaces, in: Proceedings of the CVPR, IEEE Computer Society Press, New York, 1999.

[13] B. Moghaddam, A. Pentland, Probabilistic visual learning for object representation, Trans. Pattern Anal. Mach. Intell. (1997).

[14] H. Schneiderman, T. Kanade, Probabilistic modeling of local appearance and spatial relationships for object recognition, in: Proceedings of the CVPR, IEEE, New York, 1998.

UNCORRECTED PROOF

[15] A. Colmenarez, B. Frey, T.S. Huang, A probabilistic framework for embedded face and facial expression recognition, in: Proceedings of the CVPR, IEEE Computer Society Press, New York, 1999.
[16] I.J. Cox, J. Ghosn, P.N. Yianilos, Feature-based face recognition using mixture-distance, in: Proceedings of the CVPR, 1996.
[17] G.H. Golub, C.F.V. Loan, Matrix Computations, 2nd Edition, The Johns Hopkins University Press, Baltimore, MD, 1989.
[18] S. Lang, Linear Algebra, Springer, Berlin, 1987.
[19] S.E. Umbaugh, Computer Vision and Image Processing, A Practical Approach Using CVIP tools, Prentice-Hall, Englewood Cliffs, NJ, 1998.
[20] P. Phillips, H. Moon, P. Rauss, S. Rizvi, The feret evaluation methodology for face-recognition algorithms, in: Proceedings of the CVPR, 1997.
[21] S.Z. Li, Face recognition based on nearest linear combinations, in: Proceedings of the CVPR, IEEE, New York, 1998.

About the Author—ZHONGFEI ZHANG is an Assistant Professor in the Computer Science Department at the State University of New York (SUNY) at Binghamton. He was a research scientist in the Center of Excellence for Document Analysis and Recognition (CEDAR) and Research Assistant Professor in the Department of Computer Science and Engineering, both at SUNY at Buffalo before he joined the faculty at SUNY at Binghamton. His research interests include multimedia information indexing and retrieval, image understanding, and pattern recognition. He has published extensively in national and international journals and conferences in these areas. He received a BS with honors in Electronics Engineering and an MS in Information Sciences both from Zhejiang University, China, and a Ph.D. in Computer Science from the University of Massachusetts.

About the Author—ROHINI K. SRIHARI is Associate Professor of Computer Science and Engineering at the State University of New York at Buffalo. She is also the founder and chief scientist at Cymfony Inc., a company developing software products based on information extraction technology. She has worked in both natural language processing as well as computer vision. Her current research interests include multimedia information retrieval, multimodal image annotation systems, information extraction and question answering systems, including multimedia question answering. She received a BMath in Computer Science from the University of Waterloo, Ontario, Canada, and MS and Ph.D. in Computer Science from the State University of New York at Buffalo.

UNCORRECTED PROOF