Sample Sufficiency and PCA Dimension for Statistical Shape Models

Lin Mei*, Michael Figl, Ara Darzi, Daniel Rueckert, and Philip Edwards*

Dept. of Biosurgery and Surgical Technology Imperial College London, UK
{l.mei,eddie.edwards}@imperial.ac.uk

Abstract. Statistical shape modelling(SSM) is a popular technique in computer vision applications, where the variation of shape of a given structure is modelled by principal component analysis (PCA) on a set of training samples. The issue of sample size sufficiency is not generally considered. In this paper, we propose a framework to investigate the sources of SSM inaccuracy. Based on this framework, we propose a procedure to determine sample size sufficiency by testing whether the training data stabilises the SSM. Also, the number of principal modes to retain (PCA dimension) is usually chosen using rules that aim to cover a percentage of the total variance or to limit the residual to a threshold. However, an ideal rule should retain modes that correspond to real structural variation and discard those that are dominated by noise. We show that these commonly used rules are not reliable, and we propose a new rule that uses bootstrap stability analysis on mode directions to determine the PCA dimension.

For validation we use synthetic 3D face datasets generated using a known number of structural modes with added noise. A 4-way ANOVA is applied for the model reconstruction accuracy on sample size, shape vector dimension, PCA dimension, and the noise level. It shows that there is no universal sample size guideline for SSM, nor is there a simple relationship to the shape vector dimension (with p-Value=0.2932). Validation of our rule for retaining structural modes showed it detected the correct number of modes to retain where the conventional methods failed. The methods were also tested on real 2D (22 points) and 3D (500 points) face data, retaining 24 and 70 modes with sample sufficiency being reached at approximately 50 and 150 samples respectively. We provide a foundation for appropriate selection of PCA dimension and determination of sample size sufficiency in statistical shape modelling.

1 Introduction

Statistical shape modelling (SSM) is a technique for analysing variation of shape and generating or inferring unseen shapes. A set of sample shapes is collected and PCA is performed to determine the principal modes of shape variation. These modes can be optimised to fit the model to a new individual, which is the familiar active shape model (ASM) [1,2,3]. Further information, such as texture, can be included to create an active appearance model [4] or morphable model [5].

* We would like to thank Tyco Healthcare for funding Lin Mei's PhD studentship. We are also grateful to many other members of the Department of Computing and the Department of Biosurgery and Surgical Technology at Imperial College.

D. Forsyth, P. Torr, and A. Zisserman (Eds.): ECCV 2008, Part IV, LNCS 5305, pp. 492–503, 2008.

Despite its popularity, PCA-based SSMs are normally trained from datasets for which the issue of sufficiency is not considered. The PCA dimension for an SSM is often chosen by rules that assume either a given percentage or level of noise. As will be shown later in this paper that these two methods are highly dependent on sample size.

In this paper, we review the discussions on sample size sufficiency for a closely related field, common factor analysis (CFA), and design a mathematical framework to investigate the source of PCA model error. This framework provides a theoretical evaluation of the conventional rules for retaining PCA modes, and enables analysis of sample size sufficiency for PCA. We then propose a rule for retaining only stable PCA modes that uses a *t*-test between the bootstrap stability of mode directions from the training data and those from pure Gaussian noise. The convergence of the PCA dimension can then be used as an indication of sample sufficiency.

We verify our framework by a 4-way ANOVA for reconstruction accuracy is applied to the models trained from synthetic datasets generated under different conditions. Our PCA dimension rule and procedure for sample sufficiency determination are validated on the synthetic datasets and demonstrated on real data.

2 Background

2.1 Minimum Sample Size for CFA

There is little literature on the issue of minimum sample size for PCA. In the related field of CFA, however, this issue has been thoroughly discussed. CFA is commonly used to test or discover common variation shared by different test datasets. Guidelines for minimum sample size in CFA involve either a universal size regardless of the data dimension or a ratio to the data dimension. Recommendations for minimum size neglecting the sample dimension and the number of expected factors vary from 100 to 500 [6]. Such rules are not supported by tests on real data. Doubts have been raised about a universal sample size guideline since it neglects the data dimension. Size-variable ratios (SVR) may be more appropriate and values of between 2:1 to 20:1 have been suggested [7]. There have been a number of tests using real data, but no correlation was found between SVR and the mode stability [8], nor has any minimum value for SVR emerged [9]. The minimum sample size needed in these real tests is not consistent either, varying from 50 [8], to 78-100 [9], 144 [10], 400 [11] and 500 or more [12].

The inconsistency among these results shows that the minimum size depends on some nature of the data other than its dimension. MacCallum et al. [13,14] proposed a mathematical framework for relating the minimum sample size for CFA with its communality and overdetermination level. They then designed an experiment using 4-way ANOVA to study the effects of communality, overdetermination level, model error and sample size on the accuracy in recovering the genuine factors from synthetic data. The results showed that communality had the dominant effect on the accuracy regardless of the model error. The effect of overdetermination level was almost negligible when communality is high. In low communality tests, accuracy improves with larger sample size and higher accuracy was found in tests with lower overdetermination levels.

There is no equivalent to communality and overdetermination level for PCA. Instead, the factors we consider are the data dimension and the number of genuine structural modes that are retained.

2.2 Number of Modes to Retain for SSM

Many rules choosing the number of modes to retain for SSM and PCA have been proposed [6,15,16,17]. The most popular rule used in SSM for structural mode determination is simply to take the leading modes covering a percentage of the total variance in the sample set. The percentage is arbitrarily set, which equivalent to simply assuming a corresponding percentage of noise.

Another popular rule is to discard the least principal modes until the sum of total variance, which is the model residual on the training data, reaches a certain threshold. This threshold is normally set according to the error tolerance of the application.

Stability measurements for PCA have been proposed to determine the number of modes. Given two shape models trained from different sample sets, Daudin et al [18] used a sum of correlation coefficients between pairs of principal components; Besse et al [19] used a loss function derived from an Euclidean distance between orthogonal projectors; Babalola et al [20] used the Bhattacharya Metric to measure the similarity of PCA models from different sample sets. Resampling techniques such as bootstrapping [18] and jackknifing [19] can be used. The distribution of PCA modes across the replicates reflects their distribution in the population, allowing stability analysis to be performed. The selected principal modes span a subspace. Besse et al. proposed a framework for choosing the number of modes based on their spanned-space stability [21]. This method differentiates structural modes and noise-dominated modes when the sample set is large. However, as will be shown in the section 4.3, this method can only provide a estimation of the number of modes when the sample size is sufficient.

3 Theories

3.1 Sources of PCA Model Inaccuracy

We propose the following mathematical framework to examine the characteristics affecting the sufficiency of a sample set drawn from a population with genuine modes of variation, listed in the column of A. Due to the presence of noise, we have $\widehat{X}$ instead of X, and the PCA modes from $\widehat{X}$ are $\widehat{A}$. The model inaccuracy can be expressed as the difference between the covariance matrices $\Delta = \widehat{X}\widehat{X}^T - XX^T$.

Let $X = AW$ and $\widehat{X} = \widehat{AW}$, we have:

$$\widehat{X} = \widehat{AW} = AA^T\widehat{AW} + (I - AA^T)\widehat{AW} \tag{1}$$

Since A is orthonormal, $(I - AA^T)$ is a diagonal matrix with only 1s and 0s. Hence $(I - AA^T) = NN^T$. Equation 1 becomes:

$$\widehat{X} = AA^T\widehat{AW} + NN^T\widehat{AW} = A\widehat{W}_A + N\widehat{W}_N \tag{2}$$

Applying equation 2 on the covariance matrix of $\boldsymbol{X}$:

$$\begin{aligned}\widehat{X}\widehat{X}^T &= A\widehat{W_A}\widehat{W}_A^T A^T + A\widehat{W_A}\widehat{W}_N^T N^T + N\widehat{W}_N\widehat{W}_A^T A^T + N\widehat{W}_N\widehat{W}_N^T N^T \\ &= A\widehat{\Sigma_{AA}}A^T + A\Sigma_{AN}N^T + N\Sigma_{NA}A^T + N\Sigma_{NN}N^T \end{aligned} \quad (3)$$

The model inaccuracy becomes:

$$\begin{aligned}\Delta &= \widehat{X}\widehat{X}^T - XX^T \\ &= A(\widehat{\Sigma_{AA}} - \Sigma_{AA})A^T + A\Sigma_{AN}N^T + N\Sigma_{NA}A^T + N\Sigma_{NN}N^T \\ &= (A(\Sigma_{EE})A^T + A\Sigma_{AN}N^T + N\Sigma_{NA}A^T) + N\Sigma_{NN}N^T \end{aligned} \quad (4)$$

A PCA model error consists of two parts:

$E_N = N\Sigma_{NN}N^T$, the error introduced by sampling noise modes that are orthogonal to $\mathbb{A}$. E_N depends only on the noise level introduced by human interaction or measurement error during the process of building an SSM. Increasing sample size would cause little reduction in this error if noise level remains the same.

$E_l = A(\Sigma_{EE})A^T + A\Sigma_{AN}N^T + N\Sigma_{NA}A^T$, the error along the subspace spanned by structural modes, $\mathbb{A}$. This is due to noise affecting the sample coefficients and insufficient coverage of the dimensions in $\mathbb{A}$. Therefore, E_l increases with PCA dimension, $rank(A)$. It also affected by noise level because at high noise level some structural modes with small variances may be swamped by noise.

Rather counter-intuitively, E_l is not dependent on the shape vector dimension, as will be shown in the section 4.2. However, higher E_N can result from higher shape vector dimension, which therefore increases Δ.

3.2 Sample Size Requirement

According to the framework in section 3.1, the sample size requirement for PCA only depends on two factors: number of structural modes in the dataset, and the level of noise. Hence we propose the following procedure for sample sufficiency determination. For a sample set, X, of n samples:

PCA Sample Size Sufficiency Test

1) Apply PCA on X, to get a set of modes B.

2) Starting with a reasonably small number, n^*, construct a set X_j^* of n^* samples randomly drawn, allowing repeats, from X.

3) Apply PCA to X_j^* to get a set of modes B_j^* and resolve mode correspondence with respect to B.

4) Find the number of structural modes in B_j^*, k.

5) Repeat 2-4 with an increased n^*. If k converges before n^* reaches n, we have sufficient samples. Otherwise, further sample data is required.

Step 4 in this procedure requires determination of the number of structural modes, which is a common question for PCA. These rules are sometimes is called *stopping rules*, since they determine where to stop including the principal modes.

3.3 PCA Dimension by the Stability of Mode Direction

It is generally assumed that most of the structural variation is modelled by the leading principal modes. Noise affects the lower ranking modes and dominates those after a cut-off point further down the rank. Going back to our framework, a cut-off point is chosen for retaining principal modes in order to reduce E_N. However, since genuine structural variation may still be present in the least principal modes, discarding them would increase E_l. There is trade-off between E_N and E_l. Stopping rules should aim at discarding only modes that are sufficiently dominated by noise.

Assuming noise is randomly spread across all the dimensions, mode instability can be a good indication of the point where noise begins to dominate. There is a risk with tests using the magnitude of the variance that stopping rules will be dominated by the first few modes and fail to identify the correct cut-off point. Also, it is the mode directions that define the basis of a shape model for fitting or synthetic shape generation. Therefore we propose a stopping rule based on the stability of the mode direction only.

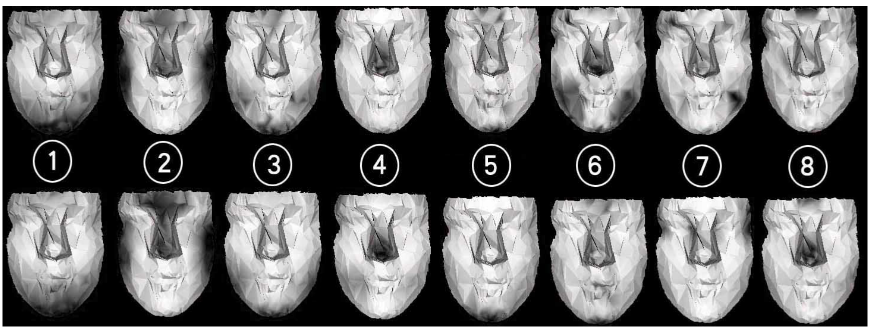


Fig. 1. Comparison of Leading 8 Eigenmodes from two mutually exclusive sets of 50 samples from our 3D face mesh database, aligned according to eigenvalue ranks. Darker texture implies larger variation, showing many mismatched after the 4th mode.

3.3.1 Establishing Mode Correspondence

Examining individual modes requires mode correspondence. Normally, this is done by matching those with the same eigenvalue ranks. Significant variation can be found between individual modes drawn from different sample sets with the same ranking, as shown in figure 1. Although leading modes may correspond, mode 5 on the top seems to correspond with mode 6 on the bottom, and modes after 6 on the top seems to correspond to none at the bottom. However, the combined modes from different sample sets may still span similar subspaces. Mode alignment can be achieved by minimising the distance between these subspaces.

For the leading PCA modes $\{(\boldsymbol{a}_i, \lambda_i) | |\boldsymbol{a}_i| = 1\}$ of an n-dimensional distribution, we define the principal spanned space (PSS) as the subspace $\mathbb{S}^k$ spanned by $\{\boldsymbol{a}_i\}$, where the distance measure used by Besse et al.[19] can be applied:

$$d(\mathbb{A}^k, \mathbb{B}^k) = k - trace(AA^T BB^T) \tag{5}$$

where the columns of A and B are the modes spanning PSS $\mathbb{A}^k$ and $\mathbb{B}^k$.

For two sets of PCA modes, $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$, trained from different sample sets of a common distribution, the following rule can be used to establish correspondence. The first mode in $\boldsymbol{a}_i$ corresponds to the mode of a replicate that minimises $d(\mathbb{S}_{\mathbb{a}}{}^1, \mathbb{S}_{\mathbb{b}}{}^1)$, and we proceed iteratively. Assume we have already aligned $\mathbb{S}_{\mathbb{a}}{}^k$, the PSS from the first k modes in $\boldsymbol{a}_i$, to the spanned space $\mathbb{S}_{\mathbb{b}}{}^k$ from k modes in the replicate $\boldsymbol{b}_i$. The mode in $\boldsymbol{b}_i$ that corresponds to the k+1$^{\text{th}}$ mode in $\boldsymbol{a}_i$ will be the one that minimises $d(\mathbb{S}_{\mathbb{a}}{}^{k+1}, \mathbb{S}_{\mathbb{b}}{}^{k+1})$.

3.3.2 Bootstrap Stability of PCA Modes

Bootstrap stability analysis can be used to analyse mode stability. We use the angles between mode directions as the measurement of distance between corresponding modes from different replicates. The instability, ξ, of mode $\boldsymbol{a}_i$ is given by:

$$\xi(\boldsymbol{a}_i) = \frac{\sum_{j=1}^{m} arccos(\boldsymbol{a_i}'_j \cdot \widehat{\boldsymbol{\alpha_i}}) \cdot}{m\pi} \tag{6}$$

where $\widehat{\boldsymbol{\alpha_i}}$ is the mean mode vector and m is the number of bootstrap replicates.

3.3.3 Stopping Rule Based on a *t*-Test against Synthetic Gaussian Noise

Since noise-dominated modes should have higher instability than structural modes, a threshold on ξ can be used to differentiate them from structural modes. However, the choice for the threshold is arbitrary and is found to be sensitive to the size of replicates. Instead, assuming the distribution of angles between corresponding modes is Gaussian, a one-tailed *t*-test can be used to establish whether a mode is dominated by noise to a given significance level.

We generate a pure Gaussian noise dataset to compare with the test dataset. All conditions must be the same – the dimensionality, the number of samples in the dataset, the number of replicates, and the number of samples in each replicate. Since we are only interested in mode directions, the level of noise is not important. Let the angle for the first pure noise mode to be $\boldsymbol{\alpha}_1$ and the angle for the test samples to be $\boldsymbol{a}_i$, The null hypothesis of the *t*-test is $H_0 : \xi(\boldsymbol{\alpha}_1) > \xi(\boldsymbol{a}_i)$. By rejecting H_0 at a given confidence level, one can safely conclude that a mode is not dominated by noise.

4 Experiments

We demonstrate the correctness of our theories with three sets of experiments. First a 4-way ANOVA is performed on synthetic datasets to show how the PCA model accuracy is affected by different features as it is discussed in section 3.1. Then we show that our stopping rule is able to identify the correct number of modes in the synthetic samples for which commonly used rules fail. This shows that our rule can be used to determine PCA sample sufficiency, by following the procedure presented in section 3.2. This is applied to two different sets of real samples.

4.1 Real Datasets

Two real shape datasets are used in the experiments. The first one comprises 150 samples of 3D faces with 5090 points each from University of Notre Dame [22] preprocessed using Papatheodorou's method [23], and the second one consists of 135 samples

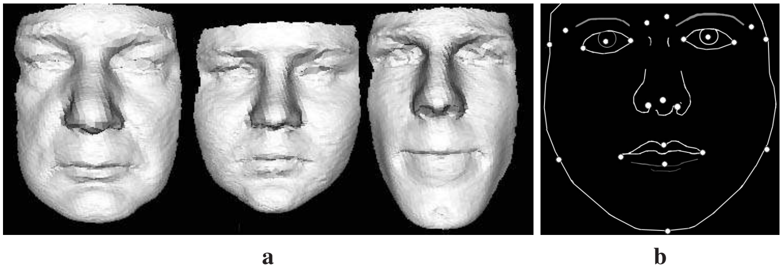


Fig. 2. Examples from real 3D Face database (a) and landmarks of 2D AR Face database

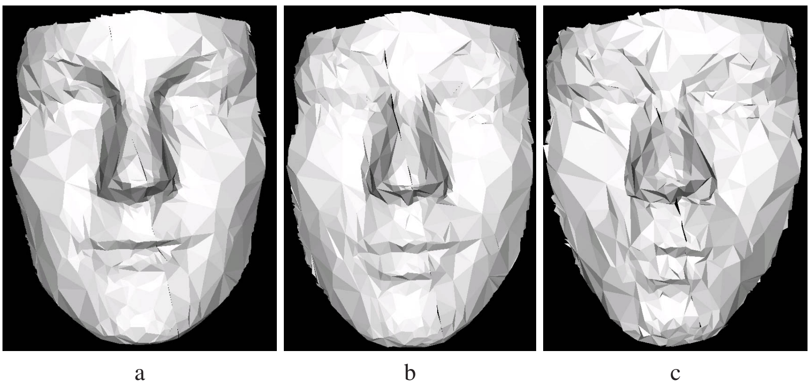


Fig. 3. Examples of three synthetic faces generated with 70 modes with shape vector dimension being 2100. Different noise levels are applied: 0.1mm (a), 0.25mm (b) and 0.5mm (c). Noise starts to become visible in (b) and (c).

from the landmarks (22 points) [24] of 2D AR face database [25]. Examples from these two datasets are shown in figure 2.

4.2 ANOVA Results

For validation of our framework, we generate a dataset consists of 8960 subsets, each having different combinations of: sample sizes, numbers of modes to generate, levels of Gaussian noise and decimated to different number of points. A list of choices for different characteristics are shown as follows

Sample Sizes (SS): 50, 100, 150, 200, 250, 300, 350, 400
450, 500, 550, 600, 650, 700, 750, 800
Shape Vector Dimension: 300, 600, 900, 1200, 1500, 1800, 2100
Number of Genuine Modes: 10, 20, 30, 40, 50, 60, 70, 80
Gaussian Noise Levels (in mm): 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5
Examples of faces generated under different conditions are given in figure 3.

PCA is applied to each of the 8960 subsets of the first synthetic dataset. Results are compared to the original modes that used to generate the data. Measurement described in equation 5 is used to calculate the error of the models trained from the subsets. A 4-way ANOVA was performed to find out which characteristics influence the model accuracy. As shown in table 1, the results confirm the correctness of our framework introduced in section 3.1. Sample size and number of genuine modes in the dataset act as the major source of influence on the model accuracy. Noise also has a significant but small influence. Also the result showed that the effect of sample dimension is negligible.

Table 1. Result of 4-Way ANOVA

Source	Sum of Squares	DoF	Mean Squares	F-Statistic	p-Value
Sample Size	194.534	15	12.9689	11315.66	¡0.03
Sample Dimension	0.008	6	0.0014	1.22	0.2932
Number of Genuine Modes	83.055	7	11.865	10352.48	¡0.03
Gaussian Noise Level	0.513	9	0.057	49.74	¡0.03

4.3 Number of Modes to Retain for SSM

We have validated previous stopping rules and our method using synthetic data generated with a known number of background modes. These shapes are generated using the leading 80 modes of the model built from all the 150 3D Faces, decimated to 500 points for faster experiments. Gaussian noise with 1mm standard deviation is added to each element of the shape vector. Example faces from the synthetic set are shown in figure 4. Stopping rules applied to this dataset should not retain more than 80 modes.

We validated the rule which retains 95% of the cumulative variance using synthetic datasets sized from 100 to 600. Compactness plots are shown in figure 5(a). With increasing sample size, the number of modes retained by this rule increases beyond 80, where the noise dominates the variance. These noise modes contribute to an increasing proportion of the total variance with increasing sample size, and the number of modes covering 95% of the total variance increases accordingly. A similar trend was also found

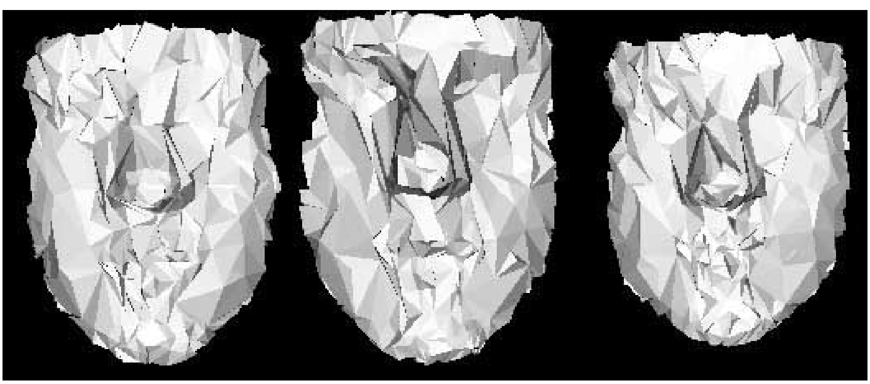
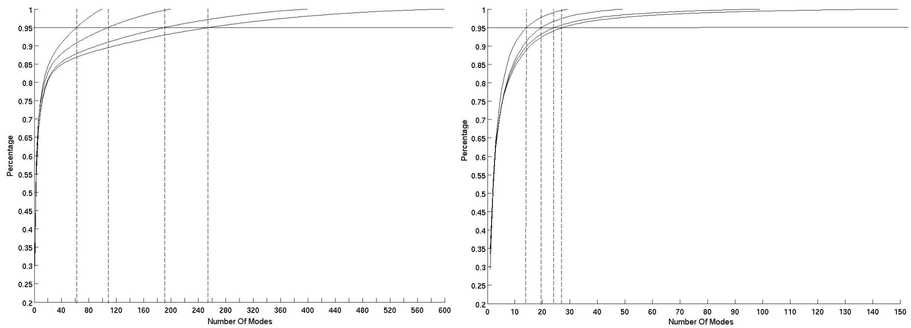


Fig. 4. Synthetic faces generated using 80 modes, added 1mm Gaussian noise on to each element of the shape vector with dimension 1500



a b

Fig. 5. 95% thresholded compactness plots of synthetic 3D face datasets (a) with 100, 200, 400 and 600 samples and real 3D face datasets (b) with 30, 50, 100 and 150 samples. The number of retained modes is clearly dependent on sample size.

Table 2. Number of modes to keep the point error below 1mm

Number of Samples	50	100	150	200	250	300	350	400	450	500
Number of Modes	32	60	95	108	120	140	169	186	204	219

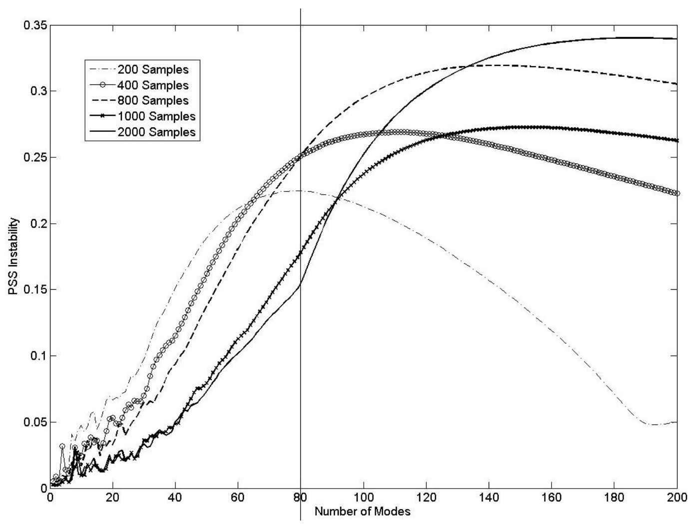


Fig. 6. Instability of PSS for synthetic datasets for synthetic datasets sized from 200 to 2000

for the real data as shown in figure 5(b), which strongly suggests that this rule is unreliable and should not be used. A similar effect, as shown in table 2, was found for the stopping rule that discards the least principal modes until the average error of each point reaches 1mm.

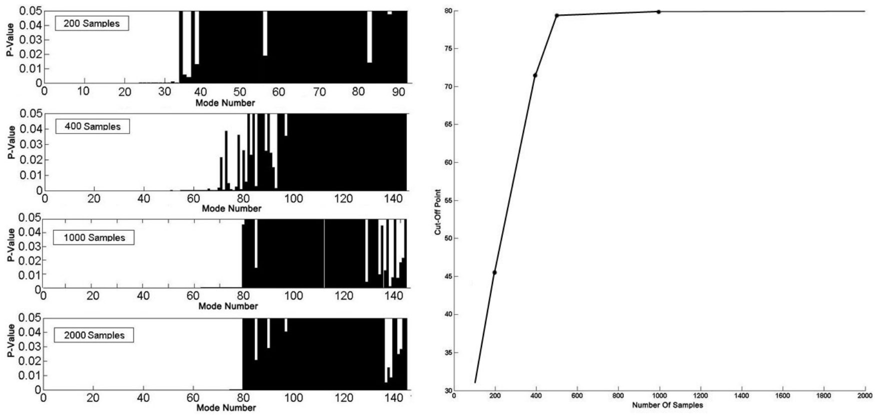


Fig. 7. *t*-Test Based stopping rule on synthetic datasets

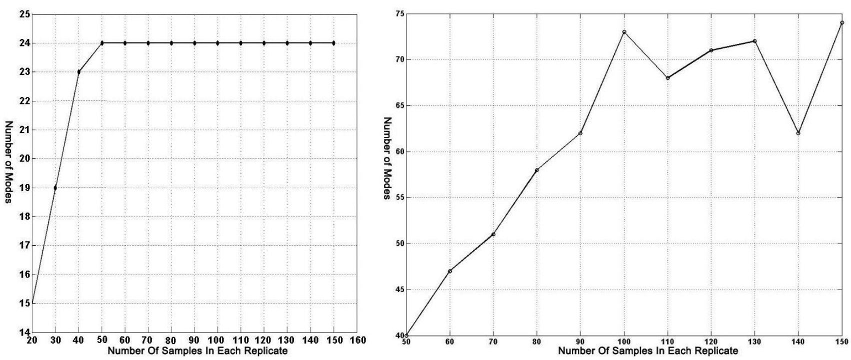


Fig. 8. Result of real datasets sufficiency test. Left: 2D faces; Right: 3D faces.

The method of Besse et al [21] was validated with synthetic datasets sized from 200 to 400. A plot of instability, measured as the distance between subspaces spanned by different replicates, is shown in figure 6. Although this method provides a visible indication of the correct number of modes to retain when the sample size is sufficiently large, it cannot identify the lower number of modes that should be retained when the sample size is insufficient.

Our method was validated with synthetic datasets sized from 100 to 2000. Figure 7 shows the number of modes to retain versus the sample size is also shown. Our stopping rule does not have the tendency to go beyond 80 with large sample sizes. It also identifies a lower number of stable modes to retain for smaller sample sizes. It appears a sample size of around 500 is sufficient.

4.4 Sample Size Sufficiency Test for SSM

Figure 8 shows the results of sample size sufficiency tests on the three real datasets we have. For the 2D dataset, the plot obviously converges at 24 modes with 50 samples.

With the 3D faces, the graph appears close to convergence at around 70 modes for the 150 samples. These results suggest both face datasets are sufficient.

5 Conclusion and Discussion

We propose a mathematical framework to analyse the sources of inaccuracy in PCA, which suggests that only the number of genuine structural modes in the dataset and the noise level will influence the minimum sample size required to stabilise the model. There is no trivial relationship between the shape vector dimension and the required sample size. These results are confirmed by a 4-way ANOVA on synthetic data.

We propose a stopping rule that retains modes according to the stability of their directions. We also establish mode correspondence from different training sets by minimising the distance between the principal spanned spaces rather than simply by the rank of their eigenvalues. For a synthetic dataset generated with known structural modes plus added noise, our method converges correctly where conventional methods did not.

The number of genuine structural modes is not known, but the number of modes to use for a given dataset can be estimated using our stopping rule. The convergence of this rule can then be used as an indicator of sample size sufficiency.

Resulting sample size sufficiency suggest 50 samples is sufficient for 2D face landmarks(22 points), retaining 24 modes in total, and 150 samples is sufficient for the 3D face meshes (500 points), where around 70 modes are retained. We believe this is the first principled test for sample sufficiency and determination of the number of modes to retain for SSM. It can also be applied to other applications of PCA and related fields.

References

1. Cootes, F., Hill, A., Taylor, C., Haslam, J.: The use of active shape models for locating structures in medical images. In: Proc. IPMI, pp. 33–47 (1993)
2. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models and their training and application. Comput. Vis. Image Underst. 61(1), 38–59 (1995)
3. Sukno, F.M., Ordas, S., Butakoff, C., Cruz, S.: Active shape models with invariant optimal features: Application to facial analysis. IEEE Trans. Pattern Anal. Mach. Intell. 29(7), 1105–1117 (2007) (Senior Member-Alejandro F. Frangi)
4. Cootes, T., Edwards, G., Taylor, C.: Active appearance models. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(6), 681–685 (2001)
5. Blanz, V., Vetter, T.: Face recognition based on fitting a 3D morphable model. IEEE Transactions On Pattern Analysis And Machine Intelligence 25, 1063–1074 (2003)
6. Osborne, J., Costello, A.: Sample size and subject to item ratio in principal components analysis. Practical Assessment, Research and Evaluation 9(11) (2004)
7. Guadagnoli, E., Velicer, W.: Relation of sample size to the stability of component patterns. Psychological Bulletin 103, 265–275 (1988)
8. Barrett, P., Kline, P.: The observation to variable ratio in factor analysis. Personality Study and Group Behavior 1, 23–33 (1981)
9. Arrindell, W., van der Ende, J.: An empirical test of the utility of the observations-to-variables ratio in factor and components analysis. Applied Psychological Measurement 9(2), 165–178 (1985)

10. Velicer, W., Peacock, A., Jackson, D.: A comparison of component and factor patterns: A monte carlo approach. Multivariate Behavioral Research 17(3), 371–388 (1982)
11. Aleamoni, L.: Effects of size of sample on eigenvalues, observed communalities, and factor loadings. Journal of Applied Psychology 58(2), 266–269 (1973)
12. Comfrey, A., Lee, H.: A First Course in Factor Analysis. Lawrence Erlbaum, Hillsdale (1992)
13. MacCallum, R., Widaman, K., Zhang, S., Hong, S.: Sample size in factor analysis. Psychological Methods 4, 84–99 (1999)
14. MacCallum, R., Widaman, K., Hong, K.P.S.: Sample size in factor analysis: The role of model error. Multivariate Behavioral Research 36, 611–637 (2001)
15. Jackson, D.: Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. Ecology 74, 2204–2214 (1993)
16. Jolliffe, I.: Principal Component Analysis, 2nd edn. Springer, Heidelberg (2002)
17. Sinha, A., Buchanan, B.: Assessing the stability of principal components using regression. Psychometrika 60(3), 355–369 (2006)
18. Daudin, J., Duby, C., Trecourt, P.: Stability of principal component analysis studied by the bootstrap method. Statistics 19, 341–358 (1988)
19. Besse, P.: PCA stability and choice of dimensionality. Statistics& Probability 13, 405–410 (1992)
20. Babalola, K., Cootes, T., Patenaude, B., Rao, A., Jenkinson, M.: Comparing the similarity of statistical shape models using the bhattacharya metric. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 142–150. Springer, Heidelberg (2006)
21. Besse, P., de Falguerolles, A.: Application of resampling methods to the choice of dimension in PCA. In: Hardle, W., Simar, L. (eds.) Computer Intensive Methods in Statistics, pp. 167–176. Physica-Verlag, Heidelberg (1993)
22. University of Notre Dame Computer Vision Research Laboratory: Biometrics database distribution (2007), http://www.nd.edu/~cvrl/UNDBiometricsDatabase.html
23. Papatheodorou, T.: 3D Face Recognition Using Rigid and Non-Rigid Surface Registration. PhD thesis, VIP Group, Department of Computing, Imperial College, London University (2006)
24. Cootes, T.: The AR face database 22 point markup (N/A), http://www.isbe.man.ac.uk/∼bim/data/tarfd_markup/tarfd_markup.html
25. Martinez, A., Benavente, R.: The AR face database (2007), http://cobweb.ecn.purdue.edu/~aleix/aleix_face_DB.html