Automatic Estimation of the Number of Deformation Modes in Non-rigid SfM with Missing Data*

Carme Julià[1], Marco Paladini[2], Ravi Garg[2], Domenec Puig[1], and Lourdes Agapito[2]

[1] Department of Computer Science and Mathematics
Universitat Rovira i Virgili, Tarragona, Spain
[2] School of Electronic Engineering and Computer Science
Queen Mary University of London, UK
{carme.julia,domenec.puig}@urv.cat,
{paladini,rgarg,lourdes}@dcs.qmul.ac.uk

Abstract. This paper proposes a new algorithm to estimate automatically the number of deformation modes needed to describe a non-rigid object with the well-known low-rank shape model, focusing on the missing data case. The 3D shape is assumed to deform as a linear combination of K rigid shape bases according to time varying coefficients. One of the requirements of this formulation is that the number of bases must be known in advance. Most non-rigid structure from motion (NRSfM) approaches based on this model determine the value of K empirically. Our proposed approach is based on the analysis of the frequency spectra of the x and y coordinates corresponding to the individual image trajectories, which are seen as 1D signals. The frequency content of the 2D trajectories is encoded using the modulus of the Discrete Cosine Transform (DCT) of the signals. Our hypothesis is that the value of K that gives the best prediction of the missing data also provides the best 3D reconstruction. Our proposed approach does not assume any prior knowledge and is independent of the 3D reconstruction algorithm used. We validate our approach with experiments on synthetic and real sequences.

Keywords: non-rigid SfM, Discrete Cosine Transform, frequency content.

1 Introduction

The Structure from Motion (SfM) problem is defined as the simultaneous estimation of the 3D coordinates of some scene points and the relative motion between the camera and the world purely from 2D trajectories of tracked features. Tomasi and Kanade [11] introduced the factorization technique to tackle the SfM problem in the case of rigid objects viewed by an orthographic camera

* This work was partially funded by the European Research Council under ERC Starting Grant agreement 204871-HUMANIS.

A. Heyden and F. Kahl (Eds.): SCIA 2011, LNCS 6688, pp. 381–392, 2011.

by imposing the rigidity constraint. This assumption was since relaxed to extend structure from motion algorithms to the non-rigid domain. Bregler et al. [4] were the first to propose a factorization approach based on a low-rank shape model to represent the deforming shape as a linear combination of K basis shapes which encode its main modes of deformation.

However, the non-rigid structure from motion problem (NRSfM) is severely under-constrained. Recent approaches to NRSfM have focused on the use of different optimization schemes and the definition of priors to overcome the problems caused by the inherent ambiguities and degeneracies [5,2,12,9]. The linear basis shape model has also allowed the formulation of closed form solutions both for the affine [14] and the perspective [15,13,6] cases. However, closed form solutions are known to be very sensitive to noise [3,12] and cannot deal with missing data.

Akhter et al. [1] depart from the low-rank shape model and instead describe the time varying 3D trajectories as a combination of trajectory bases for which they choose the Discrete Cosine Transform (DCT). The advantage of their approach is that the bases are generic and do not need to be estimated for each sequence.

So far there has been little work on the automatic estimation of the number of deformation modes needed to represent the time varying shape. Most of the aforementioned approaches estimate the number of deformation modes from the rank of the measurement matrix (e.g., [14]) or empirically (e.g., [12,9,1]). Roy-Chowdhury [10] introduces the *deformability index* (DI) to estimate the number of basis shapes by taking into account the statistics of the underlying noise in the shape sequence. However, this approach cannot deal with missing data. In their coarse-to-fine shape model, Bartoli et al. [2] use the Cross-Validation score to automatically decide when to stop adding modes of deformation to the model.

This paper addresses the automatic selection of the number of basis shapes needed to describe a non-rigid object represented using the well known low-rank shape model, focusing on the missing data case. The goal is to select the number of bases (K) that gives the best 3D reconstruction. Our hypothesis is that the value of K that gives the best prediction of the missing data also provides the best 3D reconstruction. The key point of our proposed approach is to consider the x and y coordinates of the 2D trajectories (the columns of the matrix of trajectories W) as 1D signals and to study their *frequency content*, which is assumed to be similar after filling the missing entries in W. The missing entries are filled with a NRSfM factorization technique, for different values of K. Then, a measure of goodness of the filled-in data based on the frequency content preservation is defined. The *modulus* of the Discrete Cosine Transform (DCT) is used to study the frequency content of the signals.

This paper is closely related to the work of Julià et al. ([7], [8]), where the goal was to estimate the rank of a missing data trajectory matrix in the case of multiple moving rigid objects using the FFT to describe the frequency content of the 2D trajectories. However, in this paper we focus on the more challenging case of non-rigid motion.

2 Deformable Low-Rank Shape Model

We use the low-rank shape model introduced by Bregler et al. [4] in which they describe the time varying 3D shape of a non-rigid object as a linear combination of some basis shapes B_1, B_2,..., B_K. Each basis shape B_K is a $3 \times p$ matrix describing the 3D coordinates of p points. The 3D points deform as a linear combination of the fixed basis set according to time varying coefficients: $S_i = \sum_{d=1}^{K} l_{id} B_d$, where the matrix $S_i = [\mathbf{S}_{i1}, ..., \mathbf{S}_{ip}]$ is the 3D shape of the object at frame i and l_{id} are the configuration weights. Under an orthographic projection model, the p points of S_i are projected onto 2D image points as:

$$W_i = R_i\big(\sum_{d=1}^{K} l_{id} B_d\big) + T_i \tag{1}$$

where R_i contains the first two rows of the full 3D camera rotation matrix and T_i is the camera translation, which can be eliminated by registering the origin of the image coordinate system to the centroid of the object. If we now consider all the frames, f, in the sequence, we can rewrite the linear combination in (1) as:

$$W = \begin{bmatrix} l_{11}R_1 & \dots & l_{1K}R_1 \\ \vdots & \ddots & \vdots \\ l_{f1}R_f & \dots & l_{fK}R_f \end{bmatrix} \begin{bmatrix} B_1 \\ \vdots \\ B_K \end{bmatrix} = \begin{bmatrix} M_1 \\ \vdots \\ M_f \end{bmatrix} \begin{bmatrix} B_1 \\ \vdots \\ B_K \end{bmatrix} = MS \tag{2}$$

Since M is a $2f \times 3K$ matrix and S is a $3K \times p$ matrix, the rank of W is constrained to be at most $3K$ and it can be factorized into two matrices: M contains the camera pose R_i and configurations weights l_{i1},...,l_{iK} for each frame i, while S contains the K basis shapes B_d.

3 Proposed Approach

This section describes our new algorithm to estimate the number of deformation modes K automatically in the NRSfM problem in the case of missing data.

Our goal is to select the value of K that yields the best 3D reconstruction of the deformable object. The missing data in the trajectory matrix W are filled with a NRSfM factorization technique, considering different values for K. Our hypothesis is that the value of K that best predicts the missing data also provides the best 3D reconstruction. Thus, the aim is to define a measure of goodness of the filled-in data, when different values of K are considered. One could consider using the *root mean square error* (rms):

$$rms = \|W - MS\|_F / \sqrt{n} \tag{3}$$

where $\|\cdot\|_F$ is the Frobenius norm and n is the number of known elements in W. However the rms error would only give information about how well the initially known entries are approximated. Alternatively, our proposed approach aims to

define a measure of goodness of fitting for both the initially known data and the missing data. In this paper we propose to consider the x and y coordinates of each 2D trajectory (each column of W) as 1D signals and we define a measure of goodness based on the frequency content of these signals.

3.1 Frequency Content of 2D Trajectories

The proposed measure of goodness is based on the assumption that the frequency content of the signals (the 2D trajectories) should be preserved after filling-in the missing entries in W. The Discrete Cosine Transform (DCT) is used to study the frequency content of those 1D signals. More specifically, the *modulus* of the DCT coefficients, which encodes the amount of information (energy) of the signal contained at a given frequency, is used.

We propose to compare the energy of the original signals with the one obtained after filling in the missing entries considering different K values. The number of basis shapes K that gives the best frequency content preservation is then selected. Naturally, the problem is that the original signals are not full. This paper proposes a strategy to fill missing entries in the original signals, without assuming any K value, to give a full *reference trajectory matrix* (see details in Section 3.3). The frequency content of this *reference trajectory matrix* will then be compared with the one of the matrix filled-in using different values of K. It is therefore crucial that our *reference trajectory matrix* be a good approximation of the original, unknown, full matrix W.

To illustrate the key idea behind our approach, we consider one of the datasets used in Section 4 and we show the way in which the missing data in a single trajectory is recovered, assuming different values of K. It consists of a sequence of 37 3D points on a face tracked along 74 frames. A percentage of 30% missing data is randomly generated from the full matrix of original trajectories. Fig.1 shows the single trajectory studied in this section, when different values for K are considered. Specifically, the full trajectory (black line), the data filled-in with different values for K (red line) and the *reference* trajectory (blue-dashed line) are plotted. The corresponding x and y coordinates of this trajectory are plotted in Fig. 2 (a) and (b), respectively. In addition, the modulus of the DCT of each of the x and y signals for each K considered are plotted in Fig. 2 (c) and (d) respectively. The modulus of the DCT of the full signal (black line), the filled-in signal for different values of K (red line) and the *reference* signal (blue-dashed line) are shown. It can be seen that the *reference* matrix is a good approximation of the full data both in terms of the 2D image coordinates (see Fig. 2 (a) and (b)) and of the frequency content (see Fig. 2 (c) and (d)). Notice that most of the energy of the signal is contained in the lowest frequencies (left part of the frequency content plot). The proposed approach takes into account the frequencies containing about 99.99% of the energy of the signal. Fig. 2 (e) and (f) shows only the lowest frequencies corresponding to the current example.

Fig. 1 and Fig. 2 (a) and (b) show that the data are better filled when $K \geq 6$. Furthermore, the filled-in signal is most similar to the original one when $K \geq 6$

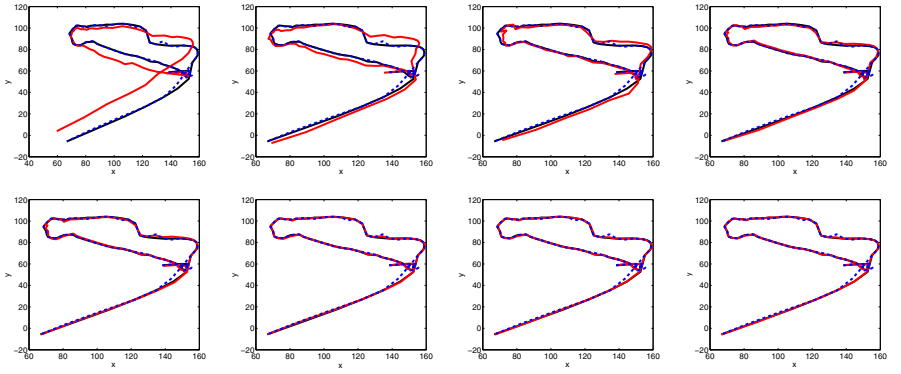


Fig. 1. Single trajectory plotted in the image plane: full data (black line), filled-in data with values of K ranging from 1 to 8 (red line) and *reference* data (blue-dashed line)

(see Fig. 2 (e) and (f)). In fact, the best 3D reconstruction is obtained for $K = 6$ in this example, as we will show later in Fig. 3 (b).

3.2 Algorithm

The proposed algorithm is based on the preservation of the frequency content (or energy) of the signals. First the *reference* matrix is built from the visible tracking data in W following the method described in Section 3.3. The missing data in W are then filled with a NRSfM factorization technique, considering different values of K. The modulus of the DCT of each filled-in matrix (referred to as DCT_K) is compared with the one given by the *reference trajectory* matrix (denoted as DCT_{ref}). In fact, only a small number of the lowest DCT frequencies, which contain about 99% of the energy of the signal, are considered. The proposed measure of goodness of fit compares the signals corresponding to the x and y coordinates separately and is defined as follows:

$$e_{DCT}(K) = e_x(K) + e_y(K) \tag{4}$$

where

$$e_x(K) = \|DCT_{ref|x} - DCT_{K|x}\|_F/\sqrt{l}, \tag{5}$$

$$e_y(K) = \|DCT_{ref|y} - DCT_{K|y}\|_F/\sqrt{l}, \tag{6}$$

l is the length of the signal and, $DCT_{ref|x}$ and $DCT_{ref|y}$ are the modulus of the DCT of the x and y coordinates of the *reference* matrix. At the same time, $DCT_{K|x}$ and $DCT_{K|y}$ are the modulus of the DCT of the x and y coordinates of the matrix filled for different values of K. Notice that the DCT is applied to the x and y coordinates of each individual trajectory (each column of the matrix). Therefore, l is the number of the lowest frequencies taken into account. We propose to stop increasing K when either e_{DCT} increases or when it decreases below a given threshold.

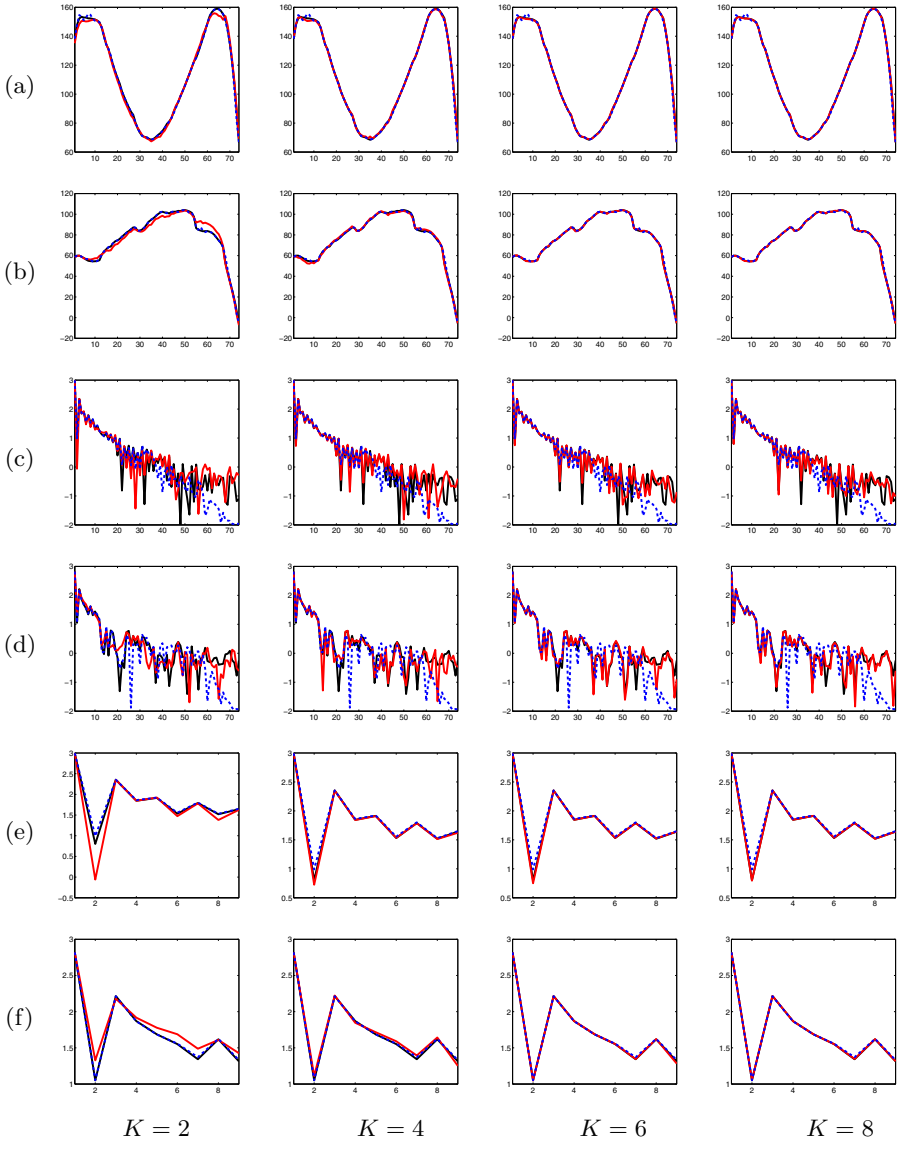


Fig. 2. (a) and (b): x and y coordinates of the studied trajectory; (c) and (d): modulus of the DCT (logarithmic scale) of the above signals; (e) and (f): modulus of the lowest frequencies of the DCT. Legend: full data (black line), filled-in data for different values of K (red line) and *reference* trajectories (blue-dashed line).

The proposed algorithm is summarized below:

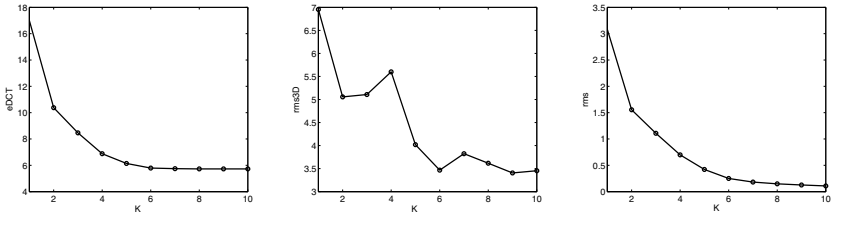
Algorithm: Automatic estimation of the number of deformation modes in NRSfM with missing data

Input: A matrix of trajectories $W_{2f\times p}$ with missing data, where f is the number of frames and p the number of feature points.

1. Missing data entries of the original W are filled to obtain the *reference* matrix W_{ref} using the algorithm in Section 3.3.
2. Set $K = 1$ and the threshold τ. Compute $DCT_{ref|x}$ and $DCT_{ref|y}$ (modulus of the lowest frequencies of the DCT of W_{ref}).
3. Factorize the missing data matrix W into the structure and motion matrices using a non-rigid factorization technique: $\widetilde{W}_K = M_{2f\times 3K} S_{3K\times p}$.
4. Compute $DCT_{K|x}$ and $DCT_{K|y}$ (modulus of the lowest frequencies of the DCT of $\widetilde{W}_K$).
5. Compute the error value: $e_{DCT}(K) = e_x(K) + e_y(K)$,
6. Stop if $e_{DCT}(K) \geq e_{DCT}(K-1)$ or $(e_{DCT}(K-1) - e_{DCT}(K)) \leq \tau$. Otherwise, increase $K = K + 1$ and go back to step 3.

Solution: $K = K - 1$ is the estimated number of deformation modes.

Fig. 3 (a) and (b) show the proposed measure of goodness of fit (e_{DCT}) and the 3D error (rms_{3D}, defined in the next section) obtained for the current example (30% of missing data) for increasing values of K. If the proposed stopping criterion was used with $\tau = 0.09$, the selected number of deformation modes would be $K = 6$, which yields the best 3D reconstruction. Fig. 3 (c) shows the obtained rms. It would be more difficult to define a stopping criterion by studying the rms, since its value does not stabilize. The rms decreases as K increases, in general.



(a) (b) (c)

Fig. 3. (a) e_{DCT}; (b) rms_{3D}; (c) rms, all for different values of K (results correspond to the current example, 30% of missing data)

3.3 The *reference* Matrix

As mentioned above, missing data in the original matrix W should be filled in order to study the frequency content of the original signals. This section presents a strategy to obtain a full matrix that will be used as a *reference trajectory* matrix in the proposed algorithm. The main advantage of the proposed strategy is that it does not assume any K value. However, other *reference* matrices could be considered. In [7] and [8], for instance, the *reference* matrix was obtained by filling the missing entries with zeros. This paper proposes to express the original 1D signals using the DCT basis. That is, given the signal w_x^j (x coordinates of the jth-column of W), we express it as the following product:

$$w_x^j = \mathbf{\Phi}_{f\times d}\mathbf{x}_{d\times 1} \tag{7}$$

where $\mathbf{\Phi}$ contains a predefined set of d DCT basis vectors and $\mathbf{x}$ are the unknown coefficients. Specifically, each element in the matrix $\mathbf{\Phi}$ is the jth-frequency cosine term at time i:

$$\phi_{ij} = \frac{\sigma j}{\sqrt{f}} \cos\left(\frac{\pi(2i-1)(j-1)}{2f}\right) \tag{8}$$

with $\sigma_1 = 1$ and $\sigma_j = \sqrt{2}$, for $j \geq 2$, and f is the number of frames.

The following expression gives the solution to find the coefficients $\mathbf{x}$:

$$\mathbf{x} = (\mathbf{\Phi}^t\mathbf{\Phi})^{-1}(\mathbf{\Phi}^t w_x^j) \tag{9}$$

Once the coefficients $\mathbf{x}$ have been computed, the missing entries in w_x^j are filled with the product (7). It should be remarked that only the known entries in w_x^j and the corresponding rows in $\mathbf{\Phi}$ are used to compute $\mathbf{x}$. Due to that fact, the matrix $\mathbf{\Phi}$ may be close to singular when working with missing data and equation (9) may give incorrect results. In order to avoid this situation, we propose an incremental strategy to compute the DCT coefficients. In a first step, a small number of coefficients of the DCT basis are computed, by using only the initially known data in W. Then, missing entries in W are filled with the product (7). The following steps consist in computing a larger number of coefficients of the DCT basis by using the data filled in the previous step. This is repeated until the maximum number of coefficients is achieved (the number of frames). Therefore, we only work with missing data in the first step, where the number of DCT coefficients (that is, the number of unknowns in equation (9)) is very low.

4 Experiments

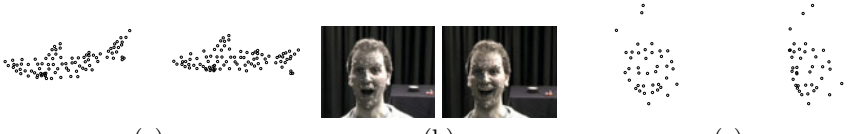
The goal of this section is to show that the proposed algorithm estimates the number of modes of deformation K that yields the best 3D reconstruction of the deformable object or a very close one. The algorithm is tested for different percentages of missing data in the initial matrix W—from 10% up to 40%. Missing entries are randomly generated, as in most of the works that deal with

missing data (e.g., [12,2,9]). In order to obtain more robust results, 50 runs are carried out for each hypothesis. Although any NRSfM method can be used in step 3 of the proposed algorithm, in this paper we chose the EM algorithm proposed by Torresani et al. [12]. The quality of the 3D reconstruction (S) is measured, when the ground truth (S_{GT}) is available, by computing the rms_{3D} error:

$$rms_{3D} = \frac{\|S_{GT} - S\|_F}{\|S_{GT}\|_F} = \frac{\sqrt{\sum_{i,j} |(S_{GT})_{ij} - S_{ij}|^2}}{\|S_{GT}\|_F} \tag{10}$$

4.1 Synthetic Data

In our experiments with synthetic data, the 3D animation of a shark data-set, *Shark*, also tested in [12], is used. It consists of 91 3D points tracked along 240 frames and the orthographic projection is obtained by discarding the third coordinate of each 3D point. The object undergoes rigid motion and deformation corresponding to 2 basis shapes. Thus, $K = 3$ in our formulation. Two frames of the sequence are shown in Fig. 4 (a).



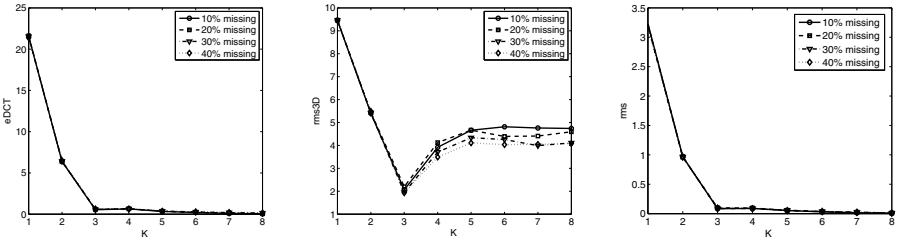
(a) (b) (c)

Fig. 4. Sequences: (a) *Shark* data-set: frames 1 and 50; (b) *Face1* data-set: frames 45 and 70. (c) *CMU face* data-set: frames 1 and 62.

Fig. 5 (a) and (b) shows the e_{DCT} and the rms_{3D} for different values of K and different percentages of missing data. In addition, the rms is plotted in Fig. 5 (c). In this synthetic example, the rms plot is similar to the one obtained with the proposed measure of goodness of fit. These plots correspond to a single run. Fig. 5 (a) shows that $e_{DCT}(4) > e_{DCT}(3)$, for any percentage of missing data. Therefore, the algorithm would stop at $K = 3$. The rms_{3D}, on the other hand, takes its minimum value at $K = 3$ in all the cases. The estimated value of K for 50 runs of the algorithm with different percentages of missing data are plotted in Fig. 6 (a). The threshold τ that defines the stopping criterion for K is empirically set to 0.09 in both synthetic and real data experiments and for all percentages of missing data. The number of deformation modes is correctly estimated ($K = 3$) for every percentage of missing data. Only a few outliers are obtained in the cases of 30% and 40% of missing data.

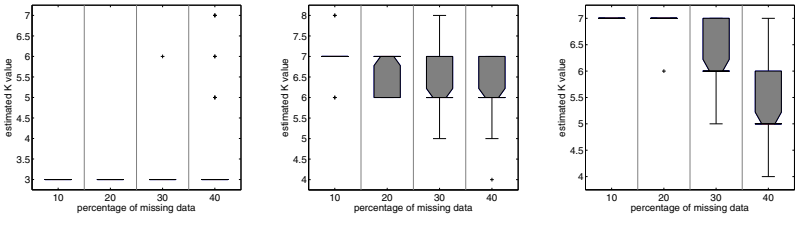
4.2 Real Data

Two different data-sets are used in the experiments with real data. The first data-set, *Face1*, is a motion capture sequence with 3D ground truth that is



(a) (b) (c)

Fig. 5. *Shark* data-set: (a) e_{DCT}; (b) rms_{3D}; (b) rms, all for different K values and different percentages of missing data (single run)

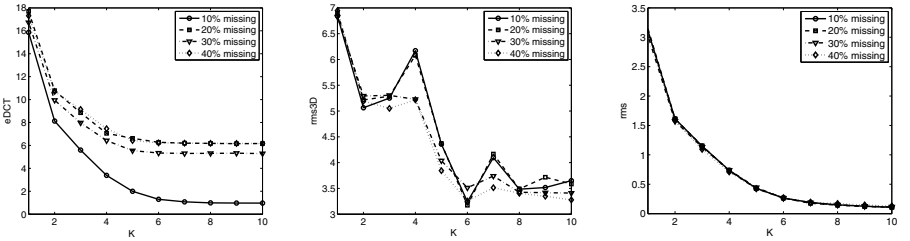


(a) (b) (c)

Fig. 6. Estimated K values for different percentages of missing data (50 runs): (a) *Shark* data-set; (b) *Face1* data-set; (c) *CMU face* data-set

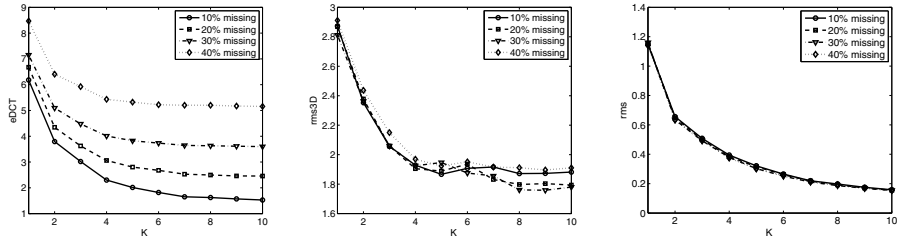
also tested in [9] and has been used as an example in Section 3. It consists of 37 3D points on a face tracked with a motion capture system and projected synthetically onto a 74 frame long sequence using an orthographic camera model. Two frames of the sequence are shown in Fig. 4 (b). The second data-set, *CMU face*, also tested in [12], consists of 40 points tracked by a motion capture system along 316 frames. Data is obtained by orthographic projection. Fig. 4 (c) shows two frames of the sequence.

Fig. 7 (a) shows the e_{DCT} obtained with the *Face1* data-set for increasing values of K and different percentages of missing data (values are given for a single run of the algorithm). The e_{DCT} stabilizes at $K = 7$ for percentages of missing data below 20%. For percentages of missing data equal or higher than 20%, its value stabilizes at $K = 6$. Fig. 7 (b) shows that the rms_{3D} takes it minimum value at $K = 6$ for any percentage of missing data. It can be seen that the rms decreases for increasing values of K (see Fig. 7 (c)). Fig. 6 (b) shows the estimated K values with the defined threshold ($\tau = 0.09$) at the 50 runs and for different percentages of missing data. The median (horizontal line in the thinner region) of the estimated K is 7, for percentages of missing data below 30%. For percentages of missing data equal or higher than 30%, the median of the estimated K is 6. Therefore, the estimated K is equal or very close to the one



(a) (b) (c)

Fig. 7. *Face1* data-set: (a) e_{DCT}; (b) rms_{3D}; (c) rms, all for different K values and different percentages of missing data (single run)



(a) (b) (c)

Fig. 8. *CMU face* data-set: (a) e_{DCT}; (b) rms_{3D}; (c) rms, all for different K values and different percentages of missing data (single run)

that gives the smallest 3D error reconstruction ($K = 6$) for any percentage of missing data (see Fig. 7 (b)).

Fig. 8 shows the results corresponding to the *CMU face* data-set. Fig. 8 (a) shows that the e_{DCT} stabilizes at $K = 7$ for percentages of missing data below 30%. For percentages of missing data equal or higher than 30%, the e_{DCT} stabilizes at $K = 6$. Notice in Fig. 8 (c) that the rms decreases as K increases. Fig. 6 (c) shows the estimated K values at the 50 runs, for different percentages of missing data. The median of the estimated K value is 7 for percentages of missing data below 30%. For percentages of missing data of 30% and 40% the median of the estimated K is 6 and 5, respectively. The estimated K for each percentage of missing data corresponds to the one that gives the smallest 3D reconstruction error, or an error very close to that (see Fig. 8 (b)).

5 Conclusions

This paper proposes an algorithm to estimate the number of deformation modes of a non-rigid shape K in the case of missing data entries in the matrix of trajectories W. The missing data are filled with a NRSfM algorithm for different values of K. The modulus of the Discrete Cosine Transform (DCT) is used to

compare the frequency content of the original trajectories with each filled matrix. Experimental results show that the estimated value of K gives, in general, the best 3D reconstruction or, at least, a 3D error very close to the best one.

References

1. Akhter, I., Sheikh, Y., Khan, S., Kanade, T.: Nonrigid structure from motion in trajectory space. In: Neural Information Processing Systems (2008)
2. Bartoli, A., Gay-Bellile, V., Castellani, U., Peyras, J., Olsen, S., Sayd, P.: Coarse-to-fine low-rank structure-from-motion. In: CVPR (2008)
3. Brand, M.: A direct method for 3D factorization of nonrigid motion observed in 2D. In: CVPR, pp. 122–128 (2005)
4. Bregler, C., Hertzmann, A., Biermann, H.: Recovering non-rigid 3D shape from image streams. In: CVPR. pp. 690–696 (2000)
5. Del Bue, A., Lladó, X., Agapito, L.: Non-rigid metric shape and motion recovery from uncalibrated images using priors. In: CVPR, pp. 297–310 (2006)
6. Hartley, R., Vidal, R.: Perspective nonrigid shape and motion recovery. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part I. LNCS, vol. 5302. Springer, Heidelberg (2008)
7. Julià, C., Sappa, A.D., Lumbreras, F., Serrat, J., López, A.: Rank estimation in 3D multibody motion segmentation. Electronics Letters 44(4) (2008)
8. Julià, C., Sappa, A.D., Lumbreras, F., Serrat, J., López, A.: Rank estimation in missing data problems. Journal of Mathematical Imaging and Vision 39, 140–160 (2010)
9. Paladini, M., Del Bue, A., Stošic, M., Dodig, M., Xavier, J., Agapito, L.: Factorization for non-rigid and articulated structure using metric projections. In: CVPR, pp. 2898–2905 (2009)
10. Roy-Chowdhury, A.K.: Towards a measure of deformability of shape sequences. Pattern Recognition Letters 28, 2164–2172 (2007)
11. Tomasi, C., Kanade, T.: Shape and motion from image streams under orthography: a factorization method. IJCV 9(2), 137–154 (1992)
12. Torresani, L., Hertzmann, A., Bregler, C.: Non-rigid structure-from-motion: Estimating shape and motion with hierarchical priors. IEEE Transactions on PAMI 30(5), 878–892 (2008)
13. Vidal, R., Abretske, D.: Nonrigid shape and motion from multiple perspective views. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3952, pp. 205–218. Springer, Heidelberg (2006)
14. Xiao, J., Chai, J., Kanade, T.: A closed-form solution to non-rigid shape and motion recovery. In: IJCV, vol. 67(2) (2006)
15. Xiao, J., Kanade, T.: Uncalibrated perspective reconstruction of deformable structures. In: ICCV (2005)