A Blind Source Separation Approach to Structure From Motion

Jeff Fortuna and Aleix M. Martinez
Department of Electrical & Computer Engineering
The Ohio State University
Columbus, Ohio, 43210

Abstract

We present an alternate approach to the problem of structure from motion (SfM) with noisy point measurements. With no information available about the joint density of three-dimensional points, the assumption of independence is the only reasonable one. With this assumption alone, the process of the factorization of the observed projections of inaccurately measured 3 dimensional points into motion and shape matrices is a blind source separation (demixing) problem in the presence of noise. This approach is very general, allowing the extension of all previous work on source separation to be applied to SfM in an information theory context. A significant reduction in the error of the estimation of the motion and shape matrices over other methods is possible.

1. Introduction

Structure from motion (SfM) is a common approach to recover the 3D shape of an object observed over multiple frames of an image sequence. Given a set of 2D point correspondences across the frames, it is possible to recover both the 3D positions of the points and the underlying motion between frames [3]. A major concern with this approach is the ability to accurately co-locate the same point in all frames [4], since obtaining mere pixel-accurate correspondences (let alone sub-pixel accurate) is intractable. Further, some objects such as faces do not have a large number of hard boundaries or corners at which to locate points. Additionally, as the objects move, illumination direction can change, causing moving shadows, further confounding the correspondence problem. It is generally accepted, then, that some amount of noise in the correspondence must be accounted for in the shape and motion recovery process.

In the noise-free case, there have been a number of approaches to SfM [9], [16]. Of particular note is the method of Tomasi and Kanade [13], which uses the SVD to factor the measured point matrix into rank 3 motion and structure matrices. In the absence of noise, the SVD provides an exact solution, up to a scaling and a rotation. A similar solution which resolves the rotation can be obtained with ICA, up to a scaling and permutation.

In the presence of noise, the SVD approach provides a minimum mean-square error solution. Unfortunately, the noisy points most influence the minimization of the mean of the squared errors. This situation can only be alleviated with prior knowledge of the nature of the shape, or the noise, or both. Further, in the absence of prior shape knowledge, the only option is to model the noise. Previous attempts to include noise in the problem definition include a maximum likelihood (ML) method [12] (the maximum likelihood method is an information theoretic implementation of the "bundle adjustment" paradigm [14] for zero-mean Gaussian noise) and a covariance weighted SVD [8]. *However, the lack of shape knowledge justifies the assumption that the 3D point measurements in each dimension are independent*. In fact, the assumption of independence allows an extension: by deciding whether the data is best described as super-Gaussian or sub-Gaussian, which can be determined from the noisy data, the use of maximum a posteriori (MAP) estimation is possible. This extension is by definition a blind source separation (BSS) problem in the presence of noise, where the 2D point measurements are considered as random observation vectors, the motion as a mixing process and the shape as an unobserved random vector of sources. Priors have been considered on the motion and shape matrices previously [10] utilizing any available "prior information," whereas we assume we have none available.

In the BSS area, considerable attention has been given to noise [6], [2]. Often the moniker "independent component analysis" (ICA) is used interchangeably with BSS. These methods use the assumption of independence to motivate the existence of a solution. This assumption is entirely consistent with the fundamental nature of SfM as previously stated: the position of points in each of the 3 dimensions of an object can be considered to be independent. In the application of BSS to SfM with F frames of observed 2D point

data, each estimate of the mixing matrix and the sources is overdetermined, as will be seen in the next section, since there are $2F$ mixtures and 3 sources.

We will derive a MAP estimation technique for SfM consistent with BSS using only the assumption of shape independence. The resulting approach will be shown to be superior to existing techniques and provides a general framework within which advancements to SfM can be made.

2. Determining structure and motion with BSS

2.1. SfM in the presence of noise

A sequence of F frames f with P annotated points p can be represented in matrix form as

$$\mathbf{W} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_F \end{bmatrix} \quad \text{where} \quad \mathbf{X}_f = \begin{bmatrix} \mathbf{x}_f \\ \mathbf{y}_f \end{bmatrix}, \tag{1}$$

with each $\mathbf{X}_f \in \mathbb{R}^{2\times P}$ composed of the centered x_p and y_p co-ordinates of the points in the frame:

$$\begin{aligned} \mathbf{x}_f &= \begin{bmatrix} x_1 - \mu_{x\,f} & x_2 - \mu_{x\,f} & \cdots & x_P - \mu_{x\,f} \end{bmatrix} \\ \mathbf{y}_f &= \begin{bmatrix} y_1 - \mu_{y\,f} & y_2 - \mu_{y\,f} & \cdots & y_P - \mu_{y\,f} \end{bmatrix}, \end{aligned} \tag{2}$$

where $\mu_{x\,f}$ and $\mu_{y\,f}$ are the means of the x and y co-ordinates respectively for the f^{th} frame. In the structure from motion problem, this set of points is assumed to have resulted from the projection of a set of 3D points which undergo an affine transformation between frames:

$$\mathbf{W} = \mathbf{MS}, \tag{3}$$

where $\mathbf{S} \in \mathbb{R}^{3\times P}$ and $\mathbf{M}$ is composed of 2×3 orthogonal affine transformation matrices $\mathbf{R}_f$ for each frame:

$$\mathbf{M} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \vdots \\ \mathbf{R}_f \end{bmatrix}. \tag{4}$$

When the points in the 2D images are inaccurately measured, the SfM model becomes

$$\mathbf{W} = \mathbf{MS} + \mathbf{N}, \tag{5}$$

where $\mathbf{N} \in \mathbb{R}^{2F\times P}$ is a matrix of the noise in measurement at each point for each frame. In what follows, we will assume that the motion matrix is composed of orthogonal rows, thus employing a weak-perspective model.

2.2. SVD solution

In the noise-free case, the motion and shape matrices can be recovered exactly (up to an arbitrary orthogonal transformation) from the SVD of $\mathbf{W}$ for all frames by

$$\mathbf{W} = \mathbf{UDV}^T. \tag{6}$$

As such, $\mathbf{M}$ and $\mathbf{S}$ can be found by the rank 3 approximation of $\mathbf{W}$ – the first 3 columns of $\mathbf{UD}^{\frac{1}{2}}$ and 3 rows of $\mathbf{D}^{\frac{1}{2}}\mathbf{V}^T$. Each $\mathbf{R}_f$ in $\mathbf{M}$ must have an orthogonal row space, so some constraints must be applied simultaneously for all frames to find a corrective transformation matrix $\mathbf{G} \in \mathbb{R}^{3\times 3}$, to be applied to $\mathbf{R}_f$ for all frames. This transform can be found from a least-squares fit of an overdetermined set of linear equations. Written concisely from [1]:

$$\begin{bmatrix} \text{vc}(\mathbf{x}_f, \mathbf{y}_f) \\ \text{vc}(\mathbf{x}_f - \mathbf{y}_f, \mathbf{x}_f + \mathbf{y}_f) \end{bmatrix}^T \text{vech}(\mathbf{GG}^T) = \mathbf{0}, \tag{7}$$

where $\text{vech}(\mathbf{A})$ is a vector representation of the lower triangular elements of a symmetric matrix $\mathbf{A}$ and $\text{vc}(\mathbf{x}, \mathbf{y}) = \text{vech}(\mathbf{xy}^T + \mathbf{yx}^T - \text{diag}(\mathbf{x} \circ \mathbf{y}))$. Element-wise product is denoted by $\circ$. This defines the SfM method in [13].

2.3. SfM and BSS

The structure from motion problem with noisy point measurements can be described as a "blind source separation" problem in the presence of noise. The factorization of a set of 2D points measured from multiple frames of an image sequence into an affine transformation from 3D to 2D and a set of 3D points for each frame f of F frames is then

$$\mathbf{w}_f = \mathbf{R}_f\mathbf{s} + \mathbf{n}_f. \tag{8}$$

In this representation, $\mathbf{w}_f$ is a two dimensional random vector of observed points, $\mathbf{R}_f \in \mathbb{R}^{2\times 3}$ is an affine transformation matrix, $\mathbf{s} = \begin{bmatrix} s_1, & s_2, & s_3 \end{bmatrix}^T$ is a three dimensional random vector of 3D points and $\mathbf{n}_f$ is a random vector representing the noise on each measurement for each frame. Over all observed frames, this blind separation problem needs to recover the 3 sources $\mathbf{s}$ from $2F$ mixtures $\mathbf{w}$, resulting in a $2F \times 3$ mixing matrix $\mathbf{M}$, using only the assumption that the sources are independent:

$$P(\mathbf{s}) = \prod_{i=1}^{3} P_i(s_i). \tag{9}$$

2.4. Maximum likelihood solution

When the noise in (8) is Gaussian, a simple expression for the pdf of the observed points results:

$$P(\mathbf{w}_f|\mathbf{R}_f, \mathbf{s}) \propto C \exp\left(-\frac{1}{2}(\mathbf{w}_f - \mathbf{R}_f\mathbf{s})^T\mathbf{\Sigma}(\mathbf{w}_f - \mathbf{R}_f\mathbf{s})\right) \tag{10}$$

where $C = \frac{1}{2\pi|\mathbf{\Sigma}|^{P/2}}$ and $\mathbf{\Sigma}$ is the noise covariance.

Thus a maximum likelihood solution can be found from maximizing

$$\log L(\mathbf{M}, \mathbf{s}) = -\sum_{f=1}^{F}\sum_{p=1}^{P}\left[\|\mathbf{R}_f\mathbf{s}(p) - \mathbf{w}_f(p)\|^2_{\mathbf{\Sigma}^{-1}}\right]. \quad (11)$$

The ML solution was found bilinearly in [12] with partial derivatives $\frac{\partial L}{\partial \mathbf{m}_f}$, $\frac{\partial L}{\partial \mathbf{k}_f}$ and $\frac{\partial L}{\partial \mathbf{s}}$ where

$$\mathbf{R}_f = \begin{bmatrix}\mathbf{m}_f \\ \mathbf{k}_f\end{bmatrix}, \quad (12)$$

with $\mathbf{m}_f$ and $\mathbf{k}_f \in \mathbb{R}^{1\times 3}$.

2.5. MAP solution

Maximum a posteriori estimation, as it applies to SfM can be used to maximize the joint probability of the posteriori density of the shape and the motion:

$$P(\mathbf{M}, \mathbf{s}|\mathbf{w}) \propto P(\mathbf{w}|\mathbf{M}, \mathbf{s})P(\mathbf{M}, \mathbf{s}). \quad (13)$$

Under the assumptions that the shape is independent of the motion and that we have no prior knowledge of the motion, a solution for $\mathbf{M}$ and $\mathbf{s}$ can be found with:

$$arg\max_{\mathbf{M},\mathbf{s}} P(\mathbf{w}|\mathbf{M}, \mathbf{s})P(\mathbf{s}). \quad (14)$$

Priors on the mixing matrix could be included by considering $P(\mathbf{M})$ as non-uniform. The log of the MAP estimate provides an objective function to maximize:

$$\begin{aligned}\log L(\mathbf{M}, \mathbf{s}) \quad = \quad & -\sum_{f=1}^{F}\sum_{p=1}^{P}\Bigg[\|\mathbf{R}_f\mathbf{s}(p) - \mathbf{w}_f(p)\|^2_{\mathbf{\Sigma}^{-1}} \\ & + \sum_{i=1}^{3} g_i(s_i(p))\Bigg], \end{aligned} \quad (15)$$

where $g_i(.) = -\log P_i(.)$ are the individual densities of unit variance sources.

For SfM, this would appear to be intractable, since we do not know the densities of the sources a priori. As mentioned previously, it is, in fact, not necessary to know the source densities exactly. The key point is that it is simply necessary to know in which half-space of densities the sources lie in (sub-Gaussian or super-Gaussian). Intuitively, this can be accomplished by simply considering the kurtosis of each s_i during the current iteration of the MAP estimation process. A kurtosis > 3 would be considered super-Gaussian while kurtosis < 3 is sub-Gaussian. A stable criterion [11] can be found from the sign of

$$E\{\text{sech}^2(s_i)\}E\{{s_i}^2\} - E\{[\tanh(s_i)]s_i\} \quad (16)$$

over all points in $\mathbf{s}$. A complete discussion of the techniques of BSS can be found in [5]. To recapitulate, it is only necessary to determine whether the sources are sub or super-Gaussian *and* this can be reliably accomplished with the current estimate of the sources. Hence, blind source separation provides the ability to recover motion and shape matrices without any knowledge of the motion or shape.

A number of techniques exist to maximize the objective function in (15). Some iterative techniques resulting from direct differentiation can be found in [6], provided that the source distribution approximation is differentiable. Both of these utilize an alternating variables method. In contrast to BSS problems, the number of points is often far less than the number of samples in a signal. As a result, general non-linear optimizers can be applied directly, as was done herein.

We will describe two shape densities as examples. These densities are illustrative only, but will provide analytic properties useful for demonstration of our idea. Consider i sources which are distributed according to

$$h_i(y_i) = \frac{1}{2\sigma_i}\exp\left(-\frac{|y_i|}{\sigma_i}\right). \quad (17)$$

This defines a Laplacian distribution with variance ${\sigma_i}^2$, appropriate for the description of points which lie around a point or on a line. If the shape densities were known a priori and scaled to unit variance, the MAP objective function for SfM would include the log sum of these sources, or

$$\sum_{i,p}|s_i(p)|, \quad (18)$$

where the absolute value function is evaluated for each element of $s_i(p)$.

A second, and more flexible example of source distributions is a mixture of Gaussians (MoG). For this case, the source pdf of N mixtures n is described by

$$h_i(y_i) = \sum_{n=1}^{N}\frac{1}{2\pi\sigma_n}\exp\left(-\frac{(y_i-\mu_n)^2}{{\sigma_n}^2}\right), \quad (19)$$

where σ_n and μ_n are the standard deviation and mean respectively of the nth mixture. The MAP objective function would include

$$\sum_{i,p,n}\log h_i(s_i(p)). \quad (20)$$

The mixture of Gaussian model has general application in modeling arbitrary shape distributions.

Of course, these shape densities are not generally known ahead of time. As a result, we use a generalized Gaussian density to approximate either super-Gaussian (for example, Laplacian) or sub-Gaussian (for example, mixtures of Gaussians) densities. The generalized Gaussian density for i unit

variance sources is defined by [5]

$$h_\alpha(y_i) = \frac{\alpha_i}{2\Gamma(1/\alpha_i)} \exp\left(-\frac{|y_i|}{r}\right)^{\alpha_i}, \qquad (21)$$

where Γ is the gamma function. The corresponding description in the MAP estimation is

$$\sum_{i,p} \frac{1}{2\Gamma(1/\alpha_i)} |s_i(p)|^{\alpha_i}, \qquad (22)$$

where α is a positive constant controlling the shape of the density. With $\alpha < 2$, we are describing super-Gaussian densities. For $\alpha = 1$ we have a Laplacian density. For $\alpha > 2$ we have sub-Gaussian densities.

2.6. The orthogonality constraint on A

An additional constraint exists in SFM on the motion matrix – each row-pair of $\mathbf{M}$ corresponding to $\mathbf{R}_f$ must be orthogonal. This can be accomplished in a couple of ways. First, the orthogonality can be enforced at each step in an iterative algorithm. Second, it can be enforced after the computation of $\mathbf{M}$. If the posteriori method is used, the rows are only approximately orthogonal, as determined by a least-squares solution. A significant improvement to the ML algorithm described in [12] can be achieved by orthogonalizing the motion matrix at each step in the algorithm. This was done in the experiment we present in this paper. For MAP estimation, a convenient starting point is the ML estimate of $\mathbf{M}$ and $\mathbf{S}$. As such, $\mathbf{M}$ approximately complies with the orthogonality constraints initially. Although it is possible to include the constraint in the objective function in MAP estimation, it adds significant complexity to the optimization. As such, it is preferable to enforce the constraint a posteriori, and the resulting solution is not significantly different from constraining $\mathbf{M}$ within the optimization.

3. An example of SfM with MAP estimation

As an illustration of the use of BSS techniques for SfM, experiments with points generated according to Laplacian densities, mixtures of Gaussian densities and a face model were conducted. Specifically, the two distributions described in (17) and (19) were used to generate statistically independent 3D sources. In the Laplacian case, the source variances were 1000, 100, and 10 for the x, y and z directions respectively. For the mixture of Gaussian sources, 2 mixtures for each direction were considered. The x direction had means of -100 and 300 with variances 2000 and 1000. The y direction had means of 200 and -200 with variances 200 and 100. The z direction had means of -30 and 90 with variances of 100 and 100.

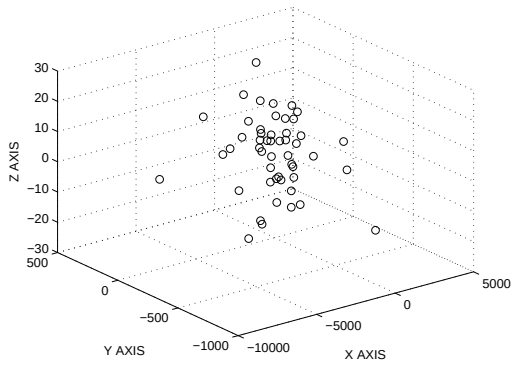

Figure 1. Laplacian Distribution of 3D points

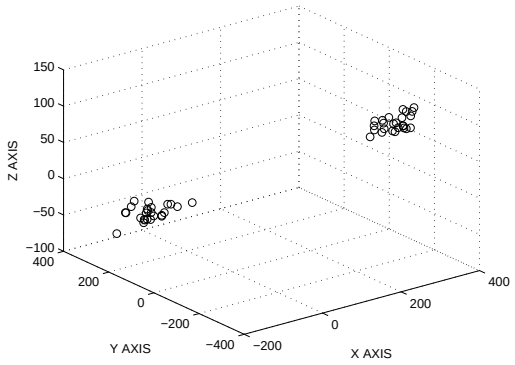

Figure 2. MoG Distribution of 3D points

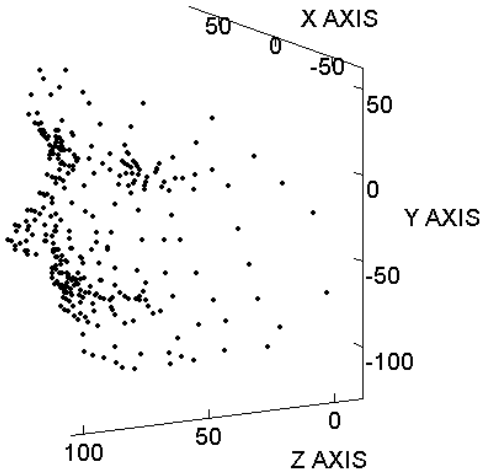

Figure 3. Face model

Prior	Noise Σ		$\begin{bmatrix}1 & 0\\ 0 & 0.1\end{bmatrix}$	$\begin{bmatrix}100 & 0\\ 0 & 10\end{bmatrix}$	$\begin{bmatrix}200 & 0\\ 0 & 20\end{bmatrix}$	$\begin{bmatrix}300 & 0\\ 0 & 30\end{bmatrix}$	$\begin{bmatrix}400 & 0\\ 0 & 40\end{bmatrix}$
Laplacian	Mean Error (motion)	SVD	1.44	14.77	20.25	34.71	39.07
		ML	0.90	9.26	12.19	23.13	26.63
		MAP	0.80	8.57	10.98	20.89	24.71
	Mean Error (shape)	SVD	0.50	4.56	5.97	12.80	13.37
		ML	0.41	4.00	4.86	9.77	11.82
		MAP	0.24	3.27	3.06	6.50	8.41
MoG	Mean Error (motion)	SVD	2.18	22.50	31.88	48.67	52.63
		ML	1.45	14.03	20.20	31.12	36.28
		MAP	1.36	12.88	19.16	31.44	35.61
	Mean Error (shape)	SVD	0.77	8.47	11.15	17.87	18.08
		ML	0.66	6.33	8.04	12.26	14.20
		MAP	0.46	3.92	5.83	9.63	10.91
Super-Gaussian	Mean Error (motion)	SVD	1.47	14.17	21.89	36.35	42.68
		ML	0.92	9.51	14.25	24.26	28.80
		MAP	0.82	8.57	12.74	22.06	25.79
	Mean Error (shape)	SVD	0.51	5.05	7.03	11.94	15.53
		ML	0.44	4.62	6.65	9.77	11.77
		MAP	0.28	3.60	4.26	6.97	8.14
Sub-Gaussian	Mean Error (motion)	SVD	2.03	22.50	31.35	48.75	53.45
		ML	1.34	13.62	18.16	32.60	36.25
		MAP	1.24	12.67	17.24	31.26	35.77
	Mean Error (shape)	SVD	0.81	7.91	8.40	19.60	19.68
		ML	0.61	5.65	8.05	13.99	15.33
		MAP	0.43	4.01	5.50	10.62	12.28

Table 1. Errors in motion and shape matrices – super and sub-Gaussian distributed shapes

Noise Σ		$\begin{bmatrix}1 & 0\\ 0 & 0.1\end{bmatrix}$	$\begin{bmatrix}20 & 0\\ 0 & 2\end{bmatrix}$	$\begin{bmatrix}40 & 0\\ 0 & 4\end{bmatrix}$	$\begin{bmatrix}60 & 0\\ 0 & 6\end{bmatrix}$	$\begin{bmatrix}80 & 0\\ 0 & 8\end{bmatrix}$
Mean Error (motion)	SVD	1.34	5.60	7.62	9.08	10.74
	ML	1.15	4.64	6.29	7.41	8.92
	MAP	1.10	4.41	5.98	6.75	8.07
Mean Error (shape)	SVD	2.91	6.11	8.28	10.02	11.59
	ML	2.88	5.92	8.05	9.67	11.09
	MAP	2.86	5.80	7.93	9.43	10.74

Table 2. Errors in motion and shape matrices – face model data

Figures 1 and 2 provide samples of the point distributions. Figure 3 shows the 3D face model consisting of approximately 350 points. For the Laplacian and MoG densities, a random transformation matrix was generated for each frame. For the face model, a rotation of 90 degrees about the y axis was generated as a geometric transformation (Figure 4). Gaussian noise was then added to the 2D point data resulting from the projection of the 3D points. The resulting 2D points were then centered for the x and y directions.

For the Laplacian and MoG points, the noise variance was parameterized to 1, 100, 200, 300 and 400 for the x direction and .1, 10, 20, 30 and 40 for the y direction, creating 5 independent experiments for each of the two source distributions. For the face data, the noise variance was parameterized to 1, 20, 40, 60 and 80 for the x direction and .1, 2, 4, 6 and 8 for the y direction. The motion and the structure matrix were recovered for each experiment with 3 different techniques – SVD, ML estimation, and MAP estimation. All estimates were referenced to the co-ordinate system of the first frame of the ground-truth.

As is typical in SfM problems, the number of points and frames were kept to a relatively small number – 25 frames

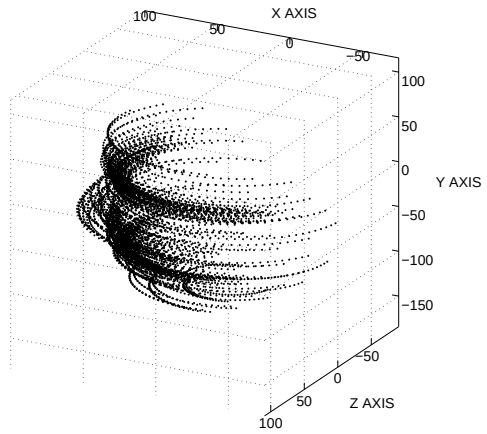

Figure 4. Face points rotated about the y axis

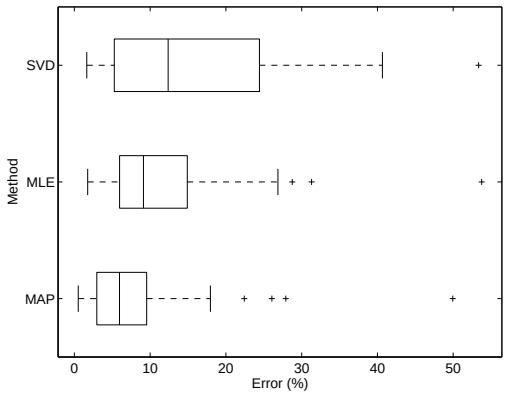

Figure 5. Super-Gaussian shape error $\Sigma = \begin{bmatrix} 400 & 0 \\ 0 & 40 \end{bmatrix}$ (Segments represent data quartiles, points lying outside 1.5 IQR are labeled with +).

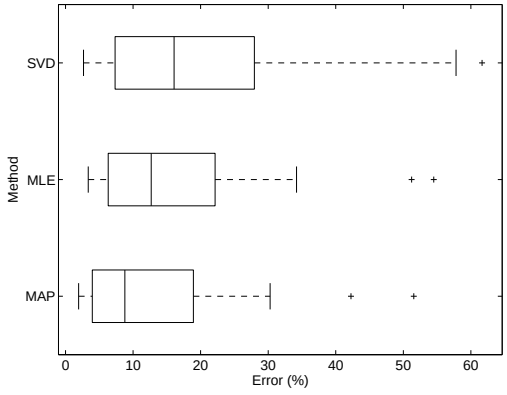

Figure 6. Sub-Gaussian shape error $\Sigma = \begin{bmatrix} 400 & 0 \\ 0 & 40 \end{bmatrix}$

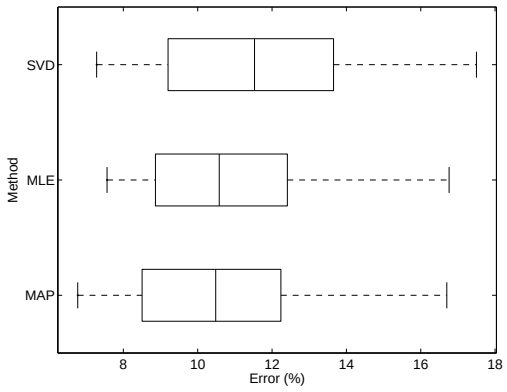

Figure 7. Face model shape error $\Sigma = \begin{bmatrix} 80 & 0 \\ 0 & 8 \end{bmatrix}$

and 50 points (small in this context applies to a comparison with the number of samples from an audio source, for example). For the face model, the 50 points were selected randomly from the totality of points in the model. Typically, BSS is conducted in a context where a great many samples are available for each source. As will be seen from our results, MAP estimation works well despite having few available samples.

Five examples of MAP estimation are given. In Table 1, the first two priors (Laplacian and mixture of Gaussian) illustrate the use of MAP estimation when the source distributions are known a priori. The third and fourth examples illustrates the main idea, recovering the motion and shape matrices of a Laplacian shape density with a super-Gaussian prior and a mixture of Gaussian shape density with a sub-Gaussian prior. The priors were a generalized Gaussian density where (16) was employed during the optimization to determine which form of density to employ. Specifically, α in (21) can be switched between 1 and 3, corresponding to super and sub-Gaussian densities respectively. Table 1 describes the mean errors in the recovered motion and shape matrices over 50 runs. The mean errors are reported as the norm of the error in the estimated matrix expressed as a percentage of the norm of the ground truth matrix. Note that these errors are decidedly non-Gaussian in distribution. Figures 5 - 6 are illustrative of the distributions of the largest measured estimation errors.

The fifth example of the use of MAP estimation is its

application to the face model. Table 2 provides the mean motion and shape estimation error over 50 runs of the randomly selected 50 face model points. Here, as in examples 3 and 4, (16) was used to determine the prior during optimization. Figure 7 provides a boxplot.

4. Discussion

We have shown that a significant improvement in the performance of SfM in the presence of noisy point measurements can be obtained by the use of MAP estimation with no prior knowledge of the sources. No significant change in performance occurs when exact knowledge of the sources is at hand. We have also shown that the method can be applied to real-world shapes, such as a human face. In the face model case, although the improvement in mean square error was not as pronounced as in the Laplacian and MoG cases, it was roughly equivalent to the improvement supplied by the use of ML over the SVD. It was interesting to note that for the face model, the point distributions in x and y were observed to be sub-Gaussian, while in z, the distribution was super-Gaussian. This is a simple statement of the fact that faces have clusters of significant points in x and y which are spread across the face (making a MoG or sub-Gaussian model appropriate), while in z, most points, with the notable exception of the nose, lie near a plane in x and y.

Perhaps more significantly, the general nature of this approach makes it appropriate for addressing a number of issues in the SfM problem. An immediate extension can be be applied in the case of non-Gaussian noise. Real noise resulting from inaccurately measured points is decidedly non-Gaussian. Typically, points are measured along linear features in the 2D projections. As such, their distributions are more appropriately modeled by a Laplacian distribution. This can be easily modeled in the MAP estimator.

Further, priors on the motion can be readily included. In the SfM problem, motion priors are reflected by constraints on the motion. For example, when it is known a priori that the dominant motion includes a scaling in one direction only, a sparse (Laplacian) prior can be applied to the rows of the motion matrix. BSS with incomplete data [15], a common problem in SfM, is also possible.

5. Conclusion

We have presented a novel approach to structure from motion using blind source separation. In doing so, we have shown that MAP estimation in the context of SfM can be conducted without prior knowledge of shape (source) densities. The method works due to an intrinsic property of blind source separation techniques – namely that the original sources can be recovered even with considerable mismatch between the model and the actual source densities. The lack of need for prior knowledge of the shape densities allows application to any shape, such as the human face example shown above. The blind source separation approach provides a significant improvement in shape estimation as well as providing an information theoretic framework for further advances to the study of SfM in the presence of noisy point measurements.

References

[1] M. Brand, "A direct method for 3D factorization of nonrigid motion observed in 2D", *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, vol. 2, pp. 122–128, June 2005, San Diego, CA, USA.

[2] A. Cichocki, S. Douglas and S. Amari, "Robust techniques for independent component analysis (ICA) with noisy data", *Neurocomputing*, vol. 22, pp. 113–129, 1998.

[3] O. Faugeras, *Three-Dimensional Computer Vision*, MIT press, 2001.

[4] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000.

[5] S. Haykin, ed., *Unsupervised Adaptive Filtering vol. 1: Blind Source Separation*, John Wiley & Sons, Inc., 2000.

[6] A. Hyvarinen, "Independent component analysis in the presence of Gaussian noise by maximizing joint likelihood", *Neurocomputing*, vol. 22, pp. 49–67, 1998.

[7] A. Hyvarinen, J. Karhunen and E. Oja, *Independent Component Analysis*, John Wiley & Sons, Inc., 2001.

[8] M. Irani and P. Anandan, "Factorization with uncertainty", *International Journal of Computer Vision*, vol. 49, no. 2, pp. 101–116, 2002.

[9] F. Kahl and A. Heyden, "Affine structure and motion from points, lines, and conics", *International Journal of Computer Vision*, vol. 33 no. 3, pp.163–180, 1999

[10] F. Kahl and A. Heyden, "Auto-Calibration and Euclidean Reconstruction from Continuous Motion", *Proceedings of the International Conference on Computer Vision (ICCV'01)*, pp. 572–577, Vancouver, Canada, 2001.

[11] T Lee, M. Girolami and T. Sejnowski, "Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources", *Neural Computation*, vol. 11, pp. 417–441, 1999.

[12] D. Morris and T. Kanade, "A unified factorization algorithm for points, line segments and planes with uncertainty models", *Proceedings of Sixth IEEE International Conference on Computer Vision (ICCV'98)*, pp. 696–702, Bombay, India, Jan., 1998.

[13] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: a factorization method", *International Journal of Computer Vision*, vol. 9, no. 2, pp. 137–154, 1992.

[14] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon, "Bundle adjustment – A modern synthesis", In *Vision Algorithms: Theory and Practice*, pp. 298–375. Springer Verlag, 2000.

[15] M. Welling and M. Weber. "Independent component analysis of incomplete data", *Proceedings of the 6th joint Symposium on Neural Computation*, May 22, 1999, University of California, San Diego.

[16] Li Zhang, B. Curless, A. Hertzmann and S. Seitz, "Shape and motion under varying illumination: unifying structure from motion, photometric stereo, and multi-view stereo", *Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV)*, pp. 618–625, Nice, France, Oct., 2003.