Addressing the Correspondence Problem: A Markov Chain Monte Carlo Approach

Frank Dellaert

January 2000

CMU-RI-TR-00-11

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

We gratefully acknowledge the support of the following sponsoring institutions: SAIC corporation; KIRIN brewery company; SONY corporation; the Robotics Institute and the Computer Science Department at Carnegie Mellon University. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing official policies or endorsements, either expressed or implied, of the United States Government or any of the sponsoring institutions.

Abstract

Many vision and AI techniques assume that some form of the infamous correspondence problem has been solved. Typically, a best mapping between sets of features is found either as a pre-processing step or as a side-effect of applying the technique. In this paper we argue that it is incorrect to insist on a single 'best' mapping between features in order to estimate a property that depends on this correspondence. Instead, one should take into account the *posterior distribution* over all possible mappings, given the measured feature data. The estimate thus obtained can differ substantially from the one where a best mapping is first singled out. The main contribution in this paper is to show how Markov Chain Monte Carlo methods can be used to efficiently sample the correct distribution over correspondences, and how this sample can subsequently be used to estimate a property of interest. We will show examples and results for several applications, including pose estimation and structure from motion. The method we propose can be used in any application where the correspondence problem is a central component.

1 Introduction

There are many algorithms in computer vision and other disciplines of AI that assume that some form of the ’correspondence problem’ has been solved. Examples from computer vision - which we will focus on - include: pose estimation, object recognition, stereo vision and 3D reconstruction, structure from motion, and motion estimation. Those algorithms that attempt to solve these problems using sparse features measured in substantially separated images (spatially or temporally), typically assume the correspondence between features across different views is known.

If the correspondence is not known, the ’best’ mapping between feature sets is found either as a pre-processing step or as a side-effect of running the algorithm. The algorithms in the ’pre-processing’ class typically construct a cost matrix that encodes the cost of associating two features in different sets. In many cases, the construction of this cost matrix is based on Ullman’s ’minimal mapping’ principles of proximity, similarity and mutual exclusion [22]. In [16, 17, 14] an eigenvalue analysis on this cost matrix leads to a mapping that approximately satisfies mutual exclusion and can be shown to be optimal under certain - rather restrictive - assumptions. In [3, 4] finding the best mapping is formulated as a maximum-weight bipartite matching problem. [18] use simulated annealing to find the best mapping between two point sets. In other algorithms the correspondence problem is more integrated. In [2] the correspondence problem is posed as a hypothesis testing problem. [9] use neural networks in an EM type algorithm that alternates between estimating pose and correspondence.

In this paper we argue that it is incorrect to insist on a single ’best’ mapping between features in order to estimate a property that depends on the correspondence, such as pose. Instead, one should take into account the *posterior distribution* over all possible mappings, given the measured feature data. The estimate thus obtained can differ substantially from the one where a best mapping is first singled out. The main contribution in this paper is to show how Markov Chain Monte Carlo methods can be used to efficiently sample the correct distribution over correspondences, and how this sample can subsequently be used to estimate a property of interest. We will show examples and results for several applications, including pose estimation and structure from motion. However, the methods we advocate can be used in any application where the correspondence problem is a central component.

Unbeknownst to us at the time of writing, MCMC has been used for a similar purpose in [13], where it used in the context of a traffic surveillance applica-

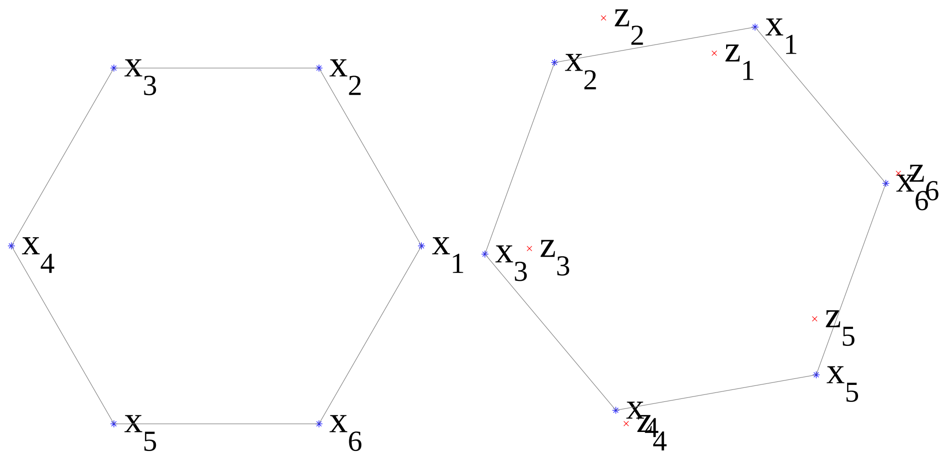

Figure 1: Left: A 2D model shape, defined by points on a circle with radius 1. Right: Measured feature points. The true rotation is 70 degrees, measurement noise $\sigma = 0.2$.

tion. They, like us, propose to use EM in order to perform maximum likelihood inference for quantities that depend on the outcome of an assignment problem. Our formulation is different than theirs, however, in that we assign image features to 3D entities, rather than pairwise associating them between images. Our approaches are very similar in spirit, however.

The paper is laid out as follows: in Section 2 we motivate our main point with a simple example from pose estimation. Section 3 explains how we can use the Metropolis sampler to sample from a joint distribution over correspondence mappings and the parameters to be estimated. If only a MAP or ML estimate is needed in a given application, we can use the EM algorithm to efficiently obtain this, as explained in Section 4. The latter approach is illustrated in the results section, Section 5, with an example from the *structure from motion* domain.

2 Motivation

In order to motivate the need for obtaining a *distribution* over correspondence mappings, we will look at a simple example from pose estimation. Assume we have a 2D model shape, given in the form of a set of 2D points $\{x_i \in \mathbb{R}^2; i \in 1..N\}$, as shown in Figure 1. We observe an image of this shape which has undergone an unknown rotation, and the problem is to recover the unknown parameters θ of this transform. As data we are given noisy measurements $\{z_k \in \mathbb{R}^2; k \in 1..N\}$ on the feature points defined in the model shape, as shown at right in Figure 1.

2.1 Known Correspondence

In the Bayesian probability paradigm, if we are interested in characterizing our complete knowledge about θ, given the data z, we need to obtain the posterior distribution $P(\theta|z)$. By Bayes law, we can express this in terms of a likelihood $P(z|\theta)$ of θ given z, and a prior distribution on θ:

$$P(\theta|z) = C_z P(z|\theta) P(\theta) \tag{1}$$

where $C_z = \frac{1}{P(z)}$ is a constant that only depends on z.

We can easily obtain an expression for the likelihood *if* the correspondence between measurements z_k and model features x_i is known. Assuming i.i.d. normally distributed noise $n_i \in \mathbb{R}^2$ on each measurement z_i, we have for all i

$$z_i = f(x_i, \theta) + n_i$$

where $f : \mathbb{R}^2 \times \mathbb{R}^3 \to \mathbb{R}^2$ is the measurement function that models the 2D transform. The likelihood term $P(z|\theta)$ can then be written as

$$P(z|\theta) = \prod_i N(z_i; f(x_i, \theta), \sigma) \tag{2}$$

where $N(.; \mu, \sigma)$ denotes the radially symmetric 2D normal distribution with mean $\mu \in \mathbb{R}^2$ and standard deviation σ.

It is just as easy to obtain a single *maximum a posteriori* (MAP) or *maximum likelihood* (ML) estimate. Assuming we have no prior information about θ, the ML estimate θ^* can be found by a simple least-squares fit:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \sum_i \|z_i - f(x_i, \theta)\|^2$$

These concepts are illustrated in Figure 2, where the likelihood of the rotation parameter is plotted, given the measurements from Figure 1. In this case, the likelihood function is unimodal and the ML estimate is located at its global maximum.

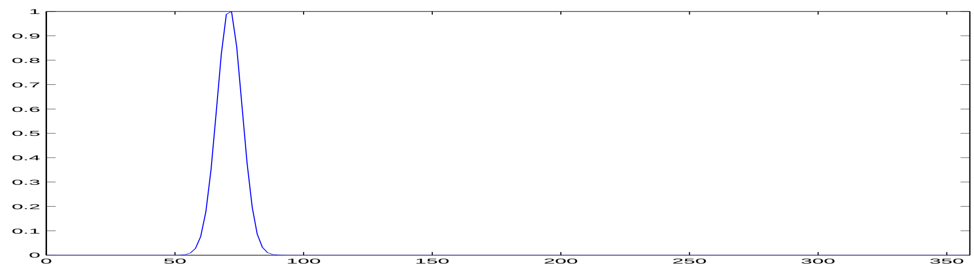

Figure 2: The likelihood function (2) (normalized so maximum=1) of rotation θ for the example in Figure 1, with known correspondences.

2.2 Unknown Correspondence

The difficulty we face in this paper is that the correspondence mapping c that maps measurements z_k to the model features x_i is unknown. *In that case, to obtain $P(\theta|z)$, we need to integrate over all possible mappings:*

$$P(\theta|z) = \sum_c P(\theta, c|z) = C_z P(\theta) \sum_c P(c)P(z|\theta, c) \tag{3}$$

Note that we have used the identity $P(c|\theta) = P(c)$, which says that in the absence of measurement information, the model parameters θ do not provide any information about the mapping c.

We can formalize this by viewing the mapping c as a function $c : \mathbb{Z} \rightarrow \mathbb{Z}$ from measurement indices to model feature indices. For simplicity's sake we will assume for now that there are no occlusions or spurious measurements, so that c is a one-on-one mapping, i.e. a permutation of the indices $1..N$. In that case, each term in (3) can be written as

$$P(c)P(z|\theta, c) = \frac{1}{N!} \prod_i N(z_i; f(x_{c(i)}, \theta), \sigma) \tag{4}$$

when c is a permutation, and 0 otherwise. The factor $N!$ in the denominator arises because there are $N!$ such permutations, and each mapping c is assumed equally likely a priori.

Thus, in the example of Figure 1, the full posterior will be a mixture of 6!=720 functions of the form (4). Note that as σ is small compared to the distance between model features, only 6 of the 720 mixture components will have any appreciable probability mass in them, each associated with one of the six rotations that are consistent with the ambiguous measurement data. This is illustrated in Figure 3.

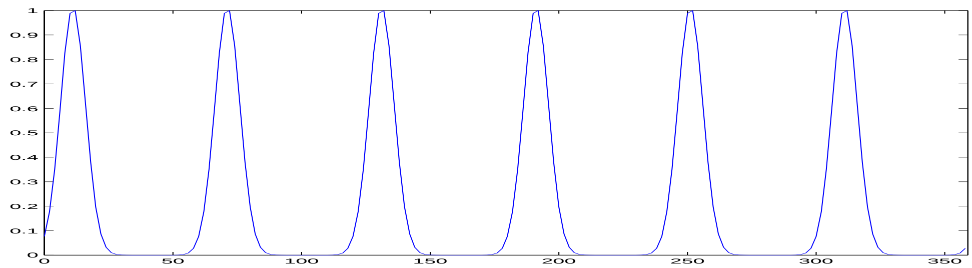

Figure 3: In the case of unknown correspondence, the posterior distribution (normalized so maximum=1) of rotation θ corresponding to the situation in Figure 1 is a mixture distribution with 720 components.

2.3 Summary

As illustrated above, the fact that we do not know the correspondence a priori can drastically alter what knowledge we can derive about the parameters θ. In a sense, the simple example presented above is a worst case scenario, as the model shape whose pose we want to estimate is radially symmetric. In this case there is a strong inherent ambiguity in the rotation part of the pose. However, ambiguity can arise in any situation in which the measurements and the noise are such that the identity of the measurement can be mistaken. When estimating a property of interest, the correct approach is to take this ambiguity into account, and factor it into the estimate, rather than choosing one or the other possibility.

3 Sampling the Posterior

Now that is established that we indeed want to consider the full distribution over correspondence mappings, we need a practical way to do this. Unfortunately, evaluating the total posterior probability (3) can be a challenge. This is because the number of possible mappings c is combinatorial in the number of features to be matched, making a direct computation of $P(\theta|z)$ in (3) intractable in most cases. In some cases, e.g. when all mappings are equally likely and certain independence assumption hold, (3) can be simplified into a manageable number of terms. However, in typical correspondence problems the factor $P(c)$ provides hard to encode knowledge about what type of mappings c are allowable. For example, in the pose estimation case above, only permutations are allowed. Since this is hard to specify analytically, it seems the only way to calculate $P(\theta|z)$ is to explicitly enumerate all allowable mappings.

As a solution to this computational challenge we propose is to instead obtain a sample from the posterior distribution $P(\theta|z)$. Again, since (3) is hard to evaluate, sampling directly from $P(\theta|z)$ is probably intractable. However, if we can obtain a sample $\{(\theta^r, c^r); r \in 1..R\}$ from the *joint* distribution $P(\theta, c|z)$, we can construct a sample from the marginal distribution $P(\theta|z)$ simply by discarding the mapping part c^r from the tuples (θ^r, c^r). This gives us the the sample $\{\theta^r; r \in 1..R\}$ from $P(\theta|z)$. A way to sample from the joint distribution will be given below.

MCMC Sampling

To sample from the joint distribution $P(\theta, c|z)$ we propose to use the Metropolis algorithm, an instance of the Monte-Carlo Markov-Chain methods (abbreviated MCMC), which involve a Markov chain in which a sequence of samples is generated [8, 11]. If we set up the transition probabilities correctly, the equilibrium distribution of the Markov chain will be equal to the posterior distribution we would like to sample from. In our case, we would like to generate a sequence of samples from $P(\theta, c|z)$, and the Metropolis algorithm can be formulated in the current context as follows (adapted from the general description in [11]):

1. Start with a initial tuple (θ^0, c^0).
2. Propose a new tuple (θ^p, c^p), which is probabilistically generated from (θ^r, c^r).
3. Compute the ratio
$$a = \frac{P(\theta^p, c^p|z)}{P(\theta^r, c^r|z)} \tag{5}$$
4. **If** $a >= 1$ then accept (θ^p, c^p) as (θ^{r+1}, c^{r+1}).
   Otherwise, accept (θ^p, c^p) with probability a. If the proposal is rejected, then we keep the previous sample.

The sequence of tuples (θ^r, c^r) thus generated will be a (correlated) sample from the joint distribution $P(\theta, c|z)$. Note that we have assumed here that the *proposal density* used in step 2 is symmetric, which leads to the simple algorithm above. If this density is asymmetric, the more general Metropolis-Hastings algorithm should be used [7].

What is left to do is specify how the proposal step is implemented and how a is to be calculated. In addition, care is required so that no invalid mappings c are proposed. These implementation details are necessarily application dependent, but we will provide an example below in the context of the pose estimation example.

3.1 Detailed Example

In the pose estimation example, we will start the chain with a random value θ^0 for θ and a randomly chosen permutation c^0 of the indices. In the proposal step, we choose randomly between two strategies: (a) in a 'small perturbation step' we keep the mapping c but add a small amount of noise to θ. This serves to explore the continuous values of θ within a mode of the posterior probability. (b) in a 'long jump', we completely randomize both θ and c, as in the initialization phase. This provides a way to jump between probability modes.

After we have proposed the new tuple (θ^p, c^p), we calculate a. In this example, and in many other applications, the noise on the feature measurements is normally distributed and isotropic. By invoking Bayes law and eq. (4), we can rewrite a from (5) as

$$a = \frac{P(\theta^p, c^p|z)}{P(\theta^r, c^r|z)} = \frac{\prod_i N(z_i; f(x_{c^p(i)}, \theta^p), \sigma)}{\prod_i N(z_i; f(x_{c^r(i)}, \theta^r), \sigma)}$$

Here we have assumed that we have uniform prior on θ and all allowable mappings c are equally likely, so that the posterior ratio is equal to the likelihood ratio. We can simplify the notation by defining $f_i^r \triangleq f(x_{c^r(i)}, \theta^r)$. By further manipulation we obtain

$$a = \exp\left[\frac{1}{2\sigma^2}\sum_i(\|z_i - f_i^r\|^2 - \|z_i - f_i^p)\|^2)\right] \tag{6}$$

The result of this procedure is shown in Figure 4, where we have plotted the histogram of the rotation parameter θ as a non-parametric approximation to the analytic posterior shown in Figure 3. The figure shows the results of running a sampler for 100,000 steps, the first 1000 of which were discarded as a transient. We can deduce from the uneven mass in each of the six modes that the 'long jump' was not accepted in many cases, as otherwise the modes would have evened out. Ongoing work attempts to quantitatively evaluate different proposal densities in terms of their efficiency. Note that this will depend on the application considered.

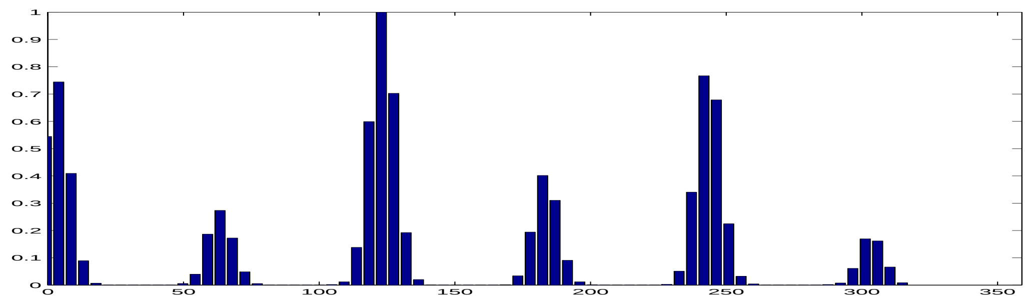

Figure 4: Histogram for the values of θ obtained in one MCMC run, for the situation in Figure 1. The MCMC sampler was run for 100,000 steps, the first 1000 of which were discarded as a transient. This figure should be compared with Figure 3, which is the analytic version of the posterior.

3.2 Flip Proposals

A different and appealing way to propose a new correspondence mapping c is through a local operation such as flipping two assignments. A substantial benefit of this approach is that we can drastically simplify (6), as a local operation will only involve a limited set of features. Whether this strategy leads to efficient sampling depends on the application, but we have used it with success for structure from motion (see below).

In a 'flip proposal', we propose to obtain c^p by switching the mapping c^r on two randomly chosen indices, j and k.Formally, we have $c^p(j) = c^r(k)$ and $c^p(k) = c^r(j)$. Because this proposal only changes c, and not θ, i.e. $\theta^p = \theta^r$, we will have $f_j^p = f_k^r$ and $f_k^p = f_j^r$ after the switch, and furthermore $f_i^p = f_i^r$ for all other indices $i \notin \{j, k\}$. With these facts, the sum in (6) will consist of only these four terms:

$$\|z_j - f_j^r\|^2 - \|z_j - f_k^r\|^2 + \|z_k - f_k^r\|^2 - \|z_k - f_j^r\|^2$$

Noting that $\|a-b\|^2 - \|a-c\|^2 = b^Tb - 2a^T(b-c) - c^Tc$, and further simplifying leads to the following simple expression for a:

$$a = \exp\left[\frac{1}{\sigma^2}(z_k - z_j)^T(f_j^r - f_k^r)\right]$$

The expression above is cheap to evaluate and is applicable in a wide variety of applications. It asserts that, when we switch the correspondence mapping of two randomly chosen indices, the acceptance ratio a only depends on the dot product between two vectors related to the switch. This holds whenever the noise on the measurement is normally distributed and isotropic.

4 MAP or ML Estimates

If we are interested only in a single MAP or ML estimate, we can avoid having to sample over the large joint space of mappings c and parameters θ by making use of the Expectation-Maximization (EM) algorithm. To obtain the MAP estimate, we need to maximize $P(\theta|z)$. Since this is given by (3), we can attempt to directly maximize that expression, but this is likely to be as intractable as sampling from it. Fortunately, the EM algorithm provides a more tractable alternative, and it can be proven that the EM algorithm converges to a local maximum of $P(\theta|z)$ [12]. The idea of EM is to maximize the *expected* log posterior

$$Q^t(\theta) \triangleq E\{\log P(\theta|z, c)\}$$

where the expectation is taken with respect to the posterior distribution $P(c|z, \theta^t)$ over all possible mapping c given the data z and a current guess θ^t for the parameters. The EM algorithm then iterates over [20]:

1. **E-step:** Calculate the expected log posterior$Q^t(\theta)$:

$$Q^t(\theta) = \sum_c P(c|z, \theta^t) \log P(\theta|z, c) \tag{7}$$

2. **M-step:** Re-estimate θ^{t+1} by maximizing $Q^t(\theta)$:

$$\theta^{t+1} = \underset{\theta}{\operatorname{argmax}} \;\; Q^t(\theta)$$

However, instead of calculating $Q^t(\theta)$ exactly using (7), which again involves summing over a combinatorial number of terms, we can replace it by a Monte Carlo approximation:

$$Q^t(\theta) \approx \frac{1}{R} \sum_{r=1}^{R} \log P(\theta|z, c^r) \tag{8}$$

where $\{c^r; r \in 1..R\}$ is a sample from $P(c|z, \theta^t)$ obtained by MCMC sampling. Formally this can be justified in the context of a *Monte Carlo EM* or MCEM, a version of the EM algorithm where the E-step is executed by a Monte-Carlo process [20, 12]. The sampling proceeds exactly as in the previous section, using the Metropolis algorithm, but now with a fixed parameter $\theta = \theta^t$. Note that θ^t changes at each iteration of EM, and in each iteration we sample from a different posterior distribution $P(c|z, \theta^t)$ over mappings c .

5 Results

We have applied the concepts above with success in the domain of *structure from motion* (SFM). The problem is as follows: given that n 3D features are observed in m images, estimate both the 3D position of the n features and the m camera poses at the time the measurements were taken. Here we restrict ourselves to the case were images are taken from separated viewpoints and the a small number of sparse measurements are taken in each image.

SFM is one of the most successful applications of computer vision, and there have been numerous papers with algorithms that work under a wide variety of assumptions [21, 15, 19, 10]. However, the applicability of these methods is limited by by their reliance on error-prone correspondence techniques. Feature-tracking techniques often fail to produce correct matches due to large motions, occlusions, or ambiguities. Furthermore, errors in one frame are likely to propagate to all subsequent frames of the sequence. Outlier rejection techniques [1, 23, 6] can ameliorate these problems, but at the cost of eliminating valid features from the reconstruction, resulting in an incomplete model that does not take into account all available image measurements.

By working with the distribution over correspondence mappings from 3D features to 2D measurements, we avoid the reliance on a separate correspondence finding algorithm. Instead, structure and motion estimates are obtained that incorporate all the knowledge that can be inferred from the measurement data, including ambiguity between different correspondence mappings. The way this is done is through the use of the the Monte-Carlo EM algorithm as described in the previous section. The implementation details are beyond the scope of this paper, but are detailed in a separate paper [5].

In practice, the algorithm converges consistently and fast to an estimate for the structure and motion where the correct correspondence is the most probable one, and where most if not all assignments in the different images agree with

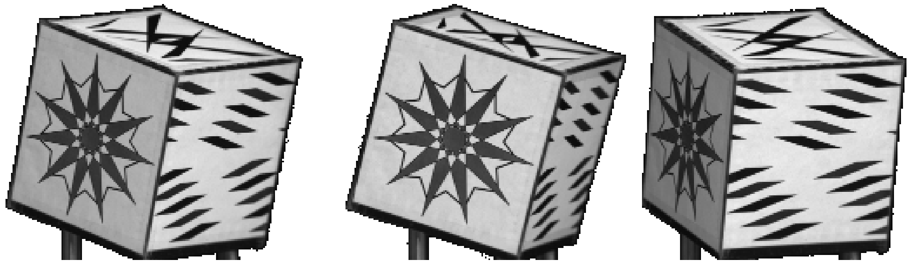

Figure 5: Three out of 11 *cube* images. Although the images were originally taken as a sequence in time, the ordering of the images is irrelevant to our method.

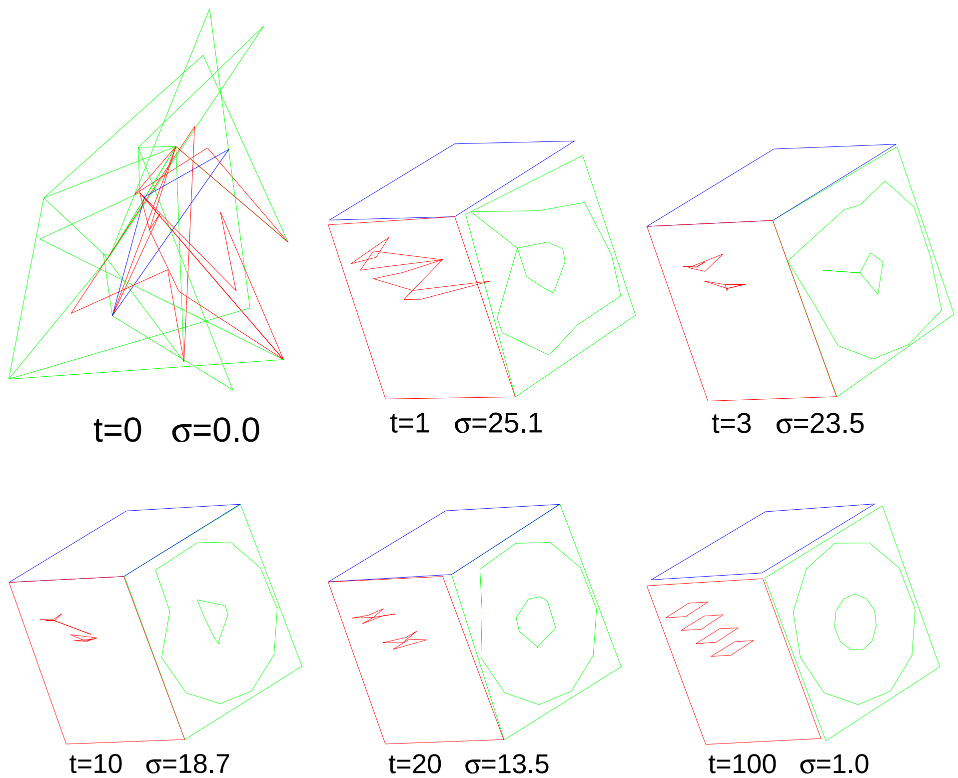

Figure 6: Starting from a random guess for the structure (t=0) we recover gross 3D structure in the very first iteration (t=1). As the annealing parameter σ is gradually decreased, successively finer details are resolved (iterations 3,10,20, and 100 are shown).

each other. We illustrate this using the image set shown in Figure 5, which was taken under orthographic projection. The complete set consists of 11 images, and 55 features were manually measured in each image, for a total of 550 2D measurements. To initialize, the 55 unknown 3D points were generated randomly in a normally distributed cloud around a depth of 1, whereas the 11 camera poses were all initialized at the origin. We ran the EM algorithm for 100 iterations, and ran the sampler each time in each image for 10000 steps. The algorithm converges to the solution in about a minute on a standard PC, but may have to be restarted on occasion when it gets stuck in a local minimum, as can happen with EM. One reason for its computational efficiency is use of flip proposals to propose new mappings, which only involves a simple local calculation to evaluate, as discussed.

To avoid getting stuck in local minima, it is important in practice to add *annealing* to the basic EM scheme. In annealing we artificially increase the noise parameter σ for the early iterations, gradually decreasing it to its correct value. This has two beneficial consequences. First, the posterior distribution $P(c|z, \theta^t)$ will be less peaked when σ is high, so that the Metropolis sampler will explore the space of assignments more easily, and avoid getting stuck on islands of high probability. Second, the expected log likelihood $Q^t(\theta)$ is smoother and has less local maxima at higher values for σ. We use a logarithmically decreasing annealing scheme, but found that the algorithm is not sensitive to the exact scheme used.

The typical evolution of the algorithm is illustrated in Figure 6, where we have shown a wireframe model of the recovered structure at successive instants of time. There are two important points to note: (a) *the gross structure is recovered in the very first iteration, starting from random initial structure*, and (b) finer details of the structure are gradually resolved as the parameter σ is decreased. The estimate for the structure after convergence is almost identical to the one found by factorization when given the correct correspondence. Incidentally, we found the algorithm converges less often when we replace the random initialization by a 'good' initial estimate where all the points in some image are projected onto a plane of constant depth.

6 Conclusions and Future Directions

In this paper we show that taking into account a distribution over possible correspondence mappings gives more complete and more accurate knowledge about the properties we want to estimate. To obtain these distributions, we propose the use of the Metropolis sampler, an MCMC method, which can be implemented efficiently in many applications. We demonstrate the validity of the approach with some simple examples from pose estimation, and a fairly elaborate structure from motion application. We have many more results for the SFM application, that are discussed in more depth and with detailed implementation notes in an accompanying paper (reference omitted). However, we have only explored the joint sampling approach discussed in Section 3. Ongoing work is evaluating the applicability of the method in object recognition and motion estimation. Specifically, we would like to see whether the approach has merit in estimating quantities such as the fundamental matrix [23].

References

[1] P.A. Beardsley, P.H.S. Torr, and A. Zisserman. 3D model acquisition from extended image sequences. In *Eur. Conf. on Computer Vision (ECCV)*, pages II:683–695, 1996.

[2] A.S. Bedekar and R.M. Haralick. Finding correspondence points based on bayesian triangulation. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 61–66, 1996.

[3] Y. Cheng, R. Collins, A. Hanson, and E. Riseman. Triangulation without correspondences. In *DARPA Image Understanding Workshop (IUW)*, 1994.

[4] Y. Cheng, V. Wu, R. Collins, A. Hanson, and E. Riseman. Maximum-weight bipartite matching technique and its application in image feature matching. In *Proc. SPIE Visual Comm. and Image Processing, Orlando, FL.*, 1996.

[5] F. Dellaert, S.M. Seitz, C.E. Thorpe, and S. Thrun. Structure from motion without correspondence. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2000.

[6] D.A. Forsyth, S. Ioffe, and J. Haddon. Bayesian structure from motion. In *Int. Conf. on Computer Vision (ICCV)*, pages 660–665, 1999.

[7] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter. Introducing Markov chain Monte Carlo. [8].

[8] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, editors. *Markov chain Monte Carlo in practice*. Chapman and Hall, 1996.

[9] S. Gold, A. Rangarajan, C. Lu, S. Pappu, and E. Mjolsness. New algorithms for 2D and 3D point matching. *Pattern Recognition*, 31(8):1019–1031, 1998.

[10] R.I. Hartley. Euclidean reconstruction from uncalibrated views. In *Application of Invariance in Computer Vision*, pages 237–256, 1994.

[11] D.J.C. MacKay. Introduction to Monte Carlo methods. In M.I.Jordan, editor, *Learning in graphical models*, pages 175–204. MIT Press, 1999.

[12] G.J. McLachlan and T. Krishnan. *The EM algorithm and extensions*. Wiley series in probability and statistics. John Wiley & Sons, 1997.

[13] H. Pasula, S. Russell, M. Ostland, and Y. Ritov. Tracking many objects with many sensors. In *Int. Joint Conf. on Artificial Intelligence (IJCAI)*, Stockholm, 1999.

[14] M. Pilu. A direct method for stereo correspondence based on singular value decomposion. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 261–266, 1997.

[15] C. Poelman and T. Kanade. A paraperspective factorization method for shape and motion recovery. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(3):206–218, March 1997.

[16] G.L. Scott and H.C. Longuet-Higgins. An algorithm for associating the features of two images. *Proceedings of Royal Society of London*, B-244:21–26, 1991.

[17] L.S. Shapiro and J.M. Brady. Feature-based correspondence: An eigenvector approach. *Image and Vision Computing*, 10(5):283–288, June 1992.

[18] J.P.P. Starink and E. Backer. Finding point correspondences using simulated annealing. *Pattern Recognition*, 28(2):231–240, February 1995.

[19] R. Szeliski and S.B. Kang. Recovering 3D shape and motion from image streams using non-linear least squares. Technical Report CRL 93/3, DEC Cambridge Research Lab, 1993.

[20] M.A. Tanner. *Tools for Statistical Inference*. Springer, 1996.

[21] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: a factorization method. *Int. J. of Computer Vision*, 9(2):137–154, Nov. 1992.

[22] S. Ullman. *The interpretation of visual motion*. MIT Press, Cambridge, MA, 1979.

[23] Z.Y. Zhang. Determining the epipolar geometry and its uncertainty - a review. *Int. J. of Computer Vision*, 27(2):161–195, March 1998.