Multiframe Motion Segmentation via Penalized MAP Estimation and Linear Programming

Conference Paper · January 2009

DOI: 10.5244/C.23.61 · Source: DBLP

Multiframe Motion Segmentation via Penalized MAP Estimation and Linear Programming

Han Hu
huh04@mails.tsinghua.edu.cn
Quanquan Gu
gqq03@mails.tsinghua.edu.cn
Lei Deng
dally211@163.com
Jie Zhou
jzhou@tsinghua.edu.cn

State Key Laboratory on Intelligent Technology and Systems
Tsinghua National Laboratory for Information Science and Technology (TNList)
Department of Automation, Tsinghua University, Beijing, China, 100084

In this paper, we present a novel algorithm for multi-frame motion segmentation problem under the affine camera model. The algorithm is based on a *mixture-of-subspace* model, which describes the multi-body motions in a unified style. Using this model, the motion segmentation problem can be formulated as a maximum a posteriori estimation (MAPE) problem taking account of model complexity. To solve this problem, a list of candidate motion models is generated by a certain scheme. Then the problem is converted into a linear programming problem and therefore can be effectively solved.

Ref. [1] assumes that trajectories from a single rigid motion lie in a linear subspace of $\mathbb{R}^{2F}$ with dimension between two and four. Denote the dimension of the motion by r, then we can use a set of orthonormal vectors $C = \{\mathbf{u_i}\}_{i=1}^{r}$ to completely represent a motion model. The orthonormal vectors can be obtained from the first r left singular vectors of U by SVD factorization [2],

$$W = U_{2F\times 2F}\Sigma_{2F\times P}V_{P\times P}^{T}. \quad (1)$$

To describe multibody motions, we define indicator variables $L_{ij}, (i = 1,\dots,p; j = 1,\dots,K)$, of which $L_{ij} = 1$ means the i^{th} trajectory belongs to the j^{th} model, and $L_{ij} = 0$ otherwise. Then a *mixture of subspace* model can be formulated for each trajectory $\mathbf{w}_i$ as follows,

$$\begin{aligned} &\textstyle\sum_{j=1}^{K} L_{ij} d(\mathbf{w}_i, C_j) = 0, \\ &s.t.\ \textstyle\sum_{j=1}^{K} L_{ij} = 1, L_{ij} \in \{0,1\}, \end{aligned} \quad (2)$$

If the number of motions K is known, then the maximum a posteriori estimation (MAPE) method can be used to solve C_j and L_{ij}.

Assume the noises are of Gaussian distribution, and the standard deviation of noise on the j^{th} model is σ_j . Given motion models $\{C_j\}_{j=1,\dots,K}$, memberships L_{ij}, and the prior probability of C_j and L_{ij}, under the independent assumption, we can reach the following log-MAPE cost function,

$$\begin{aligned} &\max_{C,L,K} \ln L = \max_{C,L,K} \ln p(C_j, L_{ij}, i=1,...,P; j=1,...,K|\mathbf{w}_i, i=1,...,P) \\ &= \max_{C,L,K} \textstyle\sum_{i=1}^{P} (\ln(p(\mathbf{w}_i|C_j, L_{ij}, j=1,...,K)) + \ln(p(C_j, L_{ij}))) \\ &= \min_{C,L,K} \textstyle\sum_{i=1}^{P}\sum_{j=1}^{K} L_{ij}(\frac{d^2(\mathbf{w}_i,C_j)}{2\sigma_j^2} + \ln(\sigma_j) - \ln p(C_j, L_{ij})) \end{aligned} \quad (3)$$

Using MAPE assumes that the number of motions K is known as a priori and motions with different dimensions are equal. However, in practice, K is usually unknown, and the motions with different dimensions cannot be treated as the same. The log-likelihood cost function certainly rises when K grows or the dimension of motions increases. Hence, a tradeoff between fitting error and model complexity should be adopted. Several criteria can touch this purpose, e.g. Akainke information criterion (AIC) and Bayesian information criterion (BIC) [4]. Then we can obtain the cost function as,

$$\mathscr{J} = -\ln L + \alpha \textstyle\sum_{j=1}^{K} Pr_j \quad (4)$$

where Pr_j indicates the complexity of the j^{th} model, and α is a penalizing factor. In this work, since the two parts of the cost function are at similar scale level, α may be chosen from 0.1 to 10. Therefore, we reach the following penalized MAPE problem:

$$\begin{aligned} &\min_{C,L,K,r} \sum_{i=1}^{P}\sum_{j=1}^{K} L_{ij}(\frac{d^2(\mathbf{w}_i,C_j)}{2\sigma_j^2} + \ln(\sigma_j) - \ln p(C_j, L_{ij})) + \alpha \sum_{j=1}^{K} Pr_j \\ &s.t. K \geq 1, K \in \mathbb{Z}, \textstyle\sum_{j=1}^{K} L_{ij} = 1, L_{ij} \in \{0,1\}, r_j = \{0,2,3,4\}. \end{aligned} \quad (5)$$

Eq.(5) is a combinatorial optimization problem, which is NP hard. In this paper, a method based on linear programming relaxation is adopted to solve the problem.

Revisit the cost function of Eq.(5). Suppose that somehow we have already obtained a list of candidate motion models $\Phi\{C_1,\dots,C_N\}$, and the K true motions are contained by the list. Define indicating variables x_j with $x_j = 1$ if the j^{th} candidate motion is a true motion and $x_j = 0$ otherwise. Obviously we have $x_j = \max_{1\leq i\leq P}\{L_{ij}\}$. Denote the rank of the j^{th} candidate motion by r_j, and then the cost function of Eq.(5) can be rewritten as,

$$\begin{aligned} &\min_{C,L,x} \textstyle\sum_{i=1}^{P}\sum_{j=1}^{N} L_{ij}(\frac{d^2(\mathbf{w}_i,C_j)}{2\sigma_j^2} + \ln(\sigma_j) - \ln p(C_j, L_{ij})) \\ &\quad + \alpha P(2\textstyle\sum_{\{j:r_j=2\}} x_j + 3\sum_{\{j:r_j=3\}} x_j + 4\sum_{\{j:r_j=4\}} x_j) \end{aligned} \quad (6)$$

The terms of $d^2(\mathbf{w}_i, C_j)$, σ_j and prior probability $p(C_j, L_{ij})$ can all be pre-computed and they together can be considered as coefficients of the variables L_{ij}. Now the problem becomes linear to the unknown variables L and x.

Then we exploit the idea of continuous relaxation, which is a popular technique in Operational Research [3]. Thus we obtain the ultimate optimization problem:

$$\begin{aligned} &\min_{C,L,x} \sum_{i=1}^{P}\sum_{j=1}^{N} L_{ij}\hat{d}^2(\mathbf{w}_i,C_j) + \alpha P(2\sum_{\{j:r_j=2\}} x_j + 3\sum_{\{j:r_j=3\}} x_j + 4\sum_{\{j:r_j=4\}} x_j) \\ &s.t.\quad \textstyle\sum_{j=1}^{N} L_{ij} = 1, \forall i; \\ &\qquad\ L_{ij} \leq x_j, \forall i, j; \\ &\qquad\ 0 \leq L_{ij} \leq 1,\ 0 \leq x_{ij} \leq 1, \forall i, j, \end{aligned} \quad (7)$$

where $\hat{d}^2(\mathbf{w}_i,C_j) = \frac{d^2(\mathbf{w}_i,C_j)}{2\sigma_j^2} + \ln(\sigma_j) - \ln p(C_j, L_{ij})$ is *normalized distance* between $\mathbf{w}_i$ and the C_j. And the constraint $L_{ij} \leq x_j$ is from $x_j = \max_{1\leq i\leq P}\{L_{ij}\}$ by certain algebraic transformation.

Some highlights of the proposed method include:

1. Using *mixture of subspace* model to describe the multi-body motions makes the following unified formulation possible.
2. The proposed method uses penalized MAPE as the segmentation criterion. Since the prior probabilities of the candidate models are introduced, the MAP estimator is potentially more effective than the conventional maximum likelihood estimator. In addition, the number of motions can be automatically estimated using model complexity penalizing.
3. The assumption that the noises are not the same among candidate motion models is more reasonable than the constant noise assertion.
4. Linear programming can guarantee that the solutions are the best combination from the candidate motion models. Also it can easily incorporate other prior knowledge, e.g. the number of motions.

[1] J. Costeira and T. Kanade. A multibody factorization method for independently moving objects. *IJCV*, 29(3):108–121, 1998.

[2] G. H. Golub and C. F. Van-Loan. *Matrix Computation*. 2nd edn, The John Hopkins University Press, Baltimore, 1989.

[3] D. Hochbaum. *Approximation Algorithms for NP-Hard Problems*. PWS Publishing Company, Boston, MA, 1997.

[4] K. Kanatani. Geometric information criterion for model selection. *IJCV*, 26(3):171–189, 1998.