Motion Segmentation: a Review

Luca ZAPPELLA [a] Xavier LLADÓ [a] and Joaquim SALVI [a]

[a] *Institute of Informatics and Applications, University of Girona, Girona (Spain)*

Abstract. Motion segmentation is an essential process for many computer vision algorithms. During the last decade, a large amount of work has been trying to tackle this challenge, however, performances of most of them still fall far behind human perception. In this paper the motion segmentation problem is studied, analyzing and reviewing the most important and newest techniques. We propose a classification of all these techniques into different categories according to their main principle and features. Moreover, we point out their strengths and weaknesses and finally we suggest further research directions.

Keywords. Computer Vision, Motion Analysis, Motion Segmentation

Introduction

Motion segmentation aims at decomposing a video in moving objects and background. In many computer vision algorithms this decomposition is the first fundamental step. It is an essential building block for robotics, inspection, metrology, video surveillance, video indexing, traffic monitoring and many other applications. A great number of researchers has focused on the segmentation problem and this testifies the relevance of the topic. However, despite the vast literature, performances of most of the algorithms still fall far behind human perception.

In this paper we present a review on the main motion segmentation approaches with the aim of pointing out their strengths and weaknesses and suggesting new research directions. Our work is structured as follows. In the next section, a general overview of motion segmentation is presented. We describe common issues and the main attributes that should be considered when studying this type of algorithms. Furthermore, a classification among the different strategies is proposed. In section 2 the main ideas of each category are analyzed, reviewing also the most recent and important techniques. Finally, in section 3 general considerations are discussed for each category and conclusions are drawn.

1. Motion Segmentation: Main Features

In this section Common issues are described, a possible classification of motion segmentation algorithms is proposed, and a description of the main attributes that should be considered when studying this type of algorithms is presented.

In order to obtain an automatic motion segmentation algorithm that can work with real images there are several issues that need to be solved, particularly important are:

noise, missing data and lack of a priori knowledge. One of the main problem is the presence of noise. For some applications the noise level can become critical. For instance, in underwater imaging there are some specific sub-sea phenomena like water turbidity, marine snow, rapid light attenuation, strong reflections, back-scattering, non-uniform lighting and dynamic lighting that dramatically degrade the quality of the images [1,2,3]. Blurring is also a common issue especially when motion is involved [3]. Another common problem is caused by the fact that moving objects can create occlusions, or even worst, the whole object can disappear and reappear in the scene. Finally, it is important to take into account that not always it is possible to have prior knowledge about the objects or about the number of objects in the scene [3].

The main attributes of a motion segmentation algorithm can be summarized as follows.

- *Feature-based* or *Dense-based*: In feature-based methods, the objects are represented by a limited number of points like corners or salient points, whereas dense methods compute a pixel-wise motion [4].
- *Occlusions*: it is the ability to deal with occlusions.
- *Multiple objects*: it is the ability to deal with more than one object in the scene.
- *Spatial continuity*: it is the ability to exploit spatial continuity.
- *Temporary stopping*: it is the ability to deal with temporary stop of the objects.
- *Robustness*: it is the ability to deal with noisy images (in case of feature based methods it is the position of the point to be affected by noise but not the data association).
- *Sequentiality*: it is the ability to work incrementally, this means for example that the algorithm is able to exploit information that was not present at the beginning of the sequence.
- *Missing data*: it is the ability to deal with missing data.
- *Non-rigid object*: it is the ability to deal with non-rigid objects.
- *Camera model*: if it is required, which camera model is used (orthographic, paraperspective or perspective).

Furthermore, if the aim is to develop a generic algorithm able to deal in many unpredictable situations there are some algorithm features that may be considered as a drawback, specifically:

- *Prior knowledge*: any form of prior knowledge that may be required.
- *Training*: some algorithms require a training step.

Motion segmentation literature is wide. In order to make the overview easier to read and to create a bit of order, algorithms will be divided into categories which represent the main principle underlying the approach. The division is not meant to be tight, in fact some of the algorithms could be placed in more than one group. The categories identified are: *Image Difference*, *Statistical* (further divided into Maximum A posteriori Probability, Particle Filter and Expectation Maximization), *Optical Flow*, *Wavelets*, *Layers* and *Factorization*. For each category some articles, among the most representative and the newest proposals are analyzed. Table 1 offers a compact at-a-glance overview of the algorithms examined in this work with respect to the most important attributes of a motion segmentation algorithm.

Table 1. Summary of the examined techniques with respect to the most important attributes. Note the methods are classified into 6 categories: Image Difference, Statistical, Optical Flow, Wavelets, Layers and Factorization methods. When an attribute is not relevant for a technique the symbol “-” is used.

			Features (F) / Dense (D)	Occlusion	Multiple Objects (S static camera)	Spatial Continuity	Temporary Stopping	Robustness	Sequentiality	Missing data	Non-rigid objects	Camera Model(O Orthographic, p para-persp., P Persp.)	Prior knowledge (X)/Training (T)
Image	**Differ.**	*Cavalaro et al. 2005* [1]	F/D	√	√	√	√		-	-	√	-	
		Cheng et al. 2006 [5]	D	√	S	√		√	-	-	√	-	X
		Li et al. 2007 [6]	D		√	√	√		-	-	√	-	
		Colombari et al. 2007 [7]	D	√	√	√	√	√	-	-	√	-	
Statistical	MAP	*Rasmussen et al. 2001* [8]	D	√	√	√	√		-	-	√	-	X
		Cremers et al. 2005 [9]	D	√	√	√	√		-	-	√	-	X
		Shen et al. 2007 [10]	D	√	√	√	√	√	-	-	√	-	X
	PF	*Rathi et al. 2007* [11]	D	√	√	√	√		-	-		-	X
	EM	*Stolkin et al. 2008* [3]	D	√	√	√	√	√	-	-		-	X
Wavelets		*Wiskott 1997* [12]	F		√	√			-	-		-	
		Kong et al. 1998 [13]	F	√	√	√		√	-	-	√	-	
O.F.		*Zhang et al. 2007* [14]	F		√	√				-		-	X
Layers		*Kumar et al. 2008* [4]	F	√	S	√	√	√	-	-	√	-	T
Factorization		*Costeira et al. 1998* [15]	F		√		√					O	
		Ichimura et al. 2000 [16]	F		√		√					O	
		Kanatani et al. 2002 [17]	F		√		√					O	
		Brand 2002 [18]	F	√	√		√	√		√	√	O	
		Zelnik-Manor et al. 2003 [19]	F		√		√	√				O	
		Zhou et al. 2003 [20]	F	√	√		√	√		√	√	p	
		Sugaya et al. 2004 [21]	F		√		√	√				O	
		Zelnik-Manor et al. 2004 [22]	F		√		√	√				O	
		Gruber et al. 2004 [23]	F	√	√		√			√		O	
		Vidal et al. 2004 [24]	F	√	√		√			√		O	
		Yan et al. 2006 [25]	F		√		√				√	O	
		Gruber et al. 2006 [26]	F	√	√	√	√	√		√		O	
		Del Bue et al. 2007 [27]	F		√		√	√			√	P	
		Goh et al. 2007 [28]	F		√		√			√	√	P	X
		Julià et al. 2007 [29]	F	√	√		√	√		√		O	

2. Main Motion Segmentation Techniques

In this section we review the most important motion segmentation categories.

Image difference is one of the simplest and most used technique for detecting changes. It consists in thresholding the intensity difference of two frames pixel by pixel. The result is a coarse map of the temporal changes. An example of an image sequence and the image difference result is shown in figure 1. Despite its simplicity, this technique cannot be used in its basic version because it is really sensitive to noise. Moreover, when the camera is moving the whole image is changing and, if the frame rate is not high enough, the result would not provide any useful information. However, there are a few techniques based on this idea. The key point is to compute a rough map of the changing areas and for each blob to extract spatial or temporal information in order to track the region. Usually different strategies to make the algorithm more robust against noise and light changes are also used. Examples of this technique can be found in [1,5,6,7].

Statistic theory is widely used in the motion segmentation field. In fact, motion segmentation can be seen as a classification problem where each pixel has to be classified as background or foreground. Statistical approaches can be further divided depending on the framework used. Common frameworks are Maximum A posteriori Probability (MAP), Particle Filter (PF) and Expectation Maximization (EM). Statistical approaches provide a general tool that can be used in a very different way depending on the specific technique.

MAP is based on Bayes rule:

$$P(w_j|x) = \frac{p(x|w_j)P(w_j)}{\sum_{i=1}^{c} p(x|w_i)P(w_i)}$$

where x is the object to be classified (usually the pixel), $w_1..w_c$ are the c classes (usually background or foreground), $P(w_j|x)$ is the "a posteriori probability", $p(x|w_j)$ is the conditional density, $P(w_j)$ is the "a priori probability" and $\sum_{i=1}^{c} p(x|w_i)P(w_i)$ is the "density function". MAP classifies x as belonging to the class w which maximizes the "a posteriori probability". MAP is often used in combination with other techniques. For example, in [8] is combined with a Probabilistic Data Association Filter. In [9] MAP is used together with level sets incorporating motion information. In [10] the MAP frame-

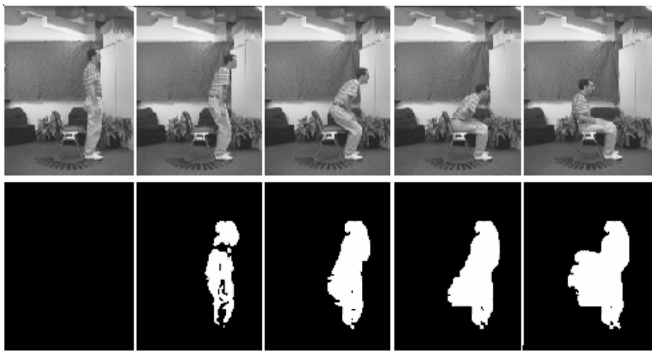

Figure 1. Example of image sequence and its image difference result. Sequence taken from [30].

work is used to combine and exploit the interdependence between motion estimation, segmentation and super resolution.

Another widely used statistical method is PF. The main aim of PF is to track the evolution of a variable over time. The basis of the method is to construct a sample-based representation of the probability density function. Basically, a series of actions are taken, each of them modifying the state of the variable according to some model. Multiple copies of the variable state (particles) are kept, each one with a weight that signifies the quality of that specific particle. An estimation of the variable can be computed as a weighted sum of all the particles. PF is an iterative algorithm, each iteration is composed by prediction and update. After each action the particles are modified according to the model (prediction), then each particle weight is re-evaluated according to the information extracted from an observation (update). At every iteration particles with small weights are eliminated [31]. An example of PF used in segmentation can be found in [11] where some well known algorithms for object segmentation using spatial information, such as geometric active contours and level sets, are unified within a PF framework.

EM is also a frequently exploited tool. The EM algorithm is an efficient iterative procedure to compute the Maximum Likelihood (ML) estimate in presence of missing or hidden data. ML consists in estimating the model parameter(s) that most likely represent the observed data. Each iteration of EM is composed by the E-step and the M-step. In the E-step the missing data are estimated using the conditional expectation, while in the M-step the likelihood function is maximized. Convergence is assured since the algorithm is guaranteed to increase the likelihood at each iteration [32]. As an example, in [3] an algorithm which combines EM and Extended-Markov Random Field is presented.

Another group of motion segmentation algorithms that we have identified is the one based on wavelets. These methods exploit the ability of wavelets to perform analysis of the different frequency components of the images, and then study each component with a resolution matched to its scale. Usually wavelet multi-scale decomposition is used in order to reduce the noise and in conjunction with other approaches, such as optical flow, applied at different scales. For instance, in [12] Wiskott combines Optical Flow with Gabor-wavelets in order to overcome the aperture problem. Furthermore, he extracts and tracks edges using the information provided by the Mallat-wavelets. Finally, the results are merged in order to obtain a robust segmentation. A different approach is presented

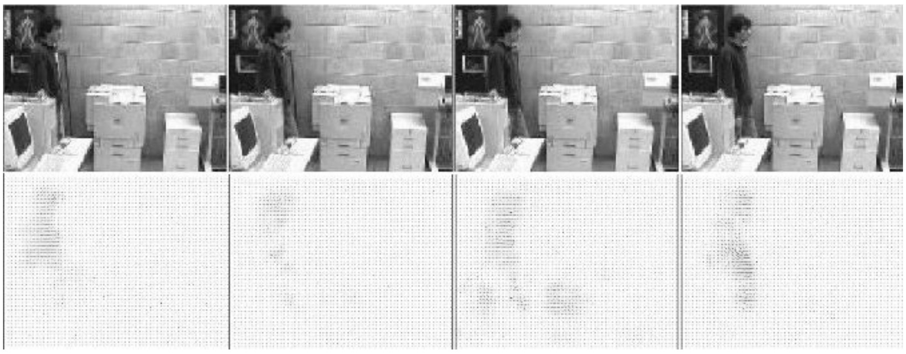

Figure 2. Example of Optical Flow: darker areas are the vectors of apparent velocities with length grater than zero

in [13] where the motion segmentation algorithm is based on Galilean wavelets. These wavelets behave as matched filters and perform minimum mean-squared error estimations of velocity, orientation, scale and spatio-temporal positions. This information is finally used for tracking and segmenting the objects.

Optical Flow (OF) is a vector motion field which describes the distribution of the apparent velocities of brightness patterns in a sequence, figure 2. Like image difference, OF is an old concept greatly exploited in computer vision. It was first formalized and computed for image sequences by Horn and Schunck in the 1980 [33], but the idea of using discontinuities in the OF in order to segment moving objects is even older. Since the work of Horn and Schunck, many other approaches have been proposed. In the past the main limitations of such methods were the high sensitivity to noise and the high computational cost. Nowadays, thanks to the high process speed of computers and to improvements made by research, OF is widely used. In [14] a method to segment multiple rigid-body motions using Line Optical Flow is presented.

The key idea of layers based techniques is to understand which are the different depth layers in the image and which objects (or which part of an articulated object) lie on which layer. This approach is often used in stereo vision as it is easier to compute the depth distance. However, without computing the depth it is possible to estimate which objects move on similar planes. This is extremely useful as it helps to solve the occlusion problem. In [4] a method for learning a layered representation of the scene is proposed. They initialize the method by first finding coarse moving components between every pair of frames. They divide the image in patches and find the rigid transformation that moved the patch from one frame to the next. The initial estimate is then refined using two minimization algorithms: $\alpha\beta$-swap and α-expansion [34]. Figure 3 gives an example of how frame sequence can be used to learn the layers and the segments (objects or part of an objects) that lie on a specific layer.

Since *Tomasi and Kanade (1992)* [35] introduced a factorization technique to recover structure and motion using features tracked through a sequence of images, factorization methods have become very popular especially thanks to their simplicity. The idea is to factorize the trajectory matrix W (the matrix containing the position of the P features tracked throughout F frames) into two matrices: motion M and structure S, figure 4. If the origin of the world coordinate system is moved at the centroid of all the feature points, and in absence of noise, the trajectory matrix is at most rank 3. Exploiting this constraint, W can be decomposed and truncated using singular value decomposition

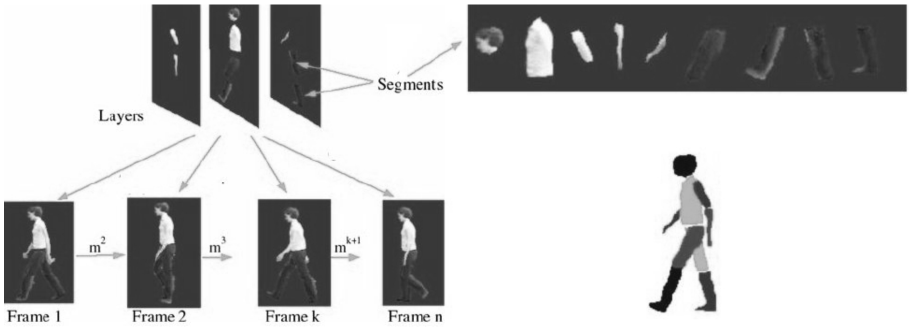

Figure 3. Example of Layers representation [4]

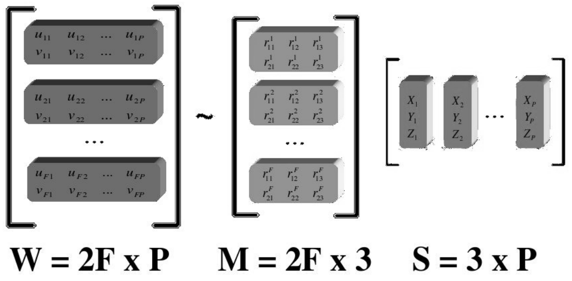

Figure 4. Basic idea of structure from motion: factorize matrix W into M and S

(SVD), figure 5. Because the rows of the motion matrix are orthonormal, the matrix D

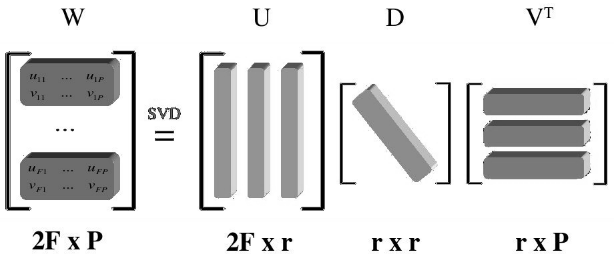

Figure 5. W can be decomposed and truncated using SVD and exploiting its rank deficiency

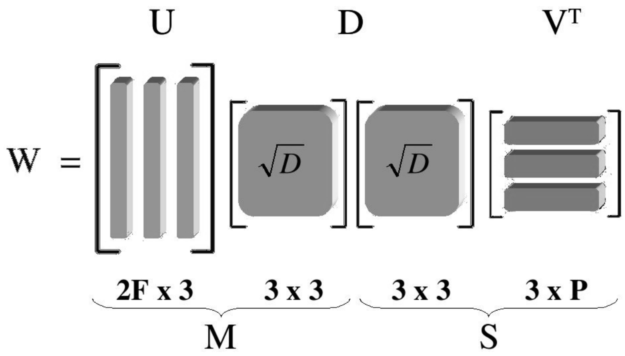

Figure 6. Exploiting orthogonality constraints of the motion matrix, W can be decomposed into motion and structure, up to a scale factor

can be evaluated using orthogonality constraints, and finally decomposed, up to a scale factor, as shown in figure 6.

This initial algorithm works for one static object viewed from a moving camera which is modeled with the simplest affine camera model: the orthographic projection. Despite the fact that this method gives the 3D structure of the object and the motion of the camera, it has evident limits: it cannot really segment (it assumes that the features belong to the same object), it can deal only with a single rigid object, it is very sensitive to noise and it is not able to deal with missing data and outliers. However, it was the first method of this family and the solution is mathematically elegant. From this initial structure from motion approach, and following the same idea of forcing the rank constraint, many approaches have been proposed in the field of motion segmentation. These methods are based on using the dimensionality of the subspace in which the image trajectories lie to perform the motion segmentation. For example, Costeira and Kanade proposed a factorization framework able to deal with multiple objects moving independently [15]. The assumption of independence between motion subspaces was also used in other approaches such as [16,17]. Moreover, Yan and Pollefeys [25] proposed a new general segmentation framework able to deal with different types of motion: independent, rigid, articulated and non-rigid motions. More recently, Julia et al. [29] extended Yan and Pollefeys approach to deal also with missing data in the image trajectories. Other remarkable factorization approaches are [19,21,22,24,26].

3. Conclusions

This review should have given an idea of how vast the motion segmentation literature is, and the fact that research in this field is still active (most of the papers presented were published after 2005) is a sign of the importance of the solution to this problem. On the other hand, effervescent research activity signifies that an outstanding solution has yet to be found.

As can be seen from Table 1, image difference is mainly based on dense representation of the objects. It combines simplicity and good overall results being able to deal with occlusions, multiple objects and non-rigid objects. The main problems of these techniques are the difficulty to deal with temporary stopping and with moving cameras. In order to be successful in these situations a history model of the background needs to be built. Furthermore, these algorithms are still very sensitive to noise and light changes.

Statistical approaches also use mainly dense based representation. These methods work well with multiple objects and are able to deal with occlusions and temporary stopping. In general they are robust as long as the model reflects the actual situation but they degrade quickly as the model fails to represent the reality. Finally, most of the statistic approaches require some kind of a priori knowledge.

Wavelets solutions seem to provide good results, wavelets were in fashion during the 90s and now it seems that the research interest is decreased. Their ability in performing multi resolution analysis could be exploited in order to extract information about the different depth planes in the scene, thus helping to solve the occlusion problem.

Optical flow is theoretically a good clue in order to segment motion. However, OF alone it is not enough since it cannot help to solve occlusions and temporal stopping. Moreover, these methods are sensitive to noise and light changes.

The layer solution is very interesting. It is probably the more natural solution for the occlusions. Note that human beings also use depth concept to solve this problem. The main drawback of this strategy is the level of complexity of the algorithm and the high number of parameters that need to be tuned.

Factorization methods are an elegant solution based on feature points. Besides, they provide not only the segmentation but they can be naturally connected to a structure from motion algorithm in order to recover the 3D structure of the objects and the motion of the camera. Furthermore, they do not have any problem with temporary stopping because the features can be tracked even if the object is not moving (provided that this is a temporary situation). With respect to other approaches, factorization methods are particularly weak in terms of ability to deal with noise and outliers, and they have more problems to deal with non rigid objects because the non rigidity has to be explicitly taken into account. A quick glance at the table may catch the attention on two elements. The first is that, to the best of our knowledge, there is no factorization algorithm which is able to provide segmentation in an incrementally way. A technique able to provide a good object segmentation exploiting only few frames, and refining the solution with the time, it would be an important step ahead. The second is that with the exception of [26], spatial continuity is not exploited. These may suggest that the usage of this information may help to improve factorization method performances, especially in terms of robustness and ability to deal with occlusions.

Considering the aspects emerged in this review, we believe that a factorization approach could be a good starting point for proposing a novel segmentation technique. They are based on a powerful mathematical framework and they can provided easily segmentation and structure of the objects. In order to obtain more robust results it would be interesting to study different ways to merge factorization with spatial information and to exploit also the ability of statistical frameworks to find and use hidden information. These are the directions that we intend to follow in the near future.

References

[1] A. Cavallaro, O. Steiger, and T. Ebrahimi, "Tracking Video Objects in Cluttered Background," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 4, pp. 575–584, 2005.

[2] E. Trucco and K. Plakas, "Video tracking: A concise survey," *IEEE Journal of Oceanic Engineering*, vol. 31, no. 2, pp. 520–529, 2006.

[3] R. Stolkin, A. Greig, M. Hodgetts, and J. Gilby, "An em/e-mrf algorithm for adaptive model based tracking in extremely poor visibility." *Image and Vision Computing*, vol. 26, no. 4, pp. 480–495, 2008.

[4] M. P. Kumar, P. H. Torr, and A. Zisserman, "Learning layered motion segmentations of video," *International Journal of Computer Vision*, vol. 76, no. 3, pp. 301–319, 2008.

[5] F.-H. Cheng and Y.-L. Chen, "Real time multiple objects tracking and identification based on discrete wavelet transform," *Pattern Recognition*, vol. 39, no. 6, pp. 1126–1139, 2006.

[6] R. Li, S. Yu, and X. Yang, "Efficient spatio-temporal segmentation for extracting moving objects in video sequences," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 3, pp. 1161–1167, Aug. 2007.

[7] A. Colombari, A. Fusiello, and V. Murino, "Segmentation and tracking of multiple video objects," *Pattern Recognition*, vol. 40, no. 4, pp. 1307–1317, 2007.

[8] C. Rasmussen and G. D. Hager, "Probabilistic data association methods for tracking complex visual objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 560–576, 2001.

[9] D. Cremers and S. Soatto, "Motion competition: A variational approach to piecewise parametric motion segmentation," *International Journal of Computer Vision*, vol. 62, no. 3, pp. 249–265, May 2005.

[10] H. Shen, L. Zhang, B. Huang, and P. Li, "A map approach for joint motion estimation, segmentation, and super resolution," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 479–490, 2007.

[11] N. Vaswani, A. Tannenbaum, and A. Yezzi, "Tracking deforming objects using particle filtering for geometric active contours," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 8, pp. 1470–1475, 2007.

[12] L. Wiskott, "Segmentation from motion: Combining Gabor- and Mallat-wavelets to overcome aperture and correspondence problem," in *Proceedings of the 7th International Conference on Computer Analysis*

of Images and Patterns, G. Sommer, K. Daniilidis, and J. Pauli, Eds., vol. 1296. Heidelberg: Springer-Verlag, 1997, pp. 329–336.

[13] M. Kong, J.-P. Leduc, B. Ghosh, and V. Wickerhauser, "Spatio-temporal continuous wavelet transforms for motion-based segmentation in real image sequences," *Proceedings of the International Conference on Image Processing*, vol. 2, pp. 662–666 vol.2, 4-7 Oct 1998.

[14] J. Zhang, F. Shi, J. Wang, and Y. Liu, "3d motion segmentation from straight-line optical flow," in *Multimedia Content Analysis and Mining*, 2007, pp. 85–94.

[15] J. P. Costeira and T. Kanade, "A multibody factorization method for independently moving objects," *International Journal of Computer Vision*, vol. 29, no. 3, pp. 159–179, 1998.

[16] N. Ichimura and F. Tomita, "Motion segmentation based on feature selection from shape matrix," *Systems and Computers in Japan*, vol. 31, no. 4, pp. 32–42, 2000.

[17] K. Kanatani and C. Matsunaga, "Estimating the number of independent motions for multibody motion segmentation," in *Proceedings of the Fifth Asian Conference on Computer Vision*, vol. 1, Jan 2002, pp. 7–12.

[18] M. Brand, "Incremental singular value decomposition of uncertain data with missing values," in *European Conference on Computer Vision*, 2002, pp. 707–720.

[19] L. Zelnik-Manor and M. Irani, "Degeneracies, dependencies and their implications in multi-body and multi-sequence factorizations," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. II–287–93 vol.2, 18-20 June 2003.

[20] H. Zhou and T. S. Huang, "Recovering articulated motion with a hierarchical factorization method," in *Gesture Workshop*, 2003, pp. 140–151.

[21] Y. Sugaya and K. Kanatani, "Geometric structure of degeneracy for multi-body motion segmentation," in *Statistical Methods in Video Processing*, 2004, pp. 13–25.

[22] L. Zelnik-Manor and M. Irani, "Temporal factorization vs. spatial factorization," in *European Conference on Computer Vision*, vol. 2. Springer Berlin / Heidelberg, 2004, pp. 434–445.

[23] A. Gruber and Y. Weiss, "Multibody factorization with uncertainty and missing data using the em algorithm," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. I–707–I–714 Vol.1, 27 June-2 July 2004.

[24] R. Vidal and R. Hartley, "Motion segmentation with missing data using powerfactorization and gpca," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. II–310–II–316 Vol.2, 27 June-2 July 2004.

[25] J. Yan and M. Pollefeys, "A general framework for motion segmentation: Independent, articulated, rigid, non-rigid, degenerate and non-degenerate," in *European Conference on Computer Vision*, 2006, pp. IV: 94–106.

[26] A. Gruber and Y. Weiss, "Incorporating non-motion cues into 3d motion segmentation," in *European Conference on Computer Vision*, 2006, pp. 84–97.

[27] X. Llado, A. D. Bue, and L. Agapito, "Euclidean reconstruction of deformable structure using a perspective camera with varying intrinsic parameters," *18th International Conference on Pattern Recognition*, vol. 1, pp. 139–142, 2006.

[28] A. Goh and R. Vidal, "Segmenting motions of different types by unsupervised manifold clustering," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–6, 17-22 June 2007.

[29] C. Julià, A. Sappa, F. Lumbreras, J. Serrat, and A. Lopez, "Motion segmentation from feature trajectories with missing data," in *Iberian Conference on Pattern Recognition and Image Analysis*, 2007, pp. I: 483–490.

[30] A. Bobick and J. Davis, "An appearance-based representation of action," *IEEE International Conference on Pattern Recognition*, pp. 307–312, 1996.

[31] I. Rekleitis, "Cooperative localization and multi-robot exploration," PhD in Computer Science, School of Computer Science, McGill University, Montreal, Quebec, Canada, 2003.

[32] S. Borman, "The expectation maximization algorithm – a short tutorial," Jul. 2004.

[33] B. K. Horn and B. G. Schunck, "Determining optical flow," Cambridge, MA, USA, Tech. Rep., 1980.

[34] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," in *International Conference on Computer Vision*, 1999, pp. 377–384.

[35] 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.