PTDC/EEA-ACR/72201/2006
Project MODI – 3D Models from 2D Images
Final Report

Pedro M. Q. Aguiar

February 2011

This document summarizes the research achievements under the FCT grant MODI – PTDC/EEA-ACR/72201/2006. This grant was warded according to a proposal addressing fundamental aspects of the problem of inferring three-dimensional (3D) models from images. Each of the following sections focuses on a specific aspect, pointing out the main contributions of the research team.

Dealing with occlusion – missing data

A major achievement of our work was the development of globally optimal method to deal with occlusion. In fact, the widely used SVD-based matrix factorization approaches to the recovery of 3D rigid structure from motion (SFM), require a set of feature points to be visible in a set of images. However, when there is occlusion, several feature points disappear, thus the observation matrix misses some entries. There is not equivalent to the SVD in this scenario and only suboptimal solutions have been proposed. We proposed a method to complete the trajectories that correspond to a rigid scene, in an optimal way. Our algorithm is not iterative (thus avoiding problems like sensitivity to initialization and local optima); it rather computes in a finite number of steps the globally optimal completion of the observation matrix [1, 2, 3, 4][1].

In the same topic, we also developed new iterative algorithms based on the Augmented Lagrangean method [5] and able to deal with particularly degenerated image sequences [6].

3D modeling of non-rigid scenes

When the 3D scene to recover from the images is non-rigid, the matrix factorization methods referred above have to incorporate a model for the 3D shape deformations. We have contributed to problem, by proposing approaches to deal

[1] Reference [2] was awarded the best paper prize in IEEE Int. Conf. on Image Processing'08.

with quadratic deformation models [7], stereo-motion [8], to model illumination changes using Spherical Harmonics [9], and to deal with the case of isometrically deformable flat surfaces [10, 11, 12]. While the majority of methods use the affine camera model, we also addressed the case of perspective cameras [13].

A particular case of non-rigid models are the articulated ones. We developed matrix factorization methods to deal with this kind of models [14]. The specific case of human body modeling has been the focus of many research teams, motivated by the obvious applications in medical diagnosis and sports research. We have developed matrix factorization techniques for this scenario, particularly when using also data from Motion Capture systems [15, 16, 17, 18, 19]

The topic of non-rigid modeling was also the motivation for a seminar/tutorial organized by one of the project members and two invited researchers [20].

Image matching, registration, and recognition

A key step in almost all approaches to the problem of inferring 3D from video is the image matching, or registration. Within the project, we developed efficient methods for point matching [21, 22]. When matching complex two-dimensional (2D) patterns, a common simplifying strategy is to work with 2D shapes, *i.e.*, binary images. When representing these 2D shapes, the challenge is to obtain, simultaneously, invariance with respect to transformations that may occur in practice, and discriminative power. This is usually referred by obtaining a *maximal invariant*. We proposed a new way to represent 2D shape that results maximally invariant to point permutation [23, 24, 25, 26, 27] and extended this representation to also accommodate maximal invariance with respect to geometric transformations, *e.g.*, rotation [28, 29, 30]. Naturally, the applications of these shape representations go beyond 3D modeling and include several image recognition tasks such as the ones illustrated in the references above.

We also dealt with the case of incorporating new views in previously acquired models [31] and proposed a new method to the necessary registration step, *i.e.*, matching 2D images with 3D models, using manifold projections [32].

Featureless methods

The attractiveness of research pathes that avoid the feature matching step was also mentioned in the project proposal. In these so-called featureless approaches, the information about the 3D shape of the scene come from single image cues, like textures or arrangements of line segments, rather than from multiple-view geometric constraints. Global voting schemes such as the Hough Transform (HT) have been widely used to robustly detect lines in images. However, they fail when the underlying images are cluttered, because the votes do not take connectivity into account. Local methods address this issue but lack robustness to deal with challenging situations, e.g., when line segments cross. We addressed the critical limitations of the HT as a line segment extractor by incorporating

connectivity in the voting process [33, 34]. Besides line segment detection, we also contributed to the problem of texture discrimination [35].

Conclusion

Our research under FCT grant MODI – PTDC/EEA-ACR/72201/2006 proceeded according to what was outlined in the project proposal. In what respects to the impact of the results, we would like to emphasize the new algorithm to factorize incomplete matrices in a globally optimal way. In fact, although the factorization with missing data has originated a huge number of papers in the recent past, only sub-optimal solutions (*i.e.*, iterative ones that converge to a local optimum) were proposed. In opposition, our approach computes the global optimum in a finite number of steps [1, 2, 3, 4]. We foresee a wide range of applications for our algorithm, since it can be used to extend any method where an SVD is used to approximate, in an optimal way, a rank deficient matrix, to the case of matrices with missing entries.

References

[1] P. Aguiar, M. Stosic, and J. Xavier. Spectrally Optimal Factorization Of Incomplete Matrices. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR2008)*, Anchorage AL, USA, 2008.

[2] P. Aguiar, J. Xavier, and M. Stosic. Globally Optimal Solution To Exploit Rigidity When Recovering Structure From Motion Under Occlusion. In *IEEE Int. Conf. on Image Processing (ICIP2008)*, S. Diego CA, USA, 2008.

[3] P. Aguiar, M. Stosic, and J. Xavier. On Singular Values of Partially Prescribed Matrices. *ELSEVIER Linear Algebra and its Applications*, 429(8-9), 2008.

[4] P. Aguiar, J. Xavier, and M. Stosic. MATLAB code for SPOC. Available from `http://www.isr.ist.utl.pt/~aguiar/spoc.m.`, 2009.

[5] A. Del Bue, J. Xavier, L. Agapito, and M. Paladini. Bilinear Factorization Via Augmented Lagrange Multipliers. In *European Conf. on Computer Vision (ECCV2010)*, volume 6314 of *Lecture Notes in Computer Science*, Crete, Greece, 2010. Springer-Verlag.

[6] M. Marques and J. Costeira. Estimating 3D Shape From Degenerate Sequences With Missing Data. *ELSEVIER Computer Vision and Image Understanding*, 113(2), 2009.

[7] J. Fayad, A. Del Bue, L. Agapito, and P. Aguiar. Non-Rigid Structure From Motion Using Quadratic Deformation Models. In *British Machine Vision Conf. (BMVC2009)*, London, UK, 2009.

[8] X. Lladó, A. Del Bue, and L. Agapito. Recovering Euclidean Deformable Models From Stereo-Motion. In *Int. Conf. on Pattern Recognition (ICPR2008)*, Tampa, USA, 2008.

[9] J. Buenaposada, A. Del Bue, E. Mu noz, and L. Baumela. A Model Of Brightness Variations Due To Illumination Changes and Non-Rigid Motion Using Spherical Harmonics. In *British Machine Vision Conf. (BMVC2008)*, Leeds, UK, 2008.

[10] R. Ferreira, J. Xavier, and J. Costeira. Shape From Motion Of Nonrigid Objects: The Case of Isometrically Deformable Flat Surfaces. In *British Machine Vision Conf. (BMVC2009)*, London, UK, 2009.

[11] R. Ferreira, J. Xavier, and J. Costeira. Reconstruction Of Isometrically Deformable Flat Surfaces in 3D from Multiple Camera Images. In *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP2009)*, Taipei, Taiwan, 2009.

[12] R. Ferreira. *Reconstruction Of Isometrically Embedded Flat Surfaces From Scaled Orthographic Image Data*. PhD thesis, Instituto Superior Técnico, Lisboa, Portugal, 2010.

[13] X. Lladó, A. Del Bue, and L. Agapito. Non-Rigid Metric Reconstruction From Perspective Cameras. *ELSEVIER Image and Vision Computing*, 28(9):1339 – 1353, 2010.

[14] M. Paladini, A. Del Bue, M. Stošic, M. Dodig, J. Xavier, and L. Agapito. Factorization For Non-Rigid And Articulated Structure Using Metric Projections. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR2009)*, Miami, Florida, 2009.

[15] J. Fayad, A. Del Bue, and P. Aguiar. Articulated Motion Analysis From Motion Capture Data. In *Int. Symp. on Computer Methods in Biomechanics and Biomedical Eng. (CMBBE2008)*, Porto, Portugal, 2008.

[16] J. Fayad. *Articulated 3D Human Modeling from Motion Capture Systems*. MSc thesis, Instituto Superior Técnico, Lisboa, Portugal, 2008.

[17] J. Fayad, A. Del Bue, and P. Aguiar. A Weighted Factorization Approach For Articulated Motion Modelling. In *ECCOMAS Multibody Dynamics Conf.*, Warsaw, Poland, 2009.

[18] J. Fayad, A. Del Bue, L. Agapito, and P. Aguiar. Human Body Modeling Using Quadratic Deformations. In *EUROMECH Solid Mechanics Conf.*, Lisboa, Portugal, 2009.

[19] J. Fayad, A. Del Bue, and P. Aguiar. Articulated and Deformable Motion Analysis from Motion Capture Data. In *Congresso Nacional de Biomecânica*, Bragança, Portugal, 2009.

[20] L. Agapito, A. Bartoli, and A. Del Bue. Computer Vision in a Non-Rigid World. Tutorial at Institute for Systems and Robotics / Instituto Superior Técnico, 2009.

[21] R. Cabral. *Joint Estimation of Correspondence and Motion using Global Rigidity and Local Descriptors.* MSc thesis, Instituto Superior Técnico, Lisboa, Portugal, 2009.

[22] J. Mota and P. Aguiar. Efficient Methods for Point Matching with Known Camera Orientation. In *Image Analysis and Recognition*, volume 6111 of *Lecture Notes in Computer Science*. Springer-Verlag, 2010.

[23] J. Rodrigues, P. Aguiar, and J. Xavier. ANSIG An Analytic Signature For Permutation-Invariant Two-Dimensional Shape Representation. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR2008)*, Anchorage AL, USA, 2008.

[24] J. Rodrigues, J. Xavier, and P. Aguiar. Classification of Unlabeled Point Sets using ANSIG. In *IEEE Int. Conf. on Image Processing (ICIP2008)*, S. Diego CA, USA, 2008.

[25] J. Rodrigues. *ANSIG – An Analytic Signature for Permutation-Invariant 2D Shape Representation.* MSc thesis, Instituto Superior Técnico, Lisboa, Portugal, 2008.

[26] J. Rodrigues, P. Aguiar, and J. Xavier. System and Method for Shape Recognition. Patent filed with Portuguese Patent Office, PT 104003, 2008.

[27] J. Rodrigues, J. Xavier, and P. Aguiar. ANSIG An Analytic Signature for Arbitrary 2D Shapes (or Bags of Unlabeled Points). Working paper – to be submitted.

[28] J. Crespo, G. Lopes, and P. Aguiar. Principal Moments for Efficient Representation of 2D Shape. In *IEEE Int. Conf. on Image Processing (ICIP2009)*, Cairo, Egypt, 2009.

[29] J. Crespo and P. Aguiar. The 2D Orientation is Unique Through Principal Moments Analysis. In *IEEE Int. Conf. on Image Processing (ICIP2010)*, Hong Kong, 2010.

[30] J. Crespo and P. Aguiar. Revisiting Complex Moments for 2D Shape Representation and Image Normalization. Submitted to IEEE Trans. on Image Processing, 2010.

[31] A. Del Bue. A factorization Approach To Structure From Motion With Shape Priors. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR2009).*

[32] A. Del Bue, M. Stosic, M. Dodig, and J. Xavier. 2D-3D Registration Of Deformable Shapes With Manifold Projection. In *Int. Conf. on Image Processing (ICIP2009).*

[33] R. Guerreiro and P. Aguiar. Incremental Local Hough Transform For Line Segment Extraction. Submitted to IEEE Int. Conf. on Image Processing (ICIP2011), 2010.

[34] R. Guerreiro and P. Aguiar. Segment Extraction By Connectivity-Enforcing Hough Transform. Working paper – to be submitted.

[35] R. Guerreiro and P. Aguiar. Learning Simple Texture Discrimination Filters. In *IEEE Int. Conf. on Image Processing (ICIP2010)*, Hong Kong, 2010.