Overview of 3D registration techniques including loop minimization for the complete acquisition of large manufactured parts and complex environments

Batlle E. , Matabosch C. and Salvi J.

Institut d'Informàtica i Aplicacions, University of Girona
Campus Montilivi, 17071 Girona, Spain

ABSTRACT

3D modelling is becoming an important research topic for visual inspection in automatic quality control. Through visual inspection it is possible to determine whether a product fulfills the required specifications or whether it contains surface or volume imperfections. Although some process such as color analysis can be achieved by 2D techniques, more challenging tasks such as volume inspection of large and complex objects/scenes may require the use of accurate 3D registration techniques. 3D Simultaneous Localization and Mapping has become a very important research topic not only in the computer vision community for quality control applications but also in the robotics field for solving problems such as robot navigation and registration of large surfaces. Although their techniques differ slightly depending on the application, both communities tend to solve similar problems by means of different approaches. This paper presents a survey of the techniques used by the robotics and computer vision communities in which every approach has been compared pointing out their pros and cons and their potential applications.

Keywords: Vision, Registration, Surface, Large, Three-dimensional, Quality control.

1. INTRODUCTION

Since early 80's, computer vision and image processing have become very important to improve visual inspection in industrial quality control. One of the first applications concerned the assessment of manufactured product correctness with respect to a known model by means of 2D imaging and using wide range of techniques starting from image morphology up to complex image filters. Nevertheless, it is still difficult to introduce a reliable quality control in terms of 3D information, especially when the objects to observe are complex. However, more challenging tasks can be faced nowadays thanks to the evolution of computer vision in the last decades. For instance, recent contributions in 3D registration permit quality assessment tasks based on the comparison of a manufactured product and its model in terms of volume. These techniques are not only focused in the detection of object imperfections (such as surface bumps, cracks and fissures) in tiny or medium size objects, but also in large scenarios such as the inspection of automotive and avionics manufactured parts, submerged parts of harbours and dams, and even the complete acquisition of ancient remains and large sculptures. These environments are usually inaccessible by human beings and require the use of complex robotic systems to acquire completely the object or scene to measure.

Thus far, robot navigation has been focused on 2D mapping in flat terrains and usually restricted to indoor structured scenarios.[1] Recently, the need to explore complex and unstructured environments has increased,[2] requiring 6DOF movement for dealing with the unevenness of the terrains and environment complexity. Therefore, 6DOF localization and 3D mapping has become an important research field in the robotics community. Besides, in the field of computer vision, the growing interest in 3D modelling of large objects and scenes has forced the scientific community to face new challenges with the aim of reducing the propagation error present in registration.[3] In both situations, the robot/camera pose is estimated in order to be used in a further alignment of the 3D map/surface. Although the techniques differ slightly depending on the application, both communities tend to solve similar problems by means of different approaches[4].[5]

*e-mail: (bbatlle,cmatabos,qsalvi)@eia.udg.es

In general, a good estimation of the initial position is always required independently of the approach or technique used. Hence, section 2 provides a classification of the most important methods used to obtain a coarse pose estimation, including inertial navigation, visual odometry and surface-to-surface based matching, among others. Then, fine pair-wise registration approaches such as the Iterative Closest Point are used to refine the alignment between two clouds of points, see section 3. Finally, any error accumulated between correlated views is minimized by means of cycles and overlapping regions common among the acquired views. Hence, section 4 discusses a new classification of these techniques including analytic methods such as bundle adjustment and the well known ICP multi-view approach, and statistical methods such as Simultaneous Localization And Mapping (SLAM), among others. These techniques are compared and discussed analyzing their pros and cons and potential applications. The article ends with conclusions.

2. COARSE ONE-TO-ONE POSE ESTIMATION

The initial pose can be obtained using two well-known approaches as shown in Table 1. The former is based on benefiting by using some sort of device: a) sensors, such as odometers, compasses or inertial systems; or b) mechanisms, such as rotating tables, robot arms or conveyors. The latter is based on directly analyzing the visual images (given by cameras) or the surface views (given by scanners) looking for correspondences which are used to solve the alignment and consequently the pose.

2.1. Initial pose estimation by mechanical devices

In the field of mobile robotics, odometry is the most extended position system. When dealing with 6DOF movement, inertial data must be added in order to provide the attitude variations and compensate the errors caused by navigating through rough terrains.[4] Some authors proposed to estimate the egomotion of the mobile robot by using a feature matching algorithm in which odometry is used to predict the region where the search of features starts.[2] Data provided by sensors is also used for Kalman Filter initialization in many SLAM approaches.[4]

Besides, when considering the 3D reconstruction of small objects in the field of computer vision, cameras and scanning lasers are mounted on rotating tables or mechanical structures in order to scan the whole object. Herein, the initial pose can be estimated by means of the calibration of these mechanisms and the use of encoders[5].[6]

2.2. Initial Pose estimation by computer vision

When sensors or mechanical devices can not be used or when their measure is rough or inaccurate, an estimation of the initial position by means of computer vision may be a good choice. Two main groups of techniques are proposed: a) Image-to-image correspondences, dealing with 2D images directly acquired by a stereo-head or a moving camera; and b) Surface-to-surface correspondences, dealing with 3D features or clouds of points acquired by any 3D acquisition technique such as stereo, laser triangulation or time-of-flight lasers, among others.

All these methods process the 2D/3D points of the given images/surfaces to extract significant points which are used in the matching process. Hence, the techniques are classified according to: a) feature-to-point approach when the significant points are only those that satisfy a given feature. Some of the most common methods are the corner detector proposed by Harris for 2D images[23] and the straight line-based method proposed by Stamos[3] for 3D views; and b) point-to-feature approach when an arbitrary group of points are characterized obtaining a set of features that differ one to another depending on point neighborhood. Points with similar features are potential matchings in the registration process. Here, some of the most used methods are the SIFT algorithm proposed by Lowe[12] for 2D images and the point signature proposed by Chua[17] for 3D views.

Image-to-image correspondences Some 3D pose estimation approaches are based on 2D image-to-image matching, which also concerns camera egomotion. Since 1980s, methods based on both the discrete and the differential epipolar constraint have been proposed. The discrete case is used mainly in self-calibration of stereo-heads (both monocular and binocular), whereas the differential case deals with a unique moving camera at high image rate. The discrete case is based on the so-called Essential matrix when the intrinsic camera parameters are known, or the Fundamental Matrix in the uncalibrated case.[24] The differential case is based on the optical flow

Table 1. Classification of coarse one-to-one pose estimation techniques.

Technique				author	DOF	sensor	scene
Coarse one-to-one pose estimation	mechanical devices	sensors		Nüchter, 2004[2]	6	TOF	outdoor
				Folkesson, 2003[4]	6R	TOF	outdoor
				Kohlhepp, 2004 [7]	6R	TOF	indoor
		mechanisms		Pulli, 1999[5]	6	LT	object
				Bernardini, 2002[6]	6	LT	object
	Computer vision	Image to image	Feature to point	Huang, 1989[8]	6	monocular	indoor
				Shang, 1998[9]	6	binocular	indoor
				Davison, 2003[10]	6	monocular	indoor
				Burschka, 2004[11]	6	monocular	outdoor
			Point to feature	Lowe, 1999[12]	6	binocular	indoor
				Se, 2002[13]	6R	trinocular	indoor
		Surface to surface	Point to feature	Chen, 1998[14]	6	DLP	object
				Johnson, 1999[15]	6	DLP	object
				Carmichael, 1999[16]	6	DLP	object
				Chua, 1997[17]	6	database	object
				Kim, 2002[18]	6	database	object
				Huber, 2003[19]	6	LT	object
			Feature to point	Nister, 2004[20]	6	monocular	outdoor
				Stamos, 2003[3]	6	TOF	outdoor
				Wyngaerd, 2003[21]	6	DLP	object
				Triebel, 2005[22]	6R	TOF	outdoor

R: Restricted (some DOF are constrained in a limited range); TOF: Time-of-flight; LT: Laser Triangulation; DLP: Digital Light Projector

and the differential epipolar constraint.[25] An early work was developed by Huang, who proposed a linear matching algorithm based on the Essential Matrix for determining 3D motion by using eight point correspondences in two views.[8] Another method for the motion estimation of a moving uncalibrated stereo ring was proposed by Zhang in 1996.[9] The proposal was based on computing the fundamental matrix and then the motion up to a scale factor was estimated by solving the well-known Kruppa equations computing a perspective reconstruction. The Euclidean reconstruction is obtained by taking any metric measure from the scene that allows the determination of the scale factor, usually a distance between two 3D features[26].[10]

In summary, techniques based on the discrete epipolar geometry have been widely studied and nowadays robust solutions are available even in 6DOF.[11] Besides, the differential movement estimators are quite sensitive to noise. Therefore, these methods are, in general, adapted to the application constraining the number of DOF with the aim of reducing the error in the estimation.

Note that image-to-image methods are commonly based on feature-to-point approaches. However, in 1999 Lowe proposed a new feature extraction algorithm called SIFT (Scale Invariant Feature Transform), which used has increased in the last years especially dealing with 2D images. The method is classified as a point-to-feature approach since the main contribution of the method is the characterization of significant points according to scale invariant features. However, the method first selects the significant points by using a feature-to-point approach, such as Harris Corner Detector, so both approaches are combined in the SIFT technique.[12]

Surface-to-surface correspondences Many authors proposed techniques that process directly the clouds of points provided by any 3D acquisition system. Here, the main difference is in the way of selecting the matching points.

As explained before, some approaches are based on searching for points that satisfy a given feature which are then used to solve the matching, so that we have classified them as feature-to-point. Herein, some authors propose to extract 2D features from the images, such as corners. Then, features are tracked over time to solve the matching with their corresponding 3D points already acquired, obtaining an initial estimate of the movement.

Table 2. Classification of fine one-to-one pose estimation techniques.

Technique		author	DOF	sensor	scene
Fine one-to-one pose estimation (Pair-wise)	Point to point	Besl, 1992[27]	6	LT	outdoor
		Greenspan, 2001[28]	6	DLP	object
		Jost, 2002[29]	6	database	object
		Guidi, 2004[30]	6	DLP	object
		Triebel, 2005[22]	6R	TOF	outdoor
		Trucco, 1999[31]	6	synthetic data	object
	Point to plane	Chen, 1991[32]	6	DLP	object
		Gagnon, 1994[33]	6	monocular	object
		Park, 2003[34]	6	database	object
		Pulli, 1999[5]	6	LT	object

R: Restricted (some DOF are constrained in a limited range); TOF: Time-of-flight; LT: Laser Triangulation; DLP: Digital Light Projector

Following this idea a new concept called Visual Odometry was introduced by Nister in 2004. The author proposed a new algorithm for stereo camera pose estimation based on the well known preemptive RANSAC algorithm.[20] Other techniques deal directly with 3D features extracted from the clouds of points, such as the straight line-based method proposed by Stamos[3] and the curved line-based method proposed by Wyngaerd.[21]

Other approaches are based on characterizing the points by using their neighborhood information to obtain a set of features. Then, these features are sought within the cloud of points in subsequent views to solve the matching, so that we have classified them as point-to-feature. Two of the most used approaches are the Spin Image and the Point Signature. Spin Image is based on projecting neighboring points onto a plane tangent at a given point obtaining a 2D Image (feature) for that point. Then, a region around the given point is considered in which two distances are computed to determine the spin image.[15] A variant of this method was proposed by Carmichael, carrying out an interpolation of a set of points inside every triangular mesh with the aim of normalizing the number of points in every spin image.[16] Point Signature was proposed by Chua in 1997.[17] A given point is characterized by computing the distance to all the neighbors, obtaining a distance vector. This vector is then compared with the points in the second view to find correspondences. Other authors proposed the use of the Principal Component Analysis in order to estimate the main axis of the whole cloud of points. Subsequent views are aligned assuming that the main axis in two consecutive acquisitions do not vary significantly.[18]

Overall, the main problem of most point-to-feature algorithms is the computing time involved in obtaining a solution. This is because once some points in the first surface are characterized, they must be compared with all the points in the second surface in order to find correspondences. Feature-to-point approaches are more selective, reducing the potential matching points to only those that satisfy a given feature, so that the computing time is drastically reduced. Besides, the accuracy of the registration highly depends on the selected feature which should ensure that the selected points are really significant and also widely spread in the whole view.

3. FINE ONE-TO-ONE POSE ESTIMATION

Once an initial 3D pose is estimated by any coarse registration technique, an iterative minimization should be applied to obtain a refined pose and hence a better alignment between both views. Herein, the methods are classified according to the minimization function, which is usually the distance between corresponding points (point-to-point) or the distance between points and their corresponding plane (point-to-plane) as shown in table 2 and discussed in the following paragraphs.

3.1. Point-to-point alignment

Point-to-point alignment such as the Iterative Closest Point (ICP) focus on finding the distance between point correspondences.[27] When an initial estimation is known, all points are transformed to a reference frame. Then, every point in the first image is taken into consideration to search for its closest point in the second image. A new motion is estimated by the minimization of the distances between corresponding points, and the process is iterated until convergence. Some modifications of ICP have been presented in recent years to improve the efficiency of the algorithm[28,22] and decrease the computing time.[29] In addition, other authors proposed to increase the robustness of ICP by removing correspondences whose distances are higher than a threshold.[31]

Overall, ICP is the most common fine registration method and the results provided by authors are good. However, the method can not cope with non-overlapping regions because outliers are barely removed. In addition, this method usually presents problems of convergence, many iterations are required and, in some cases, the algorithm converges to local minima. Moreover, unless a robust implementation is used, the algorithm can only be used in surface-to-model registration.[35]

3.2. Point-to-plane alignment

The algorithm proposed by Chen[32] is an alternative to ICP. Summarizing, given a point in the first image, the intersection of the normal vector at this point with the second surface determines a second point in which the tangent plane is computed. The distance between this plane and the initial point is the function to minimize. Despite the difficulty of determining the intersection point between a line and a plane in a cloud of points, some techniques are presented to speed this process up.[33] Compared to ICP, this method is more robust to local minima and, in general, better results are obtained. Moreover, the method is less influenced by the presence of non-overlapping regions. The reason is that only the control points whose normal vector intersects the second view are considered in the matching, differing from ICP, where all points in the first cloud are used in the registration. Moreover, Chen's approach usually requires fewer iterations compared to ICP.

4. CYCLE MINIMIZATION

One-to-one alignment of views in a sequence causes a drift that is propagated throughout the sequence. Hence, some techniques have been proposed to reduce the propagating error benefiting from the existence of cycles and re-visited regions and considering the uncertainty in the alignment. This sort of techniques is classified into analytic and statistic, as shown in Table 3 and explained in the following paragraphs.

4.1. Analytic minimization

In order to minimize the propagating error, some authors have improved their algorithms by adding a final step that aligns all the acquired views at the same time. This approach spreads one-to-one pair-wise registration errors throughout the sequence of views, being known as multi-view registration. Early approaches proposed the aggregation of subsequent views in a single metaview which is progressively enlarged each time another view is registered.[32] Here, the main constraint is the lack of flexibility to re-register views already merged in the metaview. In 1999, Pulli proposed an ICP relaxation method based on the previous metaview approach but considering all the potential alignments between views before proceeding with the multi-view registration. In addition this method takes into account the information of all the overlapping areas and the already registered regions can be analyzed again for further transformations.[5] Later on, N*ü*chter proposed a global relaxation method based on Pulli's approach with the main difference that no iterative pair-wise alignment is required. However the success of this method depends on a correct pose estimation of the vehicle.[2]

A different approach was proposed by Bergevin,[36] who presented a multi-view registration technique based on the graph theory: views are associated to nodes and transformations to edges. The authors consider all views as a whole and align all them simultaneously. The same idea was proposed later on by Silva[37] and Huber.[19] Besides, Masuda presented a multi-view registration algorithm based on the Matching Signed Distance Fields in which outliers are automatically removed obtaining a more robust method.[38] Overall, multi-view techniques suffer two main drawbacks: a) the whole set of 3D views have to be previously acquired before the algorithm starts; b) an accurate estimation of the motion between views is needed as initial guesses to ensure convergence. Thus, multi-view techniques are not considered for real-time applications such as mobile robot navigation.

Few authors have faced the challenge of registering 3D views in a sequence while they are acquired avoiding or at least controlling error propagation. For instance, Sharp[39] proposed the registration of pairs of consecutive views until a cycle is found. Since only pair-wise registration is required, the method becomes very fast. Here, the

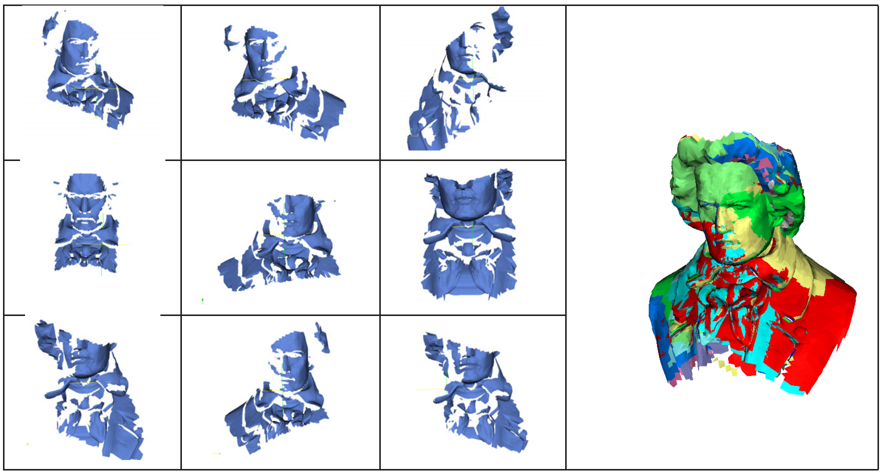

Figure 1. Example of the multi-view registration of multiple 3D views of Beethoven out in our lab.

interest is the way of distributing the motion (and hence the propagation error) among the different views. The author proposed to use weights directly related to the residue obtained in the pair-wise registration. Actually, this is not very accurate especially in the presence of misalignments between end views in the cycle as a matter of noise and object occlusions. In this case, the whole motion of such a cycle is also distributed to all the views increasing the error in the registration.

In the last few years, a photogrammetric technique called Bundle Adjustment has increased popularity in the computer vision community and it is growing in interest in robotics. Bundle adjustment is the problem of refining a visual reconstruction to produce jointly optimal 3D structure and viewing parameters (camera pose and/or calibration) estimates.[40] Therefore, bundle adjustment techniques can be used in both robot/camera localization and 3D mapping in many fields such as camera calibration, robot navigation and scene reconstruction. Since bundle adjustment is a non-linear minimization problem, it is solved by means of iterative non-linear least squares or total squares methods such as Levenberg-Marquard or M-estimator techniques[41].[42] Although bundle adjustment is commonly classified as a multiview technique, some authors have used it in consecutive pairwise alignment as a technique to reduce error propagation.[43]

In summary, we conclude that analytic methods based on the metaview approaches present good results when initial guesses are accurate and the surface to be registered does not have a large scale. Otherwise, the method suffers a large propagation error producing drift and misalignments and its greedy approach usually falls in local minima. The use of methods based on graphs has the advantage of minimizing the error in all the views simultaneously but they usually require a previous pairwise registration step, which accuracy can be determinant in the global minimization process. Besides, closing the loop strategies provide trustworthy constraints for error minimization but require a huge amount of memory and usually involve a high computational cost. Bundle adjustment techniques provide good results in the presence of outliers, but need a good enough initial guess and it is hardly used in large robot missions or large scale objects.

4.2. Statistic minimization

The same problem of registering 3D views in a sequence has been also faced by means of a probabilistic approach, especially in mobile robot navigation. The technique receives the name of Simultaneous Localization and Mapping (SLAM) since both the pose and the structure of the environment are estimated simultaneously. The main difference compared to analytic multi-view is that the uncertainty in the measure is not neglected. Hence, two main groups of techniques have been considered depending on the way of representing such uncertainty: a) Gaussian filters and b) non-parametric filters; which are discussed in the following paragraphs.

Both Kalman Filter (KF) for linear systems and Extended Kalman Filter (EKF) for non-linear systems are undoubtedly the most well-known Gaussian filters. Both consist in two main steps: a) Prediction, which

Table 3. Classification of cycle minimization techniques.

Technique			author	DOF	sensor	scene
Cycle minimization	Analytic (Multiview)	Iterative lineal	Bergevin, 1996[36]	6	monocular	object
			Huber, 2003[19]	6	LT	object
			Pulli, 1999[5]	6	LT	object
			Sharp, 2004[39]	6	DLP	indoor
		robust	Nüchter, 2004[2]	6	TOF	outdoor
			Masuda, 2001[38]	6	LT	object
			Silva, 2003[37]	6	database	object
			Pollefeys, 2000[43]	6	monocular	outdoor
	Statistic	Gaussian	Deans, 2000[44]	6R	database	indoor
			Guivant, 2000 [45]	6	TOF	outdoor
			Martinelli,2005 [46]	6R	TOF	indoor
			Liu, 2003 [47]	6R	TOF	outdoor
			Bosse, 2003[48]	6	TOF	outdoor
			Estrada, 2003[49]	6R	TOF	outdoor
		Non Parametric	Davison, 2003[10]	6	monocular	indoor
			Montemerlo, 2002 [50]	6R	TOF	outdoor
			Früh[51]	6R	TOF	outdoor

R: Restricted (some DOF are constrained in a limited range); TOF: Time-of-flight; LT: Laser Triangulation; DLP: Digital Light Projector

estimates the current state by using the temporal information of previous states; and b) Update, which uses the current information provided by robot on-board sensors to refine prediction. Whenever a landmark is observed by the on-board sensors of the robot, the system determines whether it has been already registered and updates the filter. Hence, when part of the scene is revisited, all the gathered information from past observations is used by the system to reduce the uncertainty in the whole mapping, strategy known as closing the loop.

Besides, mobile robot localization and mapping has also been tackled by using non-parametric filters such as histogram filter or particle filter. The main advantage compared to Gaussian filters is the possibility of dealing with multimodal data distribution, so that multiple values (particles) are used to represent the belief.[52] Nevertheless, note that Gaussian filters have a polynomic computational cost whereas the computational cost of a non-parametric filter may be exponential.

Another interesting approach related to SLAM arises in the presence of data provided by bearing-only sensors such as omnidirectional or unique moving cameras. Since depth information is not provided, EKF can not be directly initialized, leading to a new challenge known as Bearing-Only SLAM. An early approach was proposed by Deans,[44] who combined Kalman filter and bundle adjustment in filter initialization, obtaining accurate results at the expense of increasing filter complexity. Besides, Davison[10] proposed a top-down Bayesian framework for unique moving camera localization based on a particle filter, which benefits in the initialization of using a A4 piece of paper as a landmark to recover metric information of the scene. Then, whenever a scene landmark is observed a set of depth hypothesis are made along its direction. In subsequent steps, the same landmark is seen from different positions reducing the number of hypothesis and leading to an accurate landmark pose estimation. Recently, Lemaire[53] proposed a 3D Bearing-Only SLAM algorithm based on EKF filters, in which each feature is represented by a sum of Gaussians.

In the presence of large environments in which tons of data are gathered, Gaussian filters state vectors increase considerably leading to inefficiency in terms of computational cost. Similar problems arise using non-parametric filters such as the particle filter. Hence, some authors have proposed different techniques to cope with computational cost and memory size. For instance, the Compressed Extended Kalman Filter (CEKF) permits to update only a part of the map at every step.[45] Some techniques try to reduce complexity doing an accurate selection of the landmarks and discarding the unnecessary ones, decreasing the filter size.[46] Other approaches such as the Sparse Extended Information Filter (SEIF),[47] Fast SLAM[50] and Atlas[48] use graph based techniques to reduce the complexity. The main difference of Atlas approach compare to other methods is the use of local

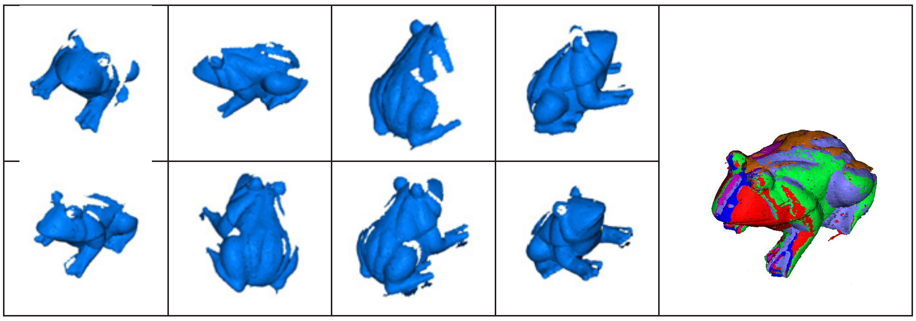

Figure 2. Multi-view registration of multiple 3D views of a ceramic frog out in our lab.

maps instead of working with a unique global map. Although, some of these techniques use global constrains, they do not impose loop consistency producing mapping misalignments in large environments. Hence, following the same idea, both Graphical SLAM[54] and Hierarchical SLAM[49] include loop consistency constraints leading to an accurate mapping.

Overall, we conclude that SLAM methods offer accurate solutions for localization and mapping in reduced environments. Since PF methods can handle with multiple hypothesis they present more robustness in periods of global uncertainty and sensor noise, but are less efficient than EKF in terms of computational costs. However, dealing with large environments both methods present problems associated to the increasing uncertainty and the huge amount of data treated. This drawback can be solved by using methods based on building submaps such as Atlas, which present more robustness against uncertainty compared to methods based on a unique global map. Some methods impose global restrictions for global map joining, providing accurate solutions in the presence of short loops. However, loop consistency constraints used in methods such as Hierarchical SLAM can be essential in order handle larger loops and prevent inconsistency and misalignments in the final map.

5. CONCLUSIONS

This paper presents a state of the art of the most representative techniques for 6DOF pose estimation and 3D registration of large objects and maps. The most referenced articles over the last few decades have been discussed analyzing their pros and cons and potential applications.

Related to coarse pose estimation, surveyed methods are classified into two main groups. The first benefits from sensors or other mechanical devices while the second focuses on computing the initial pose by solving the matching problem in both images and surface views. Although methods based on mechanical devices provide good results in flat terrains, a combination of both methods is usually required in the presence of rough and unstructured environments. Furthermore, it has been observed that image-to-image alignment presents good results in the presence of nearly planar areas where depth can be neglected. Otherwise, the alignment produces artifacts ruining the registration. Therefore, surface-to-surface alignment is more adequate for 3D scenarios, but then we have to avoid symmetries in the views to obtain accurate registrations.

Once a coarse pose between two views is estimated, a refinement step can be applied in order to provide a more accurate alignment. Two main methods and their variants have been discussed: Point-to-point and Point-to-plane. Although the first method is the most commonly used, a huge amount of iterations is required and the method may converge to a local minima. Besides, the point-to-plane method has demonstrated to work better in the presence of non-overlapping regions and usually converges faster. In addition, point-to-point is the most used in surface-to-model registration, while point-to-plane is the most accurate in surface-to-surface registration.

Regarding the minimization of the propagation error, analytic methods are the most common in high-resolution object reconstruction by means of multi-view registration techniques. Although multi-view registration methods have demonstrated to provide accurate solutions, misalignments can appear in the presence of featureless environments, symmetries and smooth objects. Note that an accurate treatment of outliers and the removal of the less confident paths in the graph are also compulsory steps to ensure an accurate registration.

Besides, statistical methods are the most used in 3D mapping in mobile robot navigation. The advantage of statistical methods is in their performance in the presence of less reliable sensors, complex environments and unstructured scenes with few features and landmarks. However, they are not recommended for handling tons of data since the manipulation of large state vectors derives to an inefficient computation. Finally, it is interesting to point out that non-parametric methods are more suitable than Gaussians when the mobile robot is in phases of global uncertainty, despite the high computational cost.

This article is intended to be a guide for any researcher interested in the field. To the best of our knowledge, this article is the first that compares the techniques present in both robotics and computer vision communities, providing new classification criteria, discussing the existing techniques, and pointing out their pros and cons and potential applications.

ACKNOWLEDGMENTS

This research has been supported by Spanish Project TIC2003-08106-C02-02.

REFERENCES

1. D. Tardos, J. Neira, P. Newman, and J. Leonard, "Robust mapping and localization in indoor environments using sonar data," *The International Journal of Robotics Research.* **21**, pp. 311 – 330, December 2002.
2. A. Nüchter, H. Surmann, K. Lingemann, J. Hertzberg, and S. Thrun, "6d slam with an application in autonomous mine mapping," in *IEEE International Conference on Robotics and Automation*, **2**, pp. 1998 – 2003, ICRA '04, May 2004.
3. I. Stamos and M. Leordeanu, "Automated feature-based range registration of urban scenes of large scale," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, **2**, pp. 555–561, 2003.
4. J. Folkesson and H. Christensen, "Outdoor exploration and slam using a compressed filter," in *IEEE International Conference on Robotics and Automation*, **1**, pp. 419–426, 2003.
5. K. Pulli, "Multiview registration for large data sets," in *International Conference on 3-D Digital Imaging and Modeling*, pp. 160 – 168, October 1999.
6. F. Bernardini, I. Martin, J. Mittleman, H. Rushmeier, and G. Taubin, "Building a digital model of michelangelo's florentine pietà," in *IEEE Computer Graphics and Applications*, **22**, pp. 59–67, Nov 2002.
7. P. Kohlhepp, P. Pozzo, M. Walther, and R. Dillmann, "Sequential 3d-slam for mobile action planning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, **1**, pp. 722 – 729, September 2004.
8. T. Huang and O. Faugeras, "Some properties of the e matrix in two-view motion estimation," *Pattern Analysis and Machine Intelligence* **11**, pp. 1310 – 1312, December 1989.
9. Z. Zhang, Q.-T. Luong, and O. Faugcras, "Motion of an uncalibrated stereo ring: Self-calibration and metric reconstruction," in *IEEE Transactions on Robotics and Automation*, pp. 103–113, 1996.
10. A. Davison, W. Mayol, and D. Murray, "Real-time localization and mapping with wearable active vision," in *Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reallity*, pp. 18–27, 2003.
11. D. Burschka and G. Hager, "V-gps(slam): Vision-based inertial system for mobile robots," in *IEEE International Conference on Robotics and Automation*, **1**, pp. 409 – 415, 2004.
12. D. G. Lowe, "Object recognition from local scale-invariant features," in *International Conference on Computer Vision ICCV*, pp. 1150–1157, (Corfu, Greece), September 1999.
13. S. Se, D. Lowe, and J. Little, "Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks," *The International Journal of Robotics Research* **21**(8), pp. 735–758, 2002.
14. C.-S. Chen, Y.-P. Hung, and J.-B. Cheng, "A fast automatic method for registration of partially overlapping range images," in *International Conference on Computer Vision*, pp. 242–248, (Bombay), January 1998.
15. A. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3d scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**, pp. 433–449, May 1999.
16. O. Carmichael, D. Huber, and M. Hebert, "Large data sets and confusing scenes in 3-d surface matching and recognition," in *International Conference on 3-D Digital Imaging and Modeling*, pp. 258–367, (Ottawa, Ont. , Canada), Oct. 1999.

17. C. Chua and R. Jarvis, “Point signatures: A new representation for 3d object recognition,” *International Journal of Computer Vision* **25**, pp. 63–85, October 1997.
18. S. Kim, J. Seo, H. H.K, C. Jho, and M. Choi, “Iterative registration of multiple 3d data sets using covariance matrix,” in *The 8th International Conference on Virtual Systems and Multimedia*, September 2002.
19. D. Huber and M. Hebert, “Fully automatic registration of multiple 3d data sets,” *Image and Vision Computing* **21**, pp. 637–650, July 2003.
20. D. Nister, O. Naroditsky, and J. Bergen, “Visual odometry,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, **1**, pp. 652–659, July 2004.
21. J. Wyngaerd, “Combining texture and shape for automatic crude patch registration,” in *Int. Conference on3-D Digital Imaging and Modeling*, pp. 179 – 186, 2003.
22. R. Triebel and W. Burgard, “Improving simultaneous mapping and localization in 3d using global constraints,” in *Proceedings of the National Conference on Artificial Intelligence*, **3**, pp. 1330–1335, 2005.
23. C. J. Harris and M. Stephens, “A combined corner and edge detector,” in *Fourth Alvey Vision Conferences*, pp. 147–151, 1988.
24. X. Armangué and J. Salvi, “Overall view regarding fundamental matrix estimation,” in *Image and Vision Computing*, **21**, pp. 205–220, February 2003.
25. H. Armangué, X. Araújo and J. Salvi, “A review on egomotion by means of differential epipolar geomety applied to the movement of a mobile robot.,” in *Pattern Recognition*, **12**, pp. 2927–2944, December 2003.
26. P. Sturm, “Critical motion sequences for monocular self-calibration and uncalibrated euclidean reconstruction.,” in *Computer Vision and Pattern Recognition.*, pp. 1100 – 1105, June 1997.
27. P. Besl and N. McKay, “A method for registration of 3-d shapes,” in *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **14**, pp. 239–256, Feb. 1992.
28. M. Greenspan and G. Godin, “A nearest neighbor method for efficient icp,” in *Third International Conference on 3-D Digital Imaging and Modeling*, pp. 161–168, (Quebec, Canada), May-June 2001.
29. T. Jost and H. Hugli, “A multi-resolution scheme icp algorithm for fast shape registration,” in *First International Symposium on 3D Data Processing Visualization and Transmission*, pp. 540– 543, 2002.
30. G. Guidi, J.-A. Beraldin, and C. Atzeni, “High-accuracy 3-d modeling of cultural heritage: The digitizing of donatello’s ”maddalena”,” in *IEEE Transactions on Image Processing*, **3**, pp. 370–380, 2004.
31. E. Trucco, A. Fusiello, and V. Roberto, “Robust motion and correspondences of noisy 3-d point sets with missing data,” *Pattern Recognition Letters* **20**, pp. 889–898, September 1999.
32. G. Chen, Y. ad Medioni, “Object modeling by registration of multiple range images,” in *IEEE Int. Conf. on Robotics and Automation*, **3**, pp. 2724 –2729, April 1991.
33. H. Gagnon, M. Soucy, R. Bergevin, and D. Laurendeau, “Registration of multiple range views for automatic 3-d model building,” in *Computer Vision and Pattern Recognition*, pp. 581–586, June 1994.
34. S.-Y. Park and M. Subbarao, “A fast point-to-tangent plane technique for multi-view registration,” in *Int. Conference on 3-D Digital Imaging and Modeling*, pp. 276 – 283, 2003.
35. C. Matabosch, J. Salvi, X. Pinsach, and R. Garcia, “Surface registration from range image fusion,” in *IEEE International Conference on Robotics and Automation*, **1**, pp. 678 – 683, (New Orleans, USA), April 2004.
36. R. Bergevin, M. Soucy, H. Gagnon, and D. Laurendeau, “Towards a general multi-view registration technique,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18**(5), pp. 540–547, 1996.
37. L. Silva, O. Bellon, and K. Boyer, “Enhanced, robust genetic algorithms for multiview range image registration,” in *3DIM03. Fourth International Conference on 3-D Digital Imaging and Modeling*, pp. 268–275, 2003.
38. T. Masuda, “Generation of geometric model by registration and integration of multiple range images,” in *Third International Conference on 3-D Digital Imaging and Modeling*, pp. 254 –261, May 2001.
39. G. Sharp, S. Lee, and D. Wehe, “Multiview registration of 3d scenes by minimizing error between coordinate frames,” in *Pattern Analysis and Machine Intelligence*, pp. 1037 – 1050, 2004.
40. B. Triggs, P. McLauchlan, and A. Hartley, R. Fitzgibbon, “Bundle adjustment - a modern synthesis,” in *International Workshop on Vision Algorithms*, pp. 298 – 372, (London, UK), 2000.
41. A. Fitzgibbon, “Robust registration of 2d and 3d point sets,” in *The British Machine Vision Conference*, pp. 662–670, 2001.

42. J. Salvi, C. Matabosch, D. Fofi, and J. Forest, "A review of recent range image registration methods with accuracy evaluation," *Image and Vision Computing* , 2006.
43. M. Pollefeys, M. Koch, M. Vergauwen, and V. G. L., "Automated reconstruction of 3d scenes from sequences of images," *Photogrammetry and Remote Sensing* (55), pp. 251–267, 2000.
44. M. Deans and M. Hebert, "Experimental comparison of techniques for localization and mapping using a bearing-only sensor," in *Seventh International Symposium on Experimental Robotics*, **271**, pp. 395 – 404, (Honolulu, Hawaii), December 2000.
45. J. E. Guivant and E. M. Nebot, "Optimization of the simultaneous localization and map building algorithm for real time implementation," *IEEE Transactions on Robotics* **3**(17), pp. 242–257, 2000.
46. A. Martinelli, N. Tomatis, and R. Siegwart, "Some results on slam and the closing the loop problem," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2917– 2922, (Lausanne, Switzerland), Aug 2005.
47. Y. Liu and S. Thrun, "Results for outdoorslam using sparse extended information filters," in *IEEE International Conference on Robotics and Automation*, **1**, pp. 1227– 1233, (USA), Sept 2003.
48. M. Bosse, P. Newman, J. Leonard, M. Soika, W. Feiten, and S. Teller, "An atlas framework for scalable mapping," in *IEEE International Conference on Robotics and Automation*, **2**, pp. 1899– 1906, (Amherst, MA, USA), Sept 2003.
49. C. Estrada, J. Neira, and J. Tardos, "Hierarchical slam: real-time accurate mapping of large environments," *IEEE Transactions on Robotics* **21**(4), pp. 588–596, 2005.
50. M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit, "Fastslam: A factored solution to the simultaneous localization and mapping problem," in *National Conference on Artificial Intelligence*, pp. 593–598, (Vancouver, BC), July 2002.
51. C. Früh and A. Zakhor, "Constructing 3d city models by merging aerial and ground views," *IEEE Computer Graphics and Applications* **23**, pp. 52 – 61, Nov 2003.
52. W. F. D. Thrun, S. Burgard, *Probabilistic Robotics*, 2005.
53. T. Lemaire, S. Lacroix, and J. Sola, "A practical 3d bearing-only slam algorithm," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 2449– 2454, Aug 2005.
54. J. Folkesson, P. Jensfelt, and H. Christensen, "Graphical slam using vision and the measurement subspace," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3383–3390, (Edmundton, Canada), Aug 2005.