Towards exploiting the advantages of colour in scan matching

Fernando Martín⋆, Jaime Valls Miró◇, and Luis Moreno⋆

⋆Carlos III University, Madrid, Spain,
{fmmonar,moreno}@ing.uc3m.es
http://roboticslab.uc3m.es
◇University of Technology, Sydney, Australia,
{Jaime.VallsMiro}@uts.edu.es
http://ims.uts.edu.au/

Abstract. Colour plays an important role in the perception systems of the human beings. In robotics, the development of new sensors has made it possible to obtain colour information together with depth information about the environment. The exploitation of this type of information has become more and more important in numerous tasks. In our recent work, we have developed an evolutionary-based scan matching method. The aim of this work is to modify this method by the introduction of colour properties, taking the first steps in studying how to use colour to improve the scan matching. In particular, we have applied a colour transition detection method based on the *delta E* divergence between neighbours in a scan. Our algorithm has been tested in a real environment and significant conclusions have been reached.

Keywords: Differential Evolution, Scan Matching, RGB-D, Colour Properties, Delta E

1   Introduction

The information provided by colours is a crucial factor in human perception. If we are approaching a dark place, it is not easy to distinguish between objects and free space. If we see fire in a house, it is a signal of danger. If we are walking in the beach, the shore can be distinguished by the blue of the sea. One traditional task in robotics consists of building maps using depth sensors such as laser range finders. This task is commonly referred to as mapping. If the sensor works in three dimensions (3D), it is possible to obtain a volumetric representation. In the last years, the development of cheap new sensors that provide depth and colour information, such as the Microsoft Kinect, has made it possible to utilize colour properties when building volumetric maps. Although colour properties have been widely used in fields such as computer vision, it is not common to use them when building volumetric maps with depth information. The objective of this work is to make an initial study about how to exploit the advantages of colour in a specific application (scan matching).

Mapping is closely related to the well-known Simultaneous Localization and Mapping (SLAM) problem, originally introduced by Leonard and Durrant-White [1] basing on an earlier work by Smith *et al.* [2]. The SLAM problem for mobile robots consists of building a map of an unknown environment while exploring the environment at the same time, using this map.

The work reported in this paper is focused on the scan matching (or registration) problem, which is a particular aspect of the mapping task. The scan matching can be defined as the estimation of the metric relation (position and orientation) between a pair of scans, using the information provided by these scans. This is a very common problem in mobile robotics because the information about the environment that is going to be represented is often given by an exteroceptive sensor with depth information.

A scan matching algorithm for 3D environments has been created in our recent work [3]. The core of this technique is the Differential Evolution (DE) method, developed by Storn and Price [4], which is a particle-based evolutionary algorithm that evolves in time to the solution that minimizes a fitness function. If the cost function is chosen in an adequate way, the scan matching problem can be solved.

The scan matching methods traditionally have failures when there is an important change in orientation between scans. In these cases, an initialization algorithm like RANSAC has been used to solve this problem [5]. There are some colour properties that are more invariant to rotation than the raw 3D data. An initial idea is to try to improve our DE-based scan matching algorithm in order to increase the robustness under sharp turns.

In this work, we incorporate colour properties into the cost function of the scan matching method. This can be done because the sensors obtain RGB information together with depth information. In particular, we propose to extract the colour transitions, which are represented by neighbours with a significant change in colour, and to use these points of interest to perform the scan matching. Due to its interesting properties, the *delta E* divergence (unit of measurement proposed by the International Commission on Illumination, CIE in french) has been chosen as a suitable measurement to obtain these points of interest.

The objective of this work is to take the first steps in studying how to use colour to improve our method. The new technique has been tested and compared to the original version in a real environment, and significant conclusions have been reached.

1.1   Related work

SLAM is one of the most popular topics in mobile robotics. The interest about 3D maps has increased within the recent past due to the availability of efficient 3D sensors. For example, Nuchter *et al.* [6] have proposed a technique based on the alignment of 3D scans using the Iterative Closest Points [7] (ICP) scan matching method combined with a heuristic for loop detection and a global relaxation method. Hähnel *et al.* [8] have proposed an algorithm "for full 3d shape reconstruction of indoor and outdoor environments with mobile robots".

Triebel *et al.* [9] have built Multi-Level Surface (MLS) maps. Cole *et al.* [10] have proven that the traditional methods used to solve the SLAM problem in planar environments can also be extended to perform 6D SLAM in more difficult conditions, e.g. undulating outdoor areas. Magnusson and Duckett [11] have designed an alignment procedure based on the Normal Distributions Transform (NDT) [12].

The scan matching techniques can work with 2D or 3D maps. There are also local [13] and global [14] methods. While the local methods match single scans, the global methods consider the current scan and the global model. It is also possible to distinguish between feature-based, point-based, or mixed approaches. The first option requires a feature extraction before the scan matching. The point-based approach does not require any distinguishable structure in the environment and the mixed method seeks the correspondence between points and features.

The ICP method [15], which is the most common scan matching technique, is an algorithm that is used to minimize the spatial distance between two scans. This method is quite simple and its computational cost makes it possible to use it in real-time. It receives two clouds of points, an initial estimate of the translation and rotation, and the stopping criteria. It iteratively estimates the transformation (translation and rotation) that minimizes the distance between the clouds. Besl and Mac Kay [7] have implemented this method to register 3D shapes.

Many variants of this method have been proposed. A very interesting comparison of several methods depending on different parameters can be found in the work by Rusinkiewicz *et al.* [16]. The cost function is changed in order to match points that belong to similar surfaces in the variant proposed by Triebel *et al.* [9]. Bosse and Zlot [17] have proposed an improvement based on "the addition of robust optimization techniques to handle outliers and imperfect correspondences between the data". A traditional improvement applied to ICP is to incorporate additional information to the points of the clouds [18, 19].

The Iterative Matching Range Point (IMRP) method, proposed by Lu and Milios [13], is based on the limitation of the maximum translation and rotation. The Iterative Dual Correspondence (IDC) method, also proposed by Lu and Milios [13], combines ICP and IMRP. The translation is computed by the ICP method and the rotation is estimated by the IMRP method. The Polar Scan Matching (PSM) method [20] does not need to find correspondences between points. It assumes that model and data are sorted in the same way and only points with the same bearing are matched. Thrun *et al.* [21] consider that the free space in the current model will remain free in the future. They increase the information that is extracted from the laser scan.

There are not many researchers who have played attention to visual information in scan matching. Henry *et al.* [5] have extracted visual features to establish correspondences in their scan matching method. Ramos *et al.* have created the CRF-matching [22], which uses conditional random fields for feature-based scan matching. May *et al.* [19] have used laser reflectance values to improve the ICP.

However, they do not take full advantage of the application of colour-based features.

An evolutionary-based scan matching algorithm for 3D environments has been developed in our recent work. In this paper, this method has been improved in order to exploit colour properties in mapping. Our method tries to solve the same problem than ICP, which is the minimization of the distance between two point clouds. It also searches corresponding points between scans.

The rest of this paper is organized as follows. The colour properties that are used here are introduced in Section 2. After that, the registration method is explained in Section 3. The experimental results are presented in Section 4 and, finally, the most important conclusions are summarized in Section 5.

2   Colour Properties - Delta E

In recent years, the development of new sensors that provide colour and depth information, especially the Kinect developed by Microsoft, has caused a significant change in the research fields that take into account this type of information. One of the most influenced areas in robotics is mapping. Their main advantage is that the colour information is now available together with the depth point cloud. In addition, their low cost has enabled a rapid introduction into the research centers.

The data obtained from these sensors is a 3D depth scan where each measurement is accompanied by its colour information in RGB. That is why this information is traditionally called RGB-D. There are many applications that use this type of data in mapping. One of the most significative approaches is the KinectFusion project [23, 24].

This paper seeks to analyze and exploit the colour characteristics for a particular application. In this section, a brief explanation about how to measure different colour properties is given. These properties will be used later to implement the cost function for the DE-based scan matching method.

In the RGB colour space, each colour is represented by three components corresponding to red, green, and blue:

$$Colour = (R, G, B). \tag{1}$$

Each value varies from zero to a defined maximum value. When the three components are equal to zero, the black colour is represented. The white colour is obtained when the maximum value is fixed in all the components.

The RGB model has been used in multiple fields such as electronics systems and photography. An important fact to be taken into account is that this colour space is highly dependent on the device. The same object will not produce the same RGB values in different devices. Therefore, a RGB value cannot be defined as an absolute colour. An additional disadvantage of this representation is that it depends on the lighting conditions.

Different interesting features can be extracted when each depth point has an associated RGB colour. One possibility is to consider the colour transitions

(significant changes in colour of nearest neighbors in the scan) as the features that will be matched. It is possible to compare the RGB values between adjacent points to measure colour differences:

$$d_c = \sqrt{\Delta R^2 + \Delta G^2 + \Delta B^2}, \qquad (2)$$

where ΔR, ΔG, and ΔB are the differences in each component between nearest neighbours in the scan, and d_c is the colour difference. If this measurement is compared to a fixed threshold, the most significant colour transitions in the depth point cloud can be discovered. These significant points can be used to match pairs of scans (only using the extracted transitions in the cost function, instead of the original point cloud without colour).

A different metric proposed to measure the difference between colours by the International Commission on Illumination (CIE, from the initials in french: *Commission Internationale de l'Éclairage*) is the *delta E* (ΔE), where E is the German word for sensation (*Empfindung*). The goal of this variable is to establish a ΔE value that means a Just Noticeable Difference (JND), which symbolizes colour differences that can be noticed by human beings. Different formulae have been proposed to compute this metric. All of them are based on the device independent CIELAB color space.

The Lab colour space can be defined as a *colour-opponent* space composed of three components: L represents the lightness, and a and b are *colour-opponent* dimensions. These dimensions are based on the CIE XYZ colour space, proposed by the CIE in 1931 and derived from the RGB space. The original coordinates of the Lab space were suggested by Hunter [25, 26]. Nevertheless, these coordinates have evolved to the more recent CIE 1976 or CIELAB colour space, where the $L^*a^*b^*$ components are defined with asterisks to make a distinction between both spaces. The main difference between them is the type of transformation that is applied to calculate the final coordinates.

Two initial objectives were followed when defining the Lab colour space. The first one was to improve the perceptual uniformity. It is important that similar changes in the numerical values produce similar changes in the visual appearance. The second one was to create a space easy to compute from the XYZ coordinates.

The colour spectrum of the CIELAB space is larger than the RGB and CMYK spaces. It includes all perceivable colours, even outside the human vision gamut. One of the most important features of this space is that it does not depend on the device. Another important characteristic is that it has been designed to imitate the human vision system. The perceptual uniformity has been increased and the L^* component is closer to the human perception of lightness.

The CIELAB colour space is composed of three variables:

- L^*: lightness. From $L^* = 0$ (black) to $L^* = 100$ (diffuse white).
- a^*: position between red/magenta and green. From green (negative values) to magenta (positive values).
- b^*: position between yellow and blue. From blue (negative values) to yellow (positive values).

The variations of L^*, a^*, and b^* are nonlinear, trying to reproduce the nonlinear response of the human eye. Furthermore, uniform changes in the values of these variables cause uniform changes in the perceived colour. For this reason the colour difference between two points characterized by their coordinates in the CIELAB space can be approximated by the euclidean distance between them.

The conversion from RGB or CMYK to CIELAB is not an easy task because these models are dependent on the device. It is necessary to make an intermediate conversion to a device-independent space, which is the sRGB. After this conversion, it is possible to obtain the CIE 1931 XYZ coordinates and then the CIELAB values.

Going back to the *delta E* difference, several definitions have been proposed throughout the years (CIE76, CIE94, CIEDE2000, CMC l:c (1984)). The reason of these changes is to minimize or reduce the problems derived from the non-uniformity of the CIELAB space. An important problem to solve is that the human eye is more sensitive to some colours than others. The objective of these metrics is to measure a JND between two colours. Some typical thresholds used to define a JND are 1.0 or 2.3. Nevertheless, the objective of this work is to apply this metric when detecting relevant colour transitions, and this threshold is out of our interest.

In this work, the CIE76 version has been implemented for simplicity (more options will be implemented in a future work). The following equation extracts the ΔE difference between two colours, $L_1^* a_1^* b_1^*$ and $L_2^* a_2^* b_2^*$, in the $L^* a^* b^*$ space:

$$\Delta E_{ab}^* = \sqrt{(L_2^* - L_1^*)^2 + (a_2^* - a_1^*)^2 + (b_2^* - b_1^*)^2}, \tag{3}$$

where $\Delta E_{ab}^* \approx 2.3$ establishes a JND. Equation 3 will be utilized to detect the most significant colour transitions in the scan, points of interest that will be used by the scan matching algorithm.

3   Six-Dimensional DE-based Scan Matching

In this section, the main ideas about the DE-based scan matching method are briefly described. For a more detailed explanation, the reader can consult our previous work [27, 28, 3].

Starting from two independently acquired scans of 3D points (named model and data), this technique computes the rotation and translation that is necessary to apply to the data set in order to maximize the matching between them.

The stochastic search of the *matching pose* is based on the DE method for global optimization problems over continuous spaces. The main concepts of this method are given in Algorithm 1.

The initial population is composed of N_P candidates. Each candidate corresponds to a possible solution. The robot's pose has six degrees of freedom (DOF) because the robot is located in a 3D world:

$$pop_i^k = (x_i^k, y_i^k, z_i^k, \phi_i^k, \theta_i^k, \psi_i^k),$$

Algorithm 1 DE-based Scan Matching using Colour Properties

1: $model\_ct, data\_ct \leftarrow extract\_ct(model, data)$ ▷ Color transitions extraction
2: **for** $i = 1 : N_P$ **do**
3:      $pop_i^1 \leftarrow init\_pop(data\_initial\_pose)$ ▷ First population generation
4:      $e_i^0 \leftarrow cost(model\_ct, data\_ct, pop_i^1)$ ▷ Cost function calculation
5: **end for**
6: **for** $k = 1 : max$ **do**
7:      **for** $i = 1 : N_P$ **do**
8:          $v_i^k = pop_a^k + F(pop_b^k - pop_c^k)$ ▷ Mutation
9:          **for** $j = 1 : D$ **do**
10:             $u_{i,j}^k = v_{i,j}^k, \forall p_{i,j}^k < \delta$ ▷ Crossover
11:             $u_{i,j}^k = pop_{i,j}^k, \forall p_{i,j}^k \geq \delta$
12:         **end for**
13:         $e_i^k \leftarrow cost(model\_ct, data\_ct, pop_i^k)$ ▷ Cost function calculation
14:         **if** $e_i^k < e_i^{k-1} - \tau$ **then**   ▷ Selection with Thresholding
15:             $pop_i^{k+1} = u_{i,j}^k$
16:         **else**
17:             $pop_i^{k+1} = pop_i^k$
18:         **end if**
19:     **end for**
20:     $pop^k = disc(pop^k)$ ▷ Discarding
21:     $ind\_best \leftarrow min(e^k)$
22:     $bestmem \leftarrow pop^k(ind\_best)$
23:     **if** $convergence = true$ **then**   ▷ Execution stops after convergence
24:         $exit(bestmem)$
25:     **end if**
26: **end for**   ▷ Return best estimate

where pop_i^k represents element i at iteration k. The position is represented by the cartesian coordinates and the orientation is given by the Euler angles.

The first population will be randomly situated in a sphere close to the pose estimate provided by the odometry information. The scan matching problem consists of rectifying the robot's pose obtained by the odometry sensors, thus a population is created around the original pose.

The fitness function is executed for each member of the population (line 4 of Algorithm 1). This cost function is a key component of this procedure. It seeks corresponding points between significant colour transitions of both frames (only considering the points of interest, not the original depth point cloud) and calculates a cost value basing on the distance between these corresponding points. The scan matching method is based on the minimization of the cost value. The main concepts and the equations will be explained below in this section.

The principal loop starts in line 6 and it is repeated until one of the convergence conditions is met or the upper limit of iterations is reached. An additional loop that contains the evolutive search begins in line 7. It consists of the generation of a new population for the next generation. In a single iteration the

algorithm is executed to obtain the next candidates, evolving in time to the correct pose.

The evolutive search starts with the mutation. Each population member is perturbed to generate a mutated vector v_i^k:

$$v_i^k = pop_a^k + F(pop_b^k - pop_c^k) \tag{4}$$

where pop_a^k, pop_b^k, and pop_c^k are three randomly chosen elements at iteration k and a, b, and c are different from running index i. F is a real and constant variable that controls the amplification of the differential variations.

To increase the diversity of the new generation, the crossover is introduced. The trial vector is defined by $u_i^k = (u_{i,1}^k, u_{i,2}^k, \ldots, u_{i,D}^k)^T$ and its parameters are selected according to the crossover probability:

$$u_{i,j}^k = \begin{cases} v_{i,j}^k; & \text{if } p_{i,j}^k < \delta, \\ pop_{i,j}^k; & \text{otherwise}, \end{cases} \tag{5}$$

where $p_{i,j}^k$ is a randomly chosen value from the interval $[0,1]$ for each parameter j of the population member i at iteration k, and δ is the crossover probability that constitutes the crossover control variable. D represents the number of chromosomes (number of components of the population element) and is equal to six because the robot's pose has six DOF.

A comparison between the new candidate u_i^k and pop_i^k is done to decide which element should become a member of generation $k+1$. If u_i^k holds a better value for the fitness function than pop_i^k, then it is replaced by u_i^k; otherwise, pop_i^k is retained for the next generation.

A thresholding band has been combined with the selection mechanism to avoid the premature convergence in noisy optimization problems. The idea is to diminish the eagerness of the algorithm by rejecting those new solutions that do not improve the previous hypothesis in a pre-specified magnitude τ.

When the selection mechanism is combined with the thresholding band described before, the following condition must be checked:

$$pop_i^{k+1} = \begin{cases} u_i^k, & \text{if } e_{pop_i}^k - e_{u_i}^k > \tau \\ pop_i^k, & \text{otherwise} \end{cases} \tag{6}$$

where $e_{pop_i}^k$ is the cost function value of the current candidate and $e_{u_i}^k$ represents the fitness function value of the trial vector.

The previous stages (mutation, crossover, and selection with thresholding) are applied to the whole population, obtaining the next generation population $(k+1)$.

Since the convergence speed is reduced by the thresholding band, a discarding mechanism has been included to increase the algorithm speed while keeping the thresholding advantages. The worst members of the new population are replaced by new solutions that are situated close to better ones.

Finally, the algorithm returns the best solution, which is the robot's pose that minimizes the difference between model and data (scan matching solution). The

data set is moved and rotated according to this solution, and the registration is successfully achieved.

3.1 Delta E-based cost function

The cost function of each particle is obtained after extracting the most significant colour transitions.

First, the *delta E* divergence is calculated for each point of the cloud with respect to its closest neighbour according to Equation 3.

After that, the colour transitions are extracted. Only the points with a *delta E* divergence greater than a threshold are selected to be considered by the fitness function.

The cost function to minimize is the derived from the Euclidean distance between the corresponding points of the colour transitions:

$$e = \sum_{c=1}^{C} d(\mathbf{m\_ct}_{i_c}, \mathbf{d\_ct}_{j_c})^2, \tag{7}$$

where the number of correspondences is equal to C and the distance between two corresponding points is denoted by $d(\mathbf{m\_ct}_{i_c}, \mathbf{d\_ct}_{j_c})$.

The result of the scan matching method will be the relation between scans that minimizes the distance between corresponding points, but only considering colour transitions.

4 Experimental Results

The performance of the implemented method has been tested in different situations in a real environment using the the *Freiburg2* data set recorded from a Kinect mounted on a Pioneer robot[1].

Configuration parameters: *delta E* threshold to extract a transition: *min* $\Delta E^*_{ab} = 10$. Reduction factor of the original data: 0.25 (The original clouds are composed of 307200 points. They are reduced in a preprocessing step to increase the computational speed). DE parameters: $N_P = 20$, $F = 0.8$, $\delta = 0.75$.

First of all, the matching results in two different cases are shown in Figure 1 (named as Map 1 and Map 2 from now on). Both maps are composed of three scans. In Figure 2, the points of interest according to the *delta E* difference (top) and the original image (bottom) are presented. The errors are given in Table 1.

As can be observed in both figures, the algorithm presented in this paper is capable of solving the scan matching problem. Besides, we have tested our method more extensively with the data set, which is composed of 2900 different scans, in order to check its performance.

Analyzing the errors in Table 1, the translation errors are in the interval $[0.03, 5.90]$ cm, and the rotation errors are in the interval $[0.010, 0.156]$ rad. The

---

[1] Thanks to Jürgen Sturm for making available this dataset (`http://vision.in.tum.de/data/datasets/rgbd-dataset/download\#freiburg2_pioneer_slam`)
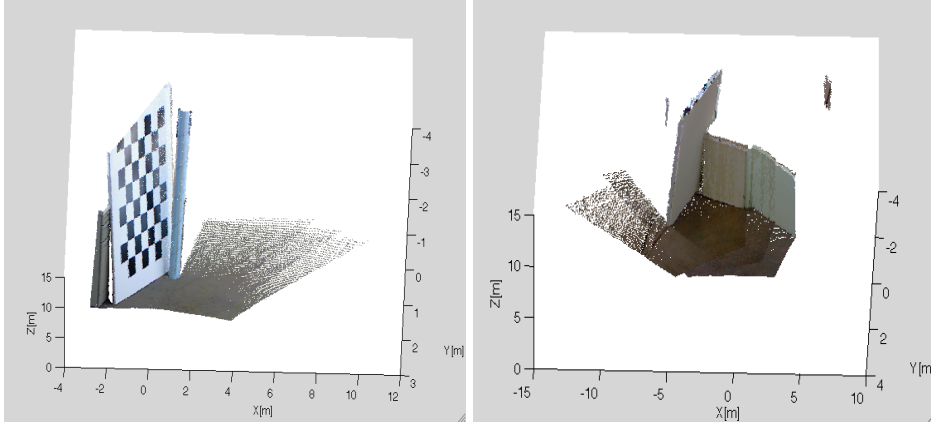
Fig. 1. Two examples of scan matching. Maps composed of three different frames.

Table 1. Scan matching errors in Maps 1 and 2. The ground truth is available. Frame 1 as origin.

	x(cm)	y(cm)	z(cm)	θ_x(rad)	θ_y(rad)	θ_z(rad)
Map 1 Frame 2	0.41	0.32	0.36	0.010	0.057	0.049
Map 1 Frame 3	1.41	0.68	0.68	0.023	0.109	0.103
Map 2 Frame 2	2.25	1.29	5.72	0.007	0.103	0.034
Map 2 Frame 3	3.32	5.90	2.73	0.034	0.156	0.044

accuracy is similar to that shown by other authors. Diosi and Kleeman [20] have reported an average error equal to 3.8 cm and 0.86° using PSM in planar maps. The error obtained by Bosse and Zlot [17] is equal to 11.3 cm, also in planar maps.

An interesting property that is often studied in scan matching is the valley of convergence. It consists of running the algorithm when one of the scans is translated and rotated, checking the results under different offsets. A threshold has to be defined to determine when the matching results are successful. In this experiment, this threshold has been fixed to 0.2 m. It is particularly important to obtain satisfactory results when there are important changes in the orientation, because this is a typical shortcoming of the scan matching methods. Figure 3 shows the valley of convergence for the colour-based scan matching algorithm (left part) with translations from −1 to 1 m and rotations from −80° to 80°. Each semicircle corresponds to a translational offset. Each individual mark of the semicircle represents a rotational offset for a fixed translation. Successful cases are drawn in green and failures in red. The results using the original method without colour are also displayed in the right part of the figure.

Analyzing the results, the *delta E*-based version has a percentage of success of 86.42%. The percentage of success of the original version without colour is
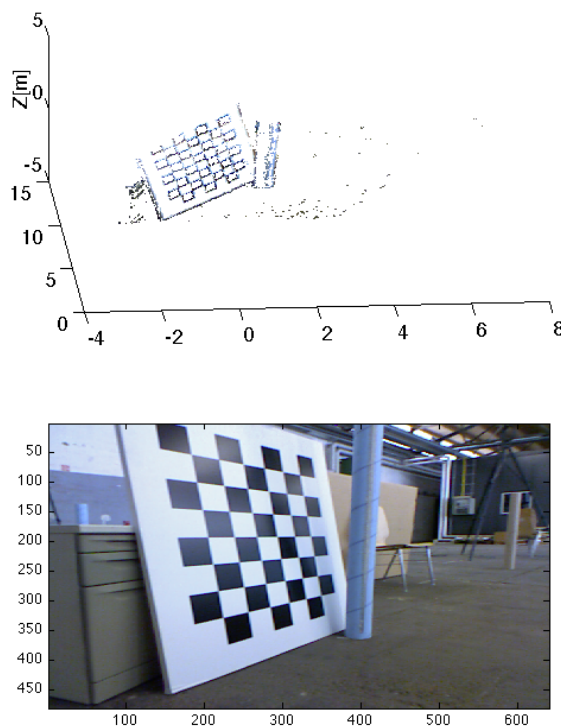
Fig. 2. Example of color transitions extraction: significant points according to the delta E between neighbors (top) and original image (bottom).

80.25%. This percentage is slightly improved, but there is not a big difference between both implementations. Nevertheless, it is a single example and more experiments are needed to reach more conclusive results. Magnusson *et al.* [29] have compared ICP to NDT (they do not consider colour). They use two thresholds, a stricter one (0.2 m) and a weaker one (1 m). We only include their results with the strict threshold because we have chosen the same value for comparison. Their success rates for the strict translation threshold are: 13% for ICP, 37% for NDT, and 95% for trilinear NDT, which is an improved version of the NDT-based method. Our results are better when compared to the ICP implementation and the basic NDT algorithm, but the trilinear NDT registration algorithm outperforms our method. However, these are the results in particular cases. In order to have an adequate comparison, all methods have to be compared in detail using the same data set, which is an interesting work to be done in the future. This comparison will let us know if our method presents a better performance
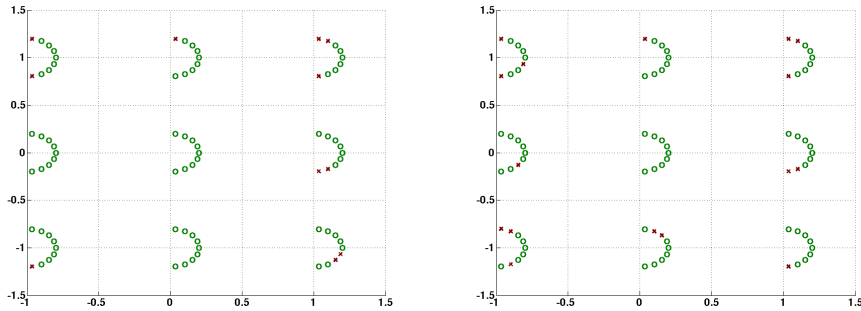
Fig. 3. Valley of convergence. Offsets in metres. Left: Version with *delta E*. Right: Version without colour. Successful matching in green. Wrong matching in red.

under sharp turns (robustness under rotation), which is one of the traditional shortcomings of the ICP method.

An additional fact has been observed in the valley of convergence experiment. The average of the fitness function final value for the *delta E*-based method is 1.1423 cm/measurement, and the average position error is equal to 0.0063 cm. In the matching version without colour, the cost function final value is equal to 1.8078 cm/measurement, and the average position error is 1.1117 cm. In this case, the accuracy is higher when the colour properties are included.

The last property analyzed in the experiments is the convergence speed. A comparison between the convergence of the new version and the original method without colour for a particular case is presented in Figure 4. The evolution of the sum of the cost values (for the whole population), in dm per measurement, against the number of iterations is displayed. In the original method, the algorithm converges after 100 iterations. In the new version, the final value is reached after 60 iterations. The convergence speed is significantly improved, which is an important result because the computational cost depends directly on the number of iterations. Due to its relation with the computational cost, a more detailed study about the improvement in the convergence speed is a necessary work to be done in the future.

5   Conclusions

An initial version of a scan matching method based on colour properties has been proposed in this paper. In particular, we have used the *delta E* divergence to extract significant colour transitions of RGB-D scans. This colour information has been included in the cost function of the DE-based scan matching method, which only matches points that belong to these significant transitions.

The algorithm has been tested in a real environment, showing a good performance in a data set composed of multiple frames recorded from a Kinect. The errors obtained are similar to those obtained by other research groups.
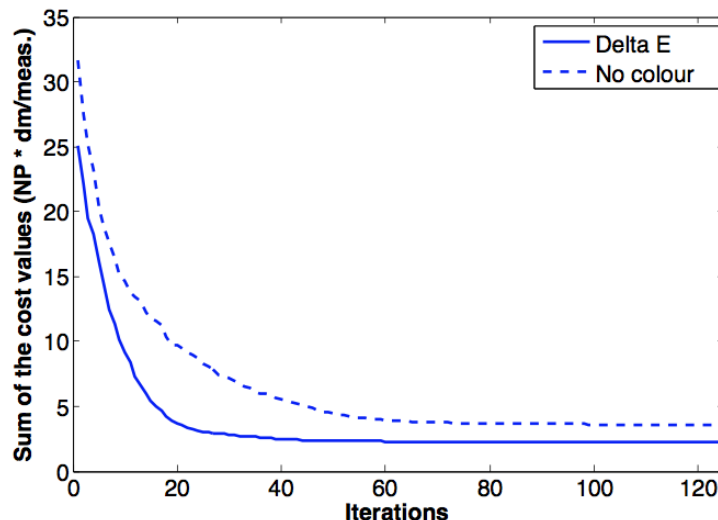
Fig. 4. Convergence comparison between the original method and the colour-based one.

Regarding the valley of convergence, this new version presents a slightly better performance. Although there are not significant differences when compared with the DE-based version without colour, both versions (with and without colour) outperform the classic ICP results shown by Magnusson *et al.* [29]. Since the lack of robustness when the change in the orientation is important is a key drawback in scan matching, a more detailed study about the valley of convergence of multiple methods/options (ΔE-DE, DE, ICP, NDT, etc.) is a challenging work to be accomplished in the future.

The convergence speed is significantly improved in the new version, which is an important advantage because the computational cost is decreased.

Finally, it is necessary to remark that this paper is an initial study about how to apply colour properties to improve the scan matching. More experiments are needed and more properties have to be considered and implemented to improve our method.

References

1. J. J. Leonard and H.Durrant-Whyte, "Mobile Robot Localization by Tracking Geometric Beacons," *IEEE Transaction on Robotics and Automation*, vol. 7, pp. 376–382, 1991.
2. R. Smith, M. Self, and P. Cheeseman, "Estimating Uncertain Spatial Relationships in Robotics," in *Proceedings of the Second Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI'86)*, (New York, NY), pp. 267–288, Elsevier Science, 1986.

3. F. Martín, R. Triebel, L. Moreno, and R. Siegwart, "Two different tools for three-dimensional mapping: DE-based scan matching and feature-based loop detection," *Robotica*, 2013.

4. R. Storn and K. Price, "Differential Evolution A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," *Journal of Global Optimization*, vol. 11, p. 341359, December 1997.

5. P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments," in *Proceedings of the Intl. Symp. on Experimental Robotics (ISER)*, 2010.

6. A. Nüchter, K. Lingemann, and J. Hertzberg, "6D SLAM-3D Mapping Outdoor Environments," *Journal of Field Robotics*, vol. 24, pp. 699–722, 2007.

7. P. J. Besl and N. D. McKay, "A method for registration of 3d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

8. D. Hähnel, W. Burgard, and S. Thrun, "Learning compact 3d models of indoor and outdoor environments with a mobile robot," *Robotics and Autonomous Systems*, vol. 44, pp. 15–27, 2003.

9. R. Triebel, P. Pfaff, and W. Burgard, "Multi-Level Surface Maps for Outdoor Terrain Mapping and Loop Closing," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'06)*, 2006.

10. D. M. Cole and P. M. Newman, "Using Laser Range Data for 3D SLAM in Outdoor Environments," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'06)*, 2006.

11. M. Magnusson and T. Ducket, "A Comparison of 3D Registration Algorithms for Autonomous Underground Mining Vehicles," in *Proceedings of the Second European Conference on Mobile Robotics*, (Ancona, Italy), 2005.

12. P. Biber and W. Straer, "The Normal Distributions Transform: A New Approach to Laser Scan Matching," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'03)*, 2003.

13. F. Lu and E. Milios, "Robot Pose Estimation in Unknown Environments by Matching 2D Range Scans," *Journal of Intelligent and Robotic Systems*, vol. 20, pp. 249–275, 1997.

14. M. Tomono, "A scan matching method using euclidean invariant signature for global localization and map building," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'04)*, 2004.

15. Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *International Journal of Computer Vision*, vol. 13, pp. 119–152, 1994.

16. S. Rusinkiewicz and M. Levoy, "Efficient Variants of the ICP Algorithm," in *Proceedings of the Third International Conference on 3D Digital Imaging and Modeling*, 2001.

17. M. Bosse and R. Zlot, "Map Matching and Data Association for Large-Scale 2D Laser Scan-Based SLAM," *The International Journal of Robotics Research*, vol. 27, pp. 667–691, 2008.

18. A. Johnson and S. B. Kang, "Registration and integration of textured 3-d data," in *International Conference on Recent Advances in 3-D Digital Imaging and Modeling (3DIM'97)*, 1997.

19. S. May, D. D. ad D. Holz, E.Fuchs, S. Malis, A. Nüchter, and J. Hertzberg, "Three-dimensional mapping with time-of-flight camers," *Journal of Field Robotics*, vol. 26, pp. 11–12, 2009.

20. A. Diosi and L. Kleeman, "Laser Scan Matching in Polar Coordinates with Application to SLAM," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'05)*, 2005.
21. S. Thrun, W. Burgard, and D. Fox, "A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping.," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'00)*, 2000.
22. F. Ramos, D. Fox, and H. Durrant-Whyte, "CRF-matching: Conditional random fields for feature-based scan matching," in *Proceedings of Robotics: Science and Systems (RSS'07)*, 2007.
23. S. Izadi, D. Kim, O. H. D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon, "KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera," in *Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST'11)*, pp. 559–568, 2011.
24. R. A. Newcombe, A. J. Davison, S. Izadi, P. Kohli, O. Hilliges, J. Shotton, D. M. S. Hodges, D. Kim, and A. Fitzgibbon, "KinectFusion: Real-Time Dense Surface Mapping and Tracking," in *Proceedings of the 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR'11)*, 2011.
25. R. S. Hunter, "Photoelectric color-difference meter," in *Proceedings of the Winter Meeting of the Optical Society of America*, 1948.
26. R. S. Hunter, "Accuracy, precision, and stability of new photo-electric color-difference meter," in *Proceedings of the Thirty-Third Annual Meeting of the Optical Society of America*, 1948.
27. L. Moreno, S. Garrido, and M. L. Muoz, "Evolutionary Filter for Robust Mobile Robot Localization," *Robotics and Autonomous Systems*, vol. 54, no. 7, pp. 590–600, 2006.
28. F. Martín, L. Moreno, S. Garrido, and D. Blanco, "High-Accuracy Global Localization Filter for three-dimensional Environments," *Robotica*, vol. 30, pp. 363–378, 2011.
29. M. Magnusson, A. Nüchter, C. Lörken, A. J. Lilienthal, and J. Hertzberg, "Evaluation of 3D Registration Reliability and Speed A Comparison of ICP and NDT," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'09)*, 2009.