3D non-rigid registration using color: Color Coherent Point Drift[☆]

Marcelo Saval-Calvo[a,*], Jorge Azorín-López[a], Andrés Fuster-Guilló[a], Víctor Villena-Martínez[a], Robert B. Fisher[b]

[a] *University of Alicante, Carretera Sant Vicent del Raspeig s/n, 03690,Spain*
[b] *School of Informatics, Univ. of Edinburgh, 10 Crichton St, Edinburgh EH8 9AB, UK*

Abstract

Research into object deformations using computer vision techniques has been under intense study in recent years. A widely used technique is 3D non-rigid registration to estimate the transformation between two instances of a deforming structure. Despite the previous developments in this topic, it remains a challenging problem. In this paper we propose a novel approach to non-rigid registration combining two data spaces in order to robustly calculate the correspondences and transformation between two data sets. In particular, we use point color as well as 3D location as these are the common outputs of RGB-D cameras, which leads to the Color Coherent Point Drift (CCPD) algorithm (an extension of the CPD method [1]). Evaluation is performed using synthetic and real data. The synthetic data includes easy shapes to evaluate the different effects of noise, outliers and missing data. Moreover, an evaluation of realistic figures obtained using Blensor is carried out. Real data acquired using a general purpose Primesense Carmine is used to validate the CCPD for real shapes. For all tests, the proposed method is compared to the original CPD showing better results in registration accuracy in most cases.

Keywords: 3D non-rigid registration, 3D deformable registration, CCPD

[☆] Authors acknowledge the project from University of Alicante (Gre16-28).
[*] Corresponding author
Email address: msaval@dtic.ua.es (Marcelo Saval-Calvo)
URL: www.dtic.ua.es (Marcelo Saval-Calvo)

1. Introduction

The study of the evolution in shapes over time is under intense study in many areas, such as biology, health, etc. During evolution, objects are affected by multiple changes, disturbing both shape and appearance. To measure all the changes is a difficult and tedious task, due to the complexity of some shapes and the large amount of data necessary to have a complete study. Computer vision techniques can help provide methods which, given a set of data from a sensor, estimate the changes. In this paper, we propose a method to robustly estimate the deformation observed in an object. Concretely, non-rigid registration methods estimate the transformation between two shapes aligning the data using non-rigid transformations.

There are several applications that require non-rigid alignment. For instance, face or body motion recovery where the different parts need to be tracked to perceive the motion or identify the action. Applications where shape evolution is studied require from deformable alignment as well, and may involve appearance changes, which commonly include color variations. Using machine intelligence to evaluate those changes mean to use methods than can perceive them regardless the nature of the change. For example intelligent farms can use these techniques to increase the quality of the products since they can be constantly supervised while growing. In health, automatic analysis of human body change will help specialists in treatment supervision (eg. for cancer therapy.

There exist various kind of deformations: isometric deformation, where both topology and distances are preserved (e.g. articulated changes or flag movements); elastic deformation, where the topology is kept but distances can vary (e.g. balloon inflation); and free deformations where both topology and distances can change (e.g. growing objects or breaking situations).

In this paper we focus on 3D point clouds without any previous filtering, only downsampling if necessary. For the specific case of this paper, the data comes from a low-cost RGB-D sensor, such as a Microsoft Kinect, which provides color and 3D information. The sensitivity of these sensors may be lower than the

requirements of the problem, and may be difficult for some tasks. Nonetheless, they are widely used and a contribution using this sensor will be useful in many research tasks and industrial applications.

The deformations considered in this work are not constrained. That is, they do not assume a prior restriction in the deformation as topology/size constraints, larger/smaller variations, etc. Then, the objective is to develop a non-rigid registration method for non-constrained free deformations.

Non-rigid registration methods for 3D point sets, such as the well-known Coherent Point Drift (CPD) [1], only use spatial 3D information (or location information) to register the data. Ignoring other information, such as color, increases the probability of misalignment. For instance, in cases where the object grows the number of points may increase or decrease in an irregular distribution. If only 3D spatial data is taken into account, those irregularities are harder to register. Those are situations where additional information can be used to robustly register. A practical example is the plant growth, where leaves change shape differently over their surface. Commonly, the central region remains similar whilst the edges enlarge significantly, but in the spatial data this variation in growth is not as obvious. It is necessary to use color information to perceive this difference. The leaf growth problem motivates our work, which improves the CPD algorithm including color information in the process of matching estimation to improve the estimation of the deformations. Although original motivated by the leaf growth problem, the developed Color Coherent Point Drift (CCPD) algorithm is a general algorithm usable for registering deforming colored point clouds.

The main contribution of this paper is a novel approach for colored point cloud non-rigid registration combining various inputs in the correspondence estimation step. To handle real and adverse situations, the method has to deal with noise, outliers and missing data, common issues in real applications. The proposal makes use of the basis proposed in the CPD algorithm [1], because of its generality and because it has shown good results in point set non-rigid registration in presence of noise and outliers.

The rest of the paper is organized as follow: Section 2 presents a review of the State-of-the-Art in 3D non-rigid registration methods for point sets. Section 3 details the proposed CCPD method. The evaluation is shown in Section 4 where synthetic and real data are used to validate the proposal. Finally, some discussion and conclusion are presented in Section 5.

2. Previous research

Recently, the increasing interest in non-rigid registration has produced much research that improves existing algorithms or introduces new methods, but this is still a challenging problem to be solved. This interest comes from the need to improve reconstruction, mapping or other computer vision problems, where dynamic objects are treated. Tam et al. [2] surveyed different methods for point cloud and mesh registration, in both rigid and non-rigid situations.

Chui and Rangarajan [3, 4] proposed the TPS-RPM non-rigid registration method for 3D point clouds based on Thin Plate Splines to stabilize the displacement of the points during the process of registration. This method uses softassign matches between each point set [5]. Softassign refers to the use of non-binary correspondences to handle noise and outliers because there is no imposing of a unique matching per point. Deterministic annealing [6] is also used in the kernel of TPS-RPM to gradually allow a less constrained movement of the individual points. Their proposal outperforms ICP in 2D, and also achieves better results in 3D than the main state-of-the-art methods. Yang revisited TPS-RPM in [7] demonstrating limited performance when outliers are present in both point sets simultaneously. He proposed a double-sided outlier handling approach obtaining better registration results.

Li et al. [8] presented a non-rigid registration method that simultaneously estimated confidence weights, that measure the reliability of each correspondence, and identified non-overlapping areas. A warping field brings the source scan into alignment with the target geometry.

Sang et al. [9] proposed the FDMM non-rigid registration method based on

GMM and the use of features, that they called Gaussian soft shape context, based on radial distribution of the neighbourhood. This feature was initially presented in [10, 11], and they modified it adding a Gaussian distribution for avoiding the problem of non-real similarities. The algorithm takes into account the relative distribution of all points with respect to the analysed point, making a histogram, which adds information to the registration process. Comparison to CPD, RPM and BEM [12] is provided using 2D data, outperforming the previous results. Yawen et al. [13] proposed also the use of this feature enhancement with CPD to handle noise and outliers with better results.

Wang and Fei [14] proposed B-spline-based point matching (BPM), an extension of RPM, using a deterministic annealing scheme to regularize the registration process. The method was evaluated in different situations with accurate results in 2D and 3D data. Yang et al. proposed in [15] the GLMD, a two step non-rigid registration method for point sets. They proposed the use of local and global distances combined to estimate the binary correspondences, and transformation using the TPS kernel. The local distances are measured using a certain neighbourhood, which is provided initially. Experiments were provided using the proposed method against CPD, TPS-RPM and GMMreg for different levels of noise, outliers and rotations.

Recently Chen et al. [16] proposed the Coherent Spatial Mapping (CSM) algorithm. They used the shape context [10] which describes the shape using a histogram of each point relative position to the others, and calculate correspondences with this information. The Hungarian method is also used to estimate the initial correspondences. The transformation is iteratively estimated with EM method using a spatial mapping function of the correct matches, and TPS to provide smooth deformations. Hence, the improvement comes from the matching estimation. They compare CSM to CS [11], CPD, COA-RPM [17] and TPS-RPM with 2D data achieving better alignment with lowest RMS error with different levels of noise and outliers. In 3D they compare against CPD achieving lower registration error.

S. Lin et al. presented in [18] a proposal for incorporating color in the regis-

tration process, both in rigid and non-rigid registration. The non-rigid approach is based on the paper of Li et al. [8], incorporating the color information in the vertex selection by evaluating 3D location and color distance in Euclidean space, using a neighbourhood to improve robustness, between the two views. Moreover, after estimating the descriptors (Gabor and HOG) from the vertex, color is also used for rejecting wrong correspondences. This paper considers small deformations mainly related to orientation of views which deform the shapes due to the RGB-D sensor pattern projection.

2.1. Coherent Point Drift variants

One of the most common algorithms used for non-rigid registration is the Coherent Point Drift (CPD) proposed by Myronenko et al. in [1, 19]. This method is based on a Gaussian Mixture Model (GMM) and Expectation Maximization (EM) to calculate the correspondences, and then the transformations, of the points to map one set of points into another. They used a GMM to represent the moving point set to be registered, and EM to evaluate the new parameters of the GMM and hence, the new position of the points. Moreover, in order to constrain the movement, they make use of Coherent Motion Theory that helps the translation of points to be regular. They compared their results to the TPS-RPM outperforming the registration for 2D and 3D cases. Wang et al. [20] proposed an extended version of CPD to automatically evaluate the outlier percentage parameter, which is manually provided in the original version. They used a combination of Nelder-Mead simplex and genetic algorithms. The genetic algorithm provides good initial values for this parameter, while the Nelder-Mead simplex optimizer attempts to find an optimal solution. The experimentation showed an improvement of the original version for different levels of noise, where they initialized the outlier parameter to 0.7.

A different approach called GMMreg was presented by Jian et al. in [21, 22]. Instead of representing a point set with a GMM and registering it to a point cloud using the EM technique, they align two GMMs each representing one of the point sets to be registered. They calculate the displacement between

Mixtures of Gaussians and iteratively align them using the L2 distance. They provided rigid and non-rigid results for 2D and 3D data compared to the original CPD, LM-ICP [23], and TPS-RPM among others, resulting in more accurate results. Additionally, they apply the L2 distance to TPS and to Gaussian radial basis functions, improving the results.

Gerogiannis et al. [24] proposed a different matching method using the Hungarian Algorithm instead of the posterior distribution used in CPD and RPM. Moreover, they used Bayesian regression for the Maximization step (i.e. the registration or transformation part). The experiments compared the proposed method with CPD, RPM and GMMreg for 2D and 3D cases.

Gao et al. studied in [25] the main drawbacks of CPD related to outliers, which are a consequence of the way CPD keeps the distribution of outliers, and the input parameter for the outlier ratio. They proposed an Expectation-Maximization solution to iteratively evaluate the outlier ratio. TPS-RPM and the original CPD algorithms show less accurate results when the outlier ratio grows. The main advantage of this method is to avoid the need to indicate the outlier ratio initially.

Ge et al. [26] presented a similar approach to the previous one, called Global-Local Topology Preservation (GLTP). The main motivation of this work is to handle non-rigid articulated deformations such as those of human movements. They added the principle of Local Linear Embedding to the original CPD to take into account local deformation coherence, apart from the global coherence intrinsic in the CPD algorithm. With large articulated deformations GLTP works better than the original CPD, which is not able to find a good registration.

De Sousa and Kropatsch [27] proposed a variant of Coherent Point Drift (CPD) by integrating centrality information, a concept initially applied in social networks. It creates a graph (e.g. Delaunay triangulation), and applies different centralities (node degree, betweenness, eigenvector ...) to evaluate which results in a better solution. The proposal shows good performance with noisy data, improving the original CPD.

Another variation of CPD was presented by Zhou et al. [28] using Student's

mixture model, which they claim to be more robust in the presence of high amounts of noise. The comparison they made against CPD and TPS-RPM shows better performance when the noise rate grows. Moreover, they automatically estimated the probability of outliers whereas in CPD it is manually indicated.

In conclusion, many studies have been done for non-rigid registration of point sets. Most of them focused their attention on outliers and noise handling. In order to do this, they proposed techniques to estimate automatically the outlier ratio or used descriptors which use point distributions to improve the matching. However, there still exist problems when there are large deformations. Another issue not studied is where the data does not have to move coherently in the whole space. For example, situations in which one set is a full model and the other is just a region. Moreover, there are no general proposals facing the problem from a generic perspective including several sources of data using individually the different spaces, e.g. using color and 3D location without using them as a 6D data set but being independent in the process for a more robust and generic combination.

3. Color Coherent Point Drift

In this section, a framework for non-rigidly registering 3D colored points based on CPD [1, 19] is presented. We use the optimization algorithm of the original CPD algorithm, only replacing the original similarity matching formulation with one that takes account of having colored 3D points.

The proposed Color Coherent Point Drift (CCPD) algorithm registers 3D points by using color and shape spaces to jointly estimate the best match. It improves upon the CPD algorithm by using the two input spaces together to handle situations where point position is not sufficient to adequately estimate the matches, *e.g.* aligning shapes with missing parts, or non-linear growth of the shape.

In any registration problem it is normal to have one point set used as the

anchor or reference point set which we will call *Anchor*, and the other as the moving points called *Moving*. The *Moving* set will be deformed and moved until it aligns with the *Anchor*. CCPD (following the basics of CPD) models the *Moving* set using a Gaussian Mixture Model (GMM) and estimates the transformation of the *Moving* set using the Expectation-Maximization (EM) technique. The use of a GMM to represent the *Moving* set will give soft correspondences, *i.e.* they are not binary, allowing a more robust estimation of the displacement by not requiring one-to-one matching. Moreover, in order to smooth the displacement, the Coherent Motion Theory is used to regularize the motion of the points in the process of the transformation.

Here, we introduce the combination of color and shape (3D positions) spaces for non-rigid registration. Let A^S, A^C, M^S and M^C (Eq. 1) be four data sets representing two spaces (shape and color) of two data sets. A^S and A^C are the shape and color values of the *Anchor* set and M^S and M^C are the shape and color values of the *Moving* set. To simplify the notation, we will refer to A^S, A^C as A, and M^S, M^C as M when we refer to both spaces together.

$$\begin{aligned} A^S &= \{a_1^S, \cdots, a_N^S\} \\ A^C &= \{a_1^C, \cdots, a_N^C\} \\ M^S &= \{m_1^S, \cdots, m_M^S\} \\ M^C &= \{m_1^C, \cdots, m_M^C\} \end{aligned} \tag{1}$$

where $a_i^S, m_i^S \in \mathbb{R}^{D_S}$ and $a_i^C, m_i^C \in \mathbb{R}^{D_C}$. N and M are the number of points in the *Anchor* and *Moving* point sets. M^S and M^C are the *Moving* to be aligned with the reference *Anchor* A^S and A^C. Each space has its own dimension *e.g.* $D_S = 3$ for shape (3D points), but $D_C = 1$ for monochrome or $D_C = 3$ if we use 3 color components. The points M^S and M^C are appended to form the centroids of the components of a Gaussian Mixture Model (GMM) (m = 1..M) that encodes the probability of the *Moving* point set, as described in Eq. 2. x

and m_i are vectors with the point's position and color appended,

$$p(x) = \sum_{i=1}^{M} w(m_i) p(x|m_i) \tag{2}$$

$w(m_i)$ is the weight of each GMM component. Here, all points are treated equally, so $w(m_i) = \frac{1}{M}$.

Let $D = D_S + D_C$ and Λ be the D dimensional covariance matrix. Then, each Gaussian is modelled using Eq. 3.

$$p(x|m_i) = \frac{1}{(2\pi)^{\frac{D}{2}}} \frac{1}{det(\Lambda)^D} e^{-\frac{1}{2}(x-m_i)'\Lambda^{-1}(x-m_i)} \tag{3}$$

Eq. 3 will be modified later as all components have equal isotropic variance σ_S^2 (for the shape components) and σ_C^2 (for the color components). The shape (S) and color (C) covariance matrices for the old and new (for $z \in \{o, n\}$ for o:old and n:new which will be defined below) Gaussian distributions are: $\Lambda_S^z = (\sigma_S^z)^2 \mathrm{I}_{D_S}$, $\Lambda_C^z = (\sigma_C^z)^2 \mathrm{I}_{D_C}$. From these, we get $(\Lambda_S^z)^{-1} = (\sigma_S^z)^{-2} \mathrm{I}_{D_S}$, $(\Lambda_C^z)^{-1} = (\sigma_C^z)^{-2} \mathrm{I}_{D_C}$, and $det(\Lambda_S^z) = (\sigma_S^z)^{2D_S}$, $det(\Lambda_C^z) = (\sigma_C^z)^{2D_C}$.

In order to handle noise and outliers, an additional probability distribution $\frac{1}{N}$, where N is the number of *Anchor* points, is included which is weighted with a predefined parameter α. Thus, Eq. 4 is the complete probability of the fit of the *Anchor* points to the *Moving* points.

$$p(x) = \alpha \frac{1}{N} + (1-\alpha) \sum_{i=1}^{M} \frac{1}{M} p(x|m_i) = \sum_{i=1}^{M+1} w(i) p(x|m_i) \tag{4}$$

where $w(M+1) = P(X|m_{M+1}) = \frac{1}{N}$ and otherwise $w(i) = \frac{1-\alpha}{M}$.

The GMM is parametrized by a set of parameters $(\theta_S, \sigma_S, \sigma_C)$ which specify the translation of the *Moving* point set (θ_S), the standard deviation (σ_S) of the points' positions, and the standard deviation (σ_C) of the points' colors.

Expectation-Maximization (EM) is used to register the *Moving* points to the *Anchor* points.

The function E finds the parameters (θ_S, σ_S) that maximize the likelihood, or equivalently, minimize the negative log-likelihood (Eq. 5). In this paper

we are registering only the shape vectors, but not the color vectors. We are using shape and color information in the similarity score to make the matching estimation more robust. Thus, the set of parameters is $(\theta_S, \sigma_S, \sigma_C)$, where θ_S are the parameters that control the position of the *Moving* points.

$$E(\theta_S, \sigma_S) = -\sum_{n=1}^{N} log(\sum_{i=1}^{M+1} w(i) p(a_n|m_i)) \tag{5}$$

Following the original formulation of CPD, the probability of correct correspondence between model point m_i and anchor point a_n is the posterior probability of the GMM centroid given the anchor point: $p(m_i|a_n)$, which by Bayes' Rule equals $p(m_i)p(a_n|m_i)/p(a_n)$. Since the objective of the registration is to find the parameters to make model M best fit anchor A, the Expectation-Maximization (EM) algorithm is used. Given the value of the "old" (superscript 'o') position and tolerance parameters, we use Bayes' theorem to estimate the posterior probability p^o (Eq. 13), known as Expectation or E-step; then we find the new parameters that Maximize (M-step) the probability. Here, we minimize the negative log-likelihood:

$$Q(\theta_S, \sigma_S) = -\sum_{n=1}^{N} \sum_{i=1}^{M+1} w(i) p^o(m_i|a_n) log(p^n(m_i) p^n(a_n|m_i)) \tag{6}$$

Before we manipulate Q, we need some useful terms. Recalling that $M+1$ refers to the background model: $p(m_{M+1}) = 1$ and otherwise $p(m_i) = 1$ and $p(x|m_{M+1}) = \frac{1}{N}$.

The multivariate Gaussian distributions that we need for the shape term is ($z \in \{o, n\}$ for o:old and n:new):

$$p_S^z(x_S|m_{i,S}) = \frac{1}{(2\pi)^{\frac{D_S}{2}}} \frac{1}{(\sigma_S^z)^{D_S}} e^{-\frac{1}{2(\sigma_S^z)^2}||x_S - \tau(m_{i,S}, \theta_S^z)||^2} \tag{7}$$

and for the color term is:

$$p_C^z(x_C|m_{i,C}) = \frac{1}{(2\pi)^{\frac{D_C}{2}}} \frac{1}{(\sigma_C^z)^{D_C}} e^{-\frac{1}{2(\sigma_C^z)^2}||x_C - m_{i,C}||^2} \tag{8}$$

where $\tau(m, \theta_S)$ transforms the position of point m given the *Moving* point set pose parameters θ_S. Here, the transformation is only a Euclidean rigid motion.

Note the color matching probability $p^z_C(x_C|m_{i,C})$ uses the distance between the colors without any transformation. Combining Eq. 7 and 8 we get $P^z(x|m_i) = P^z_S(x_S|m_{i,S}) \cdot P^z_C(x_C|m_{i,C})$.

The first manipulation addresses the background term $M+1$. We split out the $M+1$ term from the rest and analyze it:

$$Q(\theta_S, \sigma_S) = Q'(\theta_S, \sigma_S) - \sum_{n=1}^{N} w(M+1) p^o(m_{M+1}|a_n) log(p^n(m_{M+1}) p^n(a_n|m_{M+1})) \tag{9}$$

We have: $w(M+1) = \frac{\alpha}{N}$, $p^o(m_{M+1}) = p^n(m_{M+1}) = 1$, $p^o(a_n|m_{M+1}) = p^n(a_n|m_{M+1}) = \frac{1}{N}$,
$p^o(m_{M+1}|a_n) = \frac{p^o(a_n|m_{M+1}) p^o(m_{M+1})}{p^o(a_n)} = \frac{1}{N} \frac{1}{p^o(a_n)}$.
Substituting, this gives:

$$Q(\theta_S, \sigma_S) = Q'(\theta_S, \sigma_S) + \frac{\alpha log(N)}{N^2} \sum_{n=1}^{N} \frac{1}{p^o(a_n)} \tag{10}$$

The latter term becomes small as N grows. Further, there are none of the 'new' parameters to optimize in that term. So, we can ignore it and find the parameters (θ^n_S, σ^n_S) that minimizes only Q':

$$Q'(\theta^n_S, \sigma^n_S) = -\sum_{n=1}^{N} \sum_{i=1}^{M} w(i) p^o(m_i|a_n) log(p^n(m_i) p^n(a_n|m_i)) \tag{11}$$

Since $log(p^n(m_i)p^n(a_n|m_i))$
$= log(p^n(m_i)) + log(p^n(a_n|m_i))$ and $log(p^n(m_i)) = log(\frac{1}{M})$ has none of the optimization parameters, even when multiplied by $p^o(m_i|a_n)$, we can ignore this term. Similarly, $w(i) = \frac{1-\alpha}{M}$ so it is ignored. Thus, we need to optimize:

$$Q''(\theta^n_S, \sigma^n_S) = -\sum_{n=1}^{N} \sum_{i=1}^{M} p^o(m_i|a_n) log(p^n(a_n|m_i)) \tag{12}$$

By Bayes's rule:

$$p^o(m_i|a_n) = \frac{p^o(a_n|m_i) p^o(m_i)}{\sum_{j=1}^{M} w(j) p^o(a_n|m_j) p^o(m_j) + w(M+1) p^o(a_n|m_{M+1}) p^o(m_{M+1})}$$

Simplifying, we get:

$$p^o(m_i|a_n) = \frac{M}{1-\alpha} \frac{p^o(a_n|m_i)}{\sum_{j=1}^{M} p^o(a_n|m_j) + \frac{\alpha}{1-\alpha}\frac{M}{N}} \tag{13}$$

This is evaluated using the 'old' parameters and does not change with the current optimization iteration. The initial $\frac{M}{1-\alpha}$ can also be omitted as an inessential scaling factor.

Finally, we need to consider $p^o(a_n|m_i)$ and $p^n(a_n|m_i)$. We will analyze both of these together for $z \in \{o, n\}$ (for o:old and n:new).

We assume that point shape and color are independent, and that the optimization affects only the position of the points, but not the color. Therefore, $p^z(a_n|m_i) = p^z_S(a_n|m_i)p^z_C(a_n|m_i)$, and these terms were defined above. For operational reasons, we choose to weight the shape and color components with w_S and w_C. So our formula is: $p^z(a_n|m_i) = [p^z_S(a_n|m_i)]^{w_S}[p^z_C(a_n|m_i)]^{w_C}$.

Substituting these derivations into Eq. 12, we get (where the first term is evaluated before optimization using Eq. 13):

$$Q''(\theta^n_S, \sigma^n_S) = -\sum_{n=1}^{N}\sum_{i=1}^{M} p^o(m_i|a_n)\times$$
$$log([p^n_S(a_n|m_i)]^{w_S}[p^n_C(a_n|m_i)]^{w_C})$$

Applying the 'log' function and then simplifying:

$$Q''(\theta^n_S, \sigma^n_S) = -\sum_{n=1}^{N}\sum_{i=1}^{M} p^o(m_i|a_n)\times$$
$$[w_S log(p^n_S(a_n|m_i)) + w_C log(p^n_C(a_n|m_i))]$$

And then applying the substitutions from Eq. 7 and Eq. 8, and then simplifying:

$$Q''(\theta^n_S, \sigma^n_S) = -\sum_{n=1}^{N}\sum_{i=1}^{M} p^o(m_i|a_n)\times$$
$$[w_S[log(\frac{1}{(2\pi)^{\frac{D_S}{2}}}\frac{1}{(\sigma^n_S)^{D_S}}) - \frac{1}{2(\sigma^n_S)^2}||a_{n,S} - \tau(m_{i,S}, \theta^n_S)||^2]$$
$$+ w_C[log(\frac{1}{(2\pi)^{\frac{D_C}{2}}}\frac{1}{(\sigma^n_C)^{D_C}}) - \frac{1}{2(\sigma^n_C)^2}||a_{n,C} - m_{i,C}||^2]]$$

Simplifying again and removing terms not involving the optimization parameters, we get Eq; 14 to be optimised in the EM 'M' step over the parameters: $(\theta_S^n, \sigma_S^n, \sigma_C^n)$:

$$\begin{aligned} Q''(\theta_S^n, \sigma_S^n) = \sum_{n=1}^{N}\sum_{i=1}^{M} p^o(m_i|a_n)\times \\ [w_S D_S log(\sigma_S^n) + \frac{w_S}{2(\sigma_S^n)^2}||a_{n,S} - \tau(m_{i,S}, \theta_S^n)||^2 \\ + w_C D_C log(\sigma_C^n) + \frac{w_C}{2(\sigma_C^n)^2}||a_{n,C} - m_{i,C}||^2] \end{aligned} \tag{14}$$

Since the parameters of the color GMM are not optimized in the EM process, the second term in the addition in Eq. 14 becomes a constant and can be removed along with the weighting operators. Thus, we end up with a simpler Q'' as next:

$$\begin{aligned} Q''(\theta_S^n, \sigma_S^n) = \sum_{n=1}^{N}\sum_{i=1}^{M} p^o(m_i|a_n)\times \\ [D_S log(\sigma_S^n) + \frac{w_S}{2(\sigma_S^n)^2}||a_{n,S} - \tau(m_{i,S}, \theta_S^n)||^2 \end{aligned} \tag{15}$$

Therefore, the color information is involved only in the 'old' probability. Recalling that $p^z(a_n|m_i) = [p_S^z(a_n|m_i)]^{w_S}[p_C^z(a_n|m_i)]^{w_C}$ for $z \in \{o, n\}$ (for o:old and n:new) and those terms were defined in Eq. 7 and 8, we substitute Eq. 13 by Eq 16:

$$P^o(m_i|a_n) = \frac{[p_S^z(a_n|m_i)]^{w_S}[p_C^z(a_n|m_i)]^{w_C}}{(\sum_{j=1}^{M} p_S^z(a_n|m_j)]^{w_S}(\sum_{j=1}^{M} p_C^z(a_n|m_j)^{w_C} + o_C + o_L} \tag{16}$$

Outlier biases o_C are calculated with Eq. 17 and o_L with the outlier probability $\frac{\alpha}{1-\alpha}\frac{M}{N}$

$$o_C = \frac{M}{\sigma_C\sqrt{2\pi}} \cdot exp^{-\frac{1}{M}\frac{\left\|\sum_m^M P_C^o(a_C|m_i^C)\right\|^2}{2\sigma_C^2}} \tag{17}$$

The general process of registration is summarized in the next pseudo-code Algorithm 1. Since we focus on modifying the matching probability (P^o), the

general procedure is similar to the original CPD, but with modifying step E:

Data: M and A pointsets, color M and color A information

Initialization: $W = o, \sigma^2 = \frac{1}{DNM} \sum_{m,n=1}^{M,N} ||x_n - y_m||^2$;

Construct G: $g_{ij} = exp^{-\frac{1}{2\beta^2}||y_i - y_j||^2}$;

Expectation-Maximization

while *not converged* **do**

E-step: Compute P^o, (contribution)

- $P^o(m_i|a_n) =$

$= \frac{[p_S^z(a_n|m_i)]^{w_S}[p_C^z(a_n|m_i)]^{w_C}}{(\sum_{j=1}^{M} p_S^z(a_n|m_j)]^{w_S}(\sum_{j=1}^{M} p_C^z(a_n|m_j)^{w_C} + o_C + o_L}$ (see Eq. 16);

M-step:

- Solve $(G + \lambda\sigma^2 d(P1)^{-1})W = d(P1)^{-1}PX - Y$ (see [1]);

end

The result of alignment: $T = \tau(Y, W) = Y + GW$;

Algorithm 1: Pseudo-code of the proposed Color Coherent Point Drift

4. Experiments

A set of tests have been carried out to evaluate the performance of the proposed CCPD compared to the original version. First, the dataset of the original CPD (Subsection 4.1), the fish and the face, has been used (Figure 1). The implementation of the code has been done in Matlab, using part of the toolbox provided by Myronenko [1]. Color information has been added to the original data. The distribution of colors on the shape has been done in the way to distinguish its different parts, i.e. a region with same color corresponds to a specific part of the shape (e.g. mouth in the face, or tail in the fish). It is important for the non-rigid registration with color because it gives meaning to the relationship between color and shape.

The second test (Subsection 4.2) presents two synthetic datasets with realistic color and shape (Figure 9). A face and a flower are used, which have been

[1] www.bme.ogi.edu/∼myron/matlab/cpd

deformed using Blender and acquired using a plugin called Blensor [29]. This plugin emulates different sensors, including the general purpose RGB-D sensor Kinect.

Finally, a real data evaluation using data provided by a Primesense Carmine RGB-D sensor is done in Subsection 4.3 to confirm that the algorithm is able to handle real data acquired from a general purpose RGB-D sensor (Figure 19).

In this section we will use X to refer the *Anchor* set and Y to refer the *Moving* set.

The experimets evaluate different aspects:

- Outliers: points which are in *Anchor* X but do not have real matching in *Moving* Y.
- Missing data: the opposite of outliers. Points which are in *Moving* Y but do not have real matching in *Anchor* X. This situation is not taken into account in the original CPD algorithm.
- Large or non-linear deformation: deformations which involve a large displacement that may not be solved with traditional algorithms. Non-linear deformation could be seen as an abrupt change in the direction of the deformation.

The experiments, used a Windows 7, an Intel i5 processor and 8 GB of RAM. The code was implemented in Matlab vR2013b.

4.1. Synthetic data experimentation

The tests consider four issues: outliers, missing data, color distribution changes and large deformations. First, points from Y are removed. With this test the missing data handling is compared with the original CPD algorithm. Next, we remove data from *Anchor* X representing extra points, a situation which is not possible to parametrize in the original CPD (points in Y do not have a real correspondence in X). In this case, CCPD uses the color information to improve the probability evaluation to avoid wrong matches. Another test

evaluates a different displacement in the color with respect to the shape, which evaluates situations where the color distribution in X and Y are different. An example of this could be moving the eyebrows up and down, where the shape in 3D remains the same, but the color changes its position. Lastly, large deformations are evaluated to show how the color facilitates the perception when the transformation is complex or semi-coherent. It is important to highlight that the parameters have been adjusted individually to result in the best alignment for both the CPD and CCPD algorithms.

The main difference between the original CPD and the proposed CCPD method comes when the *Moving* has missing data, which cannot be modelled as outliers in the CPD. As the color is a distinctive feature, the proposal is able to evaluate the correspondences properly and then provide better results.

4.1.1. 2D fish experimentation

The 2D tests use different *Anchor* X and *Moving* Y fishes based on two initial shapes (dataset from the original work of CPD [1]). Nine colors using the H component of HSV are used to distinguish the different parts of the fish (see Figure 1).

Table 1 presents the RMS error of the registration taking into account euclidean distances of real correspondences in location space. Figure 2 shows the visual result of the tests. In general, the registration achieves better alignment (minimize the error distance) in the CCPD results. Test 1 evaluates the effect of outliers by removing in Y the top and bottom tip of the fish. In this case, the proposal returns a slightly better registration because the color feature provides a more robust matching estimation and hence registration. Test 2 and Test 3 correspond to missing data testing where points in *Anchor* X are removed, while Y remains complete. The total amount of points is 91. For Test 2, 20 points are removed ($20/91 = 21.9\%$ of outliers) and for Test 3, 53 points are removed ($53/91 = 58.2\%$ of outliers). The results demonstrate the improved performance of CCPD in the alignment against CPD. Concretely, in Test 2, CCPD achieves 0.747E-02 RMS error in registration being 4.82 times lower RMS than CPD,

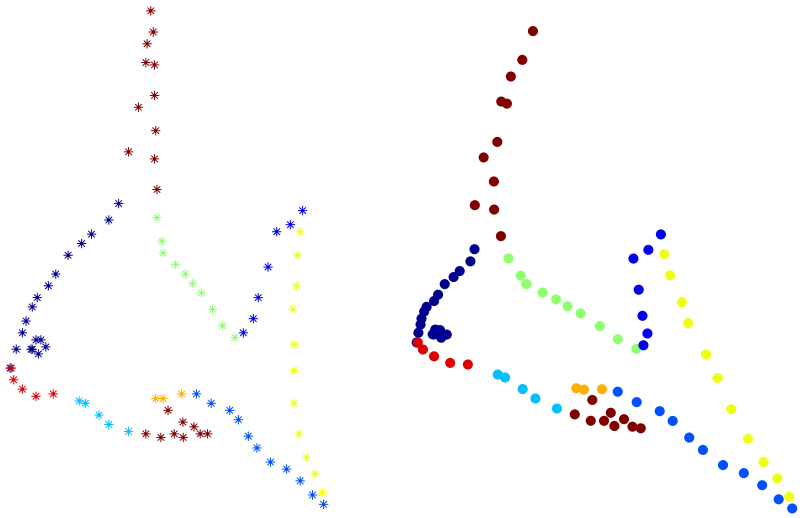

Figure 1: *Anchor X* (left) and *Moving Y* (right) fishes based on the original work of CPD [1] including color information.

while in Test 3, CCPD achieves 0.624E-02 RMS error being 23.1 times lower than the original method. CCPD is more robust against outliers in the *Moving* 380 *Y* (or missing data from the Data point of view).

A large deformation test has been considered by registering a square to the *Anchor* fish in Test 4, where Matlab *jet* colormap is used. This color map provides colors in RGB = [0,0,0.562] to [1,1,0], which in H component used here are H = [0 0.0625 0.1250 0.1875 0.2500 0.3125 0.3750 0.4375 0.5000 0.5625 385 0.6250 0.6875 0.7500 0.8125 0.8750 0.9375 1.0000]. The RMS error is 26.62E-02 in the CCPD method and 51.559E-02 in the original CPD, a 93.69% of improvement of CCPD against CPD. Furthermore, CPD on the low tip of the back tail (Figure 2 fourth-row right-image) misaligns the colors as it does not have this information, which also demonstrates the improvement in registration 390 accuracy of the proposed color feature consideration in the registration process.

The next test evaluates changes in the color distribution. In this situation both shapes have the same points as the original, but the colors are slightly

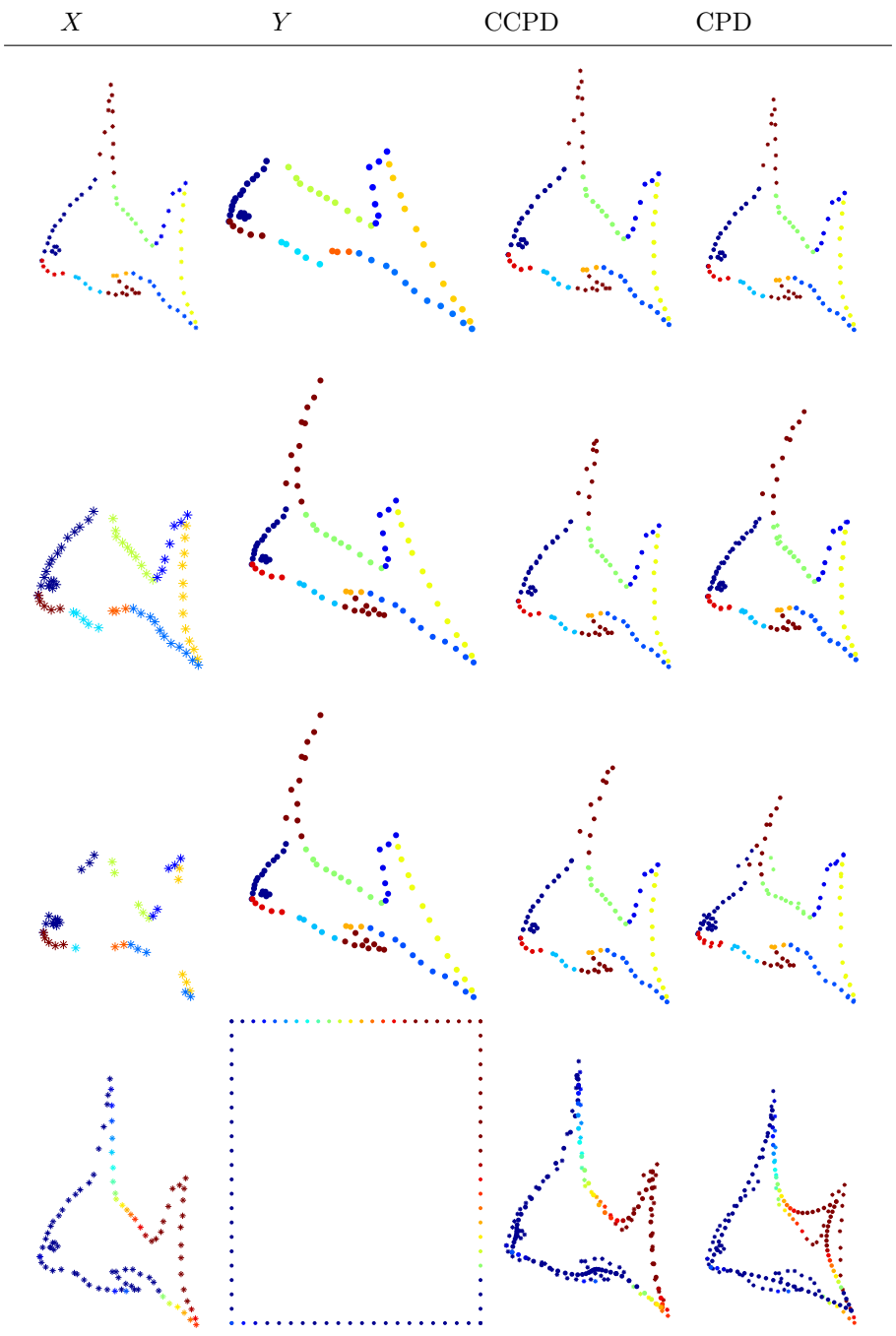

Figure 2: Tests 1 to 4 of fish shape from top to bottom respectively. The columns represent from left to right the *Anchor* X, the *Moving* Y, the CCPD registration result and the CPD result.

Table 1: RMS registration error of fish shape tests.

	CCPD	CPD
Test 1	0.52064E-02	0.53293E-02
Test 2	0.7468E-02	3.5967E-02
Test 3	0.6239E-02	14.406E-02
Test 4	26.622E-02	51.559E-02

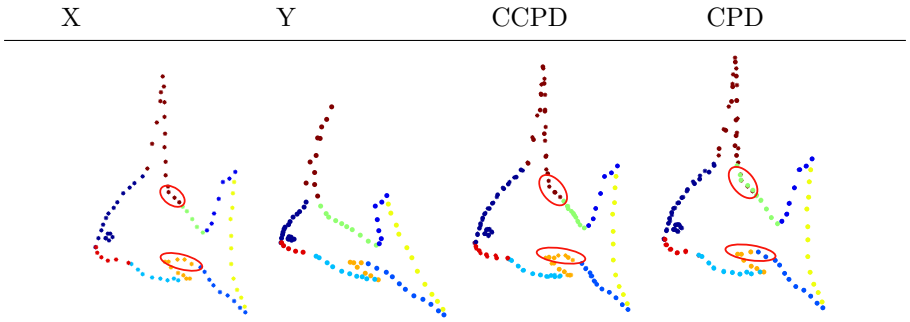

Figure 3: Registration result for different color distribution in *Anchor X* and *Moving Y* sets (see Fig. 2). The third and fourth columns are the results for the CCPD and the CPD algorithms. The red circles highlight the parts where the color distribution changes.

different. The result is visually evaluated in Figure 3. The regions where the colors do not coincide are marked with a red circle to simplify the visualization. At the lower part of the upper tip, X has larger region of brown towards the back while Y is green from the end of the tip. CCPD registers adequately this part. Similarly, the lower tip has larger part of orange on the X than in Y, and again the proposed method achieves better results.

4.1.2. 3D face experiments

The 3D face experiments are presented here. Different *Anchor X* and *Moving Y* points are used based on two initial positions (data obtained from the original work of CPD [1]). The face coloring has been done using four tones in RGB. A main black part, red lips and eyebrows, blue ears and yellow forehead (see Figure 4).

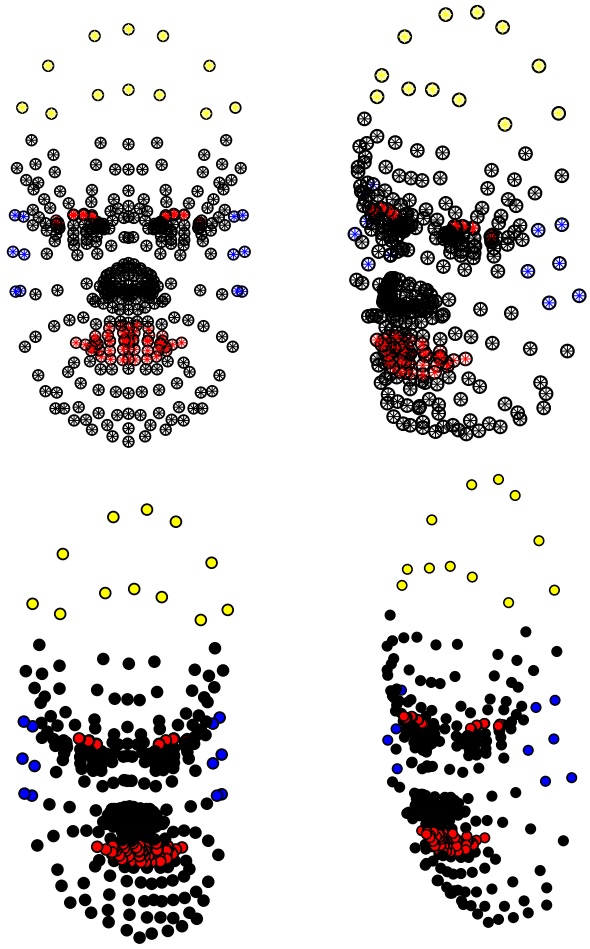

Figure 4: *Anchor X* (first row) and *Moving Y* (second row) of face shape. There appear 4 colors, yellow in the forehead, red in eyebrows and lips, blue in ears and the rest black.

Table 2 presents the RMS error for the 3D tests. In Test 1 outliers handling is evaluated by removing all data points from the forehead (yellow part) of Y. Tests 2, 3 and 4 correspond to the missing data evaluation. Test 2 is similar to Test 1, but removing the data from X. As the unmatched data cannot be parametrized as outliers, the original CPD is not able to register it properly. Test 3 removes all color parts except the black one obtaining a better results for the CCPD proposal. Finally, in Test 4 the algorithm registers the non-black parts (i.e.: forehead, ears, lips and eyebrows), in the *Anchor* X with the complete *Moving* Y. Similarly to the 2D experiments, the proposed method is able to register more accurately.

Figure 5 shows the result of the tests, where each row is a test from 1 to 4 respectively, to visually evaluate the performance of both methods. In the second row it is possible to see how CPD moves wrongly yellow points downward while the proposed method keeps the point in the top part as they do not have correspondences. The third row has only color points in the *Anchor* X, without the black part. The proposal aligns properly these remaining parts while CPD cannot align the parts properly. Similarly the fourth test is correctly aligned by CCPD as the corresponding points in the *Anchor* and *Moving* are properly aligned, while CPD returns an inaccurate result.

A large test evaluation is presented in Table 3 where a set of 50 different changes are registered (dataset available from Myronenko [1]). The average RMS errors for the Tests 2, 3 and 4 are 0,36834E-02 for CCPD and 8,8453E-02 for CPD, then the proposal is 24 times lower than original method.

Table 2: RMS registration error of face shape tests.

	CCPD	CPD
Test 1	0.37278E-02	1.62E-02
Test 2	0.28985E-02	4.078E-02
Test 3	0.21677E-02	12.051E-02
Test 4	0.5984E-02	10.407E-02

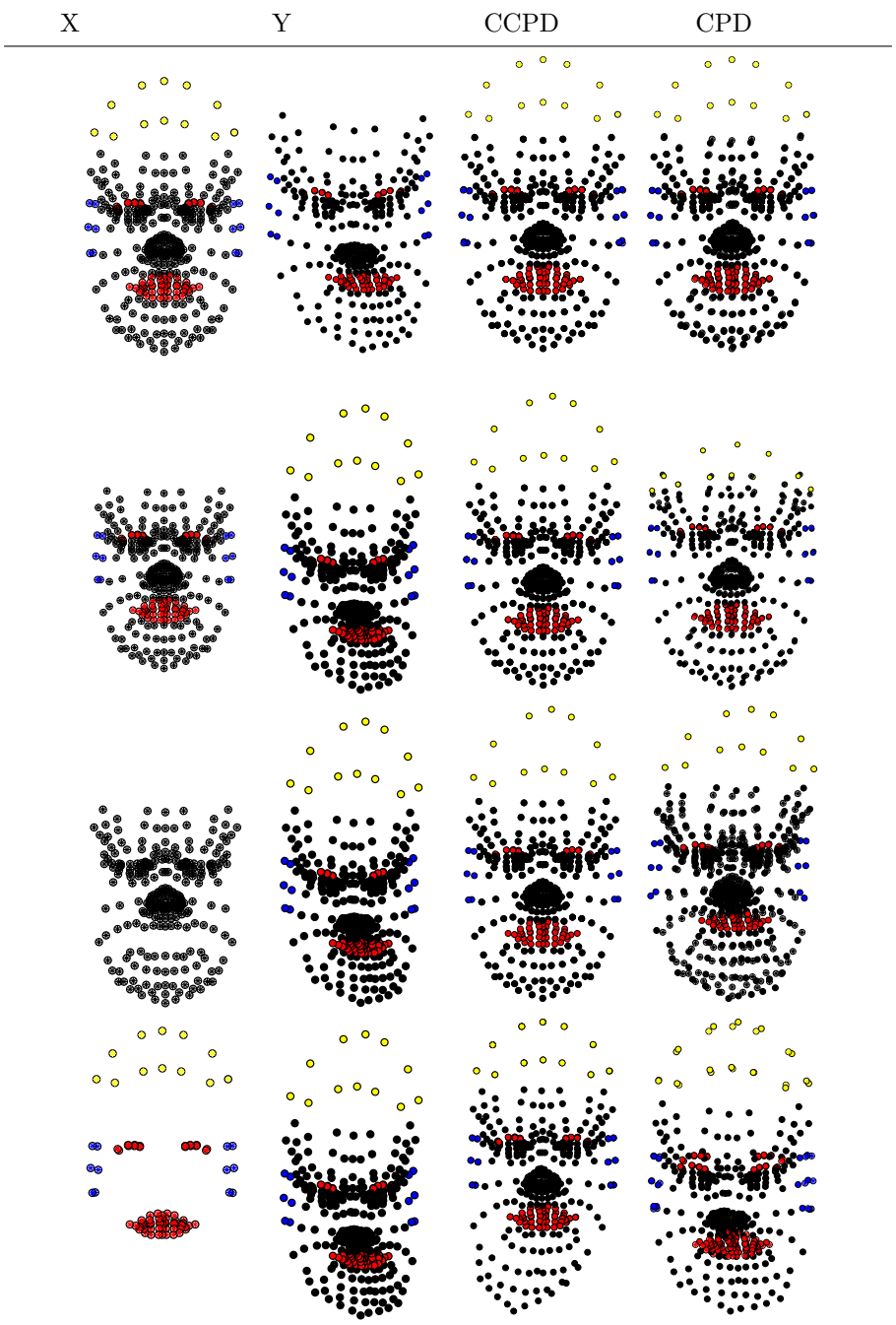

Figure 5: Rows 1 to 4 show Tests 1 to 4 of face shape. The columns represent from left to right the *Anchor* X, the *Moving* Y, the CCPD result and the CPD result.

Table 3: RMS registation error of 50 face shapes.

	CCPD	CPD
Test 2	0.26E-02	3.93E-02
Test 3	0.32E-02	8.53E-02
Test 4	1.13E-02	11.54E-02

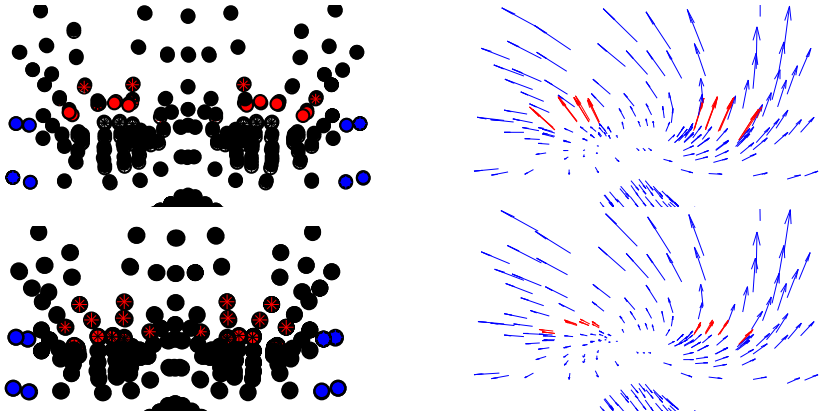

Figure 6: Eyebrow movement test. From the top to the bottom, the CCPD result, the CCPD flow, the CPD result and the CPD flow. In the flow pictures (2nd and 4th) the red arrows show the most significant displacement, i.e. the eyebrows.

A final test was carried out to evaluate a displacement of color and a large deformation. In this test, the eyebrows of Y are lower than in X. The movement should displace the eyebrows upward. This is considered a large deformation or a non-linear deformation as the movement is not coherent in the shape data space, but coherent in the color data space. Figure 6 shows this registration. In order to help in the visualization, a flow image is shown for both methods. The proposed method achieves a proper result moving up the eyebrows while the original CPD algorithm, as it does not take into account color, is not able to achieve the correct result.

4.1.3. Experiments for noise and outliers in color space

In this section we are going to evaluate the effect of noise and outliers in the color space for the non-rigid registration with the proposed CCPD. The

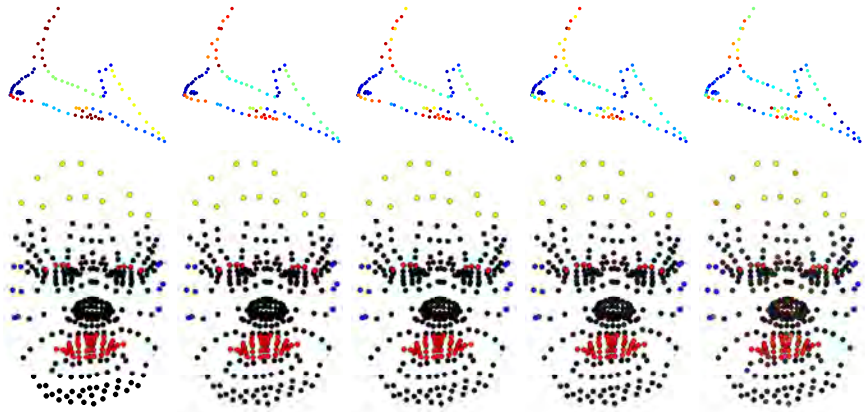

Figure 7: Noisy point-clouds of the fish and face corresponding to Y the moving data. From left to right, the data without noise, 20, 15, 10 and 5 dB of SNR.

experimentation is carried out using the fish and face data used in the previous experiments.

The first experiment considers the noise in the color space adding random gaussian noise to each R, G, and B component in 4 different levels of Signal/Noise ratio (SNR): 20, 15, 10 and 5 dB (see Fig. 7). Initially, the experiment analyses the effect of choosing the adequate parameters for CCPD to compensate the color noise using the fish data. Lately, using the face data, the parameters are fixed to those the best perform CCPD for the experiments carried out in Sect. 4.1.2 in order to analyse the color noise effects and tolerance of the proposal against that noise. Since the noise is assigned randomly, 5 iterations per level of noise have been performed to calculate the averaged RMS as the registration error.

The results for the initial experiment of color noise using the optimal set of parameters by experimentation are shown in Table 4. As can be proved by the results, even with high levels of noise, the performance of the CCPD method remains high being hardly affected by the color noise (the order of the RMS is the same regardless the SNR).

The results for the second experiment of color noise (using the optimal set of parameters obtained for the CCPD without noise presented in Sect. 4.1.2)

Table 4: RMS registration error of fish shape with color noise. The Signal-to-Noise ratios are 20, 15, 10 and 5 dB.

20	15	10	5
0.41080e-02	0.40603e-02	0.59438e-02	0.69721e-02

are shown in Table 5. The data includes noise in the same four tests shown in Table 2 and Figure 5. For Tests 2 to 4 with 15 dB of SNR the error of CCPD is similar to CPD and decreases. In the case of Test 1, the performance is lower than CPD but remains similar for every level of noise due to the outliers are modelled with the original Eq. of outliers o_L from the CPD. If we assign high σ_C or low w_C, we will have the results similar to CPD. Furthermore, Figure 8 presents the same experimentation for the set of 50 faces that is part of the original CPD synthetic dataset. This experimentation has a similar behaviour as the previous one, confirming the results in a large set of deformations. In average, the CCPD method outperforms the CPD results even for a large color noise (about 15 dB).

Table 5: RMS registration error of face shape tests with 20, 15, 10 and 5 dB of Signat-To-Noise ratio.

	CCPD	CPD	20	15	10	5
Test1	0.0037	0.0162	0.2176	0.2141	0.2150	0.2159
Test2	0.0029	0.0408	0.0308	0.0617	0.1823	0.4560
Test3	0.0022	0.1205	0.0332	0.0830	0.2606	0.4304
Test4	0.0059	0.1041	0.1365	0.1372	0.1483	0.2092

Finally, the effect of outliers in color data is evaluated. In this case, the outliers are in the color space, hence to generate them we have chosen the color that is the furthest to the rest of colors, which in this case is white. We have randomly generated, over the data set, four percentages of outliers: 5%, 25%, 50% and 75%. The results are presented in Table 6.

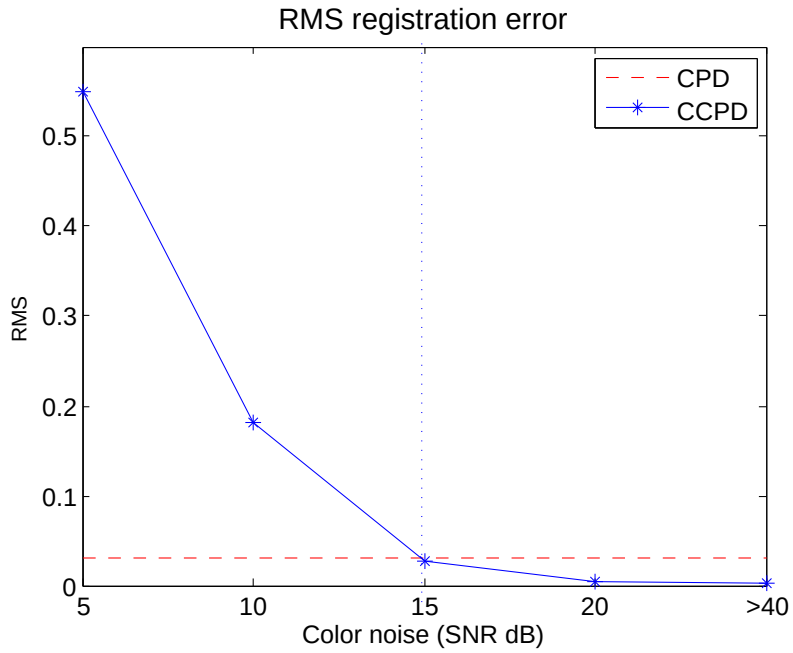

Figure 8: RMS registration error of 50 face deformations with 20, 15, 10 and 5 dB of Signat-To-Noise ratio.

Table 6: RMS registration error of face shape for four levels of color outliers, 5%, 25%, 50% and 75% compared to the CCPD without noise (BL: Baseline).

BL	5%	25%	50%	75%
0.0029	0.0222	0.0783	0.1311	0.2579

4.2. Synthetic realistic experiments

In this section, we present the experiments to evaluate the method for non-rigid registration using realistic shapes. The dataset includes two different objects: a flower[2] and a face[3]. The synthetic models have been acquired using the Blensor tool [29], a Blender plugin which simulates a Microsoft Kinect RGB-D sensor. This tool uses raytracing to simulate 3D sensors, that in this particular case is an RGB-D, providing a PCD file with all the spatial coordinates of the points and the color information. The virtual sensor is oriented in the direction as it would be done with a real one. The only preprocess is to deforme the models using the Blender tools, to have in this case three shapes, origin, small deformation and large deformation.

Figure 9 and 10 show the face and flower models used for the experiments. The images are from left to right: the target, a first deformation, and a second larger deformation. The face deformations could be seen as elastic deformations, because the face remains the same except displacement of some parts. The first deformation is a eyebrow rise and a mouth change. The second moves both eyebrows and the mouth, changes the nose and the chin. For the flower, it could be seen as growth deformations due to the size of the object changes. The first deformation enlarges a little the leaves and the second is a larger deformation.

In order to reduce and enhance the data for the registration purpose, we have used different downsampling techniques to sample the data. Figure 11

[2] `https://www.turbosquid.com/3d-models/pink-primrose-flowering-3d-obj/516226` (last access: 11/08/2017)

[3] `http://eat3d.com/forum/art-gallery/models-face` (last access: 11/08/2017)

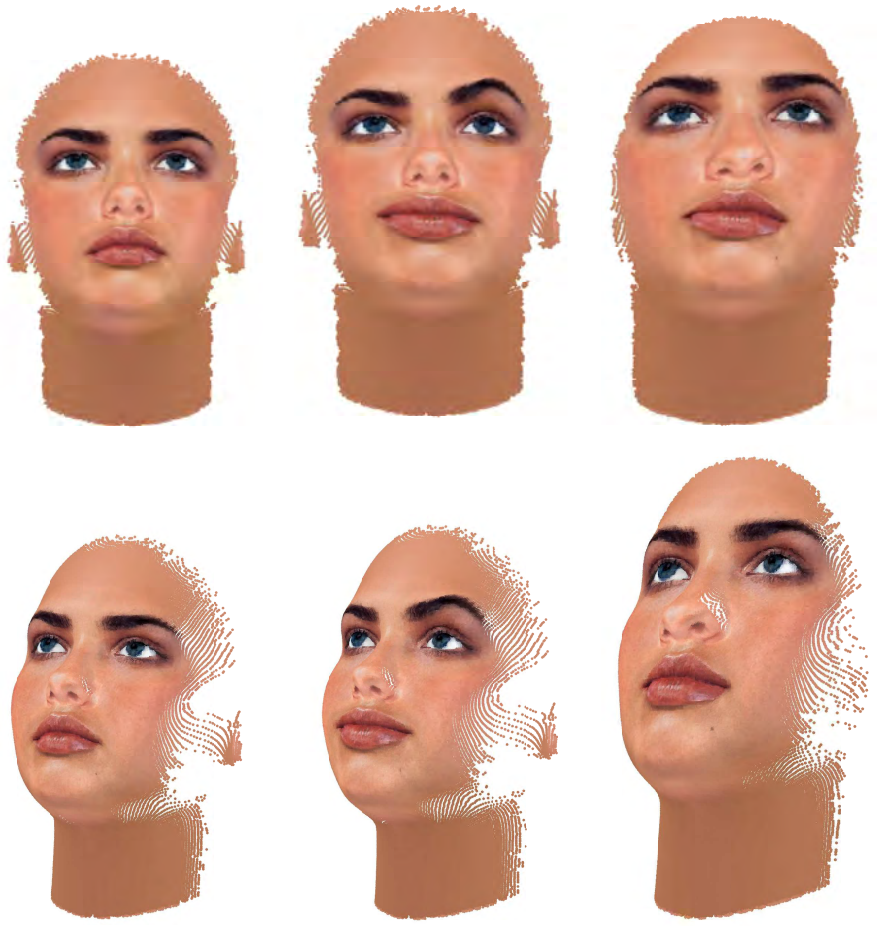

Figure 9: The face model used in the experimentation. Two viewpoints (each per row) of the faces used. From left to right, the original face shape as target for the deformations in second and third columns.

Figure 10: The flower model used in the experimentation. Two viewpoints (each per row) of the flower used. From left to right, the original flower shape as target for the deformations in second and third columns.

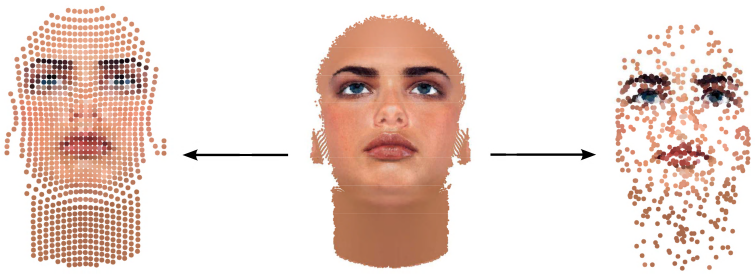

Figure 11: Two sampling examples. The image in the middle represents a point set of a face shape. At the left is a uniform sampling. At the right side is a representation of a color-based sampling, which provide higher density of points at salient features, such as eyebrows, eyes or lips.

shows two different kinds of sampling. The figure has in the middle a face example. In the left side a uniform sampling is presented, while in the right side a representation of a color-based sampling, which provide higher density of points at salient features, such as eyebrows, eyes or lips. In previous works, we have studied the use of downsampling as a method to enhance the quality of the data. This studies have been published in [30] [31]. In this paper we use the same methods, including: bilinear interpolation, normal-based sampling, color-based sampling, a combination of color and normal based technique, and GNG sampling approach proposed in [32].

4.2.1. Non-rigid registration evaluation

Here we present a comparative evaluation of CCPD and CPD registration using synthetic realistic subjects. The color information, used by CCPD, allows the registration method to achieve good results in accuracy when the surface is not very detailed where the drift of points is not constrained by the irregularities of the shape.

Using the data sampled, the non-rigid registration methods are qualitatively evaluated by visual inspection. Figure 13 shows the face shape for CCPD and the original CPD with 1000 points sampling with the different methods. More

experiments have been performed with 250 and 500 data (which correspond to the experimentation in [30] [31]), but are not included as the results are similar to the presented experiment.

Figure 16 shows the flower shape for CCPD and the original CPD with the same point sampling (similarly, more experiments have been done with similar results). The figures show the registration for the second deformation (right of Fig. 9 10) of each shape as it is the larger one, and hence, the most difficult in terms on registration procedure. For each figure, the first row presents the CCPD method and the second the original CPD. From left to right, the sampling techniques are: bilinear, normal-based, color-based, NC-based, and GNG.

To analyse the registration, we will pay special attention to a specific Region-of-Interest (ROI) for each model (i.e. those parts that are the aim of the study) depicted in Figure 12. In the face, the ROI will correspond to the mouth and eyebrows as they are the parts which are mainly displaced. The ROI in the flower will correspond to the central part, pink and yellow, as they do not deform in color unlike the rest of the leaves (i.e. the deformation produces an enlargement of the tip of leaves, but the center remains the same). This simulates the growth of a flower, where not all parts grow in the same way. Figure 13 and 16 show the registration result for all different sampling techniques using CCPD (first row) and CPD (second row). Figures 14 and 15 present a detailed view of this analysis for the face shape, and Figure 17 and 18 for the flower shape. For both shapes first figure shows the registration using a GNG sampled dataset and the second the bilinear sampled dataset.

We can conclude different aspects from the results of the experiments for the face shape. From Figures 13, 14 and 15 we can see that the proposed CCPD achieves better alignment. If we pay attention to the eyebrows area, it could be seen that the alignment of CCPD algorithm register better as it takes into account the color. In the detailed figures, it is easier to perceive this situation.

The flower shape presents a similarly behaviour to the face in the registration results. When the data comes from either color-based or NC-based, both CCPD and CPD achieves similar results. Moreover, when the data has been sampled

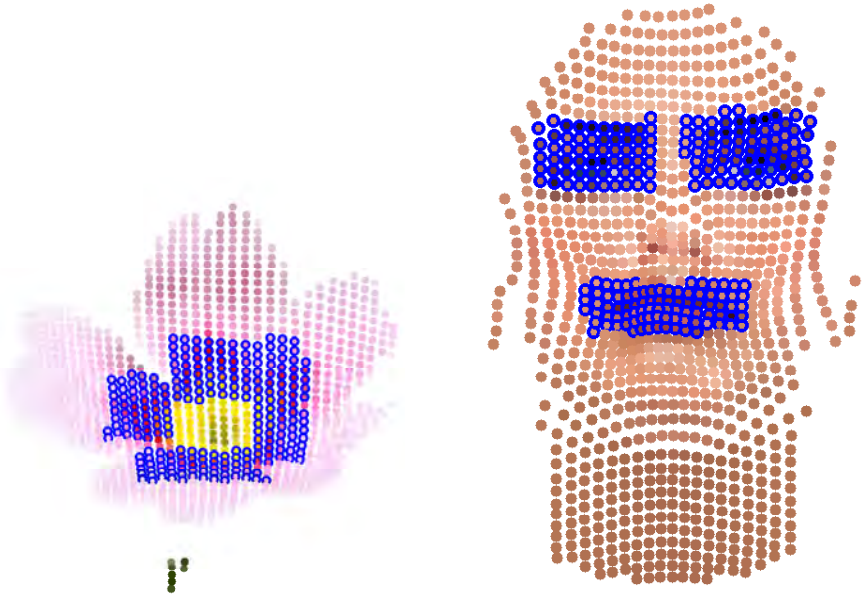

Figure 12: Example of Region of Interest for both shapes. The ROI are highlighted with blue circles.

Figure 13: Non-rigid registration result of face shape for a 1000 point sampling. The first row shows CCPD, and the original CPD in the second. Columns show different sampling algorithm that are from left to right, bilinear, normal-based, color-based, NC-based, GNG.

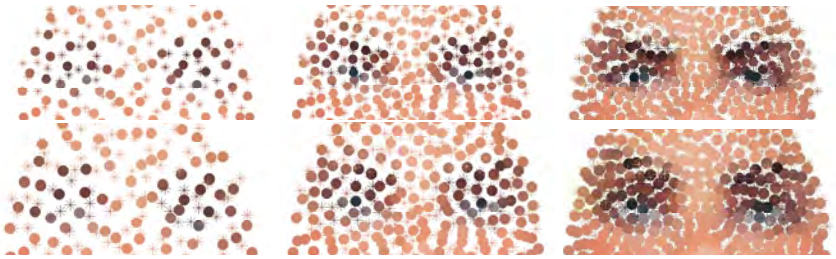

Figure 14: Enlarged example of the ROI for the face sampled with GNG. The first row shows the CCPD and the second the original CPD. The data size is, from left to right, 250, 500, and 1000 points for the GNG.

Figure 15: Enlarged example of the ROI for the face sampled with bilinear. The first row shows the CCPD and the second the original CPD. The data size is, from left to right, 250, 500, and 1000 points for the bilinear.

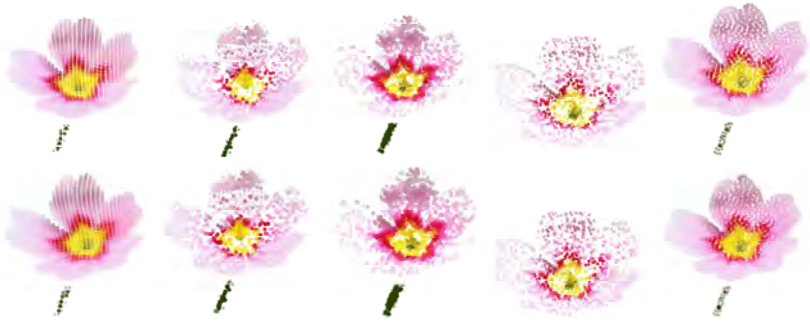

Figure 16: Non-rigid registration result of flower shape for a 1000 point sampling. The first row shows CCPD, and the original CPD in the second. Columns shows different sampling algorithms that are from left to right, bilinear, normal-based, color-based, NC-based, GNG.

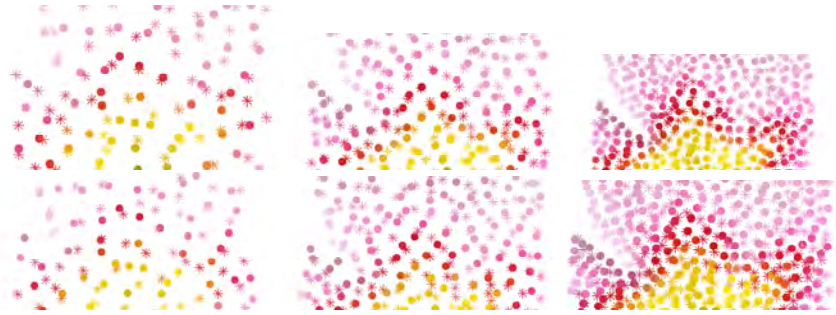

Figure 17: Enlarged example of the ROI for the flower sampled with GNG. The first row shows the CCPD and the second the original algorithm. The data size is, from left to right, 250, 500, and 1000 points for the GNG.

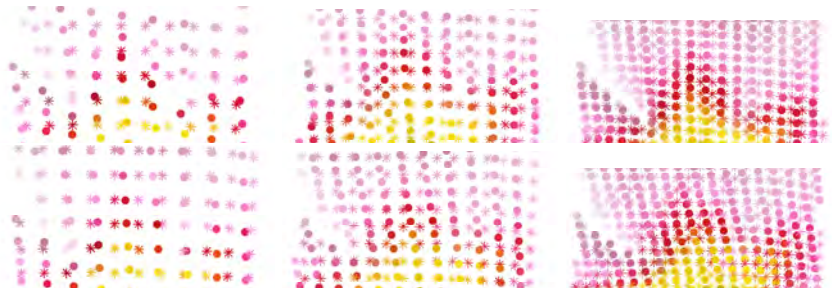

Figure 18: Enlarged example of the ROI for the flower sampled with bilinear. The first row shows the CCPD and the second the original algorithm. The data size is, from left to right, 250, 500, and 1000 points for the bilinear.

using GNG or bilinear, the proposed CCPD achieves higher registration accuracy than the original CPD. As the deformation in this shape is not isometric, the tips of some leaves are the parts that get larger compared to other, CCPD moves the points differently in the tip of the leaves than the ROI, achieving accurate results. However, as CPD moves coherently, the points shrink all together (the registration is from the larger to the original position) producing that ROI ends in a wrong color alignment. This situation is presented in Figure 17 and Figure 18.

Figures 17 and 18 show a detail of the registration to visually evaluate the accuracy of both methods. It is easy to appreciate that CCPD achieves better

Table 7: Average processing time in seconds CCPD and CPD in face shape.

	Face deformation 1			Face deformation 2		
CCPD						
	250	500	1000	250	500	1000
bilinear	11.7988	58.4475	560.6209	30.2463	98.0815	654.7234
normals	17.0563	114.3599	555.5070	24.3791	62.2356	873.7409
color	11.7048	58.2177	703.9807	19.7996	69.2475	809.8889
NC	23.2824	141.5593	465.4034	25.8104	70.3541	700.6084
GNG	35.1574	121.0136	541.7851	31.5702	181.0353	698.8645
CPD						
bilinear	3.4144	13.3339	58.8231	9.0456	44.0501	175.9423
normals	7.5732	31.0907	173.2305	12.1042	45.0169	181.4115
color	5.501	27.6411	174.1651	7.9162	29.5334	135.8508
NC	12.4572	36.2307	149.8286	7.1914	42.8539	193.599
GNG	11.3055	43.2056	171.6053	11.6725	44.3024	183.5948

results than the original version in the alignment.

Finally, processing time of the registration process has been evaluated and shown in Table 7 for the face model execution, and Table 8 for the flower. The original CPD always achieves lower times due to the number of operations. To calculate the posterior probability in CCPD, it is necessary to estimate for each point both color and location probability. Moreover, the convergence is not the same in both methods, as CCPD commonly needs more iterations to achieve a more accurate result. The time of both tables is presented in seconds, and is shown for each sampling method. The columns are: first, the sampling method; from second to fourth, the sampling rates for the first deformation; and from fifth to seventh, the three sampling rates for the second deformation.

4.3. Real data experimentation

To evaluate the method in real conditions, experimentation with data from a general purpose RGB-D sensor has been carried out. In this case, a face with

Table 8: Average time processing CCPD and CPD in flower shape.

	Flower deformation1			Flower deformation2		
			CCPD			
	250	500	1000	250	500	1000
bilinear	33.8404	100.1836	560.1712	13.4003	153.0062	562.0116
normals	18.9529	74.6561	369.0443	20.0593	121.5026	537.5028
color	13.0186	129.9349	458.0764	21.498	74.1095	361.1847
NC	52.0435	91.3152	388.7584	66.038	274.563	267.593
GNG	20.7716	112.4218	443.8667	18.5236	66.7437	615.4628
			CPD			
bilinear	9.8136	43.4829	175.2797	11.2072	43.2063	172.0116
normals	11.4153	45.7999	181.3168	11.8425	42.7337	172.5745
color	11.5049	42.8385	175.7153	9.9246	42.8615	175.9582
NC	11.79	43.2504	174.0308	11.2697	43.7694	174.9237
GNG	11.8023	45.2732	176.094	11.1835	43.7019	171.5101

different expressions is used to evaluate the non-rigid registration using CCPD, against CPD. Due to the absence of ground truth, the data will be visually evaluated to analyse the performance of both methods. Figure 19 shows the data used in this experimentation.

Figure 20 shows an eyebrow rising deformation. The target is a neutral expression and the deformation is a surprise expression. The registration results of CCPD accurately aligns the shapes. The right column shows the data flow. It clearly shows the movement of the eye region downward, from the surprise expression to the neutral one. In this case, CPD only takes into account the location, then it cannot align properly the eyebrows, resulting in a wrong homogeneous displacement.

Figure 21 shows a cheek inflating deformation. The origin inflates one cheek so the mouth also moves to the side, the target is a neutral expression. The CCPD outperforms the registration of CPD as it uses the beard color to properly align and move the points into a correct location, where correct means the color

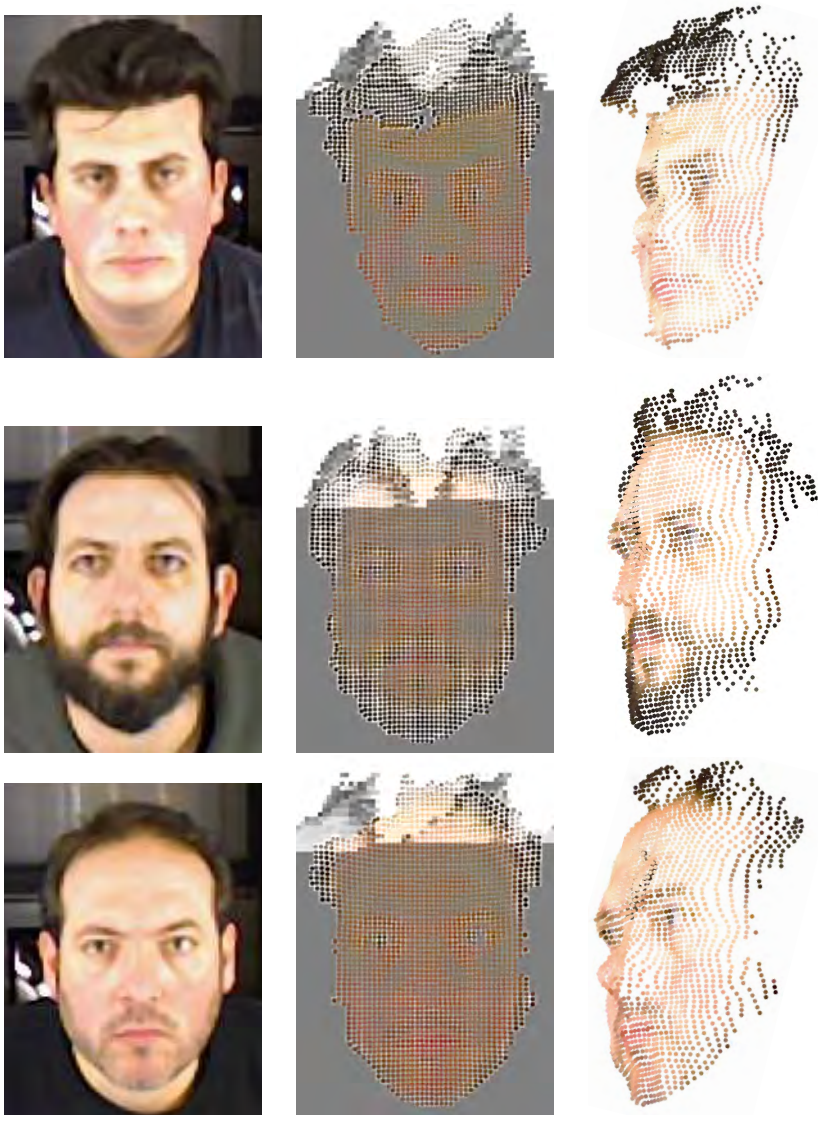

Figure 19: Real data used for the non-rigid experimentation. The first column shows the original color images, the second and third show the 3D point cloud data from front and side of the faces.

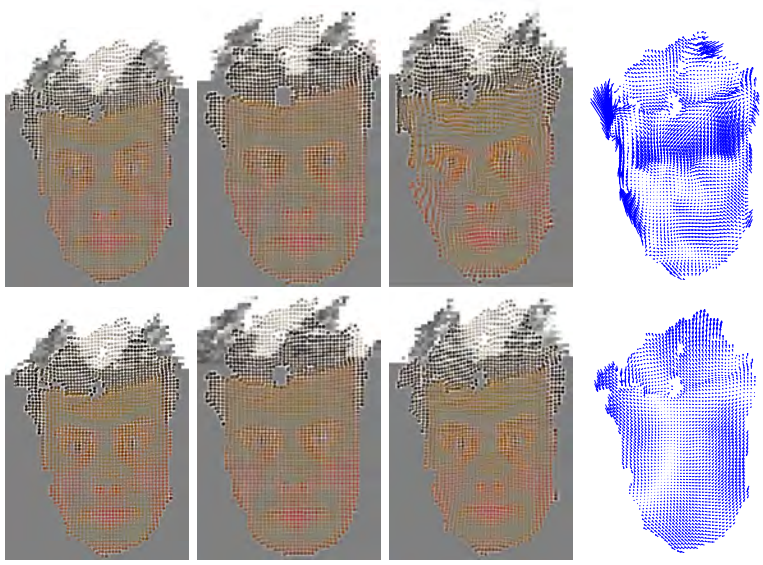

Figure 20: Real data registration, eyebrow rising. The top row is CCPD, and the second CPD. From left to right, original deformation, target shape, registered shape, final alignment and data flow

of both X and Y registered are the most similar over the data. CPD, despite the good result, ends in a non accurate registration because it can only perceive the location information.

Figure 22 presents a large deformation. Here the face is highly deformed to a side and closing an eye. CCPD aligns the points better because the registered point set results in a correct location. CPD, however, cannot move correctly the points resulting in an inaccurate result (points registered have different color).

5. Discussion and conclusions

In this paper, a novel non-rigid registration approach called Color Coherent Point Drift (CCPD) is presented. This proposal, based on the well-know CPD, introduces color information in the correspondence estimation of non-rigid registration. The combination of color and location (3D position) information in the estimation correspondence for alignment improves the result in the presence of noise, missing data and outliers.

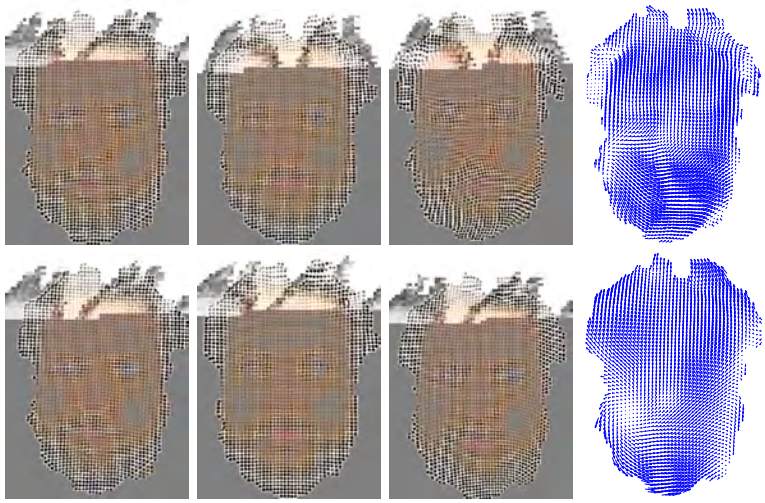

Figure 21: Real data registration, left cheek inflation. The top row is CCPD, and the second CPD. From left to right, original deformation, target shape, registered shape, final alignment and data flow

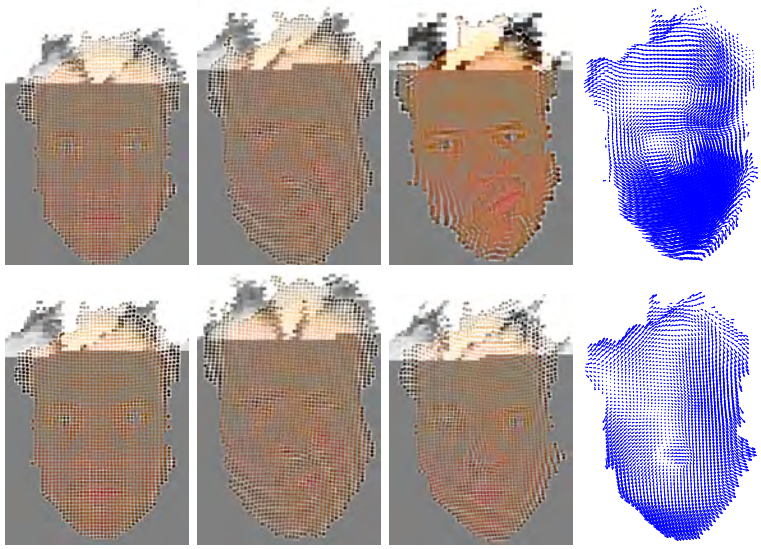

Figure 22: Real data registration, large deformation. The top row is CCPD, and the second CPD. From left to right, original deformation, target shape, registered shape, final alignment and data flow

In general terms, the proposed CCPD outperforms the original CPD in most cases. The new input, color, provides information that disambiguates situations where the 3D space provides the wrong correspondences. For example, a flower that grows is used because some parts remain the same but the tips of the leaves expand. Here, CPD returns a coherent movement which moves centre the points to a wrong position, while CCPD keeps the color in a good registration.

The experiments included three parts: synthetic simple subjects, synthetic realistic subjects and real data. The simple subjects are those used in the original CPD but with added color information (a fish and a face). The realistic subjects have been obtained using Blensor, and the real data has been acquired using a Primesense Carmine RGB-D sensor. The first experiments with a fish and a face shape show how the proposed method is able to overcome noise, outliers, missing data and large deformations. To evaluate the outliers and missing data, first the registered dataset Y is aligned to the target dataset X, this second set with outliers (points in X without correspondences in the registered set Y), providing similar result for both CCPD and CPD. Secondly, missing data evaluation has been carried out by removing points in X, so that there are points in Y without correspondences in X. In this evaluation, for the fish shape, CCPD had 4.82 times lower RMS error than CPD in registration accuracy for 21.9% missing data and 23.1 times lower for 58.2% missing data. For the face, CCPD had 24 times lower RMS error on average for all missing data tests than the original method. For a large deformation evaluation, a square shape was registered to the fish shape, obtaining better alignment by CCPD than by CPD for the RMS error (23.1 times lower RMS). In the case of the face, the large deformation moves the eyebrow upper while the rest of face remains the same, which forces a non-coherent movement in a specific region. It has been visually evaluated with CCPD outperforming CPD.

Experimental results show that a balanced adjusting of both color and location parameters, using the proposed CCPD, allows to meet the requirements of a specific problem, dealing with difficult conditions of the input data (very high level of noise and outliers in color or location space), approaching to the optimal

solution. Nevertheless, CPD including color information (CCPD) outperforms the registration process even taking into account very difficult conditions of color input data. In the worst case, in presence of corrupt color data, the CCPD can become the original CPD with either large σ_C or assigning 0 to w_C.

For realistic data experiments, two subjects have been evaluated, a flower and a face. Both subjects have two deformations, one larger than the other. The face changes the shape as expressions, then eyebrow, and mouth are the regions that mainly deform, which can be studied as an elastic deformation. The flower, performs the growth of some leaves, which can be seen as a free deformation as the subject changes both size and topology as new points appear in the deformation. CCPD has been evaluated and compared to CPD using the data provided by five downsampling methods which were used in previous works. The results have been visually evaluated, showing more accurate registration for the proposed method in most cases. The subjects, for all data (each downsampling method), are aligned not only by the point distribution, but also with a coherence in the color space (similar colors are aligned together).

The real data includes three face deformations, from smaller to larger, returning more accurate registration results for the proposed method. The deformations of the shapes are better aligned by CCPD because the flow of the points is more similar and coherent to the expected (expected by visual inspection), by aligning the points using the shape and color information.

Generalization for multiple (e.g.: include topology along with color and location) spaces combination is the next step to be done. Moreover, evaluating biological growth using CCPD is a short term future work that will provide a very useful tool for many applications. As long term future work, we are interested in modifying the method to accelerate the process by comparing neighbour points instead of the whole data set. Moreover, an implementation of the method in a massive parallel processing GPU is proposed as future work to speed up the process.

References

[1] A. Myronenko, X. Song, Point set registration: coherent point drift., IEEE transactions on pattern analysis and machine intelligence 32 (12) (2010) 2262–75. doi:10.1109/TPAMI.2010.46.

[2] G. K. L. Tam, Z.-Q. Cheng, Y.-K. Lai, F. C. Langbein, Y. Liu, D. Marshall, R. R. Martin, X.-F. Sun, P. L. Rosin, Registration of 3D point clouds and meshes: a survey from rigid to nonrigid., IEEE transactions on visualization and computer graphics 19 (7) (2013) 1199–217. doi:10.1109/TVCG.2012.310.

[3] H. Chui, A. Rangarajan, A New Algorithm for Non-Rigid Point Matching, CVPR 2 (2000) 44–51.

[4] H. Chui, A. Rangarajan, A new point matching algorithm for non-rigid registration, Computer Vision and Image Understanding 89 (2-3) (2003) 114–141. doi:10.1016/S1077-3142(03)00009-2.

[5] A. Rangarajan, H. Chui, F. Bookstein, The softassign procrustes matching algorithm, in: J. Duncan, G. Gindi (Eds.), Information Processing in Medical Imaging, Vol. 1230 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 1997, pp. 29–42.

[6] N. Ueda, R. Nakano, Deterministic annealing {EM} algorithm, Neural Networks 11 (2) (1998) 271 – 282. doi:http://dx.doi.org/10.1016/S0893-6080(97)00133-0.

[7] J. Yang, The thin plate spline robust point matching (TPS-RPM) algorithm: A revisit, Pattern Recognition Letters 32 (7) (2011) 910–918. doi:10.1016/j.patrec.2011.01.015.

[8] H. Li, R. W. Sumner, M. Pauly, Global correspondence optimization for non-rigid registration of depth scans, Eurographics Symposium on Geometry Processing 27 (2008) 1421–1430. doi:10.1111/j.1467-8659.2008.01282.x.

[9] Q. Sang, J.-Z. Zhang, Z. Yu, Robust non-rigid point registration based on feature-dependant finite mixture model, Pattern Recognition Letters 34 (13) (2013) 1557–1565. doi:10.1016/j.patrec.2013.06.019.

[10] S. Belongie, J. Malik, J. Puzicha, Shape Context: A new descriptor for shape matching and object recognition, In NIPS (2000) 831—-837.

[11] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (4) (2002) 509–522. doi:10.1109/34.993558.

[12] Q. Sang, J. Zhang, Z. Yu, Non-rigid point set registration: A bidirectional approach, in: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2012, pp. 693–696. doi:10.1109/ICASSP.2012.6287978.

[13] Y. Yawen, Z. P. Peng, Q. Yu, Y. Jie, W. S. Zheng, A Robust CPD Approach Based on Shape Context, in: 33rd Chinese Control Conference, Nanjing, China, 2014, pp. 4930–4935. doi:10.1109/ChiCC.2014.6895776.

[14] H. Wang, B. Fei, A Robust B-Splines-Based Point Match Method for Non-Rigid Surface Registration, in: 2008 2nd International Conference on Bioinformatics and Biomedical Engineering, IEEE, 2008, pp. 2353–2356. doi:10.1109/ICBBE.2008.921.
URL http://ieeexplore.ieee.org/document/4535801/

[15] Y. Yang, S. H. Ong, K. W. C. Foong, A robust global and local mixture distance based non-rigid point set registration, Pattern Recognitiondoi:10.1016/j.patcog.2014.06.017.

[16] J. Chen, J. Ma, C. Yang, L. Ma, S. Zheng, Non-rigid point set registration via coherent spatial mapping, Signal Processing 106 (2015) 62–72. doi:10.1016/j.sigpro.2014.07.004.

[17] W. Lian, L. Zhang, Robust Point Matching Revisited : A Concave Optimization Approach, in: Computer Vision ECCV 2012, 2012, pp. 259–272. doi:10.1007/978-3-642-33709-3_19.

[18] S. Lin, Y.-K. Lai, R. R. Martin, S. Jin, Z.-Q. Cheng, Color-aware surface registration, Computers & Graphics 58 (2016) 31–42. doi:10.1016/j.cag.2016.05.007.

[19] A. Myronenko, X. Song, M. A. C.-P. nán, Non-rigid point set registration: Coherent point drift, in: B. Schölkopf, J. Platt, T. Hoffman (Eds.), Advances in Neural Information Processing Systems 19, MIT Press, 2007, pp. 1009–1016.

[20] P. Wang, P. Wang, Z. Qu, Y. Gao, Z. Shen, A refined coherent point drift (CPD) algorithm for point set registration, Science China Information Sciences 54 (12) (2011) 2639–2646. doi:10.1007/s11432-011-4465-7.

[21] B. Jian, B. C. Vemuri, A Robust Algorithm for Point Set Registration Using Mixture of Gaussians., Proceedings / IEEE International Conference on Computer Vision. IEEE International Conference on Computer Vision 2 (2005) 1246–1251. doi:10.1109/ICCV.2005.17.

[22] B. Jian, B. C. Vemuri, Robust Point Set Registration Using Gaussian Mixture Models., IEEE transactions on pattern analysis and machine intelligence 33 (8) (2010) 1633–1645. doi:10.1109/TPAMI.2010.223.

[23] A. W. Fitzgibbon, Robust registration of 2D and 3D point sets, Image and Vision Computing 21 (13-14) (2003) 1145–1153. doi:10.1016/j.imavis.2003.09.004.

[24] D. Gerogiannis, C. Nikou, A. Likas, Registering sets of points using Bayesian regression, Neurocomputing 89 (2012) 122–133. doi:10.1016/j.neucom.2012.02.018.

[25] Y. Gao, J. Ma, J. Zhao, J. Tian, D. Zhang, A robust and outlier-adaptive method for non-rigid point registration, Pattern Analysis and Applications 17 (2) (2013) 379–388. doi:10.1007/s10044-013-0324-z.

[26] S. Ge, G. Fan, M. Ding, Non-rigid Point Set Registration with Global-Local Topology Preservation, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops (Ml) (2014) 245–251.

[27] S. de Sousa, W. G. Kropatsch, Graph-based point drift: Graph centrality on the registration of point-sets, Pattern Recognition (2014) 1–12doi:10.1016/j.patcog.2014.06.011.

[28] Z. Zhou, J. Zheng, Y. Dai, Z. Zhou, S. Chen, Robust non-rigid point set registration using student's-t mixture model., PloS one 9 (3) (2014) e91381. doi:10.1371/journal.pone.0091381.

[29] M. Gschwandtner, R. Kwitt, A. Uhl, W. Pree, Blensor blender sensor simulation toolbox, in: G. Bebis, R. Boyle, B. Parvin, D. Koracin, S. Wang, K. Kyungnam, B. Benes, K. Moreland, C. Borst, S. DiVerdi, C. Yi-Jen, J. Ming (Eds.), Advances in Visual Computing, Vol. 6939 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2011, pp. 199–208.

[30] M. Saval-Calvo, S. Orts-Escolano, J. Azorin-Lopez, J. Garcia-Rodriguez, A. Fuster-Guillo, V. Morell-Gimenez, M. Cazorla, A Comparative Study of Downsampling Techniques for Non-rigid Point Set Registration using Color, in: International Work-conference on the Interplay between Natural and Artificial Computation, 2015.

[31] M. Saval-Calvo, J. Azorin-Lopez, A. Fuster-Guillo, J. Garcia-Rodriguez, S. Orts-Escolano, A. Garcia-Garcia, Evaluation of sampling method effects in 3D non-rigid registration, Neural Computing and Applicationsdoi:10.1007/s00521-016-2258-z.
URL http://link.springer.com/10.1007/s00521-016-2258-z

[32] S. Orts-Escolano, V. Morell, J. Garcia-Rodriguez, M. Cazorla, Point cloud data filtering and downsampling using growing neural gas, in: The 2013 International Joint Conference on Neural Networks, IJCNN 2013, Dallas, TX, USA, August 4-9, 2013, 2013, pp. 1–8.