The Cardinalized Optimal Linear Assignment (COLA) Metric for Multi-Object Error Evaluation

Pablo Barrios[§], Ghayur Naqvi[¶], Martin Adams[‡], Keith Leung[*], and Felipe Inostroza[†]

Advanced Mining Technology Center, Universidad de Chile, Santiago, Chile
pbarrios@ing.uchile.cl[§], gnaqvi@gmail.com[¶], finostro@ug.uchile.cl[†], keith.leung@amtc.uchile.cl[*], martin@ing.uchile.cl[‡]

***Abstract*—Fundamental to any state estimation problem is the concept of estimation error. In both autonomous robotics and tracking research, the ability to assess the performance of robotic mapping and target tracking algorithms is of crucial importance. This article focusses on metrics for the automatic evaluation of target tracking and feature map estimation algorithms, in the presence of both detection and spatial uncertainty. In such realistic cases, many metrics fail to provide a meaningful and intuitive assessment of robotic map estimates. Recently the Optimal Sub-pattern Assignment (OSPA) metric provided a solution, as it was shown to provide more meaningful assessments of target tracking algorithm performance than its predecessors. This article will demonstrate that the OSPA metric still suffers various disadvantages under realistic mapping scenarios. These include its saturation to a limiting value, irrespective of the cardinality error of different estimators, and its inability to distinguish between repetitions of balanced estimates, in which single ground truth features are estimated with multiple false alarms. The Cardinalized Optimal Linear Assignment (COLA) metric is therefore introduced as a complement to the OSPA metric, and their combination is analysed in order to gauge target tracking and map estimation errors in an intuitive and meaningful manner.**

I. Introduction

Solutions to robotic mapping and Simultaneous Localization and Mapping (SLAM), in which usually the location of an unknown number of features should be estimated, are numerous, offering various degrees of performance [1]–[4]. Irrespective of the estimation methods used, while clear concepts exist for quantifying the error in the estimated pose or trajectory of a robotic vehicle [5] and/or a subset of the estimated feature locations [6], the absolute difference between all estimated and all ground-truth features in the map is rarely jointly considered, even though in SLAM, this is of equal importance to the vehicle trajectory estimate.

The primary difficulty in mathematically defining track or map estimation error is caused by the differences between the estimated and true *number* of tracks/features, and the need to satisfy the four metric axioms [7][1]. To demonstrate the problem of precisely quantifying feature-based mapping error, Figure 4 shows posterior map estimates from two separate feature mapping/SLAM filters. The true feature map, in each figure is shown as blue stars and the estimated map from algorithm 1 (left figure) as red spatial uncertainty ellipsis and that from algorithm 2 as green ellipsis (right figure). The natural question which arises is: "Which estimate is closer to the true map?" Visual intuition is difficult in such a case as this, due to the combination of missed detections, false alarms and spatial errors in both estimated feature maps. An accepted metric to answer this fundamental question is lacking in the mobile robotics community.

It will be shown in this article, that multi-object metrics, which consider both multi-target state estimation cardinality as well as spatial errors and which obey the metric axioms, can gauge robotic maps in an intuitive manner. A new method for comparing maps, called the Cardinalized Optimal Linear Assignment (COLA) metric, is introduced, and compared to its Optimal Sub-pattern Assignment (OSPA) counterpart [7]. In particular, various mathematical properties of each metric, and their consequences on their "meaningful physical interpretations" will be demonstrated.

In the robotics literature, various articles have proposed techniques to assess the quality of robotic maps in various ways [8], [9]. Recent work in SLAM and robotic map estimation has suggested that a collection of map features can be modelled as a finite set, rather than a vector [3], [4]. Indeed the mathematical definitions of the four metric axioms apply to a *set* of vectors and it will be demonstrated in this article that if a ground truth set-based map $\mathcal{M}_k$ and its set-based estimate $\widehat{\mathcal{M}}_k$, which vary with discrete time k, are modelled as *finite sets* of feature location vectors, then a mathematically consistent notion of estimation error is possible, since the 'distance', or error between sets, is a well understood concept.

The work by Schuhmacher *et al* [7] examined such metrics including the Hausdorff distance, Optimal Mass Transfer (OMAT) [10] and introduced the OSPA distance [7], which determine the "distance" between sets.

The philosophy behind the OSPA metric is continued in this article, in terms of the newly proposed COLA set-based metric.

Sections II and III overview current set-based metrics and define the COLA metric. Sections IV and V analyse the

[1]The four metric axioms can be defined as follows. Let $\mathcal{X}$ be an arbitrary, non-empty set, containing the vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$. Then the function d is a metric iff: 1) $d(\mathbf{x},\mathbf{y}) >= 0$, for all $\mathbf{x},\mathbf{y} \in \mathcal{X}$; 2) $d(\mathbf{x},\mathbf{y}) = 0$ iff $\mathbf{x} = \mathbf{y}$, $\mathbf{x} \in \mathcal{X}$ (identity axiom); 3) $d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x})$, for all $\mathbf{x},\mathbf{y} \in \mathcal{X}$ (symmetry axiom); 4) $d(\mathbf{x},\mathbf{y}) \leq d(\mathbf{x},\mathbf{z}) + d(\mathbf{y},\mathbf{z})$, for all $\mathbf{x},\mathbf{y},\mathbf{z} \in \mathcal{X}$ (triangle inequality axiom).

performance of various metrics under different simulated and real world mapping scenarios.

II. Map Set Definitions

Modelling a robotic feature map as a set $\mathcal{M}_k$ provides a general model, since the vectors within each set can contain any (spatial, color and/or other) information of relevance to the type of feature to be estimated. Further, the order in which these features are estimated, in terms of gauging overall map quality, is irrelevant. Throughout this article, the ground truth map set $\mathcal{M}_k$ is considered to contain $\mathfrak{m}_k$ vectors $\mathbf{m}_k^i$, $1 \leq i \leq \mathfrak{m}_k$. For ease of notation and explanation, and without loss of generality, $\mathbf{m}$ will be referred to as a *spatial* variable and the time index k is now dropped. Therefore the ground-truth map $\mathcal{M} = \{\mathbf{m}^1, \mathbf{m}^2, \ldots, \mathbf{m}^{\mathfrak{m}}\}$ where for spatial maps, $\mathcal{M} \in \mathbb{R}^2$ or $\mathbb{R}^3$ as appropriate. Similarly, the estimated map $\widehat{\mathcal{M}}$ is considered to contain $\widehat{\mathfrak{m}}$ vectors modelling the spatial location of map objects - i.e. $\widehat{\mathcal{M}} = \{\widehat{\mathbf{m}}^1, \widehat{\mathbf{m}}^2, \ldots, \widehat{\mathbf{m}}^{\widehat{\mathfrak{m}}}\}$ where $\widehat{\mathcal{M}} \in \mathbb{R}^2$ or $\mathbb{R}^3$ as appropriate, with spatial variable $\widehat{\mathbf{m}}$. Note that $\widehat{\mathfrak{m}}$ is itself *an estimate* of $\mathfrak{m}$ and therefore, in general, $\widehat{\mathfrak{m}} \neq \mathfrak{m}$ ($|\widehat{\mathcal{M}}| \neq |\mathcal{M}|$).

III. The OSPA and COLA Metrics

A. The Optimal Sub-Pattern Assignment (OSPA) metric

In 2008, Schuhmacher et al [7] presented a metric for gauging multi-object state estimates [11]. This is also a derivative of the Wasserstein construction. [7] demonstrated that OSPA obeys the metric axioms and that it improves most of the problems of the OMAT metric in that it is well defined when one of the sets is empty and overcomes the inconsistencies in penalising cardinality error and its geometric dependence.

1) Definition of the OSPA Metric: The OSPA metric $d^p_{\text{OSPA}}(\mathcal{M}, \widehat{\mathcal{M}})$ with power p and cut-off parameter c, for $\widehat{\mathfrak{m}} > \mathfrak{m}$, is defined as:

$$d^p_{\text{OSPA}}(\mathcal{M}, \widehat{\mathcal{M}}) = \left(\frac{1}{\widehat{\mathfrak{m}}} \min_{\sigma} \sum_{i=1}^{\widehat{\mathfrak{m}}} d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})^p \right)^{1/p} \quad (1)$$

$$= \left(\frac{1}{\widehat{\mathfrak{m}}} \left(\min_{\sigma} \sum_{i=1}^{\mathfrak{m}} d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})^p + c^p(\widehat{\mathfrak{m}} - \mathfrak{m}) \right) \right)^{1/p} \quad (2)$$

where σ is a permutation of the set $\{1, \ldots, \mathfrak{m}\}$ which minimizes $\left(\sum_{i=1}^{\mathfrak{m}} d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})^p \right)$, $1 \leq p < \infty$ and the cut-off parameter $c > 0$. The Hungarian method can be used to determine the optimal assignment σ. If both sets are empty, $\mathfrak{m} = \widehat{\mathfrak{m}} = 0$, $d^p_{\text{OSPA}}(\mathcal{M}, \widehat{\mathcal{M}}) = 0$. For $\mathfrak{m} \geq \widehat{\mathfrak{m}}$, the metric is defined as $d^p_{\text{OSPA}}(\widehat{\mathcal{M}}, \mathcal{M})$. The distance $d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})$ is defined as

$$d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)}) = \min(c, d(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})) \quad (3)$$

where $d(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})$ is any metric distance (e.g. Euclidean, Mahalanobis, Hellinger) between $\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)}$.

2) Intuitive Explanation of the OSPA Metric: For $\widehat{\mathfrak{m}} > \mathfrak{m}$ and $p = 1$, the first term of the RHS of (2), determines individual assignments between all $\mathfrak{m}$ of the feature location vectors within $\mathcal{M}$ and a subset of dimension $\mathfrak{m}$ of the feature vectors within $\widehat{\mathcal{M}}$. Due to (3), each assignment is given a value equal to its (possibly statistical) distance up to a maximum value of c (statistical) distance units. The remaining $\widehat{\mathfrak{m}} - \mathfrak{m}$ features in $\widehat{\mathcal{M}}$ which were not assigned, constitute a dimensionality error (possible false alarms and/or missed detections). Each of these is penalised with the maximal distance error c, hence yielding the RHS residual error $c(\widehat{\mathfrak{m}} - \mathfrak{m})$ in (2). To comply with other metrics ($L2$ norm etc.) the assigned distance values can be raised to the general power p. Hence the OSPA metric yields a measure of the difference between $\widehat{\mathcal{M}}$ and $\mathcal{M}$ in units of distance - e.g. meters. It can be seen from (2) and (3) that $d^p_{\text{OSPA}}(\mathcal{M}, \widehat{\mathcal{M}})$ has minimum value zero and saturates to a maximum value c for all $\mathcal{M}$ and $\widehat{\mathcal{M}}$. The effect of, and concepts for selecting c and p are discussed in detail in [7].

B. Cardinalized Optimal Linear Assignment (COLA) metric

1) Definition of the COLA Metric: As a complement to the OSPA metric, this article introduces the COLA metric, which is defined as

$$d^p_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}}) = \left(\min_{\sigma} \sum_{i=1}^{\widehat{\mathfrak{m}}} \left(\frac{d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})}{c} \right)^p \right)^{1/p} \quad (4)$$

$$= \left(\min_{\sigma} \sum_{i=1}^{\mathfrak{m}} \left(\frac{d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})}{c} \right)^p + (\widehat{\mathfrak{m}} - \mathfrak{m}) \right)^{1/p}, \quad (5)$$

where, $d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})$ is defined in (3), again with cut-off parameter c. The assignment σ is the same as that for the OSPA metric. For $\mathfrak{m} \geq \widehat{\mathfrak{m}}$, the metric is defined as $d^p_{\text{COLA}}(\widehat{\mathcal{M}}, \mathcal{M})$. From (2) and (5) it can be seen that

$$d^p_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}}) \equiv \left(\frac{\mathfrak{m}^{1/p}}{c} \right) d^p_{\text{OSPA}}(\mathcal{M}, \widehat{\mathcal{M}}). \quad (6)$$

2) Intuitive Explanation of the COLA Metric: Whereas the OSPA metric has the units of the localization error (i.e. distance), the COLA metric has the units of dimensionality error (i.e. no units). Intuitively, in the OSPA metric, when the (statistical) distance between feature i and its assigned feature j reaches a maximum value (c), feature i becomes a cardinality error, contributing one more fixed distance error c^p. In the COLA metric, when the distance between an unassigned feature i and another feature j decreases to c it changes from a cardinality error to a fractional cardinality error $\left(d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})/c\right)^p$. Although the difference between the OSPA and COLA metrics may seem trivial, Sections IV and V will demonstrate significant differences in the intuitive behavior of the COLA metric over its OSPA counterpart, when evaluating feature maps.

In contrast to the OSPA metric, it can be seen from (5) and (3) that $d^p_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}})$ has minimum value zero and maximum value $(\widehat{\mathfrak{m}})^{1/p}$ if $\widehat{\mathfrak{m}} > \mathfrak{m}$ or $(\mathfrak{m})^{1/p}$ otherwise. In the case of the

COLA metric, the question of how to select parameters c and p, and their physical interpretation, must also be addressed.

The effect of p: In a similar manner to the OSPA metric [7], as p increases, the weighting applied to smaller localization errors diminishes. Therefore the COLA metric also penalises higher localization errors for higher values of p. Also, based on the COLA metric form given in (4), it can be shown that for the same value of c, the COLA metric is also ordered with respect to p - i.e.

$$d^{p_1}_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}}, c) \leq d^{p_2}_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}}, c) \text{ for } 1 \leq p_1 < p_2 \leq \infty. \tag{7}$$

The effect of c: Analysis of the COLA metric form in (5) shows that as $c \to \infty$ the COLA metric becomes only sensitive to cardinality errors. Therefore, as in the OSPA metric, increasing c increases the penalisation of cardinality errors. However in contrast to the OSPA metric, from (5) it is evident that

$$d^{p}_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}}, c_1) \geq d^{p}_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}}, c_2) \text{ for } 1 \leq c_1 < c_2 \leq \infty. \tag{8}$$

c determines how the COLA metric penalises cardinality errors. It is the (possibly statistical) distance value beyond which it is assumed that an estimated feature no longer corresponds to a ground truth feature, so that it then remains unassigned. In a manner similar to the OSPA metric [7], c can be chosen based on "what distance (how many meters) the designer wants to penalise a false or missing estimate", which in any application should significantly aid its practical choice.

The appendix shows that $d^{p}_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}})$ is a metric.

IV. Analyses of Each Metrics' Interpretation

The natural, physical interpretation of the Hausdorff, OMAT and OSPA metrics has been analysed and compared in [7]. Therefore, in this section, similar analyses of these metrics, which focus on the OSPA and COLA metrics will be carried out for particular and general cases of the maps $\mathcal{M}$ and $\widehat{\mathcal{M}}$, in order to highlight the differences between them. In particular, four comparison scenarios will be addressed in which at least one of the maps is empty, contains a single feature, contains multiple features which are balanced[2] and imbalanced with respect to the other map, and contains outliers. The results will demonstrate the COLA metric's ability to provide meaningful error estimates as cardinality errors increase and when repetitions of balanced estimates occur.

To simplify the ensuing analyses, in the examples shown in this section, $d(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})$ will be the Euclidean distance metric. To demonstrate the generality of the COLA metric, actual SLAM performance evaluations in Section V will apply the COLA and OSPA metrics, in which $d(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})$ is the Mahalanobis distance, allowing the incorporation of estimated feature spatial uncertainties.

[2]The terms balanced and imbalanced, with respect to maps, will be explained in Sections IV-C1 and IV-C2.

A. Non Empty Set vs Empty Set

Consider a ground truth map $\mathcal{M} = \emptyset$ and its estimate $\widehat{\mathcal{M}} = \{\widehat{\mathbf{m}}^1, \ldots, \widehat{\mathbf{m}}^{\mathfrak{m}}\}$ or vice versa. The Euclidean, Hausdorff and OMAT metrics are all undefined in this case, since both sets must be non-empty. Meanwhile, the OSPA metric is given by

$$d^{p}_{\text{OSPA}}(\mathcal{M}, \widehat{\mathcal{M}}) = \left(\frac{1}{\widehat{\mathfrak{m}}} c^p (\widehat{\mathfrak{m}} - 0)\right)^{1/p} = c. \tag{9}$$

In this case, the COLA metric yields

$$d^{p}_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}}) = (\widehat{\mathfrak{m}} - 0)^{1/p} = \widehat{\mathfrak{m}}^{1/p}. \tag{10}$$

Both metrics demonstrate a desirable asset since a metric should be defined when one of the sets is empty. However, the COLA metric can be considered to provide a more intuitive result. This is because, irrespective of the difference in cardinality, the OSPA metric gives the same score (c) and is insensitive to this number, whereas the COLA metric increases with $\widehat{\mathfrak{m}}$. For $p = 1$, the COLA metric increases linearly with $\mathfrak{m}$ which is clearly the true value of the cardinality error in this case.

B. A Single Feature Ground Truth Map with Multiple Estimates

Consider the ground truth map $\mathcal{M} = \{\mathbf{m}\}$ and its estimates $\widehat{\mathcal{M}} = \{\widehat{\mathbf{m}}^1, \ldots, \widehat{\mathbf{m}}^{\widehat{\mathfrak{m}}}\}$, which could correspond to an estimator yielding multiple false alarms.

The Hausdorff metric in this case yields

$$d_H(\mathcal{M}, \widehat{\mathcal{M}}) = \min_{\widehat{\mathbf{m}}^j \in \widehat{\mathcal{M}}} d(\mathbf{m}, \widehat{\mathbf{m}}^j), \tag{11}$$

i.e. the Hausdorff metric chooses some $\widehat{\mathbf{m}}^j \in \widehat{\mathcal{M}}$ which minimizes the Euclidean distance $d(\mathbf{m}, \widehat{\mathbf{m}}^j)$.

If all distances $d(\mathbf{m}^i, \widehat{\mathbf{m}}^j) = d_0 > 0$, for $1 \leq j \leq \widehat{\mathfrak{m}}$, where d_0 is a constant, the Hausdorff and OMAT metrics both equal d_0. This implies that although both the Hausdorff and OMAT metrics are sensitive to the spatial error between the ground truth and the estimated landmarks, which is d_0 for each $\widehat{\mathbf{m}}^j$, they are completely insensitive to the number of features in $\widehat{\mathcal{M}}$ and hence to the cardinality error between $\mathcal{M}$ and $\widehat{\mathcal{M}}$ which is $\widehat{\mathfrak{m}} - 1$. This is evident in Figures 1(a) to (d) in which the number of estimated feature maps include an increasing number of false alarms.

From (2), for the same scenario, assuming that the estimate $\widehat{\mathbf{m}}^l$ is the closest point to $\mathbf{m}$ - i.e.

$$d^{(c)}(\mathbf{m}, \widehat{\mathbf{m}}^l) < d^{(c)}(\mathbf{m}, \widehat{\mathbf{m}}^i) \quad \forall \quad i \neq l, \tag{12}$$

the OSPA metric gives the following result

$$d^{p}_{\text{OSPA}}(\mathcal{M}, \widehat{\mathcal{M}}) = \left(\frac{1}{\widehat{\mathfrak{m}}} \left(d^{(c)}(\mathbf{m}, \widehat{\mathbf{m}}^l)^p + c^p(\widehat{\mathfrak{m}} - 1)\right)\right)^{1/p}. \tag{13}$$

The COLA metric then gives the following result

$$d^{p}_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}}) = \left(\left(\frac{d^{(c)}(\mathbf{m}, \widehat{\mathbf{m}}^l)}{c}\right)^p + (\widehat{\mathfrak{m}} - 1)\right)^{1/p}. \tag{14}$$

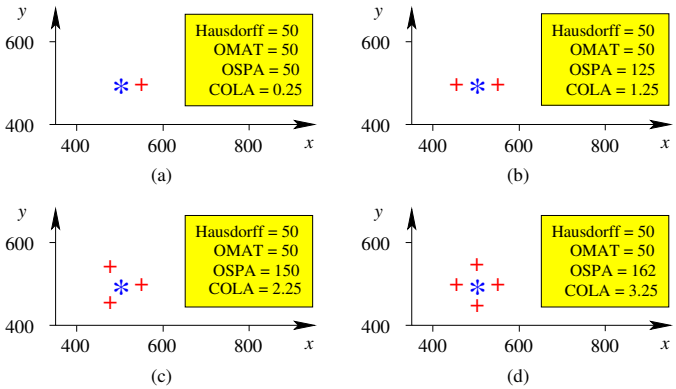


Fig. 1: *A comparison of the Hausdorff, OMAT, OSPA and COLA metrics. The blue stars represent the ground truth map and the red crosses its estimate. Scenarios (b) to (d) show an increasing number of false alarms. The distance between the ground truth landmark and every estimate is* $d = 50$*. The parameters used were* $c = 200$*,* $p = 1$*.*

In contrast to both the Hausdorff and OMAT metrics, both the OSPA and COLA metrics depend on $\widehat{\mathfrak{m}}$ and they penalise such errors in an intuitive manner since both increase with increasing $\widehat{\mathfrak{m}}$. This is verified in Figure 1.

C. Multiple Ground Truth and Estimated Features

[7] demonstrated the non-intuitive behavior of the Hausdorff and OMAT metrics when multiple ground truth and estimated feature scenarios exist in the form of balanced and imbalanced scenarios, which are defined as follows. The discussions in this section will therefore focus on the OSPA and COLA metrics only.

1) Balanced maps: These maps have the following properties[3]:

1) Each feature in the ground truth map has in its proximity the same number of estimated (possibly > 1) features.
2) The distance between each ground truth feature and its nearest feature(s) in the estimated set is the same, referred to as d.
3) The distances between the elements of the ground truth map are larger than d, and referred to as d'.

Consider a ground truth map $\mathcal{M} = \{\mathbf{m}^1, \ldots, \mathbf{m}^{\mathfrak{m}}\}$ and its balanced estimated map $\widehat{\mathcal{M}} = \{\widehat{\mathbf{m}}^{1,1}, \ldots, \widehat{\mathbf{m}}^{1,q}, \ldots, \widehat{\mathbf{m}}^{\mathfrak{m},q}\}$ (i.e. each feature in $\mathcal{M}$ appears to be associated with q estimates) as follows:

$$\begin{aligned} d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i),l}) &= d \leq c,\ 1 \leq i \leq \mathfrak{m},\ 1 \leq l \leq q \\ d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{j}) &> d \quad \forall \quad i \neq j \end{aligned} \tag{15}$$

This means that every ground truth landmark has q estimates at distances d. Then, the OSPA distance is given by

$$d_{\text{OSPA}}^{\text{balanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p = \left(\frac{1}{q}(d^p + c^p(q-1))\right)^{1/p} \tag{16}$$

This means that the OSPA metric performs the same as the case when the ground truth has one element, which is evident by substituting $\widehat{\mathfrak{m}} = q$ into (13). In fact if the scenarios in Figures 2(a) and 1(c) and Figures 2(c) and 1(d) are compared, the OSPA metric evaluation is the same.

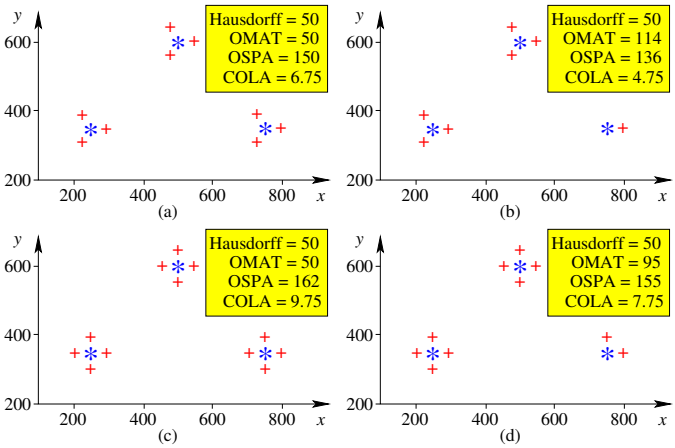


Fig. 2: *Four scenarios highlighting the performance of OSPA when cardinality errors exists. The distance between the center of each ground truth landmark and its neighboring estimates is* $d = 50$*. The parameters used in these calculations* $c = 200$*,* $p = 1$*.* $|\mathcal{M}|$ *indicates the cardinality of the ground truth set* $\mathcal{M}$ *and* $|\widehat{\mathcal{M}}|$ *the cardinality of its estimated set* $\widehat{\mathcal{M}}$*.*

In this balanced scenario, the COLA metric yields

$$d_{\text{COLA}}^{\text{balanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p = \left(\mathfrak{m}\left(\left(\frac{d}{c}\right)^p + (q-1)\right)\right)^{1/p} \tag{17}$$

Note that this result *is not* the same as the single ground truth feature case, which from 14, with $\widehat{\mathfrak{m}} = q$ yields

$$d_{\text{COLA}}^{\text{balanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p = \left(\left(\frac{d}{c}\right)^p + (q-1)\right)^{1/p} \tag{18}$$

Note that the multi feature balanced map result (17) is $\mathfrak{m}^{1/p}$ times larger than the single feature case.

The scenarios in Figures 2(a) and 2(c) contain repetitions of the scenarios in Figures 1(c) and 1(d) respectively, and the OSPA metric demonstrates its ability to determine the average error as it gauges each equally. In contrast, the extra cardinality (false alarm) errors in these scenarios are exactly reflected by the COLA metric for $p = 1$. This demonstrates the COLA metric's strict penalisation of cardinality errors.

2) Imbalanced maps: Imbalanced maps have the same properties as balanced maps except for property 1 in Section IV-C1, which becomes:

1) At least one feature in the ground truth map has in its proximity a different number of estimated features, than the others.

Now consider the OSPA and COLA metrics' performances in a more general imbalanced case with ground truth map $\mathcal{M} = \{\mathbf{m}^1, \ldots, \mathbf{m}^{\mathfrak{m}}\}$ and estimated map

$$\begin{aligned} \widehat{\mathcal{M}} = \{&\widehat{\mathbf{m}}^{1,1}, \ldots, \widehat{\mathbf{m}}^{1,q}, \ldots, \widehat{\mathbf{m}}^{2,1}, \ldots, \widehat{\mathbf{m}}^{2,q}, \\ &\ldots, \widehat{\mathbf{m}}^{\mathfrak{m}-1,1}, \ldots \widehat{\mathbf{m}}^{\mathfrak{m}-1,q}, \widehat{\mathbf{m}}^{\mathfrak{m},1}, \ldots, \widehat{\mathbf{m}}^{\mathfrak{m},q-s}\} \end{aligned} \tag{19}$$

[3]Note that due to the symmetry property of the metrics, the ensuing arguments also apply if the ground-truth and estimated map sets are swapped.

(i.e. the ground truth landmark $\mathbf{m}^{\mathfrak{m}}$ has in its neighborhood $q-s$ estimates. The subset $\mathcal{M}' = \{\mathbf{m}^1, \ldots, \mathbf{m}^{\mathfrak{m}-1}\} \in \mathcal{M}$ is balanced and has in its neighborhood q estimates), where:

$$\begin{aligned} &d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i),l}) = d \leq c, \quad 1 \leq i \leq \mathfrak{m}-1, \quad 1 \leq l \leq q \\ &d^{(c)}(\mathbf{m}^{\mathfrak{m}}, \widehat{\mathbf{m}}^{\sigma(\mathfrak{m}),l}) = c, \quad 1 \leq l \leq q-s \\ &d^{(c)}(\mathbf{m}^i, \mathbf{m}^j) > d, \quad \forall \quad i \neq j. \end{aligned} \tag{20}$$

In this case

$$d_{\text{OSPA}}^{\text{imbalanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p = \left(\frac{1}{q\mathfrak{m}-s} (\mathfrak{m} d^p + c^p(\mathfrak{m}(q-1)-s)) \right)^{1/p} \tag{21}$$

and comparing (16) and (21) gives

$$\begin{aligned} d_{\text{OSPA}}^{\text{balanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p &\geq d_{\text{OSPA}}^{\text{imbalanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p, \quad s \geq 0, \\ d_{\text{OSPA}}^{\text{imbalanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p &\geq d_{\text{OSPA}}^{\text{balanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p, \quad s < 0 \end{aligned} \tag{22}$$

meaning that the OSPA metric penalises false alarms in an intuitive manner. This can be seen in Figure 2 since the estimate in Figure (2b) is better than that of Figure (2a) because it has less cardinality error but the same spatial error.

The general imbalanced scenario described in (19) yields a COLA metric value

$$d_{\text{COLA}}^{\text{imbalanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p = \left(\mathfrak{m} \left(\frac{d}{c} \right)^p + (\mathfrak{m}(q-1)-s) \right)^{1/p}. \tag{23}$$

If this is compared with $d_{\text{COLA}}^{\text{balanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p$ in (17), it can also be seen that $d_{\text{COLA}}^{\text{balanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p > d_{\text{COLA}}^{\text{imbalanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p$ for $s > 0$ and vice versa for $s < 0$, which again complies with intuition. It should also be noted that, true to its nature of having units of cardinality error, for $p = 1$, $d_{\text{COLA}}^{\text{balanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p - d_{\text{COLA}}^{\text{imbalanced}}(\mathcal{M}, \widehat{\mathcal{M}})^p = s$, yielding the exact dimensionality error between the sets.

D. Outliers

Consider an estimated map with only one outlier - i.e. $\mathcal{M} = \{\mathbf{m}^1, \ldots, \mathbf{m}^{\mathfrak{m}}\}$ and $\widehat{\mathcal{M}} = \{\widehat{\mathbf{m}}^1, \ldots, \widehat{\mathbf{m}}^{\mathfrak{m}+1}\}$, and assume that every single ground truth landmark has a perfect estimate:

$$d(\mathbf{m}^i, \widehat{\mathbf{m}}^i) = 0 \text{ for } 1 \leq i \leq \mathfrak{m} \tag{24}$$

$$d(\mathbf{m}^i, \widehat{\mathbf{m}}^j) > c \text{ for } i \neq j \tag{25}$$

The Hausdorff distance in this case is:

$$d_H(\mathcal{M}, \widehat{\mathcal{M}}) = \min_{\mathbf{m}^i \in \mathcal{M}} d(\mathbf{m}^i, \widehat{\mathbf{m}}^{\mathfrak{m}+1}) \tag{26}$$

where it can be seen that if a large distance exists between $\widehat{\mathbf{m}}^{\mathfrak{m}+1}$ and its its closest feature $\mathbf{m} \in \mathcal{M}$, the Hausdorff metric yields a large value, even if $\mathfrak{m}$, the number of perfectly matched estimates and ground truth features, is large.

The OMAT metric in this case is:

$$d_{\text{OMAT}}^p(\mathcal{M}, \widehat{\mathcal{M}}) = \left(\sum_{i=1}^{\mathfrak{m}} \frac{d(\mathbf{m}^i, \widehat{\mathbf{m}}^{\mathfrak{m}+1})^p}{\mathfrak{m}(\mathfrak{m}+1)} \right)^{1/p}. \tag{27}$$

From a physical stand point, the OMAT metric's performance is more desirable since if $\mathfrak{m} \to \infty$, $d_{\text{OMAT}}^p(\mathcal{M}, \widehat{\mathcal{M}}) \to 0$. It can also be seen from Equation 27 however, that when $\mathfrak{m} = 0$, OMAT is undefined.

In this same scenario, from (2), the OSPA metric is equivalent to a distance error as follows

$$d_{\text{OSPA}}^{(c)}(\mathcal{M}, \widehat{\mathcal{M}})^p = c \left(\frac{1}{\mathfrak{m}+1} \right)^{1/p}. \tag{28}$$

As in the case of the OMAT metric, if $\mathfrak{m} \to \infty$ (large numbers of perfect estimates), the OSPA metric $d_{\text{OSPA}}^{(c)}(\mathcal{M}, \widehat{\mathcal{M}})^p \to 0$ distance units, again giving an intuitive result.

In the case of the COLA metric,

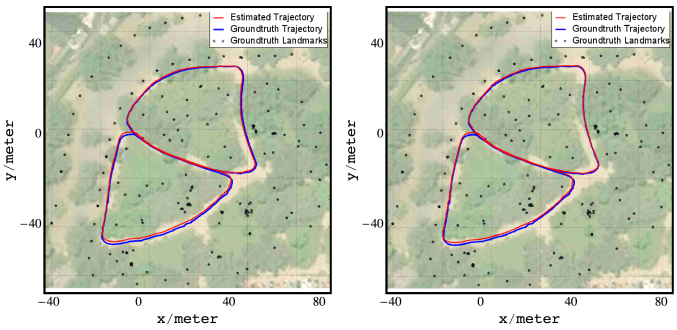
$$d_{\text{COLA}}^{(c)}(\mathcal{M}, \widehat{\mathcal{M}})^p = 1, \tag{29}$$

which, since this metric yields cardinality, as opposed to average distance units, is intuitive. In this case there is a single outlier, and the COLA metric correctly reports it. It should be noted here that the COLA metric is somewhat unforgiving to cardinality errors. For example, if a mapping algorithm estimates a large number of perfectly located estimates, with just one false alarm, the COLA metric always penalises the algorithm, even though as $\mathfrak{m} \to \infty$ the algorithm can be argued to be approaching perfection. In this sense the OSPA metric behaves more intuitively.

V. Results - Evaluating Metric Performance with Real SLAM Maps

In this section the performance of the Hausdorff, OMAT, OSPA and COLA metrics will be analysed by comparing their ability to score real estimated SLAM results in a physically meaningful manner.

Figure 3 shows the SLAM trajectory estimates from two different SLAM algorithms[4], which were designed to estimate vehicle trajectories and maps corresponding to the x, y location of tree trunks. Each algorithm is referred to as "SLAM



(a) SLAM Alg. 1. (b) SLAM Alg. 2.

Fig. 3: *Ground truth and estimated trajectories from SLAM algorithms 1 (a) and 2 (b). Each result also shows the ground truth feature locations (stars) superimposed onto a satellite image of the park in which the experiments were carried out.*

Alg. 1" and "SLAM Alg. 2" respectively and each result is superimposed onto a satellite image of the area[5] to show the tree coverage and trunk locations. The ground truth trajectories

[4]The estimated SLAM solutions are based on Multi-Hypothesis (MH)-FastSLAM [12] and Rao-Blackwellized (RB)-PHD-SLAM [3].

[5]The experiments were carried out in Parque O'Higgins, Santiago, Chile.

(blue lines) were obtained via manual scan-matching, and the red lines represent the estimated trajectories. The blue stars represent the ground truth features, centered at their locations, again obtained through independent, manual scan matching procedures. Although the differences in the trajectories appear minor, it will be shown that their corresponding map estimates are significantly different.

To highlight the estimated maps, and their spatial uncertainties, the red and green ellipses in the left and right plots respectively of Figure 4, correspond to the means and covariances for each estimated feature. Each covariance ellipse is calculated from the eigenvalues and eigenvectors of the error covariance matrices corresponding to each estimated feature. The ellipses shown correspond to "5-Sigma ellipses", which from 2 degree-of-freedom Chi-squared tables, correspond to a probability mass within each ellipse of 0.999996. Hence, $d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})$ in (2) can be the Mahalanobis metric such that

$$d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)}) = \min\left(c, \sqrt{(\mathbf{m}^i - \widehat{\mathbf{m}}^{\sigma(i)})^T (\mathbf{P}^i)^{-1} (\mathbf{m}^i - \widehat{\mathbf{m}}^{\sigma(i)})}\right) \quad (30)$$

where $c = 5$ and $\mathbf{P}^i$ is the sum of the error covariance sub-matrices corresponding to estimated feature $\widehat{\mathbf{m}}^{\sigma(i)}$ and ground truth feature $\mathbf{m}^i$. In this analysis it is assumed that the error associated with all ground truth features is zero[6] and that the estimated feature covariance values are available from the SLAM estimator. In each figure, the Hausdorff, OMAT, OSPA, COLA and the estimated cardinality errors are provided. In all of the experiments, $p = 2$, which according to [7] yields smooth distance curves, and is commonly used in other metrics, such as the L2 distance.

A. Analyses of the Complete SLAM Experiments

1) Performance of the Hausdorff Metric: The max-min function of the Hausdorff metric is represented in each figure by the dashed line and circular regions. The dashed line connects estimated and ground truth feature locations which satisfy the max-min function and hence give the Hausdorff distance d_H. It is evident that if the number of estimated features (possibly false alarms) increased outside of the circle in Figure 4 (left), they have no affect on the value of the Hausdorff metric. Conversely, if there were more ground truth features (stars) outside the circle in Figure 4 (right), possibly indicating more missed detections, these would also be ignored by the metric. The Hausdorff metric is essentially insensitive to cardinality errors outside of these circles.

2) Subdividing the Maps for Assessing Metric Performance: Assessing the intuitive performance of the map metrics is clearly a difficult task based on the complete map estimates of SLAM Alg. 1 and SLAM Alg. 2 in Figure 4, due to complex combinations of both spatial and cardinality feature errors. Therefore, to simplify the analysis, Figure 5 shows both ground truth and estimated maps divided into cells (sub-maps), based on a Voronoi partitioning of the ground truth map. This provides a means of isolating single ground truth features and analysing the corresponding estimates in that cell. Considering ground truth and estimated maps in single and multiple Voronoi cells, rather than the complete maps, simplifies the assessment of the metrics in terms of their intuitive behavior.

[6]In the target tracking literature, [13] applied the OSPA metric based on a Hellinger distance metric, in which the ground truth target covariances were replaced with their Cramer Rao lower bound values.

B. Analyses of the Sub-maps

Based on the sub-map corresponding to the orange shaded cell in Figure 5, Figure 6 shows the ground truth and estimated maps for cut-off parameter $c = 1$. Each metric's value is given

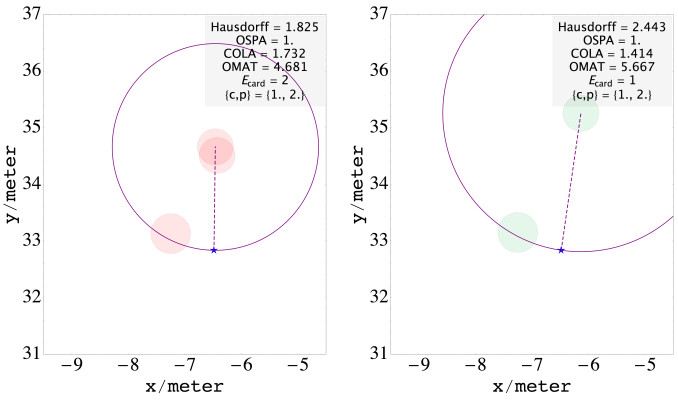


Fig. 6: *Metric performance for the single orange shaded cell of Figure 5 for OSPA and COLA metric values* $c = 1$ *and* $p = 2$.

in the top right corner, along with the cardinality error E_{card} between the estimated and ground truth maps and the values of c and p used in both the COLA and OSPA metrics. In Figure 6 ($c = 1$), from (30) no features are gated (assigned) and arguably, intuitively SLAM Alg. 2 has produced a better map estimate than SLAM Alg. 1, since the spatial error between the ground truth features and their closest estimates are approximately the same, but SLAM Alg. 2 yields one less cardinality error. Note that the only metric which correctly reports this is the COLA metric. The OSPA metric judges each sub-map to be of the same quality, as it saturates to its limit ($c = 1$) for both SLAM Algs. 1 and 2.

Due to its max-min function, the Hausdorff metric contradicts intuition as it assigns a lower value to SLAM Alg. 1, again clarified by the dashed lines and circles in the figure.

The OMAT metric also provides a non-intuitive result, penalising SLAM Alg. 2 (5.667) more than SLAM Alg, 1 (4.681), due to the nature of its fractional assignments, as explained in Section IV-B. Note that the OMAT metric made assignments only within the considered cell.

Note that if c is increased to 3, as shown in Figure 7, one of the feature estimates in each map is now assigned to a ground truth feature by both the OSPA and COLA metrics. Therefore, the OSPA metric now gauges SLAM Alg. 2 to have a superior mapping performance than SLAM Alg. 1, in

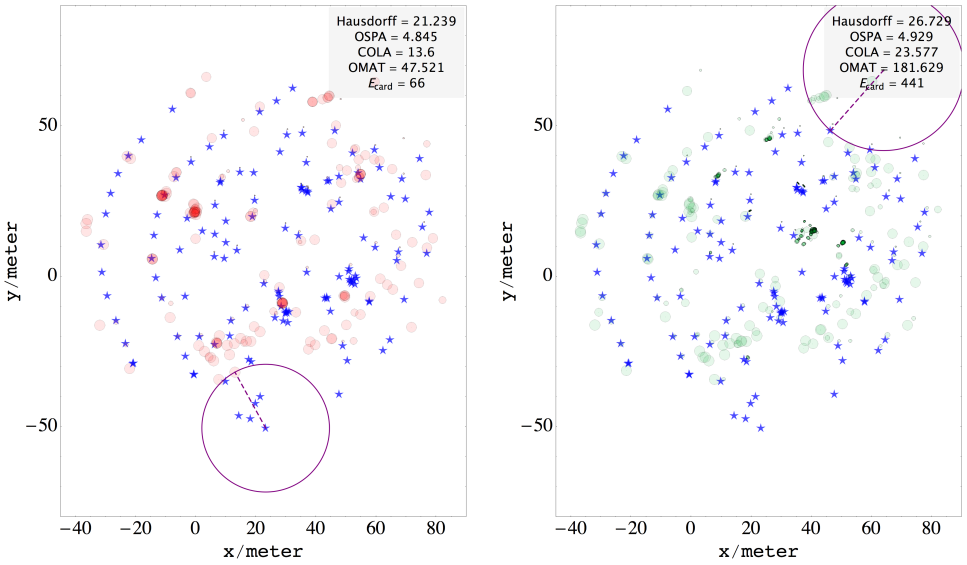


Fig. 4: *The ground truth (blue stars) and estimated maps produced by SLAM Alg. 1 (Left: red "5-sigma" confidence interval ellipsis) and SLAM Alg. 2 (Right: green "5-sigma" confidence interval ellipsis).*

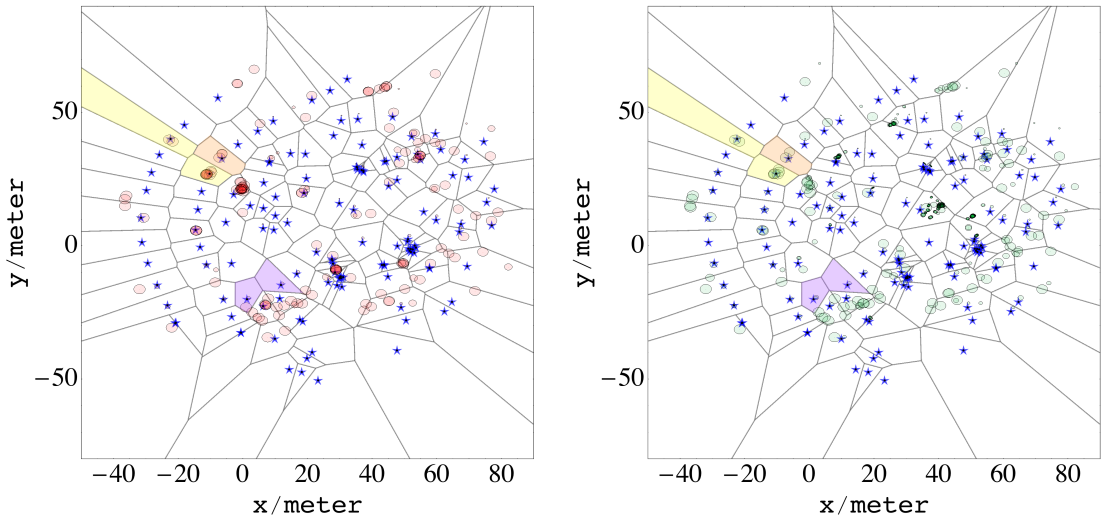


Fig. 5: *A Voronoi partitioning of the maps with respect to the ground truth features, allows individual cells and their combinations to be considered for comparing intuitive map estimation performance with that of the metrics.*

line with intuition. The COLA metric agrees with this result. However, in contrast to the OSPA metric, the COLA metric continues to decrease in each case, such that if the estimates from SLAM Alg. 1 in left sides of Figures 6 and 7 are compared, the estimates in Figure 7 are favored. The same applies to the estimates of SLAM Alg. 2 in the right sides of the figures. Importantly, in contrast to the OSPA metric, the COLA metric continues to decrease as c increases (Equation (8)) and an estimated feature changes status from ungated to gated, arguably following intuition.

Figures 8 and 9 demonstrate the metrics' performances in the combination of the yellow and orange Voronoi partitions of Figure 5, for COLA and OSPA metric parameters $c = 1$ and $c = 6$ respectively. Note again in Figure 8 that, contrary to intuition, the Hausdorff metric gauges SLAM Alg. 1 to be superior in its map estimation ability than SLAM Alg. 2, despite the larger cardinality error committed by SLAM Alg. 1, and the fact that no features are gated. As expected, the OSPA, non-intuitively gauges both maps equally, whereas the COLA metric, and in this case the OMAT metric, favor SLAM Alg. 2.

In Figure 9, c is increased to 6, reducing the Mahalanobis

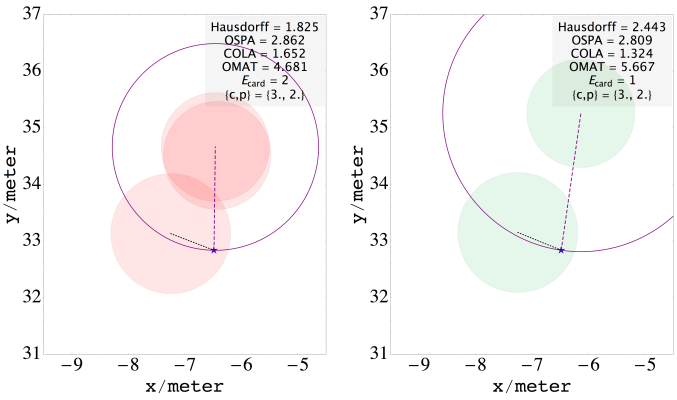


Fig. 7: *Metric performance for the single orange shaded cell of Figure 5 for OSPA and COLA metric values $c = 3$ and $p = 2$.*

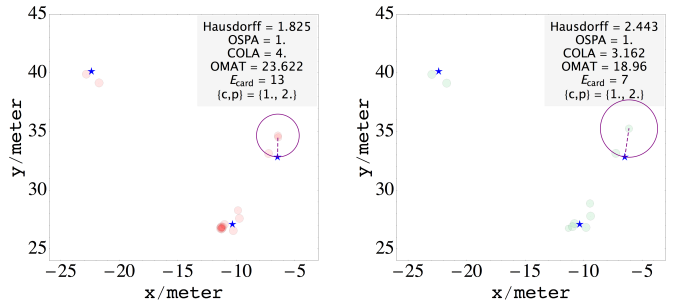


Fig. 8: *Metric performance for the yellow and orange cells of Figure 5, for OSPA and COLA metric values $c = 1$ and $p = 2$.*

distance between 3 of the estimated features and ground truth. This allows them to gate and the OSPA metric therefore reduces its value below $c = 6$. Once again however, the COLA metric demonstrates its continuity over the OSPA metric as it intuitively lowers its values, when compared with Figure 8, favoring SLAM Alg. 2 in the right side of Figure 9, due to its 3 gated features and lower cardinality error.

VI. Summary

This article introduced the COLA metric for the automatic evaluation of robotic feature map estimators. In contrast to its OSPA predecessor, it was shown that the metric can provide more physically intuitive evaluations of map errors in situations which cause the OSPA metric to saturate to its limiting value c, and in repetitions of multiple balanced

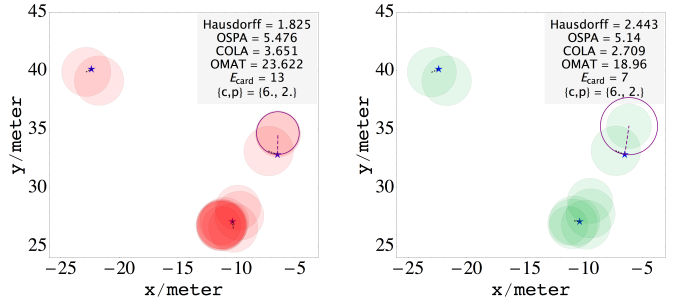


Fig. 9: *Metric performance for the yellow and orange cells of Figure 5, for OSPA and COLA metric values $c = 6$ and $p = 2$.*

map estimates. Further, in contrast to the OSPA metric, the continuity of the COLA metric is preserved as c changes, for a given map and its estimate, when estimated features change status from ungated to gated. This provides meaningful comparisons between estimated maps when the allowable statistical distances between ground truth features and their estimates are changed.

Appendix: Proof that $d^p_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}})$ is a Metric

The proof that $d^p_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}})$ is a metric follows a similar procedure to the the proof that $d^p_{\text{OSPA}}(\mathcal{M}, \widehat{\mathcal{M}})$ is a metric in [7]. Clearly $d^p_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}}) \geq 0$ for all $\mathcal{M}, \widehat{\mathcal{M}}$ because metric $d^{(c)}(\mathbf{m}^{\mathfrak{i}}, \widehat{\mathbf{m}}^{\mathfrak{j}}) \geq 0$ for all $\mathbf{m}^{\mathfrak{i}}$ and $\widehat{\mathbf{m}}^{\mathfrak{j}}$. Similarly $d^p_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}}) = 0$ iff $\widehat{\mathcal{M}} = \mathcal{M}$ - proof: From (5), if $d^{(c,p)}_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}}) = 0$,

$$\min_{\sigma} \sum_{i=1}^{\widehat{\mathfrak{m}}} \frac{d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)})^p}{c^p} = -(\mathfrak{m} - \widehat{\mathfrak{m}}) \leq 0 \tag{31}$$

Since (5) is defined for $\mathfrak{m} \geq \widehat{\mathfrak{m}}$, the RHS of (31) implies that $\mathfrak{m} = \widehat{\mathfrak{m}}$. The LHS must then also be zero, implying that $d^{(c)}(\mathbf{m}^i, \widehat{\mathbf{m}}^{\sigma(i)}) = 0 \quad \forall i$. $d^p_{\text{COLA}}(\mathcal{M}, \widehat{\mathcal{M}}) = d^p_{\text{COLA}}(\widehat{\mathcal{M}}, \mathcal{M})$ because $d^{(c)}(\mathbf{m}^{\mathfrak{i}}, \widehat{\mathbf{m}}^{\mathfrak{j}})$ satisfies the symmetry property. It remains to be verified that the triangle inequality is satisfied.

Consider the set $\widehat{\mathcal{N}} = \{\widehat{\mathbf{n}}^1, \ldots, \widehat{\mathbf{n}}^{\widehat{\mathfrak{n}}}\}$, with cardinality $\widehat{\mathfrak{n}} \in \mathbb{N}_0$. Consider the following sets of dummy points $\mathcal{U} = \{\mathbf{u}^{\mathfrak{i}}\}_{\mathfrak{i} \in \mathbb{N}_0}$ and $\mathcal{V} = \{\mathbf{v}^{\mathfrak{j}}\}_{\mathfrak{j} \in \mathbb{N}_0}$ in $\mathbb{R}^N$ where

$$d(\mathbf{u}^{\mathfrak{i}}, \mathbf{x}) \geq c, \; d(\mathbf{v}^{\mathfrak{i}}, \mathbf{x}) \geq c, \; d(\mathbf{u}^{\mathfrak{i}}, \mathbf{v}^{\mathfrak{j}}) \geq c$$

For all $\mathbf{x}$, and for all $\mathfrak{i}, \mathfrak{j}$

Case 1: $(\mathfrak{m} \leq \widehat{\mathfrak{m}} \leq \widehat{\mathfrak{n}})$: In order to raise the cardinality of sets $\mathcal{M}$ and $\widehat{\mathcal{M}}$ to $\widehat{\mathfrak{n}}$, consider the following dummy points:

$$\mathbf{m}^{\mathfrak{m}+\mathfrak{i}} = \mathbf{u}^i, \quad 1 \leq \mathfrak{i} \leq \widehat{\mathfrak{n}} - \mathfrak{m} \tag{32}$$

$$\widehat{\mathbf{m}}^{\widehat{\mathfrak{m}}+j} = \mathbf{v}^j, \quad 1 \leq \mathfrak{j} \leq \widehat{\mathfrak{n}} - \widehat{\mathfrak{m}} \tag{33}$$

Then choose $\sigma, \tau \in \Pi_{\widehat{\mathfrak{n}}}$ such that,

$$\min_{\pi \in \Pi_{\widehat{\mathfrak{n}}}} \sum_{i=1}^{\widehat{\mathfrak{n}}} \left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}}, \widehat{\mathbf{n}}^{\pi(\mathfrak{i})})}{c} \right)^p = \sum_{i=1}^{\widehat{\mathfrak{n}}} \left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}}, \widehat{\mathbf{n}}^{\sigma(\mathfrak{i})})}{c} \right)^p \tag{34}$$

$$\min_{\pi \in \Pi_{\widehat{\mathfrak{n}}}} \sum_{i=1}^{\widehat{\mathfrak{n}}} \left(\frac{d^{(c)}(\widehat{\mathbf{n}}^{\mathfrak{i}}, \widehat{\mathbf{m}}^{\pi(\mathfrak{i})})}{c} \right)^p = \sum_{i=1}^{\widehat{\mathfrak{n}}} \left(\frac{d^{(c)}(\widehat{\mathbf{n}}^{\mathfrak{i}}, \widehat{\mathbf{m}}^{\tau(\mathfrak{i})})}{c} \right)^p \tag{35}$$

Therefore

$$d_{\text{COLA}}^p(\mathcal{M},\widehat{\mathcal{M}}) = \left(\min_{\pi\in\Pi_{\mathfrak{m}}}\sum_{i=1}^{\widehat{\mathfrak{m}}}\left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}},\widehat{\mathbf{m}}^{\pi(\mathfrak{i})})}{c}\right)^p\right)^{1/p} \quad (36)$$

$$\leq \left(\min_{\pi\in\Pi_{\mathfrak{m}}}\sum_{i=1}^{\mathfrak{m}}\left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}},\widehat{\mathbf{m}}^{\pi(\mathfrak{i})})}{c}\right)^p + (\widehat{\mathfrak{n}}-\mathfrak{m})\right)^{1/p} \quad (37)$$

$$\leq \left(\min_{\pi\in\Pi_{\widehat{\mathfrak{n}}}}\sum_{i=1}^{\widehat{\mathfrak{n}}}\left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}},\widehat{\mathbf{m}}^{\pi(\mathfrak{i})})}{c}\right)^p\right)^{1/p} \quad (38)$$

$$\leq \left(\sum_{i=1}^{\widehat{\mathfrak{n}}}\left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}},\widehat{\mathbf{n}}^{\sigma(\mathfrak{i})}) + d^{(c)}(\widehat{\mathbf{n}}^{\sigma(\mathfrak{i})},\widehat{\mathbf{m}}^{\tau(\sigma(\mathfrak{i}))})}{c}\right)^p\right)^{1/p} \quad (39)$$

$$\leq \left(\sum_{i=1}^{\widehat{\mathfrak{n}}}\left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}},\widehat{\mathbf{n}}^{\sigma(\mathfrak{i})})}{c}\right)^p\right)^{1/p} + \left(\sum_{i=1}^{\widehat{\mathfrak{n}}}\left(\frac{d^{(c)}(\widehat{\mathbf{n}}^{\sigma(\mathfrak{i})},\widehat{\mathbf{m}}^{\tau(\sigma(\mathfrak{i}))})}{c}\right)^p\right)^{1/p} \quad (40)$$

$$\leq d_{\text{COLA}}^p(\mathcal{M},\widehat{\mathcal{N}}) + d_{\text{COLA}}^p(\widehat{\mathcal{N}},\widehat{\mathcal{M}}). \quad (41)$$

In (37) $\widehat{\mathfrak{n}}-\mathfrak{m}$ dummy points were added to the set $\mathcal{M}$ yielding (38). In (38) the triangular inequality on the metric $d^{(c)}$ and the application of (34) and (35) resulted in (39). Finally, Minkowski's inequality yielded (40).

Case 2: $(\mathfrak{m},\widehat{\mathfrak{n}}\leq\widehat{\mathfrak{m}})$: In order to raise the cardinality of sets $\mathcal{M}$ and $\widehat{\mathcal{N}}$ to $\widehat{\mathfrak{m}}$, consider the following dummy points:

$$\mathbf{m}^{\widehat{\mathfrak{m}}-\mathfrak{i}+1} = \mathbf{u}^i, \quad 1\leq\mathfrak{i}\leq\widehat{\mathfrak{m}}-\mathfrak{m} \quad (42)$$

$$\widehat{\mathbf{n}}^{\widehat{\mathfrak{m}}-\mathfrak{j}+1} = \mathbf{u}^j, \quad 1\leq\mathfrak{j}\leq\widehat{\mathfrak{m}}-\widehat{\mathfrak{n}} \quad (43)$$

where $d(\mathbf{m}^{\mathfrak{i}},\widehat{\mathbf{n}}^{\mathfrak{i}})=0, \quad \max(\mathfrak{m},\widehat{\mathfrak{n}})\leq i\leq\widehat{\mathfrak{m}}$.

Then choose $\sigma,\tau\in\Pi_{\widehat{\mathfrak{m}}}$ such that,

$$\min_{\pi\in\Pi_{\mathfrak{m}\vee\widehat{\mathfrak{n}}}}\sum_{i=1}^{\mathfrak{m}\vee\widehat{\mathfrak{n}}}\left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}},\widehat{\mathbf{n}}^{\pi(\mathfrak{i})})}{c}\right)^p =$$

$$\min_{\pi\in\Pi_{\widehat{\mathfrak{m}}}}\sum_{i=1}^{\widehat{\mathfrak{m}}}\left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}},\widehat{\mathbf{n}}^{\pi(\mathfrak{i})})}{c}\right)^p = \sum_{i=1}^{\widehat{\mathfrak{m}}}\left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}},\widehat{\mathbf{n}}^{\sigma(\mathfrak{i})})}{c}\right)^p \quad (44)$$

$$\min_{\pi\in\Pi_{\widehat{\mathfrak{m}}}}\sum_{i=1}^{\widehat{\mathfrak{m}}}\left(\frac{d^{(c)}(\widehat{\mathbf{n}}^{\mathfrak{i}},\widehat{\mathbf{m}}^{\pi(\mathfrak{i})})}{c}\right)^p = \sum_{i=1}^{\widehat{\mathfrak{m}}}\left(\frac{d^{(c)}(\widehat{\mathbf{n}}^{\mathfrak{i}},\widehat{\mathbf{m}}^{\tau(\mathfrak{i})})}{c}\right)^p \quad (45)$$

where $\mathfrak{m}\vee\widehat{\mathfrak{n}}=\max(\mathfrak{m},\widehat{\mathfrak{n}})$. Therefore, finally

$$d_{\text{COLA}}^p(\mathcal{M},\widehat{\mathcal{M}}) = \left(\min_{\pi\in\Pi_{\widehat{\mathfrak{m}}}}\sum_{i=1}^{\widehat{\mathfrak{m}}}\left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}},\widehat{\mathbf{m}}^{\pi(\mathfrak{i})})}{c}\right)^p\right)^{1/p} \quad (46)$$

$$\leq \left(\sum_{i=1}^{\widehat{\mathfrak{m}}}\left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}},\widehat{\mathbf{n}}^{\sigma(\mathfrak{i})}) + d^{(c)}(\widehat{\mathbf{n}}^{\sigma(\mathfrak{i})},\widehat{\mathbf{m}}^{\tau(\sigma(\mathfrak{i}))})}{c}\right)^p\right)^{1/p} \quad (47)$$

$$\leq \left(\sum_{i=1}^{\widehat{\mathfrak{m}}}\left(\frac{d^{(c)}(\mathbf{m}^{\mathfrak{i}},\widehat{\mathbf{n}}^{\sigma(\mathfrak{i})})}{c}\right)^p\right)^{1/p} + \left(\sum_{i=1}^{\widehat{\mathfrak{m}}}\left(\frac{d^{(c)}(\widehat{\mathbf{n}}^{\sigma(\mathfrak{i})},\widehat{\mathbf{m}}^{\tau(\sigma(\mathfrak{i}))})}{c}\right)^p\right)^{1/p} \quad (48)$$

$$\leq d_{\text{COLA}}^p(\mathcal{M},\widehat{\mathcal{N}}) + d_{\text{COLA}}^p(\widehat{\mathcal{N}},\widehat{\mathcal{M}}) \quad (49)$$

In (46) the triangular inequality on the metric $d^{(c)}$ and the application of (44) and (45) resulted in (47). Again, Minkowski's inequality yielded (48).

ACKNOWLEDGEMENTS

Effort sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA9550-15-1-0032. The US Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright notation thereon. The authors acknowledge "Becas Conicyt - Doctorado Nacional, 2012 and 2014", CSIRO - CORFO project 10CEII-9007-F1-L3-P2, The AMTC and Conicyt Anillo project No. ACT 1120, Chile. Invaluable advice from Ba-Ngu Vo is also acknowledged.

REFERENCES

[1] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit. Fastslam: A factored solution to the simultaneous localization and mapping problem. In *Proc. AAAI Nat. Conf. on Art. Intel.*, pages 593–598, 2004.

[2] M. Kaess, A. Ranganathan, and F. Dellaert. iSAM: Incremental smoothing and mapping. *IEEE Transactions on Robotics*, 24(6):1365–1378, December 2008.

[3] J. Mullane, B.N. Vo, M.D. Adams, and B.T. Vo. A random-finite-set approach to Bayesian SLAM. *IEEE Transactions on Robotics*, 27(2):268–282, April 2011.

[4] C.S. Lee, D.E. Clark, and J. Salvi. SLAM with dynamic target via single cluster PHD filtering. *IEEE Selected Topics on Signal Processing*, 7(3):543–552, June 2013.

[5] W. Burgard, C. Stachniss, G. Grisetti, B. Steder, R. Kümmerle, C. Dornhege, M. Ruhnke, A. Kleiner, and J.D. Tardos. A comparison of SLAM algorithms based on a graph of relations. In *IEEE/RSJ Int'l. Conf. on Intel. Robots and Systems (IROS)*, St. Louis, USA, October 2009.

[6] M.W.M.G. Dissanayake, P. Newman, S. Clark, H.F. Durrant-Whyte, and M. Csorba. A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Trans. Robotics and Automation*, 17(3):229–241, June 2001.

[7] D. Schuhmacher, B.T. Vo, and B.N. Vo. A consistent metric for performance evaluation of multi-object filters. *IEEE Transactions on Signal Processing*, 86(8):3447–3457, 2008.

[8] M. Chandran-Ramesh and P. Newman. Assessing map quality using conditional random fields. *Field and Service Robotics - Springer Tracts in Advanced Robotics*, 2008.

[9] A.I. Wagan, A. Godil, and X. Li. Map quality assessment. In *PerMIS '08, Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems*, pages 278 – 282, New York, USA, 2008.

[10] J.R. Hoffman and R. Mahler. Multitarget miss distance via optimal assignment. *IEEE Transactions on Systems, Man and Cybernetics*, 34(3):327–336, May 2004.

[11] B. Ristic, B.-N. Vo, D. Clark, and B.-T. Vo. A metric for performance evaluation of multi-target tracking algorithms. *IEEE Transactions on Signal Processing*, 59(7):3452–3457, July 2011.

[12] J. Nieto, J. Guivant, E. Nebot, and S. Thrun. Real time data association for fastSLAM. In *IEEE Internationlal Conference on Robotics and Automation*, volume 1, pages 412–418, September 2003.

[13] D. Clark, B. Ristic, Ba-Ngu Vo, and Ba-Tuong Vo. Bayesian multi-object filtering with amplitude feature likelihood for unknown object snr. *Signal Processing, IEEE Transactions on*, 58(1):26–37, Jan 2010.