dwolter@sfbtr8.uni-bremen.de

Qualitative Reasoning for Spatial Matching

Diedrich Wolter

Universität Bremen,
Transregional Collaborative Research Center SFB/TR 8 Spatial Cognition

Abstract. Matching two representations against one another is central to many applications, e.g., matching observation against models in object or scene recognition or matching a local perception against a global map in self-localization if a mobile robot. In both examples, *spatial* information structures are associated, i.e., the representation includes spatial arrangement information of individual parts, either in a relative or an absolute frame of reference.
In this paper the utility of qualitative configuration information for matching is investigated. Qualitative configuration information expresses relative arrangement information. It is argued that such information is helpful for making side conditions of plausible matchings explicit, which can improve matching efficiency as well as quality of the matching.

Keywords. Matching, Qualitative Spatial Information

1 Introduction

Many applications involve the task of matching two representations against one another. For example scene recognition involves identifying a set of spatially distributed objects. Many objects bear rich information which can be exploited the recognition task, but it is information about the spatial arrangement of the involved objects that allows for important distinctions. Consider you are observing a scene that contains dishes, glasses, and sets of cutlery. These objects, when nicely arranged on a table, gives rise to the hypothesis that you are about to enjoy a nice meal, whereas piled up dishes, glasses, and clutter of cutlery suggests dish washing to be your destiny.

This article investigates into matching tasks which are involved with matching spatially embedded representations, i.e. representations whose parts are arranged in a spatial domain (e.g., picture space, tabletop, geographic space). The term *spatial matching* is coined to summarize different facets of matching tasks, aiming at the design of a common theoretic framework. One particular goal of this approach is to device subclasses that can be dealt with efficiently. The role of qualitative information in expressing confident side conditions will be highlighted. Qualitative information can even be exploited for identifying efficient subclasses automatically.

There are several application areas for spatial matching, most notably visual object recognition or robot localization. In the case study presented in this article spatial matching in robot localization is highlighted, but the overall approach presented here is independent of the application area.

Dagstuhl Seminar Proceedings 08091
Logic and Probability for Scene Interpretation
http://drops.dagstuhl.de/opus/volltexte/2008/1610

2 Spatial Matching

In this section a general theoretical approach to spatial matching is developed. The overall goal is to compute an *optimal* matching between two representations whereby optimality is evaluated by application-dependent measures. To enable focusing on the spatial information the underlying representation is characterized by two distinct layers, namely the *features description layer* and the *feature configuration layer*. It is the feature configuration layer that exhibits similarities through different fields of applications and provides a starting point for realizing a universally applicable efficient matching framework.

2.1 Feature Description Layer

The feature description layer is introduced to abstract from the individual features used in different applications. Features may be extracted by a computer vision system (e.g., blobs, SIFT features) but they may also come from completely different information sources sources such spatial databases. Common to all features is that one can assume availability of a feature distance measure $F \times F \rightarrow \mathbb{R}_0^+$ where F stands for the set of all possible features. Feature distance may be a probability measure for feature identification, but it may also be a similarity measure (e.g., shape similarity measure). In the degenerate case of dealing with features not comprising any feature-intrinsic information, feature distance may be just a constant value. Examples for features not comprising any information are reflection points measured by radar-like sensors: it is only known that there is some obstacle, but no properties other than its location are available. Yet, location is not a feature-intrinsic feature but is part of the feature configuration layer.

2.2 Feature Configuration Layer

Information about the spatial arrangement of features is gathered in the feature configuration layer. Possible ways of utilizing configuration information include local coordinate frames, relative information, etc. There are two ways in which the value of configuration knowledge manifests itself:

- obtaining more plausible matchings
- computing matchings more efficiently

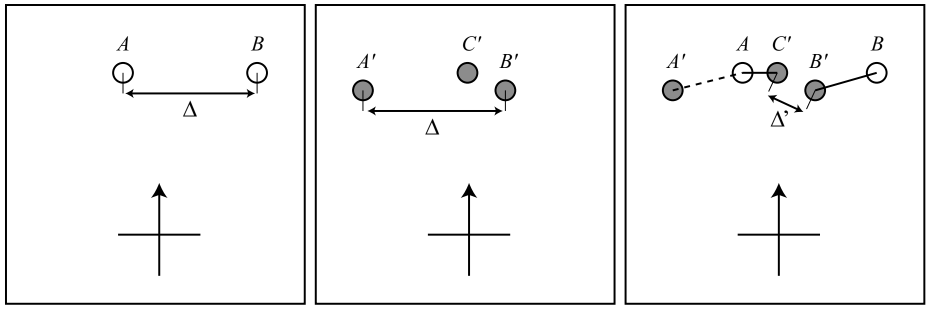
Regarding an improved plausibility of matchings, notice that configuration information links together individual features and thereby establishes a larger context than a single feature presents. This makes configuration information valuable for matching—in the case of features lacking distinctive properties, configuration knowledge is the only clue to feature identification. There two types of configuration knowledge: absolute and relative information.

Using absolute configuration knowledge usually means to localize individual features within a local coordinate system. When assigning features, congruency of local coordinates is examined and features may only get matched if their local coordinates are agreeable. This requires that the alignment of the two local coordinate systems involved is known or can be estimated. Such approach is particularly popular in robot localization contexts, where landmarks observed by the robot are registered in a local coordinate system which is defined by the current position and orientation of the robot. Tracking landmarks from observation to observation allows the robot to keep track of its position[1]. Measuring the movement of the robot provides an estimate on how local coordinate systems between movements need to be aligned. The drawback of this approach is that a reliable estimate of alignment is required. Moreover, matching is performed in an iterative framework, which repeatedly alternates a matching and an alignment phase to update the alignment estimate. This way the discrete, combinatorial matching problem is transformed into an optimization task that aims at determining the best fitting alignment of two feature sets. Such approaches (also termed *non-matching* [4]) are closely related to iterative optimization, including the problems of local minima.

Relative configuration knowledge, on the other hand, does not require prior alignment information and is, thus, useful when no estimate exists how feature sets need to be aligned. This is for example the case in rotation invariant recognition. The utility of relative configuration knowledge has already been investigated in the context of object recognition (see[4]), but not taken advantage of in robot mapping. Yet, the use of relative configuration knowledge can help to obtain more plausible matching results. Consider an example from the domain of robot localization depicted in Fig. 1. A single observer (the robot) takes two sensor readings and needs to recognize objects in the second reading. In other words, it needs to track the objects it observes. In the first observation (Fig. 1 (a)) the robot detects two features A, B which—for the sake of simplicity—are assumed to be indistinguishable. In a second observation, the observer detects three features A', B', C'. The task of determining the most plausible correspondence should yield the correspondences $A \sim A'$ and $B \sim B'$. If individually considering differences of position the result may be counter-intuitive as illustrated in Fig. 1 (c): since the spurious detection of C is closer to the expected position of A, the correspondence $A \sim C$ is erroneously determined. Simultaneously considering the mapping $A \sim C$ and $B \sim B'$ indicates that A and B are indeed mapped to features much closer to one another than agreeable with the initial observation. To overcome such problems, Neira and Tardós [5] proposed an algorithm that jointly considers differences in position in a probabilistic model.

Matching is a computational costly operation when aiming to determine the optimal feature correspondence with respect to a given optimality criterion. First, let us consider the problem space of feature-feature correspondences faced in matching tasks. When matching two representations comprising n and

[1] Among others, see [1,2,3].



(a) (b) (c)

Fig. 1. (a) An observer located as marked by the cross detects two features, A and B. (b) In a second observation from the same pose, three features are detected at different positions. This can be the result of measurement noise. (c) Both observations are matched using the compatible local reference system. Individually considering the distances of associated features (nearest neighbor) can result in a counter-intuitive matching (solid lines). Considering the relative position of features (Δ) allows handling this situation (dashes lines).

m features respectively, there are

$$\sum_{i=0}^{n} \binom{n}{i} \cdot \binom{m}{i} \cdot i! \tag{1}$$

potential correspondences to evaluate if not all features should necessarily be matched (e.g., to account for changes). This is already infeasibly complex, so additional knowledge must be exploited to reduce the search space and computation time. Respecting configuration knowledge in the approach by Neira and Tardós [5] still yields a exponential time worst-case complexity. However, respecting relative configuration knowledge can in some cases lead to efficient polynomial-time algorithms though. The idea is to interpret relative configuration knowledge as confident knowledge, pruning the search space of feature correspondences. Before identifying tractable spatial matching problems, we now formulate the general matching framework.

3 Optimal Homomorphic Matching

Spatial matching is challenging in many regards: obtaining a plausible solution, obtaining it efficiently, responding to potential changes that can happen, and handling of uncertain knowledge. Unfortunately, plausibility of matching is hard to define. Generally speaking, one desires that only similar features are associated and that spatial configurations are respected, i.e. any matching needs to respect to the overall spatial structure. Changes, measurement noise, and uncertainty make it necessary to balance all contributing factors. Consequently, it is doubtful whether an indisputable definition of plausible matching exists at all. However, a formal framework to express plausible matchings can very well be introduced. The framework should exhibit these key features:

- Multiple assignments to handle alias effects in feature detection
- Even very similar sets of features sets are likely to be somewhat incongruent and some features may not get associated; matchings may not be of maximum cardinality
- Matchings shall maximize plausibility of features to correspondence modeled by negation of a feature distance measure
- Side conditions to introduce confident configuration knowledge shall be expressible

The theoretical framework of matching in hypergraphs offers an elegant way to include these features.

Definition 1 (Hypergraph) *A graph $\mathcal{G} = (V, E)$ with a finite set of vertices V and set of hyperedges $E \subseteq 2^V$ which is a subset of the powerset of vertices is called hypergraph.*

Matching problems in graph theory are usually studied in the context of bipartite graphs and there is one subtle difference between matching in bipartite graphs versus matching in hypergraphs. In bipartite graph matching the task is to select a subset of graph edges which connect the two subsets of vertices into which each bipartite graph can be decomposed naturally. In hypergraphs, however, where edges span over multiple vertices there is no natural decomposition and the little more elaborate notion of *balanced hypergraphs* [6] are required.

Definition 2 *A hypergraph $\mathcal{G}$ is balanced if each odd cycle in $\mathcal{G}$ has an edge containing at least three vertices of the cycle. Hereby, a cycle is defined as closed path $v_i \overset{e_i}{\to} v_{i+1} \overset{e_{i+1}}{\to} \dots \overset{e_j}{\to} v_{j+1} = v_i$, whereby $v_i \in e_i \wedge v_{i+1} \in e_i$. Analogously to bipartite graphs, $\mathcal{G} = (V', V'', E)$ is used as notation, whereby V', V'' model two disjunct sets of vertices to be associated.*

Balanced hypergraphs allow for a straightforward generalization of matching.

Definition 3 (Balanced hypergraph matching) *Let $\mathcal{G} = (V, E)$ be a balanced hypergraph; then, $M \subseteq E$ is a matching if M consists of pairwise disjoint edges. If M is a vertex cover of $\mathcal{G}$ then M is called* perfect.

The key feature is that matching in balanced hypergraphs allows us to assign multiple features simultaneously, see Fig. 2 for an example.

Notably, this definition already includes the case of not-associated vertices by allowing for hyperedges comprising single vertices. Therefore, the generalized matching task can simply be formulated as the maximal weight vertex cover, i.e., we need to determine a set of edges which covers all vertices while maximizing the weight associated with the edges. The idea is to use negated feature distance

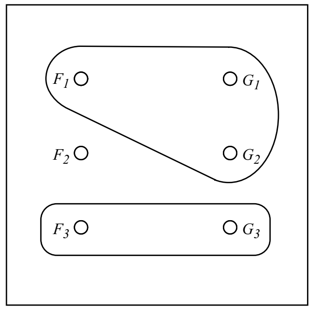


Fig. 2. Hypergraph modeling a matching that comprises a two-to-one correspondence; edges in hypergraphs are depicted to enclose adjacent vertices.

(cf. Sec. 2.1) as edge weight, thereby minimizing the overall feature distance. The formulation using vertex covers ensures that all features are regarded.

For incorporating side conditions to the matching, the notion of homomorphic matchings that obey constraints has been introduced [7]. Since constraints are usually defined for tuples of individual objects, but multiple correspondence partners are allowed in the desired correspondence relation $\sim \subseteq 2^F \times 2^G$, it is useful to first define an elementary correspondence relation $\approx \subseteq F \times G$ with respect to the correspondence relation $\sim$:

$$x \approx y :\Leftrightarrow \exists X, Y : X \sim Y \wedge x \in X \wedge y \in Y \quad (2)$$

This relation maps n-to-m-correspondences to 1-to-1 correspondences and eases the definition of a homomorphic matching.

Definition 4 (Homomorphic matching) *Let $\mathcal{G} = (V', V'', E)$ be a balanced hypergraph and $\sim$ a matching. Let further C be an n-ary constraint on V', i.e. a n-ary relation over the vertices. A matching defined by its correspondence relation $\sim$ and induced elementary correspondence relation $\approx$ is homomorphic, if the mapping from V' to V'' (and vice versa) respects the given constraint:*

$V' \to V''$ $\forall v'_1, \ldots, v'_n \in V' : C(v'_1, \ldots, v'_n) \Rightarrow \exists v''_1, \ldots, v''_n \in V'' : v'_1 \approx v''_1 \wedge \ldots \wedge v'_n \approx v''_n \wedge C(v''_1, \ldots, v''_n)$

$V'' \to V'$ $\forall v''_1, \ldots, v''_n \in V'' : C(v''_1, \ldots, v''_n) \Rightarrow \exists v'_1, \ldots, v'_n \in V' : v'_1 \approx v''_1 \wedge \ldots \wedge v'_n \approx v''_n \wedge C(v'_1, \ldots, v'_n)$

Homomorphic matching with respect to a set of constraints is defined analogously.

As shown in [7] computing the optimal homomorphic matching with respect to constraints C which present a total ordering is still polynomial time and can be tackled within the Dynamic Programming (DP) paradigm [8]. The only prerequisite is that the feature similarity measure exhibits a *local optimality criterion*. Roughly speaking, the criterion demands that if in the globally optimal association of feature sets $F = \{F_1, \ldots, F_m\}$ and $G = \{G_1, \ldots, G_n\}$ features

$F_i \subset F$ are associated with $G_i \subset G$, $i = 1, \ldots, k$, then this association must also be optimal in the subtask of associating $F_i \cup F_j$ and $G_i \cup G_j$. It then holds that the maximum weight match can be computed in $O(|F|^2 \cdot |G|^2)$ time.

4 Cyclic orientation knowledge in robot localization

Mobile robots can learn about there whereabouts by matching their local perception against a global map. This is called the *correspondence problem* or the problem of *data association* in robotics. The area of robot localization is particularly interesting as usually the features observable by a robot bear little information only. Nevertheless, we expect the robot correctly to identify its position. This requires to exploit the spatial arrangement information effectively.

TBD Summarize approach presented in [9]

5 Conclusions

TBD

References

1. Gutmann, J.S., Weigel, T., Nebel, B.: A fast, accurate and robust method for self-localization in polygonal environments using laser range finders. Advanced Robotics **14** (2001) 651–667
2. Lu, F., Milios, E.: Robot pose estimation in unknown environments by matching 2D range scans. Journal of Intelligent and Robotic Systems (1997)
3. Besl, P., McKay, N.: A method for registration of 3D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence **14** (1992) 239–256
4. Grimson, W.E.L.: Object Recognition by Computer: The Role of Geometric Constraints. MIT Press, Cambridge (MA), USA (1990)
5. Neira, J., Tardós, J.D.: Data association in stochastic mapping using the joint compatibility test. IEEE Transactions on robotics and automation **17** (2001) 890–897
6. Berge, C.: Sur certains hypergraphes generalisant les graphes bipartites. In Erdős, P., Rény, A., Sós, V., eds.: Combinatorial Theory and its Applications I. Colloq. Math. Soc., North Holland, Amsterdam (1970)
7. Wolter, D.: Spatial Representation and Reasoning for Robot Mapping—a Shape-based Approach. Volume 48 of Springer Tracts in Advanced Robotics. Springer (2008)
8. Bellman, R.: Dynamic Programming. Princeton University Press (1957)
9. Wolter, D., Freksa, C., Latecki, L.J.: Towards a generalization of self-localization. In Jefferies, M.E., Yeap, W.K., eds.: Robot and Cognitive Approaches to Spatial Mapping. Springer (2008)