

Weighted Walkthroughs in Retrieval by Contents of Pictorial Data

E.Vicario and W.X.He

Dipartimento di Sistemi e Informatica, Università degli Studi di Firenze
E-mail: {vicario|wenxue}@aguirre.ing.unifi.it

Abstract. An original framework is presented which supports symbolic representation and visual querying of images based on spatial arrangements of typed objects. Theoretical properties of the framework are expounded and an algorithm is presented which supports approximate matching of example queries against image descriptions.

1 Introduction

In the access to image databases, queries based on the appearing visual features of searched data reduce the gap between the user and the engineering representation of data. To support this access modality, image contents can be modeled in terms of different types of features such as shape, texture, color, and spatial arrangement. In this paper, we focus on visual modeling based on spatial relationships among typed objects.

Modeling of spatial arrangements among pixel regions occupied by visual objects has been addressed in a variety of applications, which can be roughly reduced to the contexts of geographic information systems and image databases. While in the first case, modeling is focused on topological relationships, capturing set-theoretical concepts such as inclusion, overlapping or adjacency [7] [11], modeling for image databases mainly insists on directional relationships describing concepts associated with linguistic constructs such as *above*, *below*, *left of*, independently from the possible existence of intersections or adjacencies [4] [10].

In the *symbolic projection* approach, both directional and topological relationships are reduced to the composition of the ordering relationships among intervals representing object projections on two reference axes [2] [4]. Since projections on each axis are independent, the overall representation subtends the assimilation of objects to their minimum embedding rectangles. As in Allen's logic, on each axis, 13 exclusive ordering conditions are possible, so that the relationship among two objects is always encoded as one out of 13x13 mutually exclusive cases. Classification among the cases is inherently Boolean, so that slightly differing spatial distributions may be cast into different cases independently from their possible perceptual similarity [5].

In [10], directional information is represented through the numeric value of the orientation of the line connecting object centroids. The critical condition for the applicability of the approach is the replacement of an object through a

single representative point, which may lose soundness in the presence of complex shaped objects or when object distances are small with respect to object dimensions. This may prevent distinction between perceptually dissimilar configurations.

In [6], objects are replaced through a set of representative points, and the relationship between an observing and an observed object is represented by a set of *walkthroughs*. Developing on the style of the symbolic projection, and in particular on the *point of view* concept proposed in [3], walkthroughs provide a symbolic representation of the set of different paths leading from any representative point of the observing object to any representative point of the observed object. The lack for numeric values still yields a discontinuous relation the spatial configuration and its symbolic representation.

In this paper, the walkthrough approach is now extended so as to associate walkthrough relationships with a set of weights accounting for the different degree of truth by which each single walkthrough represents the actual relationship between the objects. The resulting representation, that we call *weighted walkthroughs*, is expounded in the rest of this section.

2 Weighted Walkthroughs

Given a pair of objects A and B within a Cartesian reference system, the set of *Cartesian paths* from A to B is the set of paths leading from any point of A to any point of B along two subsequent horizontal and vertical steps. *Walkthroughs* from A to B are equivalence classes on the set of Cartesian paths from A to B , which abstracts from: the two connected points of A and B ; the distance covered by each step; and the order in which the two steps are taken. For any two objects A and B , there exist at most 9 independent walkthroughs, which can be represented by a pair of indices $\langle i, j \rangle$, taking values -1 , 0 or $+1$: index i is equal to -1 , 0 or $+1$, whether the step along the horizontal axis has negative orientation or null length or positive orientation, respectively; in a similar manner, index j encodes the step along the vertical axis. Fig.1 reports an example helping intuitive understanding.

In order to account for its perceptual relevance, each walkthrough $\langle i, j \rangle$ is associated with a *weight* $w_{i,j}(A, B)$, which measures the number of pairs of points belonging to A and B which are connected through a Cartesian path equivalent to the walkthrough $\langle i, j \rangle$ itself. This weight is evaluated as a four-dimensional integral over the pairs of points belonging to objects A and B :

$$w_{i,j}(A, B) = K_{ij}(A, B) \int_A \int_B C_i(x_b - x_a) C_j(y_b - y_a) dx_b dy_b dx_a dy_a \quad (1)$$

where: $i) C_{-1}(\cdot)$ and $C_{+1}(\cdot)$ denote the characteristic functions of negative and positive real numbers \mathcal{R}^- and \mathcal{R}^+ , respectively, while $C_0(\cdot)$ denotes the *Dirac* function $\delta(\cdot)$ which here acts as characteristic function of the singleton

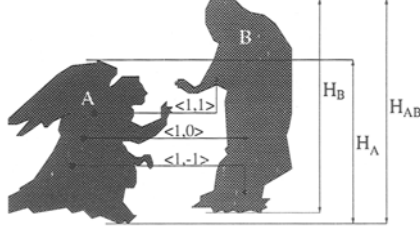


Fig. 1. Cartesian paths from A to B are collected in three independent walkthroughs $\langle 1, 1 \rangle$, $\langle 1, 0 \rangle$, and $\langle 1, -1 \rangle$. The figure reports normalization factors referred to in Eq.2.

set $\{0\}$; *ii*) $K_{ij}(A, B)$ is a dimensional normalization factor which is defined differently whether indexes i and j are equal to zero or take non-null values I and J :

$$\begin{aligned} K_{I,J}(A, B) &= \frac{1}{|A||B|} & K_{0,0}(A, B) &= \frac{1}{\sqrt{|A||B|}} \\ K_{I,0}(A, B) &= \frac{1}{H_{AB}L_A L_B} & K_{0,J}(A, B) &= \frac{1}{L_{AB}H_A H_B} \end{aligned} \quad (2)$$

where: $|A|$ and $|B|$ are 2-dimensional measures of A , and B ; L_A , H_A , L_B and H_B are the width and the height of A and B , respectively; L_{AB} , and H_{AB} are the width and height of the minimum embedding rectangles of the union of A and B .

The use of Dirac function in the expression of the characteristic function of the origin of real numbers ensures a finite non-null measure also for middle and central weights which are obtained by the integration of a quasi-everywhere-null function. Different normalization factors $K_{i,j}$ are used to compensate the effect on dimensionality of the integrals which derives from the use of such Dirac functions: while $K_{I,J}$ is a space length to the fourth power, space dimensionalities of $K_{I,0}$, $K_{0,J}$, and $K_{0,0}$ are reduced to the third, the third, and the second power, respectively. The 9 coefficients in a weighted walkthrough representation are invariant with respect to shifting and zooming of the image, and they are anti-commutative with respect to the objects A and B to which they refer (i.e. $w_{i,j}(A, B) = w_{-i,-j}(B, A)$)

Weighted walkthroughs are compositional, i.e. for any object A and for any two disjoint objects B_1 and B_2 , the weight $w_{i,j}(A, B_1 \cup B_2)$ can be derived by linear combination of weights $w_{i,j}(A, B_1)$ and $w_{i,j}(A, B_2)$:

$$\frac{w_{i,j}(A, B_1 \cup B_2)}{K_{ij}((A, B_1 \cup B_2))} = \frac{w_{i,j}(A, B_1)}{K_{ij}((A, B_1))} + \frac{w_{i,j}(A, B_2)}{K_{ij}((A, B_2))} \quad (3)$$

Weighted Walkthroughs between Rectangles

The property of compositionality of Eq.(3) permits to reduce the relationships between any two objects to a linear combination of the relationships among

the parts of any their decomposition. In particular, if objects are assimilated to multi-rectangles, the evaluation of the four-dimensional integral of Eq.(1) can be reduced to the linear composition of a set of sub-integrals taken over rectangular domains. This permits to derive the relationship between any two multi-rectangular objects A and B by linear composition of close-form terms computed in a strict set of independent variants of the mutual position between two rectangles. These variants are 9 and correspond to the 3 basic cases shown in Fig.2.

Case 1: Projections of the two objects are disjoint on both axes. Four variants are possible. Fig.2 considers the cases that B is in the right-lower quadrant of A (similar results are obtained for the remaining three variants). *Case 2:* Projections of objects A and B are completely superposed on one axis and disjoint on the other axis. Again, four variants of the case are possible, and one of them is reported in Fig.2. *Case 3:* Projections of objects A and B are completely superimposed on both the axes.

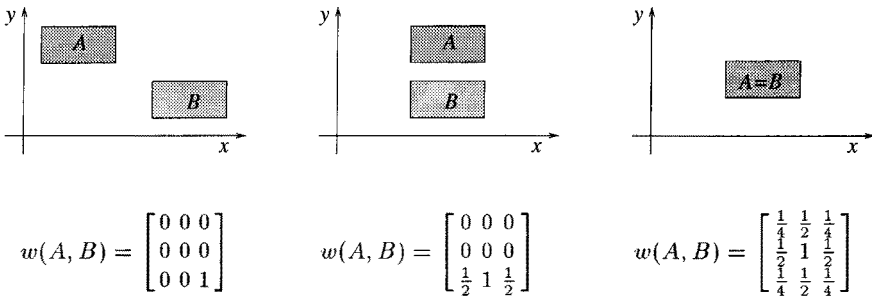


Fig. 2. Basic cases representing the 9 possible basic variants in the positioning of two rectangles.

It is worth stressing the fact that the reduction of multi-rectangular compositions to a strict nucleus of 9 basic variants not only simplifies analytical treatment and effective implementation of spatial modeling based on walkthroughs. Rather, it also ensures consistency in quantitative weights associated with relationships which may be decomposed in different manners.

A Metric of Similarity for Weighted Walkthroughs

Similarity $\mathcal{S}(w, \bar{w})$ between a pair of 9-tuples w and \bar{w} can be evaluated as the the 1 complement of a numeric distance $\mathcal{D}(w, \bar{w})$ in the space of 9-tuples:

$$\mathcal{S}(w, \bar{w}) = 1 - \mathcal{D}(w, \bar{w}) \tag{4}$$

In turn, $\mathcal{D}(w, \bar{w})$ is evaluated as a convex normal combination of four *components* which evaluate the distance between homologous coefficients in corner positions, in middle horizontal and vertical positions, and in the central position of the weighted walkthrough 9-tuples w and \bar{w} :

$$\mathcal{D}(w, \bar{w}) = \frac{\alpha}{2} \sum_{I=\pm 1, J=\pm 1} |w_{I,J} - \bar{w}_{I,J}| + \frac{\beta}{2} \sum_{I=\pm 1} |w_{I,0} - \bar{w}_{I,0}| + \frac{\gamma}{2} \sum_{J=\pm 1} |w_{0,J} - \bar{w}_{0,J}| + \delta |w_{0,0} - \bar{w}_{0,0}| \quad (5)$$

where α , β , γ , and δ are non-negative numbers with sum equal to 1.

Distance \mathcal{D} can be proven to exhibit the five properties that are commonly assumed as axiomatic basis of a dis-similarity measure, i.e. *positivity* ($\mathcal{D}(w, \bar{w}) \geq 0$), *normality* ($\forall w, \bar{w}, \mathcal{D}(w, \bar{w}) \leq 1$), *autosimilarity* ($\mathcal{D}(w, \bar{w}) = 0$ iff $w \equiv \bar{w}$), *simmetry* ($\mathcal{D}(w, \bar{w}) = \mathcal{D}(\bar{w}, w)$), and *triangularity* ($\mathcal{D}(w, \bar{w}) + \mathcal{D}(\bar{w}, \hat{w}) \geq \mathcal{D}(w, \hat{w})$). The properties are proven by inductive extension through Eq.(3) of the properties holding in the 9 basic variants. Each property is proven to separately hold for each of the four components d_{xy} , d_{xo} , d_{oy} , and d_{oo} , and it is then extended to the sum \mathcal{D} by the assumption that α , β , γ , and δ comprise a convex and normal combination.

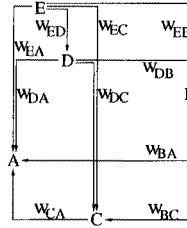
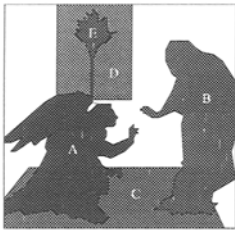
Beside the five properties of *positivity*, *normality*, *autosimilarity*, *simmetry*, and *triangularity*, the plausibility of the metric is also supported by computational experience and by a property of continuity in the relation between object arrangements and their weighted walkthrough representation. Detailed treatment of this property is not in the scope of this paper. Roughly speaking, this continuity means that slight changes in the mutual positioning or in the mass distribution of two objects A and B themselves, result into slight changes in the weighted walkthroughs between A and B .

3 Modeling Spatial Arrangements with Weighted Walkthroughs

Representation of spatial relationships based on weighted walkthroughs can be embedded within a graph theoretical framework for the representation of spatial arrangements among set of typed objects. In this section, we describe one such framework as it was designed and implemented within a prototype system supporting retrieval by contents from a digital library of images reproducing renaissance paintings. The system follows a three-stage operation model distinguishing *archiving*, *querying* and *retrieval*.

Image Description: At archiving time, each image is associated with a description capturing identity and types of imaged objects along with their mutual weighted walkthroughs. One such description can be represented as a triple *description* = $\langle Obj, Type, WW \rangle$, where: *Obj* is a set of objects; *WW* associates each pair of objects with a weighted walkthrough capturing their mutual spatial relationship; *Type* associates each object with a value within an infinite support of types.

A description can be regarded as complete graph in which vertices are objects and edges are spatial relationships. In this perspective, each vertex is labeled with a type, and each edge is labeled with the weighted walkthrough associated with the relationship that it represents. Fig.3 shows the sketch of an image and its description in the form of labeled graph.



$$w(A, B) = \begin{bmatrix} 0 & 0 & .69 \\ 0 & 0 & .33 \\ 0 & 0 & .31 \end{bmatrix}$$

$$w(A, E) = \begin{bmatrix} .23 & .12 & .77 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Fig. 3. The sketch of an image, its representation as a graph, and weighted walk-throughs between objects A and E , and between objects A and B . In the graph, nodes labeled with the object type and edges are labeled with weighted walk-throughs representing the mutual relationship between connected objects.

Query Specification: Queries are formulated as properties of the spatial relations among typed objects in the picture. These properties are specified by drawing and arranging typed icons on a composition screen. The interpretation of the arrangement can be captured within a querying graph which basically shares the same syntactic structure of the description graph used in image descriptions. The only difference is that, in the querying stage, the user is allowed to express a concept of relevance by selecting which object relationships appearing in the example are actually relevant for the query. In the syntactic perspective, this implies that the querying graph is not complete. Fig.4 shows the sketch of a query and its description in the form of a non-completely-connected labeled graph.

Automated Model Checking: Model checking of a query against an image description can be formulated as a problem of weighted sub-graph matching. The algorithm reduces the comparison between the description and the query to the problem of checking a graph against a sequential formula. To this end, the graph of the query is replaced through a sequential formula which represents a linear path traversing exactly once all the edges of the querying graph.

The algorithm operates through two subsequent stages performing a top-down syntactic decomposition of the formula and a bottom-up semantic labeling of the image description graph. In this checking, the nodes of the image graph are recursively labeled with the degree of similarity of sub-formulae that are satisfied in the node itself. The labeling proceeds incrementally in a bottom-up manner starting with sub-formulae of minimum length. On the n^{th} step, the labeling of formulae with length lower than n is used to compute the labeling of the formulae with length $n + 1$. The maximum degree of similarity obtained on any object for the main formula represents the degree of similarity between the image description and the querying sketch.

Checking complexity mainly depends on the number of labels that the algo-

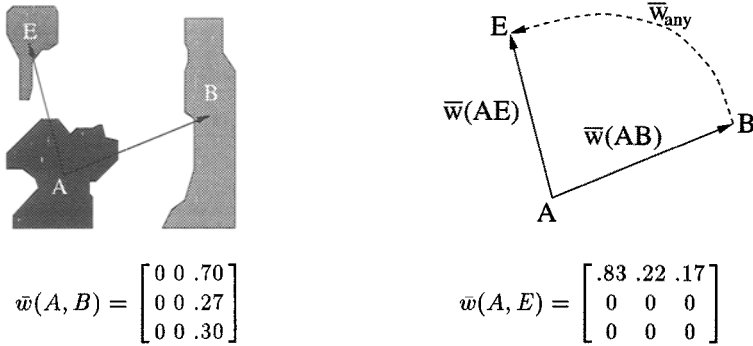


Fig. 4. The sketch of a query, its representation as a graph according to the style of Fig. 3, and the weighted walkthroughs between objects A and B and between objects A and E . In the sketch, relationships between A and E , and between A and B have been marked as relevant to the purposes of the query. In the graph, a fictitious edge, drawn as a dotted line, is added between objects B and E to permit application of the chinese postman algorithm. The fictitious edge is labeled with the conventional relationship \bar{w}_{any} which is assumed to exactly match any relationship w (i.e. $\mathcal{S}(\bar{w}_{any}, \cdot) = 1$).

rithm computes. In the worst case that both the query and the description contain objects of the same type, if n_d and n_q are the number of objects in the description and the query, respectively, this number is in the order of $D_{n_d n_q} = \frac{n_d!}{(n_d - n_q)!}$.

By relying on the fact that the similarity degree is monotonically decreasing with the execution of the algorithm, a heuristic reduction of complexity can be obtained by enabling early discard of four-tuples scoring spatial or type similarities under minimum accepted thresholds μ^{min} .

4 Discussion

An original technique for modeling of spatial relationships between imaged objects has been introduced. The model, referred to as weighted walkthroughs, represents the mutual positioning of two objects as a 9-tuple of weights, each providing integral measure for the significance of one out of nine alternative directional relationships. The use of multiple weights encompasses vague relationships between complex and near objects near objects without requiring their replacement through centroids or minimum embedding rectangles. Besides, the use of quantitative weights enables metric evaluation of similarity supporting imprecise matching. Finally, integral derivation of weights supports analytical treatment and ensures continuity in the relation between spatial arrangements and their descriptions. Weighted walkthroughs have been embedded within an original graph-theoretical framework supporting representation of spatial arrangements among multiple typed objects, each type allowing for multiple instances. A model checking approach permits quantitative evaluation of similarity among different

spatial arrangements with corresponding types. This opens the way to the use of the entire framework as theoretical basis of engines for visual retrieval by contents supporting similarity matching between image descriptions and user's sketches.

Acknowledgement

The authors are indebted with Alberto Del Bimbo for his continuous support and guidance in the development of their research. His help is gratefully acknowledged.

References

1. S.K.Chang, S.H.Liu, "Picture indexing and Abstraction Techniques for Pictorial Databases", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.6, No.4, pp.475-484, July 1984.
2. S.K.Chang, Q.Y.Shi, C.W.Yan, "Iconic Indexing by 2-D Strings", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.9, No.3, pp.413-427, July 1987.
3. S.K.Chang, C.W.Yan, D.C.Dimitroff, T.Arndt, "An Intelligent Image Database System," *IEEE Transactions on Software Engineering*, Vol.14, No.5, pp.681-688, May 1988.
4. S.K.Chang, E.Jungert, "Pictorial Data Management Based upon the Theory of Symbolic Projections," *Journal of Visual Languages and Computing*, Vol.2, No.2, pp.195-215, June 1991.
5. A.Del Bimbo, E.Vicario, D.Zingoni, "Symbolic Description of Image Sequences with Spatio-Temporal Logic", *IEEE Transactions on Knowledge and Data Engineering*, Vol.7, No.4, Aug.1995.
6. A.Del Bimbo, E.Vicario, "Specification by-Example of Virtual Agents Behavior," *IEEE Transactions on Visualization and Computer Graphics*, Vol.1, No.4, Dec.1995.
7. M.J.Egenhofer, R. Franzosa, "On the Equivalence of Topological Relations," *International Journal of Geographical Information Systems*, Vol. 9, n. 2, 1992
8. M.J. Egenhofer, "Spatial Query by Sketch," *Proc. IEEE International Symposium on Visual Languages*, Boulder CO, 1996.
9. H.A.Eiselt, M.Gendreau, G.Laporte, "Arc Routing Problems, Part I: the Chinese Postman Problem," *Operations Research*, Vol.43, No.2, March-April 1995.
10. V.N.Gudivada, V.V.Raghavan, "Design and Evaluation of Algorithms for Image retrieval by Spatial Similarity," *ACM Trans. on Information Systems*, Vol.13, No.2, April 1995.
11. P.D.Holmes, E.Jungert, "Symbolic and Geometric Connectivity Graph Methods for Route Planning in Digitized Maps," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 14, n. 5, May 1992
12. E.Jungert, "The Observer's Point of View, an Extension of Symbolic Projections," *Proc. of Int.Conf.on Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, Pisa, Sept. 1992, Lecture Notes in Computer Science, Springer Verlag, 1992, pp.179-195.