An Asymmetric Approach to Signature Matching

Tatiana Jaworska

Abstract The image signature concept can be a successful approach to image comparison in content-based retrieval, but it is a very challenging task. Fundamental for this purpose is defining signature similarity. There exist a lot of similarity models which measure similarity of images or their objects in multimedia databases. In order to deal with semantic retrieval, we have introduced a three stage search engine. At each stage, we need different but equally effective similarity measures. Here, we have analysed asymmetric and symmetric approaches to signature comparison. In our experiment, we present an extensive comparison of some similarity measures dedicated to image retrieval.

Keywords Similarity measure · Image signature · Search engine · CBIR

1 Introduction

In the recent years, many researchers have intensively analysed similarity evaluations between whole images, their fragments, or some image elements, such as contours. Content-based similarity models have been developed [1] so as to comply with the system needs and user requirements regarding semantic multimedia retrieval.

According to Beecks et al. [1], a similarity model between the query and image or a group of image objects can be determined, even for a large multimedia database, by working out only the distance between their corresponding feature representations. We claim that, even though a query is automatically generated by a CBIR system or is prepared manually by a user as, for instance, we proposed in our system introducing a special dedicated GUI [2], the introduction of the spatial object location is strongly recommended.

T. Jaworska (✉)
Systems Research Institute Polish Academy of Sciences,
01-447 6 Newelska Street, Warsaw, Poland
e-mail: Tatiana.Jaworska@ibspan.waw.pl

Therefore, we provide in this paper a comparison between the Beecks' concept of feature signatures and their similarity measure, and our search engine concept where image signature and spatial object location are treated as global image information, but object features are only local. We decided to compare two kinds of signatures in order to check what gain it would bring if we found objects and compare their locations, as we proposed in our search engine [3, 4]. The afore-mentioned knowledge is crucial for the effectiveness of multimedia retrieval systems.

2 Signature Matching

2.1 Metrics Properties

Generally, when we analyse a metric space we assume by default that four basic conditions are satisfied:

- Non-negativity: $d(x, y) \geq 0$;
- Identity: $d(x, y) = 0 <=> x = y$;

- Symmetry: $d(x, y) = d(y, x)$;   (1)

- Triangle inequality: $d(x, y) + d(y, z) \geq d(x, z)$ for any points x, y, z of the set.

These conditions express our common notions of distance. For example, the distance between distinct points is positive and the distance from point A to B is equal to the distance from B to A.

We may also need to find the distance between two vectors, namely, feature vectors. Then, in a normed vector space $(X, \|\bullet\|)$ we can define a metric on X by

$$d(x, y) = \|x - y\|$$ (2)

A metric defined in such a way is translation invariant and homogeneous. The most widely used similarity measure is the Euclidean measure. It can be applied to measure the distance between two points or between two feature vectors.

However, in real life the symmetry is questionable, for example, the way up a hill and down a hill takes different time. A similar situation is when we compare images. We can imagine various criteria, for instance, the number of particular elements or segments which constitute a query. Hence, when we select as a query an image among the previous matching images, we obtain a different set of matching images because the symmetry is incomplete.

In such a situation a quasimetric may be needed. A quasimetric is defined as a function that satisfies the previously mentioned axioms for a metric without symmetry:

$$d(x, y) \neq d(y, x). \tag{3}$$

This notion is more often used in practice than in mathematics, and that is why sometimes it is called a semimetrics [5].

2.2 Signatures

In our system [3], at the first stage, objects o_{ij} are extracted from an image I_i based on low-level features. These features are used for object classification. Additionally, the objects' mutual spatial relationship is calculated based on the centroid locations and angles between vectors connecting them, with an algorithm proposed by Chang and Wu [6] and later modified by Guru and Punitha [7], to determine the first three principal component vectors (PCV_{oi}, $i = 1, \ldots, 3$ for each object o_{ij}). Spatial object location in an image is used as the global feature [4].

Definition 2.1 (*Image signature* [3])
Let the query be an image I_q, such as $I_q = \{o_{q1}, o_{q2}, \ldots, o_{qn}\}$, where o_{ij} are objects. An image in the database is denoted as I_b, $I_b = \{o_{b1}, o_{b2}, \ldots, o_{bm}\}$. Let us assume that in the database there are, in total, M classes of the objects marked as L_1, $L_2, \ldots, L_M$. Then, as the image signature I_i we denote the following vector:

$$\text{Signature}(I_i) = [\text{nobc}_{i1}, \ \text{nobc}_{i2}, \ \ldots, \text{nobc}_{iM}] \tag{4}$$

where: nobc_{ik} are the number of objects o_{ij} of class L_k segmented from an image I_i. Note that the length of a signature is always the same and is equal to M.

As the second kind of signature we adopt the feature signature defined by Beecks et al. [8] in 2009 who aggregated features into a compact feature representation, for the purpose of effective calculation of similarity between two data objects. Rubner et al. [9], used two common feature representation types: feature histograms and feature signatures, which were worked out from global partitioning of the feature space and local feature clustering for each data object. Contrary to our approach, these authors applied global partitioning to the feature space, regardless of feature distributions of single objects, in order to create feature histograms which in turn correspond to the number of features located in the global partition.

According to Beecks et al. feature signature is defined as follows:

Definition 2.2 (*Feature signature* [1])
Let $FS \subseteq R^k$ be a feature space and $C = C_1, \ldots, C_k$ be a local clustering of the features $f_1, \ldots, f_n \in FS$ of object o_{ij}. Then a feature signature S^o of length M^i can be defined as a set of tuples from a $FS \times R^+$ such as:

$$S^o = \{c_k^o, w_k^o, k = 1, \ldots, M\}. \tag{5}$$

where: $c_k^o = \frac{\sum_{f \in c_k} f}{|c_k|}$ is a centroid of similar objects o_{ij} of image I_i and $w_k^o = \frac{|c_k|}{n}$ is its weight.

It means that a feature signature S^o of object o_{ij} is a set of centroids $c_k^o \in FS$ with the corresponding weights $w_k^o \in R^+$.

According to Definition 2.2. carrying out the feature clustering individually for each data object reflects aggregation of feature distribution in a better way than any feature histogram. However, feature histograms are a special case of feature signatures, whose centroids stay the same for the whole database and the information about objects is reflected only via weights, which results in a limitation of object representation.

By this approach, Beecks et al. aggregated the objects' location in the feature space which is substituted only by grouping similar feature values in signature and histogram form. They proposed only seven basic features: two coordinates, three components of colour and two texture descriptors, whereas we offered 45 features for a particular object, for example: moments of inertia and Zernike's moments [10].

In our adaptation of their method, a number of objects of a particular class were interpreted as weights. Object centroids represent locations of real, early segmented, objects in the image space. Here, class centroids are situated in the geometrical centre among particular object centroids. We also use different methods to determine the similarity in these two approaches.

2.3 Similarity Functions

Asymmetry is one of the most controversial properties of similarity. In this subsection we describe the asymmetric approach to image signature matching and a signature quadratic form distance in comparison with standard similarity measures, such as Euclidean, absolute difference or Hamming. All the measures are implemented in our search engine [3, 4].

In order to answer the query I_q, we compare it with each image I_b from the database in the following way. A query image is obtained from the GUI, where the user constructs their own image from selected DB objects. First of all, we determine a similarity measure sim_{sgn} between the signatures of query I_q and image I_b:

$$\text{sim}_{\text{sgn}}(I_q, I_b) = \sum_i (\text{nob}_{qi} - \text{nob}_{bi}) \tag{6}$$

computing it as an equivalent with the Hamming distance between two vectors of their signatures (cf. (4)), such that $\text{sim}_{\text{sgn}} \geq 0$ and $\max_i(\text{nob}_{qi} - \text{nob}_{bi}) \leq \text{thr}$, thr is the limitation of the quantity of elements of a particular class by which I_q and I_b can

differ. It means that we select from the DB images with the same classes as the query. The above comparison is asymmetric because if we interchange the query and the image, we obtain negative similarity value, that is: $\text{sim}_{\text{sgn}}(I_q, I_b) = -\text{sim}_{\text{sgn}}(I_b, I_q)$. Then, the condition of non-negativity of similarity is incomplete. This fact is crucial from the semantic matching point of view because the human brain recognizes things in context with others.

If the maximum component of (6) is bigger than a given threshold (a parameter of the search engine set by the user), then image I_b is discarded. Otherwise, it goes to the next step and we find the spatial similarity sim_{PCV} (7) of images I_q and I_b, based on the City block distance between their PCVs as:

$$\text{sim}_{\text{PCV}}(I_q, I_b) = 1 - \sum_{i=1}^{3} |PCV_{bi} - PCV_{qi}| \tag{7}$$

Definition 2.3 (*Signature Quadratic Form Distance* [1])

If $S^o = \{c_k^o, w_k^o, k = 1, \ldots, M\}$ and $S^q = \{c_k^q, w_k^q, k = 1, \ldots, N\}$ are two feature signatures, then the Signature Quadratic Form Distance (SQFD) between S^o and S^q is defined as:

$$\text{SQFD}(S^o, S^q) = \sqrt{(w_o| - w_q)A(w_o| - w_q)^T} \tag{8}$$

where: $A \in R^{(M+N) \times (M+N)}$ is the similarity matrix, $w_q = (w_1^q, \ldots, w_m^q)$ and $w_o = (w_1^o, \ldots, w_n^o)$ are weight vectors and $(w_o| - w_q) = (w_1^o, \ldots, w_n^o, -w_1^q, \ldots, -w_m^q)$.

The similarity matrix **A** can be constructed assuming that there exists a similarity function $f: FS \times FS \rightarrow R$. The $a_{k,l}$ components of **A** are calculated as follows:

$$a_{k,l} = f\left(c_k^o, c_l^q\right) = \frac{1}{1 + d\left(c_k^o, c_l^q\right)} = \frac{1}{1 + \left[\left(c_{k,x}^o - c_{l,x}^q\right)^2 + \left(c_{k,y}^o - c_{l,y}^q\right)^2\right]} \tag{9}$$

where: $k, l = 1, \ldots, N + M$.

In our approach, there is the same number of classes for each image signature ($N = M$), hence we decided to assume the length of vectors w_q and w_o equal to M which implies the size of a square matrix $A_{[M \times M]}$. Then the signature form distance (8) can be simplified to the form:

$$\text{SQFD}(S^o, S^q) = \sqrt{w_o A w_q^T} \tag{10}$$

and $a_{k,l}$ components are computed only for $k, l = 1, \ldots, M$, according to (9). Here, in Definition 2.3 and in our approach, S^q means a query signature, whereas S^o means image signatures in the database. The signature similarity, computed according to SQFD (cf. (10)), gives more information than the one computed it according to sim_{sgn} (cf. (6)) which is seen in the results.

3   Results

Below we present examples of matching results obtained for the above-mentioned similarity measures. We applied two queries designed in our GUI. The former is a semidetached building with a hip roof and the latter is a terraced house with three gable roofs. From the semantic matching point of view, the best results should be images of houses with the same kinds of roofs and similar number of building segments. All figures present query (far left picture in each) and 11 best matched images which are ordered decreasingly, according to the similarity to the query. Figures 1 and 2 present results found according to the asymmetric image signature (cf. (6)). We can see that both results contain buildings with two flat roofs and one with a semicircular type, which are not desired.

Figures 3 and 4 present matching for both queries computed according to our modification of signature form distance (cf. (10)) and all results fulfil the semantic requirements. Figures 5 and 6 present matching for these same queries, computed according to the signature quadratic form distance (cf. (8)).

Generally, in the case of semantic comparison, it is difficult to compare these results in a quantitative way, so that is why we present the result in full form. This gives us the opportunity for a qualitative evaluation. Hence, as we can see, especially in Fig. 6, where there are no fully correct matchings because separate doors,
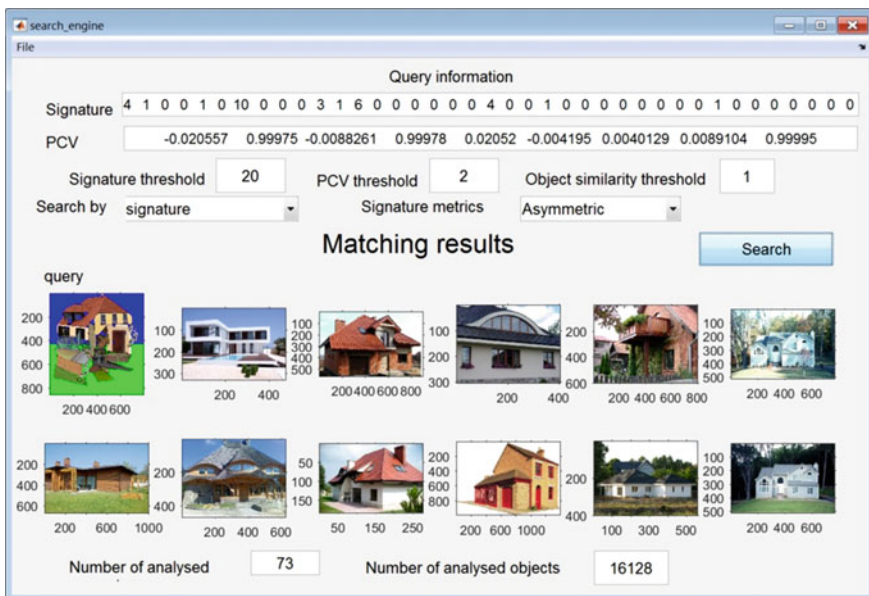

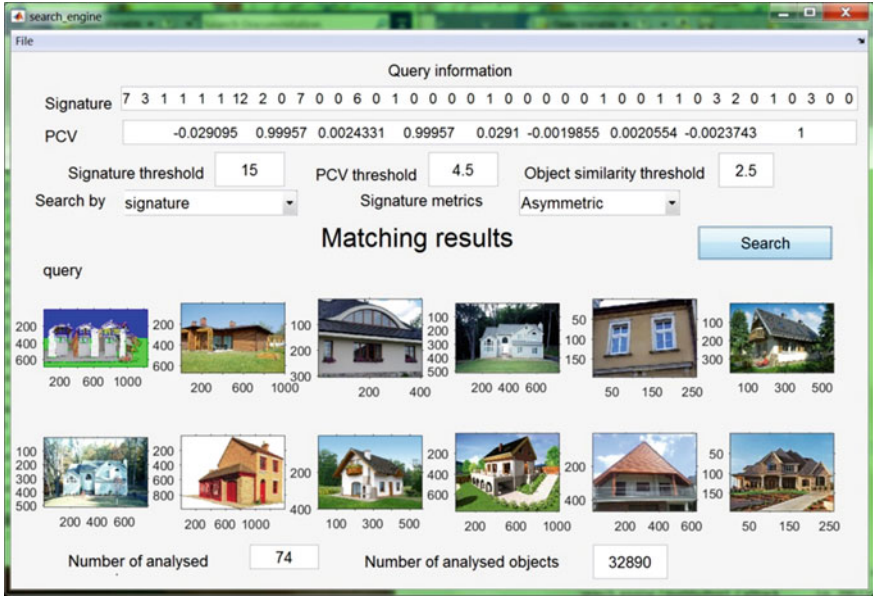**Fig. 1** Matching results for image signature for query 1

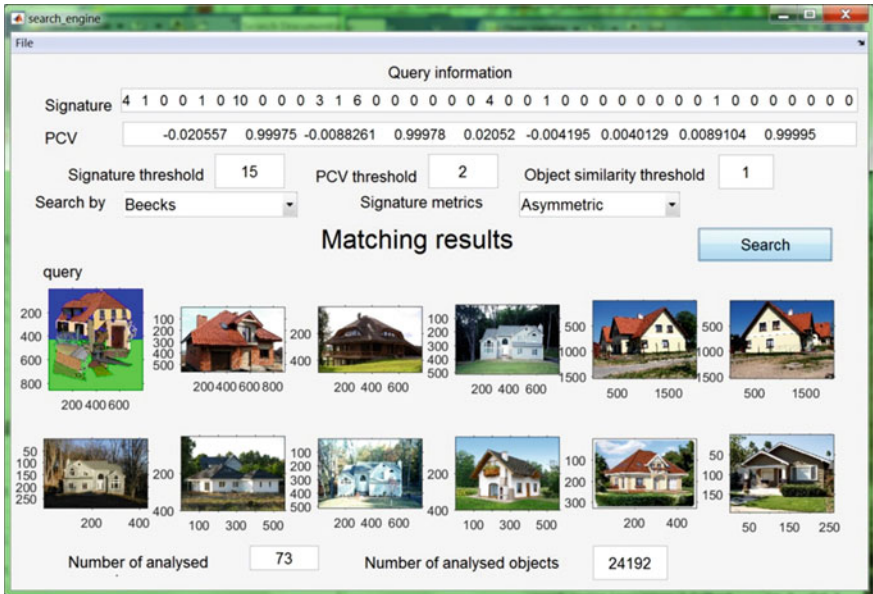**Fig. 2** Matching results for image signature for query 2

Fig. 3 Matching results for the signature form distance for query 1

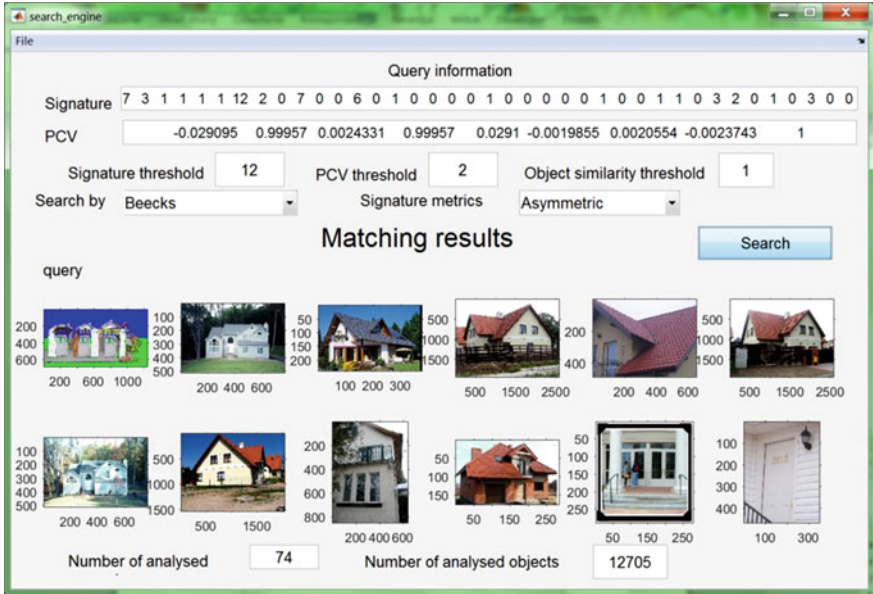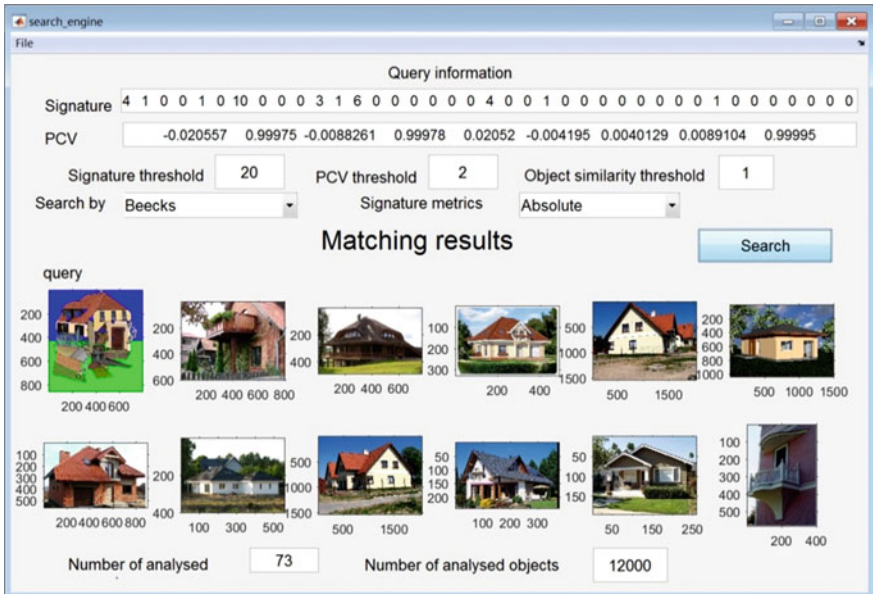**Fig. 4** Matching results for the signature form distance for query 2

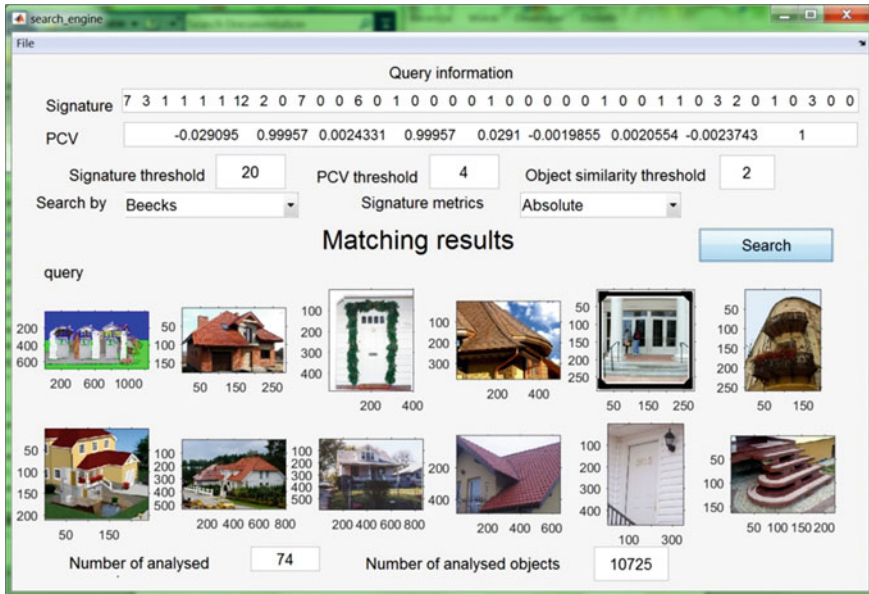**Fig. 5** Matching results for signature quadratic form distance for query 1

Fig. 6 Matching results for signature quadratic form distance for query 2

stairs, balconies appear strongly undesirable. It happens because the Beecks' team analysed less information about object spatial location than we did. Even though we used the simplified signature form distance, we obtained better results thanks to the fact that we added the separate object spatial location similarity (cf. (7)).

4 Discussion

In our analysis we have not decided to add such a popular approach as the SIFT method [11], because it mainly concentrates on finding a particular object similarity without a deep object spatial location analysis. The example of such a matching is shown in Fig. 7 where to the three terraced buildings seven houses with balconies were matched because a gutter or another less important element added in the query were found in the DB.
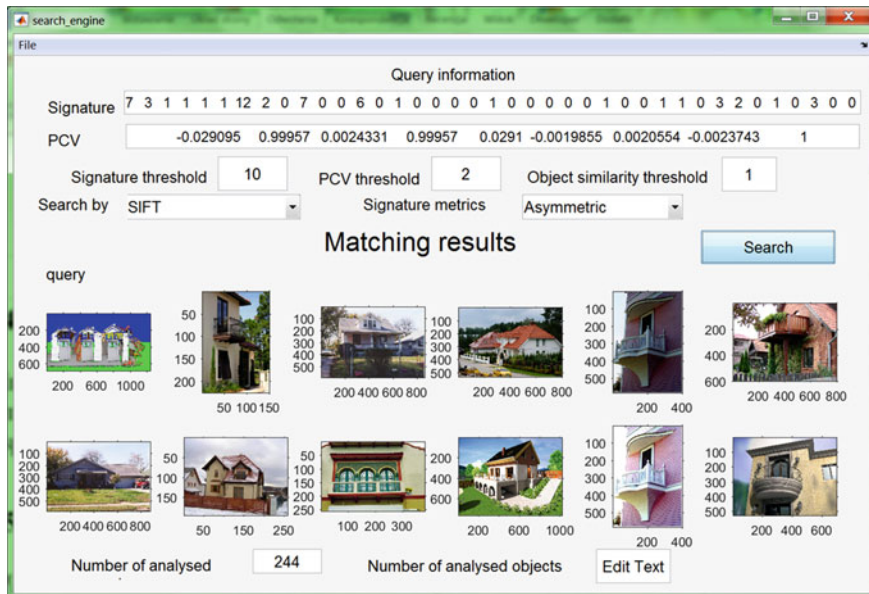
Fig. 7 Matching results based on the SIFT method for query 1

5 Conclusion

In this paper we compare the asymmetric and symmetric similarity measures applied to two kinds of signatures implemented in our content-based image retrieval system. In order to present the evaluation of the above-mentioned similarity measures, we used the database created in our institute containing mainly images of houses coming from the Internet.

We can observe that in a situation when a signature similarity is enhanced by object spatial location, the quality of semantic matching is better. All these similarity measures are applicable to signatures of different size and structure.

References

1. Beecks, C., Uysal, M.S. Seidl, T.: A comparative study of similarity measures for content-based multimedia retrieval. In: Multimedia and Expo (ICME), Suntec City, 19–23 July 2010
2. Jaworska, T.: Query techniques for CBIR. In: Andreasen, T., Christiansen, H., Kacprzyk, J., Larsen, H., Pasi, G., Pivert, O., De Tre, G., Vila, M.A., Yazici, A., Zadrożny, S. (eds.) Flexible Query Answering Systems, vol. 400, pp. 403–416. Springer, Cracow (2015)
3. Jaworska, T.: A search-engine concept based on multi-feature vectors and spatial relationship. In: Christiansen, H., De Tré, G., Yazici, A., Zadrożny, S., Larsen, H.L. (eds.) Flexible Query Answering Systems, vol. 7022, pp. 137–148. Springer, Ghent (2011)

4. Jaworska, T.: Spatial representation of object location for image matching in CBIR. In: Zgrzywa, A., Choroś, K., Siemiński, A. (eds.) New Research in Multimedia and Internet Systems, vol. 314, pp. 25–34. Springer, Wrocław (2014)
5. S. Rolewicz, Functional Analysis and Control Theory: Linear Systems, vol. Series: Mathematics and its applications, PWN-Polish Scientific Publishers, Warsaw 1987
6. Chang, C.-C., Wu, T.-C.: An exact match retrieval scheme based upon principal component analysis. Pattern Recogn. Lett. **16**, 465–470 (1995)
7. Guru, D.S., Punitha, P.: An invariant scheme for exact match retrieval of symbolic images based upon principal component analysis. Pattern Recogn. Lett. **25**, 73–86 (2004)
8. Beecks, C., Uysal, M.S. Seidl, T.: Signature quadratic form distances for content-based similarity. In: ACM Multimedia, Beijing, China, 19–24 Oct 2009
9. Rubner, Y., Tomasi, C., Guibas, L.J.: The earth mover's distance as a metric for image retrieval. Int. J. Comput. Vision **40**(2), 99–121 (2000)
10. Teague, M.R.: Image analysis via the general theory of moments. J. Opt. Soc. America **70**(8), 920–930 (1980)
11. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)