Pyramidal Objects and Comparing Objects Using Similarity Measures

D. Klimešová, J. Konopásek, E. Ocelíková

***Abstract*—**In the paper the problems of objects similarity measures and detection of similarity of objects are addressed. Currently, when using the object-based representation and object oriented approaches we frequently meet the problem of object similarity. This problem spans many different application domains in which it is necessary to apply decision making process. Given a pair of objects, it is of interest to know how they are related to each other. Similarity measures can be used in many types of data retrieval, data mining and many analysis tasks. Very often we can group the objects of a given application into clusters based on their similarity values. Sophisticated methods use multiple levels of objects in the frame of one task and different types of similarity: attribute, correlation or behaviour based similarity measures.

Keywords—Associated objects, pyramidal objects, object-based representation, similarity measures.

I. INTRODUCTION

The object-based representation and object oriented approaches deal with the problem of object similarity in the different application domains. It can be viewed as problem of classification or generally decision making process. The task is to select the best solution, most appropriate alternative of solution from the set of available solutions with respect to the defined goal. The goal determines the way how the alternative solutions are investigated to have enough information for decision making. The formulation of goals is often associated with significant problems connected with the ability to refine the definition.

D. Klimešová is with the Czech University of Life Sciences, Prague, Faculty of Economics and Management, Department of Information Engineering, CZECH REPUBLIC, (e-mail: klimesova@pef.czu.cz), and with the Institute of Information Theory and Automation, Czech Academy of Sciences, Department of Image Processing , email: klimes@utia.cas.cz) as associate professor. Education: Master Degree in Numerical Mathematics, Faculty of Mathematics and Physics, Charles University in Prague, 1973. Ph.D. Degree in Computer Science, 1987. Research interests: Spatial modelling, geographical information management, context mapping, GIS, image processing. Education activities: Geographical information systems and image processing, remote sensing, system engineering.

J. Konopásek is doctoral student on Czech University of Life Sciences, Prague, Faculty of Economics and Management, Department of Information Engineering, Kamýcká 129, 165 21 Prague, CZECH REPUBLIC, (e-mail: konopasek@pef.zu.cz).

E. Ocelíková is working at the Department of Cybernetics and Artificial Intelligence of the Faculty of Electrical Engineering and Informatics at Technical University in Košice as professor. Her research work includes problems of decision processes, especially the problems of multicriterial classifications designing and high dimensionality reduction of feature space of multidimensional data in decision (e-mail: ocelike@ccsun.tuke.sk).

The decision-making criterion is the rule that explicitly allows us to choose the best solution. Decision rule is exactly defined when it prescribes the algorithm determining the solution that will be optimal with respect to the specific criteria.

We have to take into consideration the external conditions – a complete system of events affecting the final solution. The concept of similarity has been used and studied in many application domains like Data Analysis, Pattern Recognition, Data Mining, Cognitive Sciences and many others [1], [2].

The usual approach is to use metric definitions and evaluation. The measures of similarity are compared with some distance in the feature space, and also influence setting of feature weights to make objects more similar to those in the same category and dissimilar to those in different categories [3], [4].

Many similarity measures are based on comparing the internal feature values of the object or objects correlation. Generally speaking these two approaches are feature-based and correlation-based similarity. Most previous studies on correlation consider the co-occurrence based correlation, where two objects are considered correlated if they occur together in transactions, in definite distances and directions or in the defined arrangement [5].

The geographical database contains a lot of objects and we cannot only compare theirs properties we need to set up different requirements to detect or discover at first not apparent context relationship between objects.

Usually, we tend to think that similar objects should be those objects whose internal features are very similar to each other or those which co-occur often. Many similarity measures have been designed to capture such thinking [6], [7].

These measures use different ways to check internal feature values of objects or co-occurrences of objects. Such similarity can be discovered from data sets of vectors of attribute values type or transaction dataset type evaluation.

II. OBJECT DESCRIPTION

A. Feature Space

We will assume that each object is described with the attributes. The attribute is a variable describing a characteristic of the investigated object. In most cases, the objects are characterized by a number of the attributes. Attributes can be assigned the weight of significance but usually they are elementary (simple) and without weights coefficients.

Let us assume *m*-dimensional feature space, then every *i*-th object, where $i = 1,2,...,n$ is defined by *m*-dimensional vector $\boldsymbol{x}^i = \left(x_{i1}, x_{i2}, \cdots, x_{im}\right)$. Then the input matrix of the objects $\boldsymbol{X}$ can be written as a matrix of dimensions *n* x *m* with elements x_{ij}, where $i = 1,\ 2,\ \ldots,\ n$ and $j = 1,\ 2,\ \ldots,\ m$ *in* shape (1)

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \qquad (1)$$

The lines of matrix *X* represent data vectors of objects and the columns represent the values of their attributes [8].

Attributes may be qualitative or quantitative. The values of qualitative attributes are expressed with words, the value of quantitative attributes we set the measurement units. These attributes are often incomparable to each other just because of the different units of measure.

In such cases it is necessary to adjust the value of the attributes so that they become comparable. The most common way to do this is to use methods of data normalization. The data normalization aim is to transform values of all the attributes into the interval [0,1].

B. Normalization Procedures

There are many methods for data normalization [8]. If the highest attribute value is the best, then we subtract from the value of each attribute its minimum value and then we divide with variation range *Rj* of this attribute.

If the lowest value is the best, then we subtract from the maximum value of attribute the real value and then we divide with the variation range *Rj* of this attribute. Min-max normalization performs a linear transformation on the original data. Each attribute is normalized by scaling its values so that they fall within a small specific range, such as 0.0 and 1.0. When the actual minimum and maximum of attributes are unknown, or when there are outliers that dominate the min-max normalization, than the z-score normalization should be used.

In z-score normalization, the values for an attribute are normalized based on the mean and the standard deviation. Almost all similarity measures are indicators of attributes based on the comparison of the attribute-value expressing the basic properties of the objects.

C. Dissimilarity Matrix

The dissimilarity matrix stores a collection of proximity values for all pairs of objects. This matrix is often represented by an $n \times n$ table. We can see the dissimilarity matrix D_m corresponding to the data matrix *X* in (1), where each element $d(i, j)$ represents the difference or dissimilarity between objects *i* and *j*.

$$Dm = \begin{bmatrix} 0 & & \cdots & & d(1,n) \\ d(2,1) & 0 & & \cdots & d(2,n) \\ d(3,1) & d(2,2) & 0 & & d(3,n) \\ & \vdots & & \ddots & \vdots \\ d(n,1) & d(n,2) & d(n,3) & \cdots & 0 \end{bmatrix} \qquad (2)$$

In general, $d(i, j)$ is a nonnegative number that is close to zero when the objects *i* and *j* are very similar to each other, and becomes larger the more they differ. To calculate the dissimilarity between objects *i* and *j* we can use many different kinds of distance measures. The most popular distance measure is Euclidean distance

If $i = (x_{i1}, x_{i2}, \ldots x_{in})$ and $j = \left(x_{j1}, x_{j2}, \ldots x_{jn}\right)$ are *n*-dimensional data objects, the Euclidean distance between *i* and *j* is given by:

$$d(i,j) = [\sum_{k=1}^{n} (x_{ik} - x_{jk})^2]^{1/2} \qquad (3)$$

III. OBJECT SIMILARITY

A. Metric

For the determination of object similarity are used the measures of similarity (as well as dissimilarity). Similarity between two objects $\boldsymbol{x}^i$ and $\boldsymbol{x}^j$ can be expressed using the distance function d that assigns for each pair of objects ($\boldsymbol{x}^i$, $\boldsymbol{x}^j$) a certain number with properties as follow:

$$\begin{aligned} &d\left(\boldsymbol{x}^i,\ \boldsymbol{x}^j\right) = 0 \ \rightarrow \boldsymbol{x}^i = \boldsymbol{x}^j && \text{identity} \\ &d\left(\boldsymbol{x}^i,\ \boldsymbol{x}^j\right) = d\left(\boldsymbol{x}^j, \boldsymbol{x}^i\right) && \text{symmetry} \\ &d\left(\boldsymbol{x}^i,\ \boldsymbol{x}^j\right) \geq 0 && \text{positivity} \\ &d\left(\boldsymbol{x}^i,\ \boldsymbol{x}^j\right) = 0 && \text{minimality} \end{aligned} \qquad (4)$$

In case we are working with the quantitative attributes, then the expression similarity (dissimilarity) relationship between two objects is the distance between them. Less distance between objects means that these objects are more similar. If the function d satisfies also the triangular inequality (5) then we talk about metric [1], [7].

$$d\left(\mathbf{x}^i, \mathbf{x}^j\right) + d(x^j, x^k) \geq d\left(\mathbf{x}^i, x^k\right) \tag{5}$$

where $i, j, k \in \langle 1, n \rangle$.

The similarity/distance measure reflects the degree of closeness or dissimilarity of the target objects and should correspond to the characteristics that are believed to distinguish them. In many cases, these characteristics are dependent on the data or the context of a given problem. There is no measure that is universally best for all kinds of problems [9].

Various distances have been developed according to the type of attributes (Euclidian distance, Mahalanobis distance, Minkowski, Manhattan, and others). Nominal variables can be encoded either by asymmetric binary variables or by mapping them into a numeric domain.

Quantitative variables can be discreet or continuous, dichotomous (symmetric or asymmetric) or multi-categorical. In case, we are working with real values, the mostly used metric is the Minkowski measure [10].

Minkowski distance is a generalization of Euclidean distance where λ is a parameter, n is the number of dimensions (attributes) and x_i and y_i express the value of *i-th* attribute describing the individual object *x* and *y*.

$$d_\lambda(\mathrm{x}, \mathrm{y}) = \left[\sum_{i=1}^{n} W_i \ |x_i - y_i|^{\lambda}\right]^{1/\lambda} \tag{6}$$

with $\lambda > 0$.

W_i is the numerical weight correlated with this attribute.

According to the value of the parameter λ we can retrieve some well-known distances such as Manhattan distance (λ = 1) and Euclidian distance (λ = 2).

To transform the Minkowski distance (6) into a similarity measure S_λ (x, y) we just need to introduce a value *Di* corresponding to the difference between the upper and the lower bounds of the range of the *i-th* attribute:

$$S_\lambda(\mathrm{x}, \mathrm{y}) = \left[\sum_{i=1}^{n} W_i \ \frac{D_i - |x_i - y_i|^{\lambda}}{D_i}\right]^{1/\lambda} \tag{7}$$

with $\lambda > 0$.

In practice very often the mathematical properties defined as minimality, symmetry and triangular inequality are not verified. This is the reason why another ways are used to evaluate the degree of similarity *S (x, y)* between two individuals *x* and *y*, respectively they are described by a set of attributes *A* and *B* as

$$S_\lambda(\mathrm{x}, \mathrm{y}) = \frac{dm(A \cap B)}{dm(A \cup B) + \varphi\, dm(A - B) + \omega\, dm(B - A)}$$

with $\varphi, \omega \geq 0$, (8)

and dm stays instead of just used model.

B. Coefficients of Association

Association coefficients are calculated from the so-called association table that describes the similarity of each object pair x^i and x^j. It is used only in cases where objects are characterized by dichotomous attributes [11].

Because for each symptom there are admissible only two values 0 and 1, the association table has only two columns and two rows. Its shape is given in Table 1.

		x^j	
		1	0
x^i	1	a	b
	0	c	d

Table 1. Association table

a - the positive coincidence e.g. number of attributes, in which objects x^i and x^j take a value 1,

b - the number of disagreement cases, where the object x^j is 0 and the object x^i has value 1,

c - the number of disagreement cases, where the object x^j is 1 and the object x^i has value 0,

d - the negative coincidence e.g. number of attributes, in which objects x^i and x^j take a value 0 and $a + b + c + d$ is the number of attributes.

If we quantify the association table for all pairs of objects then we get the table showing the association between pairs of objects.

We can quantify the association coefficients S_{ij} for each table. These coefficients are expressing the similarity of alternatives x^i and x^j. All coefficients of association S_{ij}, where i, j = 1, 2, ..., n are elements of the matrix of associations **S** for which $S_{ij} = S_{ji}$, relationship (2).

$$S = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1n} \\ S_{21} & S_{22} & \cdots & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{n1} & S_{n2} & \cdots & S_{nn} \end{bmatrix} \tag{9}$$

C. Most Widely Used Coefficients

Sokal and Michener coefficient - S_{SM}

$$S_{SM} = \frac{a+d}{a+b+c+d} \qquad (13)$$

Jaccard coefficient - S_J

$$S_J = \frac{a}{a+b+c} \qquad (14)$$

The Jaccard coefficient, which is sometimes referred to as the Tanimoto coefficient, measures similarity as the intersection divided by the union of the objects.

Russell and Rao coefficient - S_{RR}

$$S_{RR} = \frac{a}{a+b+c+d} \qquad (15)$$

The disadvantage of this coefficient is the similarity of objects with one's self assessed differently. The coefficient S_{RR} is equal to 1 if the object has all of attribute values equal to 1. On the other hand coefficient has a value of zero if all the signs are zero.

Dice coefficient, Rogers and Tanimoto or Hamann coefficient - S_H

$$S_H = \frac{(a+d)-(b+c)}{a+b+c+d} \qquad (18)$$

The coefficient of association S_H has different range of values $\langle -1,1 \rangle$, while the other field values of association coefficients are in the interval $\langle 0;1 \rangle$.

Coefficient S_H shall take the negative value -1 if the two objects do not correspond in any symptom, the value 1 if the pair coincides in all attributes and it shall take the value 0 if the number of positive and negative coincidence is equal.

IV. EXPERIMENTAL PART

A. Car Factories and Car Attributes

The coefficients of association referred to in the previous sections were applied to case study where compared objects were set of car factories. The aim was to determine which of them have similar product lines. Data on car factories were taken from their actual product catalogs.

The following car factories were compared. Their serial numbers are used as identification numbers when relating to them in other tables:

Opel

1. BMW
2. Mercedes - Benz
3. Volkswagen
4. Porsche
5. Audi

Each carmaker was assessed according to the following 10 dichotomous symptoms.

1. urban mini – length to 3,5 meters, engine capacity up to 1 liter
2. small cars - 4 meters in length, engine capacity to 1,4 liters
3. lower middle class - a length of 4,0 to 4,3 meters, engine capacity from 1,4 to 1,8 liters
4. middle class – 4,5 meters in length, engine capacity to 2,5 liters
5. upper middle class - 5 m in length, powerful engine, luxury trim
6. luxury cars - the length over 5 meters, powerful engine, luxury trim
7. SUV / Off road - better ground clearance chassis, all-wheel drive
8. MPV - multi-purpose vehicle, the maximum space and comfort
9. sport cars - high speed and performance, powerful engine
10. utility vehicles - delivery of cargo, large cargo space.

The association tables with values *a, b, c, d* are expressed for each pair of automakers and they are arranged separately in the Table 2, from which shall be expressed the association coefficients.

For illustration we show the Jaccard coefficient values for the all pairs of automakers which are calculated from the Table 2 in Table3.

The smallest is the similarity between object 1 and 5 with the lowest coefficient of association 0.125, the biggest similarity is between objects 2 and 6 with the highest association coefficient 0,857.

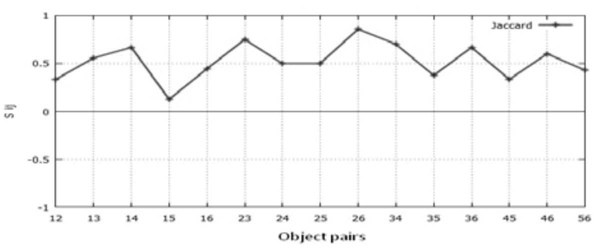

Fig. 1. Similarity of object pairs expressed by Jaccard coefficients.

	1	2	3	4	5	6
1	6 0 0 4	3 3 3 1	6 0 3 1	5 1 2 2	1 5 2 2	4 2 3 1
2	3 3 3 1	6 0 0 4	5 1 4 0	5 1 2 2	3 3 0 4	6 0 1 3
3	6 3 0 1	5 4 1 0	9 0 0 1	7 2 0 1	3 6 0 1	6 3 1 0
4	5 2 1 0	5 2 1 5	7 0 2 1	7 0 0 3	3 4 0 3	5 2 2 1
5	1 2 5 2	3 0 3 4	3 0 6 1	3 0 4 3	3 0 0 7	3 0 4 3
6	4 3 2 1	6 1 0 3	6 1 3 0	5 2 2 1	3 4 0 3	7 0 0 3

Table 2. Association tables of all objects.

	1	2	3	4	5	5
1	1.000	0.333	0.667	0.625	0.125	0.444
2	0.333	1.000	0.500	0.625	0.500	0.857
3	0.667	0.500	1.000	0.778	0.333	0.600
4	0.625	0.625	0.778	1.000	0.429	0.556
5	0.125	0.500	0.333	0.429	1.000	0.429
6	0.444	0.857	0.600	0.556	0.429	1.000

Table 3. Matrix of Jaccard coefficients of association.

The similarity of object pairs according Jaccard association coefficient from the Table 3 is shown in Fig.1.

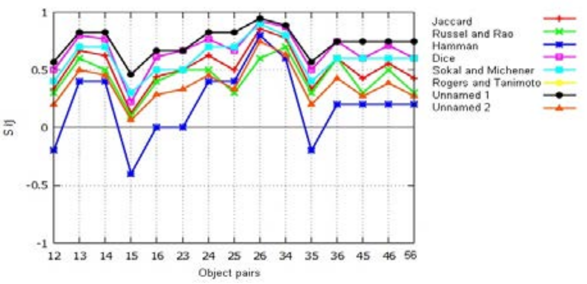

Fig. 2. The similarity of object pairs

V. COMMON CORRELATION MEASURES

A. Transactions

We know that as long as pair of objects occurs in at least one transaction, then there is a co-occurrence based correlation relationship between these two objects [12], [13].

We define the concepts of co-occurrence based similarity. Let $L = \{I_1, I_2, \dots I_n\}$ be a set of n binary attributes called items. These items will also be referred to as objects.

Let $D = \{T_1, T_2, \dots T_m\}$ be a set of transactions where each transaction T is a set of items from L. Each transaction is associated with an identifier, called *TID*, and contains a subset of the items in L. A set of items is called an itemset. An itemset that contains k items is a k-itemset.

Transaction T is said to contain an itemset A if and only if $A \in T$. A correlation relationship is a pair of itemsets (A, B), where $A \in L$, $B \in L$, and $A \cap B = \{ \}$. When A and B are both single items, we sometimes refer to (A, B) as an object pair. A special type of correlation between A and B is association, denoted by $A \Rightarrow B$.

Transaction ID	**Items**
1	milk, bread
2	bread, butter
3	honey
4	milk, bread, butter
5	bread

Table 4. Market data set

We will use a small example from the supermarket domain to illustrate the concept of correlation by co-occurrence. The set of items is $I = \{milk, bread, butter, honey\}$ and a small transactional database is shown in Table 4.

In this table, each row is a transactional record. The first column is the transactional ID, the second contains the items that were bought for the transaction identified by their ID. Most previous studies on correlation consider the co-occurrence based correlation, where two objects are considered correlated if they occur together in transactions. By checking the dataset in Tab. 4, we can find out correlation relationships like: Both milk and bread co-occur in transactions 1 and 4, so there is a co-occurrence based correlation relationship between milk and bread. Both bread and butter co-occur in transaction

2, so we can say that bread and butter have a co-occurrence based correlation relationship between them.

For the same reason, we have found that milk, bread, butter are correlated (by co-occurrence) with each other based on transaction 3 [20], [23].

B. Support and Confidence

We would like to know how intensely two objects are correlated to each other. To achieve this goal, we can introduce two concepts: supporting identifier *SP* and confidence *CF*.

The support *SP (X)* of an item set X is defined *as the proportion of transactions* in the data set which contain the item set X [18], [22], [26].

For example, in the sample database in table 2, the support count for the item bread is 4, since bread appears in transactions 1, 2, 4, 5.

The support value for bread, *SP (bread),* is 4 / 5 * 100 = 80%. The support count for {milk, bread} is 2, because they occur in transactions 1and4 and the support value *SP (milk, bread)* is 2 / 5 * 100 = 40%. 40% of all the transactions (2 out of 5 transactions) show that milk and bread were bought together.

Once we calculate the support values, we can use them to calculate the confidence values. The confidence of an association relationship/rule X=>Y is defined as:

$$CF\ (X\ => Y)\ =\ \frac{SP\ (X \cup Y)}{SP\ (X)} \qquad (19)$$

Confidence can be interpreted as an estimation of the probability *P(Y | X),* the probability of finding the Rhs (Right hand side) of the association rule in transactions under the condition that these transactions also contain the Lhs.

For example, the correlation relationship Milk => Bread has a confidence of 0.4 / 0.4 = 1 in Table 3, which means that all the transactions that contain milk also contain the bread as well. Also, we can get the confidence value for Bread => milk which is 0.4 / 0.8 = 0.5, and this means that among all the transactions that contain bread, only 50% of them also contain milk.

Support and confidence reflect the applicability and certainty of the association rule and can stay as criterions for evaluating the strength of an association rule of a correlation relationship. The supporting identifier and confidence concepts are used to introduce commonly favoured correlation measures and evaluate the correlation relationship between two objects like Cosine measure or Coherence measure.

B. Cosine Measure

Cosine is a simple correlation measure that is defined as follows. The occurrence of item set *A* is independent of the occurrence of item set *B* if

$$P\ (AB) = P\ (A) * P\ (B) \qquad (20)$$

which means that there is no correlation relationship between *A* and *B*. Otherwise, attribute-sets *A* and *B* are dependent and correlated to each other. The Cosine between the occurrence of *A* and *B* can be measured as follows:

$$Cosine\ (A, B)\ =\ \frac{P\ (AB)}{\sqrt{P(A) \times P\ (B)}}\ =\ \frac{SP\ (A \cup B)}{\sqrt{SP\ (A) \times SP\ (B)}} \qquad (21)$$

In the cosine equation, we take the square root on the product of the probabilities of *A* and *B* in the denominator because the cosine value should only be influenced by the supports of *A*, *B*, and *A* ∪ *B*, and not by the total number of transactions. The value range for the cosine measure is [0, 1].

If *A* and *B* are positively correlated, it means that the correlation relationship between *A* and *B* is strong than the resulting value of the cosine measure is larger or equal to 0.5 and smaller than 1. If the result value is larger or equal to 0 and smaller than 0.5, then the occurrence of *A* is negatively correlated with the occurrence of *B* which means that the correlation relationship between *A* and *B* is weak.

C. Coherence Measure

Coherence is another measure that is commonly used to evaluate the correlation relationship between a pair of objects. This measure is similar to the Jaccard similarity coefficient. The coherence measure i**s** calculated as follow:

$$Coherence\ (A, B) = \frac{SP\ (A \cup B)}{SP\ (A) + SP\ (B) - SP\ (A\ \cup B)} \qquad (22)$$

The meaning of this formula is that given two objects *A* and *B*, if they are strongly dependent on each other, then the value for *SP (A∪B)* should be very large, which is close to *min (SP (A), SP (B)).* In that case, the value for *(SP (A) + SP (B) – SP (A∪B))* should be close to the value of *max (SP (A), SP (B)).*

VI. BEHAVIOUR-BASED SIMILARITY

A. Similar Relationships

We tend to think that similar objects should be those objects whose internal features are very similar to each other or those which co-occur often.

These measures use different ways to check each object's internal feature values or co-occurrences of objects. Such similarity can be discovered from data sets of the attribute values or transaction dataset type.

In our real world, we have a lot of objects which do not have similar internal feature structures and which do not co-occur together often, but their relationships with other objects are very similar and their behaviour is similar [14], [15], [23].

Behaviour-based similarity can help us find more surprising similar object pairs and this can provide us more interesting information to use for further analysis.

The reason for this is because, if two objects co-occur a lot, then they ought to share a lot of correlated associated objects. When two objects share a lot of correlated associated objects, they can share behaviour-based similarity [19], [21].

B. Context and Similarity

Traditional approaches to object identification use features as the main source of information to evaluate similarity between objects.

We can also use different types of relations between objects and also different types of objects on one scene: pixels, segments, regions, scene objects and relation between them.

In case of surroundings: interposition, support, probability, position and familiar size. It means: Semantic context (probability), Spatial context (position), Scale context (size). We are dealing with context information from a global and local image level including interactions between pixels, regions, objects and object-scene interactions.

Context can be any information that is not directly produced by the appearance of an object. It can be obtained from the nearby image data, image tags or annotations and the presence and location of other objects.

Semantic context (probability): co-occurrence with other objects and in terms of its occurrence in scenes. This information is commonly obtained from training data, or from an external knowledge base [21], [23].

Fig. 3 Global and local interactions on the image

C. Local and Global Contextual Interactions

As seen from above we have in disposal a lot of different measures of similarity which may be used in a specific context of the task.

Highly sophisticated problems, for instance the scene recognition where the goal is to locate and identify instances of an object category within an image, use local interactions: pixel interactions, region and object interactions and global interactions: object-scene interactions, Fig. 3. Traditional approach is to use features as the main source of information: colour, edge responses, texture and shape [16], [17], [25].

We can use also different levels of relations between the objects and its surroundings: interposition, support, probability, position and others and to take into consideration suitable context: semantic context (probability), spatial context (position) and scale context (size). The semantic context means the co-occurrence with other objects and in terms of its occurrence in scenes [15].

.

The problem of effectively integrating context information is a challenging task.

VII. CONCLUSION

In this paper the problem of object similarity measures is addressed for different purposes in modern decision processes. This topic is closely connected with selected feature spaces, various context applying on local and global level and changing role of object. We are dealing with context information from a global and local image level. And also the content of object is temporally changing because it reflects changing interactions: pixel, region, object and object-scene interactions. Something like pyramidal objects that are finalized using some integrated context.

ACKNOWLEDGMENT

The paper was supported by the grant project IGA of the Czech University of Live Sciences: Methodology for Designing Relational Database and Web Interface for Dynamic GIS, no. 20131027. In Slovakia the work is supported by the grant project VEGA No. 1/0286/11, Dynamic hybrid architectures in multi-agent network control systems. Both supports are very gratefully acknowledged.

REFERENCES

[1] B. Verma, A. Rahman. Impact of Multicluster Characterization on Ensemble Classifier Learning. *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 4, 2012, pp. 605-618.

[2] F. S. Khan, M.A. Rao, J.v.d. Weijer, A.D. Bagdanov, M. Vanrell, A. Lopez, Color attributes for object detection, CVPR 2012.

[3] D. Klimešová, E. Ocelíková. GIS and image processing, International *NOUN Journal of Mathematical Models and Methods in Applied Sciences*, 2011, 5 (5), pp. 915-922.

[4] J. Andrade, J. Ares, R. Garcia, S.Rodriguez, S. Suarez. A Knowledge-Based System for Knowledge Management Capability Assessment, *WSEAS Transactions on Computers*, Issue 5, Volume 9, 2010.

[5] R. J. Hilderman and H. J. Hamilton, Knowledge discovery and measures of interest, Kluwer Academic, 2001.

[6] J. Zou, Ch. Liu, Y. Zhang, G. FuLu. Object recognition using Gabor co-occurrence similarity, *Pattern Recognition* 46 (2013).

[7] Z. Falomir, L. Museros, L. Gonzalez-Abril, M.T. Escrig, J.A. Ortega, A model for qualitative description of images based on visual and spatial features, *Computer Vision and Image Understanding* 116 (2012).

[8] Y. Liu, D. Zhang, G. Lu, and W. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognition*, vol. 40, no. 1, pp. 262–282, 2007.

[9] J.Figueira, S. Greco, M. Ehrgott. Multiple Criteria Analysis. USA, 2005, ISBN 0-387-23067-5.

[10] B. Wang, X. Zhang, M. Wang, and P. Zhao, Saliency distinguishing and applications to semantics extraction and retrieval of natural image, Machine Learning and Cybernetics 2010, vol. 2, July 2010, pp. 802 –807.

[11] S. K. Shandilya, N. Singhai, A survey on: Content based image retrieval systems, *International Journal of Computer Applications*, vol. 4, no. 2, pp. 22–26.

[12] R. Biederman, R. Mezzanotte, J.Rabinowitz. Scene perception: detecting and judging objects undergoing relational violations. *Cognitve Psychology*, 14(2):143–77, 1982.

[13] Y. Cao, C. Wang, Z. Li, L. Zhang, and L. Zhang. Spatialbag-of-features. CVPR, 2010.

[14] K. Murphy, A. Torralba, D. Eaton, W. Freeman, Object detection and location using local and global features, *Lecture Notes in Computer Science* 4170 (2006).

[15] C. Palm, Color texture classification by integrative co-occurrence matrices, *Pattern Recognition* 37 (5) (2004) 965–976.

[16] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, *IEEE Transactions on Systems,Man, and Cybernetics 3* (1973) 610–621.

[17] Y. Liu, H. Hel-Or, C. S. Kaplan, L. V. Gool. Computational Symmetry in Computer Vision and Graphics, *Foundations and Trends in Computer Graphics and Vision*, vol. 5, no. 1-2, pp. 1-195, 2010.

[18] J. M. Abello, P. M. Pardalos, Handbook of massive data sets, *Springer, 2002.*

[19] M. Choi, J. Lim, A. Torralba, and A.Willsky. Exploiting hierarchical context on a large database of object categories. In CVPR, 2010.

[20] D. Klimešová, E. Ocelíková, Study on Context Understanding, Knowledge Transformation and Decision Support Systems. WSEAS Transactions on Information Science and Applications, 7 (7) 2010, pp. 985-994.

[21] D. Klimešová, E. Ocelíková, Spatial-temporal modeling and visualisation, NOUN International Journal of Mathematical Models and Methods in Applied Sciences , 2012, 6 (1) , pp. 149-156.

[22] A. Prest, C. Leistner, J. Civera, C. Schmid, V. Ferrari. Learning object class detectors from weakly annotated video. In CVPR, 2012.

[23] F. Yang, K. Z. Mao. Robust Feature Selection for Microarray Data Based on Multicriterion Fusion. *IEEE Transactions on Computational Biology and Bioinformatics*, vol. 8, no. 4, pp. 1080-1092, 2011.

[24] Y. Gong and S. Lazebnik. Iterative quantization: A procrustean approach to learning binary codes. In CVPR, 2011.

[25] S. Borza, C. Simion, I. Bandrea, GIS application with geospatial database for improving the waste management in Sibiu surrounding area, NOUN International Journal of Energy and Environment , Issue 5, Volume 5, 2011, pp. 653-660.

[26] Ch. Chalkias, A. Faka , GIS-based spatiotemporal analysis of the exposure to direct sun light on rural highways, *WSEAS Transactions on Information Science and Applications*, Issue 1, Volume 7, January 2010.