Discovering Spatial Associations in Images

Osmar R. Zaïane[a] and Jiawei Han[b]

[a]Department of Computing Science, University of Alberta, Edmonton, AB, Canada
[b]School of Computing Science, Simon Fraser University, Burnaby, BC, Canada

ABSTRACT

In this paper, our focus in data mining is concerned with the discovery of spatial associations within images. Our work concentrates on the problem of finding associations between visual content in large image databases. Discovering association rules has been the focus of many studies in the last few years. However, for multimedia data such as images or video frames, the algorithms proposed in the literature are not sufficient since they miss relevant frequent item-sets due to the peculiarity of visual data, like repetion of features, resolution levels, etc. We present in this paper an approach for mining spatial relationships from large visual data repositories. The approach proceeds in three steps: feature localization, spatial relationship abstraction, and spatial association discovery. The mining process considers the issue of scalability and contemplates various feature localization abstactions at different resolution levels.

Keywords: Multimedia Mining, Association Rules, Resolution Refinement, Spatial Relationships

1. INTRODUCTION AND MOTIVATION

Knowledge discovery from data, or data mining, is a very popular and active research field. Many approaches and algorithms have been proposed for descriptive as well as predictive data mining.[1,2] However, most of the research is typically applied on alphanumeric data, whether on transactional flat files or on structured relational databases. Very little in comparison has been done for mining multimedia data such as image and or video collections, although, visual information is collected in many applications, such as medical imaging, satellite picture analysis, weather understanding, geological and astronomical examination, etc. Mining these large collections of images can yield interesting knowledge that can improve these applications and probably help the understanding of some of these studied phenomena. While successful data mining applications involving visual data have been reported in the early days of the data mining field, such as labeling volcanoes in pictures from the surface of Venus[3] , there is a new trend in data mining towards non traditional data such as images and sound[4–6] . However, these applications often necessitate new domain specific approaches. The conventional data mining algorithms are not sufficient for the acquisition of knowledge from multimedia data. Current database mining technologies are still not capable of extracting knowledge from images and videos, although some researchers are starting to investigate how to determine interesting patterns in multimedia. For example,[4] describes the CONQUEST system that combines satellite data with geophysical data to discover patterns in global climate change. The SKICAT system described in[5] integrates techniques for image processing and data classification in order to identify "sky objects" (i.e., patterns) captured in a very large satellite picture set. Tucakov and Ng in[7] used a method for outlier detection to identify suspicious behaviour from videos taken by surveillance cameras.

The motivation for our recent research is inspired by the existence of large collections of medical images, such as Positron Emission Tomography (PET scans) to produce internal images of physiological activity of the human body, and Magnetic Resonance Imaging (MRI) to produce internal images of soft tissue of the human body, and the potential of multimedia data mining in medical applications. Medical images are usually accompanied by medical records containing diagnoses etc. The possible correlations between the content of the medical records and visual features in the medical images for a large collection of images is very appealing. For instance, we want to study the spatial relationship between lesioned structures in brain scans and pathological characteristics found in the patient records. A similar association has been studied in[8] where the automatic discovery of associations between structures and functions of the human brain has been examined. In our study, in addition to the presence of lesions, we want to take into account their relative positions and spatial relationships.

Mining from visual data can either solely focus on knowledge discovery from the image collection, or combine the mining from images with mining of traditional alphanumeric data. After integrating image processing with

Correspondence: E-mail: zaiane@cs.ualberta.ca

database mining techniques, we have implemented a multimedia data mining system prototype, MultiMediaMiner[9], which uses a data cube structure for mining characteristic, association, and classification rules. The early version uses descriptors of images but does not exploit image content. In[10], in addition to the presence of colours and textures in images, we have also used localization of visual features, their spatial relationships, their motion in time (for video), etc. to discover interesting patterns within and between images. In this paper we discuss an algorithm that combines the visual data extracted from images and conventional data associated with the images to discover interesting associations between spatial relationships and alphanumeric data related to individual images. We also discuss various feature localization abstractions at different resolution levels, and present an algorithm for counting occurrences of objects in image transactions.

The remainder of the paper is organized as follows. In Section 2, we discuss our three-step approach for mining spatial relationships in images and introduce the feature localization step. Section 3 introduces the concept of multimedia association rules and discusses the spatial relationship abstraction, the second step of our algorithm. A method for mining association rules with spatial relationships is presented in Section 4. Finally, in Section 5, we briefly describe our performance analysis and conclude our study.

2. STARTING WITH FEATURE LOCALIZATION

In this study, we investigate efficient methods for mining associations in image databases between visual features in the images, as well as associations between visual features and descriptors relevant to the images. Finding association rules in databases is one of the typical and most studied data mining problems. The classic example of association rules is usually given in a retail transactional database in which the mining problem consists of discovering rules that state that, based on all available transactions, if a customer purchases a certain product A, that same customer in the same transaction might also purchase another product B with a given certitude. In order to discover associations in images, we want to express images in transactional form with visual features as well as descriptors attached to the images (e.g. medical diagnosis attached to medical images). However, it is necessary to first extract the visual features. Given the large collections of images, this visual feature extraction must be done automatically.

To discover spatial associations in images we propose a three-step approach as defined in Figure 1. The first step segments images by feature localization and retains only the interesting regions. This is the extraction of distinctive areas in the image based on colours and textures. The second step consists of identifying spatial relationships between the extracted areas in the images, such as disjoint, inside, contains, equal, meets, covered by, covers and overlaps as defined in[11,12], and generates transactions representing the images. The final step processes the transactions to discover association rules in the images.

Instead of image segmentation such as the segmentation of lesions in MRI brain images for surgery planning presented in[13], we have selected to segment images by feature localization. Feature localization is a new concept of rough image segmentation introduced in[14]. Image segmentation is a process which segments an image into disjoint regions. A region consists of a set of pixels that share certain properties, e.g., similar colour (or gray-level intensity), similar texture, etc. The traditional segmentation algorithms assume (1) regions are mostly *connected*; (2) regions are *disjoint* ($R_i \cap R_j = \emptyset$, for $i \neq j$); and (3) segmentation is *complete* in that any pixel will be assigned to some region, and the union of all regions is the entire image ($\cup_{k=1}^{m} R_k = I$). Although regions do not have to be connected, most available region-based and/or edge-based segmentation methods would yield connected regions, and it is error-prone to merge some of them into non-connected regions. Such a segmentation algorithm applied on an MRI of a human tissue may yield more disjoint regions than necessary. A web of vessels in a region with an abnormal texture may create a 2-dimensional separator that can be interpreted as borders between different segments, unless some really effective algorithm can identify these areas as belonging to a non-connected region. This is like a letter 'B' on a white label. The two white spots inside the letter 'B' are indeed part of the white label and not two separate segments. The above simple example indicates that the traditional image segmentation does not yield useful grouping and representation for object recognition or lesion detection. A more useful and attainable process is feature localization that identifies features by their locality and proximity.

As defined in[14,12], a *locale* $\mathcal{L}_x$ is a local enclosure (or locality) of feature x. $\mathcal{L}_x$ has an envelope L_x which is a set of tiles to represent the locality of $\mathcal{L}_x$, and some geometric parameters: *mass* $M(\mathcal{L}_x)$, *centroid* $C(\mathcal{L}_x)$, *variance* $\sigma^2(\mathcal{L}_x)$, and shape parameters for the locale, etc. A tile is a square area in an image. Its size is arbitrarily chosen as 16×16, but could be bigger or smaller such as 2×2 for a finer resolution. The tile is the building-unit for envelopes. A tile is 'red' if a sufficient number of pixels within the tile are red. It follows that a tile can be both 'red' and

```
INPUT: Set of images with associated alphanumeric descriptors
OUTPUT: Association rules with spatial relationships

BEGIN
  /* Step 1 Feature Localization */
  For each image do
    Extract locales by image tiling and envelope Growing based on colours and textures
    Keep set of dominant locales
  Endfor

  /* Step 2 Spatial Relationship Abstraction */
  For each image do
    For each pairs of locales do
      Compare centroid coordinates and locale envelopes to determine spatial relationships
      (disjoint, inside, contains, equal, meets, covered by, covers, overlaps)
    Endfor
    Keep set of spatial relationships
    Create transaction with image associated alphanumeric descriptors, set of spatial
    relationships and visual features.
  Endfor

  /* Step 3 Spatial Association Discovery */
  Apply association rule mining with recurrent items on transaction set.

END
```

Figure 1. A three-step algorithm for discovering associations with spatial relationships in images.

'blue' if some of its pixels are red and some are blue. While a pixel is the unit for image segmentation, a tile is the unit for feature localization. Thus, feature localization is a kind of rough segmentation where overlap is possible and completeness is not necessary. The notion of overlap is important because it allows us to express the idea of containment and overlap. Because a tile can have more than one value of a visual feature (i.e. more than one colour, more than one edge orientation, or more than one edge density), it is possible to have a locale containing another locale, or a locale intersecting another locale. This is not possible with traditional segmentation.

Tiles, if they are geometrically close, are grouped into an envelope. The tiles inside the envelope form the locale with the envelope growing by pyramidal linking procedure presented in[14] . $M(\mathcal{L}_x)$ is the number of pixels in L_x that actually have feature x. $M(\mathcal{L}_x)$ is usually less than the area of L_x, although it could be equal to it. $C(\mathcal{L}_x)$ is simply the centroid of the mass. $\sigma^2(\mathcal{L}_x)$ is the variance of the Cartesian distance from pixels in L_x to the centroid, and it measures the eccentricity of $\mathcal{L}_x$. Note, M, C, σ^2, etc. are measured in unit of pixels, not in tiles. This guarantees the granularity. Hence the feature localization is not merely a low-resolution variation of image segmentation.

The first step of our mining process is completed by selecting dominant locales from each image in the database. Dominant locales are locales that have a mass $M(\mathcal{L}_x)$ larger that a given threshold. In other words, they are large enough to be considered interesting. Common locales, locales that have a large support in the database are also ignored. These locales, identified as being similar based on their centroid coodinates and their envelope, appear in most images and are concidered normal phenomenon, thus ignored in the mining process.

Figure 2 shows an example of feature localization in an MRI brain image. Figure 2(a) is the original MRI. Figure 2(b) is an example of some locales retained from the MRI. Some locales smaller than a given threshold or common locales are eliminated. Figure 2(c) shows examples of locales that are not retained, such as the eyes and the nasal cavity.

3. SPATIAL RELATIONSHIP ABSTRACTION

Spatial relationships are essential components in query languages for geographic information systems and spatial databases, and describe topology of areas or regions in maps. In[11,15] Max Egenhofer presents a formal derivation for eight spatial relationships namely disjoint, inside, contains, equals, meets, covered by, covers, and overlap. The first row in the tables of Figure 6 illustrates these relationships. The relationships are formulated for areas based

(a) (b) (c)

Figure 2. Lesion localisation on an MRI with a 2×2 pixel tile locales: (a) original MRI. (b) Locales to consider. (c) Examples of locales to be ignored.

on intersections of the boundary of an area A denoted ∂A, the interior of the area denoted A° , and the exterior of the area denoted A^-. It is proven by using the intersections of all combinations of ∂A, A^- and A° that the eight spatial relationships named above suffice to describe all topological combinations. However, since we are mining homogeneous images (i.e. images from the same application domain), and we can impose a direction to the images, we can add lower level concepts of spatial relationships. Indeed, an MRI of a human brain, for example, can easily be oriented (i.e. face towards the top). This allows us to enhance the set of spatial relationships with new, more specific, relationships subsumed by the previously enumerated spatial relationships. For instance, the spatial relationship disjoint, given some threshold, subsumes relationships such as far-from and next-to, which in turn subsumes relationships such as vertical-next-to and horizontal-next-to. The same concept of subsumption of directional spatial relationships applies to most of the eight high-level spatial relationships. For example, meets encompasses right-meets, left-meets, top-meets and bottom-meets, with right and left belonging to the horizontal direction, and top and bottom belonging to the vertical direction. Covers and covered-by also have the notion of direction right, left, top and bottom. These new spatial relationships give more details about the topology of the regions obtained.

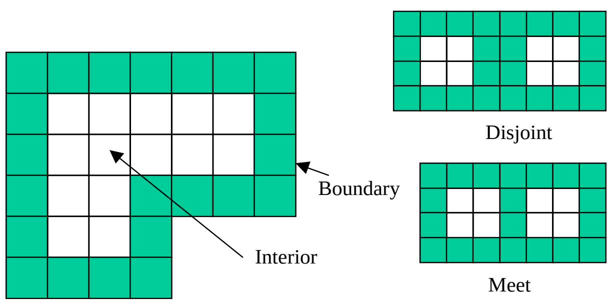

Figure 3. Locale envelope with boundary tiles.

Figure 3 shows a locale with boundary and interior. The intersection between boundaries and interiors of locales is based on shared tiles. For two locales A and B, ∂A is intersecting ∂B if there exists a tile belonging to the boundary of A and the boundary of B. A° is intersecting B° if there exists a tile belonging at the same time to the interior of A and the interior of B. The same applies to $\partial A \cap B^\circ$ and $A^\circ \cap \partial B$. Neighbouring tiles are not intersecting (See Figure 3).

To extract the spatial relationships from the images, once the relevant locales are determined, we only iterate through the set of locales associated to each image and compare for each pair of locales the coordinates of their centroids as well as their boundaries. The coordinates of the centroides reveal their horizontal-proximity or vertical-proximity, while the boundaries hint at the overlap, containment, closeness, etc. To improve the efficiency of this phase of the algorithms we define a minimum bounding circle around a locale to approximate the locale when evaluating topological relationships. With the bounding circle model, locales are first roughly estimated by a circle that comprises the totality of the locale. A minimum bounding circle is the smallest circle that could contain the

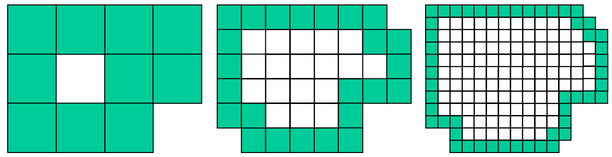

Figure 4. Progressive tile shrinking.

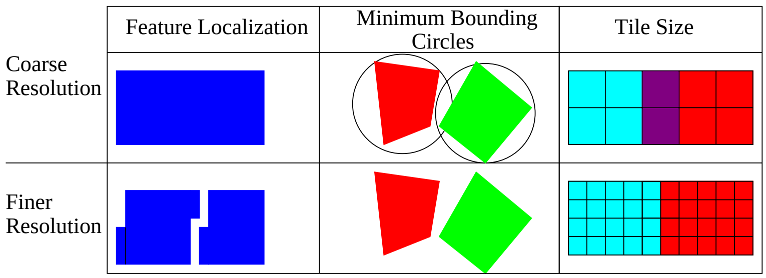

Figure 5. Relativity of visual feature and topology concepts at different resolution levels.

whole locale. The centroid of the locale is taken as the centre of the circle and the longest distance across the locale is the diameter of the minimum bounding circle. While there could be many different minimum bounding rectangles for a polygon, there is only one unique minimum bounding circle. The definition of the spatial relationships is then based on the perimeters formed by the minimum bounding circles and may differ from the original definition using the locale tiles like in Figure 3. Disjoint and meet are solely based on the boundary of the minimum bounding circles. The approximation of a locale by a minimum bounding circle can speed up the process, however, it is prone to errors if it is not alleviated in a second stage. Figure 5 illustrates an example depicting the relativity of some spatial relationships, like overlap, based on the resolution used for defining locales. While two locales may appear overlapping because their minimum bounding circles intersect, considered at the locale envelope level, they do not. Notice also that depending upon the size of the tiles in the locale resolution, different spatial relationships can be obtained with the same lesions or abnormal structures retained in the feature localization step. Figure 4 illustrates the progressive refinement in the case of shrinking tiles. The roughest resolution level uses a 8×8 tile size. Each finer resolution level divides the tile size by four. Since the child tiles of peripheral tiles may lose colour, the boundary of a locale may "retreat" inward which may result in a topological change from rough resolution to finer. The diagram in the left of Figure 5 shows an example where locales may appear to meet but with a smaller tile size they are disjoint. This clearly indicates that a revision stage is necessary to verify the knowledge discovered from one resolution level to the other. For instance, the minimum bounding circles are removed from the images relevant to selected spatial associations discovered, and the topology amended. The approximation with minimum bounding circles substantially improves the performance of the mining process, and the revision of the spatial relationships is done only for a subset of the association rules discovered in the final stage.

By eliminating the minimum bounding circles or reducing the tile size, the preservation or change in topological features is well defined and limited. The tables in Figure 6 graphically summarize the possible topological changes when minimum bounding circles are eliminated for resolution refinement, and the possible topological changes when the tile size is reduced for resolution refinement. The preservation and changes in topological features presented in Figure 6 are exact and complete. The proofs are given in[12] . In a progressive resolution refinement approach for mining association rules from images, these tables can be used to guide the filtering of candidate frequent sets during the mining, and avoid eliminating occurrences of topologies which may be relevant to the final association rules to discover[10] .

After extracting the localized visual features (i.e. locales) and determining the spatial relationships, each locale in an image is given an Object ID and spatial relationships between Object IDs are stored in a table with the Images ID where they occur, along with the locale characteristcs such as the centroid, size, colour, structure, etc. The information related to an image, such as the visual features automatically extracted (i.e. presence of colours and textures), the locales, as well as the alphanumeric information (i.e. medical diagnosis for the case of MRI), form a table of image descriptors. Both tables are denormalized to form a transactional database where each transaction

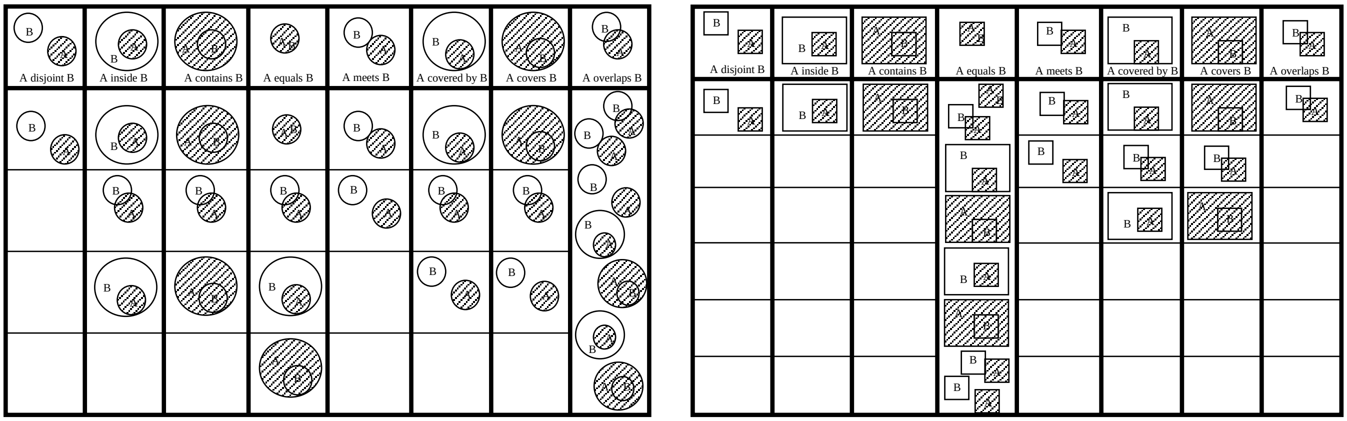

Figure 6. Topology and resolution increase with minimum bounding circles (left) and with tile size shrinking (right).

represents an image and its content. This collection of transactions is used in the next step to discover associations.

4. SPATIAL ASSOCIATION DISCOVERY

Association rules have been extensively studied in the literature. The efficient discovery of such rules has been a major focus in the data mining research community. Many algorithms and approaches have been proposed to deal with the discovery of different types of association rules discovered from a variety of databases[16–23] . The problem of discovering association rules is to find relationships between the existence of an object (or characteristic) and the existence of other objects (or characteristics) in a large repetitive collection. Such a repetitive collection can be a set of transactions for example, also known as the market basket. Typically, association rules are found from sets of transactions, each transaction being a different assortment of items. Association rules would give the probability that some items appear with others based on the processed transactions. Essentially, the problem consists of finding items that frequently appear together. However, in the context of images, the previous definitions and usage of association rules are restrictive. For instance, binary association rules and quantitative association rules are concerned with the presence of an item in a transaction or the presence in the transaction of a value defined in an interval. Images indeed contain these characteristics, like the presence of a texture or a particular spatial relationship, and the presence of a colour quantized in a given colour interval or space. However, visual data has other peculiarities proper to images and videos. For example, some visual features can be repeated in an image, and the repetition of the feature can carry more information than the existence of the feature itself. Thus, items in the antecedent of the rule repeating in the consequent can be an interesting factor in image analysis applications. For example, in an MRI, the existence of a localized abnormal texture may suggest the existence of another identical abnormal texture. Moreover, recurrent objects or features in images are very frequent. Thus, the count of occurrences of the same visual feature is a relevant attribute. Counting identical visual features occurring in the same image allows the discovery of spatial associations between these identical objects, for example, two red locales horizontally close to each other.

4.1. Discovering association rules with recurrrent items.

A method for enumerating sufficiently strong multimedia association rules that are based on recurrent atomic visual features was presented in[10] . The idea is to apply the apriori algorithm[18] but by replacing the filtering step and the support calculation. In addition, the notion of maximum support is introduced to eliminate items that are too common in the image collection. The algorithm (Algorithm 4.1) retains the information about the maximum occurrence of a visual feature in an image and use the information to form candidate sets.

ALGORITHM 4.1. (**MaxOccur**) Find sufficiently frequent item-sets for enumerating content-based multimedia association rules in image collections.

Input: (i) $\mathcal{D}_1$ a set of transactions representing images, with items being the visual and non-visual descriptors of the images; (ii) the minimum and maximum support thresholds σ' and Σ' for each conceptual level.

Output: Sufficiently frequent item-sets with repetitions allowed.

Method: The pseudo-code for generating sufficiently frequent item-sets is as follows:

begin
(1) $C_1 \leftarrow \{$Candidate 1 item-sets and their *support*$\}$
(2) $F_1 \leftarrow \{Sufficiently\ frequent$ 1 item-sets and their *support*$\}$
(3) $M \leftarrow \{Maximum\ occurrence$ in an image of $frequent$ 1 item-sets$\}$
(4) Count # of k-item-sets (total$[1..k]$)
(5) for $(i \leftarrow 2; F_{i-1} \neq \emptyset; i \leftarrow i+1)$ $do\{$
(6) $\quad C_i \leftarrow (F_{i-1} \bowtie F_{i-1}) \cup \{y \oplus X \mid X \in F_{i-1} \wedge y \in F_1 \wedge Count(y, X) < (M[y]-1)\}$
(7) $\quad C_i \leftarrow C_i - \{c \mid (i-1)$ item-set of $c \notin F_{i-1}\}$
(8) $\quad \mathcal{D}_i \leftarrow FilterTable(\mathcal{D}_{i-1},\ F_{i-1})$
(9) $\quad$ foreach image I in $\mathcal{D}_i$ do {
(10) $\quad\quad$ foreach c in C_i do {
(11) $\quad\quad\quad c.support \leftarrow c.support\ +\ Count(c,\ I)$
(12) $\quad\quad$ }
(13) $\quad$ }
(14) $\quad F_i \leftarrow \{c \in C_i \mid \frac{c.support}{total\ i\ itemset} > \sigma\prime\}$
(15) }
(16) Result $\leftarrow \bigcup_i \{c \in F_i \mid i > 1 \wedge c.support < \Sigma\prime\}$
end $\square$

Lines 1, 2, 3 and 4 are done doing the same initial scan. M contains the maximum number of times an object appears in the same image. This counter is used later to generate potential k-item-sets. The total number of k-item-sets is used for the calculation of the item-set support in line 14.

In lines 6 and 7, the candidate item-sets are generated by joining (i-1) frequent item-sets and the use of M to generate repetitive objects ($M[y] > 1$). The pruning process (line 7) eliminates infrequent item-sets based on the *apriori* property.

Line 8 filters the transactions in $\mathcal{D}$ to minimize the data set scanning time.

In line 14, only the frequent item-sets that are higher than the minimum support $\sigma\prime$ are kept. It is only at the end of the loop (line 16) that maximum support $\Sigma\prime$ is used to eliminate item-sets that appear too frequently.

4.2. Counting occurrences.

The calculation of the support for one item-set is based on the occurrence of the item-set in the images. Line 11 of Algorithm 4.1 cumulates this count. A particular precaution has to be taken when counting appearances of k-item-sets in an image, especially that objects and features can be repeated. A simple k-permutation ($\mathcal{C}_n^k = \frac{n!}{n!(n-k)!}$ where $n = \mid t \mid$) can lead to miscalculations. For example, let the transaction t be composed of repeated four objects such as $t = \{\diamondsuit\spadesuit\spadesuit\spadesuit\heartsuit\heartsuit\clubsuit\clubsuit\clubsuit\clubsuit\}$. $\mathcal{C}_{10}^2 = 45$ while we have only 9 possible unique 2-item-sets as shown below. There are also 14 possible 3-item-sets while $\mathcal{C}_{10}^3 = 240$.

{♢}	1
{♠}	3
{♡}	2
{♣}	4

{♢♠}	1	{♢♡}	1
{♢♣}	1	{♠♡}	2
{♠♣}	3	{♡♣}	2
{♠♠}	2	{♣♣}	2
{♡♡}	2		

{♢♠♡}	1	{♢♠♣}	1
{♢♡♣}	1	{♠♡♣}	2
{♠♠♠}	1	{♣♣♣}	1
{♢♠♠}	1	{♢♡♡}	1
{♢♣♣}	1	{♠♠♡}	1
{♠♠♣}	1	{♠♣♣}	1
{♡♣♣}	2	{♡♡♣}	2

Possible one, two and three item-sets and their occurrences in t.

The correct calculation of the repetitions of these item-sets in the transaction requires caution in order not to calculate occurrences more than necessary. The algorithm for enumerating the k-item-sets and counting their occurrences in the images transaction is given in Algorithm 4.2.

ALGORITHM 4.2. Counting occurrences of k-item-sets in an image transaction.

Input: (i) Image transaction $\mathcal{I}$; (ii) item-set size k.

Output: Set of k-item-sets and number of times they appear in $\mathcal{I}$.

Method. Generate all combinations from the unique objects in $\mathcal{I}$ and verify if they can be replicated (*Combination and Replication*); Generate item-sets with k times the same objects (*Twinning*); Generate item-sets with combinations of repeated objects (*Combination of twinned objects*). The pseudo-code for generating and counting the item-sets is as follows:

```
begin
    U ← {unique 1-item-sets and their count in 𝓘}
    C ← {k-combinations of u in U}
    foreach c in C do {   /* counting combinations and replications */
        c.count ← 1
        do CountReplication(c)
    }
    V ← U
    foreach u in V do {   /* Twinning */
        while V[u].count > k do {
            c ← ⊗_k u /* repeating u k times */
            V[u].count ← V[u].count − k
            Add c in C if not in set; c.count ← c.count + 1
        }
    }
    foreach u in U do {   /* Combination of twinned objects */
        for(n = 2; n < k − 1 ∧ n ≤ U[u].count; n + +) do {
            d ← ⊗_n u /* repeating u n times */
            B ← {k-combinations of d and d′ | v ∈ d′ ∧ v ≠ u ∧ U[v].count > 0}
            foreach c in B do {
                c.count ← 1
                Add c to C
                do CountReplication(c)
            }
        }
    }
    Result ← C
end
begin CountReplication(c)
        V ← U
        foreach 1-item-set i in c do {V[i].count ← V[i].count − 1}
        while V[j].count > 1(∀j in c) do {
            c.count ← c.count + 1
            foreach 1-item-set i in c do {V[i].count ← V[i].count − 1}
        }
end
```

□

4.3. Adding spatial relationships in the picture.

Algorithm 4.1 discovers sufficiently frequent itemsets from transactions that exclusively use visual atomic features. Spatial relationships are binary relationships between two visual features (i.e.locales). For a spatial relationship to be frequent, it has to be between two frequent locales. In other words, a frequent spatial relationship is associated with a frequent pair of frequent locales in the images database. A method for finding sufficiently frequent spatial relationships consists of using the set of frequent pairs using Algorithm 4.1, and associating these pairs with all potential spatial predicates (i.e. horizontal-next-to, top-meets, etc.). The frequent spatial relationships formed this way are used as input with Algorithm 4.1 to discover association rules with topology information. The association rules per se are generated from the frequent itemsets[18] .

5. CONCLUSION

We have presented in this paper a method for discovering spatial associations in image collections. The method consists of three steps for extracting visual features from the images, for determining spatial relationships, and finally for discovering association rules. The same technique is appropriate for a variety of application domains such as medical imaging, solar surface activity understanding, global weather analysis like tornado movements, etc.

The advantage of our approach is that it does not need any domain knowledge to perform the mining of spatial association from the images. The only knowledge necessary in order to choose a common orientation for all images and take advantage of lower lever spatial relationships is a simple template for systematic common direction. This is not indispensable since we can still use the high-level spatial relationships. However, orienting all images the same way brings a new meaning to the association rules discovered. In other words, they are easier to interpret. Another benefit of our approach is the possibility to apply both first steps as a pre-processing phase while the images are received. The mining step can then become an interactive process where the user refines the knowledge discovered interactively or adjusts and restricts the meta-rules describing the type of knowledge to be discovered[20] . In addition, a progressive resolution refinement approach[10] can be used to iterate along the three steps to discover knowledge from visual media at different resolution levels.

Many improvements could still be forthcoming, such as the addition of shape enumeration from images and the categorization of locales based on some visual features such colour (i.e. gray scale), texture and position. The shape can be a important attribute in some application domains such as medical imaging. Categorizing locales is a step towards object recognition. This could increase the meta-rule expressive power and significantly reduce the number of association rules discovered, and thus, help the user pin-point interesting and relevant rules.

Our main goal is to apply the algorithm on a huge collection of MRI brain scans with real patient records. However, accessing large collections of real MRI scans is a difficult problem due to privacy issues. Communication is ongoing with the University Hospital at the University of Alberta to access a large compilation of medical images for our experiments. Experimenting on a small set of images is not significant enough to test the scalability of our approach. We have generated large sets of synthetic images with simulated visual features and have tested our algorithms on the data set to demonstrate the scalability of the algorithms. Although our preliminary experiments were done on synthetic images, the results were very promissing and the algorithms scaled linearly. The association rules are discovered at different resolution levels for the visual features (i.e. locales) as well as spatial relationships at different conceptual levels.

ACKNOWLEDGMENTS

The research is supported in part by the Natural Sciences and Engineering Research Council of Canada, and the Canadian Networks of Centres of Excellence IRIS-3.

REFERENCES

1. G. Piatetsky-Shapiro and W. J. Frawley, *Knowledge Discovery in Databases*, AAAI/MIT Press, 1991.
2. U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. U. (eds.), *Advances in Knowledge Discovery and Data Mining*, AAAI/MIT Press, 1996.
3. J. Smyth, U. Fayyad, M. Burl, and P. Perona, "Modeling subjective uncertainty in image annotation," in *Advances in Knowledge Discovery and Data Mining*, U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, eds., pp. 517–539, AAAI/MIT Press, 1996.
4. P. Stolorz, H. Nakamura, E. Mesrobian, R. Muntz, E. Shek, J. Santos, J. Yi, K. Ng, S. Chien, C. Mechoso, and J. Farrara, "Fast spatio-temporal data mining of large geophysical datasets," in *Proc. First Int. Conf. On Knowledge Discovery and Data Mining*, pp. 300–305, August 1995.
5. U. M. Fayyad, S. G. Djorgovski, and N. Weir, "Automating the analysis and cataloging of sky surveys," in *Advances in Knowledge Discovery and Data Mining*, U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, eds., pp. 471–493, AAAI/MIT Press, 1996.
6. A. Czyzewski, "Mining knowledge in noisy audio data," in *Proc. Second Int. Conf. on Knowledge Discovery and Data Mining*, pp. 220–225, 1996.

7. V. Tucakov and R. Ng, "Identifying unusual spatio-temporal trajectories from surveillance videos," in *Proc. of 1998 SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery (DMKD'98)*, (Seattle, Washington), June 1998.
8. K. Megalooikonomou, C. Davatzikos, and E. Herskovits, "Mining lesion-deficit associations in a brain image database," in *Proc. Int. Conf. Knowledge Discovery and Data Mining (KDD'99)*, pp. 347–351, (San Diego, California), August 1999.
9. O. R. Zaïane, J. Han, Z.-N. Li, and J. Hou, "Mining multimedia data," in *CASCON'98: Meeting of Minds*, (Toronto, Canada), November 1998.
10. O. R. Zaïane, J. Han, and H. Zhu, "Mining recurrent items in multimedia with progressive resolution refinement," in *Proc. 2000 Int. Conf. Data Engineering (ICDE'2000)*, February 2000.
11. M. J. Egenhofer, *Spatial Query Languages.* PhD thesis, University of Maine, 1989.
12. O. R. Zaïane, *Resource and Knowledge Discovery from the Internet and Multimedia Repositories.* PhD thesis, School of Computing Science, Simon Fraser University, March 1999.
13. Y. Tao, W. Grosky, L. Zamorano, Z. Jiang, and J. Gong, "Segmentation and representation of lesions in the MRI brain images," in *Proc. SPIE Medical Imaging*, (San Diego, CA), 1999.
14. Z.-N. Li, O. R. Zaïane, and Z. Tauber, "Illumination invariance and object model in content-based image and video retrieval," *Journal of Visual Communication and Image Representation* **10**(3), pp. 219–244, 1999.
15. M. J. Egenhofer and J. Sharma, "Topological relations between regions in r^2 and z^2," in *Advances in Spatial Databases (SSD'93)*, (Singapore), 1993.
16. R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large databases," in *Proc. 1993 ACM-SIGMOD Int. Conf. Management of Data*, pp. 207–216, (Washington, D.C.), May 1993.
17. H. Mannila, H. Toivonen, and A. I. Verkamo, "Efficient algorithms for discovering association rules," in *Proc. AAAI'94 Workshop Knowledge Discovery in Databases (KDD'94)*, pp. 181–192, (Seattle, WA), July 1994.
18. R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proc. 1994 Int. Conf. Very Large Data Bases*, pp. 487–499, (Santiago, Chile), September 1994.
19. J. Han and Y. Fu, "Discovery of multiple-level association rules from large databases," in *Proc. 1995 Int. Conf. Very Large Data Bases*, pp. 420–431, (Zurich, Switzerland), Sept. 1995.
20. Y. Fu and J. Han, "Meta-rule-guided mining of association rules in relational databases," in *Proc. 1st Int. Workshop Integration of Knowledge Discovery with Deductive and Object-Oriented Databases (KDOOD'95)*, pp. 39–46, (Singapore), Dec. 1995.
21. S. Brin, R. Motwani, and C. Silverstein, "Beyond market basket: Generalizing association rules to correlations," in *Proc. 1997 ACM-SIGMOD Int. Conf. Management of Data*, pp. 265–276, (Tucson, Arizona), May 1997.
22. R. Ng, L. V. S. Lakshmanan, J. Han, and A. Pang, "Exploratory mining and pruning optimizations of constrained associations rules," in *Proc. 1998 ACM-SIGMOD Int. Conf. Management of Data*, pp. 13–24, (Seattle, Washington), June 1998.
23. S. Ramaswamy, S. Mahajan, and A. Silberschatz, "On the discovery of interesting patterns in association rules," in *Proc. 1998 Int. Conf. Very Large Data Bases*, pp. 368–379, (New York, NY), August 1998.