Mining Image Datasets Using Perceptual Association Rules

Jelena Tešić, Shawn Newsam and B. S. Manjunath *
Electrical and Computer Engineering Department
University of California, Santa Barbara
Santa Barbara, CA 93106-9560
{*jelena, snewsam, manj*}*@ece.ucsb.edu*

Abstract

This paper describes a framework for applying traditional data mining techniques to the non-traditional domain of image datasets for the purpose of knowledge discovery. In particular, *perceptual association rules*, a novel extension of traditional association rules, are used to distill the frequent perceptual events in large image datasets in order to discover interesting patterns. The focus is on spatial associations although the method is equally applicable to associations within or between other dimensions; i.e., spectral, or in the case of video, temporal. A primary contribution is the derivation of image equivalents for the traditional association rule components, namely the items, the itemsets, and the rules. The proposed approach is modular, consisting of three steps that can be individually adapted to a particular application. First, the image dataset is labeled in a perceptually meaningful way using a visual thesaurus. Second, the first- and second-order associations are tabulated in a scalable data structure termed a spatial event cube. Finally, the higher-order associations and rules are determined using an adaptation of the Apriori algorithm. The proposed approach is applied to an aerial video dataset to demonstrate the kinds of knowledge perceptual association rules can help discover.

1 Introduction

Multimedia data is being acquired at an increasing rate due to technological advances in sensors, computing power, and storage. The value of these sizable datasets extends beyond what can be realized by traditional "focused" computer vision solutions, such as face detection, object tracking, etc. Instead, new methods of analysis based on data mining techniques are required to discover the implicit patterns, relationships and other knowledge that is not readily observable. Such knowledge is useful for a variety of applications, ranging from data summarization and visualization in scientific experimentation, to query refinement in multimedia data management systems.

Data mining techniques have been used for some time to discover implicit knowledge in transaction databases. In particular, methods are available for determining the interesting associations among itemsets over large numbers of transactions, such as the products that are most frequently purchased together in market basket analysis. Achieving similar success with multimedia datasets remains a challenge, however, not only due to the size and complexity of image and video data, but also the lack of image equivalents for the association rule components, namely the items, the itemsets, and even the rules. It is not straightforward to define, let alone detect, the items and itemsets appropriate for discovering the implicit spatial knowledge contained in large collections of aerial images. The main contribution of this work is a framework for applying a specific set of traditional data mining techniques to the non-traditional domain of image datasets. In particular, *perceptual association rules* are proposed as a novel, multimedia extension to traditional association rules.

1.1 Motivation

The objectives for applying association rule algorithms to traditional transaction databases are clear. A primary objective of market basket analysis is to determine optimal product placement on store shelves. However, the objectives for mining association rules in multimedia datasets are less obvious at this early stage in research on perceptual data mining when the limit of what is technically feasible are not known. Ideally, the rules would provide insight into the prominent trends in the dataset, such as interesting but non-obvious spatial or temporal causalities.

*This research was supported in part by the following grants/awards: The Institute of Scientific Computing Research (ISCR) award under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under contract No.W-7405-ENG-48, Department of Transportation grant to the NCRST infrastructure at UCSB, ONR# N00014-01-1-0391, NSF Instrumentation #EIA-9986057, and NSF Infrastructure #EIA-0080134.

A spatial association rule derived from remote sensed imagery might help discover that two particular crops have a higher yield when planted near to each other. A strong motivation for the research presented in this paper is to investigate the *kinds of knowledge* perceptual association rules can help discover. This, in turn, will allow data mining practitioners to work with domain experts in identifying objectives that are both interesting and feasible.

1.2 Overview of the Proposed Approach

An association rule [1] is an expression of the form $A \Rightarrow B$ meaning the presence of itemset A implies the presence of itemset B. An association rule algorithm discovers the rules that have support and confidence larger than a specified threshold. The bottom-up approach proposed by this work transforms the raw image data into a form suitable for such analysis in three steps. First, image regions are labelled as perceptual synonyms using a visual thesaurus that is constructed by applying supervised and unsupervised machine-learning techniques to low-level image features. The region labels are analogous to items in transaction databases. Second, the first- and second-order associations among regions with respect to a particular spatial predicate are tabulated using spatial event cubes (SECs). The SEC entries are analogous to first- and second-order itemsets. Finally, higher-order associations and rules are determined using an adaptation of the Apriori association rule algorithm. These modular steps can be individually tailored, making the framework applicable to a variety of problems and domains

The rest of the paper is organized as follows. Section 2 presents related work and Section 3 provides a general description of association rules and outlines a widely-used algorithm for discovering them. The proposed approach is described in Section 4 and experimental results for an aerial video dataset are presented in Section 5. Section 6 concludes with a discussion.

2 Related Work

Several approaches to applying association rules to image datasets have been proposed. Ordonez and Omiecinksi [2] use segmentation results from the Blobworld system [3] to mine the co-occurrence of image regions that have been labeled as similar using an empirically determined distance measure and threshold. The segmented regions are viewed as items and the images are viewed as transactions so that the resulting rules are of the form, "The presence of regions A and B imply the presence of region C with support X and confidence Y". It is not clear, however, that the results from applying the technique to a dataset of synthetic images composed of basic colored shapes would generalize to real images for which segmentation and notions of region similarity present a significant challenge.

Ding et al. [4] extract association rules from remote sensed imagery by considering set ranges of the spectral bands to be items and the pixels to be transactions. They also consider auxiliary information at each pixel location, such as crop yield, to derive association rules of the form "Band 1 in the range $[a, b]$ and band 2 in the range $[c, d]$ results in crop yield Z with support X and confidence Y." However, analysis at the pixel scale is susceptible to noise, unlikely to scale with dataset size, and limited in its ability to discover anything other than unrealistically localized associations-i.e., in reality, what occurs at one pixel location is unlikely to be independent of nearby locations.

3 Association Rules

This section provides a general description of association rules and outlines a widely used algorithm to discover them. Association rule approach was first introduced in [1] as a way of discovering interesting patterns in transactional databases. An association rule tells us about the association between two or more items.

Let $U = \{u_1, ...u_N\}$ be set of items. A set A is a $K-itemset$, if $A \subseteq U$ and $|A| = K$. An association rule is an expression $A \Rightarrow B$, where A and B are itemsets that satisfy $A \cap B = \emptyset$. Lets D as a superset, i.e. $D = \{T|T \subseteq U\}$. Elements of database D are called *transactions*. Transaction $T \subseteq D$ supports an itemset A if $A \subseteq T$. Support of itemset A over all database transactions T is defined as:

$$supp(A) = \frac{|\{T \in D|A \subseteq T\}|}{|D|} \tag{3.1}$$

Apriori algorithm discovers combination of items that occur together with greater frequency than might be expected if the values or items were independent. The algorithm selects the most "interesting" rules based on their support and confidence. Rule $A \Rightarrow B$ expresses that whenever a transaction T contains A, it probably contains B also.

Support measures statistical significance of a rule:

$$supp(A \Rightarrow B) = \frac{|\{T \in D|A \subseteq T \wedge B \subseteq T\}|}{|D|} \tag{3.2}$$

Confidence is a measure of a strength of a rule:

$$conf(A \Rightarrow B) = \frac{|\{T \in D|A \subseteq T \wedge B \subseteq T\}|}{|\{T \in D|A \subseteq T\}|} \tag{3.3}$$

The probability of rule confidence is defined as conditional probability $p(B \subseteq T|A \subseteq T)$. Association rules can be between more than 2 items, i.e. $A, B \Rightarrow C$ where $A,\ B,\ C \subseteq U$. Association rule is strong if its confidence is larger than user's specified minimum support.

Several improvements have been proposed for mining association rules [5]. They deal with complexity of rule mining and separating interesting rules from the generated rule set in a more efficient way.

The user has to specify minimum support of frequent itemsets. Every subset of a frequent itemset is also frequent. An itemset can be frequent only if it is frequent in at least one of these partitions and algorithm takes advantage of that property. The Apriori algorithm identifies frequent itemsets as following:

Algorithm 1 Apriori Algorithm

1. Find frequent item sets;
$F_1 = \{u_i | \; \|u_i\| > \text{minimum support}\}$
for $(K = 2;\; F_{K-1} \neq \emptyset;\; K++)$ **do**
  $C_K = \{c_k | c^{(a)} \wedge c^{(b)} \in F_{K-1}\}$, where:
  $c_k = (u_{i_1}, ..., u_{i_{k-2}}, u_{i_{k-1}}, u_{i_k})$
  $c^{(a)} = (u_{i_1}, ..., u_{i_{k-2}}, u_{i_{k-1}})$
  $c^{(b)} = (u_{i_1}, ..., u_{i_{k-2}}, u_{i_k})$
  $\|c_k\| = 0$;
  for $(\forall T, T \subseteq D)$ and $(\forall c_k, c_k \in C_K)$ **do**
    if $(c_k \in T)$ **then**
      $\|c_k\| = \|c_k\| + 1$;
    end if
  end for
  $F_K = \{c_k | \|c_k\| > \text{minimum support}\}$
end for
$F = \bigcup_K F_K$
2. Use the frequent itemsets to generate strong association rules.

4 Perceptual Association Rules

This section describes the three steps of the proposed approach: 1) perceptual labeling of the image regions using a visual thesaurus; 2) tabulation of first- and second-order associations using SECs; and 3) mining of higher-order associations and rules.

4.1 Perceptual Labeling using a Visual Thesaurus

A visual thesaurus is used to label the image regions in a perceptually meaningful way. The visual thesaurus is constructed in two stages using the low-level region features and a manually labeled training set. First, the dimensionality of the feature space is reduced by feeding the training set into a self-organizing map (SOM), a clustering technique that is known to preserve the topology of the input space. The resulting clusters are assigned class labels using a majority-vote rule and the SOM is used to classify the entire dataset. The second stage fine-tunes these classes using an improved Learning Vector Quantization (LVQ) algorithm. The resulting sub-classes group the region features into the perceptual synonyms of the visual thesaurus. The image dataset is then labeled by assigning each region the codeword of its entry in the thesaurus. More details on the visual thesaurus and its construction can be found in [6].

4.2 Spatial Event Cubes

The visual thesaurus is used to label the image regions based solely on their distribution in the feature space. Knowledge of the spatial arrangement of the regions is incorporated through SECs, a scalable data structure that tabulates the region pairs that satisfy a given binary spatial predicate, see [7].

Define the raster space R for an image partitioned into $M \times Ntiles$ as $\mathbb{R} = \{(x, y) | \; x \in [1, M],\; y \in [1, N]\}$. Let the set of T of thesaurus entries u_i be: $T = \{u_i | u_i$ is a thesaurus entry/codeword$\}$. Let τ be a function that maps image coordinates to thesaurus entries, $\tau(P) = u$, where $P \in \mathbb{R}$ and $u \in T$; and let ρ be the given binary spatial predicate, $P\rho Q \in \{0, 1\}$, where $P, Q \in \mathbb{R}$. Then, an SEC face is the co-occurrence matrix $C_\rho(u, v)$ of all pairs of points that have codeword labels u and v, and satisfy the binary predicate ρ:

$$C_\rho(u, v) = \|(P, Q)| \; (P\rho Q) \wedge (\tau(P) = u) \wedge (\tau(Q) = v)\|$$

These co-occurrences are computed over all the images in the dataset. Note that it is the relation ρ that determines the particular spatial arrangement tabulated by the SEC. The choice of ρ is application dependent and can include spatial relationships such as adjacency, orientation, and distance, or combinations thereof.

4.3 Perceptual Association

An attribute value set T contains N thesaurus entries u_i. Spatial Event Cube face entries $C_\rho(u, v)$ mark frequency of codeword tuples that satisfy binary relation ρ. Define F_K^ρ as a set of frequent itemsets of size K. Also define S_ρ as a minimum support value for frequency of $C_\rho(u_{i_1}, ..., u_{i_K})$. Our goal is to find sets of tuples that show some dependency among tile spatial configurations. $F^\rho = \bigcup_K F_K^\rho$. Since the Spatial Event Cube captures spatial relationship only between two tiles, we need to implement some clever processing to build higher order candidates. Spatial relationship for $K--$itemsets depends on the characteristics of binary function ρ.

Perceptual Association Rule Algorithm supports atomic pattern set approach to candidate itemset generation in multimedia datasets [8]. We test occurrences of interesting patterns in a dataset, thus avoiding the formulation of transactions. A set of atomic patterns for our representative dataset D is $F_1 = \{u_i | \; \|u_i\| > S_\rho^{(1)}\}$. For F_2 we build the conjunction of two atomic patterns

(u_i, u_j) and look for the corresponding SEC entries. If $C_\rho(u_i, u_j) > S_\rho^{(2)}$, then $(u_i, u_j) \in F_2$. Spatial Event Cubes are built on a binary relationship ρ. For higher order candidate sets we are imposing a clever processing rule i.e. ordering. Then, we go back to representative dataset D and record the occurrences of new candidates, as explained in the previous paragraph. Only the ones with support larger that the user specified minimum $S_\rho^{(K)}$ qualify for the frequent itemset of size k. Note that the clever processing rule can differ for a different itemset size.

If we have more elements in an itemset, there are more ways to spatially organize those elements. Association rules can be between 3 items is in the form $u_i, u_j \Rightarrow u_k$ where $u_i, u_j, u_k \subseteq D$. If $C_\rho(u_i, u_j, u_k)/C_\rho(u_i, u_j)$ is larger than minimum confidence required, the rule $u_i, u_j \Rightarrow u_k$ is a valid rule. For the right neighbor example, this rule could be formulated as "If codeword u_j is right neighbor of codeword u_i, that might imply that u_k is on the right side of u_j". Extended association rule algorithm for spatial relationship ρ is:

Algorithm 2 Perceptual Association Rule

1. Find frequent item sets
$F_1^\rho = \{u_i|\ \|u_i\| > S_\rho^{(1)}\}$;
$F_2^\rho = \{(u_i, u_j)|C_\rho(u_i, u_j) > S_\rho^{(2)}\}$;
for $(K = 3;\ F_{K-1} \neq \emptyset;\ K++)$ **do**
  Candidate K-item frequent itemset C_K is formed of K
    joint elements from any frequent F_{K-1}^ρ item set;
  $C_K = \{c_k|c^{(a)}, c^{(b)} \in F_{K-1}\}$, where:
  $c_k = (u_{i_1}, ..., u_{i_{k-2}}, u_{i_{k-1}}, u_{i_k})$
  $c^{(a)} = (u_{i_1}, ..., u_{i_{k-2}}, u_{i_{k-1}})$
  $c^{(b)} = (u_{i_1}, ..., u_{i_{k-2}}, u_{i_k})$
  for $(\forall k, c_k \in C_K)$ and $(\forall i, \sigma_i \in R^K)$ **do**
    $\|c_k\| = \|\{\sigma_i|\sigma_i \text{ satisfies spatial relationship}\}\|$
    AND $(\tau(P_1) = u_1) \wedge ... \wedge (\tau(P_K) = u_K)$,
    where $\sigma_i = (P_1, ..., P_K), (P_1, ..., P_K) \in R^K$
  end for
  $F_K = \{c_k|\|c_k\| > S_\rho^{(K)}\}$
end for
$F^\rho = \bigcup_K F_K^\rho$
2. Use the frequent itemsets to generate association rules;

5 Experimental Results

The proposed technique is applied to a collection of aerial videos of Amazonia made available by the Institute for Computational Earth System Science (ICESS) at UCSB. Aerial videography is an affordable alternative to expensive high-resolution aerial and satellite imagery. It is particularly attractive for areas plagued by cloud cover, such as the Amazon, since the aircraft are flown at low altitudes. The sample dataset was captured using a high-end commercial video camera and is geo-referenced. The aerial videos are temporally subsampled to create a sequence of just-overlapping frames, which are treated as a collection of images. One hour of video results in approximately 450 frames of size 720 by 480 pixels. Figure 3(a) shows a sample frame. The results presented here are for one hour of the 40-hour dataset. The proposed method can alternately be applied to image mosaics created from the video although care must be taken that the mosaicking process does not introduce unwanted artifacts.

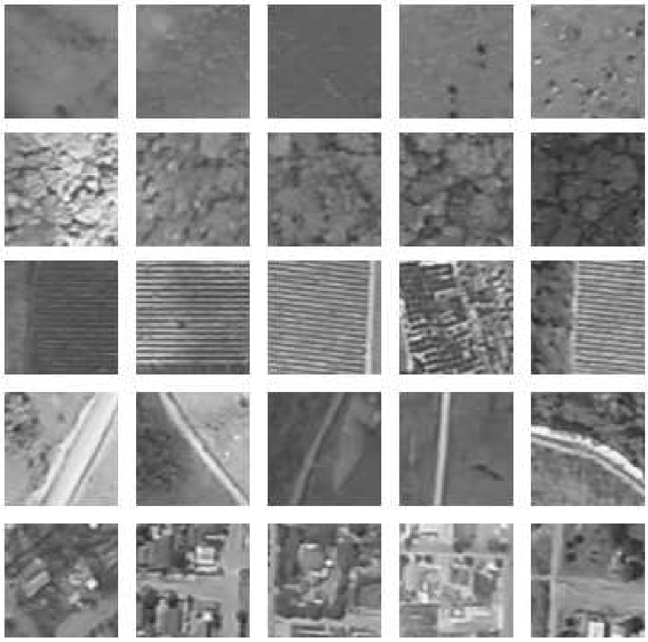
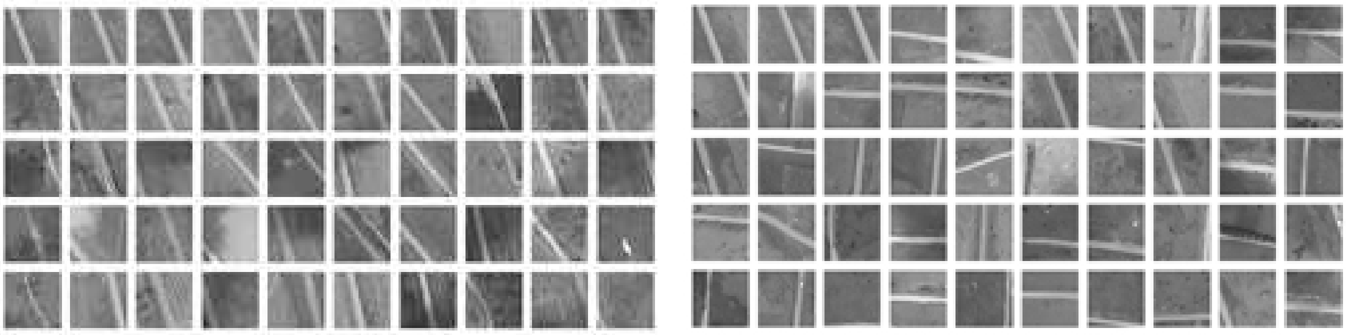


Figure 1: Training set samples. From top to bottom row: pasture, forest, agriculture, road, and urban

5.1 Homogeneous Texture Descriptor

The raw image data is perceptually characterized using a low-level descriptor based on homogeneous texture. The Vision Research Lab at UCSB has extensive experience with using texture to analyze remote sensed imagery. In particular, an MPEG-7 [9] compliant descriptor based on the outputs of orientation– and scale–selective Gabor filters [6] has shown to effectively characterize a variety of land cover types, such as agriculture, forest, and urban development. Since the texture descriptor captures the spatial distribution of relative pixel values, it is less sensitive to changes in lighting conditions than spectral descriptors. This turns out to be a significant advantage for analyzing the aerial



(a) Rotation sensitive retrieval

(b) Rotation invariant retrieval

Figure 2: Similarity retrieval based on a distance function

Class	Description	Training Set Size
0	Pasture	285
1	Forest	238
2	Agricultural	116
3	Road	185
4	Urban	116

Table 1: Training set for 5 class manual labeling. Total Number of Training Tiles is 940

videos of Amazonia, which contain a mixture of sunny and cloud-shaded regions.

The texture descriptors are extracted in an automated fashion by dividing the video frames into non-overlapping 64 by 64 pixel tiles and applying Gabor filters tuned to combinations of six orientations and five scales. The first– and second–order moments of the filter outputs form the texture feature vector. Thus, one hour of video results in approximately 35,000 tiles each characterized by a 60-dimension vector. The visual similarity between tiles is computed by defining a distance function on the high-dimensional feature space. Euclidean distance results in an *orientation– and scale–selective* similarity measure.

Invariant similarity is measured using a distance function that exploits the structure of the feature vectors [10]. Figure 2 shows examples of both orientation-selective and orientation-invariant similarity retrieval.The visual thesaurus is constructed using the invariant similarity measure since the image regions occur at arbitrary orientations in the aerial videos.

5.2 Visual Thesaurus

The visual thesaurus is used to label the video frame tiles in a perceptually meaningful way by clustering the high-dimensional texture feature vectors using supervised and unsupervised learning techniques. The training set required by the supervised learning stage is manually chosen with the help of domain experts. The primary land cover types are identified as pasture, forest, water, agricultural, and urban.

Water is not considered separately since it does not occur often and is similar to clear pasture with respect to the texture. Roads occur frequently and are distinct so the final basic land cover types are pasture, forest, agricultural, road, and urban. Table 5.1 lists the size of the five classes in the training set. Note that the training set measures less than three percent of the entire dataset.

Figure 1 shows sample tiles from the training set. The LVQ stage of the training algorithm results in a visual thesaurus with 73 codewords. These codewords can be considered as fine tuned subclasses of the manually chosen training classes. Figures 3(b) and 3(c) show the results of using the visual thesaurus to label the frame in Figure 3(a). Figure 3(b) shows the class assignments, which are mostly correct. Figure 3(c) shows the codeword labels, which are subclasses of the training set labels. For example, codeword labels 0 through 19 correspond to subsets of the pasture class, and labels 49 through 63 correspond to subsets of the road class.

This demonstrates a key feature of the visual thesaurus: the final codeword labeling represents a finer and, therefore, more consistent partitioning of the high dimension feature space than the manually chosen training set.

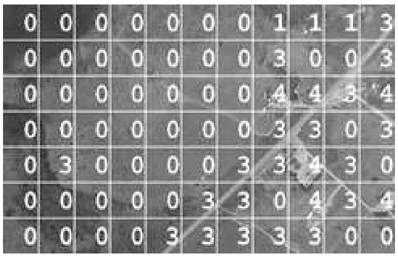
5.3 Spatial Event Cubes

An SEC is computed from the codeword labeled frames using 8-neighbor adjacency as the spatial predicate. The diagonal entries of the SEC indicate the number of times a codeword appears next to itself.

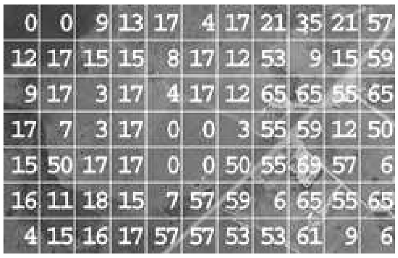
Thus, the largest diagonal values correspond to the homogeneous regions of the dataset. Table 2 lists the diagonal entries with the largest values. These entries correspond to pasture and forest codewords.

A second SEC is computed this time using a direc-

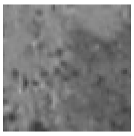
(a) Extracted frame No. 238



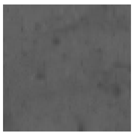
(b) Labeled frame - 5 classes



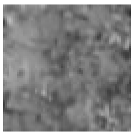
(c) Labeled frame - 73 codewords



(d) 21

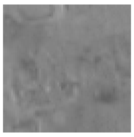


(e) 0

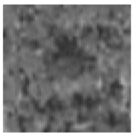


(f) 35

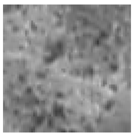
(g) 12



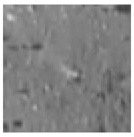
(h) 17



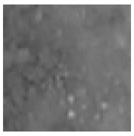
(i) 33



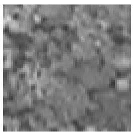
(j) 24



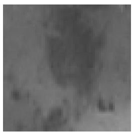
(k) 41



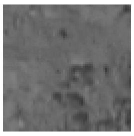
(l) 4



(m) 38



(n) 9



(o) 13

Figure 3: Thesaurus codewords for the most frequent elements of F_2^ρ

tional spatial adjacency predicate ρ:

$$(x_1, y_1)\rho(x_2, y_2) \Leftrightarrow (x_1 = x_2) \wedge (y_1 + 1 = y_2),$$

where x_i and y_i are the horizontal and vertical coordinates of tile i. This predicate allows spatial analysis along the direction of flight of the aerial video. Table 4 lists the diagonal entries with the largest values. These entries correspond to pasture and forest codewords.

5.4 Itemsets and Rules

The support 3.2 and confidence 3.3 of constructed rule $u_i \Rightarrow u_j$ for 8-connectivity neighborhood can be derived from the SEC entries:

$$supp(u_i \Rightarrow u_j) = \frac{C_\rho(u_i,u_j)}{|D|}$$
$$conf(u_i \Rightarrow u_j) = \frac{C_\rho(u_i,u_j)}{2*|u_i|}$$

A factor of two is needed in the denominator in this case since ρ is symmetric.

Constructed associated rules from F_2^ρ are listed in tables 2 and 3, respectively. The corresponding confidence values are also indicated. Figure 3 shows the corresponding codewords to the most frequent elements of the second order item set F_2^ρ. Codewords in the range 0 through 19 correspond to pasture subclasses and codewords in the range 20 through 39 correspond to forest subclasses.

i	$\|u_i\|$	$C_\rho(u_i, u_i)$	$C_\rho(u_i, u_i)/\|u_i\|$
21	3621	2366	0.653411
0	2555	2273	0.889628
35	2330	1608	0.690129
12	2903	1566	0.539442
17	2081	1054	0.506487
33	1183	796	0.672866
24	884	508	0.575226
41	943	508	0.538706
4	1289	492	0.382079
38	728	450	0.618819

Table 2: Corresponding Frequencies of largest diagonal SEC entries for 8-connectivity neighborhood

i	j	$C_\rho(u_i, u_j)$	conf($u_i \Rightarrow u_j$)	conf($u_j \Rightarrow u_i$)
12	0	2461	0.423872	0.481605
17	12	1764	0.423835	0.303824
35	33	1095	0.234979	0.462806
17	4	939	0.225613	0.364236
12	9	876	0.150878	0.344611
35	24	781	0.167597	0.441742
38	24	601	0.412775	0.339932
21	13	571	0.078846	0.402113

Table 3: Confidence of generated rules from F_2^ρ set

While it is no surprise that forest and pasture are the most frequently occurring land types, table 3 indicates that specifically forest codeword 21 and pasture codeword 0 occur most frequently. Pasture codeword 13 is more likely to occur next to forest codeword 21 than vice versa.

For the second SEC, constructed associated rules from F_2^ρ are listed in tables 4 and 6, respectively. note that the confidence of constructed rule $u_i \Rightarrow u_j$ for "righthand neighborhood rule" is:

$$supp(u_i \Rightarrow u_j) = \frac{C_\rho(u_i,u_j)}{|D|}$$
$$conf(u_i \Rightarrow u_j) = \frac{C_\rho(u_i,u_j)}{|u_i|}$$

i	$\|u_i\|$	$C_\rho(u_i, u_i)$	$C_\rho(u_i, u_i)/\|u_i\|$
21	3621	1159	0.320077
0	2555	1093	0.427789
35	2330	677	0.290558
12	2903	1760	0.261798
17	2081	425	0.204229
33	1183	329	0.278107
24	884	191	0.216063
41	943	209	0.221633
4	1289	225	0.174554
38	728	167	0.229396

Table 4: Corresponding Frequencies of largest diagonal SEC entries for the "righthand neighbor" spatial rule

(12, 0, 0)	$conf((12, 0) \Rightarrow 0) = 0.23077$
(9, 12, 12)	$conf((9, 12) \Rightarrow 12) = 0.23737$
(21, 35, 35)	$conf((21, 35) \Rightarrow 35) = 0.27708$
(21, 17, 0)	$conf((12, 17) \Rightarrow 0) = 0.095$

Table 5: Confidence of generated rules from F_3^ρ set, for the "righthand neighbor" rule

Since these results are almost identical to the ones for 8-neighborhood adjacency, it can be concluded that the dataset is isotropic with respect to adjacency.

Mining rule for third order itemset is formulated as "If codeword u_j is right neighbor of codeword u_i, that implies that u_k is on the right side of u_j with confidence conf$((u_i, u_j) \Rightarrow u_k)$". We constrained the candidates in C_3^ρ to form a "righthand neighbor" chain. Constructed associated rules from F_3^ρ are listed in table 5.

6 Conclusion

In this paper, we introduce an association rule mining framework that supports spatial event discovery in large image datasets. Feature vectors are extracted from image tiles and summarized into visual thesaurus. Visual thesaurus allows us to record spatial relationships among labelled features using Spatial Event Cube (SEC) approach. SEC support perceptual association rule mining approach and provide efficient pruning for generating higher order candidate itemsets. We demonstrate the use and possible applications of proposed framework on a large collection of aerial videos of Amazonia and large collection aerial photos of Santa Barbara region. Future research include implementation of mining more complex spatial rules and a cost function that will prove the efficiency of the proposed method. We are also planning to use Spatial Event Cubes as an index structure for multimedia database access.

i	j	$C_\rho(u_i, u_j)$	conf$(u_i \Rightarrow u_j)$	conf$(u_j \Rightarrow u_i)$
0	12	521	0.203914	0.179470
12	0	507	0.174647	0.198434
21	35	480	0.132560	0.206009
35	21	442	0.189700	0.122066
17	12	346	0.166266	0.119187
12	17	335	0.115398	0.160980
35	33	227	0.097425	0.191885
33	35	208	0.175824	0.089270
12	9	202	0.069583	0.158930
9	12	198	0.155783	0.068205
17	4	185	0.088900	0.143522

Table 6: Confidence of generated rules from F_2^ρ set, for the "righthand neighbor" rule

7 Acknowledgments

The authors would like to thank Chandrika Kamath and Imola K. Fodor for many fruitful discussions.

References

[1] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proc. 20th Int'l Conf. Very Large Data Bases (VLDB)*, 1994.

[2] C. Ordonez and E. Omiecinski, "Discovering association rules based on image content," in *Proc. IEEE Advances in Digital Libraries Conference (ADL)*, 1999.

[3] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Region-based image querying," in *Proc. CVPR '97 CBAIVL Workshop*, 1997.

[4] Qin Ding, Qiang Ding, and William Perrizo, "Association rule mining on remotely sensed images using p-trees," in *Proc. PAKDD*, 2002.

[5] J. Han and Y. Fu, "Discovery of multiple-level association rules from large databases," in *Proc. 21st Int'l Conf. Very Large Data Bases (VLDB)*, 1995.

[6] W. Ma and B. S. Manjunath, "A texture thesaurus for browsing large aerial photographs," *Journal of the American Society of Information Science*, 1998.

[7] J. Tešić, S. Newsam, and B.S. Manjunath, "Scalable spatial event representation," in *Proc. IEEE Int. Conf. on Multimedia and Expo (ICME)*, 2002.

[8] D. J. Hand, H. Mannila, and P. Smyth, *Principles of DataMining*, MIT Press, Cambridge, MA, 2000.

[9] B.S.Manjunath, P. Salembier, and T. Sikora, Eds., *Introduction to MPEG7: Multimedia Content Description Interface*, John Wiley & Sons, 2002.

[10] S. Newsam and B.S.Manjunath, "Normalized texture motifs and their application to statistical object modeling," in *Submitted to ICCV*, 2003.