

Content Analysis using Shape and Spatial Layout with Markov Random Field

Riaz Ahmed Shaikh*, Jian-Ping Li, Asif Khan and Kamlesh Kumar

School of Computer Science & Engineering, University of Electronic Science and Technology of China, Chengdu - 611731, Sichuan, China; 201314060107@std.uestc.edu.cn, jpli2222@uestc.edu.cn, 201314060113@std.uestc.edu.cn, 201314060145@std.uestc.edu.cn

Abstract

Background/Objectives: Unstructured random scene perception and understanding is a challenging problem in the field of computer vision and image processing. **Methods/Statistical analysis:** Two images with very different appearance may have the same color, texture and shape features. To recognize two different pictures having similar color, texture and shape we can apply spatial investigation. **Findings:** A methodology in the light of Markov random field has been proposed in this work to recuperate the fundamental spatial layout from a solitary image and start to examine its use as a foundation for scene understanding and content analysis in content based image retrieval. Our representation comprises of three key components (1) coarsely depicting the orientation of significant scene surfaces, (2) an examination of low level features for the understanding of image, and (3) a shape obliged Markov random field definition that enforces shape priors over the regions. **Application/Improvements:** We experimentally assess different Markov random field formations and exhibit the adequacy of our proposed approach in scene understanding and content analysis.

Keywords: CBIR, Content Analysis, MRF, Shape, Spatial Layout

1. Introduction

When humans look at any image, they see not just a pattern of texture and color, but the world behind the image. Content Based Image Retrieval (CBIR) systems shows phenomenal execution at computing low level elements from pixel representation however its output does not mirror the overall need of the client. The system performs ineffectively in the extraction of high level features that incorporates object and their definitions, feelings and actions. This referred to as the semantic gap, has required current research in CBIR frameworks towards image retrieval by the type of object or scene delineated.

For instance, in Figure 1, a computer may be able to detect the road surface and some of the pedestrians, but it will have no idea that the pedestrians are walking on the road.

In image processing, image analysis is the important stage; it is the process of extracting significant information from digital images. Image analysis consists of some

tasks such as edge detection, shapes detection, objects counting, or measuring object properties. General image analysis algorithms comprise detection of edge, shape detectors, segmentation based on color, and image thresholding. By integrating these techniques with functions of region analysis, comprehensive statistics can be getting from digital images to give human analysis with additional quantitative and qualitative data.

Our objective is to get the rough spatial layout of a picture, an illustration of major surfaces from image and their connection to each other. The spatial areas are the key focuses in a picture convey critical data for arranging the picture. For instance, a picture demonstrating a shoreline scene ordinarily comprises of sky object focuses on the top and sands object focuses at the base.

1.1 Visual Content Level

Images are naturally endowed with attributes or information content that can help in resolving the image

*Author for correspondence

retrieval problem. The information content that can be derived from an image is classified into three levels, shown in Figure 2.

- Low level – They include visual features such as texture, colour, spatial layout, shape and motion.
- Middle Level – Examples includes presence or arrangements of specific types of objects, roles and scenes.
- High Level – Include impressions, emotions and meaning associated with the combination of perceptual features. Examples include objects or scenes with emotional or religious significance.

The image content level is also a measure of level of feature extraction. At the low level, also regarded as primary level the features extracted (color, shape, texture,



Figure 1. Scene understanding requires knowledge of spatial layout.

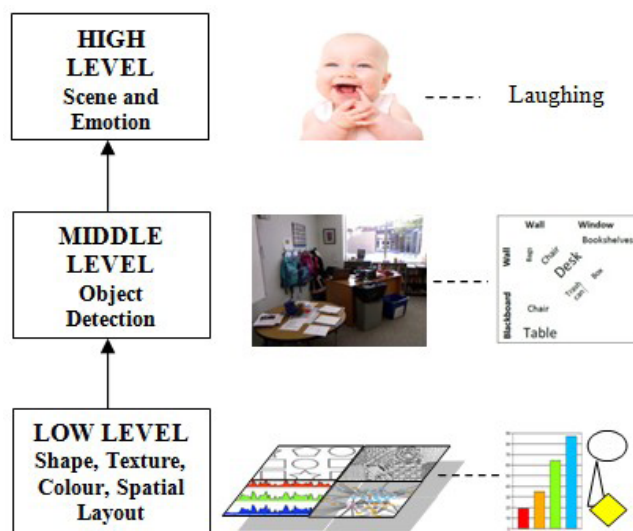


Figure 2. Image content levels.

spatial information and motion) are called primitive features because they can only be extracted by information obtained at the pixel level, that is pixel representation of the images. The middle level features are features that can be extracted by collection of pixels that make up the image, while high level features go beyond the collection of pixels. It identified the impressions meanings and emotions associated with the collection of pixels that make the object.

1.2 Spatial Layout

Spatial information is any information that can be geologically referenced by portraying an area or any information that can be connected to a location. Spatial data is the spatial relationship existing among properties portraying picture regions inside of the global image, and is normally used to addresses the issue of segregating comparative pictures in homogenous or non-assorted image repository. A case is appeared in Figure 3.

1.3 Shape

Shape is the most important and one of the primitive visual features of an image. Shape recognition is one of the modes through which human observation of the environment is executed. It is essential in CBIR systems in light of the fact that it compares to region of interest in digital images. Shape is the binary image consisting of contour

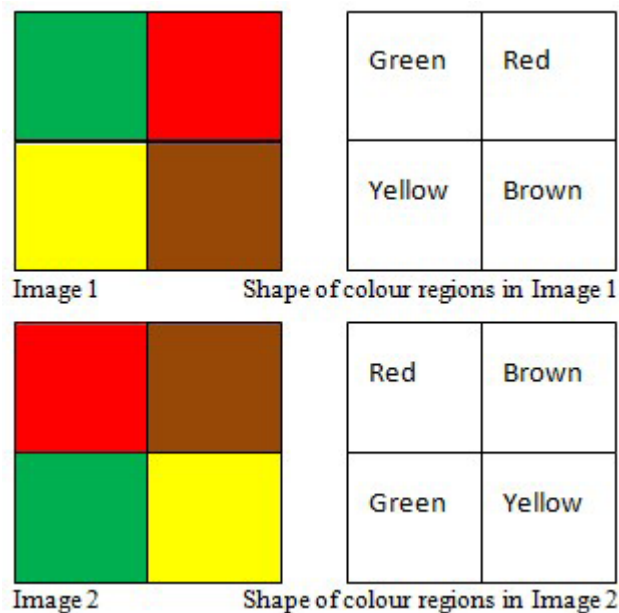


Figure 3. Demonstration of spatial information.

our outline of objects, obtained after segmentation. Shape feature representations are boundary based and region based as shown in Figure 4.

1.4 Markov Random Field (MRF)

A MRF, otherwise called a Markov system or an undirected graphical model has an arrangement of hubs $X = \{X_1, \dots, X_n\}$ and an arrangement of edges interfacing sets of hubs. Markov Random Field hypothesis gives an advantageous and steady method for demonstrating connection subordinate elements, for example, picture pixel and other spatially corresponded highlights. This is accomplished through characterizing common impact among such elements utilizing MRF probabilities.

On the lattice S , let $X = \{X_1, X_2, \dots, X_M\}$ be a family of arbitrary variables. In the discrete space $\Lambda = \{1, 2, \dots, K\}$, every random variable gets values. As a random field the family X is with configuration set $\Omega = \Lambda^M$. X as an arbitrary field is said to be a MRF on S concerning to a neighborhood framework $V(S)$ if:

$$\forall x \in \Omega, P(x) > 0$$

$$\forall S \in S, \forall x \in \Omega P(X_s = x_s | X_t = x_t, t \neq s) = P(X_s = x_s | X_t = x_t, t \in V_s(S))$$

In Panchal and Tiwari¹ discussed, analyzed and compared various techniques of image retrieval system based on contents. Some features extraction techniques also described for extraction primitive features like colour, shape and texture. Combinations of these three low level features are also described. Three approaches namely Genetic Algorithm (GA), an Energy Based Model (EBM), and a Binary Integer Programming (BIP), to utilizing spatial contextual information at object level for semantic

image analysis are discussed by Papadopoulos et al.² A framework based on a number of various combinations of primitive features and algorithms of classification was developed for the gain of good insight on the use of spatial context. Choudhary et al.³ described the integration of three methods that are Gabor Filter, Spatiogram, and Edge Histogram as a new method. This methodology fundamentally automates the process of retrieving by utilizing image analysis techniques taking into account primitive visual features for example shape, texture and colour with spatial information.

Detailed review of various methods has been focused by Priyatharshini et al.⁴ Evaluation techniques of such methods used in the current research based on spatial features of image in content based image retrieval systems. In⁵ Balasubramani and Kannan discussed about two powerful features of CBIR systems. 1. Edge Histogram Descriptor used for edge distribution 2. Spatial distribution of colours represents by Colour Layout Descriptor (CLD) in an image. More user friendly prototype NWCIBR has been developed and implemented both EHD and CLD features. Various concepts and applications of CBIR have been reviewed by Hirwane⁶. He also presented a technique on the basis of automatically derived primitive features for retrieval of images. Efficient and new algorithm for object tracking has been discussed by Naraghi et al.⁷ The developed algorithm has been analyzed, described and implemented in real time.

A method based on a probabilistic spatial context model, as well as individual material detection algorithm was proposed by Singhal et al.⁸ to determine the scene contents. Spatial context aware material detection method used to reduce misclassification and increase the accuracy of initial classification. Construction of the surface layout of objects, aging a label to the image into geometric classes is discussed by Hoiem et al.⁹ Multiple segmentation frameworks were used to provide robust spatial support to contribute to the confidence in each geometrical label. Yadav et al.¹⁰ presented a study of the behaviour of several content based image retrieval systems. Various feature extraction texture analysis with representation was also presented. Shaikh et al.¹¹ proposed a dual layer model SHODHANI, which integrates features and processes and fused neural network approach to sense a particular image and easily retrieved from complex database. Chakraborty et al.¹² firstly estimated camera view point and position of the peoples in 3D scene then extraction spatial information from 3D people position. Neural



Figure 4. Shape representations based on boundary and region.

network based a novel approach proposed by Khan et al.¹³ which contributes a novel concept for the understanding and recognition of the contents from image, which can be important to give vision intelligence to cloud robot. Object analysis and tracking is one of the most vital issues in computer vision. Altaf and Raeisi¹⁴ presented an efficient and effective algorithm based on support vector machine for object tracking in video images using texture and color features. Proposed framework is able to track object.

2. Problem Domain

Basic features like shape, colour and texture can be used to describe the simple objects in the scene for satisfactory results. It is still conceivable that two images with very different appearance may have the same colour, texture and shape data. To recognize two different pictures having similar texture, color and shape features and to describe the complex objects in the scene, spatial relation and/or position of an object help to detect the target object and content analysis. Accuracy, retrieval and efficiency are the critical issues in designing image retrieval systems based on contents. The shape and spatial features are very simple and easy to obtain and efficient. For reducing the semantic gap researchers are moving towards discovering spatial components and the extent of implementing these features into the image retrieval systems.

3. Methodology

A region in a digital image has a steady texture if a local statistics set or other local properties of the image method are consistent, gradually changing, or periodic. Texture is an evidently confusing thought. It is generally utilized as a part of the early processing of visual data, particularly for characterization purposes. On other side, nobody has succeeded in creating a common definition of texture so everyone can accept. The determination of this spatial situation, we feel, will rely on upon a MRF way, more advanced model for the processing of early visual information, a core part of which will be representational frameworks at various distinct levels of abstraction. Figure 5 and Figure 6 represents the content domain and spatial analysis ratio respectively.

These levels will most likely incorporate real intensities at the base and will progress through edge and descriptors of orientation to surface, and maybe volumetric

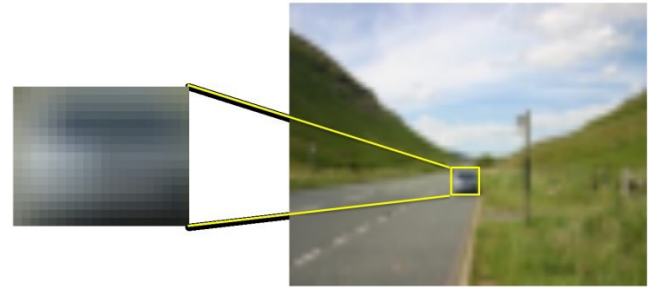


Figure 5. Content domain.

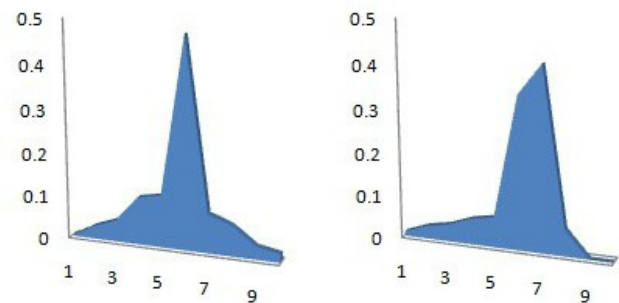


Figure 6. Spatial analysis ratio.

descriptors. Given these multi-level structures, it appears to be clear that they ought to be incorporated into the definition of, and in the computation of, item descriptors. The combined thoughts in utilizing MRFs for vision are the accompanying, images are analyzed into a gathering of hubs that may relate to pixel agglomerations of pixels, hidden variables connected with the nodes are brought into a model outline “clarify” the values such as colours of the considerable number of pixels, a joint probabilistic model is the direct statistical conditions between concealed variables are communicated by explain grouping hidden variables; most often these groups are pairs delineated as edges in a graph.

Figure 7 shows the image database, in order to apply the continuity criterion given for matching toward random image $I_1(X, Y)$ with shape condition from some spatial sub image features $I_2(X, Y)$ where

$$X, Y \in \{1, \dots, N\} \quad (1)$$

So common vision for sub image are

$$DX^a(X, Y) = X_b - X_a \quad (2)$$

and

$$DY^a(X, Y) = Y_b - Y_a \quad (3)$$

So that one of partial position ϕ of sub image where different fact line noise are closely compare require for object Z so, from where

$$\phi(X, Y) = \begin{cases} \{(X, Y) - Z_1\}^2, & \text{if } Z \text{ observe} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

So that random field for partial position ψ for first image are consider as

$$\psi(X_1 - X_2) \equiv \min(X_1 - X_2)^2 \cdot \phi \max \quad (5)$$

When spatial case as ideal consider for $\phi \max = 1$.

SIFT Flow forces stronger smoothness by MRF associations between adjacent pixels, visually providing more satisfying results Figure 8 demonstrate the precision of images returned. In actuality, DSP joins the qualities of the other two. While the task of matching objects in related with foreground or background matches, in the task of scene matching every pixel in an exemplar is commented with one of different class names. At that point, match pixels between the query image and the selected one. At

the point when measuring label exchange precision, just consider the match able pixels that relate to the common classes to both digital images.

4. Conclusions

Markov Random Field based an integrated approach combining shape and spatial layout for content analysis has been designed and implemented. Our method recovers the rough spatial layout from the image and it can use as fundamental for scene investigation and content analysis. Our representation comprises of three key components (1) coarsely depicting the introductions of significant scene surfaces, (2) an examination of low level features for the understanding of image, and (3) a shape obliged Markov random field definition that enforces shape priors over the regions. Experimental results show the effectiveness of our approach in content analysis.

5. Acknowledgment

This paper was supported by the National Natural Science Foundation of China (Grant No.61370073), the National High Technology Research and Development Program of China (Grant No.2007AA01Z423).

6. References

1. Panchal CK, Tiwari RA. A survey on CBIR using low level feature combination. International Journal of Emerging Technology and Advanced Engineering. 2014 Nov; 4(11):211–20.
2. Papadopoulos GT, Saathoff C, Escalante HJ, Mezaris V, Kompatsiaris I, Srintzis MG. A comparative study of object level spatial context techniques for semantic analysis. Computer Vision and Understanding. 2011 Sep; 115(9):1288–307.
3. Choudhary K, Pundlik M, Choukse D. An integrated approach for image retrieval based on content. International Journal of Computer Science. 2010 May; 7(3(7)):42–7.
4. Priyatharshini R, Surendernath SP, Chitrakala S. Applications of spatial features in CBIR: A survey. International Conference on Computer Science, Engineering and Applications; 2013. p. 303–11.
5. Balasubramani R, Kannan V. Efficient use of MPEG-7 colour layout and edge histogram descriptor in CBIR systems. Global Journal of Computer Science and Technology. 2009; 9(5):157–63.

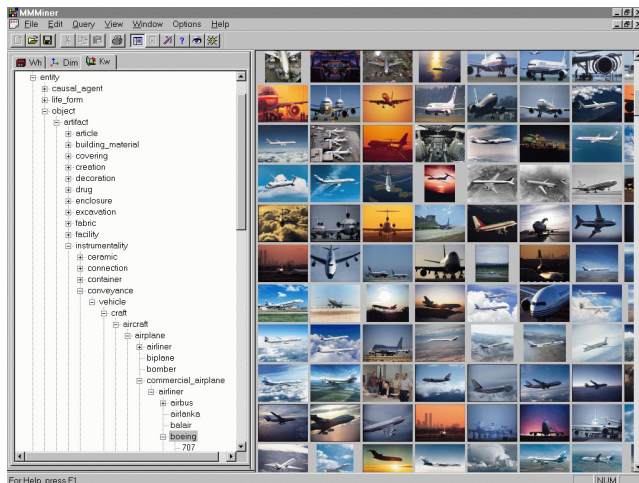


Figure 7. Dynamic database.

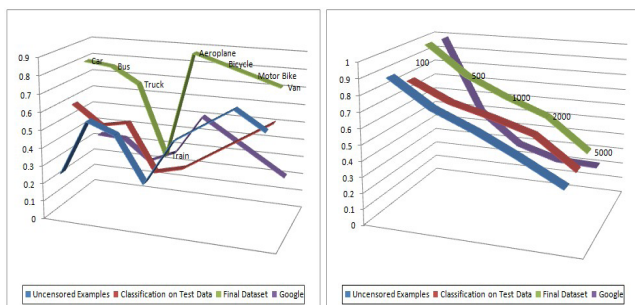


Figure 8. Precision of categories at 100 images returned.

6. Hirwani R. Fundamental of content based image retrieval. *International Journal of Computer Science and Technology*. 2012 Mar; 3(1):114–16.
7. Naraghi MG, Koohi M, Sarbazy H. Object tracking based on estimation mesh method. *Indian Journal of Science and Technology*. 2012 Jan; 5(1):1851–54.
8. Singhal A, Luo J, Zhu W. Probabilistic spatial context models for scene content understanding. *Computer Society Conference on Computer Vision and Pattern Recognition*; 2003 June 18-20. p. 235–41.
9. Hoiem D, Efros AA, Hebert M. Recovering surface layout from an image. *International Journal of Computer Vision*. 2007 Oct 16; 75(1):151–72.
10. Yadav AK, Vaishali RR, Kumar AP. Survey on content based image retrieval and texture analysis with applications. *International Journal of Signal Processing, Image Processing and Pattern Recognition*. 2014; 7(6):41–50.
11. Shaikh RA, Jian-Ping L, Khan A, Deep S, Kumar K, Memon I. Contemporary integration of content based image retrieval. *IEEE 11th International Computer Conference on Wavelet Active Media Technology and Information Processing*; 2014 Dec 19-21. p. 301–04.
12. Chakarborty I, Cheng H, Javed O. 3D Visual Proxemics: Recognizing human interactions in 3D from a single image. *IEEE Conference on Computer Vision and Pattern Recognition*; 2013 Jun 23-28. p. 3406–413.
13. Khan A, Jian-Ping L, Shaikh RA, Deep S, Kumar K, Hasan F. Vision prehension with CBIR for cloud robo. *IEEE 11th International Computer Conference on Wavelet Active Media Technology and Information Processing*; 2014 Dec 19-21. p. 293–96.
14. Altaf A, Raeisi A. Presenting an effective algorithm for tracking of moving object based on support vector machine. *Indian Journal of Science and Technology*. 2015 Aug; 8(17):1–7.