Automatic Image Structure Analysis

A.W Wardhani and R Gonzalez
School of Information Technology, Griffith University.
PMB 50 Gold Coast Mail Centre, QLD 4217 Australia
E-mail: {A.Wardhani,R.Gonzalez@eas.gu.edu.au}

ABSTRACT

The rapid growth of multimedia technology has resulted in an enormous amount of data that needs to be managed and indexed efficiently to provide effective labeling for an image indexing system requires all objects in the image to be identified. To perform better and effective object identification, the process needs to be performed automatically and without priori knowledge of the image content.

This paper presents an approach in automatic object identification scheme, by analysing the image structural information. The image is segmented using image automatic segmentation techniques and components of objects are obtained by grouping the segments together. In this paper, we present the issues and problems involved in providing such identification scheme. Some experiment results will be presented.

1. Introduction

In distributed environments, locating and retrieving images, is very demanding. Often, users want to find certain features or objects in a collection of images, but looking through all the data is tedious and inefficient. Since extensive work has been done to support retrieval of text documents, most current multimedia information systems are limited to indexing textual image annotations.

Recently, the attention has focused on the development of content-based indexing to support resource discovery for images. Statistical attributes such as colour and texture are used to generate descriptions of images and video. Various systems such as QBIC and the Virage image database system have performed indexing quite successfully using this scheme [1]. The drawback of these statistical methods is their inability to support object based access or queries, without relying on manual annotation or highly constrained semantic methods.

To provide object-based access to the content of the image requires a method to identify the objects contained in the image. Traditionally this identification problem has been approached using two main techniques: manual annotation [2] and object recognition [3]. Manual annotation has been the most common labeling method for supporting image content-based retrieval. Apart from being tedious, providing textual annotation to every image and video segment is limited to annotator's choice of vocabulary and context. This limits the scope of retrieval and is not suitable for high volume work.

Object recognition was seen as a possible technique to automatically generate labels to use for identifying the image content by extracting semantic descriptions. The main problem with object recognition is the need for the objects to be recognised, to be specified a priori. This limits the application of identification / labeling only to simple opaque polyhedra in highly constrained environments. Object recognition techniques also suffer from drawbacks such as computational expense.

In this paper, we present an approach for providing an automatic object identification scheme. This paper is organised as follows. Section 2 discusses automatic object identification schemes. Four issues will be described in this section: image segmentation, feature extraction, region grouping and image structure analysis. In section 3, the experiment procedure will be presented. Section 4 describes the experiment results. Section 5 discusses conclusion and future work.

2. Automatic Object Identification

With the drawbacks of existing techniques (object recognition and manual annotation), it is evident that we need to find a new scheme that allows the objects in an image to be identified automatically but without the requirement of interpreting the contents of the data.

The difficulty in automating the identification process is due to the complexity of information contained in images. One feasible identification method is to extract the components of the objects, followed by a grouping process that combines these together.

In the literature, the concept of object component analysis to obtain object descriptions has been suggested [4]. Rothwell [5] proposed the need for computing local features and combining them together in forming complete object descriptors. This method was suggested in order to identify objects in an image. Clemens and Jacobs [6] proposed that using components of an image and accompanied by the grouping of such components will provide significant changes to the efficiency of indexing systems. In this paper, we propose to use a similar approach to enable the object in an image be identified. This scheme is illustrated in Figure 1.

The system accepts an input image or representative image frame from a scene change detection scheme. The automatic object identification scheme consists of three main functions: feature extraction, image segmentation and region grouping.

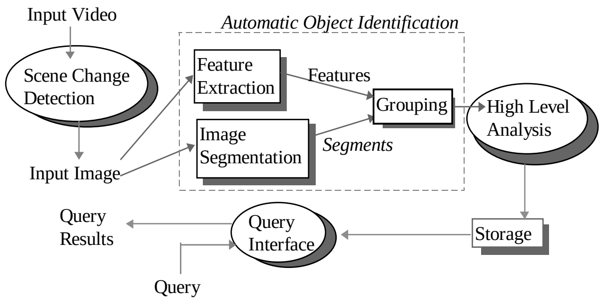


Fig.1, Proposed Object Identification Method

First, components of the object are obtained from the image using automatic image segmentation techniques. Features from the image also need to be extracted to measure the similarity or "togetherness" between the components or image segments. This is performed through the use of a feature extraction process. Grouping rules are then generated based on analysing the luminance distribution and the structure of the objects contained in the image. Using this feature information and the grouping rules, the segments are grouped together to form object descriptions. These descriptions will be stored as index information. Queries sent through the query interface will be compared and matched with the stored image indexes.

To summarise the discussion, the proposed automatic object identification scheme is divided into four topics: image segmentation, feature extraction, region grouping and image structure analysis.

2.1. Image Segmentation

Image segments can be defined as group of pixels which are statistically correlated. Techniques to identify these segments can be broadly categorized into four main groups: region based segmentation, histogram splitting based segmentation, texture based segmentation and motion based segmentation.

Region based segmentation produces segments by grouping some pixels according to some constraints such as spatial proximity and intensity similarity [7]. Some of the existing region-based segmentation techniques are region growing, split and merge and histogram splitting techniques. Region growing technique groups pixels or sub-regions into larger regions, by merging the initial pixel or region with its surrounding pixels or regions [8]. To use this region based segmentation technique, the initial pixels or regions need to be selected. Also, some similarity measures need to be chosen to guide the growing process. An alternative to region growing technique is to use region splitting and merging techniques. The image is subdivided initially into a set of arbitrary, disjointed regions. These regions are then merged and split to satisfy a given similarity predicate. This predicate is used achieve the resulting region uniformity. Some predicates commonly used are listed in [9]. A split and merge technique works iteratively and one commonly used representation is in the form of quadtree. The root of the tree corresponds to the entire image and each node corresponds to a subdivision.

For regions that do not have smooth homogenous luminance distribution, texture-based segmentation is used. This technique groups pixel which has consistent texture properties. Properties based on the mean and standard deviation of pixels in a region are used to quantify the texture of a region. Texture based segmentation is based on using measures of texture for some predicates. Any existing methods for region based segmentation can be used to perform texture based segmentation by specifying predicates based on the texture content.

Another technique that can be used is the histogram splitting technique. To use this technique, a histogram of the image pixel values is formed. The segments are formed by clustering areas in the histogram. This technique produces segments which have close values in the colour and representative colours are chosen for each segments.

For video data, the pixels in each frame can also be grouped based on the similarity in their motion characteristics, using motion-based segmentation techniques. Some existing motion based segmentation techniques are segmentation using a difference between two image frames and segmentation using optical flow characteristics.

The problem with using segmentation techniques alone for image indexing is that there are too many segments resulting from this process. These segments contain not only perceptually salient object components but also incidental segments that do not yield any natural descriptions. These arise due to reflection, shadows and uneven illumination. These undesired segments increase the complexity of the labeling process and provide no useful information.

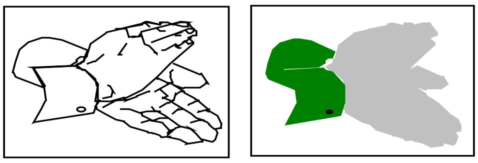
The aim of this work is to identify objects in an image. This initially performed by producing "good" or "useful" segmentation results non-recursively. Good segmentation means that the process will be directed so that the segmentation will generate significant regions. Hence, the results will provide some description about the objects in the image, therefore object identification can be performed automatically.

To achieve this goal, using statistical correlation alone in segmentation process is insufficient. Our view is that segmentation results using the existing methods can be used as initial regions. These need to be reorganised and grouped together in a latter processing stage. In this paper, we discuss this stage in our proposed region grouping method.

2.2. Feature Extraction

After the image is segmented, the resulted regions need to be grouped to form descriptions of the object in the image. To perform region grouping, information is required to be able to measure the similarity between these regions. This information is extracted from each segment in the form of various types of features. In most existing methods these have been statistical low-level features, such as average luminance, luminance variance, colour and texture. These features are compared and the grouping decision is made based on how close the values between features of each region are. An example of this operation can be seen in [10].

The disadvantage of using low level features alone is that the grouping will not follow the underlying object structure, therefore being semantically inconsistent. However, to extract these features does not require intensive computation. Apart from using low level statistical features, we also need to use higher level information that can reveal the semantic or structural relations between segments.



(a) Objects defined by lines and no colour. We perceive the image as contained objects of human hands.

(b) Objects defined with colour but no lines. We cannot perceive the hands as objects.

Fig. 2. Lines as high level information

The problem now lies in which features possess high level properties that can be extracted without any intensive computations and can be done automatically. As illustrated in Figure 2, line information in an image possesses these high level properties. Lines can describe the structure of an object, hence describing what the content of the image is. Also lines are relatively simple to extract from the image.

By incorporating the line structure information, we can apply some high level grouping rules. This will be described in section 2.4.

2.3. Region Grouping

2.3.1 Related Work. While there is much previous work in image segmentation, this is not the case for region grouping schemes. Work related to region grouping mostly concentrates on images with block objects, such as the work performed by Guzman in [11]. This technique performs grouping on simple polyhedral objects. Objects are formed by exploiting the characteristics of the line junctions in images. The image is then decomposed into objects based on this information.

For natural images, the research in region grouping is very limited. The grouping results have been limited only to producing a better low level segmentation. Nazif and Levine in [9] incorporate a rule-based system to segment an image into uniform regions and connected lines.

Feldman and Yakimovsky in [12] use Bayesian decision theory and semantic information to produce image segments. The lack of concrete region grouping techniques that are not based on heuristics, has made the use of decision theory important and well accepted.

Bayesian decision theory based grouping exhibits the statistical distribution of the data. With this tool, we can make decision under conditions of uncertainty, of which two regions are most similar and can be grouped. This is illustrated in the following diagram.

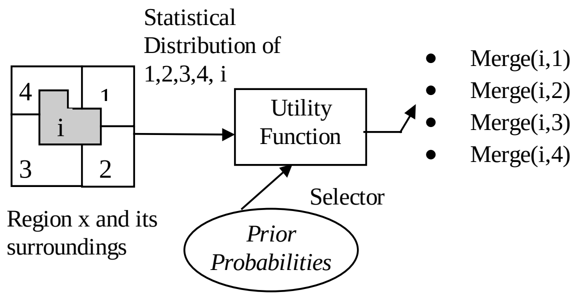


Fig. 3. Merging decision system

The core of Bayesian decision making is the determination of some form of utility function. Given some measurements from the regions, the function ranks the order of which merging action can take place between two regions. The higher rank describes the most probable merger that may take place. As shown in Fig.3, depending on the input to the function, the action can have four possible actions, the merging between region *i* and either of the neighborhood regions {1,2,3,4}. The measurement from region *i* is performed by taking some samples, representing the statistical distribution of the brightness value of region *i*. The value of the rank is computed using Bayes rule:

$$P(\omega_i|\chi) = \frac{P(\chi|\omega_i)P(\omega_i)}{\sum_{j=1}^{k} P(\chi|\omega_j)P(\omega_j)} \quad (1)$$

Where: region labeled i is surrounded by a region in the set j = {1,2,..,k} and:

- χ is a set of samples measured at region i.
- ω_j and ω_i are the state of nature of region i grouped with one of the regions j = {1,2,...,k}. Since these states are unpredictable, we assume that ω is a random variable.
- P(ω_i) and P(ω_j) are the prior probabilities that determine the next state of nature is ω_i or ω_j.
- $P(\chi|\omega_i)$ is the state-conditional probability density function of sample χ, given that the state of nature is ω_i.

Feldman and Yakimovsky used this decision theory in conjunction with semantic information to partition the image into meaningful regions. So far, there has been no automatic techniques that can successfully group segments into significant objects.

2.3.2. Perceptual Grouping. The difficulty in performing grouping is centered in the following questions: how do we determine which segments belong to the same object? Is the similarity measure appropriate? If so, under what situations and which one?

The definition of an object and the elements of the object depend largely on the way humans perceive objects. We believe that it is important to understand this grouping problem from human perception point of view. One work involving the use of perceptual grouping in image analysis, is the work by Brooks [13 BRO83] who developed an image understanding system called ACRONYM to describe and model predefined scenes and objects. We suggest that in order to have a grouping procedure that can organise segments into sets that have similar "perceptual" content, the grouping procedure needs to be based on perceptual grouping laws, known as Gestalt laws. These properties are explained as follows [14]:

Proximity: Closer elements belong to the same object. The proximity principle allows closely spaced elements to be grouped, such as in clustering.

Similarity: Similar elements belong to the same object. This principle allows grouping to be performed on the basis of how similar the segments are. We can compare the intensity, brightness, colour and texture, as a basis for comparing the similarity.

Common Fate: Elements that appear to move together belong to the same object. This allows us to segment video data into segments of moving objects.

Good Continuation: Elements are inclined to be grouped so that the results have smooth and continuous characteristics, rather than yielding abrupt changes. This principle reflects the characteristics of most natural image elements. The brightness values of regions or edges belong to natural images, change smoothly. We rarely see edge shape shown in figure 4(a), whereas the shape in figure 4(b) is most likely to occur.

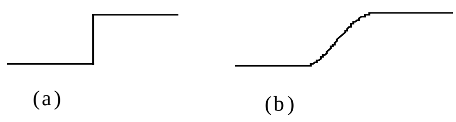


Fig. 4. Edge Profiles

Closure: Of several possible grouping the one which produces a "closed" rather than an "open" object, is more likely to occur. This principle can be used to solve the problem in noisy segmented image, where often we cannot obtain the whole boundary, forming a complete object.

Surroundedness and *Relative Size*: All things being equal, the smaller of two areas will be seen as figure against a larger background. This principle can be implemented to separate an object and the background.

Orientation and *Symmetry*: There seems to be a preference for symmetry and orientation to be seen as figures.

2.3.2 Proposed Grouping Scheme. We propose to use Gestalt laws as a basis for region grouping to do automatic object identification, as described in the previous section. Our task now, is how to implement these laws in the region grouping procedure. The implementation of Gestalt laws into image analysis has been identified as difficult, even for simple similarity measurement [15]. Hochberg and McAlister [16] attempted to quantify Gestalt laws by defining some strict measures. Some segmentation techniques have used principles of proximity and similarity such as in [9]. So far, no techniques have addressed or used other Gestalt principles such as closure or good continuation.

In this section, we list example implementations of Gestalt principles. These are as follows.

Proximity: If line A is at close to line B, extend line A to line B

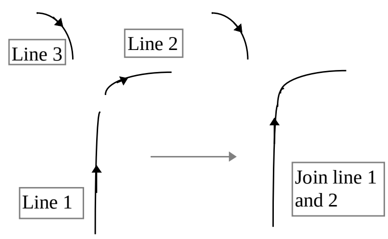


Fig. 5. Merging line elements using proximity

Closure: a line that forms a boundary almost enclosing a region can be joined together.

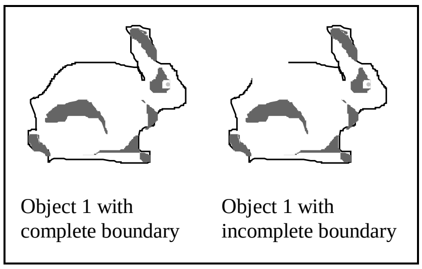


Fig. 6. Closure Principle

Similarity: if region A has a similar intensity/contrast distribution to region B, then A is merged with B.

Good Continuation: two lines close together with similar direction / gradient can be joined together.

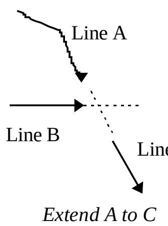


Fig. 7. Grouping Using Good Continuation

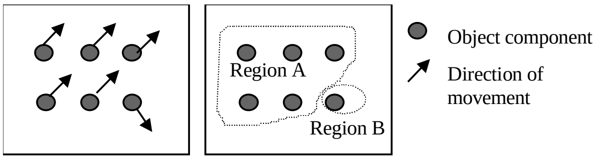


Fig. 8. Grouping Using Common Fate

Common Fate: this principle is useful in analysing video data. Existing schemes segment moving object by locating similar segment in consecutive frames. With common fate principle, we can obtain these objects by analysing the movement, rather than matching the segment template.

Symmetry, Relative Size and Surroundedness: in an image consisting of two regions, where one surrounds the other, and the size of the first region is much larger than the second, with this principle, we can set the first region as an object, and the second region as a background.

To implement these Gestalt laws into our proposed grouping scheme, we separate them into two categories: local and global measure rules. Local measure rules are based on Gestalt principles of proximity, similarity and good continuation. These rules will be used in grouping initial segments, to produce significant regions in the image. After significant regions are achieved, the regions will be combined to form objects using global measure rules that are based on implementing the rest of the Gestalt principles. Therefore, the object description will be formed by a hierarchical region grouping procedure. Figure 9 describes this procedure.

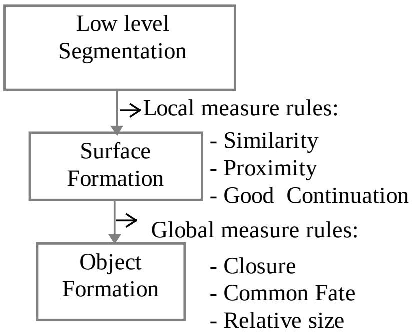


Fig. 9. Hierarchical grouping procedure

2.4 Image Line Structure Analysis

2.4.1 Perception of Line Drawings. When people look at a line drawing, they tend to perceive it as divided into several simpler figures. This area is referred as figure segregation and completion. Although at this stage we do not include figure segregation in our line analysis, we believe it is important to investigate some measures that can describe some properties of lines. These measures will be useful for analysis in the future. Shimaya in [17][18] describes some measures to describe the following line properties: relative number of corners, good continuation, symmetry, curvature constancy, convexity, coincidence, similarity. These measures used for simplified line figures can be computed analytically and will be useful to analyse the line structure. This will be further investigated in the future.

2.4.2 Line Analysis for Region Grouping. To generate the grouping rules, we analyse the image line structure. Based on the Gestalt principle of good continuation, natural image segments are inclined to have smooth and continuous characteristics, rather than yielding abrupt changes. This can mean that when two segments are separated by a line, they cannot belong to the same object, hence should not be grouped, even if their statistical distribution is very similar. This is the first grouping rule that we generated using the line information. This is illustrated in Figure 10(a).

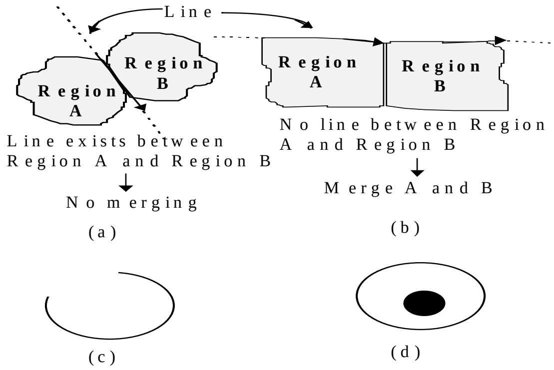


Fig. 10. Analysing line information for grouping

The second rule as shown in Figure 10(b), is based on the condition that when two segments share one continuous line side by side, they belong to the same object, hence can be grouped together.

Using the principle of closure, we can generate the third rule using the line information. Occasionally an image contains objects with no clear boundary from the background or from the other objects, due to the luminance or colour similarity between these objects. Standard region based segmentation can not separate these objects as separate entities. The only information we can use to separate the objects is the line information, but edge detection results in boundaries with many openings, as illustrated in Figure 10(c). In this situation, we can enclose the open boundary and assign the boundary-defined area as a separate object from the background or from the other objects.

Another rule that can be generated by analysing the line structure in an image is in handling concentric objects, as illustrated in Figure 10(d). Using surroundness principle, we can assign the smaller element as an object surrounded by another object or background.

2.4.3 Forming A Structure Tree. After the structural content of the image has been identified, it is very useful to form a structure tree. This description of the image content is useful to describe the relationships between image segments or object components. This can be described by the following example:

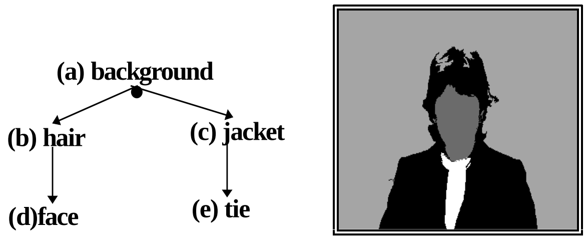


Fig.11. A structure tree example

In this example, the image contains areas resulting from grouping the initial segments together. These areas represent some significant regions in the image such as the object's face, hair, jacket, tie and the background. The relation between these components is formed by building the structure tree as described above. This structure tree can be used to store the relation between these components to label the image content.

3. Grouping Experiment

To test our proposed image structure analysis method, we set up a region grouping experiment. The layout is shown in Figure 12. At this stage we are aiming to produce significant regions in an image (first grouping stage). Higher level grouping is required to identify semantic objects. In this experiment, we need to perform three processes: quantisation, line extraction and the grouping procedure. The quantisation process aims to produce image segments that serve as initial object components. Line extraction aims to produce useful lines that are used to determine which segments are grouped together. The way on which this grouping is performed is described in the grouping procedure.

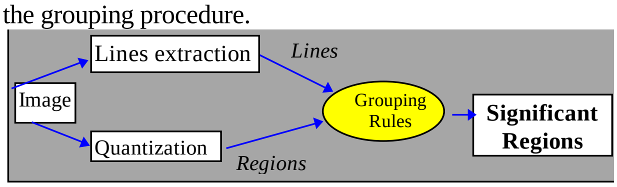


Fig. 12. Region Grouping Implementation

3.1 Quantisation Process

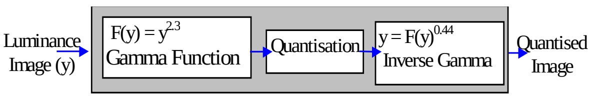


Fig. 13. Quantisation Using JND Threshold

To segment the image, we use the companding technique using JND (Just Noticeable Difference) threshold [19]. This technique reduces the number of colours or intensity values present in an image. Using this method, the pixels with a small difference in colour can be grouped as one segment under one colour representation. This process is illustrated as follows.

Before the quantisation process is taken place, all values are transformed to Gamma function domain, which is described as follows:

$$F(y) = y^{k} \qquad (1)$$

Where y is luminance value and k ranges from 2.2-2.5. In this experiment, we select the value $k = 2.3$.

This compresses low intensity values in accordance with the perceptual characteristics. This domain is used to model what we "visually" perceive. The values in the new domain are quantised to 5 bits. Changes in value below this threshold are visually indistinguishable. After each value is assigned according to this threshold, it is transformed back to original domain, using the following equation.

$$y = F(y)^{1/k} \qquad (2)$$

The segments resulted using this method is shown in Figure 16(b), 17(b) and 18(b).

3.2 Line Extraction Process

To extract lines from the image, the following extraction process is performed.

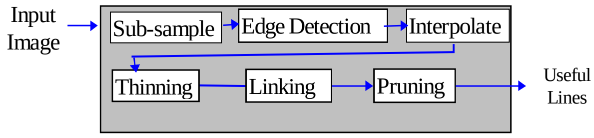


Fig. 14. Line Extraction Process

Before the edge detection process, the image is sub-sampled by 2, to reduce some unnecessary line details. The image edges are detected using Sobel operators as described in the following masks.

-1	-2	-1
0	0	0
1	2	1

(a) To compute the vertical gradient

-1	0	1
-2	0	2
-1	0	1

(b) To compute the horizontal gradient

Then, after the edges are detected, the image is interpolated back to its original size. The resulting edges are not one pixel wide, therefore, to make it easy to manipulate them, the edges are thinned to a width of one pixel using the following 3x3 thinning iterative algorithm [20]. On every iteration, every image pixel is inspected and single pixel wide boundaries, that are not required to maintain connectivity, are erased. The erasing decision is made, based on the thinning criteria: connectivity, maintaining end lines and preventing end lines from being eroded.

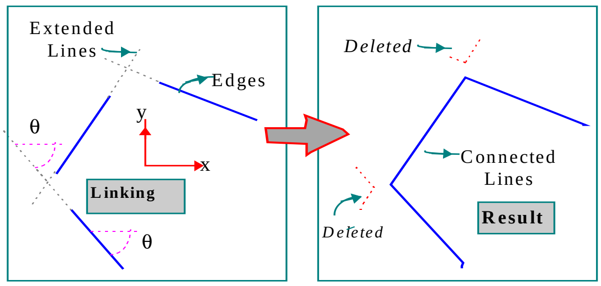


Fig. 15. Edge Linking and Pruning

After the edges are thinned, they need to be joined to form lines. To form the line, the edges are linked by the following method. The edge gradient is used to determine which edges should be linked. The edges are initially extended by 5 pixels according to the direction of the edges. When these extended edges meet other edges, the two edges are linked. The extended lines are then "pruned". In the last step, edges that have a length less than some threshold, are deleted. In this case, only lines longer than 20 pixels are considered to be useful for grouping. The result of this line extraction process is shown in Figures 16(c), 17(c) and 18(c).

3.3 Grouping Rules

The grouping rules generated for this experiment are described as follows.

(1) IF intensity of region A and B are <u>similar</u> THEN
Group(A,B)

(2) IF region A and B share a common <u>continuous</u> line THEN
Group(A,B)

(3) IF region A and B is separated by a line THEN
Do not group(A,B)

Rule (1) is based on Gestalt theory of similarity and the similarity of two regions is computed by taking the difference between the mean of the regions.

Rule (2) and (3) are based on Gestalt good continuity principle and are explained in section 2.4.

3.4 Grouping Procedure

Using these grouping rules, the following grouping procedure is implemented:

```
Get Image Segments;
Get Image Line information;
FOR all segments in the image DO:
    Get Next Segment Pair (A,B);
    Measure the statistical properties of region A, B;
    IF Similarity Test < Threshold THEN
        IF region A and B not separated by a line THEN
            Group(A,B);
        ELSE
        Get Next Segment Pair (A,B);
    ELSE
        IF region A and B not separated by a line THEN
            IF region A and B share a common line THEN
                Group(A,B);
            ELSE
                Get Next Segment Pair (A,B);
        ELSE
        Get Next Segment Pair (A,B);
    END
END procedure
```

After the image is segmented, a component labeling algorithm [21] is applied to label the regions. This information is stored in a region map array. The line information is stored in a separate edge map array. When two regions are compared, the region map array is inspected to measure the luminance distribution of the regions. The edge map is also used to determine the existence of lines around the regions.

The experimental results of our grouping procedure are described in the next section.

4. Experiment results

The following are the experiment results of the schemes presented in this paper. The luminance image "Claire" is used, as shown in Figure 16(a). This image is segmented and the results are shown in the image in Figure 16(b). Initially 842 segments were produced. The image in Figure 16(c) depicts the lines extracted from the original image.

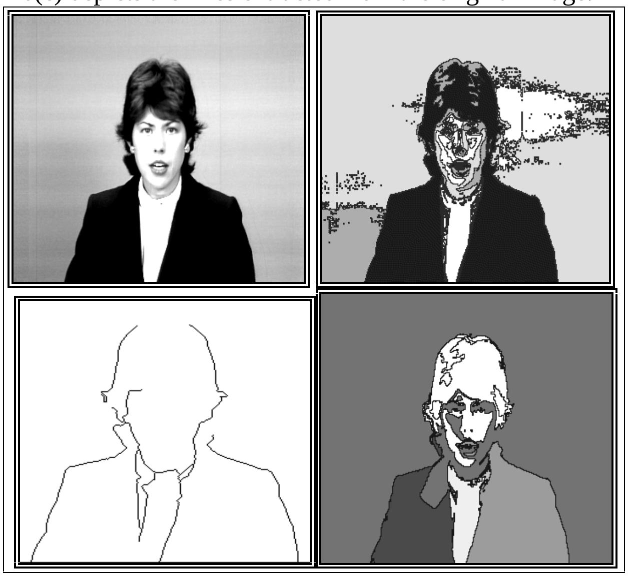


Fig. 16. Example 1 using image "Claire": (a) Original image, (b) Segmented image, (c) Line structure used, (d) Region grouping result (regions are labeled in different colours)

Fig. 17. Example 2 using image "Suzie"

The grouping is performed using the above similarity measures and the edge information. The grouping results are shown in Figure 16(d). The number of regions produced is reduced from 842 to 11 object components.

Comparing the images in this example, it is clear that after the grouping process, the result contains more significant areas. (Refer to image example 1: areas of face, hair, jacket, scarf and background). Example 2 exhibits a large textural area (in the hair) and low contrast between various areas. The number of segments produced is reduced from 1106 to 11 main object components. The grouping result for this image shows some significant areas are captured, such as face, hair, the hand and some part of the telephone.

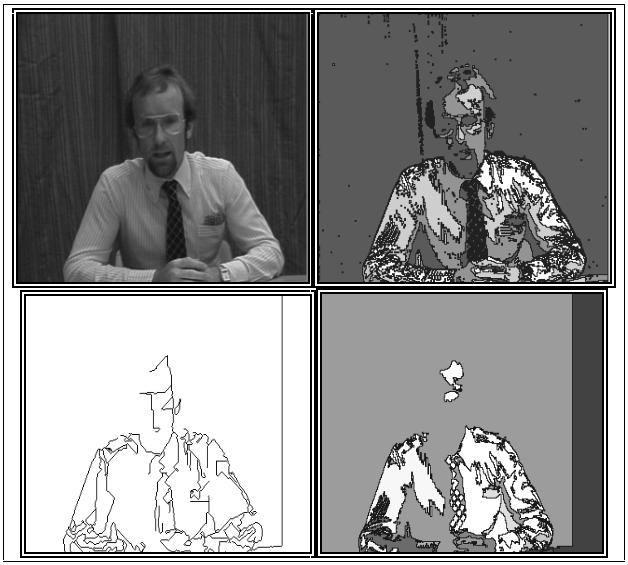


Fig. 18. Example 3 using image "Trevor"

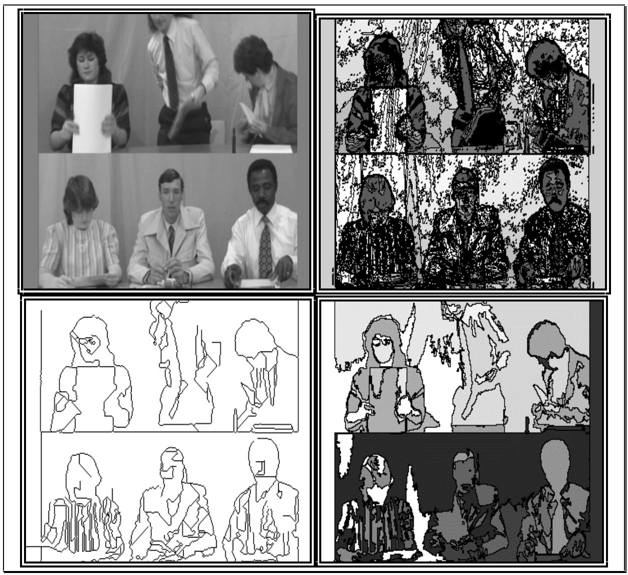


Fig. 19. Example 4 using image "Split"

In example number 3 (Fig.18), both the background and the object are textured. There is also a low contrast between the object and the background that generates errors in the edge detection process. For this image the grouping process detects the body of the actor, and the number of areas formed are reduced from 990 to 19.

In example number 4 (Fig. 19), the image contained multiple objects with a variety of colour and texture. The initial segmentation produces 1424 regions. The grouping results in 37 significant areas.

5. Future Work and Conclusion

In this paper, the study of issues involved in providing automatic object identification was presented. Recent work in the area of content-based retrieval has identified the need for automating the process of object identification. Since manual annotation is a tedious process, we have proposed to use image segmentation techniques as a partial solution to the problem.

While image segmentation techniques may operate automatically, the granularity of the resulting regions are too small to define meaningful objects. We solve this problem by using a region grouping procedure to obtain more meaningful and consistent image regions. Structural information in the form of line description is used as image high level descriptions.

We investigated Gestalt principles as a first step to understanding how to form objects from image regions. We proposed some possible implementation of Gestalt principles for region grouping. Our experimental results demonstrate an improvement in the regions obtained over simple segmentation. These can be used to provide better labels for the image and can support a better and more accurate image content based retrieval system.

Future work includes:

To use other grouping technique: We will improve the grouping method by implementing Bayesian Decision theory in the initial grouping process.

To implement other Gestalt rule implementation: So far we have implemented the Gestalt principle of similarity, proximity and good continuation. We will investigate more rules that can be obtained using lines as structural information and implement the other Gestalt principles in the grouping process.

To extend the proposed system to video data: Object identification is also required for video content-based retrieval, hence we will implement this same basic approach for analysing video data.

Develop image / video indexing system: To validate the proposed grouping schemes, we will develop an object-oriented image / video database, as described in the first section.

6. References

[1] M.Flickner, H.Sawhney, W.Niblack, J.Ashley, Q.Huang, B.Dom, M.Gorkani, J.Hafner, D.Lee, D.Petkovic, D.Steel, P.Yanker., "Query by Image and Video Content: The QBIC system," *IEEE Computer*, Sept.1995, pp.23-31, 1995.

[2] R.K.Shihari, "Use of Captions and Other Collateral Text in Understanding Photos," *Artificial Intelligence Review*, Special Issue on Integrating Language and Vision, Vol.8, pp.409-430, 1994-95.

[3] W.L Grosky, R.Mehrotra, "Index based Object Recognition in Pictorial Data Management," *Computer Vision, Graphics and Image Processing*, Vol.52, No.3, pp.416-436, 1990.

[4] P Halvadar, G Medioni, F Stein, Perceptual Grouping for Generic Recognition, *IJCV*, 20 (1/2), pp.59-80, 1996.

[5] C. Rothwell, "Hierarchical Object Descriptions using invariants", Application of Invariants in Computer Vision II, Pub.Azores, 1993, pp287-302.

[6] D.Clemens, D.Jacobs, "Space and Time Bounds on Indexing 3D Models from 2D Images" *IEEE Trans. PAMI*, No.12, pp.1007-117, 1991

[7] T.R. Reed, "Segmentation-based image processing," Proc. Int'l conf. Digital Signal Processing, Vol.1, pp.308-313, 1995.

[8] R.Hart, P.Duda, Pattern Classification and Scene Analysis, Wiley-Interscience Publication, 1973.

[9] D.H Ballard and C.M Brown, Computer Vision, Prentice-Hall, New Jersey, 1982

[10] A.M.Nazif, M.D.Levine, "Low Level Image Segmentation, *IEEE PAMI*," Vol. PAMI-6 No.5, pp.555-577, 1984.

[11] A.Guzman , "Decomposition of a Visual Scene into three Dimensional Bodies," Proc. FJCC 33, pp.291-304, 1968.

[12] J.A.Feldman, Y.Yakimovsky, "Decision Theory and Artificial Intelligence: I.A Semantics-Based Region Analyser," *Artificial Intelligence*, Vol.5, pp.349-371, 1974

[13] R.Books, "Model Based three Dimensional Interpretation of two Dimensional Images." *IEEE Trans.PAMI*, No.5, pp.140-150, 1983.

[14] V.Bruce and P.R.Green, Visual Perception: Physiology, Psychology and Ecology, 2nd Ed., Lawrence Erlbaum Assoc., Hove and London, 1990

[15] A Tversky, Features of Similarity, *Psychological Review*, Vol.84, No.4, 1977.

[16] R.Hochberg, A.McAlister, "A Quantitative Approach to Figure Goodness." Journal of Experimental Psychology, Vol.46, pp.361-364, 1953.

[17] A.Shimaya, "A Cognitive Model of Figure Segregation", Proc. 12th Int. Conf. Artificial Intelligence, Vol.1, *IJCAI-91*, 1991.

[18] A Shimaya and I Yoroizawa, Perception of Complex Line Drawing, *Journal of Experimental Psychology: Human Perception and Performance*, Vol.23, No.1, 25-50, 1997

[19] G.Murch, N.Weiman, "Assessing Visual Grey Scale Sensitivity On A CRT," *Proc. of the SPIE*, Vol. 1249, Human Vision and Electronic Imaging: Models, Methods and App. 1990, pp.214-223

[20] L.O'Gorman, "k x k Thinning", Computer Vision, Graphics and Image Processing," v.51, pp.195-215, 1990.

[21] A.F Lochovsky, "Algorithms for Real time Component Labelling of Images," *Image and Vision Computing*, Vol.6, No.1, pp.21-27, 1988