Self-Similar Convolution Image Distribution Histograms as Invariant Identifiers

C.E. Tuke, S.E.M. O'Keefe, J. Austin
Department of Computer Science,
University of York

Abstract

This paper investigates the practical application of a new method called self-similar convolution masks to binary trademark images to produce affine invariant one dimensional histogram descriptions. Because the convolution mask is a scaled version of the original image, the normalised distribution histogram of the resulting grey scale image is an affine invariant description based purely upon the image structure, which can be then be used for database search using a database of binary images.

The effectiveness of this approach to trademark image identification is tested. Multiple methods of representing the distribution histograms within the database are also tested for robustness to noise, generalisation and reliability.

1 Introduction

The reliable identification of images from large databases indexed directly upon image content has been the topic of much research over the past decade. Content based searches usually take the form of queries specified by means of colour, sketches, example images or icons which are then encoded into a more compact description that can be used in the database. One of the first examples of this approach, which actually consisted of a combination of different search criteria, was the "Query By Image Content" [1] project. A major issue, which limits mainstream acceptance of these approaches, is the rather subjective concept of image similarity. Images that possess obvious similarities to a human operator may still be extremely different in terms of apparent pixel content to each other. These difficulties, along with the ambiguity of the initial choice of salient features in the images, results in techniques being developed that are not totally automatic, often requiring substantial human guidance in both database creation and search.

Despite a wealth of research interest in the area of content-based image search over recent years that has largely been fuelled by the proliferation of imagery over the Internet and the subsequent need to search this information, content-based search engines are still in their infancy. Trademark research, which often will involve a large range of worldwide search engines of different types, still relies almost exclusively on the verbal categorisation of image content for database indexing. Although traditional image classification is a well established subject [2][3], there is scope for ambiguity in such subjective classification methods, along with the inevitable difficulties where a trademark falls into multiple

BMVC 2001 doi:10.5244/C.15.52

categories or has no obvious verbal equivalent. Examples of this description-based approach can be found in the "U.S. Trademark Electronic Search System" [2], which covers U.S. patented trademarks, along with its Australian equivalent, ATMOSS[3].

Any system designed to search such huge repositories of images would clearly need to utilise an image description which, whilst being robust to image transformations and noise, is also detailed enough to allow a unique representation for each trademark. Such a system would also require some ability to identify trademarks that, whilst not being the same, represent similar image content or form.

2 Common Image Content Description Techniques

Many of the most successful image content approaches have made extensive use of colour distribution information, usually stored as a histogram. Such colour based approaches prove to be particularly effective at bypassing geometric invariance problems in complex or natural images. Examples of colour based search approaches can be found in the [5] and the colour histogram approaches of Stricker [6]. Unfortunately, colour based approaches are highly unsuitable for the encoding of trademark images, which are primarily simple binary images. For this reason, description techniques that encapsulate shape or texture are more appropriate.

Another popular method of encoding image content is through the use of contour information. The relative change in contour curvature represents a good source for a similarity invariant description. Approaches that rely on object contours include chain codes [7], polar descriptions, curvature scale space [8] and edge direction histograms [9]. These approaches suffer from difficulties attaining sufficiently unique similarity invariant descriptions for database entry and can be very sensitive to errors in border extraction, exspecially when dealing with small trademark images.

Much recent research, especially in the natural image domain, has been devoted to the application of invariant features to form descriptions or image signatures [10] which can be used as invariant image references. In the case of trademark identification, similarity invariance would be minimum level of invariance required. Although well suited to the more complex image domains, these approaches usually require more information than is available in small binary trademark images.

A variation of this approach can be found in [11], where trademark images are processed in accordance to human perception cues and multiple image representations used. In many ways, this approach comes closest to mimicking human recognition processes.

Texture based descriptions are often attained through the use of masks applied to the image in question. An extensive overview of texture representations can be found in [12]. A major problem in the application of mask filters is their sensitivity to scale and transformation changes in the source images. In many respects, self-similar convolution masks are closely related to texture techniques.

3 Self-Similar Convolution Image Histograms

This novel approach to assigning similarity invariant descriptions to binary images, in the context of indexing a trademark image database, is adopted and examined in this paper.

Self-similar convolution histograms are generated from the interaction of the original binary image with a convolution mask generated from a scaled version of itself. An example of the generation of a self-similar convolution mask, at a scale of 0.5, follows.

```
width= Original binary image width
height=Original binary image height
f()=Original binary image
m()=Convolution mask
c()=Convolution grey-scale image
mask_width=width/2
mask_height=height/2

for y=0 to mask_height
    for x=0 to mask_width
        m(x,y)=0
for y=0 to height
    for x=0 to width
        if  (f(x,y)==1)  m(x/2,y/2)=1
```

And now, to generate the convolution grey-scale image:

```
for y=0 to height+height
    for x=0 to width+width
         for s=0 to mask_height
              for t=0 to mask_width
                  c(x,y)+=m(s,t)*f(x-mask_width+s,y-mask_height+t)
```

Results of this convolution process can be seen in Figure 1.

Because the convolution mask is directly derived from the original image, the resulting image contains richer grey-scale information unique to the structural and textural qualities of the original binary image.

In order to achieve a Euclidean level of invariance, the distribution histogram of the grey-scale image is generated.

As it stands, our histogram description is still not a similarity invariant representation. In order to eliminate scaling effects, as much as possible, the distribution histogram on both its axes must be self-normalised. In theory this should make the histogram scale invariant although, in reality, pixel aliasing effects will cause increasing noise as the trademark image reduces in size.

The paper tests combinations of four approaches to normalised representations, see Figure 2.

3.1 Base (x axis) Histogram Normalisation

1. 'Not Stretched' - Base normalised by maximum POSSIBLE grey-scale convolution mask output.

2. 'Stretched' - Base normalised by maximum grey-scale convolution mask output present.

3.2 Value (y axis) Histogram Normalisation

1. 'Normalised by highest value' - Histogram values are normalised by the HIGHEST value present in the histogram.

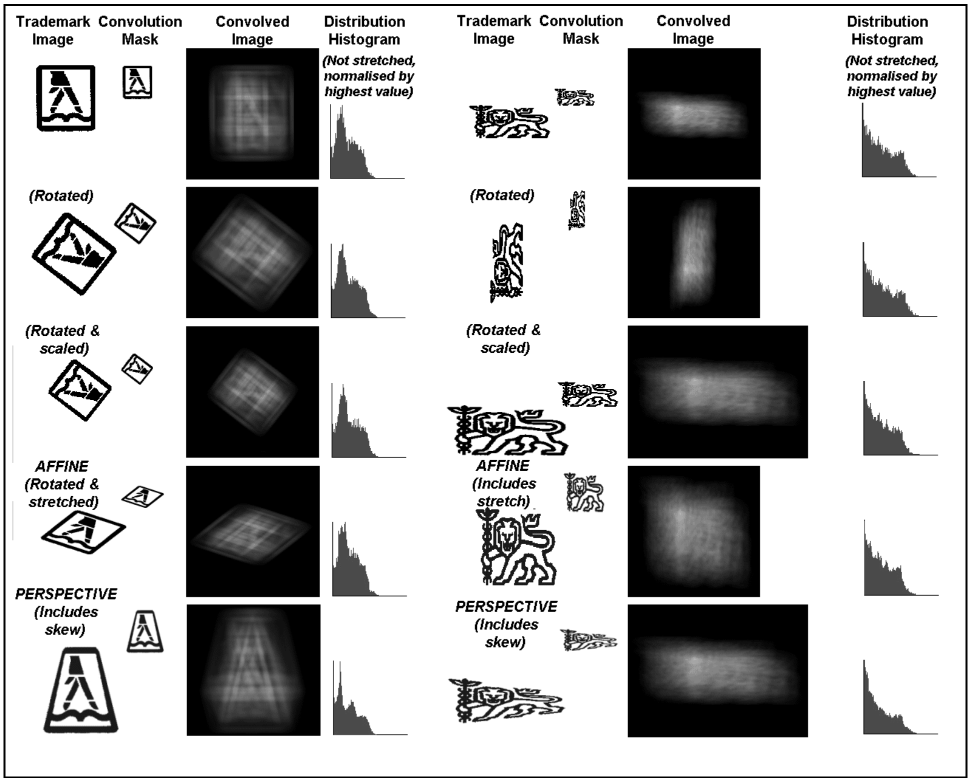

Figure 1: Showing the invariance of self-similar convolution image histograms to image transformations

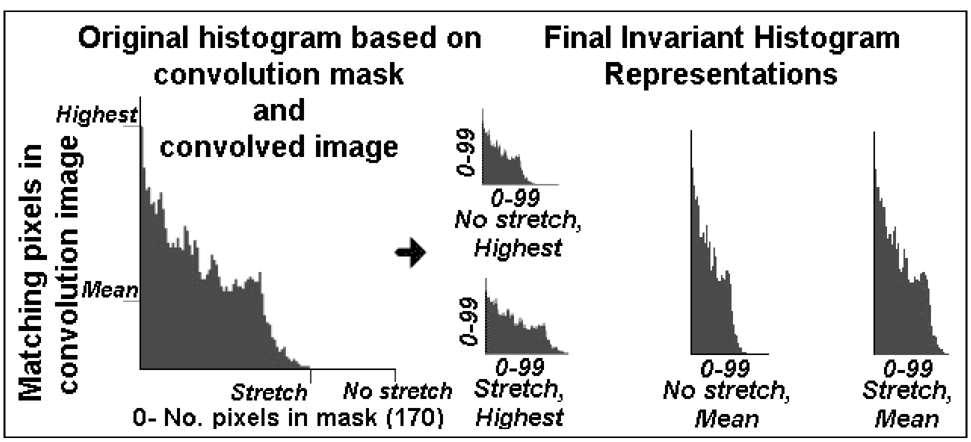

Figure 2: The 4 histogram representations tested in this paper

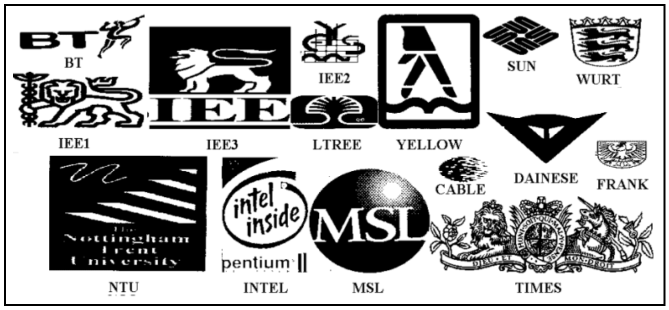

Figure 3: The 15 main trademark image types used in the database

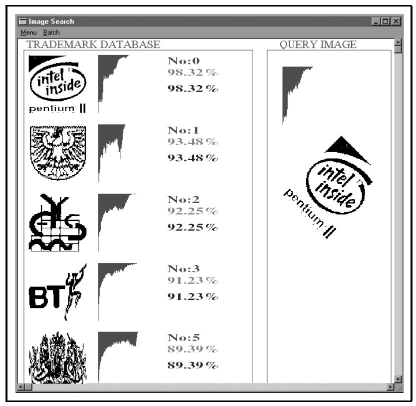

Figure 4: Screenshot of the database created to test self-similar convolution image histograms

2. 'Normalised by mean value' - Histogram values are normalised by the MEAN value present in the histogram.

As can be seen in Figure 1, the resulting histogram representation is invariant to all similarity transformations, a quality we require for successful identification. It also proves to be highly tolerant to affine transformations, although it can be seen to fail when subjected to perspective transformations. This histogram representation can now be used as an invariant database description for the original trademark.

4 Implementation

Before describing the process with which self-similar convolution distribution histogram descriptions will be tested it is essential to first clarify the problem domain. The aim is to achieve a system whereby, given a binary image of a single trademark, a database of other binary trademarks can be queried to output a list of trademark images suitably ranked in accordance with their similarity to the query. Given this narrow problem domain it is assumed that the image is primarily of a single and complete trademark that has already been separated from any background context but may exist at a variable orientation, scale

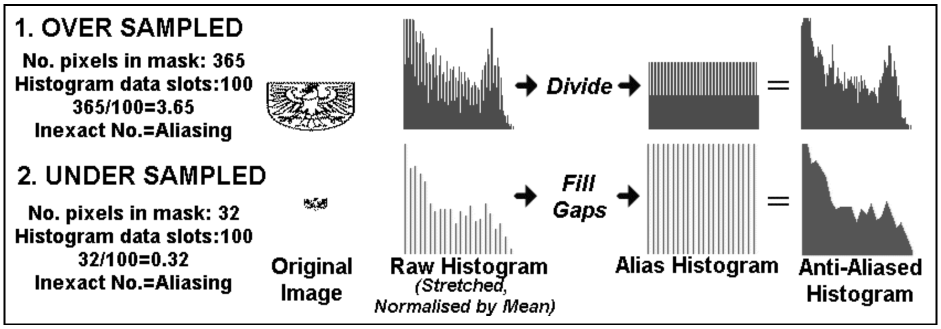

Figure 5: Anti-aliasing the database histograms

or position within the image and may contain a limited degree of scan noise.

A database application has been constructed to test the effectiveness of self-similar convolution histograms and the four different methods, outlined previously, of normalising them. The database contains a series of trademark images (Figure 3), results are indexed by their self-similar convolution histograms and is sorted according by their similarity to a query image's histogram.

Figure 4 shows a screen-shot of a typical query and database response, in this case the database is using a non-stretched, normalisation by highest value histogram representation. All experiments performed using this database were duplicated for the four main histogram representation types:

1. Not stretched, normalised by highest.
2. Not stretched, normalised by mean.
3. Stretched, normalised by highest.
4. Stretched, normalised by mean.

   To test the value of using the histogram representation at all:

5. Single average (non-zero) grey-scale value of convolution image.

The scale factor used when defining the convolution mask can greatly affect the resulting convolution image. A noticeable effect of decreasing the scale of the convolution mask is a greater generalisation capacity at the expense of unique identification. However, the effect of different scale values on self-similar convolution histograms is not investigated in this paper and a ratio of 0.5 is assumed for all convolution mask sizes.

4.1 Anti-Aliasing

To provide a common comparison framework and ensure similarity invariance, it is necessary to normalise the histograms to a pre-defined range, in this case a range from 0 to 99. This process inevitably results in a certain degree of aliasing in the results, which can be detected by creating a histogram of constant entries. All histogram representations use these "Alias Histograms" to correct for aliasing in the original histogram image, see Figure 5.

	No stretch, High.	Stretch, High.	No stretch, Mean	Stretch, Mean	Tot. Avg.
Normal	61.55%	68.86%	58.35%	69.67%	32.42%
Edge	48.36%	51.46%	40.81%	46.52%	31.13%

Table 1: Percentage number of successful trademark image matches between the noisy sample image set and the trademark database (Figure 3). Total results for both edge and full trademark image representations are presented.

4.2 Noise Tolerance

Any trademark image is likely to exhibit some degree of noise, so a good database should have a reasonable degree of noise tolerance built in. Figure 6 shows the effect of uniform noise upon database recognition rates. Each image in the database has uniform noise added to it and is submitted as a query to the database. The average recognition rates for all trademark images under a given degree of noise are plotted on the graphs. As can be seen in Figure 6a and 6b, recognition rates remain highest in those histogram representations that are normalised by the more stable mean histogram value. Mean-normalised histograms provide perfect recognition rates with up to 2.5 percent noise.

Figure 6c shows the difference in results caused by actual recognition scores, or the distinctiveness of the results. This is a measure of the relative score difference between false positives and good matches. As can be seen, the mean-normalised representations provide the greatest distinction between true and false matches. This distinctiveness remains largely constant when dealing with images containing less than 10% noise.

4.3 Generalisation Abilities

The ability to generalise is essential to trademark image search. Trademarks that are similar in content should feature highly where no exact match is apparent. In order to test self-similar convolution histograms for generalisation, a series of 186 test images of similar trademarks were scanned in. This new collection of images represents a selection typical of anticipated database query variations, examples are shown in Figure 7.

This set of trademark images were then submitted to the database, the classification results can be seen in Figure 8a, 8b and Table 1.

As it would also be desirable to include recognition of inverted trademark images, an identical series of tests were also conducted using the edge images of the trademarks, tested on a database derived from trademark edge images, Figure 8b. Although the use of edge images enabled the correct classification of inverted trademark images it can be seen, by comparing Figures 8a and 8b, that the identification rate of edge images is significantly lower than raw images. An interesting observation, when comparing edge results to the original results, is that when we use the scalar total average as an identifier the recognition rate of 32% is virtually the same for either type of image. It can be seen from these results that, in terms of robustness to noise and generalisation, a stretched histogram representation achieves the highest recognition rate of 69%. This rate is statistically significant and demonstrates that self-similar convolution histograms have good generalisation abilities and could form part of a practical image identification scheme.

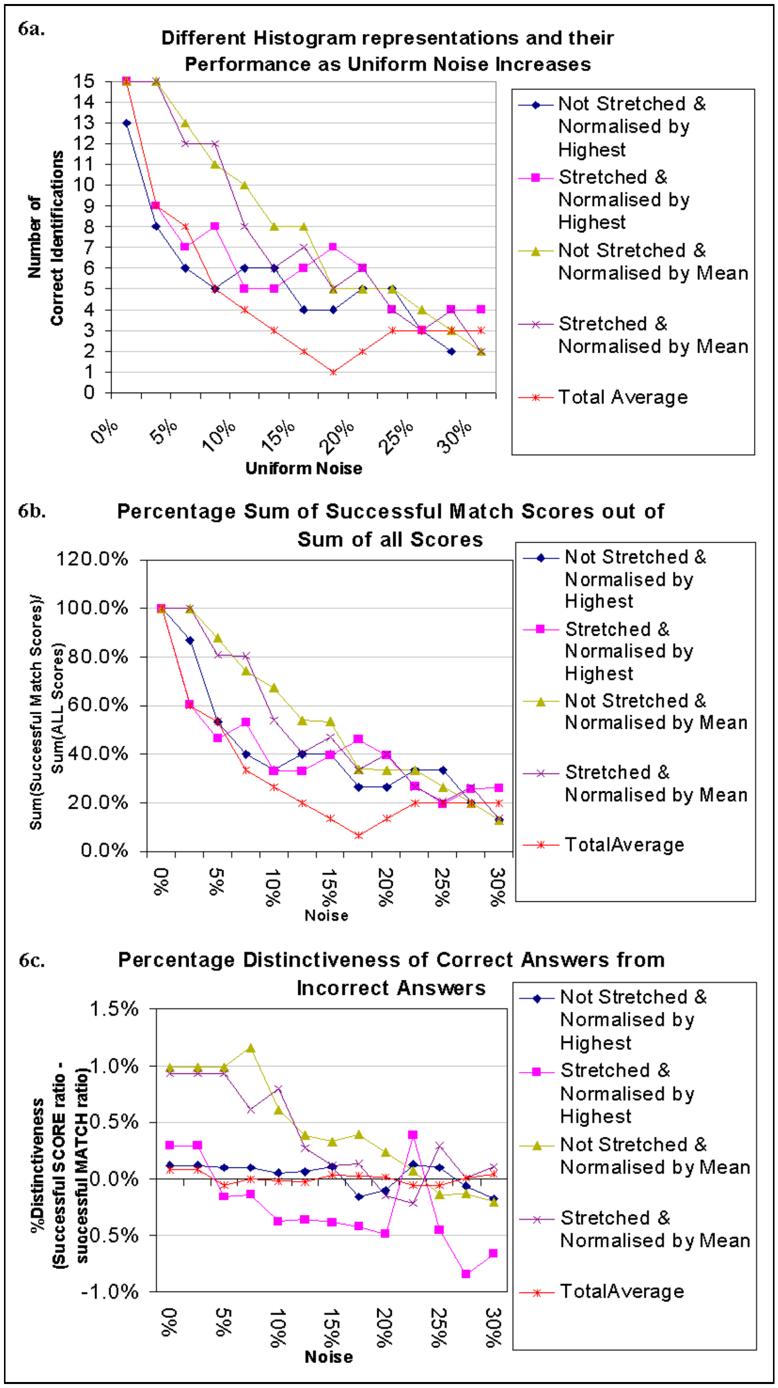

Figure 6: Performance of Self-Similar Convolution Image Histogram representations under increasing noise.

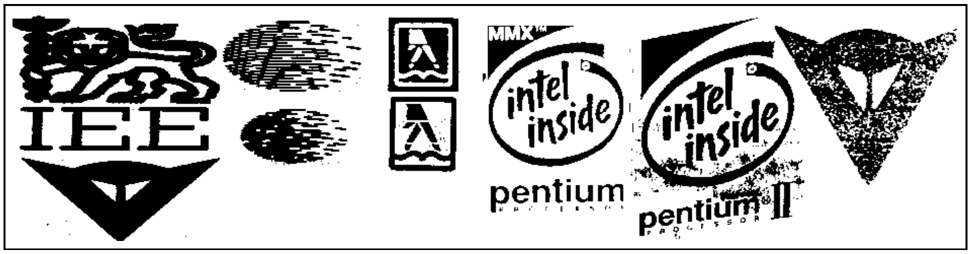

Figure 7: Typical Variations of Database Query Images Tested.

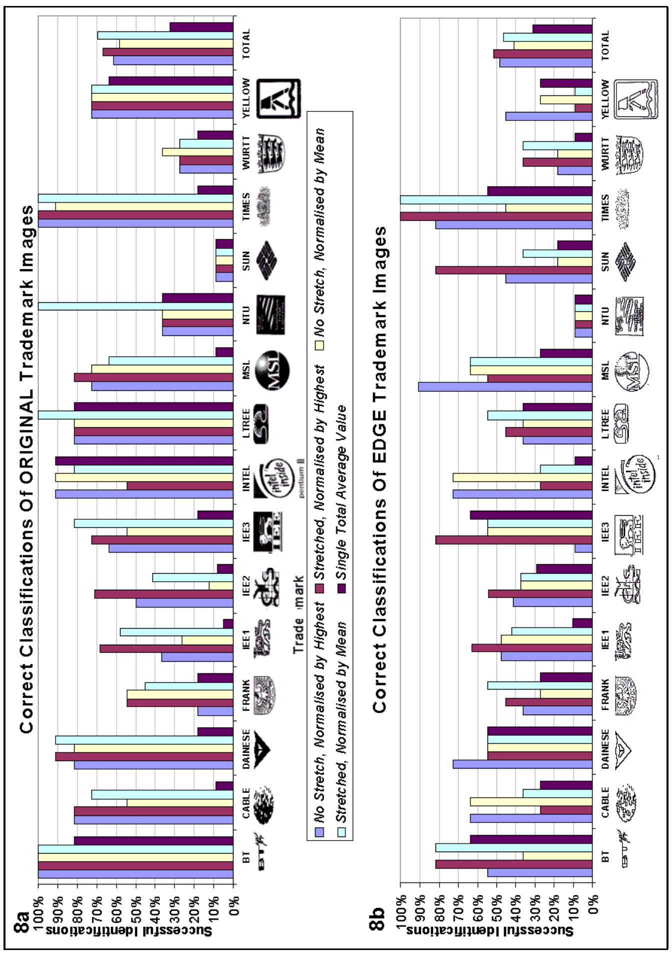

Figure 8: Results of generalisation tests over a range of 186 query trademark images.

5 Discussion

Self-similar convolution image histograms provide a simple method of invariant image identification, based solely on image content, that proves reasonably tolerant to both noise and generalisation. The time taken to generate the convolution image is directly proportionate to the size of the mask image, and therefore the original image. This results in an exponential increase in calculation time as the image size increases. This, combined with aliasing effects with very small images, suggests that a more effective approach would restrict query image sizes to a predefined range. Although both noise tolerance and generalisation properties of self-similar convolution image histograms have been tested in this paper, it is still unclear how this description will perform when applied to very large datasets and this represents an avenue for further work.

In the context of trademark image identification the technique shows much promise, although it would be most effective in an evidence accumulation framework. The implementation of a database indexed through the combination of self-similar convolution histograms and another histogram representation, such as edge direction histograms [9], seems to represent the most effective way forward.

6 References

[1] Niblack et al., The 'Query By Image Content' project: Querying images by content using color, texture and shape.
Proceedings SPIE Conference on Storage and Retrieval of Image and Video Databases. 173-187, 1993
[2] Besser, Howard. Getting the Picture on Image Databases – The Basics.
Database 18: 2 (April/May 1995): 12-19.
[3] Graeme Baxter and Douglas Anderson. Image indexing and retrieval: some problems and proposed solutions. *Internet Research,* Volume 6 Number 4, 1996
[3] T.E.S.S. U.S. Trademark Electronic Search System.
http://www.uspto.gov/web/menu/tm.html
[4] ATMOSS. Australian Trade Mark On-Line Search System.
http://pericles.ipaustralia.gov.au/atmoss/falcon.application_start
[5] Theo Gevers. PicToSeek: A Content Based Image Search System for the World Wide Web. University of Amsterdam, Intelligent Sensory Information Systems.
[6] Markus Stricker and Alexander Dimai. Similarity of color images.
Proc. SPIE Storage and Retrieval for Image and Video Databases,96
[7] RJ Schalkoff. *Digital Imaging and Computer Vision.* Wiley, 294-295, 89
[8] Mokhtarian, F; Abbasi, S and Kittler, J, University of Surrey Robust and Efficient Shape Indexing through Curvature Scale Space *The British Machine Vision Conference*, 96
[9] Jain, A.K and Vailaya, A. Image retrieval using color and shape. *Pattern Recognition*, 96, 29(8):1233-1244
[10] Jukka Livarinen, Markus Peura, Jaako Srel and Ari Visa Comparison of Combined Shape Descriptors for Irregular Objects *8th British Machine Vision Conference*, 97
[11] T.P.G.L.S. Alwis Content-Based Retrieval of Trademark Images. Thesis work. University of York. February 2000
[12] D.H. Ballard and C.M. Brown. Computer Vision. *Prentice-Hall*, 181-193, 82