

EXTRACTION OF DETAILED IMAGE REGIONS FOR CONTENT-BASED IMAGE RETRIEVAL

D. Androutsos[†], K. N. Plataniotis[‡] and A. N. Venetsanopoulos[†]

[†] Digital Signal & Image Processing Lab
Department of Electrical & Computer Engineering
University of Toronto
10 King's College Road, Toronto, Ontario, M5S 3G4, CANADA
<http://www.dsp.toronto.edu>
{zeus,anv}@dsp.toronto.edu

[‡]Ryerson Polytechnic University, Department of Math, Physics & Computer Science
350 Victoria Street, Toronto, Ontario, M5B 2K3, CANADA
kplatani@acs.ryerson.ca

ABSTRACT

We present a technique for coarsely extracting the regions of natural color images which contain directional detail, e.g., edges, texture, etc., which we then use for image database indexing. As a measure of color activity, we use a perceptually modified distance measure based on the *sum-of-angles* criterion. We then apply histogram thresholding techniques to separate the image into smooth color regions and *busy* regions where edge, texture and colour activity exists. Database indices are then created from the *busy* regions using the *directional detail histogram* technique and retrieval is performed using these.

1. INTRODUCTION

Efficient access to digital data has become an issue of utmost importance recently. In particular, the amount of digital image and video data available is staggering and the challenging issues of cataloging and retrieval has gained increasing importance. Without a doubt, efficient access to relevant data directly determines its value.

As digital acquisition and storage grow, a number of industrial fields, such as medical imagery, graphic arts, textile and paint, satellite imagery, criminology and film, require efficient access to their data. Content-Based Image Retrieval (CBIR) is a relatively new research area which is dedicated to the image retrieval problem [1] and a number of image database system have been developed [2, 3].

A key aspect of image databases is the creation of robust and efficient indices which are used for retrieval. Color and texture remain the most important low-level features which are used to index database images[4]. This is not surprising since the early stages of human vision and also human recall are closely linked to these feature sets.

Texture, in particular, proves to be quite challenging to index and numerous texture classification techniques have been considered as indices. The focus has primarily been

on applying statistical methods. Recently, multiresolution techniques have been considered, namely wavelet techniques [3, 5].

Furthermore, the spatial information contained in an image proves to be of utmost importance. For example, it is often required to find images which contain a certain color, shape or certain detail structure in a given neighborhood. Thus, it is a great advantage to be able to extract, a priori, certain spatial image information.

In this paper we present a technique for extracting natural color images into two coarse regions: one which contains smooth color and another which contains all the edge, texture and directional detail. We then apply the *directional detail histogram* [6], to essentially *fingerprint* the extracted detail regions, to build indices into an image database. By performing a coarse extraction, we can easily incorporate spatial information into our database indices regarding the location and amount of certain colours or detail and texture content. In this paper, we focus on the detailed areas. Furthermore, we are not interested with *precise* segmentation since image retrieval deals with similarity and exact spatial information is not required.

Section 2.1 discusses our extraction technique and Section 3 describes the indexing procedure. Some database query results are presented in Section 4 followed by our conclusions and future research in Section 5.

2. DETAIL EXTRACTION

2.1. Color activity measure

Color activity is an issue directly related to color texture segmentation and classification. The importance lies in the fact that an image can exhibit much wider ranges of activity levels due to color information. This can lead to much better classification and extraction of textured and detailed regions in color images as compared to performing well-established grey image texture segmentation techniques on

converted color images.

To perform extraction, we need to have an activity measure at each image location. We propose the following perceptually-based measure, from the *sum-of-angles* criterion [7]:

$$s_{ij} = \sum_{i=1}^W \sum_{j=1}^H \sum_{r=1}^{N^2} \frac{\mathbf{x}_{ij} \mathbf{x}_r^T}{|\mathbf{x}_{ij}| |\mathbf{x}_r|} \left[1 - \exp \left\{ \frac{|\mathbf{x}_{ij} - \mathbf{x}_r|}{\max(|\mathbf{x}_{ij}|, |\mathbf{x}_r|)} \right\} \right], \quad (1)$$

where W and H are image width and height, N is the number of vectors \mathbf{x}_r in an $N \times N$ window and \mathbf{x}_{ij} is the centrally located vector in that window. The first component in this measure is the *angle* between two vectors and the second is a perceptually weighted *distance* between two vectors. According to Shepard [8], visual stimuli exhibit a monotonic decrease in similarity, as a function of distance, approximating an $\exp(\cdot)$ function. Thus, color vectors which have a large magnitude difference contribute less to the measure s_{ij} . The reasoning behind using a perceptually based measure is to group together colour vectors which are perceived as similar. This is especially important for image retrieval since precise results are not the goal but results which are perceived as similar to a query image or similar to a specified characteristic.

The final activity measure is then found by taking the variance of s_{ij} in an 3×3 window:

$$\zeta_{ij} = E\{s_{ij}^2\} - E^2\{s_{ij}\}. \quad (2)$$

2.2. Histogram thresholding

After the activity measure $\xi(ij)$ is calculated for each vector, we normalize the values and build a 256-bin histogram. Inspection of these histograms reveals *peaks* and *valleys*, however we found for the majority of histograms, there exist two distinct peak groupings corresponding to the *smooth* and *busy* areas, as shown in Figure 2. For images with a great amount of detail and texture or with primarily smooth areas, the histograms did not exhibit multiple peaks, so the image is classified as either entirely *smooth* or *busy*. Figure 2 the histogram of the *zebras* image in Figure 1

After detail extraction we perform some morphological operations (CLOSE and OPEN) to remove small isolated areas and to smooth the regions. Figure 3 depicts the segmentation of the *zebra* image after thresholding.

As can be seen, the areas of the image which contain strong edge activity and significant detail are effectively grouped and separated from the smooth areas. As mentioned above, our goal is not to do a precise segmentation but an extraction which will identify areas of high detail and edge activity to be indexed into a database.

3. DIRECTIONAL DETAIL HISTOGRAM

3.1. Wavelet Decomposition

Once the regions of the image which contain the directional information have been identified, we perform a 3-level multiresolution decomposition by recursively implementing a



Figure 1: *zebras* image.

separable 2-dimensional Haar wavelet.¹

The justification for using multiresolution decomposition lies in the fact that the resulting subbands contain vertical, diagonal and horizontal edge information at the different levels of detail, as seen in Figure 4. Thus the wavelet coefficients provide a measure of directional activity at each decomposition level.

Each level in the multiresolution decomposition contains three subbands, whose wavelet coefficients contain directional information. The energy of the coefficients in each subband is an ideal measure of the their strength:

$$E_d(i, j) = \sum_{i=1}^3 \sum_{j=1}^3 |w_d(i, j)|^2, \quad (3)$$

where $E(i, j)$ is the energy of the wavelet coefficient in a 3×3 window, $w(i, j)$ are the wavelet coefficients and d is the decomposition level. Thus, we calculate the energy of the wavelet coefficients in a 3×3 window in each subband, and rescale to get energy values in the range $[0, 255]$.

3.2. Vector Representation

Next, at each decomposition level, we form 3-dimensional vectors whose components are the coefficient energy values, at the same locale in each of the horizontal, vertical and diagonal subbands, for regions which have been classified as detailed:

$$\mathbf{S}_{ij}^d = [E_{Vd}(i, j), E_{Hd}(i, j), E_{Dd}(i, j)], \quad (4)$$

where \mathbf{S} is the new 3-Dimensional vector, E is the energy at position (i, j) in each subband (\mathbf{V}, \mathbf{H} and \mathbf{D}) at decomposition level d . Thus, the vectors \mathbf{S}_{ij} contain the directional

¹The decomposition is done on a single-channel image. When dealing with color images, the luminance channel can be used.

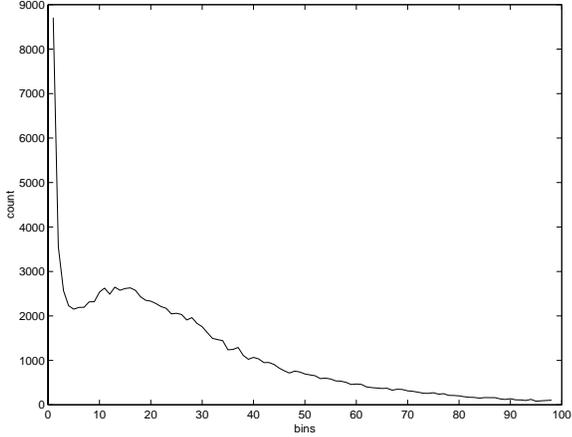


Figure 2: Histogram of activity measure for *zebras* image. Thresholding is done at bin 5.

detail at each location within the extracted detail regions. Figure 4 graphically depicts the process. The motivation behind using the vector representation is to capture the correlation between the subbands of the wavelet decomposition. Even though each subband contains vertical, diagonal or horizontal components, there are still points in the image which can exhibit strong wavelet coefficient energy in more than one subband.

3.3. Histogram Calculation

Once the vectors \mathbf{S}_{ij} for the detailed regions have been calculated for each decomposition level, the *directional detail histograms* are then calculated by counting the number of occurrences of all possible \mathbf{S}_{ij} . We then quantize \mathbf{S}_{ij} to uniform step sizes of $(7,7,7)$, for storage and computational considerations, resulting in a histogram with 512 bins. The quantization is a loss of resolution, however, we must keep in mind that our ultimate goal is image retrieval, where a *close match* is what is desired and not an *exact match*.

3.4. Histogram Similarity

For the calculation of similarity between the query index and the database index, we chose to implement the *Histogram Intersection* method [9]:

$$\xi_d(\mathbf{H}, \mathbf{Q}) = \frac{\sum_{b=1}^N \min(H_b, Q_b)}{\sum_{b=1}^N Q_b}, \quad (5)$$

where \mathbf{H} is the detail histogram of the extracted detail regions of the database image, \mathbf{Q} the query histogram, N the number of histogram bins and d is the decomposition level. The overall similarity is then computed as a weighted sum of the similarity measures of each decomposition level:

$$\Xi = \sum_{d=1}^D w_d \xi_d, \quad \sum w_d = 1, \quad (6)$$

where D is the number of wavelet decomposition levels and w_d are weights which can be customized at query time by



Figure 3: Extracted detail regions of *zebra* image.

the user or by the system, depending on what is desired, e.g fine texture, strong edges, etc. For example, if it was desired to query for images which contained regions with fine details, then a greater weight would be assigned to the highest decomposition level for the query. Thus, these weights allow for flexibility in the query based on resolution.

4. RESULTS

We tested our extraction method on an image database of 1000 images which contained a variety of images including scenery, architecture and people. Figure 5 is a query result obtained from our system when it was requested to find images which contained regions with strong horizontal detail activity. We must note here that no other characteristics were taken into account in the retrieval, only directional detail. As can be seen, all the images retrieved comply with the desired query and exhibit regions with horizontal activity, both at a fine level and at a coarser level.

5. CONCLUSIONS & FUTURE WORK

In this paper we have proposed a method for extracting the regions of an image which contain detail using a perceptually based activity measure. We then apply the directional detail histogram technique to these regions to build indices into an image database. This is achieved by first performing wavelet multiresolution decomposition and then building 3-d vectors of the energy coefficients within the classified detail regions. We have shown that our extraction technique, along with the directional detail histogram, provide effective retrieval results. We have not included any other features in our search for this paper, we concentrated only on the detail content. Currently, we are incorporating the color information of the smooth areas and spatial image information into the database indices.

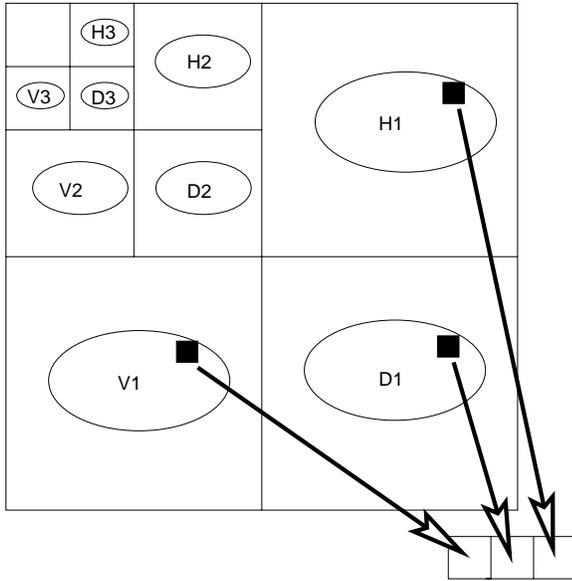


Figure 4: Mapping of detailed regions coefficients into 3-d vectors.

6. REFERENCES

- [1] V. N. Gudivada and V. V. Raghavan, "Content-based image retrieval systems," *Computer* **28**, September 1995.
- [2] W. Niblack, R. Barber, W. Equitz, M. Flickner, Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin, "The qbic project: Querying images by content using color, texture and shape," in *Storage and Retrieval for Image and Video Databases*, M. H. Loew, ed., *Proc. SPIE* **1908**, 1993.
- [3] J. Smith and S. F. Chang, "Visualseek: a fully automated content-based image query system," in *ACM Multimedia Conference*, November 1996.
- [4] M. Stricker and M. Orengo, "Similarity of color images," in *Storage and Retrieval for Image and Video Databases III*, *Proc. SPIE* **2420**, pp. 381-392, 1995.
- [5] W.Y.Ma, Y. Deng, and B. Manjunath, "Tools for texture/color based search of images," in *Human Vision and Electronic Imaging II*, *Proc. SPIE* **3016**, pp. 496-505, 1997.
- [6] D. Androutsos, K. N. Plataniotis, and A. N. Venetsanopoulos, "Image retrieval using the directional detail histogram," in *Storage and Retrieval for Image and Video Databases VI*, January 1998.
- [7] P. E. Trahanias and A. N. Venetsanopoulos, "Directional processing of color images: Theory and experimental results," *IEEE Transactions on Image Processing* **5**, pp. 868-880, June 1996.
- [8] R. N. Shepard, "Toward a universal law of generalization for psychological science," *Science* **237**, pp. 1317-1323, 1987.
- [9] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision* **7**(1), 1991.



Figure 5: Query result when it was requested for images which contained regions with strong horizontal detail. Decreasing similarity from left to right, top to bottom.