Image Content Search by Color and Texture Properties

T. Randen

Høgskolen i Stavanger
4004 Stavanger, Norway
t.randen@ieee.org

J. H. Husøy

Høgskolen i Stavanger
4004 Stavanger, Norway
jonh@hsr.no

Abstract

A new scheme for color and texture feature extraction for image content search is presented. We introduce a scheme using the two chrominance components for color information and a computationally efficient infinite impulse response (IIR) quadrature mirror filter bank (QMF) energy measure of the luminance component for texture information. The color and texture information is combined into one feature vector, and the components are balanced with respect to dimensionality. We illustrate the utility of our features with experiments in searching for a specific color texture in a large database of images. Several different sub-band decompositions are evaluated. Features extracted using a previously published Gabor filter bank are also evaluated against the proposed scheme. We conclude that the proposed scheme outperforms the Gabor features in both quality and complexity.

1 Introduction

One of the major contributions of computers to everyday life is the ability to conduct fast and reliable text pattern searches. Searching non-textual data like images based on contents (e.g., color and texture) was not previously practical, but recent research is making such an application increasingly feasible. In an image database system, features may be computed in advance and the difficult issues are feature storage and comparison. In searches on images not stored in an integrated database, like searches over a network, precomputed features will typically not be available and the feature extraction complexity is of paramount importance.

Query on image content is a relatively new research area, and only a moderate number of experiments on querying by color and texture have been reported. Picard and Minka [1] focus their work on user interaction and adaptive feature selection. Claiming that no single feature extraction scheme is suitable for all image types, they apply a "society of models". Some works on retrieving a set of most similar sub-images based on matched filters [2], Markov random field parameters [3], and Gabor features [4] have been presented. Taking a more global perspective, Pentland et al. [5] and Flickner et al. [6] present complete systems, integrating many aspects of query systems. Except from the work by Liu and Picard [7] on modeling the human visual system, not much work on feature tuning for image queries have been reported.

In this paper, a feature extraction scheme for color and texture content search is suggested and tested on a large number of real world images. In Section 2 the feature extraction and classification scheme is proposed and in Section 3 experiments indicating the utility of the approach are given.

2 Method description

Many filter bank approaches to texture recognition give robust and good classification results [9]. The most important issue in the selection of a filter bank is not the filter basis, but the way the frequency band is partitioned by the filter bank [9]. A computationally attractive critically sampled quadrature mirror bank of IIR filters [9, 10] is the back bone of the feature extraction scheme used for image content search. This filter bank is essentially a bank of band-pass filters covering the two-dimensional spatial frequency spectrum by various decompositions, and some possible decompositions are illustrated in Fig. 2. Due to space limitations, the details [10] are omitted, but since the sub-band decomposition is more important than the filter basis [9], similar results are expected from many different sub-band filter bases.

Furthermore, color is very important for many query patterns. In this approach, the color is represented in the feature space by the two chrominance components of the NTSC color representation [11], while the texture features are extracted from the in-

Figure 1: Test image set, downloaded from [8].

tensity component. The chrominance components were smoothed to reduce local variations.

How are the color and texture features combined in the classification? In our approach we append the color and texture features into one feature vector. In order to avoid that the large-dimensional texture feature sub-space dominates the two-dimensional color sub-space, a balanced Euclidean distance measure is applied. This measure (squared) is given by

$$
\begin{aligned}
d = & \frac{1}{2^2} \left(\mathbf{v_C} - \mathbf{m_C}\right)^T \operatorname{diag}\left(\mathbf{C_C}\right)^{-1} \left(\mathbf{v_C} - \mathbf{m_C}\right) \\
& + \frac{1}{N^2} \left(\mathbf{v_T} - \mathbf{m_T}\right)^T \operatorname{diag}\left(\mathbf{C_T}\right)^{-1} \left(\mathbf{v_T} - \mathbf{m_T}\right),
\end{aligned} \tag{1}
$$

where the C and T subscripts denote the sub-vectors and sub-matrices corresponding to the color and the texture features, respectively and N is the dimensionality of the texture feature vector. By this normalization, it may easily be shown that linear scaling of the distance in the color and linear scaling of the distance in the texture domain affects the total distance measure identically. Next a codebook is designed by clustering the design feature vectors by the k-means clustering algorithm [12].

In an image database system, query times of a few seconds, even if the database consists of tens of thousands of images, are desirable and have been reported [1]. Although the complexity of the proposed feature extraction scheme is very low [9] even loading all those images will violate the complexity constraints. A possible solution to this problem is to store a small amount of feature information for each image and apply a fast similarity match [1, 13, 14]. Another alternative, which will be illustrated by experiments in the next section, is to classify the regions of the images off-line and store class labels. Hence, if a user requests images with hair, the class labels are examined and images with hair are returned. The problem of integration with higher-level information, e.g. context information, will not be treated here.

3 Experiments

In the experiments herein, an example of indexing regions with grass using the 98 vacation images in the MIT VisionTexture database [8] is given. Some of the MIT images are shown in Fig. 1. One of the most challenging tasks in the experiments was the selection of good training image regions, and eventually the images 008, 013, 020, 022, 024, 036, 043, 071, 078, and 095 were used as training images. When testing on any of the training images, the leave-one-out scheme was used.

How do we determine which sub-band decomposition is most appropriate for the problem? The selection of sub-band decomposition is clearly a tradeoff between the "curse of dimensionality" [15] and the need for adequate feature separation. In the experiments, several sub-band decompositions were tested along with the dyadic Gabor features [4]. The distance distributions along with the corresponding overlaps were computed and are shown in Fig. 3. From the overlap percentages, we can see that the filter bank structure of Fig. 2(d) yields the most powerful features. We also see that it compares favorably with the

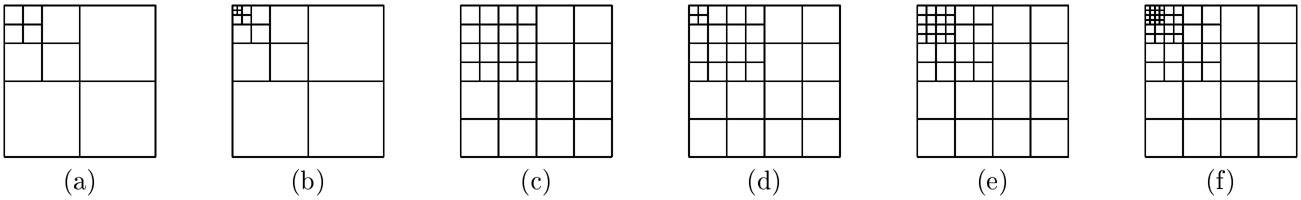

Figure 2: Sub-band decompositions of the filter banks.

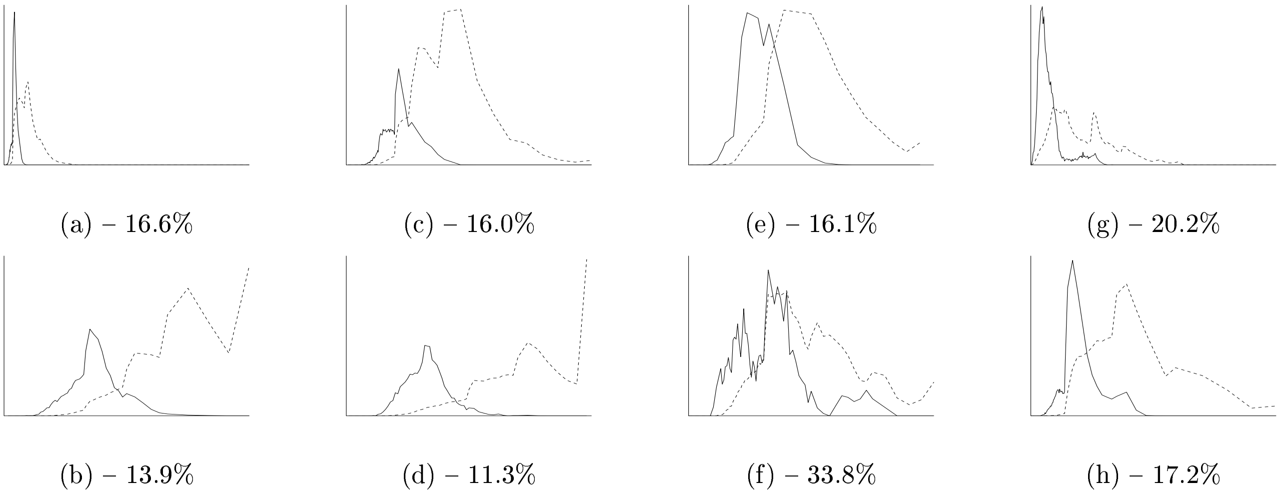

Figure 3: Histograms of feature distance to closest grass codebook vector for grass (solid) and non-grass (dashed) for different features. Minimal overlap is optimal. (a) to (f) corresponds to the filter structures in Fig. 2, (g) corresponds to no texture filter bank at all (using only the luminance in addition to the color), and (h) to the Gabor filter bank, [4]. The percentages of histogram overlaps are indicated.

computationally *much* more complex Gabor [4] features, Fig. 3(h). In Fig. 4 some classification results with this feature extractor are shown. Different intensities indicate different thresholds for the distance measure. As can be seen (if the images in Fig. 1 are studied in detail), the system detects regions with grass well. There are some false detections of walls and leaves, but from a pure texture and color point of view, the falsely detected areas are very similar to certain types of grass. The experiments were restricted to grass due to the effort in building and testing a good training dataset with the prototype tools.

The results also tell us that as the dimensionality of the texture features becomes too large (e.g. Figs. 2(f) and 3(f)), the performance degrades, indicating that it becomes increasingly difficult to design a good codebook with the increased number of feature variables – the "curse of dimensionality" [15].

4 Summary

In this paper, a feature extraction and classification scheme for color-texture pattern search in digital image databases has been proposed. Color was represented by the two chrominance components in the NTSC scheme, while texture was represented by QMF filter bank features. The color and texture features were combined and their contribution balanced using a modified Euclidean distance measure. The scheme was tested on the problem of labeling/finding regions in images with grass in a database of vacation images with success.

Acknowledgments

This work was partially supported by Norwegian Research Council grant 110914/410 and by a grant from the Norwegian Society for Image Processing and Pattern Recognition. The authors would like to thank

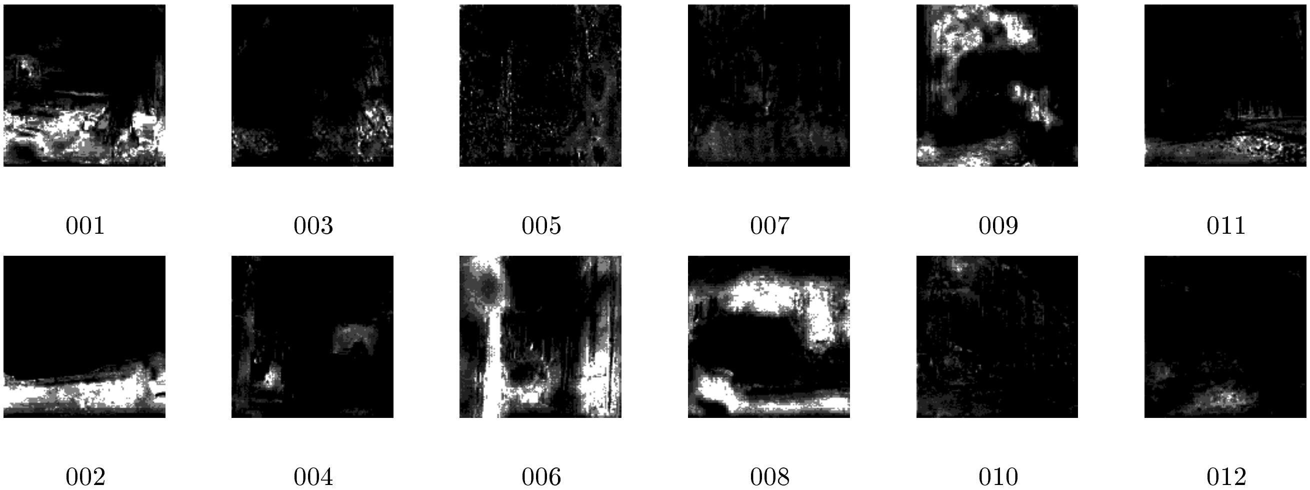

Figure 4: Segmentation results for the images of Fig. 1.

Aditya Vailaya at Michigan State University for useful discussions during the project.

References

[1] R. W. Picard and T. P. Minka, "Vision texture for annotation," *J. Multimedia Systems*, vol. 3, pp. 3–14, 1995.

[2] R. W. Picard and T. Kabir, "Finding similar patterns in large image databases," in *Proc. Int. Conf. Acoust. Speech, Signal Proc.*, vol. V, (Minneapolis, MN), pp. 161–164, Apr. 1993.

[3] G. L. Gimel'farb and A. K. Jain, "On retrieving textured images from an image database," *Pattern Recognition*, vol. 29, no. 9, pp. 1461–1483, 1996.

[4] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 18, pp. 837–842, Aug. 1996.

[5] A. Pentland, R. W. Picard, and S. Sclaroff, "Photobook: Tools for content-base manipulation of image databases," in *Proc. SPIE Conf. on Storage and Retrieval of Image and Video Databases II*, (San Jose, CA), pp. 34–47, Feb. 1994.

[6] M. Flickner, H. Sawhney, W. Niblack, J. Sahley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by image and video content: The QBIC system," *IEEE Computer*, vol. 28, pp. 23–32, Sept. 1995.

[7] F. Liu and R. W. Picard, "Peridicity, directionality, and randomness: Wold features for image modeling and retrieval," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 18, pp. 722–733, July 1996.

[8] MIT Vision and Modelling Group. http://www.media.mit.edu/vismod/.

[9] T. Randen, *Filter and Filter Bank Design for Image Texture Recognition and Automated Seismic Interpretation.* PhD thesis, The Norwegian University of Science and Technology, Oct. 1997.

[10] T. Randen and J. H. Husøy, "Multichannel filtering for image texture segmentation," *Optical Engineering*, vol. 33, pp. 2617–2625, Aug. 1994.

[11] A. K. Jain, *Fundamentals of Digital Image Processing.* Englewood Cliffs: Prentice-Hall, 1989.

[12] R. Schalkoff, *Pattern Recognition: Statistical, Structural and Neural Approaches.* New York: Wiley, 1992.

[13] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger, "The R*-tree: An efficient and robust access method for points and rectangles," in *Proc. ACM SIGMOD International Conference on Management of Data*, (Atlantic City, NJ), pp. 322–331, May 1990.

[14] D. A. White and R. Jain, "Similarity indexing with the SS-tree," in *Proc. 12th IEEE International Conference on Data Engineering*, (New Orleans, Louisiana), Feb. 1996.

[15] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis.* New York, NY, USA: Wiley, 1973.