From: KDD-97 Proceedings. Copyright © 1997, AAAI (www.aaai.org). All rights reserved.

# Image feature reduction through spoiling: Its application to multiple matched filters for focus of attention

**Timothy M. Stough and Carla E. Brodley**

School of Electrical and Computer Engineering
Purdue University, West Lafayette, IN 47906
{stough,brodley}@ecn.purdue.edu

**Abstract**

In order to be of use to scientists, large image databases need to be analyzed to create a catalogue of the objects of interest. One approach is to apply a multiple tiered search algorithm that uses reduction techniques of increasing computational complexity to select the desired objects from the database. The first tier of this type of algorithm, which is often called a focus of attention (FOA) algorithm, selects candidate regions from the image data and passes them to the next tier of the algorithm. In this paper we present a new approach to FOA that employs multiple matched filters (MMF), one for each object prototype, to detect the regions of interest. The MMF are formed using k-means clustering on a set of example image patches identified by experts. An innovation of the approach is to radically reduce the dimensionality of the feature space, used by the k-means algorithm, by spoiling the sample image patches. This research was motivated by the need to detect small volcanos in the Magellan probe data from Venus. An empirical evaluation of the approach illustrates that MMF plus average filter perform better than a single matched filter for high true detection rates.

**Keywords:** image database search, focus of attention, feature selection, k-means, multiple matched filters

## Introduction[1]

Many image databases are so large that a comprehensive search, even an automated search for the objects of interest, may prove impossible. In such cases it is useful to do a multiple tiered search, in which each tier processes the data from the previous tier and selects candidate regions to pass on to the next tier (Fayyad, Haussler & Stolorz, 1996). The final tier produces a list of detections. This approach is designed such that each tier receives a decreasing amount of data, allowing each successive tier to increase in complexity and discrimination power. Our research addresses the first tier, focus of attention (FOA), which should be computationally simple and select promising regions of the image.

[1]Copyright ©1997, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

This paper describes a new method for learning an FOA procedure that can be used to search for heterogeneous objects of interest in an image database. The approach was developed in the domain of searching for small volcanos in SAR images of Venus taken by the Magellan space probe. Magellan mapped 95% of the surface of Venus and returned 30,000, $1024 \times 1024$ pixel SAR images. Our approach is based on the hypothesis that the set of all volcanos is heterogeneous, that there are prototypes that represent each class of volcano, and that such prototypes can be learned. Using multiple matched filters (MMF) allows more volcanos to be found by the FOA algorithm, because each filter specializes in a specific type of volcano. The MMFs are produced using the k-means algorithm to cluster the training images into $k$ classes and a matched filter is produced for each class. In order to reduce the dimensionality of the feature vector describing each image patch we radically scaled down the training image patches.

In the remainder of this paper we first describe the procedure used to form the FOA algorithm. We then present the results of an empirical comparison to the original approach of a single matched filter, which illustrates that the MMF approach outperforms the single matched filter. We also present results combining the MMF and average filter. The average filter detects the bulk of the objects and leaves the less common ones to be picked up be the MMFs. Simply put, if it is necessary to find all of the objects, the MMF plus average filter finds them with fewer false hits than the single matched filter method.

## Finding Object Prototypes

The single matched filter FOA algorithm uses a normalized average of the training images to form a single filter. By using a single filter the approach assumes that one can characterize all of the training images in a single template. However, there are several classes of volcanos that appear visually dissimilar (see Figure 1). Therefore, it seems likely that separating these classes and forming a filter for each class will create an FOA algorithm that is capable of detecting more volcanos

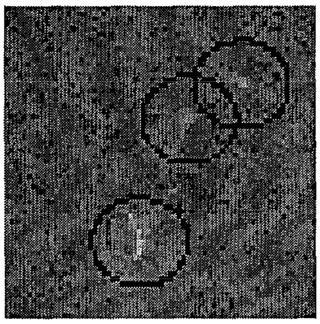

Figure 1: An Example of Heterogeneous Volcanos

than a single matched filter. The FOA algorithm operates in two stages. First the system is trained using ground truth provided by experts to create a set of MMFs. Then the filters are applied to images to produce a detection list.

## Reducing Dimensionality and Clustering

The labeled data is in the form of a number of image patches that contain the region of the training image that surrounds the object of interest. To form the MMF we can cluster these image patches into $k$ clusters and then form a filter for each cluster. Clustering using the pixel values of the $l \times l$ training image patches as a feature vector is not computationally feasible (Fayyad, Haussler & Stolorz, 1996), because an $l \times l$ image patch results in a feature space that requires a large training set to cluster with k-means (for $l = 15$, there are 225 features which require far more data to produce good clusters than we have available). We present a solution to this problem that radically scales down the image before clustering. Specifically, the $l \times l$ image is scaled down to an $m \times n$ image and then its pixel values are used as the feature vector. Scaling down is referred to as spoiling because scaling down is irreversible. Moreover, empirical tests show that clustering in the full $l \times l$ dimensional feature space produces clusters that less well separated than those produced by clustering in the spoiled feature space.

In image search, treating the pixel value features as a simple vector obscures some obvious ways of reducing the number of features. Features can be calculated based on regional properties of the image patch (Rosenfeld & Kak, 1982) and may capture some trends in a lower dimensionality feature space that were not apparent in the original high dimensionality space. One very simple regional property that can be calculated from the image patches is the average intensity of a region. We observed that, since the volcanos were illuminated from the $x$ direction, there was more intensity information in the $x$ direction. Therefore spoiling to a higher resolution in the $x$ direction preserves more information about the volcano type. Our approach combines both domain knowledge and a clustering metric to determine the degree of spoiling.

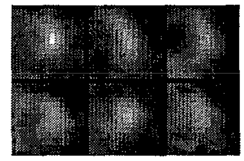

Figure 2: Filters Formed by $2 \times 3$ Spoiled Training Images

To form the object prototypes we first spoil the image and then apply k-means clustering (Duda & Hart, 1973). Applying k-means clustering to the spoiled images results in a number of cluster centers that correspond to spoiled prototype volcanos. Once the spoiled cluster centers have been found, a corresponding $l \times l$ (where $l$ is the size of the original image patch, as discussed above) matched filter must be formed for each of the $k$ classes. Because spoiling is not reversible, all of the spoiled training images are classified into one of the $k$ classes and the $l \times l$ images corresponding to each spoiled image in each class are averaged into a matched filter for that class. This results in $k$, $l \times l$ matched filters; one filter for each cluster center. Figure 2 shows the filters formed using a $2 \times 3$ feature space for clustering and $k = 6$ clusters. The number of clusters, $k$, was chosen empirically using a pair-wise cross-correlation metric.

## Producing a Single Detection List

Once trained, the FOA algorithm can be used to process images and collect regions of interest (ROIs) to be passed on to the next tier of the image search system. The product of the FOA algorithm is a detection list reporting the location of all ROIs.

To apply a filter, the system computes the normalized correlation of the filter $f$ with each $l \times l$ patch in the test image. Candidate volcano locations are placed where the matched filter response exceeds a threshold. Any threshold crossings within four pixels are considered to be due to the same object and the highest response is chosen to represent the object (Burl, et al. 1994). The result of this process is $k$ detection lists, one for each matched filter. Although each filter corresponds to a particular class of volcano, many locations are detected by several different filters. It is, therefore, necessary to combine the $k$ detection lists into a single list, dramatically reducing the number of candidate locations that are produced by the FOA algorithm. All of the detection lists are projected back into an empty (zero) image and the non-zero positions are spatially clustered around local filter response maxima. This procedure results in a single unified detection list that is much smaller than the sum of the individual detections lists.

## Experimental Design and Results

The system was tested using a set of thirty-six images labeled by experts. The thirty six images were divided into six sets of six images. A six-fold cross validation was performed using parameter values determined using different training and test sets. We used images from a set distinct from *both* the training and the test sets to tune the parameters. In the experiments we varied the detection threshold from 0.8 to 0.2. For the MMF, we used a single threshold value for all $k$ matched filters. Clearly, a better approach is to customize the threshold for each prototype filter, and future work will address how these can be learned from the training data. A detection returned by the FOA algorithm can either be a true detection (true positive) or a false detection (false positive). Ideally, the FOA procedure should return a list of all of the objects of interest with zero false positives

### MMF versus Single Matched Filter FOA

Statistics were calculated for each of the thirty-six test images for each threshold, 0.2 though 0.8. The true and false hit rates were then averaged across all thirty six of the test images for both the single and MMF methods and normalized by the number of volcanos in the image. To compare results across different test sets we computed the percentage of true detections (true hit rate) and the ratio of false hits to volcanos (false hit rate). The MMF method out-performs the single matched filter for high true hit rates on these images.

### MMF Plus the Average Filter versus the Single Filter FOA

Examination of the initial MMF results lead us to question why the single filter FOA outperformed the MMF at lower true hit rates. We hypothesized that, if the average filter was able to capture a larger percentage of the true detections at a low false hit rate, including an average filter to the MMFs would allow them to pick up the rare objects while the average filter detected the bulk of the objects. In order to test this hypothesis, we performed an experiment wherein the average filter was added to the MMFs at a static threshold of 0.6 (this is a selective threshold) and the threshold of the MMFs was varied from 0.2 to 0.8 as in the above experiment. In Figure 3, the dotted line represents the JPL baseline and the dashed line is the MMFs alone. The unbroken line is the results of combining the MMFs and the average filter. Combining the MMFs and the average filter results in an earlier breakaway from the JPL baseline system. This supports the hypothesis that, at low detection rates, the average filter picks up the majority of the detections leaving the MMFs to detect the rare cases.

### Ability to Achieve 100% Detection

Because, for many applications, it is of paramount importance for the FOA to achieve 100% detection we

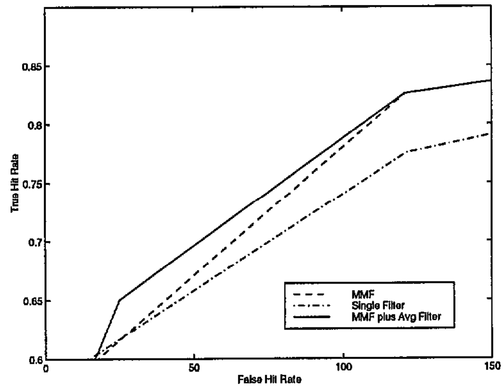


Figure 3: Single and MMF Plus Average Filter FOA

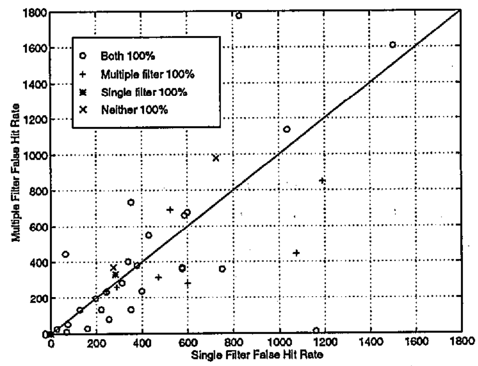


Figure 4: Single and Multiple Matched Filters at 100% Detection

constructed experiments to determine the false hit rate at the point where a true hit rate of 100% is achieved. To understand the difference between the single and MMFs at a 100% true hit rate, we compared their false hit rate. Figure 4 shows a scatter plot of the thirty-six test images; the x-axis is the false hit rate for the single matched filter and the y-axis is the false hit rate for the MMF. For the points marked with '+', '*', and '×', one or both of the methods did not detect 100% of the volcanos at any threshold. In these cases, we choose the maximum false hit rate as the rate at "100%" detection (i.e. the false hit rate for a threshold of 0.2). If a point falls on the line $x = y$, the single and MMF methods have the same false hit rate at 100% detections for that image. If a point falls below the line, the MMF method out-performed the single filter for that image. If a point falls above the line, the single matched filter is better for that image. It can be seen in Figure 4 that the number of points on or below the line is greater than the number of points above the line. This illustrates that the MMF achieves 100% with fewer false hits for more of the test images than the single matched filter does.

## Why Does Spoiling Work?

Spoiling is applied to calculate new, low dimensionality, feature vectors from the original image patches. For this domain, calculating the average intensity of regions of pixels forms a feature vector that captures the essential differences between different types of volcanos. Spoiling works for feature generation whenever the image patches show large scale pixel intensity differences which separate different classes of the objects. If, however, the majority of information necessary to identify the objects resides in individual pixel values, spoiling would obscure the information necessary to form classes.

In order to measure the quality of the clusters formed by the k-means algorithm, we used a metric that measures the ratio of the average inter-cluster distance to the average intra-cluster distance (Fukunaga, 1990). A higher metric indicates better clusters. We applied this metric to clusters formed by k-means in the full dimensionality and low dimensionality (spoiled) feature spaces. Clustering in the full dimensional feature space yielded a lower metric value than clustering in the spoiled feature space (0.92 versus 1.16). This indicates that spoiling allows the formation of better clusters than clustering in the original feature space for this domain.

The essential advantage provided by using spoiling is the reduction of the dimensionality in the training data. In high dimensionality spaces, learning algorithms like k-means clustering require a large amount of data to form an accurate generalization. Reducing the number of features, assuming the information that separates classes is preserved, allows generalizations to be formed using less data.

## Discussion and Conclusions

The MMF approach to the FOA algorithm is a general approach to object detection. While no volcano-specific domain knowledge is necessary for producing MMFs using our approach, different spoiling methods can be investigated the cross-correlation metric and an appropriate feature vector can be determined automatically. This allows our approach to be used without domain knowledge; however, we can take advantage of domain knowledge where it is available to determine the degree of spoiling. If achieving the maximum number of true hits is the most important factor in searching a set of images, the MMF method is superior to the single matched filter. While at lower true hit rates, the single matched filter has fewer false hits, at higher true hit rates, the MMF method has fewer false hits. Because each filter tends to produce a comparable number of hits for a given threshold, $n$ MMFs produce $n$ times as many hits as the single filter for a given threshold. In order for the MMF method to outperform the single matched filter, it needs to detect more volcanos, but have sufficient overlap in false hits that the combination of the detection lists will result in a higher true hit rate for a given false hit rate.

In performing the experiments, a single threshold value was used for all of the MMFs. Some filters have extremely strong responses for non-volcano features of the image. For example, the filter which detects a single point volcano (second row, middle column, in Figure 2) is very responsive to lines in the image as well as small volcanos. This example illustrates that performance can be increased by having a separate threshold for each filter. One possible method for customizing thresholds would be to characterize the response of each filter and probe the test image in order to determine the best threshold before applying the filters. Thresholds could be set based on this probing allowing each filter's threshold to be customized.

The technique of spoiling images to reduce the number of pixels, and the dimensionality of the features, can be applied to other image processing tasks as it is a computationally inexpensive method of feature reduction. Future work could entail testing the performance of spoiling with other learning methods.

## Acknowledgments

This research has been carried out with the support of the Jet Propulsion Laboratory under grant number 960628. We would like to thank P. Smyth, L. Asker, M. Burl, and P. Stolorz at JPL for their help in acquiring and understanding the data and the algorithms used by the JARtool system, C. Codrington for his help in devising the pairwise cross-correlation metric, and T. Lane for his assistance in code development.

## References

Burl, M. C., Fayyad, U. M., Perona, P., Smyth, P., & Burl, M. P. (1994). Automating the hunt for volcanoes on Venus. *Proceedings of the 1994 Computer Vision and Pattern Recognition Conference (CVPR-94)* (pp. 302–309). Los Alamitos, CA: IEEE Computer Society Press.

Duda, R. O., & Hart, P. E. (1973). *Pattern Classification and Scene Analysis.* NY: Wiley & Sons.

Fayyad, U. M., Haussler, D., & Stolorz, P. (1996). KDD for science data analysis: Issues and examples. *The Second International Conference on Knowledge Discovery & Data Mining.* Menlo Park, CA: AAAI Press.

Fukunaga, K. (1990). *Statistical Pattern Recognition (second edition).* San Diego, CA: Academic Press.

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1988). *Numerical Recipies in C: The Art of Scientific Computing.* New York: Cambridge University Press.

Rosenfeld, A., & Kak, A. C. (1982). *Digital Picture Processing.* Academic Press.