REDUCED-COMPLEXITY BIOMETRIC RECOGNITION USING 1-D CROSS-SECTIONS OF CORRELATION FILTERS

Jason Thornton and *B.V.K. Vijaya Kumar*
Dept. of ECE, Carnegie Mellon University

ABSTRACT

Correlation filters are an attractive image processing technique for object recognition. They can provide the necessary recognition accuracy for many applications, but it would be desirable to reduce the complexity of the correlation filter algorithm (in terms of computation and storage space). This is especially true for biometric identification tasks, where multiple correlation filters must be tested against a single image. We propose an algorithm for match metric computation that trades a (usually minor) degradation in accuracy for an orders-of-magnitude complexity reduction. This algorithm analyzes 1D cross-sections of the frequency domain in which the filter is applied. We compare our proposed technique to the standard technique using a dataset of face images.

1. INTRODUCTION

The design of correlation filters for object recognition within images has been well studied [1]. These filters are used for a variety of tasks, including automatic target recognition [2], road sign detection [3], real-time tracking of moving objects [4], and terrain mapping [5], among other things. Recently, correlation filters have been suggested for biometric recognition [6]. For example, the authors have applied correlation filters to the task of iris recognition [7]. One possible issue with the use of correlation filters is that at least two 2D fast Fourier transforms (FFTs) are needed to form one correlation output, and 2D FFTs are computationally expensive. In many cases, it is desirable for the recognition process to be computationally much faster and require less storage space, at an acceptable trade-off in accuracy. This is the motivation for the reduced-complexity method we present in this paper.

This paper is organized into the following sections: Section 2 gives a brief overview of how standard correlation filters work. Section 3 suggests how recognition may be accomplished using only a part of the frequency domain; in particular, when the task is limited to single object recognition, as is the case in biometric recognition. Section 4 presents the proposed algorithm. Section 5 presents test results that compare both methods, and Section 6 gives conclusions.

2. CORRELATION FILTERS

A correlation filter is designed specifically for recognition of one object class; i.e., to produce a strong correlation peak when applied to any image containing a member of its object class. The filter takes the form of a complex-valued 2D array in the frequency domain, and is applied as shown in Figure 1:

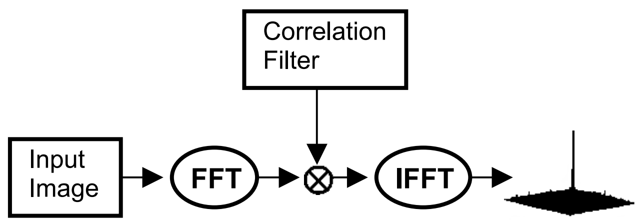

Figure 1. Applying a correlation filter

The input image is decomposed into its spatial frequencies by the computationally efficient fast Fourier transform (FFT) and multiplied by the correlation filter. The result is referred to in this paper as the filtered frequency plane. Then this filtered frequency plane is transformed back to the spatial domain by the inverse fast Fourier transform (IFFT), to produce what is called the correlation plane.

If the input image is authentic (i.e., it contains a member of the correlation filter's object class), the correlation plane should contain a sharp correlation peak, as in Figure 1. If the input image is an imposter (does not contain a member of the object class), there should be no such peak. The presence of a peak, and therefore a match score, is computed by some metric such as the ratio of peak to correlation plane energy (PCE).

There are different types of correlation filters, designed to meet different performance criteria [1]. For example, filters have been developed with built-in tolerance to additive noise, rotation, or scale. This paper does not focus on such design issues, but on a reduced-complexity algorithm that can be applied to all correlation filters.

3. FREQUENCY DOMAIN CROSS-SECTIONS

In the filtering process, let $c(r,s)$ represent the M by N correlation plane produced as output, and let $C(k,l)$ represent its frequency domain equivalent, the filtered frequency plane. Their relation can be expressed as

$$C(k,l) \;=\; \sum_{r=0}^{M-1}\sum_{s=0}^{N-1} c(r,s)e^{j\mathbf{v}_{rs}^{t}\mathbf{w}} \qquad (1)$$

where

$$\mathbf{w} = [k,\ l]^{t} \qquad \mathbf{v}_{rs} = [-r\tfrac{2\pi}{M},\ -s\tfrac{2\pi}{N}]^{t} \qquad (2)$$

The filtered frequency plane $C(k,l)$ is clearly a weighted sum of complex exponentials, with the values of the correlation plane $c(r,s)$ corresponding to the weights.

In many biometric recognition applications, the input scene is known to contain a single biometric signature. For example, in face verification applications, the input may contain one face image. In such a case, the ideal correlation plane for a match forms a single peak, with small values in the rest of the plane. We assume single-peak applications in the development of this algorithm. Let (p_1, p_2) represent the location of the peak value in the correlation plane. Define external energy E_{ext} as the energy of the plane outside of the peak:

$$E_{ext} = \sum_{(r,s)\neq(p_1,p_2)}\sum \left|c(r,s)\right|^2 \qquad (3)$$

The better the match, and the sharper the correlation peak, the smaller the E_{ext} is relative to the energy of the peak value. As the correlation plane approaches its ideal delta function shape, with zero external energy, the sum in Eq. (1) reduces to

$$\lim_{E_{ext}\to 0} C(k,l) \;=\; c(p_1,p_2)\, e^{j\mathbf{v}_{p_1p_2}^{t}\mathbf{w}} \qquad (4)$$

In the limiting case, all weights of the sum (except for the weight corresponding to the peak) decrease until they drop out of the sum, leaving a single 2D complex exponential.

In practice, the correlation peak is always surrounded by sidelobes, but they do not contribute as much to the weighted sum as the peak does. Therefore, any input that matches the filter should have one dominant complex exponential in its filtered frequency plane, as in Figure 2; input that does not match the filter should have no dominant complex exponential.

Every 2D complex exponential has the useful property that any 1D cross-section taken of it is also a complex exponential, with the same magnitude, i.e.,

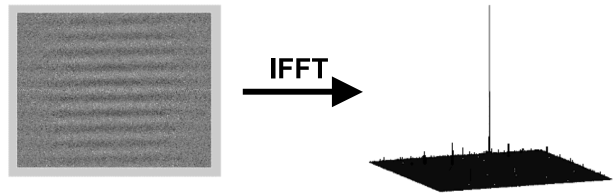

Figure 2. Example of filter output for match
Left: filtered frequency plane (real part), containing a dominant complex exponential. Right: correlation plane, with distinct peak. Both are indicative of a match.

$$e^{j(v_1w_1+v_2w_2)}\Big|_{w_1 \,=\, \alpha w_2+\beta} \;=\; \mu e^{j(v_1\alpha+v_2)w_2} \qquad (5)$$

where μ is a constant phase term. So taking a cross-section of the filtered frequency plane allows us to make the following substitution for the complex exponentials in Eqs. (1) and (4):

$$e^{j\mathbf{v}_{rs}^{t}\mathbf{w}} \xrightarrow{\text{cross-section}} \mu_{rs}e^{j\hat{v}_{rs}w} \qquad (6)$$

with μ_{rs} a constant phase term. The frequency of the complex exponential changes from 2D vector $\mathbf{v}_{rs}$ to scalar $\hat{v}_{rs}$ which is a linear projection of vector $\mathbf{v}_{rs}$ depending on the slope α of the cross-section. Substituting Eq. (6) into Eq. (4) gives the following relation for a cross-section of the filtered frequency plane:

$$\lim_{E_{ext}\to 0} C_{cs}(w) \;=\; c(p_1,p_2)\,\mu_{p_1p_2}e^{j\hat{v}_{p_1p_2}w} \qquad (7)$$

That is, when approaching the ideal filter output for a match, any cross-section of the filtered frequency plane approaches a single 1D complex exponential. As a result, the relative dominance of the largest exponential component in a cross-section can be taken as a match metric.

A simple way to measure the spectral components of a cross-section is to take its 1D FFT; then a ratio of peak energy to external energy quantifies the dominance of one component over all others. In this way, any cross-section of the filtered frequency plane can be used to quickly compute a match metric.

To clarify the effect of using frequency cross-sections, consider the Fourier slice theorem. We know that taking a cross-section in frequency is equivalent to projecting the spatial domain onto a line (the orientation of this line is determined by the orientation of the cross-section). When the correlation plane is projected onto a line, the correlation peak remains a peak, although in one dimension and with reduced sharpness. Therefore, each cross-section will yield a different match metric, depending on the angle of the correlation plane projection.

4. PROPOSED ALGORITHM

The algorithm relies on multiple cross-sections, so the number and orientation of the cross-sections must be decided beforehand, for consistency.

Correlation filter design / storage:

- Design a correlation filter from training images in the standard way, using any filter type.
- Take cross-sections of the filter according to predetermined scheme, and store these in memory.

Correlation filter application - to an image transform:

- Take cross-sections of image transform.
- For each cross-section:
  - Multiply image cross-section by the corresponding correlation filter cross-section.
  - Compute 1D FFT of the product, and take its magnitude.
  - Find the peak value, and divide the energy of the peak by the energy of all other values. This computes a match score.
- Multiply the match scores of all individual cross-sections to get the final match metric.

As with the standard correlation filter algorithm, object recognition is accomplished by comparing the match score against some threshold, or against other match scores.

5. TESTING AND RESULTS

We tested both correlation filter algorithms on the task of face recognition. The purpose of the experiments was not to find the optimal correlation filter for the task, but to evaluate the performance of the proposed algorithm relative to the standard algorithm.

The dataset of face images was taken from the Advanced Multimedia Processing Lab at CMU [8]. It consists of 13 face classes, with 74 images (each image representing a different expression) in each class. The face images are cropped and vary in expression, as shown in Figure 3.

From each class, 8 images were secluded as training images, leaving 66 for testing. The 8 training images were used to design a correlation filter that would recognize that particular face. To create the filters, we used the Optimal Trade-off Synthetic Discriminant Function (OTSDF) filter design, given in [9]. This is a common correlation filter because it gives good discrimination over the training set as well as noise tolerance. Other correlation filters could have been used as well. In our implementation of the proposed algorithm, we took 4 cross-sections of the filters: one vertical, one horizontal, and two diagonal.

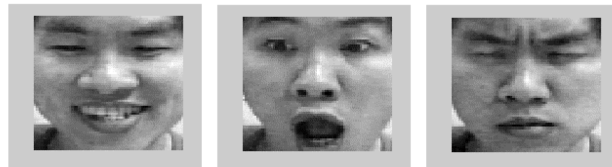

Figure 3. Sample images from one face in the dataset

5.1. Recognition results

Each of the 13 filters has 66 authentic test images (from its own face class) and 12x66 = 792 imposter test images (from all other face classes). Every filter was applied to all available test images to produce match scores. Then a single threshold for recognition determines the overall False Accept Rate (FAR) and False Reject Rate (FRR). (Note that this is a more difficult recognition task than allowing the threshold to vary by class.) Figure 4 shows the lower left corner of the ROC curve for both algorithms, which plots FAR vs. FRR as the threshold changes.

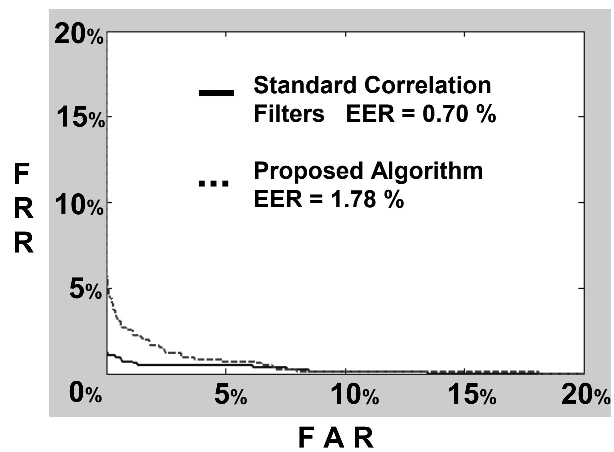

Figure 4. False Accept Rate (FAR) vs. False Reject Rate (FRR) for standard 2-D correlation filter method and the proposed method using 1-D cross-sections

As expected, the standard correlation filter algorithm shows better recognition performance. The proposed algorithm results in a (somewhat marginal) sacrifice in accuracy, with the Equal Error Rate (EER) increasing from 0.70 % to 1.78 %. However, this sacrifice allows for significant savings in computation and storage.

5.2. Timing and storage results

We also tested the computation time required per filter; i.e., the computation that must be repeated for every filter applied to an image transform. This does not include the original 2D FFT that computes the image transform, as this is a one-time computation which is not repeated for every filter.

The tests were conducted on randomly generated images of different sizes, and Table 1 shows the results. The last row of the table gives a ratio of computation time for the standard algorithm to computation time for the proposed algorithm. Note that there is a definite timing advantage when using the proposed algorithm, and that this advantage increases with image size.

	im. size 100 x 100	im. size 200 x 200	im. size 400 x 400	im. size 800 x 800
Standard algorithm	6.4×10^{-3}	0.026	0.14	0.60
Proposed algorithm	7.2×10^{-4}	9.2×10^{-4}	1.3×10^{-3}	1.9×10^{-3}
Speed-up factor	8.9	28.3	107.7	315.8

Table 1. Computation Times (in seconds)

Table 2 compares the storage requirements of both algorithms, assuming 8 bytes per stored value (using single precision complex numbers). The last row shows the storage advantage of the proposed algorithm for different image sizes. Note that the proposed algorithm has a storage advantage that increases to two orders of magnitude for larger images.

	im. size 100 x 100	im. size 200 x 200	im. size 400 x 400	im. size 800 x 800
Standard algorithm	8×10^{4}	3.2×10^{5}	1.28×10^{6}	5.12×10^{6}
Proposed algorithm	3200	6400	1.28×10^{4}	2.56×10^{4}
Compress-ion factor	25	50	100	200

Table 2. Storage Requirements (in Bytes)

6. CONCLUSIONS

We have proposed a variation of the correlation filter algorithm which processes only a fraction of the frequency domain, in the form of 1-D cross-sections. This algorithm works for single-peak recognition and is close in accuracy to the standard algorithm, but offers substantial gains in computation and storage. Depending on the image size and the number of filters, the proposed algorithm can run hundreds of times faster with much less storage space. At a minimum, the proposed algorithm may be used to produce preliminary match metrics which serve to confirm or reject images which are not borderline.

This research is supported in part by the US Army Research Office (ARO).

7. REFERENCES

[1] B.V.K. Vijaya Kumar and A. Mahalanobis, "Recent Advances in Composite Correlation Filter Designs," *Asian Journal of Physics, Vol. 8*, 1999, No. 3.

[2] L.Rahmati, L. Hassebrook, and B.V.K. Vijaya Kumar, "Automatic Target Recognition with Intensity and Distortion Invariant Hybrid Composite Filters," *Proc. Of SPIE, Vol. 1959*, 1993.

[3] E. Perez and B. Javidi, "Nonlinear Distortion-Tolerant Filters for Detection of Road Signs in Background Noise," *IEEE Trans. On Vehicular Tech., Vol. 51.*

[4] M.W. Eklund, G. Ravichandran, M.M. Trivedi, and S.B. Marapane, "Real-Time Visual Tracking Using Correlation Techniques," *Proc. Of 2nd IEEE Workshop on Applications of Computer Vision*, Dec. 1994, pp. 256-263.

[5] H. Singh and A. Mahalanobis, "Correlation Filters for Texture Recognition and Applications to Terrain Delimitation in Wide-Area Surveillance," *Int. Conf. On Acoustics, Speech, and Signal Processing (ICASSP) 1994*, Adelaide, Australia, April 1994, pp. 153-156.

[6] B.V.K. Vijaya Kumar, M. Savvides, K. Venkataramani, and C. Xie, "Spatial frequency domain image processing for biometric recognition," *Proc. of Intl. Conf. on Image Processing (ICIP), Vol. I*, 2002.

[7] B.V.K. Vijaya Kumar and J. Thornton, "Distortion-Tolerant Iris Recognition Using Advanced Correlation Filters," *Proc. of Workshop on Multimodal User Authentication (MMUA)*, Santa Barbara, CA, 2003.

[8] Advanced Multimedia Processing Lab, Carnegie Mellon University, Pittsburgh, PA; http://amp.ece.cmu.edu

[9] B.V.K. Vijaya Kumar, D.W. Carlson, and A. Mahalanobis, "Optimal Trade-off Synthetic Discriminant Function Filters for Arbitrary Devices," *Optics Letters, Vol. 19*, 1994, pp. 1556-1558.