# Novel Transform Domain Principal Component Analysis (PCA) Techniques and Some Applications: Facial and Automatic Target Recognition

W.B. MIKHAEL, MOATAZ M. ABDELWAHAB, P. RAGOTHAMAN
School of Electrical Engineering and Computer Science
University of Central Florida
4000 Central Florida Boulevard
Orlando, FL 32816, USA.

*Abstract:* - Novel transform domain formulations and implementations of PCA techniques will be given. It will be shown that this new approach results in considerable computational and storage savings while yielding very high accuracy. Useful applications such as Facial and Automatic Target Recognition are given confirming the considerable performance improvement.

*Key-Words:* - Transform Domain Techniques, Principal Component Analysis (PCA), Facial Recognition, Automatic Target Recognition (ATR)

## 1. Introduction

Many algorithms based on Principal Component Analysis (PCA) [1] have been applied to the facial recognition problem. Recently, a 2DPCA method was reported [2]. A Transform Domain Two-Dimensional Principal Component Analysis algorithm (TD2DPCA) applied to facial recognition is presented. By dealing with the images in the transform domain, this algorithm maintains the excellent high recognition accuracy of the recently reported methods. In addition, by reducing the number of coefficients representing the images, the TD2DPCA reduces the storage requirements and computational complexity by 90% and 50 % respectively, compared to the spatial 2DPCA method. Experimental results employing existing facial databases confirm these excellent properties [3].

Quadratic Correlation Filters (QCFs) have been used successfully to detect and recognize targets embedded in background clutter, [6], [7]. Recently, ATR using a Rayleigh Quotient Quadratic Correlation Filter (RQQCF) was proposed that formulates the class separation metric as a Rayleigh quotient that is optimized by the QCF solution, [8], [9]. Consequently, the means of the two classes are well separated while simultaneously ensuring that the variance of each class is small. The RQQCF technique involves the Eigenvalue Decomposition (EVD) or PCA of target and clutter autocorrelation matrices. In this paper, a recently proposed transform domain RQQCF (TDRQQCF). [10], [11], is presented. The technique formulates the autocorrelation matrices in the transform domain that enables a considerable reduction in computational and storage requirements while retaining the high recognition accuracy of the spatial domain RQQCF. The advantages of the scheme are illustrated using sample results obtained from experiments on Infrared (IR) imagery

This paper is organized as follows: Section 2 describes the TD2DPCA method for facial recognition. Section 3 describes the TDRQQCF method for Automatic Target Recognition. Section 4 presents the conclusions.

## 2. Facial Recognition

Several excellent algorithms based on Principal component analysis (PCA) have been developed and applied to facial recognition. The PCA approach is based on finding the vectors that best account for the distribution of facial images within the entire image space. In 1991 Turk and Pentland [1] developed the Eigenfaces method based on the principal component analysis (PCA) [4]. Recently Yang et al [2] proposed the two dimensional PCA (2DPCA) technique, which has many advantages over the PCA method. It is simpler for image feature extraction, better in recognition rate and more efficient in computation. However, it is not as efficient as PCA in terms of storage requirements. Here, we present the transform domain two-dimensional principal component analysis, TD2DPCA, algorithm that represents the images in an appropriate transform domain. Consequently, the computational and storage requirements are greatly simplified as will be shown.

## 2.1 The TD2DPCA Algorithm

In the TD2DPCA method the covariance matrix *S'* for *N* training images $A_i$ of dimensions *mxn* (where *i=1 to N*) is formed in the transform domain using 2DPCA. The proposed algorithm is described as follows.

*Training mode:*

In the training mode the features of the training images are extracted and stored as the feature matrices, $B_i$'s (where *i=1 to N*).

Step 1: A suitable transform (*Tr*), such as the 2D-DCT, is applied to each *m x n* image $A_i$ of the *N* training images, yielding $T_i$ (*i=1 to N*).

$$T_i = Tr\{A_i - \bar{A}\} \quad (1)$$

where $\bar{A}$ is the mean matrix of all the N training images.

Step 2: The transform (*Tr*) is chosen such that the significant coefficients of $T_i$ are contained in a submatrix, $T_i'$, (upper left part of $T_i$) of dimension *n' x n'*. Thus $T_i'$ is used to replace $A_i$ in our algorithm.

Step 3: The covariance matrix *S'* for the N training images is calculated as follows.

$$S = \frac{1}{N}\sum_{i=1}^{N}(T_i)^T(T_i) \quad (2)$$

*S'* is obtained by retaining the upper left submatrix of S containing its significant coefficients. *S'* is of dimensions *n' x n'*.

Step 4: A set of *k* eigenvectors, V= [$V_1$ , $V_2$ … $V_k$ ] corresponding to the largest *k* eigenvalues of *S' is* obtained. The dimension of each eigenvector $V_j$ ( *j =1..k*) is *n'x1*.

Step 5: The feature matrices $B_i$ of the training images are then calculated as

$$B_i = [Y_{1,i}, Y_{2,i}, \ldots Y_{k,i}] \quad (3)$$

where

$$Y_{j,i} = T_i' V_j \quad j = 1,2,\ldots k \text{ and } i = 1,2,..N \quad (4)$$

The $B_i$ matrices are stored.

*Testing mode:*

In the testing mode, a facial image $A_t$ is presented to the system to be identified. The following steps are followed

Step 1 The same transform used in the training mode is applied to $A_t$ which yields $T_t$.

Step 2 The sub matrix $T_t'$ containing the significant coefficients of $T_t$ is obtained (dimension *n' x n'* )

Step 3 The feature matrix $B_t$ for the testing image is then calculated

$$B_t = [Y_{1,t}, Y_{2,t}, \ldots Y_{k,t}] \quad (5)$$

where

$$Y_{j,t} = T_t' V_j \quad j = 1,2,\ldots k \quad (6)$$

Step 4 Distance measures, such as the Euclidean distances, between the feature matrix of the testing image and the feature matrices of the training images are measured. The image corresponding to the minimum distance is identified.

## 2.2 Simulation Results

The TD2DPCA algorithm was applied to the ORL database [5]. The ORL database consists of 400 images of 40 different individuals (10 images each), where pose and facial expressions are varying, Fig.1. Two experiments, I and II, have been applied to the ORL dataset, where all the images are grayscale with 112 x 92 pixels each.

In experiment I, 40 images of 40 different individuals are used for training, and the remaining 360 images are used for testing. A two-dimensional discrete cosine transform (DCT) is applied to the *N* training images. The dimensions of $T'_i$ and the covariance matrix *S'* are *20x20*. The 5 largest eigenvectors of *S'* corresponding to the 5 largest eigenvalues are obtained. In our approach $k$ of only 5 was needed relative to $k = 10$ in other approaches, while even achieving better recognition accuracy. The feature matrices for all the training images are obtained using (3) and (4). The procedure for the testing mode is followed for the 360 testing images. Results are listed in Tables 1.

In experiment II, 5 images per individual are used for training, and the remaining 200 images are used for testing. The Dimensions of $T'_i$ and $S'$ are the same as in the first experiment. Results are listed in Tables 1.

Table 1 shows that the proposed algorithm yields good recognition accuracy compared to the 2DPCA method. In addition, it illustrates the storage requirements, in terms of the dimensions of the feature matrix. It is seen that, for the TD2DPCA, the amount of storage is drastically reduced (by approximately 90%), compared with one of the best available algorithm, 2DPCA. Also it is worthwhile to note that the computational requirements in the training and testing modes compared to number of multiplications are reduced by a factor of 2. Similar results were obtained when the TD2DPCA was applied to the YALE and FERET databases.

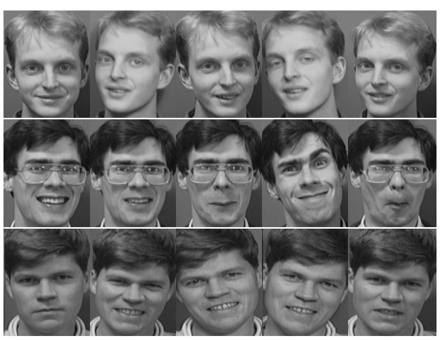


Fig.1. Five samples for 3 individuals in the ORL database.

## 3. Automatic Target Recognition (ATR)

This section describes the spatial domain RQQCF technique briefly for the sake of completeness. This is followed by a description of the TDRQQCF technique for an ATR application. In the following subsections, quantities in lowercase with an underscore are vectors, and quantities in upper case are matrices. Quantities in the transform domain are distinguished by a't' in the subscript.

### 3.1 The RQQCF Technique

In the RQQCF technique, the QCF coefficient matrix T is assumed to take the form,

$$T = \sum_{i=1}^{n} \underline{w_i}\,\underline{w_i^T} \qquad (7)$$

where, $\underline{w_i}$ , $1 \le i \le n$, form an orthonormal basis set. The objective of the technique is to determine these basis functions such that the separation between the two classes, say X and Y, is maximized. The output of the QCF to an input vector $\underline{u}$ is given by,

$$\varphi = \underline{u}^T T \underline{u} \qquad (8)$$

The objective is to maximize the ratio,

$$J(\underline{w}) = \frac{E_1\{\varphi\} - E_2\{\varphi\}}{E_1\{\varphi\} + E_2\{\varphi\}} = \frac{\sum_{i=1}^{n} \underline{w_i}(R_x - R_y)\underline{w_i^T}}{\sum_{i=1}^{n} \underline{w_i}(R_x + R_y)\underline{w_i^T}} \qquad (9)$$

where, $E_j\{.\}$ is the expectation operator over the $j^{th}$ class, and $R_x$ and $R_y$ are the correlation matrices for targets and clutter respectively. Taking the derivative of (9) with respect to $\underline{w_i}$ , we get

$$(R_x + R_y)^{-1}(R_x - R_y)\underline{w_i} = \lambda_i \underline{w_i} \qquad (10)$$

Let,

$$A = (R_x + R_y)^{-1}(R_x - R_y) \qquad (11)$$

Thus $w_i$ is an eigenvector of A with eigenvalue $\lambda_i$. It should be noted that $J(\underline{w})$ is in the form of a Rayleigh Quotient, which is maximized by the dominant eigenvector of $A$.

In practice, $M$ target and $M$ clutter training sub-images, referred to as chips, are obtained from IR imagery. Each chip, having dimensions $\sqrt{n}$ x $\sqrt{n}$ , is converted into a 1-D vector of dimensions $n$ x $1$ by concatenating its columns. Target and clutter training sets of size $n$ x $M$ each, are obtained by placing the respective vectors in matrices. The $n$ x $n$ autocorrelation matrices of the target and clutter sets, $R_x$ and $R_y$ are computed, and used to obtain A according to (10). As a result, the eigenvalues of A vary from –1 to +1. The dominant eigenvalues for clutter, $\lambda_{ci}$, are close to or equal to -1 and those for targets, $\lambda_{ti}$, are close to or equal to +1. The RQQCF coefficients, $w_{ci}$ and $w_{ti}$, are mapped to the corresponding eigenvalues. The spatial domain

RQQCF is then correlated with an input scene to obtain a correlation surface from which the existence and location of the target is deduced. An efficient method to perform the correlation is discussed in the original paper, [8]. In the TDRQQCF though, to identify a data point as target or clutter, the sum of the absolute value of the $k$ inner products of a data point with $w_{ti}$ and $w_{ci}$, $p_t$ and $p_c$, are calculated. If $p_t > p_c$, the data point is identified as a target. Otherwise, it is identified as clutter. We will refer to the absolute values of these inner products as the 'response' of that particular data point.

## 3.2 The Transform Domain RQQCF (TDRQQCF)

The RQQCF technique operates on spatial domain data. Furthermore, each two-dimensional data chip in the spatial domain is converted into a one-dimensional vector by the lexicographical ordering of the columns of the chip. This leads to two interrelated issues. Firstly, the spatial structure in the two-dimensional chip is lost by converting it into a vector as described above. Secondly, the dimensionality of the system is increased considerably. One way to tackle both these issues simultaneously is to synthesize the RQQCF in the transform or frequency domain. Transforms capture the spatial correlation in images, and de-correlate the pixels. Consequently, if the transforms are appropriately selected, they compact the energy in the image in relatively few coefficients. Thus spatial domain data is transformed into an efficient and compact representation.

The TDRQQCF technique proceeds as follows:

1. Each target chip, $C_x$, and clutter chip, $C_y$, is first transformed using the Discrete Cosine Transform (DCT) to obtain $C_{xt}$ and $C_{yt}$. It is seen that most of the energy in $C_{xt}$ and $C_{yt}$ is concentrated in the top left corner. In addition, the distribution of energy for targets and clutter differ from each other.
2. Each $C_{xt}$ and $C_{yt}$ is truncated to an appropriate size and converted to a one-dimensional vector by lexicographically ordering the columns. Thus, vectors of reduced dimensionality compared to the spatial domain case, are obtained. In addition, these vectors are very efficient representations of the spatial domain chips.
3. The autocorrelation matrices, $R_{xt}$ and $R_{yt}$ are computed, and used to obtain $A_t$ according to (5). We note that since the dimensions of the target and clutter vectors are much smaller than in the spatial domain case, the dimensionality of the autocorrelation matrices, $R_{xt}$ and $R_{yt}$, and therefore $A_t$, are correspondingly reduced.
4. The EVD is performed on $A_t$ to obtain the QCF coefficients. The QCF coefficients thus obtained are in the DCT domain.

In addition to reduced dimensionality, there is another advantage to the TDRQQCF. Often, in practice, in applications of techniques such as the RQQCF, one encounters low rank matrices, which give rise to numerical problems. This is because the number of data points available for training is much smaller than the dimensionality of each data point. On the other hand, TDRQQCF overcomes this problem by reducing the dimensionality of the data points.

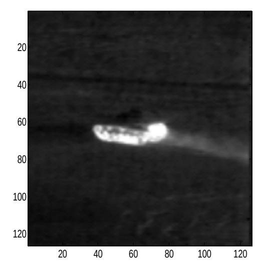



Fig. 2 Sample frame

## 3.3 Simulation Results

The TDRQQCF was tested on various Infrared video sequences provided by Lockheed Martin, Missile and Fire Control, (LMMFC). Sample results are presented from one video sequence to illustrate the performance of the proposed technique. Figure 2 shows sample frame from the video. The video has 778 frames from which 763 target and clutter chips are extracted. The size of each chip is 16x16, i.e., $\sqrt{n}$ *=16 and n=256.* Table 2 shows the average energy in sub-images of different sizes, retaining the low spatial frequency region, of the transformed 16x16 chips. From this table, 85% to 95% of the energy is concentrated in 25% of the transformed chips. Also, the target energy is slightly more compressed in the transform domain. In addition, the energy distribution for target chips is different from that for clutter chips. At this point, if no truncation is performed, and the original RQQCF technique is applied, the same set of eigenvalues as in the case of the spatial domain RQQCF is obtained. For example, twelve dominant eigenvalues (six positive and six negative) for both cases are listed below.

$\lambda_i$ (Spatial domain)

-0.9975 -0.9641 -0.9559 -0.9378 -0.9283 -0.9221
0.9934 0.9938 0.9962 0.9971 0.9981 0.9985

$\lambda_i$ (DCT domain)

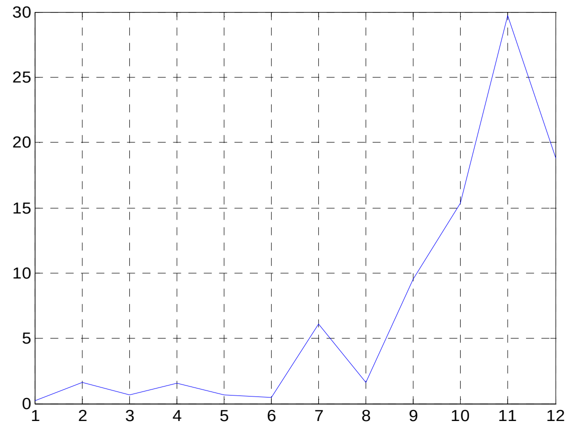
-0.9975 -0.9641 -0.9559 -0.9378 -0.9283 -0.9221
0.9934 0.9938 0.9962 0.9971 0.9981 0.9985

Sample results are presented for the case when the transformed target and clutter chips, $C_{xt}$ and $C_{yt}$ respectively, are truncated to a 8x8 size. This means that $\sqrt{n}=8$ *and* $n=64$. The twelve dominant eigenvalues (six positive and six negative) among the 64 are listed below.

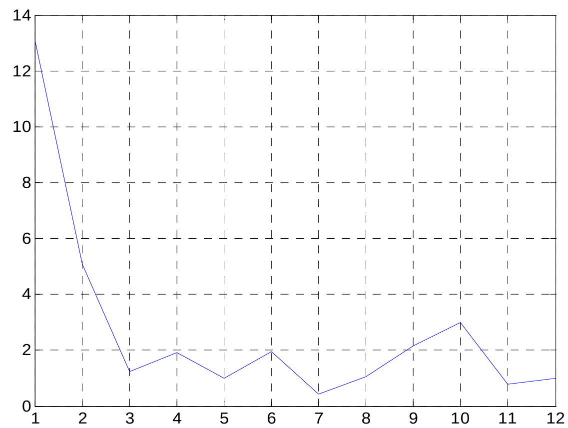
$\lambda_i$ (DCT domain)

-0.9930 -0.8648 -0.7825 -0.7642 -0.7472 -0.6427
0.9642 0.9722 0.9746 0.9878 0.9943 0.9952

Figure 2 shows the response of representative target and clutter vectors versus the index of the eigenvectors in the spatial domain. Figure 3 shows the response of the same representative target and clutter vectors versus the index of the eigenvectors in the DCT domain. A close look at the plots reveals the following: i) The magnitude of each of the responses, inner products, in the DCT domain is much higher than the corresponding magnitude in the spatial domain, ii) The magnitude of ($p_t$ - $p_c$) is also much higher in the DCT domain than in the spatial domain. In other words, separation between targets and clutter is also much higher in the DCT domain than in the spatial domain. This means that the requirements on the threshold to decide if a chip is target or clutter can be relaxed considerably. Although the plots shown are for a randomly chosen data point from the video, it was found that the TDRQQCF consistently produces much larger responses and target-clutter separation than the spatial domain RQQCF for all data points. Similar results are obtained for other test videos as well. This is achieved while maintaining the excellent recognition accuracy of the spatial domain RQQCF - all target and clutter chips in the video are identified correctly.



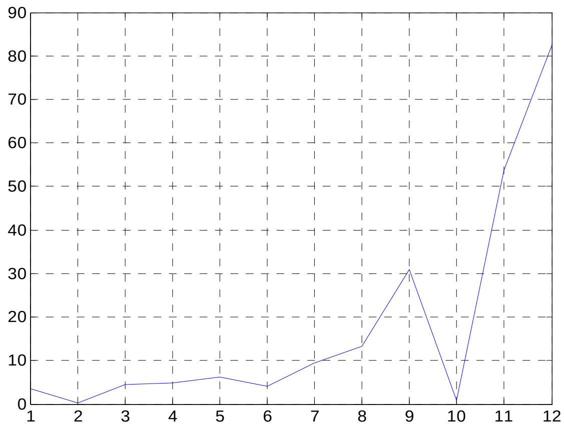

(a)



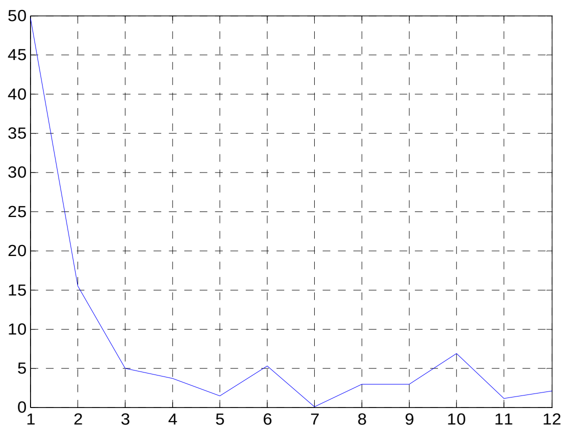

(b)

Fig. 2: Response of (a) a representative target vector, and (b) a representative clutter vector, versus the index of the dominant eigenvectors (spatial domain)

The RQQCF involves the inversion and Eigenvalue Decomposition (EVD) of large matrices. The computational complexity for each of these operations is of the order $O(n^3)$, where '*n*' is the dimensionality of the autocorrelation matrices. On the other hand, by using the TDRQQCF, where compressed representations are used for target and clutter, large savings are obtained. Table 3 compares the spatial domain RQQCF with the TDRQQCF in terms of storage and computational complexity. The computational complexity for the $2^j$ x $2^j$ DCT is approximately $2^j \times (2^{j+1} - j - 2)$, [12]. From Table 2, it can be easily shown that the overall computational complexity including computing the DCT and the storage requirements of the TDRQQCF are still much smaller than the spatial domain RQQCF. In addition, for the TDRQQCF, the storage and computational savings increase as the chip size increases.




(a)




(b)

Fig.3 Response of (a) a representative target vector, and (b) a representative clutter vector, versus the index of the dominant eigenvectors (DCT domain)

## 4. Conclusion:

Transform domain formulations, namely, the TD2DPCA and the TDRQQCF, were given. It was shown using sample results from useful applications such as Facial and Automatic Target Recognition that these formulations result in considerable computational and storage savings while yielding very high accuracy.

*References:*

[1] M. Turk and A. Pentland, "Eigenfaces for Recognition", J. Cognitive Neuroscience, Vol. 3, No. 1, 1991, pp. 71-86.

[2] Yang J., Zhang, D., Frangi, A.F. "Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.26, No(1), January 2004, pp.131-137.

[3] Moataz M. Abdelwahab, and Wasfy B. Mikhael, "Recognition of Noisy Facial Images Employing Transform –Domain Two- Dimensional Principal Component Analysis," The 49th IEEE International MWSCAS, San Juan, Puerto Rico, August 5-8,2006.

[4] M. Kirby and L. Sirovich, "Application of the KL Procedure for the Characterization of Human Faces", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.12, No. 1, Jan. 1990, pp. 103-108.

[5] http://www.orl.co.uk/facedatabase.html.

[6] X. Huo, M. Elad, A. G. Flesia, R. R. Muise, S. R. Stanfill, J.Friedman, B. Popescu, J. Chen, A. Mahalanobis and D. L. Donoho, "Optimal reduced-rank quadratic classifiers using the Fukunaga–Koontz transform with applications to automated target recognition," Automatic Target Recognition XIII, F. A. Sadjadi, ed., Proc. SPIE 5094, 2003, pp. 59–72.

[7] S. R. F. Sims and A. Mahalanobis, "Performance evaluation of quadratic correlation filters for target detection and discrimination in infrared imagery," Optical Engg., 43(8), Aug. 2004, pp. 1705-1711.

[8] A. Mahalanobis, R. R. Muise and S. R. Stanfill, "Quadratic Correlation Filter Design Methodology for Target Detection and Surveillance Applications," Applied Optics, 43(27), 2004, pp. 5198-5205.

[9] R. Muise, A. Mahalanobis, R. Mohapatra, X. Li, D. Han and W. Mikhael, "Constrained Quadratic Correlation Filters for Target Detection," Applied Optics, 43(2), 2004, pp. 304-314.

[10] P. Ragothaman, W. B. Mikhael, R. R. Muise and A. Mahalanobis, "Automatic Target Recognition Using Signal Compression," Applied Optics, (to appear, July 2007).

[11] W. B. Mikhael, P. Ragothaman, R. Muise, A. Mahalanobis, "An Efficient Quadratic Correlation Filter for Automatic Target Recognition," Proceedings of SPIE -- Volume 6566, Firooz A. Sadjadi, Editor, SPIE Defense and Security Symposium, Orlando, Florida, April 9-13, 2007.

[12] E. Feig and S. Winograd, "On the multiplicative complexity of discrete cosine transforms," IEEE Transactions on Information Theory, vol. 38, July 1992, pp. 1387–1391.

| | Experiment I: ORL database | | Experiment II: ORL database | |
|---|---|---|---|---|
| | TD2DPCA | 2DPCA | TD2DPCA | 2DPCA |
| Recognition accuracy | **73.61 %** | **72.77 %** | **92%** | **91%** |
| Dimensions of feature matrix per image | (20x5) | (112x10) | (20x5) | (112x10) |
| Storage requirements for N images | (20x5)xN | (112x10)xN | (20x5)xN | (112x10)xN |

Table 1: Recognition accuracy, Dimensions of feature matrix and number of computations required for experiments I , II. on ORL database.

| **Avg. Energy in →** | **8x8** | **9x9** | **10x10** | **11x11** | **12x12** | **13x13** | **14x14** | **15x15** | **16x16** |
|---|---|---|---|---|---|---|---|---|---|
| **Target chips** | 95.3762 | 96.409 | 97.35 | 97.9219 | 98.4602 | 98.8663 | 99.2888 | 99.5657 | 100 |
| **Clutter chips** | 87.2741 | 89.2236 | 90.841 | 92.2614 | 93.6029 | 94.8183 | 96.5067 | 98.578 | 100 |

Table 2: Avg. energy in different transformed and truncated matrices of the target and clutter sets

| | **RQQCF** | **TDRQQCF** | **% of savings using TDRQQCF*** |
|---|---|---|---|
| **No. of storage locations for chips** | $2 \times M \times \sqrt{n} \times \sqrt{n}$ | $2 \times M \times \sqrt{k} \times \sqrt{k}$ , $k < n$ | *75%* |
| **No. of storage locations for autocorrelation matrices** | $2 \times n \times n$ | $2 \times k \times k$ | *93.75%* |
| **Complexity of Inversion**** | $O(n^3)$ | $O(k^3)$ | *98%**** |
| **Complexity of EVD**** | $O(n^3)$ | $O(k^3)$ | *98%**** |

* For M=763, n=256, k=64; ** # of multiplications; ***approximately

Table 3: Storage and computational complexity of the spatial domain RQQCF versus that for the TDRQQCF, [12].