Handwritten Digit Recognition based on DCT features and SVM Classifier

Bouchra El qacimy*
Laboratory LRGE, ENSET of Rabat,
Mohamed V Souissi University,
Rabat, Morocco
elq.bouchra@gmail.com

Mounir Ait kerroum
Laboratory LARIT, Ibn Tufail University,
Faculty of Science, ENCG of Kenitra,
Kenitra, Morocco
maitkerroum@gmail.com

Ahmed Hammouch
Laboratory LRGE, ENSET of Rabat,
Mohamed V Souissi University,
Rabat, Morocco
hammouch_a@yahoo.com

***Abstract*—Handwritten digit recognition is an active topic in optical character recognition applications and pattern learning research. However, the extraction of informative features from handwritten digits for recognition task remains the most important step for achieving high accuracy. This work investigates the effectiveness of four feature extraction approaches based on Discrete Cosine Transform (DCT) to capture discriminative features of handwritten Digits and compare it to classical PCA. These approaches are: DCT upper left corner(ULC) coefficients, DCT zigzag coefficients, block based DCT ULC coefficients and block based DCT zigzag coefficients. The coefficients of each DCT variant are used as input data for Support Vector Machine Classifier to evaluate their performances. The objective of this work is to identify the optimal feature extraction approach that speeds up the learning algorithms while maximizing the classification accuracy. The results have been analysed and compared in terms of classification accuracy and reduction rate and the findings have demonstrated that the block based DCT zigzag feature extraction yields a superior performance than its counterparts.**

Keywords*—*Handwritten Digit Recognition, Feature Extraction, DCT, SVM Classifier.

I. INTRODUCTION

Handwritten digit recognition is an active topic in pattern recognition research in view of its numerous applications such as postal mail sorting, bank check processing, form data entry, etc. For these applications, the accuracy and speed of digit recognition is crucial to the overall performance [1]. However, the extraction of informative features from handwritten digits for recognition task remains one of the most important steps for achieving high accuracy. In the literature, various feature extraction methods have been explored for handwritten digit recognition [2] [3] [4], and the discrete cosine transform(DCT) has been widely used in pattern recognition problems. In his work on word based recognition system [5], Alkhateeb used DCT as feature extraction method and his results showed that DCT yield to a higher recognition rate than its counterparts. In the same vein, Lawgali [6], in his work on Arabic isolated character recognition, compared DCT to discrete wavelet transformation (DWT) and the results showed the effectiveness of DCT features to lead to a better recognition rate. DCT is also used in face recognition [7], video text detection [8], carplate recognition [9] [10] and Iris recognition [11]. However, little research addressed the issue of what is the most effective method to retain DCT coefficients for handwritten digit classification. In this work, we investigate the effectiveness of four feature extraction approaches based on Discrete Cosine Transform to capture discriminative features of handwritten digits. These approaches are:
- DCT upper left corner(ULC) coefficients: in this approach, we retain the DCT coefficients in the upper left corner of the DCT matrix.
- DCT zigzag coefficients: the DCT coefficients are extracted in a zigzag fashion.
- Block based DCT ULC coefficients: the input image is partitioned in blocks, and the DCT upper left corner coefficients are retained for each block.
- Block based DCT zigzag coefficients: in this approach we apply DCT on each block and retain the N significant coefficients in a zigzag fashion in order to form our feature vector.

To evaluate the performance of each approach in terms of classification accuracy and reduction rate (i.e. how many features are retained), the DCT coefficients of each approach are used as input data for Support Vector Machine Classifier and compared to the principal components extracted using PCA. The objectif of this work is to identify the optimal feature extraction approach that speed up the learning algorithms while maximizing the classification accuracy. The Database retained for this work is the MNIST Dataset [12] that we will describe in more details in the next section. The rest of this paper is organised as follows: in section III we give a brief overview of the DCT transformation and its different variants we used for dimension reduction. Next, in section IV we present the SVM classifier retained for this work before discussing the results in section V. Finally, the conclusion gives a brief summary and addresses some perspective for future work.

II. DATA DESCRIPTION

The Database retained for this work is the MNIST database which is a reference database of handwritten digits [2] that have been widely used in handwritten digit research [12] [13] [14]. It was developed by Yann LeCun and Corinna Cortes [15]. It contains a training set of 60000 examples, and a test set of 10000 examples. The digits have been centered and size-normalized into 28x28 gray-scale images. Some sample images are shown in figure 1 . The dataset is available at the homepage of LeCun.

III. DISCRETE COSINE TRANSFORM (DCT)

DCT initially used for image compression [16], have been of growing interest among the pattern recognition community

Figure 1: Sample images of MNIST dataset.

[5] [17] [7]. DCT is a technique to convert data of the image into its elementary frequency components [18]. It clusters high value coefficients in the upper left corner and low value coefficients in the bottom right of the array(m,n), where [m n] is the image size. DCT coefficients f(u,v) of f(m,n) are computed by:

$$f(u,v) = \alpha(u)\alpha(v)\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n)cos(\frac{(2m+1)\Pi u}{2M})$$
$$cos(\frac{(2N+1)\Pi v}{2N})$$

where

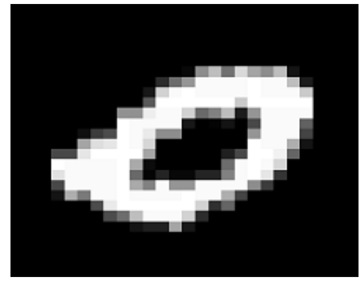
$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{M}}, & u = 0 \\ \sqrt{\frac{2}{M}}, & 1 \leqslant u \leqslant M-1 \end{cases}$$

and

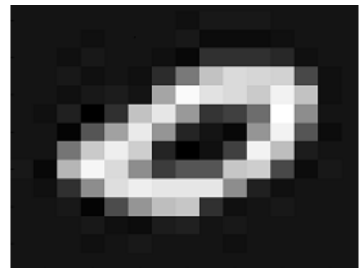
$$\alpha(v) = \begin{cases} \frac{1}{\sqrt{N}}, & v = 0 \\ \sqrt{\frac{2}{N}}, & 1 \leqslant v \leqslant N-1 \end{cases}$$

One of the main characteristics of DCT is its ability to convert the energy of the image into a few coefficients [5]. In this work we compare the performance of four variants of DCT coefficients namely:
- DCT upper left corner(ULC) coefficients;
- DCT zigzag coefficients;
- Block based DCT ULC coefficients;
- Block based DCT zigzag coefficients.

Our main goal is to identify the optimal feature extraction approach that reduces the dimension of MNIST data in order to speed up the learning algorithms while maximizing the classification accuracy.
Next, we will briefly present each method.



(a) Example of Initial Image size 28x28



(b) Image reconstructed with only 15x15 DCT ULC coefficients

Figure 2: Example of image reconstruction using 15x15 DCT ULC Coefficients

A. *DCT ULC coefficients*

This approach applies DCT on each image of MNIST Dataset and retains the N significant coefficients of the upper left corner from the transformed image. The Number N of the coefficients to retain is chosen experimentally. The higher the number of retained coefficient the better the quality of the reconstructed image characters. Figure 2 shows the original image character(a) with 28x28 pixels and the reconstructed image characters(b) with only 15x15 coefficients.

B. *DCT zigzag coefficients*

The same idea of the previous approach is applied: First, the DCT is applied to each image in the MNIST dataset, then the higher value DCT coefficients corresponding to low frequency are then extracted in a zigzag order and stored in a vector sequence as illustrated in figure 3.

C. *Block based DCT ULC coefficients*

In the literature, for block based DCT, the image is divided into blocks of 8x8, then the DCT is applied to each block. In this work, we dispose of images normalized into 28x28 size image, hence we used blocks of 7x7 then retained the most significant coefficients in the ULC of each block.

D. *Block based DCT zigzag coefficients*

The same process used for the block based DCT ULC is applied here. For each 7x7 block, we apply the DCT and retain the N significant coefficients in a zigzag fashion in order to form our feature vector. This vector will be fed to a SVM classifier for classification. A brief overview of this classifier is presented in the next section.

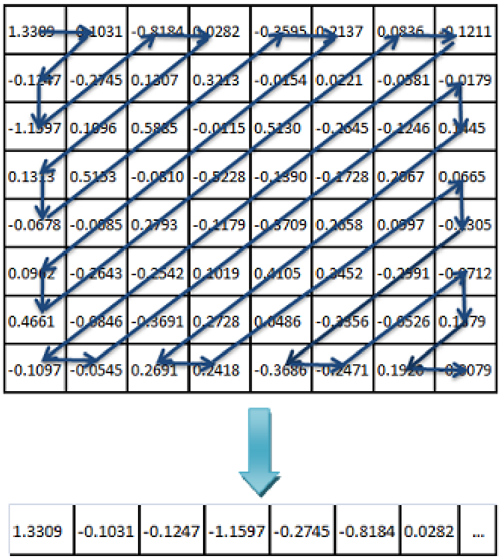


Figure 3: Selecting DCT coefficients in a zigzag fashion

IV. Classification using SVM

The Support Vector Machine (SVM) is a modern classifier which uses kernels to give optimal decision boundary to separate between classes in higher dimensional feature spaces. It was successfully evaluated on pattern recognition problems [19]. The SVM algorithm was originally introduced by Vapnik [20] in his work on structural risk minimization. For example, in two classes problem (positive and negative sets of samples), the basic form of linear SVM classifier tries to find an optimal hyperplane that separates the set of positive samples from the set of negative samples. In this work, we use the LIBSVM package [21] that support multi-class problem to classify the different handwritten digit classes.

V. Experimental Results

Experiments have been conducted in order to compare the classification performance on the MNIST dataset of the four variants of DCT coefficients presented previously in section III and the principal components extracted using PCA using an SVM classifier. The experiments were carried out in three steps:
- First we applied the feature extraction methods to the raw data MNIST dataset in order to constitute feature vectors. We retained 15x15 features for DCT ULC, 10x10 for DCT zigzag, 256 for block based DCT ULC (i.e 4x4 in each block of 7x7 pixels), 160 for block based DCT zigzag (that is 10 coefficients for each block) and 80 principal components for PCA. As shown in figure 4, 80 principle components can interpret approximately 90% of total information , which suffices to be representative and informative.
- Second we perform a grid search with a view to find the optimal parameters c and gamma for each one if the 45 SVM classifiers corresponding to the couples of the MNIST Digit classes ranging from 0 to 9.
- Finally, we train and test the 45 SVM classifiers on, respectively, the 60000 training feature vectors and the 10000 test feature vectors. The results achieved are summarized in Table I.

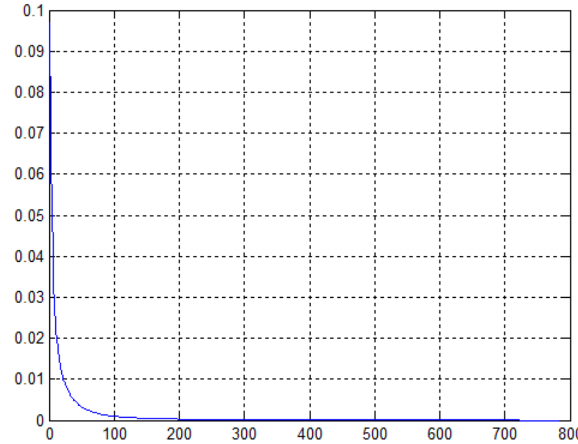


Figure 4: Spectrum of singular values

Figure 5 illustrates comparative performances of raw data features, PCA principal components and all retained DCT variants in terms of dimension reduction and global accuracy classification.

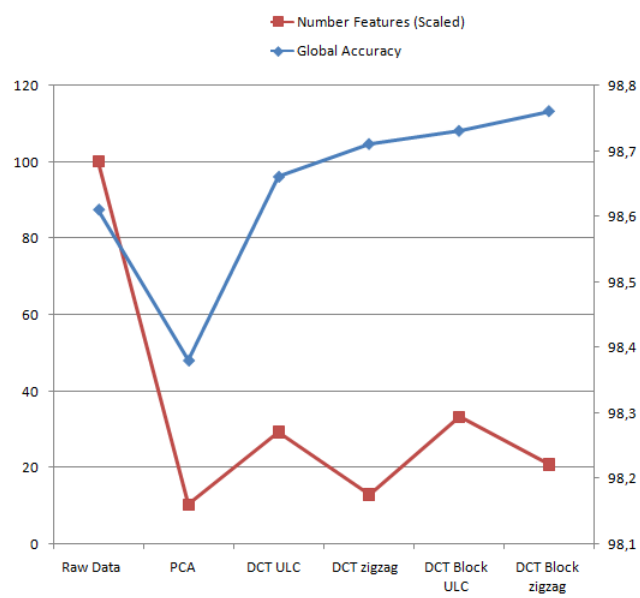


Figure 5: Classification rate and dimension reduction using the different Feature extraction techniques.

As we can see from table I, the classification of the four DCT variants outperform the results of raw data classification and the principal components feature vector. In the other hand, the dimension of MNIST dataset has been considerably reduced from 784 parameter pixels to just 80 for PCA, 225 for DCT ULC, 100 for DCT zigzag, 256 for block based DCT ULC and 160 parameters for block based DCT zigzag. Furthermore, the two block based DCT features are more discriminative than the DCT applied to the whole image. This could be due to the fact that the block based DCT retains the most significant coefficients in different regions of the image which is more accurate.

TABLE I: Performances of retained feature extraction approaches in terms of dimension reduction and accuracy classification

Feature Extraction Approach	Number of Features	Global Accuracy
Raw Data	784	98.61
PCA	80	98.38
DCT ULC	225	98.66
DCT zigzag	100	98.71
DCT Block ULC	256	98.73
DCT Block zigzag	160	98.76

VI. CONCLUSION

In this work we have evaluated the effectiveness of four variants of the discrete cosine transform to capture indiscriminate features in order to achieve a higher classification accuracy for handwritten digit recognition. These methods have been evaluated using the reference database MNIST and the SVM classifier and have been compared to classical PCA. The results have demonstrated that the block based DCT zigzag coefficients lead to a higher recognition rate. This performance is promising to integrate this method in our future work on developing an Arabic handwritten word recognition system.

REFERENCES

[1] C.-L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, "Handwritten digit recognition: benchmarking of state-of-the-art techniques," *Pattern Recognition*, vol. 36, no. 10, pp. 2271–2285, 2003.

[2] Y. LeCun, L. Jackel, L. Bottou, A. Brunot, C. Cortes, J. Denker, H. Drucker, I. Guyon, U. Muller, E. Sackinger *et al.*, "Comparison of learning algorithms for handwritten digit recognition," in *International conference on artificial neural networks*, vol. 60, 1995.

[3] C.-L. Liu, M. Koga, and H. Fujisawa, "Gabor feature extraction for character recognition: comparison with gradient feature," in *Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on*. IEEE, 2005, pp. 121–125.

[4] X.-T. Yuan and B.-G. Hu, "Robust feature extraction via information theoretic learning," in *Proceedings of the 26th Annual International Conference on Machine Learning*. ACM, 2009, pp. 1193–1200.

[5] J. AlKhateeb, J. Ren, J. Jiang, S. S. Ipson, and H. El Abed, "Word-based handwritten arabic scripts recognition using dct features and neural network classifier," in *Systems, Signals and Devices, 2008. IEEE SSD 2008. 5th International Multi-Conference on*. IEEE, 2008, pp. 1–5.

[6] A. Lawgali, A. Bouridane, M. Angelova, and Z. Ghassemlooy, "Handwritten arabic character recognition: Which feature extraction method," *International Journal of Advanced Science and Technology*, vol. 34, pp. 1–8, 2011.

[7] V. V. Kohir and U. Desai, "Face recognition using a dct-hmm approach," in *Applications of Computer Vision, 1998. WACV'98. Proceedings., Fourth IEEE Workshop on*. IEEE, 1998, pp. 226–231.

[8] C.-W. Ngo and C.-K. Chan, "Video text detection and segmentation for optical character recognition," *Multimedia Systems*, vol. 10, no. 3, pp. 261–272, 2005.

[9] M. Klima, P. Pata, K. Fliegel, and P. Hanzlik, "Image quality evaluation in security imaging systems," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 22, no. 1, pp. 22–25, 2007.

[10] R. Parisi, E. Di Claudio, G. Lucarelli, and G. Orlandi, "Car plate recognition by neural networks and image processing," in *Circuits and Systems, 1998. ISCAS'98. Proceedings of the 1998 IEEE International Symposium on*, vol. 3. IEEE, 1998, pp. 195–198.

[11] D. M. Monro, S. Rakshit, and D. Zhang, "Dct-based iris recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 4, pp. 586–595, 2007.

[12] L. Deng, "The mnist database of handwritten digit images for machine learning research," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 141–142, 2012.

[13] B. El Kessab, C. Daoui, B. Bouikhalene, M. Fakir, and K. Moro, "Extraction method of handwritten digit recognition tested on the mnist database." *International Journal of Advanced Science & Technology*, vol. 50, 2013.

[14] Y. Qian, W. Xichang, Z. Huaying, S. Zhen, and L. Jiang, "Recognition method for handwritten digits based on improved chain code histogram feature," in *3rd International Conference on Multimedia Technology (ICMT-13)*. Atlantis Press, 2013.

[15] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[16] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still image data compression standard*. Springer, 1993.

[17] M.-J. Tsai and H.-Y. Hung, "Dct and dwt-based image watermarking by using subsampling," in *Distributed Computing Systems Workshops, 2004. Proceedings. 24th International Conference on*. IEEE, 2004, pp. 184–189.

[18] A. Al-Haj, "Combined dwt-dct digital image watermarking." *Journal of computer science*, vol. 3, no. 9, 2007.

[19] V. Vapnik, *Statistical learning theory*. Wiley New York, 1998, vol. 2.

[20] V. N. Vapnik, "The nature of statistical learning theory," 1995.

[21] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.