EXPLORING THE POTENTIAL OF RECURRENCE QUANTIFICATION ANALYSIS FOR VIDEO ANALYSIS AND MOTION DETECTION

T. Kyprianidi[*], *E. Doutsi*[*], *G. Tzagkarakis*[*], *P. Tsakalides*[*†]

[*]Foundation for Research and Technology - Hellas, [†]University of Crete

ABSTRACT

This paper presents an enhanced methodology for Recurrence Quantification Analysis (RQA) designed specifically for video analysis. By utilizing image quality metrics, with a focus on the Peak Signal-to-Noise Ratio (PSNR), we determine meaningful values for the RQA threshold ε, a critical factor for successful image processing. Utilizing the False Nearest Neighbors (FNN) technique, we identify the optimal embedding dimension D for each patch within the video frames. Our approach produces a heatmap that visualizes temporal recurrence information for each video patch.

***Index Terms*—** Recurrence Quantification Analysis, Motion detection, video analysis, embedded dimension, temporal redundancy.

1. INTRODUCTION

Across diverse domains, spanning from image processing to biology, the amount of digital content is undergoing exponential expansion. An essential initial step in processing such extensive datasets is the reduction of data size by eliminating redundant information. Recurrence stands out as one of the most straightforward redundant aspects to identify. Currently, the inclination is to avoid conducting difference detection through consecutive comparisons, which can be time-consuming. This study aims to unveil recurrent properties within video analysis and put forth an alternative approach for identifying temporal redundancies.

Recurrence quantification analysis (RQA) is a method used in time series analysis to reveal recurrent properties of dynamical systems. It serves as a valuable tool for understanding the dynamics of time series data. RQA finds applications across various disciplines, including biomedical signal analysis [1] such as protein classification [2] [3], geophysics [4], meteorology [5], and psychology [6], just to name a few.

This work was funded by the TITAN ERA Chair project (contract no. 101086741) within the Horizon Europe Framework Program of the European Commission and the Hellenic Foundation for Research and Innovation (HFRI) and the General Secretariat for Research and Technology (GSRT) under grant agreement No. 330 (BrainSIM).

Our goal is to present an expanded version of RQA tailored for 2D time series data, with a particular focus on its applicability to video sequences. We aim to define the method's parameters in a manner that is meaningful for video analysis, ensuring that the detected recurrences accurately reflect real recurrent data. Through this extension, we seek to identify areas within the video that exhibit static behavior or varying degrees of motion intensity. Additionally, our goal is to produce heatmaps that illustrate the temporal stability of distinct regions within the video.

The remaining sections of the paper are structured as follows: Section 2 provides an introduction to the application of RQA in both time-series and images. Section 3 introduces the suggested mathematical framework that enables the utilization of RQA for 2D time-series data, particularly in the context of videos. In Section 4, a brief discussion on data pre-processing is provided, while Section 5 presents the experimental findings. Section 6 outlines the conclusions derived from this study and proposes potential directions for future exploration.

2. RECURRENCE QUANTIFICATION ANALYSIS

2.1. RQA for time series analysis

RQA is a method employed in the analysis of time series, to examine the recurrence properties of dynamical systems. Recurrence plots (RP), introduced by Eckmann *et. al.* [7], aim to visualize recurrences in dynamical systems. The concept of RQA relies on the reconstruction of phase space through time-delayed embedding. Takens' theorem [8] enables the reconstruction of phase-space profiles from a single one-dimensional observation.

If a system consists of multiple independent variables that drive its dynamics, and there is only one observable variable of the system, Takens' theorem suggests that plotting that single observable against itself multiple times (D) at a specific time delay (τ) can effectively reconstruct the multidimensional dynamics of the system. In other words, assuming the observable time series x with n values sampled at regular time intervals δt the recurrence properties of the system can be studied through this phase-space reconstruction technique.

$$x = (x_1, x_2, ..., x_n) \tag{1}$$

The optimal dimension of a dynamical system is the minimum number of variables or dimensions required to fully capture and describe its dynamics. If the optimal dimension of the dynamical system is D and the time-delay chosen is τ then the phase-space vector of x can be reconstructed as follows

$$\mathbf{X} = \begin{bmatrix} \mathbf{X_1} \\ \mathbf{X_2} \\ \vdots \\ \mathbf{X_{n-(D-1)\tau}} \end{bmatrix} = \begin{bmatrix} x_1 & x_{1+\tau} & \cdots & x_{1+(D-1)\tau} \\ x_2 & x_{2+\tau} & \cdots & x_{2+(D-1)\tau} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-(D-1)\tau} & x_{n-(D-2)\tau} & \cdots & x_n \end{bmatrix} \quad (2)$$

where $\mathbf{X}$ comprises a set of vectors $\mathbf{X_i}$ generated by starting at an initial point and selecting $D-1$ consecutive points with a time offset of τ. To compute the RPs, all pairwise distances $||\mathbf{X_i} - \mathbf{X_j}||$ between vectors $\mathbf{X_i}$ and $\mathbf{X_j}$ are calculated. If the distance is smaller than a threshold ε then the point $R_{i,j}$ in RP is considered recurrent, following the formula:

$$\mathbf{R_{i,j}} = \mathbf{\Theta}(\varepsilon - ||\mathbf{X_i} - \mathbf{X_j}||), \quad (3)$$

where $\Theta(x)$ denotes the Heaviside step function, taking the value 0 when $x < 0$ and the value 1 when $x \geq 0$. In the RP all recurrent points are depicted in black while non-recurrent points are shown in white. The choice of the threshold ε is pivotal, as it determines the identification of recurrences in the RP. Choosing the threshold requires achieving a balance by aiming for the smallest possible value while still ensuring a sufficient number of recurrences and recurrent structures [9]. There is no universal rule for ε, which depends on the specific application and experimental conditions.

The RP is symmetric, thus the main diagonal line inherently comprises recurrent points, as $R_{i,i} = 1$ by definition. To be able to quantify the RPs, various metrics for recurrence quantification analysis have been developed [10] [11] [12]. Unclear dynamical behaviors in the original time series can be revealed by the measures of RQA. Some of these metrics are the recurrence rate (RR), which indicates the percentage of recurrent points in the RP, the Determinism (DET), which measures the recurrent points present in diagonal structures, the maximal diagonal length (Lmax) excluding the main diagonal, and the entropy (ENT), which quantifies the signal complexity in bits per bin.

2.2. RQA for image analysis

An extension of the one-dimensional method of the RPs and their quantification to higher dimensions was proposed by Marwan et al [13] and later by Wallot et al [14]. The extension of the method to higher dimensions can serve as an effective approach to manage and analyse multiple features[15]. Their method was applied in 2D images to reveal and quantify recurrent structures within images. For a d-dimensional system, they defined an n-dimensional RP using the following formula

$$\mathbf{R_{i,j}} = \mathbf{\Theta}(\varepsilon - ||\mathbf{X_i} - \mathbf{X_j}||), \quad (4)$$

where $\mathbf{i}$ is the d-dimensional coordinate vector and $\mathbf{X_i}$ is the phase space vector at the location given by the coordinate vector $\mathbf{i}$. In other words, the vector $\mathbf{i}$ stands for the coordinates (i_1, i_2) where i_1 and i_2 are the row and column indexes of the input image, respectively. The dimension of the resulting RP is $n = 2 * d$ and even though it cannot be visualized, its quantification is still possible by computing the RQA measures in the n-dimensional RP. In the one-dimensional approach, the RP features a main diagonal line. Similarly, in the n-dimensional RP, there are diagonally oriented structures of d dimensions. When the method's input is a 2D image the resulting RP is a 4D plot and slices of that RP can be visualized as 3D subsections of the 4D RP.

The authors in [13] showed that using this extension typical spacial structures could be distinguished employing recurrences. They applied this method to biomedical images, to evaluate the bone structure from CT images of the human proximal tibia. Chomiak [16] introduced a method that utilizes a fuzzy c-means clustering approach and RQA measures (FCM-RQAS) for extracting image features, which can then be employed for training a classifier.

3. PROPOSED METHOD

This section introduces three primary aspects: (i) the mathematical framework involved in extending RQA to make it suitable for video data, incorporating both spatial and temporal information in 3D; (ii) establishing the threshold ε using image processing tools, and (iii) examining the influence of this new threshold value on defining the embedding dimension D.

3.1. Extension of RQA for video analysis

We will consider a video that consists of m frames where each frame has dimensions of M by N pixels as described below:

$$\mathbf{f} = (\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3}, ..., \mathbf{f_m}), \quad (5)$$

where $\mathbf{f}$ is the video and $\mathbf{f_i}$ the i-th frame. We divide each frame in smaller patches, where the patch size depends on the dimensions (M, N) of the frames and each patch starts after the previous one finishes without any overlap (see Fig. 1). Here we will consider specific values for certain parameters to illustrate the method: a patch size of 16×16 pixels, a video with $m = 299$ frames, and a frame size of (288×352). Eq. (6) shows the first frame of the video divided in patches $p_{i,j}$, where i denotes the number of the frame and j the number of each patch in that frame.

$$\mathbf{f_1} = \begin{bmatrix} \mathbf{p_{1,1}} & \mathbf{p_{1,2}} & \cdots & \mathbf{p_{1,22}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{p_{1,375}} & \mathbf{p_{1,376}} & \cdots & \mathbf{p_{1,396}} \end{bmatrix} \quad (6)$$

where for two representative patches, $\mathbf{p_{1,1}}$ and $\mathbf{p_{1,396}}$ that consists of the 16×16 top left and the 16×16 bottom right pixels of frame $\mathbf{f_1}$, respectively, the formula is given by:

$$\mathbf{p_{1,1}} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,16} \\ \vdots & \ddots & \vdots \\ x_{16,1} & \cdots & x_{16,16} \end{bmatrix}, \mathbf{p_{1,396}} = \begin{bmatrix} x_{273,337} & \cdots & x_{273,352} \\ \vdots & \ddots & \vdots \\ x_{288,337} & \cdots & x_{288,352} \end{bmatrix} \quad (7)$$

Each patch of size 16×16 is finally flattened resulting in a row vector of $(16 \times 16 = 256)$ pixel values, as shown in Eq. (8)

$$\begin{aligned} \mathbf{p}_{\text{flat}_{\mathbf{1,1}}} &= \left[x_{1,1}, x_{1,2}, \ldots, x_{16,16}\right] \\ \mathbf{p}_{\text{flat}_{\mathbf{1,396}}} &= \left[x_{273,337}, x_{273,338}, \ldots, x_{288,352}\right] \end{aligned} \quad (8)$$

Next, we have to determine the RQA parameters, i.e. the embedding dimension (D) and the time lag (τ): for illustration, we will choose $D = 16$ and $\tau = 1$, and we will focus on determining the optimum embedding dimension further bellow. Then, we reconstruct the phase space vector V as in one-dimensional RQA for each patch in the frames.

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \\ \vdots \\ \mathbf{V}_{284} \end{bmatrix} = \left[\begin{bmatrix} \mathbf{p}_{\text{flat}_{1,1}} \\ \vdots \\ \mathbf{p}_{\text{flat}_{16,1}} \end{bmatrix} \begin{bmatrix} \mathbf{p}_{\text{flat}_{2,1}} \\ \vdots \\ \mathbf{p}_{\text{flat}_{17,1}} \end{bmatrix} \cdots \begin{bmatrix} \mathbf{p}_{\text{flat}_{284,1}} \\ \vdots \\ \mathbf{p}_{\text{flat}_{299,1}} \end{bmatrix} \right]^T \quad (9)$$

The space vector $\mathbf{V}$ consists of vectors $\mathbf{V_i}$, containing $m - (D-1)\tau = 299 - (16-1)1 = 284$ vectors V_i, $\mathbf{V} = (\mathbf{V_1}, \mathbf{V_2}, ..., \mathbf{V_{284}})$. Each vector $\mathbf{V_i}$ consists of a 2D matrix that contains 16 flattened patches because the embedded dimension is set to $D = 16$, e.g. the first matrix V_1 contains the pixel values of flattened patches from $\mathbf{f_1}$ to $\mathbf{f_{16}}$. There is a visual explanation of the method in Fig. 1 for a hypothetical video with 4 frames, each frame divided in 9 smaller patches, and a phase space reconstruction with $D = 1$, $\tau = 1$. Then to continue with RQA, we compute all pairwise Euclidean distances between vectors V_i, $||V_i - V_j||$ and store them in a 2D symmetrical matrix where the rows and the columns correspond to the V_i, V_j used to compute the pairwise Euclidean distance. The final step is to calculate the RP according to the following formula:

$$\mathbf{R_{i,j}} = \mathbf{\Theta}(\varepsilon - ||\mathbf{V_i} - \mathbf{V_j}||) \quad (10)$$

where $\Theta(x)$ is Heaviside function and ε the threshold that determines the recurrent points.

3.2. Adopting image quality metrics in RQA

As mentioned before, the choice of threshold ε is crucial and it depends on the application. When employing the method on video data, it is essential to ascertain the presence of recurrent patches within the video frames. The Euclidean distance calculated for pairwise vectors V_i, V_j does not have any significant meaning in image processing and was not useful in the selection of the threshold ε. Thus, we decided to slightly modify the method and insert image quality metrics, such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), to be able to choose an ε value that will have a significant meaning for video analysis. Then, instead of computing all pairwise Euclidean distances of vectors $V_i, V_j, \forall i, j$, we compute all possible pairwise PSNR or SSIM values of the vectors V_i, V_j, and store them in a symmetrical matrix. According to this modification, PR is now given by Eq. (11) where $\varepsilon_{\text{PSNR}}$ is the threshold value using the PSNR approach.

$$\mathbf{R_{i,j}} = \mathbf{\Theta}\left(\varepsilon_{\text{PSNR}} - \frac{\mathbf{1}}{\text{PSRN}(\mathbf{V_i}, \mathbf{V_j})}\right), \quad (11)$$

$$\text{PSNR} = 10 \cdot \log_{10}\left(\frac{\text{MAX}^2}{\text{MSE}(\mathbf{V_i}, \mathbf{V_j})}\right), \quad (12)$$

where MAX is the maximum pixel value found in the frame (i.e. for a grayscale image this value is 255), and MSE is the mean square error between the pixels of the original and the reconstructed image.

We utilized the method for both PSNR and SSIM thresholds. However, considering the time-consuming nature of the SSIM threshold approach, we opted to proceed with the PSNR approach. This modification enables us to identify recurrent patches that fulfill the RP equation for a certain threshold $\varepsilon_{\text{PSNR}}$ of image similarity.

3.3. False Nearest Neighbors to calculate the optimal embedding dimension D

As discussed in Sec. 3.1, another RQA parameter that requires determination is the embedding dimension D. The False Nearest Neighbors (FNN) approach to determine the optimal embedding dimension D was first proposed by Kennel et al. [17]. The basic idea of this approach is to examine the change between neighboring points in phase space (vector V) as the embedding dimension is gradually increased. When two points in a one-dimensional time series have a close distance they are considered as neighbors. Then the distance of those two points is computed when they are embedded in two dimensions (for a time lag τ). If the embedding changes the distance between the neighbors, they are considered false neighbors and embedded in higher dimensions keeping the same time lag. When the distance of the two points does not change for a specific embedding dimension, they are considered as true neighbors, and that embedding dimension is considered optimal.

We apply the FNN criterion for every patch of the frames. We choose a set of embedding dimensions D starting from $D = 1$, we reconstruct the phase space vectors $\mathbf{V}$ and compute the pairwise PSNR distance matrices for every D. Then, for every row in the PSNR matrix, we keep the position of the maximum PSNR value (max(PSNR_D)) and compare it with the value in the same position of the PSNR matrix of an embedding dimension $D + 1$ (max(PSNR_{D+1}). If $||\text{max}(\text{PSNR}_D) - \text{max}(\text{PSNR}_{D+1})|| > t_{\text{PSNR}}$ where t_{PSNR} is

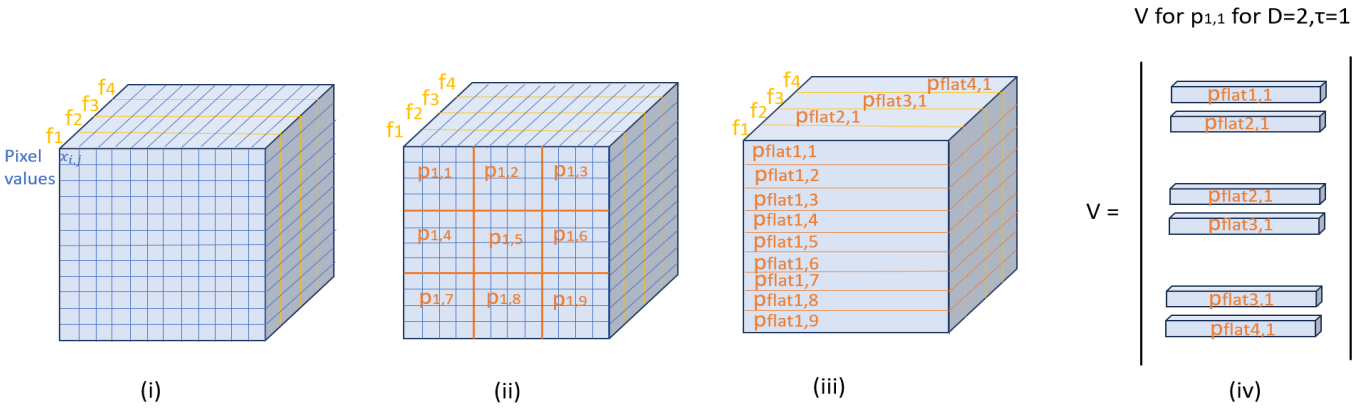


Fig. 1. Illustration to explain the proposed method (i) representation of the video as a cube containing 4 frames and pixel values for each frame, (ii) division of each frame in patches, (iii) rearrangement of the flattened patches through the frames, (iv) phase space vector V reconstruction for patch number 1 and $D = 1$, $\tau = 1$

a threshold value, they are considered false neighbors. We compute this difference for every row of the PSNR matrix and calculate the percentage of the false neighbors for that embedding dimension D. Then, we repeat the process for the next embedding dimension $D + 1$.

4. DATA PRE-PROCESSING AND PARAMETER SETTING

All the videos utilized in our experiments were obtained from derf's collection on xiph.org [18]. The experiments exclusively involved grayscale videos, necessitating the initial conversion of colored videos (in RGB format) to grayscale. The videos and the results of our analysis can be found in the following link [19]. Subsequently, each frame underwent division into smaller, non-overlapping patches of size 16×16. The determination of the optimal embedding dimension D for each patch was accomplished using the previously outlined FNN approach (refer to 3.3). The outcome of the FNN process is represented in a plot illustrating the percentage of false neighbors for the chosen set of D values. The optimal D is identified as (i) the value that reduces the percentage of false neighbors to 0, or (ii) the first of subsequent D that maintains this reduction, or (iii) the D for which the percentage of false neighbors starts to increase again. This methodology yields the optimal D for every patch in the video frames, serving as an indicator of how frequently a patch recurs. A patch with a large D suggests recurrence across a substantial number of frames, while a small D implies rapid changes in the patch over the frames.

To apply the proposed RQA method to video data, the determination of the threshold value $\varepsilon_{\text{PSNR}}$ is imperative. In the realm of image and video compression, the degradation in quality is typically reflected in PSNR values ranging from 30 to 50 dB for 8-bit data representation [20]. In this study, we opted for a threshold value of $\varepsilon_{\text{PSNR}} = 1/40$ dB. This choice aims to ensure a relatively good quality standard for patches identified as recurrent throughout the frames.

5. RESULTS

To demonstrate the efficacy of the proposed method in identifying recurrent regions within visual scenes, we have opted to utilize three distinct types of videos: (i) videos captured by fixed cameras, wherein all scene elements remain constant except for a single object, typically positioned centrally, displaying slight motion, (ii) videos recorded by fixed cameras, with most scene elements remaining static while objects or persons traverse the visual scene horizontally, and (iii) videos featuring rapid movement, where both the camera and scene elements undergo swift motion. According to the proposed theory, it is anticipated and experimentally proven in the rest of this section that constant regions will exhibit high D values, accompanied by dense RPs. Conversely, regions characterized by light or strong motion are expected to be associated with smaller D values and sparse RPs.

The first video we analyzed was *silent.cif* from derf's collection [18]. The content of the video is one woman signing against a constant background. The selection of this specific video was based on its composition of both stable and moving elements, rendering it an ideal candidate for evaluating the effectiveness of the proposed method in identifying recurrent patches. The video comprises a total of 300 frames of size (288×352) while the video length is 9 sec. The overall number of patches resulting after dividing each frame is 396 and can be observed in Fig. 2(a), whereas the Histogram of the distribution of patches' counts concerning the optimal D is depicted in Fig. 2(b).

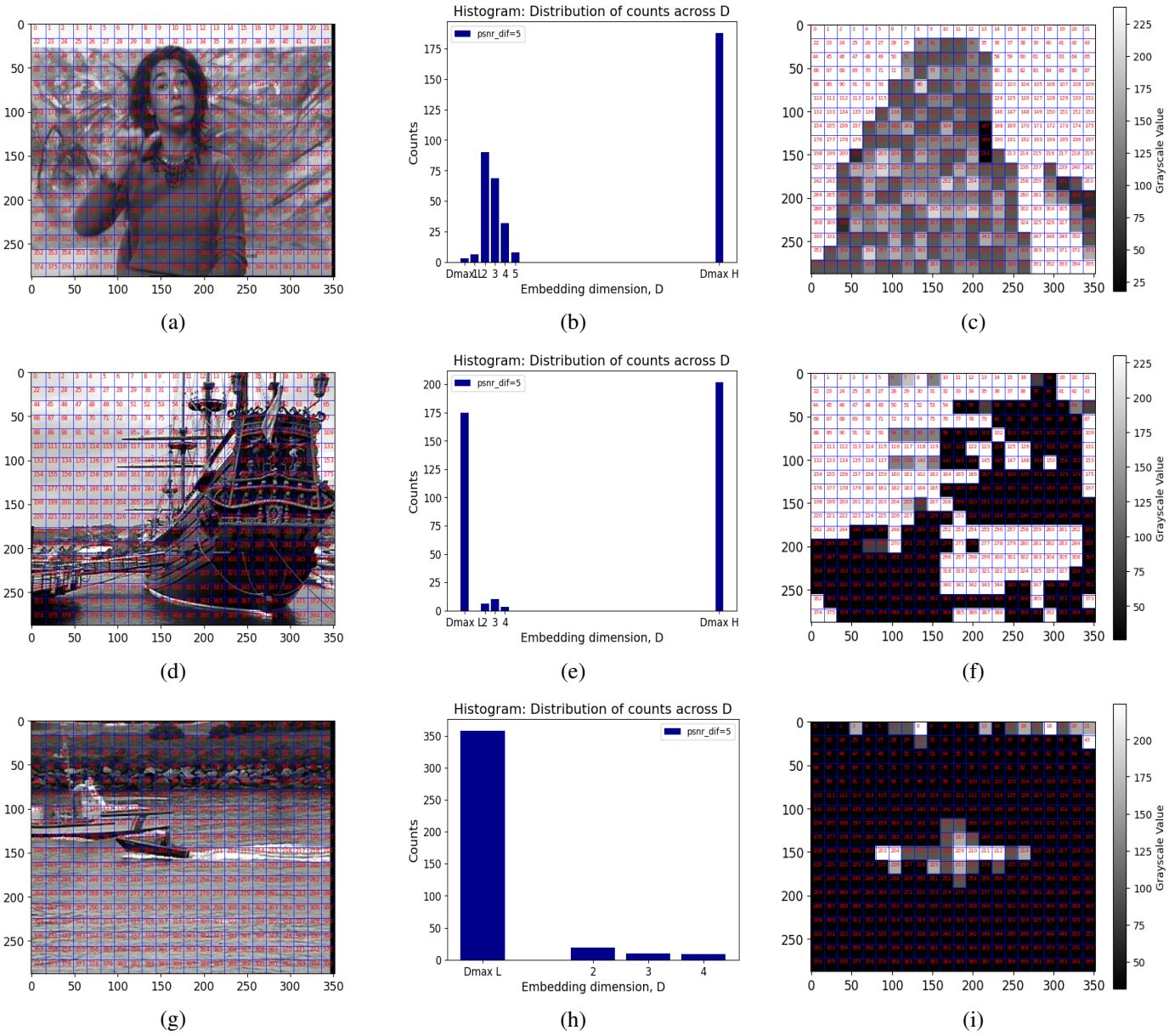


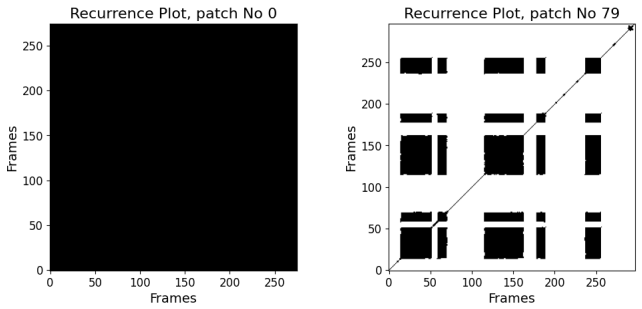
Fig. 2. Result for the videos "silent.cif" (a)-(c), "galleon.cif" (d)-(f) and "coastguard.cif" (g)-(i). (a), (c), and (g) grid of patches, (b), (e), and (h) histogram of counts against D optimal given by FNN, (c), (f) and (i) heatmap of gray values of patches according to their D optimal.

No false neighbors can happen when the values checked in the FNN method are always smaller than the threshold value for FNN. That can either mean that the patch remains stable for a substantial number of frames or that there are very slight changes in pixel intensities, resulting in minimal repetitive movements (e.g. subtle sea movements) below the FNN threshold. To discern between these possibilities for a patch with no false neighbors we compare the PSNR values of consecutive patches. If PSNR > 40 ($D_{\text{opt}} = D_{\max_H}$) it is considered to be a stable patch and D is set equal to 25. Conversely, if PSNR < 40 ($D_{\text{opt}} = D_{\max_L}$) the patch is deemed to have small movement, and D is set equal to 1.

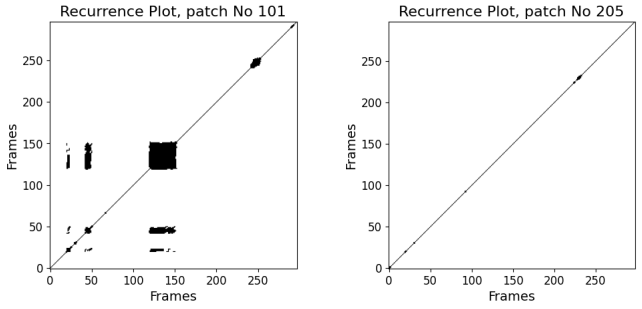
In the example video considered in Fig. 2, the optimal D values range from 1 to 5, with a significant number of patches having D optimal $D_{\max_H}$ and a smaller number having D optimal $D_{\max_L}$. As explained above, these values are obtained when there are no false neighbors for every embedding dimension D checked. Additionally, they are obtained when the FNN percentage for $D = 1$ is less than 4%, as we consider this percentage of false neighbors to be small for deter-

mining an optimal D. In such a case, the optimal solution for that patch is to have no false neighbors. This criterion is consistently applied across all our experiments.

Applying the FNN method to determine the optimal D for each patch, we collect information to create a heatmap of the patches according to their optimal D. Patches with the largest D were plotted as white and patches with the smallest D as black, and patches with D in between were assigned different values of gray according to their D value. The heatmap for the video examined is provided in Fig. 2(c). This representation of the video is highly valuable as it visually depicts different segments and their recurrence across frames.



(a) Patch 0, Dopt=25 (a) Patch 79, Dopt=3



(b) Patch 101, Dopt=3 (c) Patch 205, Dopt=2

Fig. 3. Recurrence plots for different patches of the frames

At this point, we can apply RQA with the optimal embedding dimension for each patch. The phase space vectors are reconstructed for each patch with D optimal. Then, we compute the pairwise PSNR matrix and the resulting RP for an RQA threshold value $\varepsilon_{\text{PSNR}} = 1/40$. Some RPs of different patches can be observed in Fig. 3. According to the RPs, patch 0 is recurrent for the whole video. Patch 79 is recurrent for frames around 10-50, 110-160, and 240-255. In patch 101 there are some recurrent frames around 120-150 and some smaller recurrences at different frame numbers. Finally, in patch 205 there are almost no recurrences as it is probably a patch with continuous motion. Referring to Figure 2, it is evident that patch 0 is situated in the background, which confirms its recurrence in all frames. In contrast, patch 205 is positioned around the hand moving to sign, elucidating why the RP appears empty.

We also applied the same analysis on two other videos from the same dataset, namely galleon.cif and coastguard.cif. The one video consists of 360 frames of size 288×352 with a length of 11 seconds. The content depicts a galleon docked against a backdrop with slight movements in certain areas and minor sea fluctuations. The other video consists of 299 frames of size 288×352 with a length of 9 seconds. There is continuous motion through this video, with the camera trying to capture a boat moving in the sea. The grid of the patches, the results of FFN, and the heatmap with grayscale values for the various D optimal values are found in Fig. 2 for both videos. The galleon.cif has many stable patches in the sky background and at the lower section of the galleon, along with several patches showing motion in the sea and at the upper part of the galleon. In coastguard.cif the majority of patches exhibit strong motion, with only a few patches in the middle of the frames, where the boat is located, showing lighter motion. The distribution of a number of frames for which patches are considered recurrent is illustrated in the histogram. This information is then spatially located in the heatmap and shown by grayscale values. The heatmap is a tool to detect motion through the patches of the video.

6. DISCUSSION AND FUTURE WORK

In this paper, we provided an extension of the RQA method for video analysis. Utilizing image quality metrics, such as the PSNR, to define RQA parameters, we were able to determine recurrences of each patch along the frames of the video. Additionally, the application of various quality metrics can be evaluated, including full-reference quality metrics like SSIM, and no-reference quality metrics such as DIVINE, BRISQUE, and NIQE. By employing the FNN approach, we determined the optimal embedding dimension D for each patch. Based on this, each patch was categorized as static or exhibiting either light or strong motion. This alternative approach yields a heatmap that contains information regarding the temporal recurrences of each patch. The color representation of each patch in the heatmap corresponds to the number of frames during which the patch remains static.

This information holds potential utility in video compression, as it provides insights into the frequency of changes in each patch. The embedding dimension D can be associated with the Group of Pictures (GoP), which represents the number of consecutive frames in a video stream that are encoded and compressed together. However, the utilization of this aspect in the work necessitates further investigation into how this information can be effectively employed for the encoding and compression of videos.

7. REFERENCES

[1] U. R. Acharya, S. V. Sree, S. Chattopadhyay, W. Yu, and P. C. A. Ang, "Application of recurrence quantification analysis for the automated identification of epileptic eeg signals," *International journal of neural systems*, vol. 21, no. 03, pp. 199–211, 2011.

[2] M. A. Zervou, E. Doutsi, P. Pavlidis, and P. Tsakalides, "Structural classification of proteins based on the computationally efficient recurrence quantification analysis and horizontal visibility graphs," *Bioinformatics*, vol. 37, no. 13, pp. 1796–1804, 2021.

[3] ——, "Efficient dynamic analysis of low-similarity proteins for structural class prediction," in *2020 28th European Signal Processing Conference (EUSIPCO)*. IEEE, 2021, pp. 1328–1332.

[4] N. Marwan, M. Thiel, and N. R. Nowaczyk, "Cross recurrence plot based synchronization of time series," *Nonlinear processes in Geophysics*, vol. 9, no. 3/4, pp. 325–331, 2002.

[5] B. Yan, P. W. Chan, Q. Li, Y. He, and Z. Shu, "Dynamic analysis of meteorological time series in hong kong: A nonlinear perspective," *International Journal of Climatology*, vol. 41, no. 10, pp. 4920–4932, 2021.

[6] B. N. Jenkins, J. F. Hunter, M. J. Richardson, T. S. Conner, and S. D. Pressman, "Affect variability and predictability: Using recurrence quantification analysis to better understand how the dynamics of affect relate to health." *Emotion*, vol. 20, no. 3, p. 391, 2020.

[7] J.-P. Eckmann, S. O. Kamphorst, D. Ruelle *et al.*, "Recurrence plots of dynamical systems," *World Scientific Series on Nonlinear Science Series A*, vol. 16, pp. 441–446, 1995.

[8] F. Takens, "Dynamical systems and turbulence," *Warwick, 1980*, pp. 366–381, 1981.

[9] N. Marwan and C. L. Webber Jr, "Mathematical and computational foundations of recurrence quantifications," in *Recurrence quantification analysis: Theory and best practices*. Springer, 2014, pp. 3–43.

[10] J. P. Zbilut and C. L. Webber Jr, "Embeddings and delays as derived from quantification of recurrence plots," *Physics letters A*, vol. 171, no. 3-4, pp. 199–203, 1992.

[11] C. L. Webber Jr and J. P. Zbilut, "Dynamical assessment of physiological systems and states using recurrence plot strategies," *Journal of applied physiology*, vol. 76, no. 2, pp. 965–973, 1994.

[12] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, "Recurrence-plot-based measures of complexity and their application to heart-rate-variability data," *Physical Review E*, vol. 66, no. 2, p. 026702, 2002.

[13] N. Marwan, J. Kurths, and P. Saparin, "Generalised recurrence plot analysis for spatial data," *Physics Letters A*, vol. 360, no. 4-5, pp. 545–551, 2007.

[14] S. Wallot, A. Roepstorff, and D. Mønster, "Multidimensional recurrence quantification analysis (mdrqa) for the analysis of multidimensional time-series: A software implementation in matlab and its application to group-level data in joint action," *Frontiers in psychology*, vol. 7, p. 1835, 2016.

[15] M. A. Zervou, G. Tzagkarakis, and P. Tsakalides, "Automated screening of dyslexia via dynamical recurrence analysis of wearable sensor data," in *2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*. IEEE, 2019, pp. 770–774.

[16] T. Chomiak, "Recurrence quantification analysis statistics for image feature extraction and classification," *Data-Enabled Discovery and Applications*, vol. 4, pp. 1–9, 2020.

[17] M. B. Kennel, R. Brown, and H. D. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Physical review A*, vol. 45, no. 6, p. 3403, 1992.

[18] "Xiph.org derf's test media." [Online]. Available: https://media.xiph.org/video/derf/

[19] "Videos and results in github." [Online]. Available: https://github.com/dwrakyp/icip2024_videos_results/tree/master

[20] R. G. Deshpande, L. L. Ragha, and S. K. Sharma, "Video quality assessment through psnr estimation for different compression standards," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 11, no. 3, pp. 918–924, 2018.