DATA SECURITY USING SVD BASED DIGITAL WATERMARKING TECHNIQUE

NAME:HIBA NAFEESATH
ROLL.NO:20
PRODUCT OWNER:K P BALACHANDRAN

Table Of Contents:

DESCRIPTION:

Illegal misuse of copyright information such as forgery, manipulation and duplication is not uncommon. To prevent this digital watermarking techniques are widely used thus increasing the robustness and imperceptibility properties in a digital multimedia. The main objective of developing a digital image watermarking technique is to satisfy both imperceptibility and robustness requirements.Digital watermarking appears as an efficient means of securing multimedia contents such as copyright protection and authentication. In this paper a hybrid scheme using Singular Value Decomposition (SVD) and Discrete Wavelet Transform (DWT) is being proposed. SVD and DWT are matrix based operations, this hybrid method prevents convolution which would otherwise consume a lot of resources. Computation of a larger set of data occurs faster due to the use of SVD. The watermarking scheme proposed is blind and uses a signature based authentication mechanism at the decoder which improves security. The method is subjected to various attacks and is evaluated in terms of PSNR and correlation values. A simple digital watermarking algorithm based on discrete wavelet transform and singular value decomposition has been proposed in this paper. This proposed method helps to understand basic concept of digital watermarking. Experimental results demonstrate the effectiveness of the proposed method. One of the major advantages of the proposed scheme is the robustness of the technique on wide set of attacks.

The cover image's singular values (SV's) are found and these SV's are used for embedding. Some important properties of SV'S are as follows:

(i) Singular values are stable

(ii) They represent intrinsic algebraic properties

(iii) SV's can be applied to rectangular matrices

(iv) They can survive various noise attacks , SVD preserves both one-way and non-symmetric properties of an image

(v) large portion of signal energy can be represented by just a few SV's.

Consider an image represented by a matrix Am×n, where m is the number of rows and n is the number of columns. The result of applying SVD on matrix Am×n is

SVD(Amxn) = [Umxn Smxn Vnxn]

where, U and V are the orthogonal matrices and S is the diagonal matrix. The elements of S represent the singular values (SVs).

Image assessment is done to find the quality of the watermarking technique. This is achieved by finding the PSNR value using the formula given below:

$$PSNR(db) = 10log_{10}\frac{(Max\,I)^2}{MSE}$$

Where I is the maximum pixel value of the image. MSE is the mean square error given by:

$$MSE = \frac{[\Sigma_{i=1}^{M}\Sigma_{j=1}^{N}[f(i,j) - f'(i,j)]^2]}{M \times N}$$

f(i,j) and f'(i,j) are the original and watermarked image respectively of size M×N. In order to find the robustness we calculate the correlation factor given by:

$$NC = \frac{\Sigma_{j=1}^{N} W\, W'}{\sqrt{\Sigma_{j=1}^{N} W}\sqrt{\Sigma_{j=1}^{N} W'}}$$

Where W and W' are the original watermark logo and extracted watermark logo respectively.

MODULES:

1.Data collection
2.Data Cleaning
3.Training
4.Testing
5.Result Generation

METHODOLOGY

The objective of this project is to develop a watermarking scheme which is based on cascading DWT with SVD. DWT decomposes the image into four frequency bands: LL band which represents low frequency, HL and LH representing middle frequency and HH represents high frequency band. LL band gives approximate details. In this proposal, we select LL band to embed the watermark because it contributes significantly to the robustness of an image. Thus it can survive certain image processing operations like noise addition, intensity manipulation, etc. In this SVD based watermarking scheme, instead of embedding the watermark directly on the wavelet coefficients SVD transformation is applied to the whole image and then the singular values of the host image are modified to embed the watermark.

A. Watermark Embedding

- Watermark W is decomposed using SVD

  $$W = U_W \times S_W \times V_W^T$$

- Using Haar wavelet perform first level decomposition of the cover image: LL, HL, LH, and HH. SVD is then applied to LL band.

  $$L = U_L \times S_W \times V_L^T$$

- The singular values of the LL band are replaced with the singular values of the watermark. After applying inverse SVD we obtain modified LL band.

  $$L' = U_L \times S_L \times V_L^T$$

- Inverse DWT is applied to produce the watermarked cover image.

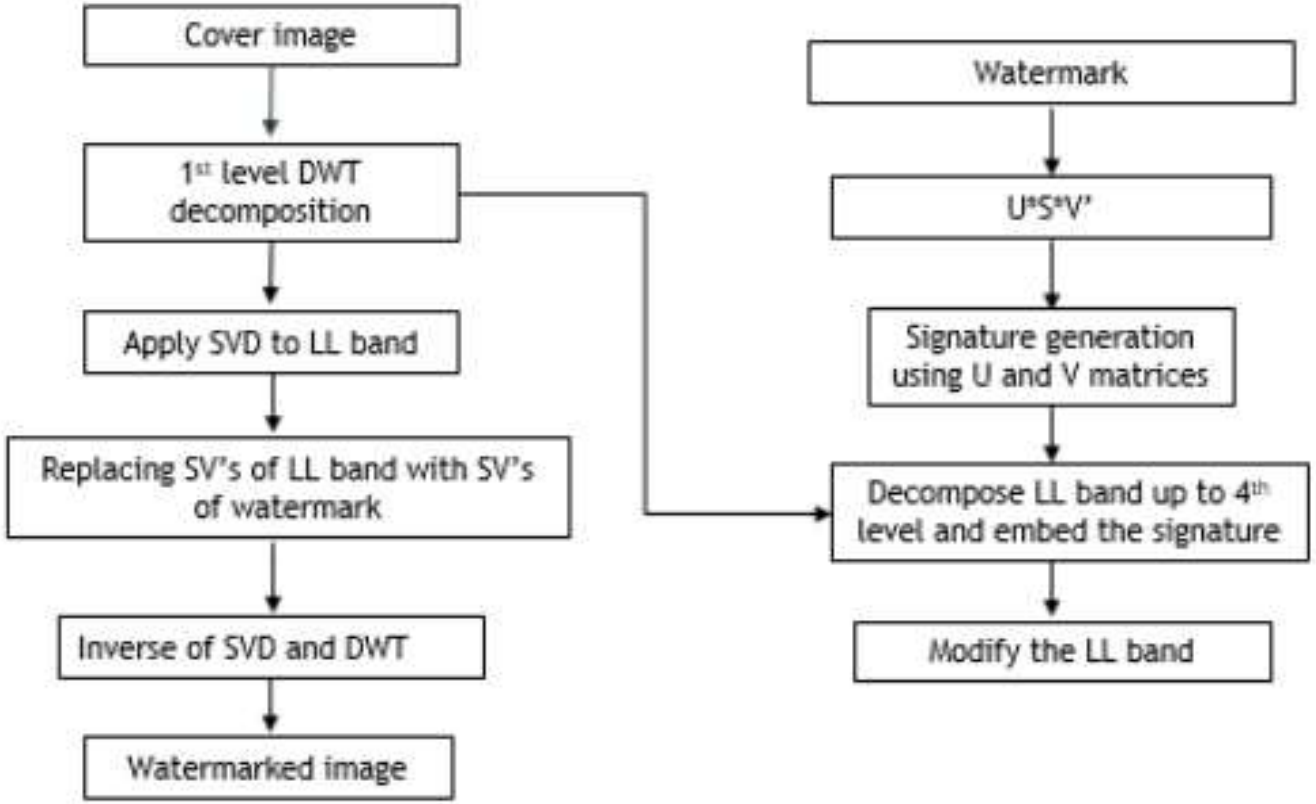

Fig 1.a Watermark embedding block diagram

B. Watermark Extraction

- Using Haar wavelet, the noisy watermarked image is decomposed.
- SVD is applied to LL band

  $L = UL \times SL \times VL^T$
- Then extract the singular values from LL band.
- The watermark is constructed using singular values and orthogonal matrices UW and VW obtained using SVD of original watermark.

  $WE = UW \times SL \times VW^T$

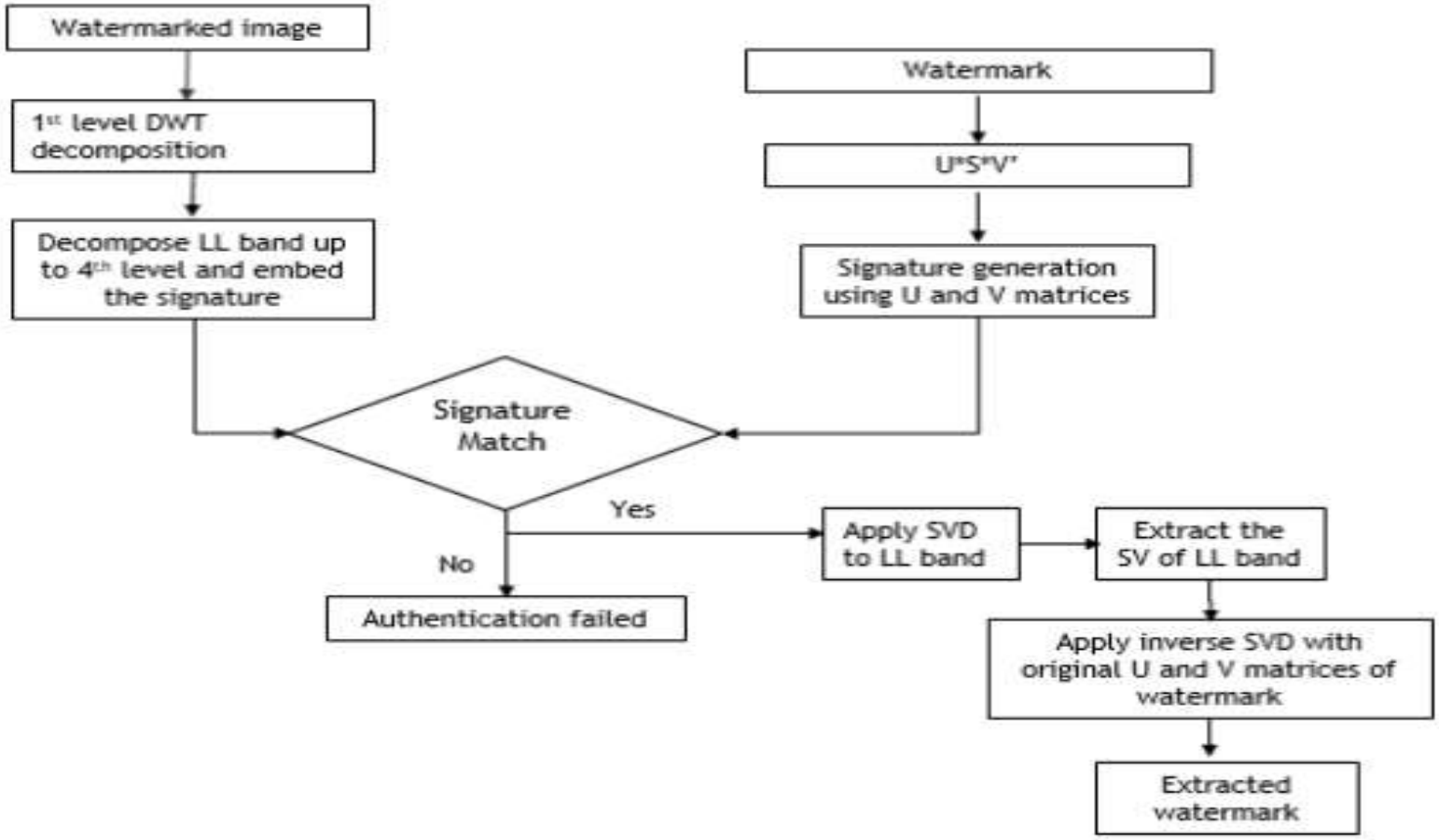

Fig 1.b Watermark extraction block diagram

C. Signature based Authentication

The U and V orthogonal matrices needs to be authenticated before extracting the watermark. A unique signature is generated using the U and V matrices and it is embedded into the cover image along with the watermark. There is one to one correspondence between the SV's and orthogonal matrices.

Digital signature for the U and V matrices is generated as follows:

- Create an array by summing the column of the U and V orthogonal matrices.
- Map the values of the array obtained with the corresponding binary digits based on threshold.
- XOR the binary digits to obtain the signature.

For embedding the signature we generate the signature of N bits for the U and V matrices of watermark. Using Haar wavelet, we perform fourth level decomposition of the cover image's LL band. N random coefficient are selected from LL4 and HH4 band with the help of secret key. Then we convert the integer part into L bits binary code. For signature extraction, perform the inverse procedure. To generate the signature use the U and V matrices of the original watermark at the receiver and compare it with the extracted signature. If they match, U and V matrices are authenticated and hence we can use them in the extraction of the watermark.

CONCLUSION

In this project a blind watermarking scheme has been proposed which combines SVD along with DWT. The watermark embedding and extraction algorithm were successfully implemented using MATLAB. From the result it is observed that this scheme yields higher PSNR value. Thus better imperceptibility is achieved. The correlation factor obtained is also high in case of attacks. Thus proving its robustness property. This hybrid technique is better than some of the existing methods such as DWT-DCT and pure SVD techniques.

FUTURE ENHANCEMENT

Tackling copyright issues, digital watermarking comes out as suitable solution. Digital watermarking is process of inserting watermark information into host image. Watermark is the copyright information which protects digital data from the illegal replication and distribution.

DEVELOPING ENVIRONMENT:

- Hardware specification:

  Processor : Intel Pentium Core i3 and above

  Primary Memory : 4 GB RAM and above

  Storage : 500 GB hard disk and above

- Software specification:

  Language :Python

  Front end : Python Django

  Back end : SQLite

  Operating system : windows 7 and above

  IDE : Visual Studio code,Jupyter Notebook

  Others : HTML,CSS

  Algorithm:Singular Value Decomposition

  Technique:Watermarking Embedding & Extraction,Signature based authentication

  Data set:coco dataset from Kaggle website

PROJECT PLAN:

User Story ID	Task Name	Start Date	End Date	Days	Status
1	Sprint 1	20/4/2022	01/05/2022	2	Completed
2		01/05/2022	09/05/2022		Completed
3	Sprint 2	15/05/2022	18/05/2022	5	Completed
4		19/05/2022	22/05/2022		Completed
5		23/05/2022	28/05/2022		Completed
6		29/05/2022	01/06/2022		Completed
7	Sprint 3	02/06/2022	05/06/2022	3	Completed
8	Sprint 4	Planned	Planned		Planned

USER STORY:

User Story ID	As a <type of user>	I want to <perform some task>	So that I can <achieve some goal>
1	User	Collection of Dataset	Coco Dataset
2	User	Preprocessing of collected data(null values elimination)	Cleaned final dataset
3	User	UI designing(a form to enter news for checking)	UI designing
4	User	Visualisation of input image	Graphical representation of data
5	User	Applying watermark on image	Watermarked image
6	User	Split data into training & testing set	80% -training data & 20%-testing
7	User	SVD-DWT Algorithm	Decomposed values & original image,watermarked image
8	User	Generation of output	Informs whether copyright occurred or not

PRODUCT BACKLOG:

User Story ID	Priority<High /Medium/Low >	Size(Hours)	Sprint	Status<Planned/Inprogress/ Completed>	Release Date	Release Goal
1	Medium	2	1	Completed	01/05/2022	Collection of datasets
2	High	3		Completed	09/05/2022	Preprocessing of collected data
3	Medium	3	2	Completed	18/05/2022	UI Designing
4	High	2		Completed	22/05/2022	Graphical representation of data
5	Medium	5		Completed	28/05/2022	Watermarked image
6	High	5		Completed	01/06/2022	Train the data,
7	High	10	3	Completed	05/06/2022	Decomposed values & original image,watermarked image
8	High	20	4	Planned		Output

SPRINT PLAN:

Backlog item	Status & Completion date	Original Estimate in hours	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13
User Story#1,2		Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours
Data collection	01/05/2022	2	2	0	0	0	0	0	0	0	0	0	0	0	0
Preprocessing	09/05/2022	3	0	3	0	0	0	0	0	0	0	0	0	0	0
User story #3,4,5,6															
Visualisation & Training	01/06/2022	15	0	0	3	3	3	3	3	0	0	0	0	0	0
User story #7															
UI Designing	05/06/2022	10	0	0	0	0	0	0	0	4	4	2	0	0	0
User story #8,9															
Testing	12/07/2022	20	0	0	0	0	0	0	0	0	0	0	8	6	6
Total		50	2	3	3	3	3	3	3	4	4	2	8	6	6

Sprint1 Actual:

Backlog item	Status & Completion date	Original Estimate in hours	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Completed<Y/N>
User Story#1,2		Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours
Data collection	01/05/2022	2	2	0	0	0	0	0	0	0	0	0	0	0	0	Y
Preprocessing	09/05/2022	3	0	3	0	0	0	0	0	0	0	0	0	0	0	Y
User story #3,4,5,6																
Visualisation & Training	Planned	15	0	0	3	3	3	3	3	0	0	0	0	0	0	N
User story #7																
UI Designing	Planned	10	0	0	0	0	0	0	0	4	4	2	0	0	0	N
User story #8,9																
Testing	Planned	20	0	0	0	0	0	0	0	0	0	0	8	6	6	N
Total		50	2	3	3	3	3	3	3	4	4	2	8	6	6	N

Sprint2 Actual:

Backlog item	Status & Completion date	Original Estimate in hours	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Completed<Y/N>
User Story#1,2		Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours
Data collection	01/05/2022	2	2	0	0	0	0	0	0	0	0	0	0	0	0	Y
Preprocessing	09/05/2022	3	0	3	0	0	0	0	0	0	0	0	0	0	0	Y
User story #3,4,5,6																
Visualisation & Training	01/06/2022	15	0	0	3	3	3	3	3	0	0	0	0	0	0	Y
User story #7																
UI Designing	Planned	10	0	0	0	0	0	0	0	4	4	2	0	0	0	N
User story #8,9																
Testing	Planned	20	0	0	0	0	0	0	0	0	0	0	8	6	6	N
Total		50	2	3	3	3	3	3	3	4	4	2	8	6	6	N

Sprint3 Actual:

Backlog item	Status & Completion date	Original Estimate in hours	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Completed<Y/N>
User Story#1,2		Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours	Hours
Data collection	01/05/2022	2	2	0	0	0	0	0	0	0	0	0	0	0	0	Y
Preprocessing	09/05/2022	3	0	3	0	0	0	0	0	0	0	0	0	0	0	Y
User story #3,4,5,6																
Visualisation & Training	01/06/2022	15	0	0	3	3	3	3	3	0	0	0	0	0	0	Y
User story #7																
UI Designing	05/06/2022	10	0	0	0	0	0	0	0	4	4	2	0	0	0	Y
User story #8,9																
Testing	Planned	20	0	0	0	0	0	0	0	0	0	0	8	6	6	N
Total		50	2	3	3	3	3	3	3	4	4	2	8	6	6	N

Thank you