



A Survey on a System to Maintain the Storage of Contentious Image in the form of Descriptor

Nilesh Kawale, Yogesh Bhute

Computer Science & Engineering, R.T.M.N.U
India

Abstract— *Child pornography is one of the rapidly increasing illegal activities in all over the world. Possession and distribution of Child pornography are considered as crimes which are prosecuted in most countries around the world. This law is so strict that it is not only just applicable to regular citizens but also to police forces which may complicate them in gathering and sequencing evidence for future reference. Due to this problem, this module presents an innovative solution. This paper presents tools for sequencing metadata of the evidence material and also provide fast and accurate search to recover such evidence from suspect's file system which would helpful in capturing offenders and provide protection to victims .*

Keywords— *Child Pornography, Forensics Analysis, Deviation & Dispersion Matrix Algorithm*

I. INTRODUCTION

Since the widespread use of the Internet, the distribution and acquisition of child pornography, namely illegal and immoral images and video files of children, or "any visual depiction of illegally explicit conduct involving children has dramatically increased. Nowadays, the exchange of information between people is mainly through the World Wide Web (WWW), the peer-to-peer networks, e-mail, and instant messaging services. Thus the analysis of computer storage media such as hard disk drives, pen drives, optical media and memory card is very much important to verify the presence and distribution of child pornography. Possession of illegal images is considered as a crime and this law is not applicable to only regular citizens but also to the police unit. Police officers have to destroy all the evidence after the investigation is completed which may complicate them in gathering and presenting evidence in investigation. Due to this is impossible for police forces to make connection between individual cases in order to find evidence in suspect's system.

To overcome this problem, this paper presents an application or tool which is able to catalogue the image of child pornography. In the proposed system, during investigation police officers can extract the metadata from suspected system and store it in the form of descriptor in the SEQUENCER component. When they have access to the suspect's system they can use SEARCHER component to find identical data or very visually similar to the samples present in the database.

This application is consist of two components. Most of the application has already demonstrated such as generic MIRROR [12] content-based image retrieval system and more specific system such as GAMA, designed for accessing libraries of media art. With this, there are again more applications such as EnCase and Forensic Tool Kit (FTK) which determine the relationship in the data and find key parts of information and generate reports that are easily consumed by other investigator. The above mentioned tools allows only the retrieval of images that are identical (they utilize MD5 sums), and also allows the retrieval of similar images, since it utilizes MPEG-7 descriptor values.

II. LITERATURE SURVEY

Iván Pau de la Cruz, Celia Fernández Aller, Sergio Sánchez García & Justo Carracedo [14] deals with the social problem of child pornography on peer-to-peer (P2P) networks on the Internet and heres an automated system with effective computer and telematic tools for seeking out and identifying data exchanges with pedophilic content on the Internet. This tool can be used only for the detection of pedophilic content and the information identifying users will remain anonymous until authorization has been received from a judicial authority.

Manjunath BS, Salembier P, Sikora Th [12] has introduced a system called Mirror: an interactive content based image retrieval system for evaluating MPEG-7 visual and developing new retrieval algorithms Web-based user interface for query by image example retrieval. A new Merged Color Palette approach for MPEG-7 dominant color descriptor similarity measure and relevance feedback is also developed in this system. This system is also very suitable for new algorithm development and evaluation. A new merged-color histogram (MCH) approach for similarity measure and relevance feedback feature generation was developed in MIRROR to enhance the retrieval accuracy on using MPEG-7 Dominant Color Descriptor (DCD).

MIRROR consists of following module:

A. Feature Extraction Module

It takes out the descriptor from the original image and converts the extracted information into MPEG-7 data stream.

B. Similarity Measure Module

It uses user's reference image as a query to search the relevant images in the database which consisting of pre-generated MPEG-7 data

C. Relevance Feedback Module

It receives user feedback on relevant retrieval and generates a new query for Similarity Measure module for a new retrieval.

Rogers MK, Goldman J, Mislan R, Wedge T, Debrota S [11] has presented the Cyber Forensic Field Triage Process Model (CFFTPM) that suggests an onsite or field approach for providing the identification, analysis and interpretation of digital evidence in a short time frame as possible. The proposed model adheres to commonly held forensic principles, and does not negate the ability that once the initial field triage is concluded, the system storage media be transported back to a lab environment for a more thorough examination and analysis. This model is used in various real world cases, and its investigative importance and pragmat approach has been amply demonstrated. The model is general enough to be used across a wide spectrum of investigations Furthermore; the derived evidence from these cases has not been challenged in the court proceedings where it has been introduced.

Edge defines the boundary between an objects and the background and it also indicates the boundary overlapping objects. The edge detection [4] is very important in computer vision for the identification and classification of objects in image as well as segmentation and object reorganization. The quality of image detection is highly dependent on lighting condition, the presence of objects of similar entities, density of edges in the scene and noise. This paper has compared various edge detection techniques in image processing like,

- The Marr-Hildreth Edge Detector
- The Canny Edge Detector
- The Local Threshold and Boolean Function Based Edge Detection
- Color Edge Detection Using Euclidean Distance and Vector Angle
- Color Edge Detection using the Canny Operator
- Depth Edge Detection using Multi-Flash Imaging

This paper have proposed and implemented Multi flash method. But this implemented method is difficult in finding the edges between object that are at almost the same depth or at depth which are very far away. Multi flash method would not work in this technique. The author tested six edge detectors that use different methods for detecting edges and compared their results under a variety of situations to determine which detector was preferable under different sets of conditions. This data could then be used to create a multi-edge-detector system, which analyzes the scene and runs the edge detector best suited for the current set of data.

Eleuterio P, Polastro M [5] has introduced a tool to optimize the automatic nudity detection provided by Nu Detective Forensic Tool by reducing the images resolution. This will be helpful for nudity detection in videos, it resolve multiple images into a single video by reducing resolution. This tool carries out automatic detection of nudity in images and also performs analysis of file names. Two evaluation experiments of the Tool were performed and showed detection rates around 95%, with low rates of false positives, combined with fast processing.

Pham DT, Ghanbarzadeh A, Koç E, Otri S, Rahim S, Zaidi M[10] has introduced a search procedure inspired by the way honey bees forage for food. Computationally fast multi-objective optimizer tool for complex engineering multi-objective optimization problems .This solve a multi objective optimization problem without any special domain information, apart from that it needed to compute objective functions.

It has several drawbacks in terms of algorithm

1. This algorithm searches complete image which consumes too much time to search. Hence it is not useful.
2. Bee's algorithm searches nearby folder for contentious images due to this complete scan of system are not possible.
3. It also takes more time for scanning video files.

Osama A. Lotfallah, Martin Reisslein, and Sethuraman Panchanathan, [13] proposes to employ MPEG-7 descriptors to improve the quality of the video delivered over best-effort networks. In addition to this he proposes a video transmission system that uses the motion activity descriptors to ensure robust video transmission.

III. CONCEPT OF APPLICATION

This tool is classified into a set of two components. In this application, SEQUENCER component is designed to be used at police head quarter. The police unit has their disposal set of information such images containing child pornography from various sources, including ongoing investigation and internal channels of cooperation. This data is given as input to SEQUENCER component. It processes the images and converts that set of images in the form of image descriptor. The descriptor set consist of MPEG-7 descriptor. This process is a one way process that means image cannot be recreated from that descriptor.

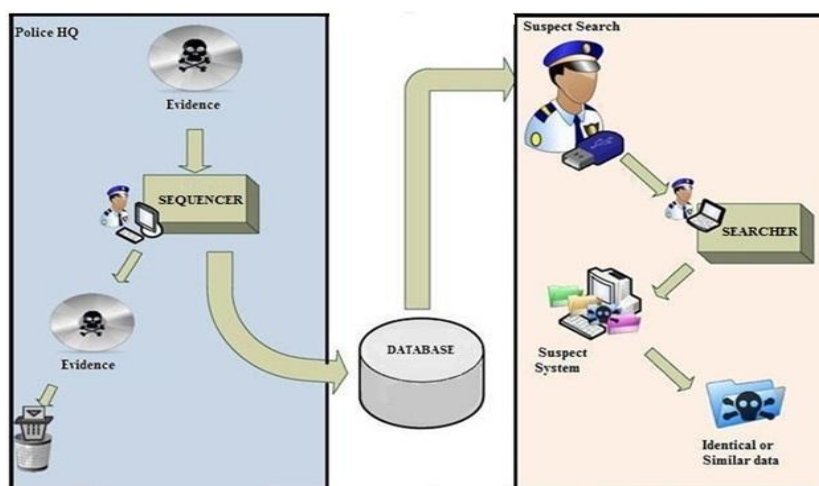


Fig. 1 System Architecture

The SEARCHER component is designed to search a suspect’s system. It uses the database of descriptor created by SEQUENCER component. All the identical and visually similar images will be retrieved and presented alongside information to the police officer performing search. This helps the police to draw conclusion regarding the possible sources of images and their distribution paths.

IV. PROPOSED ALGORITHM

Deviation and Dispersion matrix algorithm

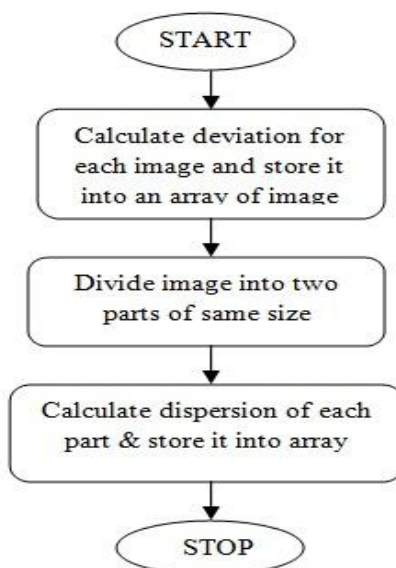
Two images with the same deviation may be the same but two images with different deviation, must be different. The deviation of the image is very light and it can save time in the search process. But it is not enough for judgment that the two images that have the same variance will be the same. So it needs another test to assure that.

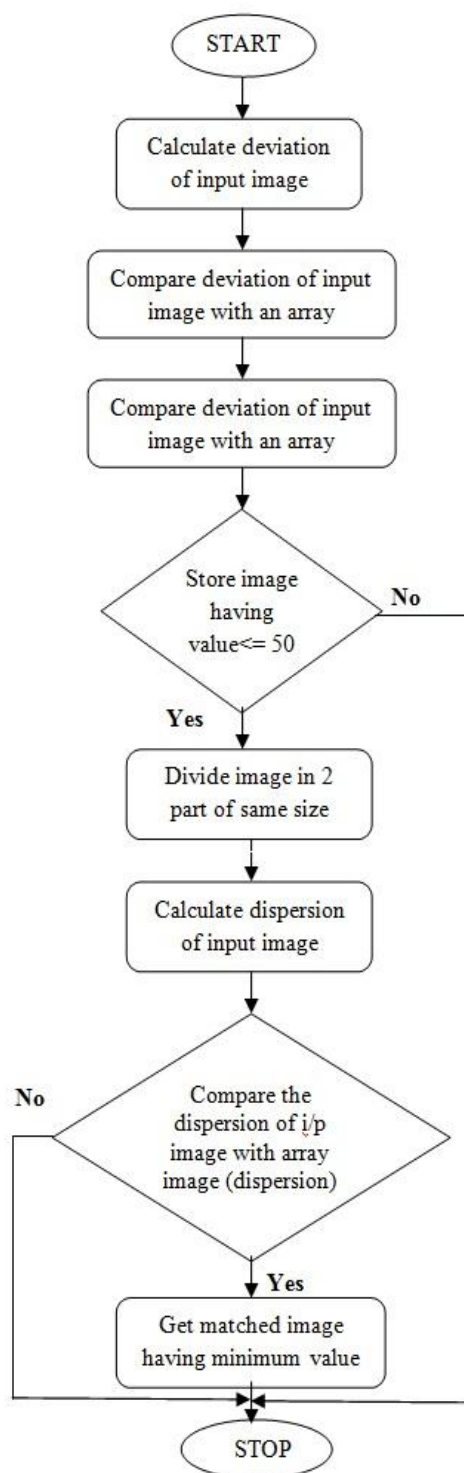
Therefore it finds the dispersion matrix of images and according it performs the matches to the sample stored in database. It searches the target image for an image having similar dispersion matrix.

The following describes the proposed algorithm

- Calculate the deviation value for each image and store it into an array of image
- Divide image in two parts of same size
- Calculate the dispersion of each part in new array of image.
- Calculate the deviation of input image
- Compare deviation of input image with array image.
- Store the image that has values less than equal to 50.
- Divide image in two parts of same size.
- Calculate dispersion of input image.
- Compare dispersion of input image with dispersion of extracted image.
- Get matched image having minimum value.

Flow Chart





V. CONCLUSION

This paper presents a tool that creates and maintains the storage of child pornographic image which would not be allowed by police to store and search those images in suspect's system in fast way. It also provides the relation between previous and current cases. Due to this it would be helpful to provide the protection to victims, capturing offenders in effective manner and also reduce the tedious work of police unit. The proposed algorithm reduces the complexity of previous image searching algorithm in very efficient manner. Hence it provides fast search based on deviation and dispersion calculation.

REFERENCES

- [1] AccessData: Forensic Toolkit (FTK) computer forensics software. <http://accessdata.com/products/forensic-investigation/ftk>. Accessed 30 Sept 2011.
- [2] ADF Solutions: The Forensic Triage Company. <http://www.adfsolutions.com>. Accessed 30 Sept 2011.

- [3] Stefan Berchtold, Daniel A. Keim, Indexing The Solution Space: A New Technique For Nearest Neighbor Search In High-Dimensional Space, *IEEE Transactions On Knowledge And Data Engineering*, Vol. 12, No. 1, January/February 2000 45.
- [4] Nadernejad E, Sharifzadeh S, Hassanpour H (2008) Edge detection techniques: evaluations and comparisons. *Appl Math Sci* 2(2–31):1507–1520.
- [5] Eleuterio P, Polastro M (2010) Optimization of automatic nudity detection in high- resolution images with the use of NuDetective Forensic Tool. In: *Proceeding of the fifth international conference on forensic computer science—ICoFCS 2010, Brasilia, Brazil*.
- [6] Mohand-Saïd Hacid, Cyril Declair, and Jacques Kouloumdjian, A Database Approach for Modeling and Querying Video Data *IEEE Transactions On Knowledge And Data Engineering*, Vol. 12, No. 5, September/October 2000 729.
- [7] Ludtke A, Gottfried B, Herzog O, Ioannidis G, Leszczuk M, Simko V (2009) Accessing libraries of media art through metadata. In: *International workshop on database and expert systems applications*, pp 269–273.
- [8] Manjunath BS, Salembier P, Sikora Th (2002) *Introduction to MPEG-7 media content description interface*. Wiley, Chichester.
- [9] Pardo A (2006) Probabilistic shot boundary detection using inter frame histogram differences. In: *CIARP'06: Proceedings of the 11th Ibero american conference on progress in pattern recognition, image analysis and applications*. Springer-Verlag, Berlin, Heidelberg, pp 726–732. doi:[10.1007/11892755_75](https://doi.org/10.1007/11892755_75).
- [10] Pham DT, Ghanbarzadeh A, Koç E, Otri S, Rahim S, Zaidi M (2006) The Bees Algorithm—a novel tool for complex optimisation problems. In: *Proceedings of the 2nd international virtual conference on intelligent production machines and systems (IPROMS 2006)*. Elsevier, Oxford, pp 454–459.
- [11] Rogers MK, Goldman J, Mislán R, Wedge T, Debrotá S (2006) Computer forensics field triage process model. In: *Conference on digital forensics, security and law*.
- [12] Wong KM, Cheung KW, Po LM (2005) Mirror: an interactive content based image retrieval system. In: *ISCAS (2)*. IEEE, pp 1541–1544.
- [13] Osama A. Lotfallah, Martin Reisslein, and Sethuraman Panchanathan, Fellow, Adaptive Video Transmission Schemes Using MPEG-7 Motion Intensity Descriptor *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 16, No. 8, August 2006 929.
- [14] Iván Pau de la Cruz, Celia Fernández Aller, Sergio Sánchez García & Justo Carracedo, (2010) A Careful Design for a Tool to Detect Child Pornography in P2P Networks *IEEE International Symposium on Technology and Society*, pp 227-233.