International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-03, Issue-04, July 2017

# Classification of Covert Photographs By Fusing Attributes and Features

[1]Vrushali C. Bhadane, [2]Prof. Nilesh R. Wankhade

[1]P.G. Student, [1,2]Kalyani Charitable Trust's Late G.N. Sapkal College of Engineering, Nashik, Maharashtra, India.

*[1]vrush.c.b@gmail.com, [2]nileshrw_2000@yahoo.com*

Abstract **The increased in technology in image acquisition technique made recording images never easier and brings a great convenience to our daily life. But it also raises the issue of privacy protection in the photographs. The particular problem addressed in this paper is about covert photographs, which are those pictures or photographs taken secretly and often violate the subject willingness. Covert photos are often privacy invasive and, if distributed over Internet, can cause serious consequences. Automatic identification of such photos, therefore, serves as an important initial step toward further privacy protection operations. The problem is, however, very challenging due to the large semantic similarity between covert and noncovert photos, the enormous diversity in the photographing process and environment of cover photos, and the difficulty to collect an effective data set for the study. To overcome these challenges three methods are followed. Initially a dataset of 2500 covert images is considered and each image is verified carefully. Secondly a user study is conducted on how human perform to distinguish between covert and non-covert images. Lastly a covert photo classification algorithm is implemented that fuses various image features and visual attributes in the multiple kernel learning framework.**

***Keywords — covert photography, image classification, multiple kernel learning, privacy protection, visual attribute.***

## I. INTRODUCTION

Proliferation of image acquisition devices and new Internet technologies provides people great convenience to shoot photos or pictures and share them publicly. Such convenience is however accompanied with a series of social and legal issues, such as adult contents controlling and privacy protection. They make use of a new type of visual privacy threat to investigate, called covert photography, in which the photographing process is intentionally concealed from the subject being photographed. Photos taken this way, named covert photographs or covert photos, often seriously threaten public or personal privacy.

Despite a large literature on image understanding, covert photo classification has never been studied as per the best of our knowledge and it challenges existing image classification methods from several aspects. First, there is a large ambiguity between covert and non-covert photos in terms of semantic content. This is because covert photos distinguish themselves by their acquisition procedure other than semantic content. Second, there is a large variation in the photographing environment and conditions for covert photos, and the variation causes serious within-class variation of covert photos. Third, there exists no covert photo dataset and it is nontrivial to collect one. For addressing these challenges three contributions are made towards automatic covert photo classification.

- First, a large covert photo dataset containing 2,500 covert photos and 10,000 non-covert ones is created. Each covert or non-covert photo is verified by checking its photographing process. The datasets are carefully designed so as to reduce potential biases as much as possible.

- To help for better understand the problem, they conduct a user study to investigate how human beings distinct covert photos from non-covert ones. The study generates an important benchmark measuring human's ability to the task.

 

- Further fusing of both low-level image descriptors and middle-level visual attributes for covert photo classification is being performed. Specifically, 8 different image features (e.g., color GIST) and 13 visual attributes (e.g., color richness) are chosen for image representation. All the features and attributes are then combined using the multiple kernel learning framework (MKL) for covertness decision.

## II. RELATED WORK

In 2008 , [1]S. Greiner and J. Yang, proposed system preliminary works on privacy protection in video for supporting automatic dietary assessment in obesity studies. An approach to protect people's identities and contents on computer screens using object detection techniques has been proposed. The Adaboost algorithm is implemented from the OpenCV framework to build a system that allows detection of faces and screens in order to obscure them and make them unrecognizable in the recorded images. Although the described approach can provide robust detection of objects concerning privacy, it needs further improvements for real working systems. The classiffiers need to be improved, especially for occluded and rotated objects.

In 2011, P. Agrawal and P. J. Narayanan [3], studied issues relating to DE identification of individuals in videos are analyzed to protect their privacy by going beyond face recognition. A basic system is presented to protect privacy against algorithmic and human recognition. The studies indicate that gait and other temporal characteristics are difficult to hide if there is sufficient familiarity with the subjects and the user. Blurring is a good way to hide the identity if gait is not involved. Limitation of these proposed system is that for the deidentification manual identification of individuals is involved.

In 2012, H. Lang and H. Ling [4], proposed a study on novel image recognition/classification task, i.e. covert photograph classification is evolved. Comparing with the existing pornography/naked photograph recognition task, covert photograph classification is more challenging. Pornography is defined by subject of the photograph, whereas covert photography is defined by the acquisition method.

It is more difficult to learn clues which might reflect images acquisition method than to learn those based on image contents only.

In 2009, Andrea Frome in his paper on "Large-scale privacy protection in Google streetview,"[14]presented a system that addresses the challenge of automatically detecting and blurring faces and license plates for the purpose of privacy protection in Google Street View. The proposed system combines a standard sliding-window detector tuned for a high recall, low-precision operating point with a fast post-processing stage that is able to remove additional false positives by incorporating domain-specific information not available to the sliding-window detector.

In 2010, A. K. Moorthy and A. C. Bovik [15], proposed a new two-step framework for no-reference image quality assessment based on natural scene statistics (NSS). Once trained, the framework does not require any knowledge of the distorting process and the framework is modular in that it can be extended to any number of distortions. The framework is unique, since it assesses the quality of an image completely blind—i.e., without any knowledge of the source distortion. This is achieved by using distorted image statistics (DIS)—an extension of natural scene statistics for distorted images. The architecture of proposed framework is modular, and although have used only a few techniques for classification and QA, one can replace any module with a better-performing one.

In Feb. 2011, [12] M. Gnen and E. Alpaydin, proposed methods to combine many kernels instead of using single one. These kernels corresponds to different notations coming from similar or different sources having different representations or different subsets. The support vector machine (SVM) is a discriminative classifier is proposed for binary classification problems and is based on the theory of structural risk minimization. The multiple kernel learning is useful in practice and that there is ample evidence that better MKL algorithms can be devised for improved accuracy, decreased complexity and training time.

## III. ANALYSIS OF OUR WORK

**Module 1: Image Features Extraction**

In these module input image is converted into a hierarchical representation. After that, image features are being extracted by using various parameters.

The different image extracted features are as follows:

***1) Bag-Of-Features***

The SIFT [5] feature is implemented in for densely sampled local image patches. A vocabulary of size 200 is build using k-means. Finally, an image is represented by the histogram of word frequencies.

***2) Pyramid Histogram of Oriented Gradients***

A. Bosch, A. Zisserman, and X. Munoz in Representing shape with a spatial pyramid kernel [2], proposed objectives of classifying images by the object categories they contain.

***3) Gray Level Co-occurrence Matrix***

David A. Clausi and M. Ed Jernigan in, A Fast Method to Determine Co-Occurrence Texture Features [4], represent a critical shortcoming of determining texture features derived

 

ISSN : 2454-9150 Vol-03, Issue-04, July 2017

from grey-level co-occurrence matrices (GLCMs) is the excessive computational burden. Texture features calculated from grey-level co-occurrence matrices (GLCMs) are often used for remote-sensing image interpretation.

***4) Hue Descriptor***

Color histogram and texture features based on a co-occurrence matrix are extracted to form feature vectors. Then the characteristics of the global color histogram, local color histogram and texture features are compared and analyzed for CBIR. Based on these works, a CBIR system is designed using color and texture fused features by constructing weights of feature vectors.

***5) Local Binary Pattern***

The method is based on recognizing that certain local binary patterns termed uniform are fundamental properties of local image texture, and their occurrence histogram proves to be a very powerful texture feature. They derive a generalized gray scale and rotation invariant operator presentation that allows for detecting the uniform patterns for any quantization of the angular space and for any spatial resolution, and present a method for combining multiple operators for multiresolution analysis.

***6) Gray Histogram***

Gray scale is the monochromatic version of a color image where value of each pixel is between 0 to 255. 0 is for complete black and 255 for complete white. It is number of pixel Vs intensity. Histogram gives information about contrast of the image. Contrast is difference between highest and lowest pixel intensity.

**Module 2: Image Attribute Extraction**

In these module, after image feature extraction process, image attributes are extracted. The image attributes are extracted from four groups that include image quality, visual property, photographs and image content. The examples include, color richness, blur image, image saturation, dept of focus, image contrast etc.

**Module 3: Feature And Attribute Fusion**

MKL (Multiple Kernel Framework) is use to fuse the heterogeneous information obtain from extracted features and attributes. The main idea is to build a compound kernel from base kernel for two features vector $x_1$, $x_2$. The combination is given as follows:

$$\kappa_{\text{mk}}(\mathbf{x}_1, \mathbf{x}_2; \eta) = \sum_{k=1}^{K} \eta_k \kappa_k(\mathbf{x}_1, \mathbf{x}_2),$$

Where $\eta$ = Combination of coefficients to be learned.

$x_1$, $x_2$ = Feature vectors.

K = Base Kernel.

The overall system implementation is shown below:

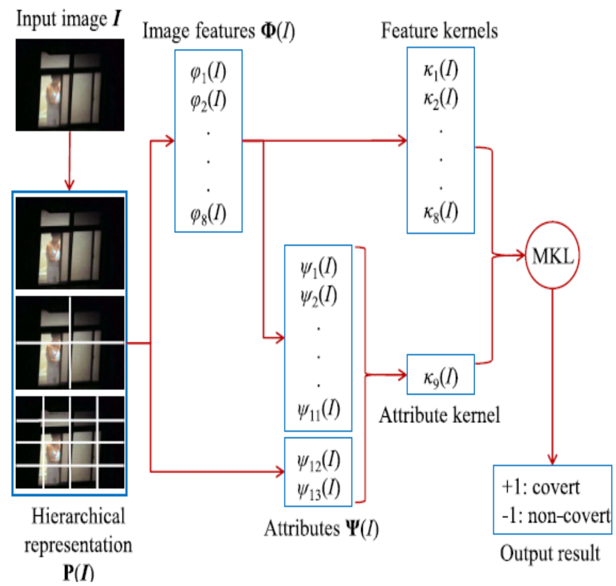


**Fig 1. Overview of System**

## IV. MATHEMATICAL MODEL

**Mathematical Model for Proposed Work**

**A: Problem Description**

Assumptions:

S: System; A system is defined as a set such that:

S = {I, F, O}.

Where,

I: Set of Input image

$I = \{I_1, I_2\}$

Where,

$I_1$: Hierarchical representation of image.

$I_2$: Gray scale image.

F: Set of functions.

$F = \{F_1, F_2, F_3\}$

Where,

$F_1$: Function to extract Features.

$F_2$: Function to extract attributes.

$F_3$: Function to train MKL

O: Set of output.

**B: Overall Algorithm**

Step 1: Input training Image.
Step 2: Represent the image into hierarchical format.
Step 3: Convert the image into gray level.
Step 4: Extract the image features.
Step 5: Extract image attributes.
Step 6: Perform feature and attribute fusion at MKL.
Step 7: Input Query image.
Step 8: Extract image feature and attribute.
Step 9: Compare the information with training dataset.
Step 10: Output result as covert or non-covert.

ISSN : 2454-9150 Vol-03, Issue-04, July 2017

**C: Description**

Initially a covert dataset has to be created by using following three steps:

**1) Initial Collection**: In these step a dataset of candidate covert photos is build. For this purpose, various image sources such as images retrieved over the Internet using relevant keywords (e.g., "voyeurs"), frames sampled from TV programs of secret investigation or surveillance records, photos contributed by volunteers for research purpose (without privacy violation), etc are to be checked.

**2) Covertness Classification**: An important issue with the initial covert collection is that some of the photos may not be truly covert, especially those collected by querying over the Internet. For each candidate in the initial collection, we need to trace carefully its acquisition process to make sure the process was hidden from the subject being photographed. A candidate photo is retained only if its covertness can be clearly verified. For the Examples, shown in Fig. 2, one of the two photos is verified by associated text description about its covert photographing process, while the other one is non-covert as indicated by the text.

**3) Bias Reduction**: Due to the special characteristic and the diversity of collecting sources, there are potential bias in the initial collection. For example, photos downloaded from a website are often captured in same scenes, photos contributed by an individual volunteer are often captured with a fixed camera model, etc. To reduce the biases in the initial dataset, following factors are taken into consideration including the race and ages of the subjects in photos, the time and environment of the capturing activities, the photo contents, and the capturing methods. On the other hand, a control is kept on the number of covert photos of any specific type to avoid its dominance. For example, in Fig. 2, some profile photos are discarded to avoid the dominance of such pose.

After following the above three steps a covert dataset is been created. Then is followed by automatic covertness classification.

1. Given an input image, firstly a spatial pyramid is generated P(I).
2. Then, for each sub-image a series of features extractor is being applied denoted as φ.
3. After that, visual attributes are being extracted denoted by ψ.
4. Finally there is a fusion of features and attributes using multiple kernel learning framework (MKL) to give final covertness classification.

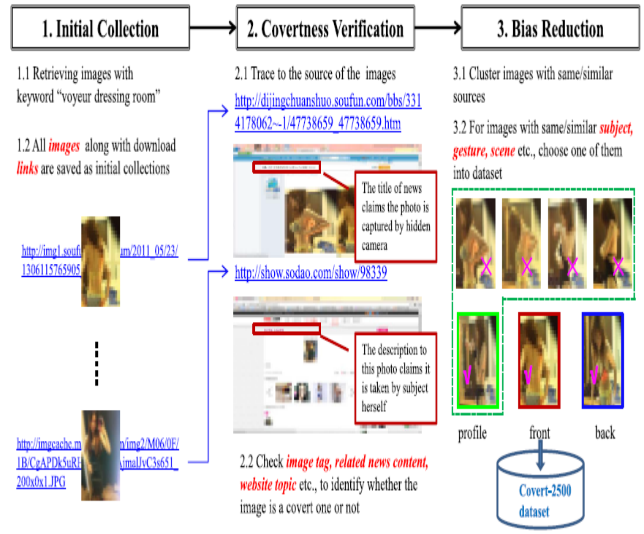


**Fig 2. Dataset Collection**

## V. Result and Analysis

The initial step is to compute a dataset of training images. After that the overall process of image feature extraction and image attribute extraction is being performed. Then the query image is given to the system for classification. Among the Bag-of-feature extraction the SIFT descriptor used in existing system is replaced by its newer version SURF in proposed system. There has been a drastic change and improvement in the proposed system result as compared to existing model.

The following graph shows the difference between existing and proposed model.

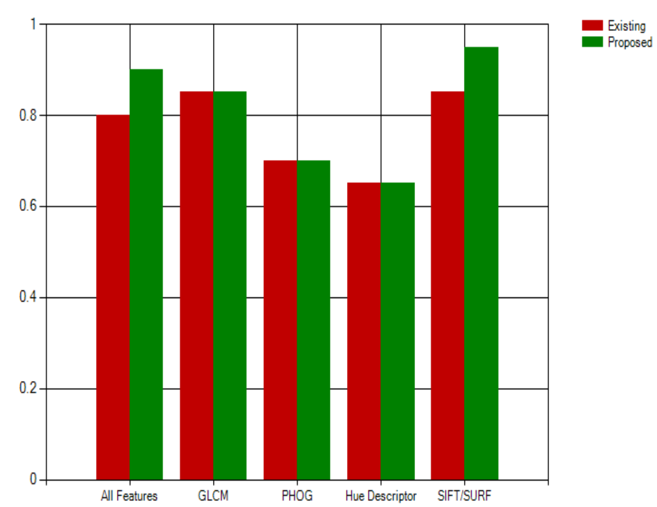


**Fig 3. Proposed and Existing system output graph**

The performance measure is computed uing precision and recal methods. Following table shows the difference between true positive and false positive rates of existing and proposed methods.

 

| Participant | TruePositiveRate | False Positive Rate |
|---|---|---|
| All Features Existing | 0.8000 | 0.2000 |
| GLCM Existing | 0.8500 | 0.1500 |
| PHOG Existing | 0.7000 | 0.3000 |
| Hue Descriptor Existing | 0.6500 | 0.3500 |
| SIFT Existing | 0.8500 | 0.1500 |
| GLCM Proposed | 0.8500 | 0.1500 |
| PHOG Proposed | 0.7000 | 0.3000 |
| Hue Descriptor Proposed | 0.6500 | 0.3500 |
| SURF Proposed | 0.9500 | 0.0500 |
| All Features Proposed | 0.9000 | 0.1000 |

**Table 1. True and false positive rate of proposed and existing system**

## VI. CONCLUSION

A mostly faced problem of classifying covert photos has been studied, which may threat public or personal privacy especially when distributed over the Internet. The challenges of the problem stem mainly from the between class semantic similarity and the intra-class photometric variation of covert photos. Firstly a dataset containing 2,500 covert photos is created. Then a human study is conducted on dataset, and laid down a benchmark human performance. Motivated by the observation in the human study, a covert photo classifier is proposed by integrating heterogeneous image features and visual attributes using the multiple kernel learning framework.

The proposed system was first given by the author. So it requires to implement some new methods in order to improve its scale of covertness classification and accuracy.

## ACKNOWLEDGMENT

First and foremost, I would like to thank my guide Prof. N. R. Wankhade for his guidance and support. I would also like to thank to my friends for listening my ideas, asking questions and providing feedback and suggestions for improving ideas. I wish to express my sincere thanks to the Head of department, Prof. N. R. Wankhade also grateful thanks to the departmental staff members for their support.

## REFERENCES

[1] S. Greiner and J. Yang, Privacy protection in an electronic chronicle system, in Proc. IEEE 34th Annu. Northeast Bioeng. Conf., 2008, pp. 8990.

[2] D. Chen, Y. Chang, R. Yan, and J. Yang, Tools for protecting the privacy of specific individuals in video, EURASIP J. Adv. Signal Process., vol. 1155, 2007.

[3] P. Agrawal and P. J. Narayanan, Person de identification in videos, IEEE Trans. Circuits Syst. Video Technol., vol. 21, no. 3, pp. 299310, Mar. 2011.

[4] H. Lang and H. Ling, Classifying covert photographs, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 11781185.

[5] D. G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis., vol. 60, no. 2, pp. 91110, Nov. 2004.

[6] J. van de Weijer and C. Schmid, Coloring local feature extraction, in Proc. 9th Eur. Conf. Comput. Vis., 2006, pp. 334348.

[7] R. M. Haralick, K. Shanmugam, and I. Dinstein, Textural features for image classification, IEEE Trans. Syst., Man, Cybern., vol. SMC-3, no. 6, pp. 610621, Nov. 1973.

[8] T. Ahonen, A. Hadid, and M. Pietikinen, Face description with local binary patterns: Application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 12, pp. 20372041, Dec. 2006.

[9] S. Lazebnik, C. Schmid, and J. Ponce, Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories, in Proc. IEEE Comput. Soc. Conf.

Comput. Vis. Pattern Recognit., vol. 2. Jun. 2006, pp. 21692178.

[10] K. Grauman and T. Darrell, The pyramid match kernel: Discriminative classification with sets of image features, in Proc. 10th IEEE Int. Conf. Comput. Vis., vol. 2. Oct. 2005, pp. 14581465.

[11] A. Oliva and A. Torralba, “Modeling the shape of the scene: A holistic representation of the spatial envelope,” *Int. J. Comput. Vis.*, vol. 42, no. 3, pp. 145–175, May 2001.

[12] M. Gnen and E. Alpaydin, Multiple kernel learning algorithms, J. Mach. Learn. Res., vol. 12, pp. 22112268, Feb. 2011.

[13] A. Vedaldi, V. Gulshan, M. Varma, and A. Zisserman, Multiple kernels for object detection, in Proc. IEEE 12th Int. Conf. Comput. Vis., Sep./Oct. 2009, pp. 606613.

[14] A. Former, A Guide to Large-scale privacy protection in Google streetview, IEEE 12th Int. Conf. Comput. Vis.,Oct. 2009.

[15] Wikipedia. (2011). Secret Photography. [Online]. Available: http://en.wikipedia.org/w/index.php?title= Secret photography; accessedOct:19; 2011.

 