Table of Contents

A Texture-based Segmentation Method for Improved Iris Recognition

Xiaojing Yuan, Ilyas Uyanik

Engineering Technology Department, University of Houston, Texas, United States

Abstract – Many research and algorithms have been developed for iris recognition in recent years due to advances in imaging acquisition and processing and renewed emphasis in biometric based security measure, given that iris recognition is still the most accurate biometrics used in security identification. The advances in imaging acquisition enabled more accurate and realistic iris images, while introducing new challenges for iris recognition due to reflection, which negatively affected the performance of the traditional segmentation methods used for iris recognition. In this paper, we present a new iris segmentation and pattern extraction algorithm that achieves high performance with or without reflection. The segmentation based iris recognition system developed in the paper streamlined while set of analysis algorithms for iris recognition from image preprocessing, feature extraction, enrollment, to recognition. The system achieved perfect recognition rate when tested on well-known iris recognition databases CASIA-IrisV1 and CASIA-IrisV3, provided by the Chinese Academy of Sciences.

Key Words – iris recognition, iris segmentation, normalization, 2D Gabor filter bank, texture feature extraction.

I. Introduction

Reliable and automatic recognition of individuals(a.k.a., personal identification) has long been a goal in law enforcement and security applications. As in other pattern recognition systems, the key challenge remains as to be able to model the relationship between interclass and intra-class variability accurately so that objects can be reliably classified with low error rate [1, 2].

Iris patterns became attractive as an alternative approach for reliable visual identification of a person with the advances in imaging acquisition. Using iris pattern as a biometric security measure has several advantages over other biometric measures such as face, finger-prints, voice-prints, or gait signature. First of all, the iris pattern variability among different person is enormous, thus satisfies the unique identification requirement. Secondly, as an internal, yet externally visible organ of the eye, the iris is well protected from the environment and stable over time. The extraordinary structure of the human iris provides abundant texture and spatial patterns that are unique to each individual. All these remain remarkably stable over many decades, making the iris pattern more reliable for personal identification [3]. Thirdly, from analysis perspective, as a planar object, the iris image is relatively insensitive to angle of illumination. Furthermore, changes in viewing angle cause only affine transformations. Even the non-affine pattern distortion caused by pupillary dilation is readily reversible [4]. Lastly, it is relatively easy to localize eyes in faces. The distinctive annular shape of the iris further facilitates reliable and precise isolation of it from the eye [5]. All these features make the iris recognition as personal identification more accurate and reliable as compared with using other biometric measures.

Furthermore, with advances in imaging acquisition, accurate iris images can be captured rapidly from a distance of about one meter nowadays, with ongoing research to increase that distance so that iris recognition can be extended to applications such as airport security monitoring from access security screening. Moreover, faster and efficient storage mechanism and search algorithms enabled searching for a match of a specific iris pattern in very large databases (and getting bigger) without incurring any false matches despite a huge number of possibilities. It is no surprise that given the increasing demand of more security and the advances in image acquisition and analysis iris recognition is getting renewed attention.

In this paper, we present our new segmentation method based iris recognition system that can deal with traditional iris image as well as more realistic iris image. Section II presents the framework of our iris recognition system that incorporates functions from image preprocessing, iris segmentation, feature extraction, enrollment, to recognition. Section III presents our segmentation method that extracts iris-pupil and iris-sclera boundary. Section IV describes the details of texture feature extraction based on two-

dimensional (2D) Gabor filter bank, including the normalization and image enhancement processes that ensure features extracted from iris images taken from different condition are comparable. Section V presents the iris database we use, the experiment setup and results. Conclusions will be drawn in Section VI with directions for future work.

II. Iris Recognition System Architecture

As shown in Figure 1, two main stages are necessary for any iris recognition system to function properly after iris images are acquired: enrollment and recognition. Each stage shares some common functions such as image pre-processing, iris segmentation, image enhancement, and feature pattern extraction. After extracting iris feature pattern, different actions are taken for enrollment and recognition stages. For enrollment, the extracted feature pattern is saved into the database, with pointers established to link to its corresponding raw iris image. For recognition, the extracted iris feature pattern will be compared to the pattern stored in the database using Euclidian distance, resulting in a similarity value. These similarity values will be ranked and the iris image with maximum similarity value or a set of iris images with similarity values above user defined threshold will be returned as the matched results. If all similarity values between the incoming iris image and those in the database are less than the threshold, the system will return "no match" result, and trigger the enrollment process automatically.
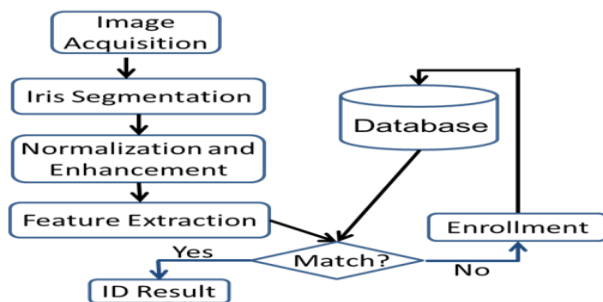

Fig. 1 Iris recognition system architecture

III. Image Preprocessing and Iris Segmentation

There are many pre-processing steps that are needed for any iris recognition system given that most iris images are acquired in the RGB color space, with little or none light and noise control. These include color mapping from RGB (red, green, and blue) to HIS (hue, saturation, and intensity) color space, noise reduction and artifacts (such as dense and intrusive eyelashes) removing.

Figure 2 shows iris images from the CASIA-IrisV1 database (left) and the CASIA-IrisV3 database (right). The iris locates between the sclera and the pupil. In order to extract iris feature pattern, iris needs to be segmented out by identifying the boundaries between iris-pupil and iris-sclera, which is not trivial. As shown in Figure 2, the iris images from CASIA-IrisV3 dataset (right) have light reflection in the pupil, making them more realistic; while the iris images from CASIA-IrisV1 dataset (left) have the light refection removed beforehand. The technical challenge lies in answering the question of "how can we accurately and efficiently localize the iris boundary without additional pre-processing step to remove the light reflection in the pupil?"[14] An efficient and accurate iris segmentation algorithm or mechanism for CASIA-IrisV3 dataset can be easily adapted for other realistic iris database such as UBIRIS-V2 and Iris Challenge Evaluation dataset used in 2005 (ICE2005) and 2006 (ICE2006).

As shown in Figure 2, iris segmentation requires identification of boundaries between the iris and pupil (inner boundary) and iris and sclera (outer boundary). In the next subsections, we will present in detail both boundary localization and illustrate the difference between CASIA-IrisV1 and CASIA-IrisV3.

3.1 Iris-Pupil Boundary Localization

For iris image from the CASIA-IrisV1 dataset, the boundary between iris and pupil can be identified simple by applying any edge detection algorithm since the pupil area has homogenous low intensity value (usually the lowest intensity value in the whole image). Figure 3 shows the process of iris-pupil boundary localization with a sample iris image from CASIA-IrisV1.

Using $I(x,y)$ to represent the pixel intensity in the image, in which x and y are the Cartesian coordination of the pixel, we identify the minimum intensity value ($\min(I(x,y))$) and its location in the iris image. For iris images from CASIA-IrisV1dataset, such a pixel always locates in the pupil. Figure 3 (b) shows the heat-map of each pixel based on its intensity value for the sample iris image. Using the location of the pixel with minimum intensity value as the new origins (as pointed by arrows in Figure 3 (c)), a threshold value of ($\min(I(x,y))+N$) is then used to identify the boundary between the iris-pupil based on the histogram in the x- and y- direction respectively. N is empirically chosen. Note that instead of counting.
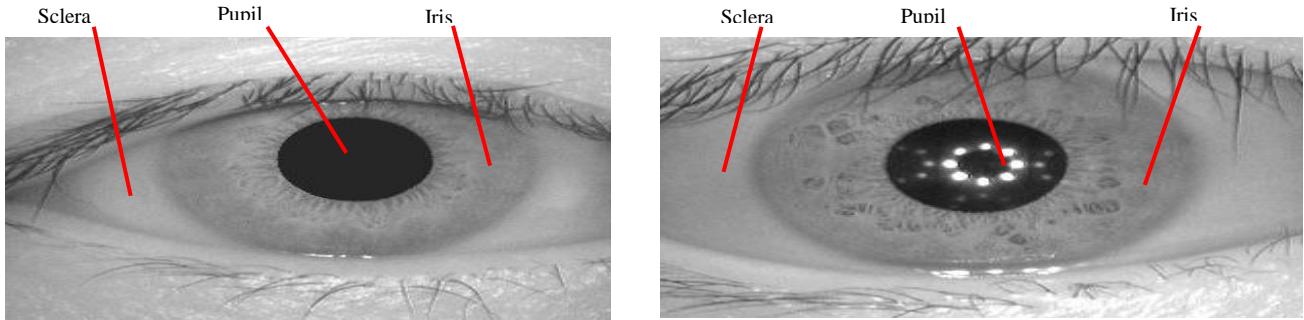
Fig. 2 Iris Image examples from CASIA-IrisV1 (left) and CASIA-IrisV3 (right) databases
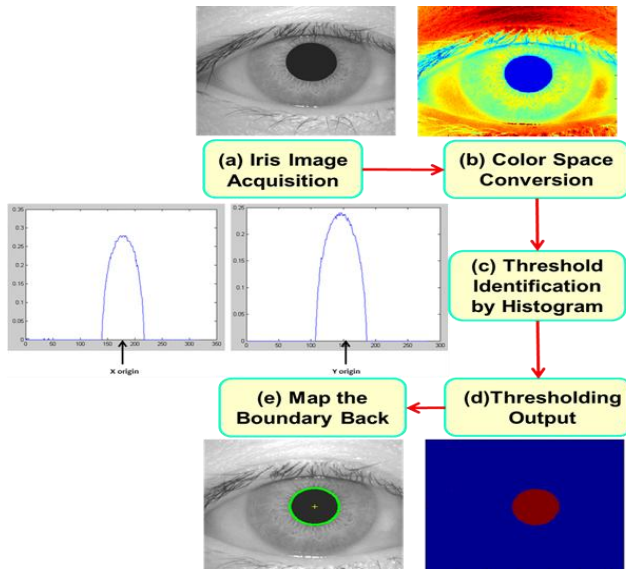

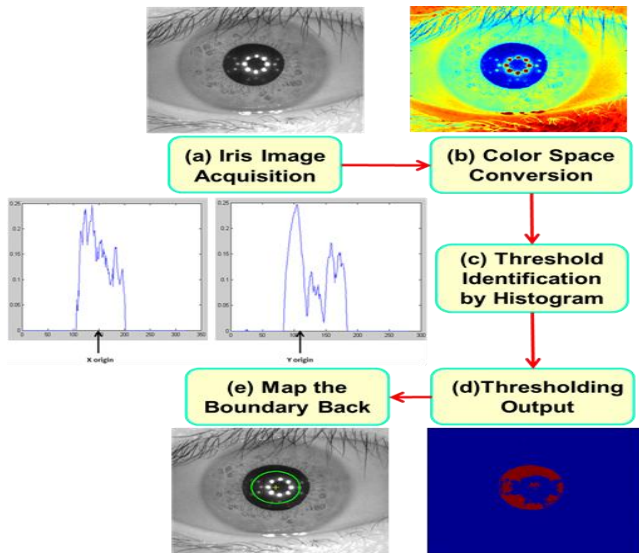Fig. 3 Iris-Pupil Boundary Localization: CASIA-IrisV1

Fig. 4 Iris-Pupil Boundary Localization: CASIA-IrisV3

based on the pixel intensity directly, Figure 3 (c) shows the percentage of the pixels below the threshold in each direction. Figure 3 (d) shows the extracted pupil area using the simple thresholding methods based on the threshold identified. Figure 3 (e) illustrated the iris-pupil boundary in green circle after mapping the identified boundary back to the original iris image

For iris images from the CASIA–IrisV3 dataset, even though the iris pattern remains clear, the iris-pupil boundary localization is much more difficult due to the light reflection within the pupil. Figure 4 shows the results when applying the same algorithm to iris image from CASIA-IrisV3 dataset. Because of the reflection captured in the pupil area, it does not provide homogenous low intensity area. Hence, the accuracy of identifying the iris-pupil boundary using histogram decreases, as shown in Figure 4(d) and (e).

To address these issues, we introduced two basic mathematical morphology (MM) operations: dilation and erosion, before threshold identification step. MM theory and techniques are used in the analysis and processing of geometrical spatial structures based on set theory, lattice theory and topology. Used in imaging processing, the MM operations consider the image as a set of pixel values (primary set), and apply pre-defined shape (secondary set) to see whether it fits or misses the region of interest (object) in the primary set. The advantage of basic MM operations is that they are translation invariant.

The dilation operation of a secondary set B on a primary set A is defined by $A \oplus B = \bigcup_{b \in B} A_b$. In Matlab, the dilation operation is implemented so that the value of the output pixel is the maximum value of all the pixels in the input pixel's neighborhood, resulting in expanded boundary.

Similarly, the erosion operation is defined by $A \ominus B = \{z \in E | B_z \subseteq A\}$, where B_z is the translation of B by the vector z, i.e., $B_z = \{b + z | b \in B\}, \forall z \in E$. In Matlab, the erosion operation is implemented so that the value of the output pixel is the minimum value of all the pixels in

the input pixel's neighborhood, resulting in shrank boundary.

To address the light reflection in the pupil, we first applied the dilation operation with the secondary set B chosen as an open disk with radius of 3, with its center moved around the identified boundary as shown in Figure 5(a). Result is shown in Figure 5(b), including the resulting binary image and the resulting image after the boundary is mapped back to the original iris image. Afterwards, erosion operation is applied to shrink the boundary with the radium of the open disk representing the secondary set B chosen as 2, with results shown in Figure 5 (c). Figure 5 also compares the histogram from original algorithm with that after the dilation and erosion operations. The resulting histogram shows smooth region within the pupil, indicating intensity homogeneity. The histogram is used to identify the accurate iris-pupil boundary as before, with resulting boundary shown in Figure 5 (c).

3.2 Iris-Sclera Boundary Localization

After identifying the boundary between the iris-pupil, boundary between the iris and sclera needs to be identified. It is generally agreed that the iris-sclera boundary localization is more difficult than iris-pupil boundary identification [6] because of the low contrast between iris and sclera and intensity heterogeneity of iris. Various segmentation algorithms [7], including active contour and Hough transform, have been applied to iris-sclera boundary detection. The main drawback of these algorithms is their computational

complexity, resulting in more computation time and memory space.

In our iris recognition system, we identify the iris-sclera boundary by applying the Sobel operator. A discrete differentiation operator, the Sobel operator convolving the image with a small, separable, and integer valued filter in horizontal and vertical direction to approximate the gradient of the image intensity function. Compared with other edge detection algorithms, it is computationally inexpensive. The gradient approximation Sobel operator produces is crude, i.e., the identified edge or boundary is not precise. However, it is good enough for iris-sclera boundary identification since even high-noised imperfect iris images do not have high- frequency variations. Figure 6 shows the iris-sclera boundary identification result, with identified iris-pupil boundary, for a sample iris image from CASIA-IrisV3 dataset.

IV. Image Enhancement and Feature Extraction

4.1 Iris Normalization and Image Enhancement

Iris images are increasingly acquired without strict constraints, resulting in more intra-class variations caused by illumination (camera gain, contrast settings), position and pose, and other external noise. To enable comparison of iris images taken under different conditions, the iris normalization process tries to minimize intra-class variance without compromise inter-class variance (difference from
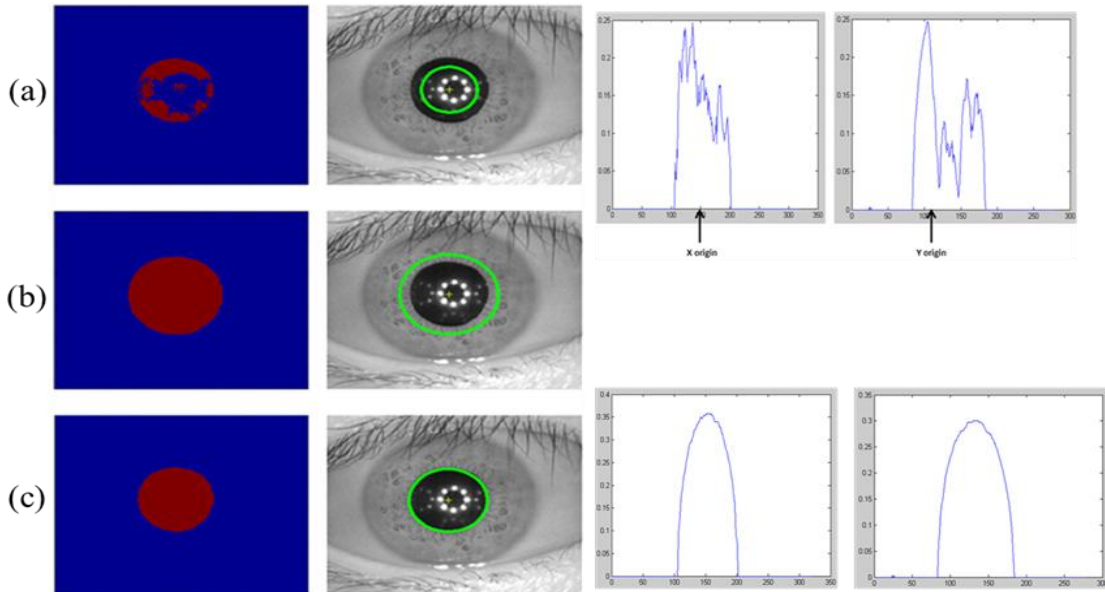

Fig. 5 Iris-Pupil Boundary Localization after MM operations: (a) original, (b) dilation, and (c) erosion
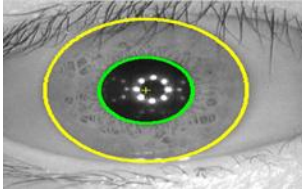
Fig. 6 Iris-Sclera boundary identification result of a sample iris image from CASIA-IrisV3 dataset

person to person) and outputs a normalized iris sheet with the same constant dimensions. The normalization step ensures that the characteristic features of iris pattern extracted will be the same at the same spatial location regardless of how the original iris images were acquired.

In our iris recognition system, we use Daugman model [5] for the normalization process. The Daugman model maps each pixel within the segmented iris region to a pair of polar coordinates (r, θ), in which r $\in [0,1]$ and $\theta \in [0,2\pi]$. The mapping process and equation are shown in Figure 7, in which the normalized iris sheet is displayed as a rectangular image with size M*N, with the radial coordinate on the vertical axis, and the angular coordinate on the horizontal axis [9].

By taking the pupil dilation and size inconsistence (intra-class variation) into consideration, the Daugman method models the iris region as a flexible rubber sheet anchored at the iris boundary with the pupil center as the reference point, resulting in a normalized representation of iris with constant dimensions. In our experiment, we quantize the radius in the range of [1, 64] and the orientation θ in the range of [1-512], resulting in normalized iris sheet with constant dimension of 512 X 64. Figure 8 shows an example of the normalized iris sheet of the sample iris image from CASIA-IrisV3 dataset.

After mapping the extracted iris region from Cartesian representation I(x,y) to polar representation I(r, θ), a low pass Gaussian filter with window size of MxM is applied to the normalized iris sheet first. The intensity of the central pixel within the window of the Gaussian filter is replaced by the mean intensity of the pixels within the window, resulting in the iris background bias sheet (Figure 8 (b)). Then, the enhanced iris sheet is obtained by subtracting the iris background bias sheet from the normalized iris sheet (Figure 8 (c)).

4.2 Iris Texture Feature Pattern Extraction

Many feature extraction algorithms have been proposed to extract unique and invariant features from the iris patterns, including variations of Gabor, wavelet filter [10-13] extracting texture and spatial features. Gabor filters are well known as texture feature extraction method that produces feature invariant to scaling, shift, and rotation [8]. Equation (1) is used to compute Gabor filter response of a two-dimension (2-D) image in the Cartesian domain.

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{1}{2}\left(\frac{x_1^2}{\sigma_x^2} + \frac{y_1^2}{\sigma_y^2}\right) + 2\pi f x_1\right) \quad (1)$$

in which $x_1 = x\cos\vartheta + y\sin\vartheta$ and $y_1 = -x\cos\vartheta + y\cos\vartheta$, and f is the frequency and θ is the orientation of the filter.

Gabor filter responses can be divided into two parts: the real part and the imaginary part, as shown in Equation (2)-(4).

$$g(x, y) = g^e(x, y) + jxg^o(x, y) \quad (2)$$

$$g^e(x, y) = \frac{1}{2\pi\sigma_y\sigma_x} \exp[-\frac{1}{2}(\frac{x_1^2}{\sigma_y^2} + \frac{y_1^2}{\sigma_y^2})]\cos 2\pi F x_1) \quad (3)$$

$$g^0(x, y) = \frac{1}{2\pi\sigma_y\sigma_x} \exp[-\frac{1}{2}(\frac{x_1^2}{\sigma_y^2} + \frac{y_1^2}{\sigma_y^2})]\sin 2\pi F x) \quad (4)$$

In our iris recognition system, we use 2-D Gabor filter with five central radial frequencies f: ([0.157, 0.346, 0.691, 1.414, 2.827]) and four orientations ϑ: ([0 45 90 135]). Each enhanced normalized iris sheet is divided into eight equal sized patches (32x64). Each patch then is fed into the Gabor filter bank and produces 20 filter responses with the same size. For the filter response of each patch, the average absolute deviation (AAD) defined in Equation (3) is used to represent its texture feature:

$$E = \frac{1}{N}(\sum_N |G(x,y) - \mu|) \quad (5)$$

in which N is the size of the patch and its filter response, μ is mean filter response of the patch, and G(x,y) is the Gabor filter response coefficient at pixel (x, y).

For one iris image, we get a feature vector with 160 ADD values, representing the texture pattern of the enhanced normalized iris sheet. Figure 10 shows the Gabor filter bank based texture feature extraction and a sample feature vector with length of 160 and ADD values ranges from 0 to 0.03. The extracted feature vector for the iris image will be stored into the database for enrollment or compared to the texture feature vectors for iris images already in the database for matching and recognition.

$$I(x,y) \rightarrow I(r,\theta)$$
$$r = \sqrt{(x - x_c)^2 + (y - y_c)^2} \, , \; 0 \le r \le r_{max}$$
$$\theta = tan^{-1}\left(\frac{y-y_c}{x-x_c}\right),$$

$(x_c, \; y_c) \; is \; the \; Cartesian \; coordination \; of \; the \; center \; pixel.$
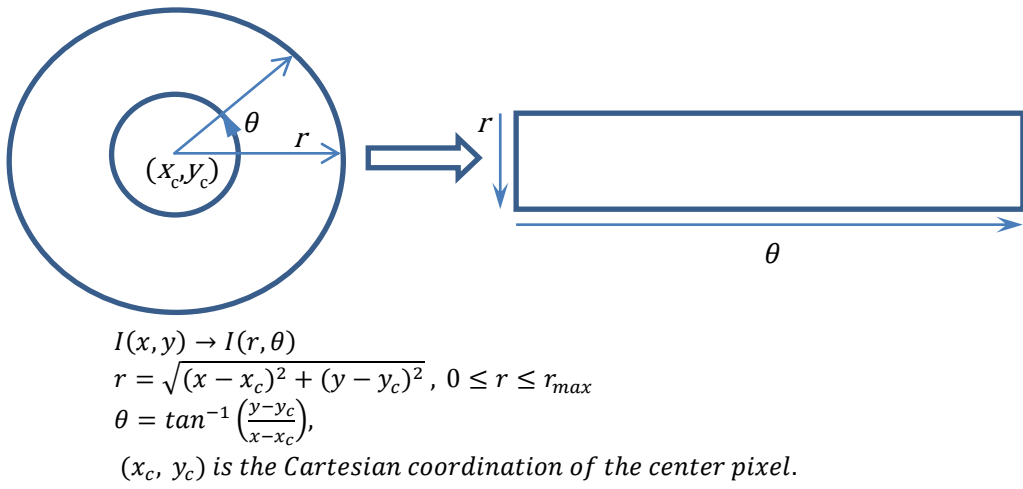
Fig. 7 Daugman Model based normalization process and equations

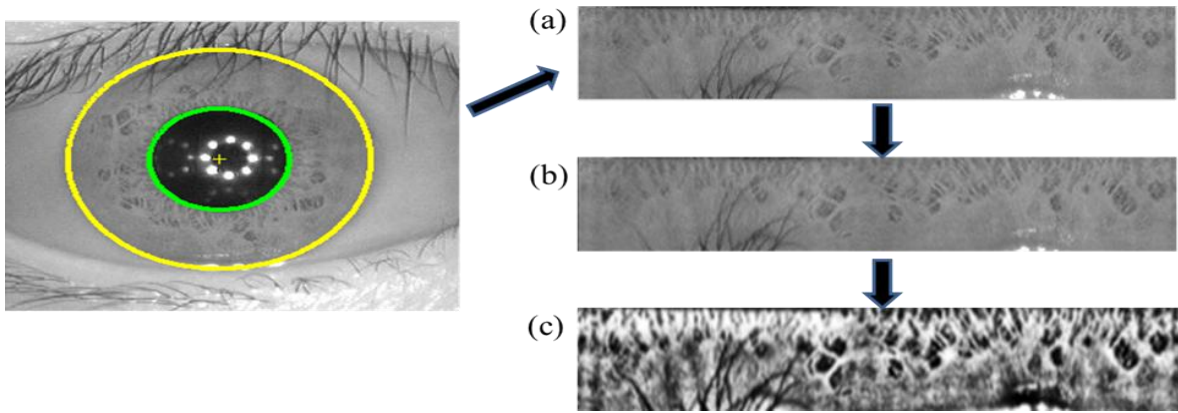Fig. 8 (a) Normalized iris sheet from Daugman normalization of a sample iris image from CASIA-IrisV3 dataset;
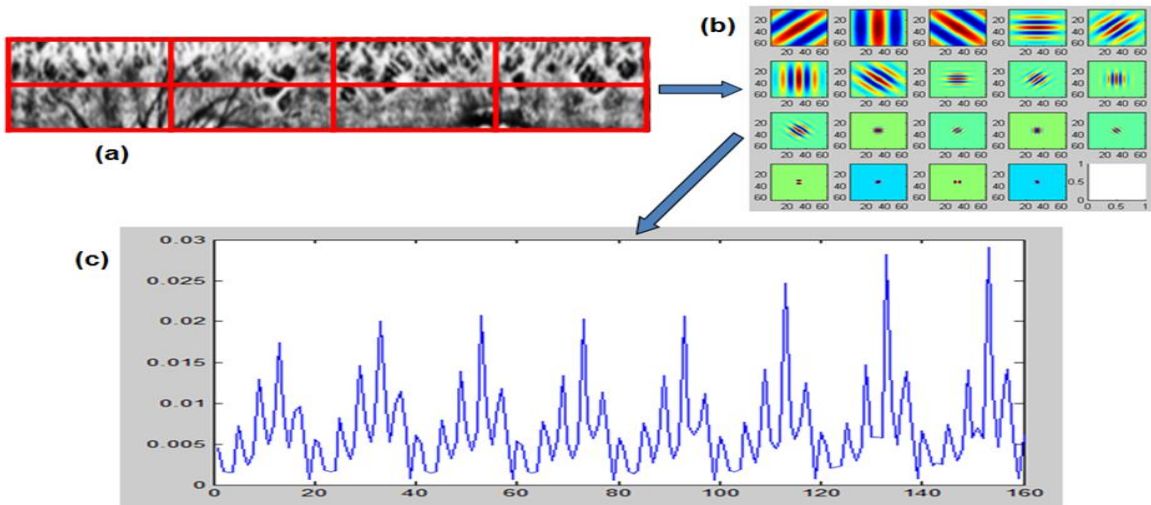(b) iris background bias sheet; and (c) enhanced iris sheet

Fig. 10   2D Gabor filter bank based texture feature extraction: (a) Enhanced Normalized Iris Sheet divided into 8 patches;
(b) 20 2D Gabor Filters; and (c) Texture feature vector with 160 average absolute deviation (ADD) values

V. Experiment Setup and Results

In this section, we describe the iris databases used to evaluate the performance of our iris recognition system, including both enrollment and matching. Then we summarize the parameters used and present the results for matching.

5.1 Iris Datasets

To evaluate the performance of our iris recognition system, we selected two iris databases, namely CASIA-IrisV1 and CASIA-IrisV3 [5]. They are chosen because the iris images in these databases encompass most of the irregularities caused by different acquisition instruments specification and various acquisition conditions. The databases also contain iris images captured from different ethnicity.

5.2 Experiment Setup and Results

Table1 summarizes the experiment setup we used in the evaluation. These values are used in the evaluation of our iris recognition system, including both enrollment and matching. All parameters were chosen empirically to best fit for iris images from both CASIA-Iris databases. Both the enrollment and matching process use the same iris segmentation and texture feature extraction mechanism described. For enrollment process, after the texture-based feature vector with 160 ADD values is extracted, it will be saved into the database with the original iris image with iris owner identity information. For the matching process, the feature vector will be extracted from the incoming new iris image before similarity value based on Euclidean distance is calculated between the new iris and the feature vector of each of the iris image already in the database. Only the iris image with the highest similarity value will be returned as the matching one.

Figure 11 shows a simple user interface we designed in Matlab to enroll and match the iris. After the iris image files are read into the system, both the enrollment and matching process take little time. Two additional functions are also included in the Matlab application. First, when there is no matching found, enrollment process will be triggered and the user will be prompted to input the iris owner identify information. Secondly, image acquisition capability is included so that in addition to reading iris image from the available database, new iris image can be acquired directly and fed into the iris recognition system to

enroll or find matching. This function simulates the real world security scenario where iris images are taken from new individuals that have never been registered in any databases. Figure 11 also shows the iris image acquired directly from a volunteer in the lab using an embedded webcam on a laptop, and the prompt asking the user to "Enter Iris Owner Identity".

Table 1 Parameters used in the Evaluation of Our Iris Recognition System

Para.	Value	Description
N	35	Iris-pupil boundary identification, used in threshold min $(I(x,y))+N$
R_D	3	Secondary disk radius for dilation operator
R_E	2	Secondary disk radius for erosion operator
r	64	Radius range of radial coordination for normalized iris sheet
θ	512	Angle range of radial coordination for normalized iris sheet
M	16	Window size of low pass Gaussian filter used in image enhancement
m x n	32x64	Patch size on which 2D Gabor filter bank is applied
ϑ	[0 45 90 135]	Four orientations used in 2D Gabor filter bank
f	[0.157, 0.346, 0.691, 1.414, 2.827]	Five central radial frequencies used in 2D Gabor filter bank

Using the dataset chosen from CASIA-IrisVx, our algorithms achieved 100% recognition rate. When the system is tested against iris images acquired directly from volunteers, the recognition rate hovers around 90%, after enrolling the first iris image for the individual. We believe the reduction in the success rate is mainly due to the uncertainty introduced by image quality decline caused by camera settings of embedded webcam on laptop, no control of the ambient lighting and no strict requirements for the pose of the volunteers.
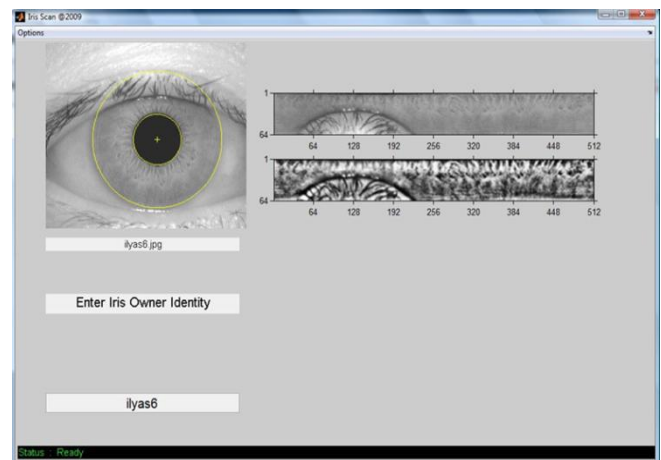

Figure 11 Iris Recognition Implementation in Matlab: IrisScan

VI.   Conclusion

In this paper, we presented an iris recognition system based on improved iris segmentation and feature extraction algorithms. Details of iris localization, normalization, enhancement, and texture based feature extraction are presented. Our iris segmentation algorithm shows comparable accuracy as those computationally more intensive algorithms such as Hough and active contour. Our iris recognition system implementation in Matlab, IrisScan, achieved high accuracy with minimum computational time.

A future direction for the iris recognition is to develop algorithms that can provide descent recognition success rate in a timely manner when dealing with iris images taken from moving subjects such as those from security surveillance video.

ACKNOWLEDGEMENT

REFERENCES

[1] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: the problem of compensating for changes in illumination direction," IEEE Trans. Pattern Anal. Machine Intell., vol. 19, pp. 721–732, July 1997.

[2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class-specific linear projection," IEEE Trans. Pattern Anal. Machine Intell., vol. 19, pp. 711–720, July 1997.

[3] A.K. Jain, R.M. Bolle, and S. Pankanti, Eds., Biometrics: Personal Identification in a networked Society, Norwell, MA: Kluwer, 1999.

[4] G.Gürkan, A. Akan, "Texture Analysis Based Iris Recognition", Istanbul University Journal of Electrical and Electronics Engineering vol. 6, No 1, 2006.

[5] J.Daugman, "How Iris Recognition Works," IEEE Trans. On Circuits and Systems for Video Technology, vol. 14, No. 1 Jan 2004.

[6] K.S.S.Kyaw, "Iris Recognition using features for biometric Identification" IEEE Electronic Computer Technology, pp. 554–556, Feb 2009.

[7] J.Daugman "New Methods in Iris Recognition", IEEE Transaction and System Man and Cybernetics: Part B vol.37 No.5 , October 2007.

[8] H.Zheng, F.Su, "An Improved Recognition System Based on Gabor Filters" IEEE Network Infrastructure and Digital Content, pp. 823–827, Nov 2009.

[9] K. W. Bowyer, K.Hollingsworth, P.J.Flynn "Image Understanding for Iris Biometrics: Survey", Computer Vision and Image Understanding, vol.110 (2), 2008.

[10] R. P.Wildes, "Iris recognition: An emerging biometric technology," Proc. IEEE, vol. 85, no. 9, pp. 1348–1363, Sep. 1997.

[11] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Personal identification based on iris texture analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 12, pp. 1519–1533, Dec. 2003.

[12] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Efficient iris recognition by characterizing key local variations," IEEE Trans. Image Process., vol. 13, no. 6, pp. 739–750, Jun. 2004

[13] J. Daugman. High confidence visual recognition of persons by a test of statistical independence. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, No. 11, 1993.

[14] X. Yuan, Segmentation of Blurry Object by Learning from Examples, SPIE Medical Imaging, San Diego, CA, Feb. 13-18, 2010

Contextual Anomaly Detection Cueing Methods for Hyperspectral Target Recognition

Kelly D. Friesen, Trevor J. Bihl, Kenneth W. Bauer, Mark A. Friend

Department of Operational Sciences, Air Force Institute of Technology, Wright Patterson AFB, OH, USA

Abstract – Detection and recognition are two primary tasks of hyperspectral imaging, with the objective of detecting and identifying materials remotely. Frequently these are done separately, with anomaly detection or signature recognition being applied, but not both in conjunction. Here, we introduce a fully automatic hyperspectral-based target recognition system through the combination of automatic anomaly detection, anomaly segmentation, pixel cueing methods for background reduction filtering, signature matching, automatic atmospheric compensation, and vegetation index thresholding for automatic atmospheric compensation. Of interest herein are the handling of pixels and groupings of pixels detected as a statistical anomaly; various methods of processing anomaly groups are examined. Results are provided using HYDICE hyperspectral radiance images and collected ground reflectance data.

Key Words – ATD, ATR, Automatic Target Recognition, Automatic Target Detection, chipping, cluster identification, cueing, hyperspectral, remote sensing.

I. Introduction

Hyperspectral imaging (HSI) systems are a form of imaging spectroscopy, collecting both spatial and spectral features; in essence HSI operates as a panchromatic sensor with individual spectrometers for each pixel [1]-[3]. Other modalities, such as multispectral imaging (MSI) cover a large portion of the spectrum; however MSI bands are collected with coarse spectral sampling [2], whereas HSI has a multitude of finely sampled spectral bands. Fine spectral sampling enables measurement of minute spectrally reflected and/or emitted features, offering the potential to remotely detect, examine, and identify materials through differing spectral characteristics. In operation, HSI sensors collect spectral characteristics sequentially, building a three dimensional image object known as an image cube. In such an image cube, notionally, the x and y axes contain spatial information, while the z axis contains spectral data. HSI sees frequent application for three primary purposes: anomaly detection, change detection, and spectral signature matching [1]. The research herein focuses on a combination of anomaly detection and spectral signature matching. Automatic anomaly detection (ATD) seeks pixels significantly different than the rest of the image; in contrast automatic target recognition (ATR) involves spectral signature matching to identify spectral similarity of a pixel to a known material.

Applications of cueing in ATD and ATR frequently include focusing an analyst's attention to a particular pixel or location; an example of automatic HSI ATD cueing includes the search and rescue operations of the Civil Air Patrol's ARCHER program [1]. Cueing methods such as this require a 'user in the loop,' and are naturally susceptible to collecting more data than can be analyzed, confirmation bias [4] and/or operator fatigue [5]; such situations can lead to a preponderance of data, yet lack of information [6]. As a variant, cueing in this paper refers to algorithm cueing, where signature matching algorithms are cued to process detected anomalies. The outcome of this autonomous method, Fig. 1 [7], would then be sent to an analyst. Some of the methods leading from ATD cues to recognition involve combining multiple spectra for a detected anomaly; this is akin to combining multiple vegetation spectra for atmospheric compensation.

Ratches [8] offered a similar generic path from input to detection to classification; however a methodology was not detailed. Chang and Chiang [9] developed an anomaly classification method based on detected statistical anomalies; however, this classification scheme was based on in-scene classes with anomalous points clustered based on statistics.

The methodology presented in Fig. 1 incorporates modules facilitating the testing of various functions and combinations, with Step 0, loading image and ground truth data, and Step 1a being automatic atmospheric compensation. Anomaly detection, Step 1b in Fig. 1, is applied as a background reduction filter to identify probable objects of interest while discarding apparent background; such a method was applied by Manolakis et al. [10] for improved signature matching. The identified anomalous pixels and groups of pixels are then classified in Step 2,

through established signature matching algorithms. Although the algorithms used in this study are from the literature, the end-to-end automation, flexibility of changing algorithms, vegetation thresholding method for automatic atmospheric compensation, and the method for selecting and grouping detected anomalies
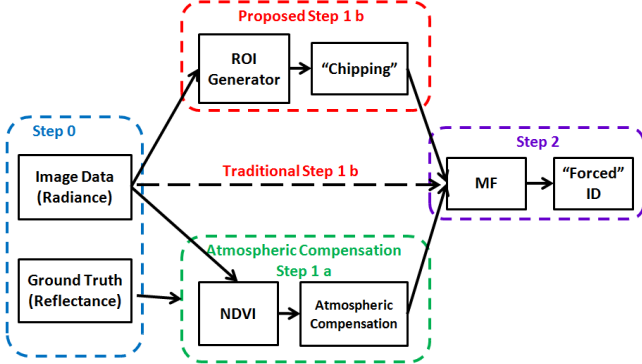

Fig. 1 Automated methodology

for signature matching are unique.

In HSI ATD, anomalies can consist of individual pixels or groupings of pixels. In this research, particular attention is given to reducing the number of points under consideration in ATR by first processing the image with an ATD algorithm. Comparisons of applying matched filters to the entire image to various means of using an anomaly detector for background reduction, removing statistical background information are provided. Emphasis is given to automating the entire process, with a minimum amount of user input.

This paper is organized as follows: Section II reviews HSI, characteristics of the data, and dimensionality reduction for creating a parsimonious dataset for analysis. Section III describes the automatic atmospheric compensation applied, while Section IV describes automatic target detection. Section V describes signature matching and aspects of the combined methodologies presented in Fig. 1. Concluding remarks and discussions appear in Section VI.

II.   Hyperspectral Remote Sensing

The HSI images used herein, Fig 2, are from the Forest I and Desert II Radiance collections from the HYDICE pushbroom sensor; these images provide radiance data for 210 bands with 10nm spectral sampling over 400 – 2500 nm, thereby covering the visible, NIR and short wave infrared regions [3] [11]. The images used for this research were both collected at approximately 5,000' AGL.

Provided with the Forest I and Desert II radiance images are 5nm spectrally sampled field spectrometer reflectance measurement files for in-scene objects and background. For modeling vegetation, reflectance data for in-scene sweet gum trees and creosote bushes were used respectively for the Forest and Desert images. Known man-made objects of interest total: 10 in the Desert and 21 in the Forest image. In this study, one object class of interest was selected for each image, with 223 and 348 target pixels respectively provided within the Forest and Desert images; yielding percentages of 0.18% (Forest) and 0.309% (Desert) of total pixels being known in-scene target pixels.

Fig. 2 Images used in analysis. the image on the left will be referred to as Forest (611 by 201 spatial pixels); the image on right (401 by 281 spatial pixels) as Desert.

2.1   Dimensionality Reduction

Fringe bands (at the edges of the spectral regions), bands associated with water absorption, and noise filled bands were identified (bands 1-9, 98-114, 133-157, 201-210) and discarded in similar fashion as in Johnson [15] and Farrell and Marsereau [16], resulting in 149 bands available for analysis. Although discarding said bands removed 61 bands from the HYDICE images (29% of the image), there is not necessarily a parsimonious dataset present. To reduce dimensionality further for efficiency in a computational sense while preserving variability (in this case energy measured), Principal Component Analysis (PCA) was used.

PCA is frequently applied to HSI due to its straightforward and predictable behavior [3][16]-[18]. PCA preserves and groups a maximal amount of variance through a linear transformation of data, using the eigenvectors of the data correlation matrix [19]. Principal component (PC) bands are orthogonal to each other, with the first PC containing the most variance, the second PC the second most variance and

so forth [19]. In terms of dimensionality assessment, the cumulative variance explained and the magnitudes of the image covariance matrix eigenvalues through a Scree plot are often used [19]. For the Forest image, 10 PC bands were retained, equivalent to 99.85% of the variance in the HSI image cube; similarly for the Desert image, 10 PC bands were retained preserving 99.78% of the variance in the image.

2.2 Normalized Difference Vegetation Index (NDVI)

Due to the presence of water, chlorophyll, and other pigments, vegetation typically appears green in visible wavelengths due to light absorption properties; also, between 700nm and 1300nm vegetation typically acts as a scatterer, causing vegetation to appear bright [3][12][13]. Linear vegetation indices take advantage of this property, measuring the slope between the absorbed visible light and the mostly reflected NIR light with high values corresponding to probable vegetation. One method, NDVI, developed by Rouse et al. [14], detects instances of probable vegetation using red and NIR bands. NDVI is regarded as having a high dynamic range and sensitivity to changes in vegetation cover [12]; it is described by a linear relationship

$$NDVI = \frac{L(NIR) - L(Red)}{L(NIR) + L(Red)}, \qquad (1)$$

where L refers to radiance at selected *NIR* and *Red* wavelengths bands, respectively [1][12][14]. Scores for NDVI vary from -1 to 1, the higher magnitudes typically being associated with probable vegetation.

III. Atmospheric Compensation for ATR

Remotely measured HSI pixels inherently include additional effects from external sources, requiring radiative transfer considerations [1]-[3]; included in radiance images are atmospheric influences, weather effects, illumination, path radiance, and adjacency aspects, among others [1][12][23]. However, for signature matching known materials are frequently analyzed by collecting reflectances, a unitless ratio of the amount of light reflected to light striking an object, requiring a known source and controlled conditions [12]. Atmospheric compensation, Step 1a in Fig. 1, is therefore a necessity when searching for a known signal collected using ground equipment to convert ground measurements of reflectance to the same unit space as angle per area [1][3][12][23].

Methods of atmospheric compensation frequently rely on *a priori* knowledge in the form of a known

target being in-scene for reference, as in Empirical Line Methods, or appropriate parameter selection for model based approaches. In a purely automatic system, such approaches are not appropriate and therefore the vegetation normalization (VN) method, which leverages vegetation detection, was used [1][3]. While VN is considered to offer poor quality results [3], an advantage is achieved in the ease of automation. The VN approach used herein applies a linear conversion from reflectance to radiance employing in-scene pixels estimated to be vegetation with appropriately selected known reflectance data for vegetation. The linear relationship between estimated pupil-plane radiance and reflectance employs equation (2),

$$\hat{L}(\lambda) = \hat{a}(\lambda)\rho(\lambda) + \hat{b}(\lambda), \qquad (2)$$

where $\hat{a}$ is a linearly computed gain vector, $\hat{b}$ is a linearly computed offset vector, with $\hat{L}$ the estimated pupil-plane radiance of the given reflectance spectra vector, ρ, each as a function of wavelength λ [1][3]. The gain and offset were calculated through linear relationships, with

$$\hat{a}(\lambda) = \frac{L_2(\lambda) - L_1(\lambda)}{\rho_2(\lambda) - \rho_1(\lambda)} \qquad (3)$$

and

$$\hat{b}(\lambda) = \frac{L_1(\lambda)\rho_2(\lambda) - L_2(\lambda)\rho_1(\lambda)}{\rho_2(\lambda) - \rho_1(\lambda)} \qquad (4)$$

where the subscripts 1 and 2 indicating appropriately corresponding radiance and reflectance data sources [1][3].

3.1 Vegetation Index Thresholding

Due to a desire to have consistent operating points among different types of images, raw NDVI thresholds for VN are possibly inappropriate given differing distributions of vegetation across environments and scenery types, an NDVI score associated with a forest image would possibly result in few if any pixels being selected as vegetation in a desert image. An alternative method of selecting a percentage bound of the total pixels to be considered as vegetation was adopted. As an example of the use, this research selects 1% of the top scoring NDVI values for retention as vegetation; this operating point equates to 1,228 points for the forest image and 1,126 points for the desert image. Pixels corresponding to these scores are then averaged for all bands to create the reference VN vegetation radiance vector.

3.2 Atmospheric Compensation Data Points

For both images equations (2)-(4) were computed by selecting appropriate vectors. For both images, ρ_1

was a vector of zeros corresponding to an ideal dark reflectance vector; correspondingly, L_1 was a constructed vector of the darkest radiance in the image, created by taking the minimum value in each wavelength [3]. For both images, L_2 vectors were computed through the *Vegetation Index Thresholding* approach using the mean vegetation radiance vector. For the Forest image, data for the ρ_2 vector was taken from radiance information from in-scene Sweet Gum Trees; whereas for the Desert image the ρ_2 vector was created using reflectance data from in-scene Creosote brush.

IV.   Anomaly Detection

Anomaly detection in HSI involves a statistical classification problem where pixels different from the background are detected; detectors apply either global or local background models with varying advantages to each [1][3][25]. The RX algorithm is a sliding window approach to anomaly detection where the local, within window, statistics are examined for anomalies [1][3][24]-[26]. RX is an extension of generalized likelihood ratio test and approximates to a Mahalanobis distance for large numbers of points in an image [1][24]-[26]. RX is expressed as

$$RX(x) = (x - \hat{\mu})^T \hat{C}^{-1}(x - \hat{\mu}), \qquad (5)$$

with

$$\hat{C} = \left(\left(\frac{N}{N+1}\right)\hat{\Sigma} + \left(\frac{1}{N+1}\right)(x - \hat{\mu})(x - \hat{\mu})^T\right), \qquad (6)$$

where x is a given pixel vector at the center of the window; $\hat{\mu}$, the estimated mean vector of pixels in the window; N, the number of pixels in the window; and $\hat{\Sigma}$, the estimate of the window's covariance matrix [3]. As N becomes large, $\hat{C}$ converges to $\hat{\Sigma}$, and equation (5) converges to the Mahalanobis distance between the pixel vector and the mean vector of the window [26].

For anomaly detection, pixels with an RX score greater than the chi-squared distribution, $\chi^2_{\alpha,(N-1)}$, are viewed as anomalous. RX was selected for the "ROI [Region of Interest] Generator" in Step 1b of Fig. 1, due to its ubiquity, simple mathematical expression, and known expectations of performance.

Performance of ATD algorithms are typically displayed through Receiver Operating Characteristic (ROC) curves [3]; ROC curves plot false positives on the x-axis and true positives on the y-axis [27]; thresholds for the ROC curve are from varying the significance level of the chi-squared distribution of the RX scores. Presented in Fig. 3 are ROC curves for the RX algorithm when applied to the two images in Fig. 2.

For the Forest and Desert images, selecting appropriate criteria for operating points from the RX detected anomalies is critical. In operation, knowledge of true positives and false positives would not be known a priori. For the example herein, an operating point with 20% false positives was selected (equating to a Chi-squared distribution setting of 0.378 for both images); although this enables many false positives to be included, the combined ATD background reduction to cue to ATR mitigates this. This operating point resulted in 11.84% of the Forest and 11.19% of the
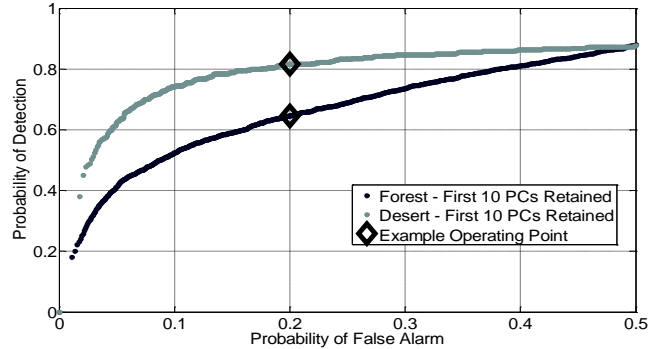

Fig. 3 Anomaly Detection ROC Curves for Forest and Desert Images using known man-made objects in-scene. The true and false positive fractions equating to a Chi squared distribution setting of 0.378 is indicated by the diamond.

Desert image identified as anomalous, with false negatives of 14.35% for the target class in the forest image and 21.55% in the desert image (which must be considered when discussing final performance).

4.1 Background Filtering/ROI Chipping

Chipping frequently refers to selecting a region around a detected anomaly for user viewing and analysis [1]. In this research, "Chipping" relates to grouping detected anomalies from the ATR process through various means, listed in Table 1, for background filtering as an additional automatic computational classification in Step 1b in Fig. 1. Through automatic target detection preprocessing/background filtering, only pixels statistically different from the background are further examined for signature matching purposes, thereby decreasing the likelihood that background pixels are grouped as possible matches for a given signature. Subsequently, detected anomalies are grouped by the five different methods presented in Table 1, offering different types of ROI clusters for evaluation. The various chipping methods listed in Table 1 permit examination of how clusters of detected anomalies relate to identifying anomalies; cluster groups are defined herein as groupings of pixels contiguously

Table 1 ROI chipping methods for background filtering

Chipping Method*	ROI Description
Averaged	Matched filter scores for pixels are averaged per group
Baseline/None	No chipping is applied, the entire image is processed for spectral similarity to a known sample
Centroid	Pixel vector of the centroid from multi-pixel anomalies used
Individual	Each ATD filtered anomalous pixel is individually examined
Majority	All pixels in a group analyzed, with score fusion used to label anomaly group identity
Mean	Averaged radiance signature from multi-pixel groups
Median	Median spectral band created from median value of each wavelength for a multi-pixel anomalies
Mode	Mode radiance band taken from a multi-pixel anomalies

*All methods are equivalent for a one pixel detected anomaly

Table 2 Example ROI chips by image for chipping methods

Chipping Method	Forest Image		Desert Image	
	Number of Chips	Avg Number of Pixels/Chip	Number of Chips	Avg Number of Pixels/Chip
Baseline/None	122,811	1	112,681	1
Individual	14,537	1	12,620	1
All Other ChippingMethods	3,547	4.098	2,707	4.662

connected to each other, diagonally connected points are also considered contiguous.

Conceptually, the chipping methods take groupings of points identified as anomalous, create matrices of each group's pixel vectors; in essence this is akin to creating vegetation spectrum in VN, only now we are creating multiple anomalous grouping spectrums. Descriptively, the *centroid* chipping method selects only the spectral from the geometric centroid of an anomaly group, a caveat being this may or may not be a pixel within the group itself; the *individual* method processes each ATD anomalous pixel and computes a matching score individually for each pixel; the *averaged* method extended the *individual* method by computing a mean matched filter score for grouped pixels, thereby incorporating some contextual knowledge of adjacent points; the *majority* method computes the spectral similarity of each pixel in a group then applying a majority voting score for the declaration of the group as the class of interest or as background; the *mean* chipping method averages the spectral signature of all pixels in a group; the *median* chipping method computes a median pupil-plane radiance of all pixels in a group; the *mode* method computes a mode pupil-plane radiance of all pixels in a group. Of note, all chipping methods are equivalent for single pixel detections. For comparison to operating with no background filtering, the *baseline* method processes the entire image through the ATR. In other words, *baseline* follows the "Traditional Step 1b" in Fig. 1.

Table 2 lists details for each HSI image and the chipping type; the absence of chipping was considered as the baseline method. The *individual* method results in much fewer pixels (7.4-7.6% of the image) to process than the baseline, however this method ignores any contextual knowledge of groupings of anomalies, in essence making each pixel its own chip; *all other chipping* methods considers the same number of chips, however the manner in which each method incorporates and handles the chips differs and the number of chips to process is considerably lower than the baseline or individual methods.

V. Automatic Target Recognition

ATR in this paper employs the commonly used Match Filter (MF) algorithm which produces similarity scores as a function of the distances between candidate target signatures and known target signatures. Higher MF scores indicate a closer match between candidate signatures and a known source. In HSI, ATR methods are frequently employed for material identification where fine spectral sampling enables detection of fine spectral features at the molecular level [3].

The baseline algorithm employed for ATR was the MF described by

$$y = \frac{(s - \hat{\mu})^T \widehat{\Sigma}^{-1} (x - \hat{\mu})}{(s - \hat{\mu})^T \widehat{\Sigma}^{-1} (s - \hat{\mu})}, \qquad (7)$$

where the MF score, y, results from examining each pixel vector, x; the mean vector of the image, $\hat{\mu}$; the known spectral source vector, s; and estimated covariance of the entire image, $\widehat{\Sigma}$ [1][28]. In processing equation 7, higher magnitude scores are viewed to be closer to the known spectral source than lower scores. For the chipping methods applied herein,

the Adaptive MF (AMF) was used as developed by Manolakis et al. [10].

The *baseline* method can also be termed 'none' as no chipping methods are applied, this method processes the entire image pixel by pixel using equation 7; many applications in literature (e.g. [1][3][29]) examined entire images in this manner, supporting the contention of this as the typical HSI ATR methodology, hence the term baseline. Whereas the baseline methodology assumes that no preprocessing has been performed, the combined methodology of automatic HSI anomaly detection cueing involves examining only the ATD detected anomalous pixels and groupings of pixels. When applying the AMF to the chipping methods, determining appropriate selections of AMF parameters was critical. For all chipping methods, the inverse

Fig. 4 Top: AMF ROC for the Forest image; Bottom: AMF ROC for the Desert image

covariance matrix and mean vector were computed using the entire image.

Fig. 4 displays ROC performance results of the chipping methods for ATR in comparison with the baseline; no initial biases were incorporated for the chipping methods due to the ATD operating point selected, Fig. 3, containing a majority of the target class and missing no whole targets (of primary importance when detecting and classifying targets). Of note, the *individual* chipping method offers similar

ATR performance for both the Forest and Desert images as examining the entire image. The *mean* and *averaged* chipping methods offered consistently better ATR performance than the baseline; the *mode* and *centroid* method offered erratic performance being among the highest performing chipping methods in the Desert image and among the lowest in the Forest image. Whereas the *median* and *majority vote* methods were seen to offer the lowest performance among the methods examined. A caveat in considering the *majority vote* method must be discussed; there are four true clusters in the Forest image and six true clusters in the Desert image. This introduces the possibility that the low performance was a factor of a low number of true class clusters being present in the images.

VI.   Conclusion

This research introduced an autonomous framework for HSI ATR through the incorporation of ATD background reduction filtering, automated atmospheric compensation, vegetation index thresholding, and variations on chipping. Results were provided using HYDICE HSI radiance images and collected ground reflectance data. The advantages of this framework are its conceptual simplicity and reduction in number of pixels processed by signature matching schemes when compared to computationally intensive methods.

The chipping methods in one sense could be viewed as increasing pixel mixing, on another hand decreasing the number of pixels to be processed at the expense of possibly decreasing signature matching performance. Advantages over the baseline were evident in many methods. While the performance of the individual method was similar to analyzing the whole image significant reduction in the number of points to analyze increases the list of possible methods when computation times are critical.

While, in operation, one would not have the luxury of knowing true performance of the ATD method, methods are available for robustly finding HSI ATD operating points prior to fielding [30]. One noticeable benefit of the chipping methods is their analysis of much fewer data points for ATR processing; this could be exploited as a benefit through algorithmic fusion of multiple ATR methods and/or more computationally intensive methods thereby leveraging the benefits of a computationally simple ATD background reduction to remove probable background points.
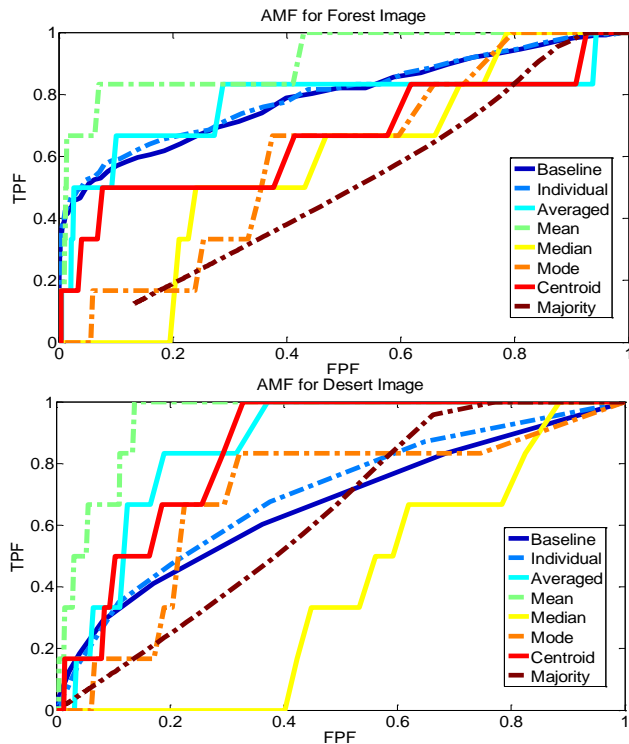
ACKNOWLEDGEMENT

REFERENCES

[1] M. T. Eismann, A. D. Stocker, and N. M. Nasrabadi, "Automated hyperspectral cueing for civilian search and rescue," Proc. IEEE, vol 97, no. 6, June 2009, pp. 1031-1055

[2] J. N. Rinker, "Hyperspectral imagery – what is it? – what can it do?" USACE 7th Remote Sensing Symposium, Portland, OR, May 1990

[3] M. T. Eismann, Hyperspectral Remote Sensing, SPIE Press, Apr. 2012

[4] P. Lehner, L. Adelman, R. DiStasio, M. Erie, J. Mittel, and S. Olson. (2008 Jan. 8) "Confirmation bias in the Analysis of Remote Sensing Data." Mitre Corporation, Rep. 06-1288. Accessed: Jan. 25, 2012, Available: http://www.mitre.org/work/tech_papers/tech_papers_07/06_1288/

[5] M. C. Mettel, "A GIS model for the determination of wetland mitigation sites," in Application of Geographic Information Systems in Hydrology and Water Resources Management, K. Kovar and H. P. Nachtnebel, Eds., IAHS Press, 1996, pp. 331-340

[6] S. Magnuson, "Military's swimming in sensor and drowning in data," National Defense Magazine, January, 2010

[7] K. D. Friesen, Automatic Target Recognition for Hyperspectral Imagery, MS Thesis, Air Force Institute of Technology, 2012

[8] J. A. Ratches, "Review of current aided/automatic target acquisition technology for military acquisition tasks," Optical Engineering, vol. 50, no. 7, July 2011

[9] C. I. Chang, and S. S. Chiang, "Anomaly detection and classification for hyperspectral imagery," IEEE Trans. Geoscience and Remote Sensing, vol. 40, no. 6, June 2002, pp. 1314-1325

[10] D. Manolakis, R. Lockwood, T. Cooley, and J. Jacobson, "Is there a best hyperspectral detection algorithm?" Proc. SPIE Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV, vol. 7334, 2009, pp. 1-16

[11] L. J. Rickard, R. Basedow, E. Zalewski, E., and P. Silverglate, P., "HYDICE: An airborne system for hyperspectral imaging," Proceedings of the SPIE , vol. 1937, 1993, pp. 173-179

[12] T. W. Ray. (1994 Oct. 13) "A FAQ on vegetation in remote sensing," Accessed: Jan. 10, 2010, Available: http://www.yale.edu/ceo /Documentation/rsvegfaq.html

[13] P. K. Connor, and D. W. Mooneyhan, "Practical applications of Landsat data," in Monitoring Earth's Ocean, Land, and Atmosphere from Space – Sensors, Systems, and Applications, A. Schnapf, Ed., AIAA, 1985, pp. 371-396

[14] J. W. Rouse, R. H. Haas, J. A. Schell, and D. W. Deering, "Monitoring vegetation systems in the great plains with ERTS," 3rd ERTS Symposium, NASA SP-351, vol. 1, 1973, pp. 309-317

[15] R. J. Johnson, Improved Feature Extraction, Feature Selection, and Identification Techniques that Create a Fast Unsupervised Hyperspectral Target Detection Algorithm, MS Thesis, Air Force Institute of Technology, Mar. 2008

[16] M. D. Farrell, and R. M. Mersereau, "On the impact of PCA dimensionality reduction for hyperspectral detection of difficult targets," IEEE Geoscience and Remote Sensing Letters, vol. 2, no. 2, Apr. 2005, pp. 192-195

[17] C. Rodarmel, and J. Shan, "Principal component analysis for hyperspectral image classification," Surveying and Land Information Systems, vol 62, no. 2, 2002, pp. 115-123

[18] D. A. Landgrebe, Signal Theory Methods in Multispectral Remote Sensing, Wiley-Interscience, New Jersey, 2003

[19] W. R. Dillon, and M. Goldstein, Multivariate Analysis: Methods and Applications, John Wiley & Sons, New York, 1984

[20] I. E. Bell, and G. V. G. Baranoski, "Reducing the dimensionality of plant spectral databases," IEEE Trans. Geoscience and Remote Sensing, vol. 42, no. 3, Mar. 2004, pp. 570-576

[21] F. Tsai, E. K. Lin, and K. Yoshino, "Spectrally segmented principal component analysis of hyperspectral imagery for mapping invasive plant species," Int'l J. Remote Sensing, vol. 28, no. 5, Mar. 10, 2007, pp. 1023-1039

[22] D. H. K. Fairbanks, and K. C. McGwire, "Patterns of floristic richness in vegetation communities of California: regional scale analysis with multi-temporal NDVI," Global Ecol. Biogeogr., vol. 13, 2004, pp. 221-235

[23] C. C. Borel, (2010) "Vegetative canopy parameter retrieval using 8-band data," DigitalGlobe 8-Band Research Challenge, Available: http://www.exelisvis.com/portals/0/pdfs/envi/8_bands_Christoph_Borel.pdf

[24] I. S. Reed, and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," IEEE Trans. Acoust., Speech, Signal Process., vol 38, 1990, pp. 1760-1770

[25] D. W. J. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," IEEE Sig. Proc. Mag., Jan. 2002, pp. 58-69

[26] T. E. Smetek, Hyperspectral Imagery Target Detection Using Improved Anomaly Detection and Signature Matching, PhD Dissertation, Air Force Institute of Technology, June 2007

[27] T. Fawcett, (2003) "ROC graphs: notes and practical considerations for data mining researchers," Technical Report HPL-2003–4, HP Laboratories, Palo Alto, CA.

[28] D. Manolakis, and G. Shaw, "Detection algorithms for hyperspectral imaging applications," IEEE Sig. Proc. Mag., Jan. 2002, pp. 29-43.

[29] L. Wood, Survey of ERIM approaches applicable to semi-automatic target detection and cueing for multispectral and multisensory exploitation, Technical Report: 215400-1-F, July 25, 1989.

[30] F. M. Mindrup, M. A. Friend, K. W. Bauer, "Selecting training and test images for optimized anomaly detection algorithms in hyperspectral imagery through robust parameter design," Proceedings of SPIE, vol. 8048, pp. 80480C-80480C-15, 2011.

HORNS: A Semi-perfectly Secret Homomorphic Encryption System

Mahadevan Gomathisankaran[a], Kamesh Namuduri[b], Akhilesh Tyagi[c]

[a] Dept. of Computer Science and Engineering, University of North Texas, Denton, Texas, USA

[b] Dept. of Electrical Engineering, University of North Texas, Denton, Texas, USA

[c] Dept. of Electrical and Computer Engineering, Iowa State University, Ames, Iowa, USA

Abstract – With the increase in computation power and communication bandwidth it has become cheaper to aggregate the computation capabilities. This new model of computation is known as cloud computing. The security of cloud computing model will determine its universal applicability. In this paper we propose a novel homomorphic encryption system which can become a building block of securing the computations performed by cloud. We introduce the notion of semi-perfect secrecy and prove that our system is secure under this setup.

Key Words – Cloud Computing, Encryption function, Homomorhpic Encryption, Residue Number System.

I. Introduction

Cloud computing refers to the provision of computational resources on demand via a computer network. Cloud computing fundamentally allows for a functional separation between the resources used and the user's computing environment. This has enabled new business models as well as computing applications. DARPA, in its recent budget, has announced that it wants to create a cloud infrastructure which can be used to provide any-time and any-where access to the military users. Success of these new cloud applications depends on the security it can guarantee.

Homomorphic encryption systems allow computations to be performed on the ciphertext directly without having to decrypt them to plaintext. Such systems can form the building block of providing security to cloud computing applications. Recently, Gentry [2] has proposed a homomorphic encryption scheme of the form $c = pq + m$, where c is the ciphertext, m is the plaintext message, p is the secret key, and q is a random number. This encryption function is homomorphic with respect to addition, subtraction and multiplication. The relationship between c and m is that m is the residue of c with respect to modulus p. In other words, the encryption function is the inverse of residue operation. While this approach provides data security it does not exploit the parallelism inherently present in the cloud.

Residue Number System (RNS) is a well-known and a well-studied number theory system [6]. RNS has been used to achieve performance improvement as the arithmetic involves smaller numbers and can be done in parallel. RNS is defined in terms of a set of relatively prime moduli.

Let P denote the moduli set, then, $P = \{p_1, p_2, \ldots, p_n\}$ and $\mathrm{GCD}(p_i, p_j) = 1$ for $i \neq j$. The dynamic range M_P of this RNS is given by

$$M_P = \prod_{i=1}^{n} p_i.$$

Any integer in the residue class Z_{M_P} can be represented in the RNS with the n-tuple, $X \rightarrow (x_{p_1}, x_{p_2}, \ldots, x_{p_n})$, where, $x_{p_i} = x \bmod p_i$. The RNS representation is homomorphic with respect to addition, subtraction, and multiplication. In other words, $X \otimes Y \leftrightarrow (x_{p_1} \otimes y_{p_1}, \ldots, x_{p_i} \otimes y_{p_i}, \ldots, x_{p_n} \otimes y_{p_n})$.

The primary application of homomorphic encryption is in the field of *Cloud Computing*. In this set-up, the *cloud*, which is untrusted, is given the task of *computing* on a client's confidential data. The client can protect its confidential data from the untrusted *cloud* if it can encrypt its data using a homomorphic encryption function and use the *cloud* to do the computing on the encrypted data.

RNS creates multiple shares of a data and the operations on these shares are homomorphic. These two properties of RNS can be used to design a homomorphic encryption function for *cloud computing*. The application of RNS, so far, has been in the fields of computer arithmetic and digital signal processing.

In this paper we design a novel encryption system, HORNS, that exploits the inherent parallelism present in the cloud. HORNS uses residue number system to create multiple ciphertext shares. These shares themselves are homomorphic hence computations on them can be done in parallel. For example, the client can partition the shares and provide different partitions

Fig. 1 Example HORNS scheme with multiple competing clouds doing the computations

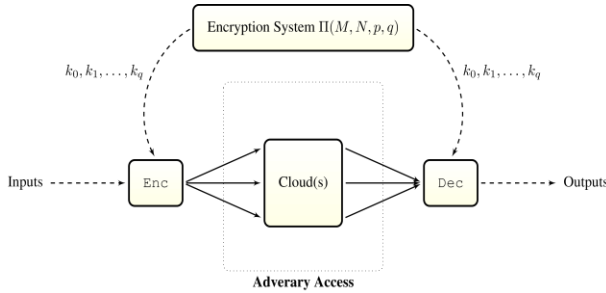Fig. 2 Example HORNS scheme with client computing some partitions

Fig. 3 Adversarial access in the HORNS encryption system

to different clouds, as shown in Figure 1. In such a scenario, all the clouds have to collude to do the attack. Another approach, as shown in Figure 2, is that the client can itself perform the computations with respect some partitions thus reducing the success probability of an attack dramatically.

The advantage of HORNS over other approaches is that it is semi-perfectly secret. In a perfectly secret encryption system the adversary does not learn any partial information about the plaintext from the ciphertext. Similarly, in a semi-perfectly secret encryption system the adversary does not gain any partial information about the plaintext from observing a subset of the ciphertext shares. We will formally define this notion and prove the security in the following sections.

The idea of splitting the information into multiple shares has been studied in the design of private circuits.

II. HORNS Encryption System $\prod(M, N, p, q)$

Let $\mathbb{N}$ be the set of natural numbers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. M, N, p, q are the parameters of the

encryption system with $M, N, p, q \in \mathbb{N}$ and $M < N$, and $p < q$.

Message Space:
$$\mu = \{m | m \in \mathbb{N}_0, m < M\}$$

Key Space:
$$\kappa = \{ \ (k_0, \dots, k_i, \dots, k_j, \dots, k_q) \qquad |$$
$$\forall 1 \leq i < j \leq q \ 1 < k_i < k_j < N,$$
$$\forall 0 \leq i < j \leq q \ \mathrm{GCD}(k_i, k_j) = 1,$$
$$\prod_{z=1}^{q} k_z > N, \prod_{z=1}^{q-p} k_z > k_0 \geq M \ \}$$

Ciphertext Space:
$$\psi = \{(c_1, \dots, c_i, \dots, c_q) | \ \forall 1 \leq i \leq q \ 0 \leq c_i < k_i\}$$

Note that both key and ciphertext are q-tuples. We refer to each element in this q-tuple as a *share*. For example, the ith share of ciphertext $c = (c_1, \dots, c_i, \dots, c_q)$ is c_i.

Generator Function (GEN):
The key generator function GEN chooses a key at random from the keyspace κ.

Encryption Function (ENC):
For any given message $m \in \mathrm{M}$ and key $k = (k_0, \dots, k_q) \in \mathrm{K}$ the encryption function proceeds as follows. **ENC** chooses a random number r_m such that $m + r_m k_0 < N$. Let $x_m = m + r_m k_0$ be the intermediate value. Then, $\forall 1 \leq i \leq q$ **ENC** computes i^{th} share of the ciphertext as $c_i = x_m \mathrm{mod} \ k_i$.

Decryption Function (DEC):
The decryption function solves the system of simultaneous congruence equations $x_m \equiv c_i \mod k_i \quad \forall \ 1 \leq i \leq q$ to find x_m. Then the message m is given by $m = x_m \mod k_0$.

III.   Semi-Perfect Secrecy

A perfectly secret encryption system is defined as one in which no adversary can learn any partial information about the message from any ciphertext. This condition can be expressed as
$P[\mu = m \mid \psi = c] \ = \ P[\mu = m]$.

Similarly, we define an encryption system as $(p, q) -$ secret if no adversary can learn any partial information about the message from any p shares out of q of the ciphertext. For example, a perfectly secret system will be $(q, q) -$ secret as per our definition. When $p < q$, such a system will be *semi*-perfectly secret.

Theorem 1: *The HORNS encryption system is $(p, q) -secret$.*

Proof: The adversarial access to the HORNS system is shown in Figure 3. The intuition behind the proof is as follows. We allow the adversary to observe any p

shares out the q ciphertext shares. In order to show that the adversary can not infer any partial information about the plaintext from these p shares we need to show that there is no correlation between them. In otherwords every possible value of the p shares should be representing any given plaintext and the vice-versa. In order to prove this theorem, we will construct the following experiment. This approach of proving the security is similar to the one proposed by Ishai et al. in [4].

A key $k \in \kappa$ is chosen at random. A message $m \in \mu$ is chosen at random. p out of q shares of the ciphertext are chosen at random. Then we show that the given p shares of cipher text can be generated from the chosen message m with the given key k. In other words, any message and key pair can generate any given p shares of the ciphertext.

Let S be the set of all p indices of the shares which were set randomly and $\bar{S}$ be the set of indices where not present in S. To prove $(p,q)-$secrecy we have to show that there exists a valid intermediate value $x_m = m + k_0 r_m$ such that $x_m \equiv c_i \bmod k_i \ \forall i \in S$, and $x_m < \prod_{z=1}^{q} k_z$.

The intermediate value x_m can be obtained by solving the following system of simultaneous congruence relations.

$$x_m \equiv c_i \bmod k_i \ \forall i \in S$$
$$x_m \equiv m \bmod k_0$$

Since the moduli in this system are co-prime with respect to each other, there exists a unique solution x_m (Chinese Remainder Theorem). The solution to this system x_m is bounded by the product of the moduli $k_0 \cdot \prod_{i \in S} k_i$.

Now we have to show that $k_0 \cdot \prod_{i \in S} k_i < \prod_{z=1}^{q} k_z$. From the properties of κ, k_1 to k_q are increasingly ordered and $\prod_{z=1}^{q-p} k_z > k_0$, we observe that

$$\prod_{i \in S} k_i \leq \prod_{z=q-p+q}^{q} k_z$$

$$\Rightarrow k_0 \cdot \prod_{i \in S} k_i \leq k_0 \cdot \prod_{z=q-p+1}^{q} k_z$$

$$\Rightarrow k_0 \cdot \prod_{i \in S} k_i < \prod_{z=1}^{q-p} k_z \cdot \prod_{z=q-p+1}^{q} k_z$$

$$\Rightarrow k_0 \cdot \prod_{i \in S} k_i < \prod_{z=1}^{q} k_z$$

IV.   Homomorphism

The encryption system $\Pi(M,N,p,q)$ is homomorphic with respect to addition, subtraction and multiplication, provided messages and their intermediate values lie within the domain. Since the ciphertext c_i is residue of the intermediate value x_m with respect to the modulus k_i, the ciphertexts derive their homomorphic properties from the residue number system.

The system $\Pi(M,N,p,q)$ uses *random* r_m to convert a message to its intermediate value. This could cause issues with overflow as $x_{m_1} \otimes x_{m_2}$ may fall outside the valid range of 0 to $N-1$. Redundancy can be used for overflow detection. Let m_1 and m_2 be the two messages such that $m_1 \otimes m_2$ to be performed in their encrypted form. Let message m_1 be encrypted to two different ciphertexts c_1 and c_2' by choosing two different intermediate values x_{m_1} and x_{m_2}'. Let message m_2 be encrypted to ciphertext c_2. Then, both $c_1 \otimes c_2$ and $c_1' \otimes c_2$ should yield the same result $m_1 \otimes m_2$ when decrypted. But if there is an overflow then the decrypted values of these ciphertexts will not match.

Binary Message Space:

When the message space is binary i.e., $\mu = \{0,1\}$, the encryption system exhibits the following homomorphisms.

$$a \text{ AND } b \quad \leftrightarrow \quad a_i \times b_i$$
$$a \text{ XOR } b \quad \leftrightarrow \quad a_i + b_i$$

Thus every binary operation in the message space can be achieved by operating on the individual share in the ciphertext space. Any arbitrary Boolean function represented in its canonical product of sum form can be implemented in HORNS.

V.   Research Issues

In order to apply HORNS for encryption, the issues of confidentiality, integrity, and cloud collusion need to be addressed. These research issues in HORNS will be illustrated by the following example.

Let a and b represent two numbers that need to be added to produce the result, $c = a + b$, by a *cloud*. Let $P = \{p_1, p_2, \dots, p_n\}$ be the moduli set that defines the RNS and M_P be its range such that $-\frac{M_P}{2} \leq a, b < \frac{M_P}{2}$. In the residue class Z_{M_P}, the range $[0, \frac{M_P}{2})$ represents positive numbers and the range $[\frac{M_P}{2}, M_P)$ represents the negative numbers. A negative number, say $-x$, is encoded as $M_P - x$, a representation analogous to 2's complement. The client generates n shares of $a \rightarrow (a_{p_1}, a_{p_2}, \dots, a_{p_n})$ and $b \rightarrow$

$(b_{p_1}, b_{p_2}, \ldots, b_{p_n})$. The client requests the *cloud* to perform modular additions over p_i on the individual shares $c_{p_i} = a_{p_i} + b_{p_i}$ independently. The client reconstructs c from the n shares, c_{p_i}, it receives from the *cloud*.

Overflow and Sign Detection:

If $a, b \geq \frac{M_P}{4}$, then the result $c = a + b > \frac{M_P}{2}$ which will imply the result is a negative number while it is not.

Confidentiality and Cloud Collusion:

The *cloud* is given access to the data shares a_{p_i}, b_{p_i} and the modulus p_i. The client can employ several mechanisms so that the *cloud* does not get access to the moduli set P. For example, the client can partition the shares and provide different partitions to different *clouds*, as shown in Figure 1. In such a scenario, all the *clouds* have to collude to reconstruct the moduli set P. Another approach, as shown in Figure 2, is that the client can itself perform the computations with respect some partitions. In any case, a *cloud* should not be given access to all the moduli in P. Even with such a restriction, the confidentiality cannot be fully guaranteed.

The *cloud* can predict M_P as it most likely will be in the neighborhood of system word size. Then it can predict a as $a = x \cdot p_i + a_{p_i}$ such that $a < M_P$. The probability that the *cloud* can find a correctly is p_i/M_P. With the cloud given access to k moduli, this probability can be increased significantly to $\prod_{i=1}^{k} p_i / M_P$. Thus, to prevent the *cloud* from inferring the data the moduli need to be protected from it. But, RNS requires moduli to be given access to the cloud for computations. The solution is to transform the moduli and let the *cloud* operate on this transformed domain.

Integrity:

The *cloud* can provide a random result c without performing the actual addition $a + b$. In other words, the client should be able to detect if the result does not correspond to the actual computation requested.

VI.   Possible Solutions

The following are the solutions we propose for the research issues presented in the previous section.

Overflow and Sign Detection:

Redundant RNS can be used for overflow and sign detection. The idea is to do the computation on multiple RN systems and compare the results. For example, let RNS_p and RNS_q be the two different RN systems used for computation. RNS_p is defined by the set of moduli $P = \{p_1, p_2, \ldots, p_n\}$ and RNS_q is defined

by $Q = \{q_1, q_2, \ldots, q_n\}$. Let $M_p = \prod_{i=1}^{n} p_i$ and $M_q = \prod_{i=1}^{n} q_i$. Let Y_p and Y_q be converted results of RNS computation in RNS_p and RNS_q respectively. Then results are valid if and only if $Y_p = Y_q$ or $M_p - Y_p = M_q - Y_q$. This solution can detect if an overflow has occurred but not correct it. While earlier research has focused on improving the performance [3] by sharing the moduli between the redundant systems, we propose to increase the redundancy by not sharing the moduli. This can increase the security as will be shown later.

Modulus Confidentiality:

In order to do computations in RNS the modulus has to be provided to the *cloud*. But, this can in-turn reduce the security of the system as the cloud can infer the range M_p if it can acquire all the moduli of the RNS by some means. In order to prevent such a possibility we want to design the HORNS in such a way that the cloud should be able to operate on the data without having to know the actual modulus. This is similar to the data confidentiality requirement hence can be achieved in a similar way by adding *confusion* to the modulus. One way to add confusion is to transform the modulus by multiplying it with a random noise, $r_{p_i} \in R_P$, to the modulus $p_i \in P$, in such way that the computations are performed using the modulus $p_i' = p_i \cdot r_{p_i}$. The result from the cloud, which operates using modulus, p_i' can be converted back to modulus p_i as shown in Lemma 1.

Lemma 1: $(x \bmod p_i') \bmod p_i = x \bmod p_i$

Proof: Let $c = x \bmod p_i$ then, for some integer $x = k \cdot p_i + c$. Then for some integer l,

$$k \cdot r_{p_i} = k \cdot r_{p_i} + l \cdot r_{p_i} - l \cdot r_{p_i}$$

$$\Rightarrow k = r_{p_i} \cdot \frac{k - l}{r_{p_i}} + l$$

$$\Rightarrow x = p_i \cdot r_{p_i} \cdot \frac{k - l}{r_{p_i}} + l \cdot p_i + c$$

$$\Rightarrow x \bmod p_i' = l \cdot p_i + c$$

$$\Rightarrow (x \bmod p_i') \bmod p_i = c$$

Montgomery Representation Variations:

In modular arithmetic, multiplications are expensive. This is so since the size of the result doubles in multiplication. If many repeated multiplications occur, as in modular exponentiation needed in RSA, the size of the result gets out of hand very quickly. The result size can be kept in check if modular reduction is performed after each multiplication. Modular reduction with its trial division is an expensive operation.

Montgomery [5] came up with an interesting scheme to maintain the multiplication products with sizes in check. For modular arithmetic modulo N, another modulus/radix $R > N$ is chosen such that R and N are coprime. R is chosen to be a power of 2 say 2^k so that multiplication and division by R on a typical machine are just a left shift or right shift by k bit positions. Somewhat similar to Lemma 1, modulus operations performed in R still maintain the modular reductions in N valid. The Montgomery representation converts each a into aR. Addition works fine as is. For multiplication of aR and bR, the size doubles. Intuitively a division by R such that the modular reduction with respect to N can be carried out later will restore the operand size. Montgomery proposed the following reduction (Montgomery reduction) for a value T − result of a multiplication: ($m \leftarrow (T \bmod R) \cdot N' \bmod R$ and $t \leftarrow \frac{T + mN}{R}$. The value t retains all the bits relevant for modular reduction with respect to N and is yet of the right size. N' is chosen such that $RR^{-1} - NN' = 1$.

Intuitively, we would like most arithmetic within the cloud to proceed in Montgomery domain with radix R (or moduli r_i). With appropriately chosen Montgomery moduli, the computation can be efficient. In the current form of Montgomery reduction, though, the reduction expression needs both N and its inverse N'. Hence if the reduction were to be performed within the untrusted cloud node, both the moduli and its multiplicative inverse need to be revealed. This goes counter to our preceding argument for modulus confidentiality.

The following are the open research questions within the context of Montgomery reduction.
1.   Current Montgomery representation maintains a value t within one subtraction of the correct value. If we allow this size to be larger, can we get away with revealing less of N and N'? There is likely to be a trade-off in computation efficiency of the modular arithmetic and its privacy. Bajard et al. [1] merge RNS and Montgomery field in an interesting manner.
2.   Another option is not to perform Montgomery reduction at the cloud node. For efficiency, up to certain number of multiplications, say k, no reduction is performed. This allows the operands to grow into k times as many bits as needed minimally. At these k − multiplication epochs, a synchronization with a trusted processor is forced wherein the Montgomery reduction and/or modular reduction is performed.

Cloud Collusion:

One way to protect the confidentiality of modulus is to distribute the computations to different clouds as shown in Figure 1. In this section, we devise a strategy for allocating a subset of the moduli to each *cloud* in such a way that it will minimize the impact on security due to collusion. Let $P = \{p_1, p_2, \dots, p_n\}$ be the moduli set and $|p_i| > S_P^n$, where S_P^n is the minimum size of modulus. Let M_P be the range of this RNS. Let k be the number of moduli given to a cloud for execution. Then, the maximum probability that any *cloud* can infer the data is S_P^k / M_P and $M_P \approx S_P^n$.

Let $f(n, k)$ represent the probability of success for any *cloud* with k out of n moduli to infer the data. Then:

$$f(n, k) = \frac{1}{S_P^{n-k}}$$

A simple way to distribute the moduli to the *clouds* is to create disjoint subsets of P and distribute it to all the clouds. In this simple approach, collusion among two clouds will exponentially increase the success probability, or in other words:

$$f(n, k_1 + k_2) = \frac{1}{S_P^{n-(k_1+k_2)}}$$

This can be achieved by the following redundant scheme in which each cloud will be given $k = q + l$ moduli, where q moduli are distinct and non-overlapping, and l moduli are redundant and overlapping with the moduli assigned to other clouds.

Thus, if two clouds collude, the number of moduli they can gather is not $2k$, instead, it is less than $2k$. Of course, if all clouds collude, they will have all the moduli required to break the RNS system.

Proposition 1: To recreate the RNS system completely, a collusion of at least $n/q + l$ clouds is needed. As long this doesn't happen, the RNS system is secure.

Proof: Consider a scenario, where y clouds collude to recreate RNS system. The collusion gives rise to $y \cdot q$ distinct moduli and $y \cdot l$ redundant moduli. In best case scenario, $y \cdot l$ moduli do not overlap with $y \cdot q$ moduli. The group of y clouds will recreate all the moduli if $n = y \cdot l + y \cdot q$. In other words, $y \geq \frac{n}{q+l}$ in order to recreate the RNS system.

Proposition 2: Assuming that $\geq \frac{n}{q+l}$, the probability of finding $(n - y \cdot q)$ distinct moduli from the $y \cdot l$ redundant moduli is given by $\frac{\binom{y \cdot q}{n - y \cdot q}}{y \cdot \binom{n-q}{l}}$.

Proof: The numerator gives the number of ways of selecting $(n - y \cdot q)$ moduli from $y \cdot q$ distinct moduli and the denominator specifies the number of

ways of selecting any l number of moduli from $(n - q)$ moduli and repeating this step for y times.

Integrity:

The integrity requirement of HORNS is to verify that the result produced by the *cloud* is indeed valid, in other words, the *cloud* has not tampered the moduli or moduli in generating the result. The solutions of overflow detection and confidentiality together provide the required integrity protection.

VII. Cloud Computing Framework

In this section, we provide a conceptual *cloud computing* framework based on HORNS. Given a program P to be dispatched to an untrusted could, there are many options for transforming it. Moreover, there are at least two possible trust models. (1) Cloud is completely untrusted. Hence any trusted operations need to be performed by the client node itself. (2) Cloud contains a kernel trusted processor. Particularly, for each group of k untrusted cores there could be one trusted core. The trusted core could have TPM [7] or Arc3D [8] like architecture. The main point to note is that this node necessarily would have significantly lower throughput than the computation cores due to its cryptographic overhead. Hence, in this model, we must carefully design the throughput load of the trusted core to be some fraction of the throughput of untrusted computing cores (say $\left(\frac{1}{10}\right)^{th}$ to $\left(\frac{1}{100}\right)^{th}$). We next describe program transformations.

RNS Coding: The protected computation of program P is threaded into k threads. All the key variables to be protected are split into k moduli using the RNS schema of HORNS. Each thread would have identical control flow, but different data. In fact, data-parallel model of NVidia Fermi GPU, CUDA, works well for such a schema. In GPU terminology, each original thread becomes a $k-$way Warp. The threads within a Warp are forced to sync either for Montgomery reduction, or for HORNS moduli reduction, or for a branch.

Threading Control: The preceding discussion highlights the fact in a secure cloud environment, some of the thread scheduling flexibility must be given to the client. A client should be able to specify scheduling parameters loosely. How this specification must be incorporated into Cloud protocols, and what degree of scheduling specifications cane be entrusted with the client is a topic for cloud computing research.

Root of trust at Cloud: It would be more efficient to have a trusted node at cloud. How would such a node be rooted in trust is still an open question.

Validation: There are many possible validation mechanisms. To name a few, the redundancy schema of Section VI could deploy multiple moduli sets for each thread. The results from all the k threads can be validated at predetermined validation points by a trusted processor. Along a similar validation schema, a $(k + 1)$ st hidden modulus could be selected. The corresponding residue can be kept as a secret with the trusted processor. When the k results from the k threads corresponding to the k moduli come back, their consistency with respect to the hidden residue can be validated with Chinese Remainder Theorem (CRT). There are many other in-between validation schema.

VIII. Cloud Computing Framework

Cloud computing refers to the provision of computational resources on demand via a computer network. With the increase in computation power and communication bandwidth it has become cheaper to aggregate the computation capabilities provision it as and when required over the network. The security of cloud computing model will determine its universal applicability. In this paper we design a novel encryption system, HORNS, that exploits the inherent parallelism present in the cloud. HORNS uses residue number system to create multiple ciphertext shares. We have defined a new notion of secrecy known as *semi-perfect secrecy* that is application to HORNS. We have also proven that, in Theorem 1, HORNS is semi-perfectly secret. We have discussed the research issues involved in applying HORNS in a Cloud Computing framework.

REFERENCES

[1] BAJARD, J.-C., DIDIER, L.-S., and KORNERUP, P. An RNS montgomery modular multiplication algorithm. IEEE TRANSACTIONS ON COMPUTERS 47, 7 (1998), 766–776.

[2] GENTRY, C. Computing arbitrary functions of encrypted data. Commun. ACM 53, 3 (2010), 97–105.

[3] GREGORY, R. T., and MATULA, D. W. Base conversion in residue number systems. Residue number system arithmetic: modern applications in digital signal processing (1986), 22–30.

[4] ISHAI, Y., SAHAI, A., and WAGNER, D. Private circuits: Securing hardware against probing attacks. In Advances in Cryptology - CRYPTO 2003, 23rd Annual International Cryptology Conference, Santa Barbara,

California, USA, August 17-21, 2003, Proceedings (2003), vol. 2729 of Lecture Notes in Computer Science, Springer, pp. 463–481.

[5] MONTGOMERY, P. L. Modular multiplication without trial division. Mathematics of Computation 44, 170 (April 1985), 519–521.

[6] SODERSTRAND, M. A., JENKINS, W. K., JULLIEN, G. A., and TAYLOR, F. J., Eds. Residue number system arithmetic: modern applications in digital signal processing. IEEE Press, Piscataway, NJ, USA, 1986.

[7] TRUSTED COMPUTING GROUP. TPM Main Specification Level 2 Version 1.2, Revision 103.

[8] GOMATHISANKARAN, M. and TYAGI, A. "Architecture support for 3D obfuscation," Computers, IEEE Transactions on , vol.55, no.5, pp.497,507, May 2006

Cluster-based Sampling and Ensemble for Bleeding Detection in Capsule Endoscopy Videos

Mohamed Abouelenien[a], Xiaohui Yuan[a], Balathasan Giritharan[a], Jianguo Liu[b], Shoujiang Tang[c]

[a]Department of Computer Science and Engineering, University of North Texas, Denton, TX, U.S.A.
[b]Department of Mathematics, University of North Texas, Denton, TX, U.S.A.
[c]University of Mississippi Medical Center, Jackson, MS, U.S.A.

Abstract – We present a cluster-based sampling and ensemble method to learn from large, imbalanced data set for bleeding detection in CE videos. Our method selects training examples randomly according to the data distributions derived from clustering. Multiple training sets are created such that data balance is restored. The sampling probability is proportional to the cluster distribution, and within each cluster the probability of a sample being selected is proportional to the distance to the center of the cluster. Classifiers are evaluated to compute performance-based weights and the prediction is made by aggregating decisions from the ensemble. Experiments were conducted using 8 annotated full-length videos. The cluster-based sampling provides training examples that preserve the innate data distribution with much less number of instances. Our experiments demonstrate that ensemble coupled with cluster-driven sampling achieves superior sensitivity and very competitive specificity. The one way ANOVA analysis reveals that our method greatly outperforms conventional SVM method.

Key Words – Capsule Endoscopy, Classification, Clustering, Video Analysis

I. Introduction

Capsule endoscopy (CE) is an imaging technology that has revolutionized our ability to visualize the entire small intestine non-invasively. The imaging component of this system is a vitamin-sized capsule that is composed of a color CMOS camera, a battery, a light source and a wireless transmitter. The camera acquires two pictures every second for approximately eight hours and generates 256×256 images transmitted to a recording device worn by the patient. It has been used to examine the entire small intestine non-invasively and is used mainly to diagnose lesions beyond the reach of conventional push endoscopy and colonoscopy. Its clinical applications have shown great improvement in diagnostic yield for bleeding sources in patients with obscure GI bleeding, and in diagnosing and localizing the source of blood loss.

More information and clinical applications of CE can be found in [1] and the reference therein.

Among many efforts in computer aided diagnosis with CE videos, bleeding detection has been investigated the most due to its clinical importance. The "Suspected Blood Indicator" function by the Given Imaging, a CE manufacturer, provides the capability of detecting blood in video frames. A study by Liangpunsakul et al. [20] showed that the overall sensitivity and accuracy of SBI were 25% and 34.8%, respectively. It exhibits better performance for active bleeding lesions in the small bowel with reported sensitivity and accuracy of 81.2% and 83.3%. This deficiency motivated studies in automatic bleeding detection. Color feature is adopted in many detection studies [13, 10] and texture features are used in applications particularly for detecting heterogeneous objects, e.g., ulcer and polyps [2, 29, 4]. The combination of color and texture has also been heavily experimented [22, 8]. On the other hand, neural networks [27, 29, 21], Support Vector Machines (SVMs) [24, 23], and thresholding [29] are used to make decisions. Despite the encouraging improvement, many previous studies were evaluated with a small number of samples and to the best of our knowledge no performance was reported with respect to entire videos. An important question awaits investigation: "Given relatively small number of positive examples from CE videos, how to train learning algorithms to achieve minimal false negative detections?"

In this article, we present a novel method to learn from large, imbalanced data set for bleeding detection in CE videos. Our method uses a cluster-based sampling strategy to select training examples and create multiple distinct training sets such that data balance is restored. Using each training set, a classifier is built and evaluated with the rest of the examples to compute a performance-based weight. The prediction to a new instance is the weighted aggregation of decisions from all classifiers. With downsampling, the size of each training set is manageable by classifier. In addition, since multiple training sets are created with randomly selected examples, the loss of information is suppressed such that the generalization performance is

greatly improved.

Our contribution to the bleeding detection in CE videos is twofold: 1) a novel sample selection method that analyzes sample distribution and intelligently selects subsets for training such that a close representation of the data distribution is reached as well as data balance is recovered; and 2) a performance-driven ensemble learning strategy that circumvents possible loss of information due to downsampling by weighting trained base classifiers with their performance measures. Our method provides a framework that integrates multiple image features and addresses the imbalance problem in the real-world CE video analysis with a statistically plausible solution. From the experimental point of view, our extensive experiments conducted with 8 full-length videos reveal the possible drawbacks of training classifier with improper manually-selected data set and demonstrated a feasible remedy using cluster-based sampling and classifier ensemble.

The rest of this article is organized as follows: Section 2 reviews the state-of-the-art methods in bleeding detection from CE videos. Section 3 describes our method that uses cluster-based sampling and ensemble (CSE) to address the difficulties arose from large imbalanced data sets. Section 4 presents the experimental results using complete CE videos and discussions. Section 5 concludes the paper.

II.   Related Work

Automatic detection of obscure bleeding in CE videos has been studied and Table 1 summarizes 12 related works. Despite different features and classification methods used, the experimental data and performances vary greatly. Among these studies, results in eight studies were generated from experiments using 1000 examples or less [18, 2, 29, 11, 22, 8, 12, 14, 15]. Four studies [17, 10, 19] used moderately larger number of examples. Comparing to the number of frames available in a CE video (approximately 50,000), however, the training data set size is much smaller. Ideally, if the training set is well-selected and unbiased, the classifier can achieve satisfactory generalization performance. It is unclear how the samples are selected and if the cohort formed represents the true data distribution.

Like many medical diagnosis applications, CE videos are full of negative examples and much less number of positive examples. That is CE videos are imbalanced in nature [29]. Classification with imbalanced data sets has been a well-known problem

in many other fields of applications. The abundant examples from the majority class and significantly inadequate number of examples from the minority class affect the classification performance when applied to examine the entire video. The challenges lie in the misrepresented data distribution.

Methods have been developed to address the challenges of imbalanced classification from both data and algorithm aspects. Data-centered methods rely on resampling to achieve equal or approximately equal number of instances from both classes [7, 16]. The Synthetic Minority Oversampling Technique proposed by Chawla et al. gained much popularity in generating instances for the minority class [7]. The arguments, however, are the increase of data size which could potentially exceed the capacity of our modem computing power; whereas the downsampling techniques are facing critiques on possible loss of information.

In algorithm-centered methods, assumptions are made in favor of the minority class. There are many real-world applications that support such assumption. For instance, in medical diagnosis and surveillance, the rare cases (samples of the minority class) carry significantly greater values than the ordinary instances. To implement this assumption in algorithms, biased decision weights are commonly employed [9]. Another thrust of efforts is to construct classifiers using training instances from only the majority class, i.e., one-class learning [26]. The rationale is that ample instances from the majority class provide a well-defined class boundary. The difficulty lies in the subjectivity of the preference and the magnitude of the bias toward the minority class.

Combination of data-centered and algorithm-centered methods has also been investigated [3]. Research has explored generating multiple data sets and aggregating cost-sensitive classification. It was claimed improved performance in both handling large data set and overall accuracy. All submissions should follow the guidelines of this journal for submission.

III.   Methodology

Our method uses cluster-based sampling and ensemble and consists of three steps: feature extraction, data rebalancing via cluster-based sampling, and ensemble classification. Fig. 1 illustrates the diagram of our method. The rationale of downsampling is that samples of the majority class are of great number and are likely to be redundant. Downsampling the majority examples balances the two classes. The possible loss

Table 1 Experimental data and detection outcomes. '-' indicates not reported in the paper.

| Reported studies	Data set size			Performance		
	Total	Abnormal	Normal	Sensitivity	Specificity	Accuracy
Kodogiannis and Boulougoura [14]	140	35	35	-	-	95.7%
Kodogiannis and Lygouras [15]	140	35	35	-	-	97.1%
Vilarino et al. [29]	400	100	300	-	-	95.5%
Coimbra and Cunha [8]	1000	-	-	-	-	87%
Lau and Correia [17]	1705	577	1128	88.3%	-	-
Li and Meng [18]	60	30	30	65.2%	82.5%	-
Li and Meng [19]	3600	1800	1800	88.8%	84.2%	-
Li and Meng [22]	400	200	200	91%	93%	-
Jung et al. [10]	2000	1000	1000	92.8%	89.5%	-
Barbosa et al. [2]	204	100	104	98.7%	96.6%	-
Karargyris and Bourbakis [11]	-	20	30	75%	73.3%	-
Karargyris and Bourbakis [12]	50	10	40	100%	67.5%	-
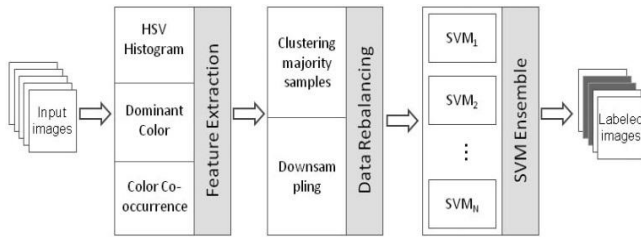
Fig. 1: Our method consists of three steps: Feature extraction, Data rebalancing, and SVM Ensemble. The arrows depict the data flow.

of information in the process of downsampling could be leveraged via bootstrap aggregating classifiers, which are trained with balanced examples from both classes.

3.1 Feature Extraction

We employ three image features in our method: color histograms, dominant color, and color co-occurrence.

Color Histogram (CH)

Color histogram is widely used due to its concise representation of color information. Among many color spaces, HSV separates the luminance from chromaticness. It is usually represented with a hexacone, the central vertical axis of which denotes the luminance. Hue is defined as an angle relative to the red and ranges in $[0, 2\pi]$. Saturation is measured as a radial distance from the central axis of the hexacone. Its chromatic components describe color in a way that is most suitable to bleeding detection [23]. Hence, video frames are converted to HSV color space and each color component is normalized to [0, 1] and sampled with 256 bins.

Dominant Color (DC)

The dominant color consists of eight representative colors, variances for each color, and their percentages in the image [25]. The descriptor is presented as a vector in the following format and the total percentages of the colors in the image sum to 1.

$$DC = \{c_i, v_i, p_i\}, \text{and } i = \{1, \ldots, 8\} \qquad (1)$$

where c_i is the i-th dominant color, p_i is its percentage, v_i is the color variance.

For each video frame, colors are clustered and the mean color is used to represent each cluster. This results in a much smaller number of colors. The variance of dominant colors is computed for bleeding and non-bleeding frames. Despite possible information overlap with CH, DC delivers a more concise color description and suppresses the color variance as well as the number of colors.

Color Co-occurrence (CC)

The color co-occurrence matrix follows the classical computation of co-occurrence matrix and contains the frequency of color pair within a pre-defined distance, i.e., (Δx, Δy). In an 8-bit color image, there are possible 2^{24} colors. To reduce the matrix size, we quantize the color into a set of representative ones. In addition, to eliminate rotation variance in the image plane, we omit the direction of the spatial location of two pixels and only keep track of the pixel distance, i.e., $d = \Delta x^2 + \Delta y^2$. Because the matrix is symmetric with respect to the major diagonal line, our feature vector only uses the components in the upper triangle matrix.

3.2 Cluster-based Sampling

The imbalance ratio in CE videos is usually significant, which can be as much as 1000:1 (refer to Table 2 for examples). Randomly downsampling the majority class to rebalance the training data could lose critical instances; whereas upsampling the minority class results in much larger data set that exceeds the capacity of modern computers. We propose a downsampling strategy based on unsupervised clustering of the data set followed by a probability-driven sampling from each cluster to preserve the geometric structure of the data set with less number of instances. Our sampling strategy is inspired by the observation that CE video frames within a temporal neighborhood are highly correlated. That is, these data points are close to each other in feature space and, hence, form a cluster. Frames that contribute to a cluster are not necessarily temporally adjacent. Retaining a number of samples from each cluster could maximize the preservation of the original data distribution with a smaller set of data points. Our hypothesis is that the instances close to the centroid of the cluster are less influential to the classifier than the ones close to the boundary of the cluster. Hence, we sample each cluster according to its innate distribution area and the sample distance to the centroid. The sampling probability is proportional to the cluster distribution, and within each cluster the probability of a sample being selected is proportional to the distance from the sample to the cluster center. Let C_j be the j-th cluster with n_j samples. The probability of a sample c, $c \in C_j$, being selected is computed as follows:

$$p(c) = \frac{n_j}{\sum_{j=1}^{J} n_j} \cdot \frac{\|c - \bar{c}_j\|}{\sum_{i=1}^{n_j} \|c_i - \bar{c}_j\|} \qquad (2)$$

where $\bar{c}_j$ is the mean of the j-th cluster and J is the number of clusters. Function $\|c - \bar{c}_j\|$ computes the distance of c to $\bar{c}_j$. The first term in Eq. (2) gives the overall probability of samples selected from C_j. The second term decides the selection probability of samples inside C_j. The denominators ensure the sum of probabilities of all samples in the majority class is unit.

Depending on the nature of data distribution, the within-cluster probability can be modified. For instance, inverse multi-variant Gaussian function provides gradual descent of probabilities from the centroid. For simplicity, we use Euclidean distance in our probability modeling.

In our method, k-means clustering is employed. For each cluster C_j, a probability, p(c), is computed and assigned to each instance c. The cumulative probability function $P(C_j)$ of the cluster integrates the probability of all instances in C_j. In our sampling process, we compare a uniform random number r with $P(C_j)$ and select the sample that defines the range. Since the overall probability of a cluster is factored by the number of its instances, the random number is generated in the range of [0, T] and $T = \frac{n_j}{\sum_j n_j}$. An instance is removed from the cluster to prevent duplication. Removing an instance changes the cumulative probability P and the random number upper bound T. Hence P is updated in the iterations. Algorithm 1 summarizes our cluster-based sampling method.

Algorithm 1: Cluster-based Sampling

1. Generate J clusters from $\mathcal{C}$ using k-means algorithm
2. **for all** $C_j \subseteq C, j \in \{1, 2, \dots, J\}$
3.   Compute n_j and $\bar{c}_j$ for cluster C_j
4.   **for all** $c_i \in C_j$
5.   Compute $p(c_i)$ using Eq. (2)
6.   **endfor**
7. $\tilde{n}_j \leftarrow n_j$\$
8. $P \leftarrow P(C_j)$ and $T \leftarrow \frac{\tilde{n}_j}{\sum_j n_j}$
9.   **for** $l \leftarrow 1$ to $\frac{N \cdot n_j}{\sum_j n_j}$
10.   Generate uniform random number $r \in [0, T]$
11. $\tilde{\mathcal{C}}(l) \leftarrow c_l$ such that $r \in P(c_l)$
12. $C_j \leftarrow C_j - \{c_l\}$ and $\tilde{n}_j \leftarrow \tilde{n}_j - 1$
13. $P \leftarrow P(C_j)$ and $T \leftarrow \frac{\tilde{n}_j}{\sum_j n_j}$
14.   **endfor**
15. **endfor**

Fig. 2 illustrates a visualization of our clustering result to the majority examples in one training video. Ten clusters are used in this example. Out of 4096 feature components, 2 are retained to accommodate 2-dimensional visualization space. Our feature selection is based on Fisher discriminant analysis and retains the two that give the best separation of the clusters. It is clear that the instances close to the origin are showing strong greenish color; whereas the far right of the x-axis depicts reddish color. The top of the space are filled with instances with wiggling folds and the lower part of the space is dominated with images with no significant texture.

3.3 Aggregation of Classifiers Trained with Multiple Features

A major concern of downsampling the majority class to rebalance the training set is the potential of
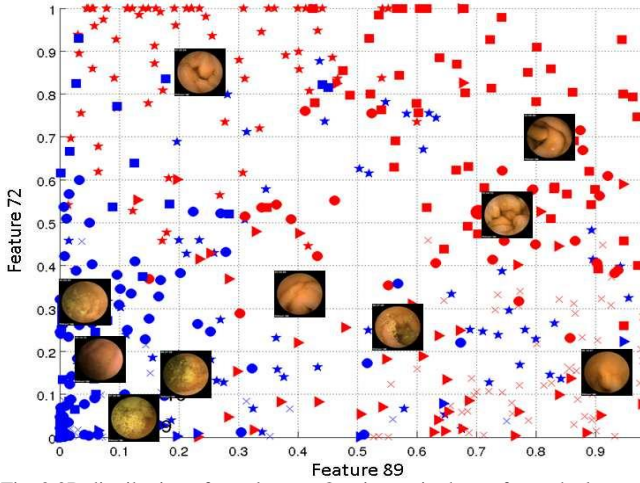
Fig. 2 2D distribution of ten clusters. One image is shown for each cluster.

missing critical instances and hence results in lower generalization performance. Without knowledge of the majority class distribution and the spatial relation of the two classes, a sampling process cannot guarantee that the downsized data set represents the information of the available data for the advantage of classification. Ensemble classifier is promising in that multiple dissimilar downsampled data sets provide balanced training set with comprehensive coverage of the majority class data distribution. The idea of bootstrap aggregating with SVM was broadly used in many problems and was originally developed to improve the estimation accuracy of weak classifiers [5,28]. The classifier ensemble labels an instance by aggregating decisions of all classifiers.

In our classifier ensemble, several SVMs are trained independently using training set created by our sampling method. Let $V_j, j = \{1, \ldots, M\}$, denotes a set of samples that are created using Algorithm 1 from a training set V. Let f_j be the discriminant function obtained through SVM learning using V_j. The label for an unseen instance, x_i, is computed by aggregating the decisions of the trained classifiers as follows:

$$L(x\_i) = \text{sign}\left(\sum_{j=1}^{M} \alpha_j \ f_j(x_i; w, b, \epsilon)\right) \qquad (3)$$

The α_j is a weight to the discriminant function f_j and is proportional to the generalization performance of f_j to the data set V:

$$\alpha_j = \frac{\mathcal{P}(f_j)}{\sum_{j=1}^{M} \mathcal{P}(f_j)} \qquad (4)$$

where $\mathcal{P}(f_j)$ denotes the metric function for evaluating the performance of f_j. The choices of $\mathcal{P}$ are many. In our implementation, we used sensitivity for $\mathcal{P}$.

An SVM is a hyperplane that maximizes the class margin. Suppose data points are represented as

$\{(x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l)\}$, $y_i \in \{1, -1\}$, and each x_i is an N-dimensional vector. The hyperplane takes the form of $w \cdot x - b = 0$. In the linearly non-separable classification problems, soft margin SVM allows, but penalizes, examples that fall on the wrong side of the decision boundary. A general form of the quadratic programming problem with soft margin and nonlinear classifier is as follows:

$$\min \frac{1}{2}\|w\|^2 + C\xi^T e$$
$$\text{subject to } y_i(w \cdot \phi(x_i) - b) \geq 1 - \xi_i \qquad (5)$$
$$\text{and } \xi_i \geq 0, \qquad 1 \leq i \leq l$$

where ξ denotes training error and C provides a weighting between regularization term and the training error. The function ϕ is a mapping from $\mathfrak{R}^n$ to a higher dimensional space. Details on SVM and its implementation can be found in [6].

IV. Experimental Results and Discussion

4.1 Data Preparation

Eight CE videos were annotated by a gastroenterologist in our team. Frames were extracted from the raw video and converted into images in JPEG format. The frame size is 256×256 or 512×512 (see Table 2) with 8-bit color depth. The bounding black region in CE frames was removed, which is outside the field-of-view of the CE camera. A 3 by 3 average filter was used to suppress random noise.

Over-exposed (usually at the start of the video before the CE device enters the digestive tract) and very dark frames (usually the ones picturing feces) were removed since these frames are easily identifiable and could present erroneous information to the classifier training. Examples of such extreme frames are shown in Fig. 3. An empirical threshold was used: if more than 65% of the pixels within the field of view are either black or white, the frame is considered as an extreme frame. The elimination of
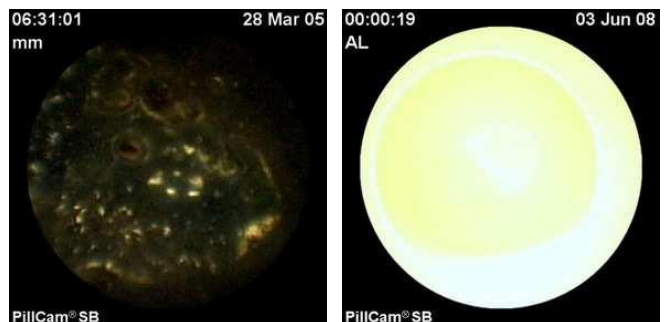

Fig. 3 Examples of over-exposed frame (a) and dark frame (b).

Table 2 Properties of the videos used in our experiments.

	Video index	Total # frames	Bleeding frames #	Ratio	Frame size
Test	1	53,526	52	1030:1	512×512
	2	55,460	691	80:1	256×256
	3	53,219	590	94:1	256×256
	4	56,721	51	1112:1	512×512
	5	54,522	588	93:1	256×256
	6	52,340	82	638:1	256×256
Train	7	51,450	465	111:1	256×256
	8	56,457	33	1711:1	256×256

extreme frames was applied to each video before image features were extracted.

Table 2 lists the number of frames and imbalance ratio of the videos. Out of the 8 videos, frames from 2 videos were used as training data for classifiers and the other 6 videos were used as testing data. The data sets from the two training videos are representative in that both videos are highly imbalanced with a large number of negative examples (i.e., frames with no bleeding signs). For performance evaluation, we adopted widely used metrics: sensitivity ($\mathcal{S}_e$) and specificity ($\mathcal{S}_p$):

$$\mathcal{S}_\mathrm{e} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \quad \text{and} \quad \mathcal{S}_\mathrm{p} = \frac{\mathrm{TN}}{\mathrm{TN+FP}} \qquad (6)$$

4.2 Effect of Manually Selected Small Training Samples

In our first experiment, we trained SVMs (based on libSVM [6]) with 800 images, in which 400 are CE frames showing obscure bleeding and the rest show normal tissues. Images of both classes were selected by our gastroenterologist. In these experiments, we used the color histogram and the raw pixel value in both RGB and HSV color spaces. Based on our previous studies [23], both polynomial and radial basis produced satisfactory results in classifying CE frames. In this study, we used radial basis function kernel with variance empirically selected as 0.0013.

Table 3 presents the results of our classifiers using manually selected balanced training data sets. The size denotes the percentile of examples used in training SVMs. The rest was used in testing. For example, for 80%, 640 images were used in training and about 160 images were used in testing. The training examples were selected randomly in each trial. The experiments were repeated for 10 trials, and the average results were reported together with standard deviations (STD).

The highlighted results illustrate the best outcomes. Between histogram and pixel value, there is no significant difference in performance. With the same color space, the two metrics give very comparable results. However, histogram-based feature provides

more concise description of the view. The HSV color representation demonstrates better results than those of the RGB color representation. The mean sensitivity and specificity as well as the standard deviations are in the close range of 90%. If we aim to maximize the performance, the HSV histogram gives the best overall detection rates.

Table 3 The performance of bleeding detection using balanced training examples. The STDs are listed in parenthesis.

Feature	Training data	$\mathcal{S}_e$	$\mathcal{S}_p$
RGB Histogram	80%	93.8% (2.4)	82.6% (4.5)
	60%	92.6% (2.3)	78.3% (3.4)
	40%	93.1% (2.5)	78.0% (3.4)
RGB raw Pixel value	80%	91.8% (1.7)	80.7% (5.2)
	60%	91.8% (2.9)	80.7% (3.3)
	40%	90.4% (3.3)	77.6% (2.8)
Mean		92.5%	79.7%
HSV Histogram	80%	96.8% (1.8)	93.8% (1.4)
	60%	96.2% (0.9)	**93.9%** (2.9)
	40%	95.1% (1.1)	89.3% (5.0)
HSV raw Pixel value	80%	**97.5%** (2.9)	86.3% (1.8)
	60%	95.6% (3.4)	86.9% (3.5)
	40%	92.6% (4.6)	87.0% (2.1)
Mean		95.6%	89.5%

In our second experiment, we applied the trained classifier to the CE videos. Table 4 reports the performance of our classification. Six videos were used to test the previously trained classifier. The mean sensitivity is at 60.6%; whereas the specificity is at 88.1%. Comparing to the previous 95.6% (sensitivity) and 89.53% (specificity), the degradation is significant. A major factor is the misrepresentation of data distribution from the training data set.

Table 4 Results HSV histogram SVM classifier, without re-balancing.

Video	$\mathcal{S}_e$	$\mathcal{S}_p$
1	61.5%	88.3%
2	60.2%	88.1%
3	59.3%	89.1%
4	61.2%	87.9%
5	62.7%	86.9%
6	58.5%	88.4%
Mean (STD)	60.6% (1.5)	88.1% (0.7)

4.3 Performance Analysis

In constructing classifier ensemble, three sets of image features were extracted: Color histogram (CH), Dominate color (DC), and Color co-occurrence (CC). The parameters used for these image features are as follow: in color histogram feature, we used HSV (Hue-Saturation-Value) space with 256 bins for each component; in dominant color, 16 most prominent colors were selected to compute the feature vectors, and the radius used in color co-occurrence is selected

Table 5 Results from average SVM using rebalanced data sets. }

Video	CH		DC		CC		CSE	
	S_e	S_p	S_e	S_p	S_e	S_p	S_e	S_p
1	69.2	93.4	75.0	91.5	76.9	90.1	82.7	92.2
2	72.2	92.1	76.6	90.3	77.4	87.9	78.3	94.3
3	70.3	92.6	75.5	94.2	73.2	90.5	82.0	93.1
4	72.5	94.9	76.5	92.3	72.5	91.5	80.6	94.1
5	68.9	93.7	70.4	93.1	76.5	90.7	81.4	92.5
6	71.1	91.2	75.6	94.2	70.7	92.3	83.1	93.6
Mean	70.7	93.0	74.9	92.6	74.5	90.5	**81.4**	**93.3**
STD	2.3	1.7	5.3	2.4	7.7	2.2	3	0.7
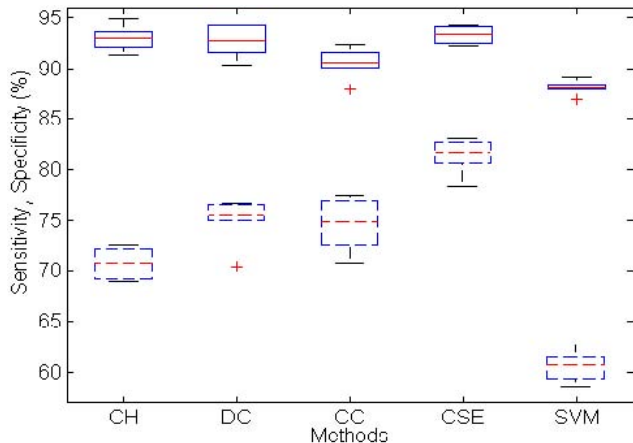
Fig. 4 Box plots of sensitivity (dash line) and specificity (solid line) of our method and the SVMs trained with manually selected data set.
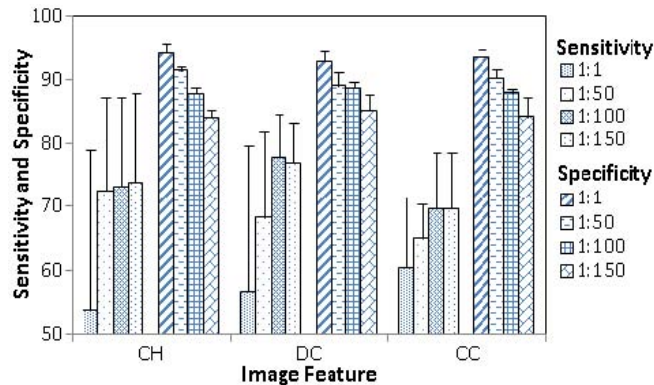
Fig. 5 Sensitivity and specificity of SVMs trained with weighted samples.

as 5. For each set of features, three SVMs were generated. The selection of training images follows our cluster-based sampling method in Algorithm 1. The number of clusters was empirically chosen as 40. In this study, we employed radial basis function as the kernel for SVMs with the variance at 0.0024.

Table 5 lists the results of method. Each row presents the test performance with one entire video. The average performance and its standard deviation are reported. The sensitivity and specificity for each image feature are average of three individually trained SVMs. In contrast to the results in our second experiment (as shown in Table 4), using the same image feature and classifier, the outcomes are better with the sensitivity in the lower 70% and specificity in the lower 90%; whereas SVM trained with manually selected balanced data resulted sensitivity in the lower 60% and specificity in the upper 80%. The testing results using the other two image features are slightly better in sensitivity. The ensemble outperformed all the classifiers with average sensitivity at 81.4% and average specificity at 93.3%.

Fig. 4 illustrates the box plots of the sensitivity and specificity between our method and the SVM trained with manually selected data set. It is interesting that even without constructing an ensemble the cluster-based sampling improves the classifier performance. The column CH and SVM was resulted from trained SVMs using the same image features (color histograms). The mean sensitivity and specificity are improved from 60.6% and 88.1% to 70.7% and 91.2%, respectively. Using one-way ANOVA analysis of the results using cluster-based sampling and manually selected balanced training data, the p-values of the sensitivity and specificity are 4.22E-7 and 1.27E-5, respectively. This indicates a fairly significant improvement originated from our sampling algorithm.

Fig. 5 shows the average performance over 6 videos of the conventional SVMs trained with weighted samples. Sensitivity and specificity are grouped for each image feature. The minority instances are assigned with greater weights to achieve a balance For instance, in the case of 1:50, all minority instances are weighted 50; whereas all majority instances are weighted 1. The average imbalance ratio in our data set is close to 1:150. With the increase of weights, sensitivity improves and specificity drops. When small weight is used, e.g., 1:1, the trained SVMs exhibit poor sensitivity and greatest variance. By applying large weight to minority instances, the sensitivity can be boosted to upper 70s with dominant color feature in sacrifice of specificity. In addition, it is evident that the improvement of sensitivity by increasing weight levels out as weight increases.

Table 6 reports the F-test results and p-values using pair-wise one-way ANOVA analysis. Each method is compared against the CSE method and SVM classifier. The results produced using CSE as reference show great improvement in sensitivity. The F-test and p-value in comparison to SVM are 479.79 and 8.8E-10, respectively. The specificity difference between CSE and the three features is not significant. However, as shown in Fig. 4, the specificity of CSE is more consistent than other methods. From the results produced using SVM as reference, it is clear that dominant color achieved most improvement in sensitivity (the F-test and p-value are 161.31 and

1.71E-7, respectively) and color histogram outperformed the other images features (the F-test and p-value are 64.15 and 1.17E-5, respectively), which is coincide with our observation from Fig. 4.

Table 6 F-test results and p-values of one way ANOVA analysis.

CSE as reference		CH	DC	CC	SVM
S_e	F-test	128.5	29.6	26.1	479.8
	p-value	5E-7	2.9E-4	4.6E-4	8.8E-10
S_p	F-test	0.3	1	15.9	129.4
	p-value	6.3E-1	3.6E-1	2.6E-3	4.8E-7
SVM as reference		CH	DC	CC	SVM
S_e	F-test	133.2	161.3	116.8	479.8
	p-value	4.2E-7	1.7E-7	7.8E-7	8.8E-10
S_p	F-test	64.2	41.4	12.3	129.4
	p-value	1.2E-5	7.5E-1	5.6E-3	4.8E-7

V. Conclusion

In this paper we describe a cluster-based sampling and ensemble method to learn from large, imbalanced data set for bleeding detection in CE videos that minimizes false negative decisions. Our method selects training examples randomly according to unsupervised clusters and creates multiple training sets such that data balance is restored. The sampling probability is proportional to the cluster size, and within each cluster the probability of a sample being selected is proportional to the distance to the center of the cluster. The prediction to a new instance is the weighted aggregation of decisions from all classifiers. With downsampling, the size of each training set is greatly reduced. In addition, since multiple training sets are created with randomly selected examples, the loss of information is greatly suppressed.

Based on our experiments, the following conclusions can be drawn. First, the cluster-based sampling provides training examples that preserves the innate data distribution with much less number of instances. Using the same number of training instances and the same image features, it is evident that the sampling algorithm contributes to the improvement of the classifier performance.

Second, the classifiers trained with different image features achieved much improved results using sampled data set. The dominant color and color co-occurrence give better sensitivity and the color histogram gives higher specificity.

Third, the ensemble integrates individually trained SVMs and achieves superior sensitivity and very competitive specificity. The one way ANOVA analysis illustrates that our method greatly outperforms conventional SVM method. The possible loss of information due to downsampling is successfully circumvented.

Last, we demonstrated the generalization degradation of using misrepresented training data set by constructing an SVM with manually selected, balanced data set of 800 images and applying the trained classifier to full-length videos. The testing performance degraded significantly. This is because the training data misrepresent the true data distribution of the CE video frames. Such misrepresentation is exaggerated when there is a large majority data set and only a small number of instances are selected for training.

REFERENCES

[1] D. G. Adler and C. J. Gostout, "Wireless capsule endoscopy," Hospital Physician, vol. 39, no. 5, pp. 14–22, 2003.

[2] D. J. C. Barbosa, J. Ramos, and L. S. Carlos, "Detection of small bowel tumors in capsule endoscopy frames using texture analysis based on the discrete wavelet transform," in the Int'l Conf. of the IEEE Engineering in Medicine and Biology Society, Aug 2008.

[3] G. E. A. P. A. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," in SIGKDD Explorations, vol. 6, 2004, pp. 20–29.

[4] N. Bourbakis, "Detecting abnormal patterns in wce im¬ages," in the 5th IEEE Symposium. on Bioinformatics and Bioenqineering, Oct 2005, pp. 223–238.

[5] L. Breiman, "Bagging predictors," Machine Learning, vol. 24, no. 2, pp. 123–140, Aug 1996.

[6] C.C. Chang and C-J. Lin, "Libsvm: a library for support vector machines, software available at http: //www.csie.ntu.edu.tw/˜cjlin/libsvm," 2010.

[7] N. V. Chawla, K. W. Bowyer, L. O. Hall, and P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," Journal of Artificial Intelligence and Research, vol. 16, pp. 321–357, 2002.

[8] M. T. Coimbra and J. Cunha, "MPEG-7 visual descriptors contributions for automated feature extraction in capsule endoscopy," IEEE Transactions on Circuits and Systems for Video Technology, vol. 16, no. 5, pp. 628–637, May 2006.

[9] Y. Huang and S. Du, "Weighted support vector machine for classification with uneven training class sizes," the Int'l Conf. on Machine Learning and Cybernetics, vol. 7, pp. 4365–4369, 2005.

[10] Y. S. Jung, Y. H. Kim, D. H. Lee, and J. H. Kim, "Active blood detection in a high resolution capsule endoscopy using color spectrum transformation," Int'l Conf. on BioMedical Engineering and Informatics, vol. 1, pp. 859–862, May 2008.

[11] A. Karargyris and N. Bourbakis, "Identification of polyps in wireless capsule endoscopy videos using log ga¬bor filters," in the IEEE Life Science Systems and Applications Workshop, Apr 2009, pp. 143–147.

[12] ——, "Identification of ulcers in wireless capsule endoscopy videos," in the IEEE Int'l Symposium on Biomedical Imaging: From Nano to Macro, Jun 2009, pp. 554–557.

[13] V. S. Kodogiannis, "Computer-aided diagnosis in clinical endoscopy using neuro-fuzzy systems," in the IEEE Int'l Conf. on Fuzzy Systems, vol. 3, July 2004, pp. 1425–1429.

[14] V. S. Kodogiannis and M. Boulougoura, "Neural network-based approach for the classification of wireless-capsule endoscopic images," in Neural Networks, vol. 4, Aug 2005, pp. 2423–2428.

[15] V. S. Kodogiannis and J. N. Lygouras, "A computerized diagnostic decision support system in wireless capsule endoscopy," in the Int'l Conf. on Intelligent Systems, Sep 2006, pp. 638–644.

[16] M. Kubat and S. Matwin, "Addressing the curse of imbalanced training sets: One sided selection," in the Int'l Conf. on Machine Learning, 1997, pp. 179–186.

[17] P. Y. Lau and P. L. Correia, "Detection of bleeding patterns in WCE video using multiple features," in the Int'l Conf. of the IEEE Engineering in Medicine and Biology Society, Aug 2007, pp. 5601–5604.

[18] B. Li and M. Q. H. Meng, "Wireless capsule endoscopy images enhancement using contrast driven forward and backward anisotropic diffusion," in the IEEE Int'l Conf. on Image Processing, vol. 2, Sep 2007, pp. 437–440.

[19] ——, "Computer aided detection of bleeding in capsule endoscopy images," in Electrical and Computer Engineering Canadian Conference, May 2008, pp. 1963–1966.

[20] S. Liangpunsakul, L. Mays, and D. K. Rex, "Performance of given suspected blood indicator," American Gastroenterology, vol. 98, no. 12, pp. 2676–2678, 2003.

[21] C. S. Lima, D. Barbosa, J. Ramos, A. Tavares, L. Monteiro, and L. Carvalho, "Classification of endoscopic capsule images by using color wavelet features, higher order statistics and radial basis functions," in the Int'l Conf. of the IEEE Engineering in Medicine and Biology Society, Vancouver, Canada, Aug 2008.

[22] B. Liu and M. Q. H. Meng, "Computer-aided detection of bleeding regions for capsule endoscopy images," IEEE Transactions on Biomedical Engineering, vol. 56, no. 4, pp. 1032–1039, Apr 2009.

[23] J. Liu and X. Yuan, "Obscure bleeding detection in endoscopy images using support vector machines," Optimization and Engineering, vol. 10, pp. 289–299, 2009.

[24] M. Mackiewicz, J. Berens, M. Fisher, and D. Bell, "Color and texture based gastrointestinal tissue discrimination," in the IEEE Int'l Conf. on Acoustics, Speech and Signal Processing, vol. 2, May 2006.

[25] B. Manjunath, J.-R. Ohm, V. Vasudevan, and A. Yamada, "Color and texture descriptors," IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 6, pp. 703–715, Jun 2001.

[26] B. Raskutti and A. Kowalczyk, "Extreme re-balancing for SVMs: a case study," SIGKDD Explorations Newslet¬ter, vol. 6, pp. 60–69, 2004.

[27] P. Spyridonos, F. Vilarino, J. Vitria, and P. Radeva, "Identification of intestinal motility events of capsule endoscopy video analysis," Lecture Notes in Computer Science, vol. 3708, pp. 302–311, 2005.

[28] V. Vapnik, The Nature of Statistical Learning Theory. New York: Springer, 1999.

[29] F. Vilarino, L. I. Kuncheva, and P. Radeva, "ROC curves and video analysis optimization in intestinal capsule endoscopy," Pattern Recognition Letters, vol. 27, pp. 875–881, 2006.

A Crowd Motion Analysis Framework Based on Analog Heat-Transfer Model

Yu Liang[a], William Melvin[b], Shane Fernandes[a], Michael Henderson[a], Subramania I. Sritharan[c], Darrell Barker[d]

[a]Dept. of Mathematics and Computer Science, Central State Univ., OH, USA
[b]Sensors and Electromagnetic Applications Laboratory, Georgia Institute of Technology, GA, USA
[c]Dept. of Water Resource, Central State Univ., OH, USA
[d]U.S. Air Force Research Laboratory, OH, USA

Abstract – Crowd motion analysis covers the detection, tracking, recognition, and behavior interpretation of a target group according to persistent surveillance video data. This project is dedicated to developing and employing a generic crowd motion analysis framework, which is based on an analog-heat-transfer model and thus denoted as CMA-AHT for simplicity, to measure and identify the anomalous pedestrians from persistent surveillance video data. Based on the hypothesis of ergodicity, the CMA-AHT framework is formulated according to the statistical analysis about the historical records of crowd's behavior, geographic information system, and crowd motion dynamics. The derivation of the CMA-AHT framework and the innovative methods involved in the framework's implementation will be discussed in detail. Using the sample video data collected by Central Florida University as a benchmark data, CMA-AHT is validated through measuring and identifying anomalous personnel or group in the video.

Key Words – crowd motion analysis, heat-transfer, partial differential equations, geographic information, statistics, video processing

I. Introduction

It is an extremely computationally intense, labor-intensive and highly unreliable job to obtain a panoramic, timely, trusted understanding about crowd behavior [1-5,7,9,14-17,19,23-25,27,29] and its future status by exploiting networked sensor assets (autonomous, heterogeneous and multi-layer sensor nodes). This project proposes a new crowd motion analysis framework, which is mainly constructed over an analog-heat-transfer model [13,18] and denoted as CMA-AHT for simplicity, to detect, track, understand and evaluate pedestrians' activity using persistent surveillance video data [6,8,10,14].

Crowd motion analysis (CMA) is critical for effective surveillance. It plays a significant role in human society. For example, in the area of homeland security, it is a challenging issue to detect and recognize those anomalous behaviors in crowd scenes (Figure 1) according to their motion features such as velocity, position and trace. In the area of civil engineering, CMA will guide us in making reasonable evacuation plans for use during an emergency. And in the area of city property management, CMA will help policy makers to optimize the exploitation of public resource.

Fig. 1 Observe and detect anomalous behavior in crowd scenes

Crowd motion modeling and simulation have been discussed in many previous works. The existing crowd motion analysis methods are classified into the following three classes of schemes: microscopic [4, 12, 25], macroscopic [14,15,17,19,20,23,24,29], and multi-scale [2].

Microscopic methods regard crowds as consisting of discrete individuals (or agents) so that the movement of each individual within the group can be identified. Microscopic methods provide a very detailed and accurate formulation about crowd's movement while suffering terribly from inhibitive computational cost. Multi-scale methods aim to provide a seamless coupling between a microscopic scheme and a macroscopic scheme.

Macroscopic methods regard pedestrian crowd as a continuum [18]. The macroscopic CMA methods are further divided into anisotropic [15, 29] and isotropic strategies [13, 17, 19]. Being applicable in both anisotropic and isotropic senior, CMA-AHT employs an analog-heat-transfer model to formulate the expected macroscopic movement of crowd in macroscopic way. Compared to alternative methods, the CMA-AHT model uses the pseudo temperature gradient, which is obtained by solving the heat-

transfer-analog partial differential equations, to represent the normal "movement trend" of pedestrians. In the implementation of CMA-AHT framework, the pseudo temperature is not explicitly defined but implicitly given by its gradient.

As a cutting-edge technology in scientific modeling and simulation, multi-scale model is dedicated to seamlessly merge the discrete individual (or particle) behavior with the continuous crowd (or continuum) behavior. It will be a major topic in our future work.

This paper emphatically discusses the following topics: (1) processing and analysis of persistent surveillance video data so as to extract the movement features of target-of-interest; (2) infrastructure of sensor-oriented applications such as crowd-motion analysis based on persistent surveillance video; (3) mathematical modeling of the expected behavior of target-of-interest. The discussions about these topics comprise the major contributions of this work.

This paper is organized as follows: Section II provides a system overview; Section III (pre-processing) discusses how to obtain pedestrian moving status such as the real-time velocity, position and movement trace using persistent surveillance video data; Section IV discusses the formulation of the pseudo temperature field according to a location's historical crowd motion data; Section V discusses how to use the pseudo-temperature field to measure and analyze the motion of an individual or crowd; Section VI introduces a macro-cell strategy, a variance of divide-and-conquer method, into the CMA-AHT model so that large-scale crowd-motion (such as a city or even a country) can be simulated efficiently. Section VII summarizes the effort.

II. Overview of the Implementation of CMA-AHT framework

As illustrated in Figure 2, CMA-AHT couples three modules: (1) crowd motion dynamics, (2) geographic information, and (3) processing and analysis of persistent surveillance video data. In other words, crowd motion dynamics, geographic information, and video data constitutes the three inputs of CMA-AHT model.

Considering the prohibitive cost to obtain accurate "situational awareness" from sensor networks [8], the proposed CMA-AHT framework for data fusion, manipulation, and prediction is supposed to own the following features:
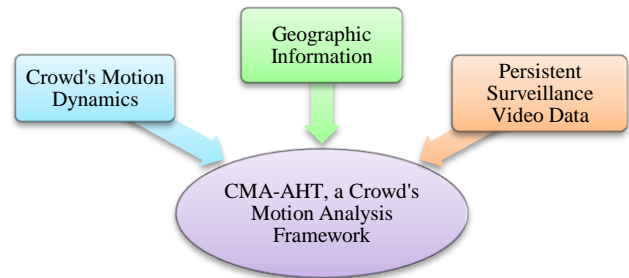

Fig. 2 Infra-structure of the CMA-AHT framework: a coupling of GIS, video data and motion dynamics

• All related data will be merged into a single shared platform. The CMA-AHT platform should provide interfaces to a variety of sensor carriers such as satellite, unmanned vehicle/flight, and on-ground cameras, etc. and be capable of handling heterogeneous sensory data such as hyper-spectral image/video and radar/LADAR signal [27, 28], etc. Putting those heterogeneous data, which might have different spatial and temporal scale and be of different format, on a single platform will help us to obtain a comprehensive situational awareness about ongoing pedestrian behavior. In addition, a single shared platform can easily coordinate cooperative sensor nodes.

• The expected or normal crowd behavior is mathematically formulated as temporal- and spatial-dependent partial differential equations (PDE) [13], which is derived from the statistical analysis on historical sensory data. Using the partial differential equations as reference, situational awareness about the observed pedestrian individual or crowd can be obtained.

• According to its current situational knowledge about crowd motion, the CMA-AHT platform will deliberately collect more detailed and specific sensory data for further analysis (we call this data scavenging or self-optimization). For example, if an anomalous pedestrian motion is detected, CMA-AHT will "zoom-in" for further detailed information corresponding to the target-of-interest, contextualized by already available relevant data. On the other hand, the non-dominant or redundant information corresponding to the target-of-interest will be ignored. The optimization of the exploitation of sensor data according to existing situational awareness is called "self-optimization".

• Besides the "self-optimization" function, according to current situational awareness, CMA-AHT should be able to collectively control the parameters of autonomous sensor nodes for a better observation about the target-of-pedestrian.

Self-optimization of the exploitation of sensory data and the collective control of sensor asset will be discussed in our future work.
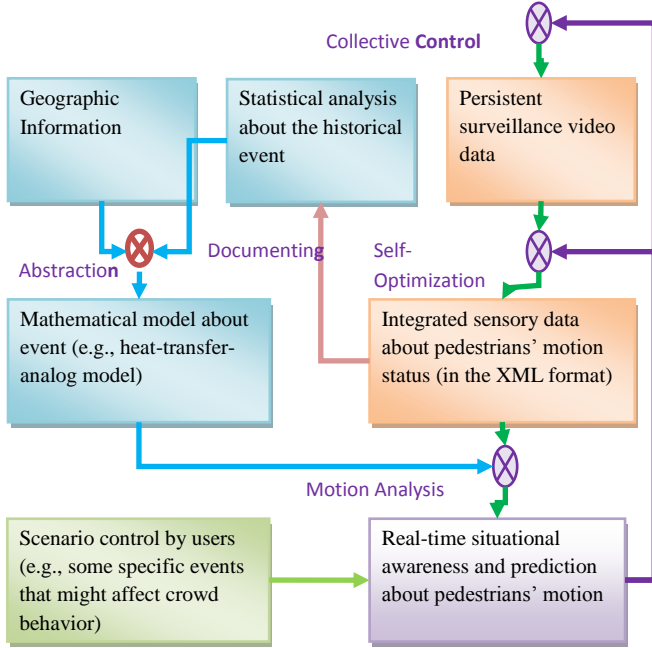

Fig. 3 Flow-chart of crowd motion analysis

Figure 3 describes the flowchart of CMA-AHT framework. It is illustrated that the implementation of CMA-AHT consists of following two threads:

• Formulating the pseudo-temperature field using historical motion data about crowd activity in the scene.

• Obtaining the motion features (velocity, location, number of pedestrians, etc.) by processing the persistent surveillance video data. The implementation of each thread will be discussed in this paper.

Based on the above two threads, a situational awareness about the behavior of observed pedestrians can be measured and evaluated. A detailed description about the implementation of the above two threads and their application in motion measurement and analysis will be covered in the following three sections.

III. Processing of Video Data

3.1 Overview of Video Processing

In this work, video data processing aims to extract a pedestrian's motion features, to include velocity, position, and motion trace, out of the persistent surveillance video data.

Processing of persistent surveillance video data includes the following operations: (1) acquisition of video data; (2) segmentation, which extracts pixels of pedestrians from background; (3) isolation of pedestrians out of noise or other moving objects; (4) translation of optical moving features (i.e., the velocity and position of moving targets within the sensor coordinate system) of detected pedestrians into their actual moving features (i.e., the velocity and position of moving targets within the geographic coordinate system); (5) documentation, which posts the output in a format suitable for post-processing and includes position, velocity, and track (optional). Step (3) and (4) will be emphatically discussed in this section.

Table 1 lists the notations involved in the processing of video data. As addressed in the "Annotation" column of the table, the specifications about the camera are predetermined. The expected human dimension $<w.l>$ is derived from the statistical analysis about human body. The optical moving features, which include the position of target in the frame ((x_{sensor}, y_{sensor})) and velocity derived from video sequence ($(v_{sensor,x}, v_{sensor,y})$) using optical-flow algorithms [6]. Based on the above parameters, the moving features about pedestrians,

Table 1: Notations involved in the processing of video data

| Variables		Definition	Annotation
Specif. about the electro-optical camera	ϕ_0	The angle between vertical and optical axis	Pre-determined
	f	Focus length	
	h	The altitude of lens from the ground	
	$(\alpha_{max}, \beta_{max})$	Maximal view angle of camera lens	
	$P_{totalView}$	Total pixel area of video frame	
Expected dimension of target	$<w>$	Expected width about target	User defined
	$<l>$	Expected height about target	
	$P_{t\arg et}$	Pixel area of expected target	Derived from the expected target dimension and lens specif.
Optical moving features of target	(x_{sensor}, y_{sensor})	Position of target in video frame	Directly derived from video sequence using optical algorithms
	$(v_{sensor,x}, v_{sensor,y})$	Optical velocity of target in video frame	
Geographic moving features of target	(x, y)	Geographic position of target (Figures 4 and 5)	To be calculated using optical moving feature and lens specif.
	(v_x, v_y)	Geographic moving velocity of target (Figures 4 and 5)	
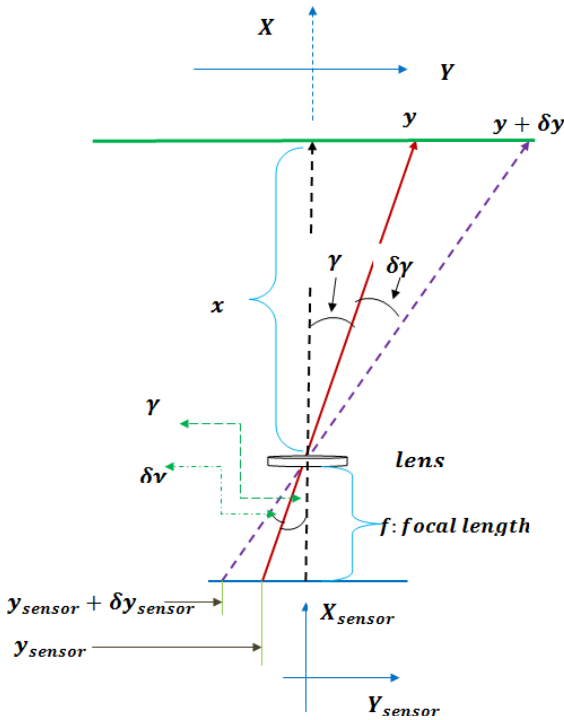
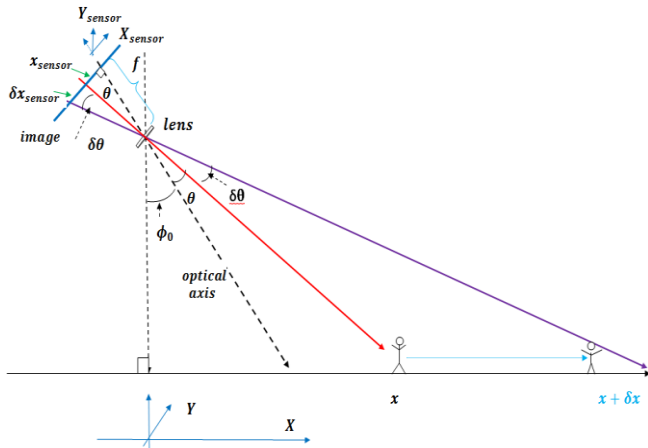Fig. 4 Side-view of the sensor coordinate system and geographic coordinate system

Fig. 5 Overview of the sensor coordinate system and geographic coordinate system

which include geographic position and geographic velocity, can be calculated.

Figures 4 and 5 provide a vivid illustration about the key notations involved in video processing.

3.2 Acquisition of Video Data

This effort supports multiple collaborative input-equipment. The resulting data will be collected and fused by CMA-AHT server.

In this work, the persistent surveillance video data comes from benchmark data collected by the University of Central Florida (www.ucf.edu/data).

3.3 Segmentation of Target-of-Interest

In this work, segmentation is accomplished using the optical-flow method [6,10] and the background-subtraction method [6,10] jointly. Optical-flow is used to segment those pixels corresponding to the moving targets or pedestrians. Background subtraction method is used to segment those pixels corresponding to the still targets or pedestrians.

3.4 Isolation of Target-of-Interest According to Their Expected Pixel Area

The pedestrian objects are detected by clustering those pixels with similar motion features. At this step, estimation about the pixel area (or the dimensional size) of the targeting pedestrian, which is denoted as $P_{t\,\mathrm{arg}\,et}$, is needed due to the following two motivations:

- Firstly, $P_{t\,\mathrm{arg}\,et}$ will help us identify the targets from those too big (e.g., moving vehicle) or too small objects (e.g., noise).
- Secondly, when several targets are clustered together closely, $P_{t\,\mathrm{arg}\,et}$ can help us obtain the number of pedestrians.

In this work, $P_{t\,\mathrm{arg}\,et}$, the expected pixel area about target is calculated via the following equations:

$$p_{t\,\mathrm{arg}\,et} = \frac{\alpha\beta}{\alpha_{\max}\beta_{\max}} P_{totalView} \qquad (1)$$

where

$$\begin{cases} x = h\left[\dfrac{f\tan(\phi_0) + x_{sensor}}{f - \tan(\phi_0)x_{sensor}}\right] \\[2mm] \alpha = \sin^{-1}\left(\dfrac{x}{\sqrt{x^2 + (h - \langle l \rangle)^2}}\right) - \sin^{-1}\left(\dfrac{x}{\sqrt{x^2 + h^2}}\right) \\[2mm] \beta = 2\tan^{-1}\left(\dfrac{\langle w \rangle}{2\sqrt{x^2 + h^2}}\right) \end{cases} \qquad (2)$$

and the parameters involved in Equation (1) and (2) are defined in Table 1.

The above formula for $P_{t\,\mathrm{arg}\,et}$ is validated using the real pixel-area about targeting pedestrian, which is illustrated in Figure 6.

Figure 7 illustrates a comparison between the theoretical pixel-area defined by Eq. (1) and real pixel area about targeting pedestrians. It is observed that theoretical pixel-area is basically consistent with the real pixel-area about targeting pedestrians. It should be
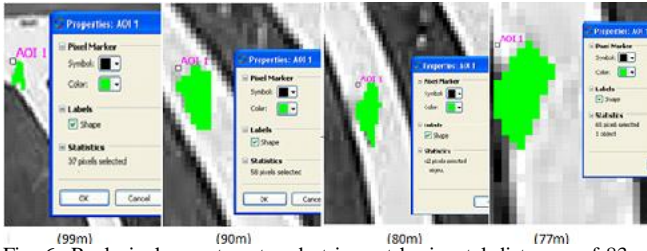
Fig. 6 Real pixel-area: target pedestrians at horizontal distances of 83m-193m from camera have variant pixel-area ranging from 34 pixels to 116 pixels (83m)

remarked that the theoretical pixel-area is derived from the following configurations: the view angle limit $(\alpha_{\max}, \beta_{\max})$, pixel-area of the whole frame $P_{totalView}$, the altitude of lens from the ground h, and the expected human dimension $(<l>,<w>)$. It is observed that:

• Eq. (1) provides a very accurate estimation about the pixel-area of target pedestrians.

• Instead of the distance between target and lens, the pixel-area of a target pedestrian is mainly dependent on the view angle about target.

• Target pedestrians at horizontal distances of 83m-193m from camera have variant pixel-area ranging from 34 pixels to 116 pixels (83m).
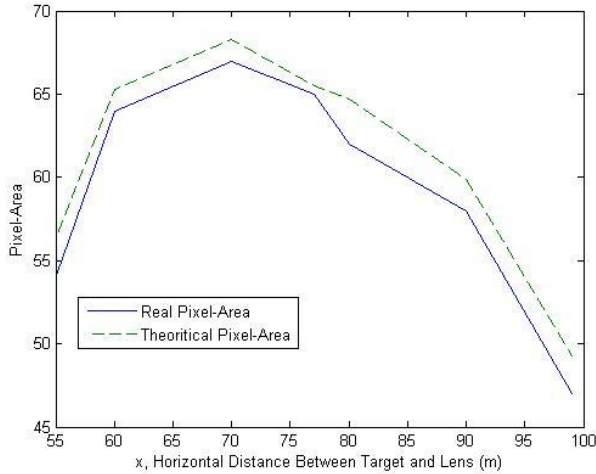
Fig. 7 Validation of the expected pixel-area about target pdestrian, with view-angle limit $\alpha_{\max} = 33.5°$ by $\beta_{\max} = 40.8°$, pixel-area of the whole frame $P_{totalView} = 1024*768 (pixel)$, expected human dimension $<l>=1.7m$ by $<w>=0..66m$, and the altitude of lens from the ground $h = 44.2m$

3.5 Translating the Optical Velocity into Geometrical Velocity

Through the detection step, the optical velocity $(v_{sensor,x}, v_{sensor,y})$ and the optical position (x_{sensor}, y_{sensor}) of each pedestrian or pedestrian block

(sometimes the pedestrians are too close to be identified) are obtained.

In the following step, it is necessary to obtain the geographic motion feature (i.e., (x,y) and (v_x, v_y)) of pedestrians according to the optical motion features (i.e., $(v_{sensor,x}, v_{sensor,y})$ and (x_{sensor}, y_{sensor})). As illustrated in Figures 4 and 5, optical motion features are formulated within the sensor coordinate system (X_{sensor}, Y_{sensor}) and geographic motion features are formulated within the geographic coordinate system (X, Y).

Geographic motion feature are computed according to the pre-determined specifications of the video camera and the observed optical motion feature. From Figures -5 and the following trigonometry formula:

$$\tan(\alpha + \beta) = \frac{\tan(\alpha) + \tan(\beta)}{1 - \tan(\alpha)\tan(\beta)} \qquad (3)$$

(x, y) is computed via the following formula:

$$\begin{cases} x = h\tan(\phi_0 + \theta) = h[\frac{f\tan(\phi_0) + x_{sensor}}{f - \tan(\phi_0)x_{sensor}}] \\ y = x\tan(\gamma) = \frac{h}{f}[\frac{f\tan(\phi_0) + x_{sensor}}{f - \tan(\phi_0)x_{sensor}}]y_{sensor} \end{cases} \qquad (4)$$

Furthermore, Figures 4 and 5 illustrates that:

$$\begin{cases} x + \delta x = h\tan(\phi_0 + \theta + \delta\theta) \\ y + \delta y = x\tan(\gamma + \delta\gamma) \end{cases} \qquad (5)$$

According to the trigonometry formula (1) and limit theory, it follows that

$$\begin{aligned} v_x &= \lim_{\delta t \to 0} \frac{\delta x(t)}{\delta t} = \lim_{\delta t \to 0} \frac{h\tan(\phi_0 + \theta + \delta\theta) - x}{\delta t} \\ &= \lim_{\delta t \to 0} \frac{h}{\delta t}\{ \frac{f\tan(\phi_0) + x_{sensor} + v_{sensor,x}\delta t}{f - \tan(\phi_0)(x_{sensor} + v_{sensor,x}\delta t)} \\ &\quad - \frac{f\tan(\phi_0) + x_{sensor}}{f - \tan(\phi_0)x_{sensor}} \} \end{aligned} \qquad (6)$$

and

$$\begin{aligned} v_y &= \frac{x(t)}{f}v_{sensory,y} \\ &= \frac{h\tan(\phi_0 + \theta)}{f}v_{sensoy,y} \end{aligned} \qquad (7)$$

Following Equations (6) and (7), (v_x, v_y) can be formulated using the following formula:

$$\begin{cases} v_x = [\dfrac{h}{f - \tan(\phi_0)x_{sensor}}]v_{sensor,x} \\ v_y = [\dfrac{h(f\tan(\phi_0)+x_{sensor})}{f(f-\tan(\phi_0)x_{sensor})}]v_{sensor,y} \end{cases} \quad (8)$$

3.6 Documenting Video Processing Results

All the pre-processing results are stored in XML-format. XML-format pre-processing data consists of (1) camera data, and (2) pedestrian data. As illustrated in Figure 3, those data will be stored and employed to formulate the mathematical model of the pedestrians' motion.

IV. Mathematical Description of CMA-AHT Model

4.1 Governing Equations for CMA-AHT Model

The mathematic model for crowd motion is formulated based on the hypothesis of ergodicity [7,13,25,26], which indicates that a dynamical system has the similar (or periodic) behavior averaged over time as averaged over the space of all the system's states.

The governing equation for the CMA-AHT model is given in Formula (9):

$$C(X)\frac{\partial T(X,t)}{\partial t} + \nabla\cdot(-K(X):\nabla T(X,t)) = Q(X) \quad (9)$$

It is observed that the CMA-AHT model shares similar governing equations with the canonical heat-transfer model [18,19,20] while having a different physical explanation for this application. In equation (9), X indicates the position of pedestrian; t indicates the time; $C(X)$ indicates the pedestrian capacity at location X; pseudo thermal conductivity tensor $K(X)$ indicates the pedestrian throughput that is derived from the historical crowd motion data, for example, grassland should have smaller $\|K(X)\|$ (i.e., the norm of tensor) value compared to that of a walkway; $Q(X)$, the pseudo energy per unit volume generated per unit time, indicates the normal behavior initiated within the scene; and pseudo temperature $T(X,t)$ indicates the empirical crowd's pseudo kinetic energy density.

An explicit quantitative definition about pseudo temperature $T(X,t)$ is not given in this work because the value of pseudo temperature itself is not our concern. Quantitatively, $T(X,t)$ is implicitly defined through the definition about its gradient $\nabla T(X,t)$:

$$K(X)\nabla T(X,t) = \left\langle \sum_i m_i v_i \right\rangle \quad (10)$$

In above equation, $\left\langle \sum_i m_i v_i \right\rangle$ indicates the expected collective momentum (m_i = mass, v_i = velocity) of the crowd. $\nabla T(X,t)$ will be employed to measure and analyze the motion of pedestrians.

Similar to a canonical heat-transfer model, which is derived from the law of conservation of energy and Fourier's law, CMA-AHT model is derived from the law of conservation of the crowd's pseudo kinetic energy, Fourier law, and the following extra assumptions.

• Anomalous action does not initiate within a boundary cell (i.e., an image is divided into a grid of scene and boundary are located at the edges of image).

• Normal behavior initiated within the boundary cell can be modeled by $Q(X)$, the pseudo-energy generated within the scene.

• Crowd motion has Gaussian distribution [11].

According to the law of conservation of crowd's pseudo-kinetic energy, where the rate of change of pseudo-kinetic energy should be equal to the sum of pseudo-energy flow across the boundary per unit time and pseudo-energy generated inside per unit time:

$$\frac{\partial T(X,t)}{\partial t} = \nabla \circ f(X,t) + Q(X) \quad (11)$$

Furthermore, based on the assumptions that (1) an entering pedestrian contributes positive kinetic energy; (2) an existing pedestrian contributes negative kinetic energy, Fourier's Law also holds true because a dominant number of pedestrians will move from a high energy position to a low energy location. Therefore,

$$f(X,t) = -K(X)\nabla T(X,t) \quad (12)$$

In CMA-AHT model, the kinetic flow of the crowd is quantitatively defined as the expected value of the collective crowd moment during a specific time period $[t-\Delta t, t+\Delta t]$:

$$f(X,t)|\{X \in \Gamma_1\} = \left\langle \sum_{t\in N, X\in\Gamma_1, \tau\in[t-\Delta t, t+\Delta t]} m_i v_i(X,\tau) \right\rangle \quad (13)$$

Coupling Equations (11) and (12), the governing equation (9) for the CMA-AHT model is formulated. In this paper, the CMA-AHT model only discusses two-dimensional problems, where the governing equation becomes:

$$C(x,y)\frac{\partial T(x,y,t)}{\partial t}$$

$$= \frac{\partial}{\partial x}(K_x \frac{\partial T(x,y,t)}{\partial x}) + \frac{\partial}{\partial y}(K_y \frac{\partial T(x,y,t)}{\partial y}) + Q(x,y) \quad (14)$$

It should be emphasized that in most cases K_x does not have to be equal to K_y because in reality the crowd motion is anisotropic. Given appropriate boundary conditions, Equation (14) can be solved using suitable numerical methods (e.g., finite difference or finite element) and finally ∇T, the pseudo temperature gradient, will be obtained.

The following section will focus on the formulation of physics domain Ω and boundary conditions.

4.2 Geometry Configuration of CMA-AHT

To formulate the mathematics model of CMA-ATH, the physics domain Ω for pseudo temperature field is directly derived from the geographic information system such as Google-Earth in this work. As an alternative way, an accurate geometrical configuration should be generated using image vectorization software.

4.3 Boundary Conditions Formulation for CMA-AHT

Once the geometry configuration of the CMA-AHT model is formulated, the temperature field about it will be determined with sufficient boundary conditions.

The boundary conditions for the CMA-AHT model are formulated based on the following assumptions:

• The crowd motion is periodic, either daily, weekly or corresponding to special event;

• The crowd motion is Gaussian-distributed.

Table 2: sample historical motion data around t=9:00am

DATE	(x, y)	(v_x, v_y)	ANNOTATION
1998.7.3, 8:59am	(345m, 200m)	(-5.0m/s, 6.0m/s)	*
1998.7.3, 9:01am	(215m, 120m)	(1.5m/s, 2.6m/s)	*
1998.7.4, 8:58am	(125m, 100m)	(2.0m/s, -3.3m/s)	Event (Firework)
1998.7.4, 9:02am	(-111m, 200m)	(0.5m/s, 0.2m/s)	Event (Firework)
…	…	…	…
2011.7.5, 9:00am	(-32m, -215m)	(2.5m/s, 7.0m/s)	*

As a result, the boundary conditions for CMA-AHT model are directly derived from the historical pedestrians' motion record. For example, to formulate the temperature field at 9:00am, we can employ the historical pedestrians' motion data during time period [9.0-0.1, 9.0+0.1] (illustrated in Table 2).

The boundary conditions for CMA-AHT model consist of Neumann boundary conditions and Dirichlet boundary conditions [18,21,22].

Neumann boundary conditions ($\Gamma_1 \subseteq \partial\Omega$) indicates expected collective momentum that moves across boundary Γ_1 during the period $[t-\Delta t, t+\Delta t]$. In this work, Neumann boundary condition is defined by the following formula:

$$-K\nabla T(X,t)|\{X \in \Gamma_1\} = \left\langle \sum_{i \in N, X \in \Gamma_1, \tau \in [t-\Delta t, t+\Delta t]} m_i v_i(X,\tau) \right\rangle \quad (15)$$

Dirichlet boundary conditions ($\Gamma_2 \subseteq \partial\Omega$) indicate pseudo kinetic energy.

$$T(X,t)|\{X \in \Gamma_2\} = T_0 \quad (16)$$

In CMA-AHT model, Dirichlet boundary conditions are only used to ensure the non-singularity of the governing equations. In Formula (16), T_0 is a meaningless arbitrary value.

Figure 9 shows the Neumann boundary (Γ_A, Γ_B and Γ_C) and Dirichlet boundary (T_0) conditions that are derived from the historical crowd motion data at around 9:00am, which is illustrated in Figure 8. It should be remarked that Neumann boundary conditions are effective only when there is dominant number of entry targets or exit targets.

A finite element method [18,21,22] is used to solve the governing equation (9) with Neumann and Dirichlet boundary conditions. Figures 10 (a) and (b) shows the resulting pseudo-temperature field and the temperature gradient, from which the expected motion velocity of crowd is illustrated.

V. Measuring the Motion of Pedestrians Using Pseudo Temperature Field

The last section mainly discusses the details of the formulation of the pseudo temperature field. In this section, the resulting temperature field will be used to measure and evaluate the motion of an individual pedestrian and a crowd.

5.1 Quantitative Measurement of a Pedestrian's Movement

Assuming that a pseudo temperature field for a specific location (X) and time-period ($[t-\Delta t, t+\Delta t]$) has been formulated and thus the resulting pseudo-

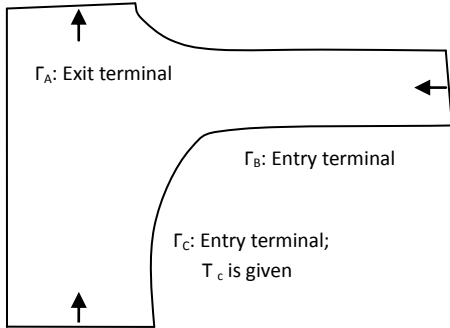Fig. 8 Historical crowd motion status at around 9:00am

Fig. 9 Boundary conditions derived from historical data around 9:00am
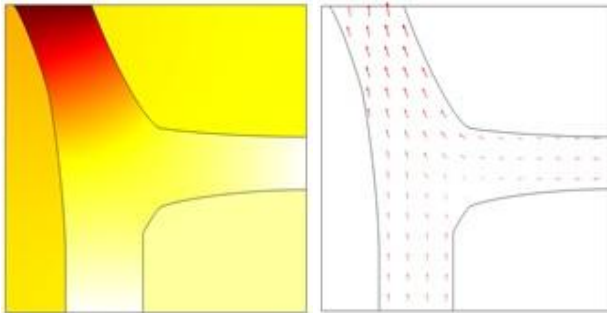

Fig. 10 Pseudo temperature field around 9:00am: (a) pseudo temperature field; (b) pseudo temperature gradient
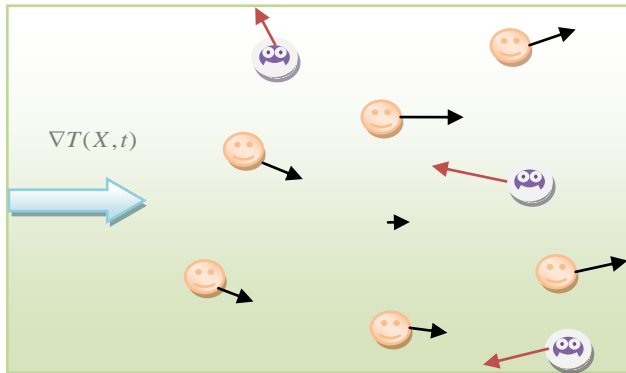

Fig. 11 Pedestrian motion within the CMA-AHT pseudo-temperature field

temperature-gradient $\nabla T(X,t)$ is known, then a pedestrian at similar location and time-period with observed velocity $v(X,t)$ can be measured by the ζ-value. , which is defined using following formula:

$$\xi(t) = \nabla T(X,t) \circ v(X,t) \qquad (17)$$

In Formula (17), "$\circ$" indicates inner-product.

Figure 11 shows that a crowd of pedestrians move within a pseudo temperature field, where the pseudo temperature gradient $\nabla T(X,t)$ is indicated by big arrow. The small arrows indicate the observed moving velocity of the individuals. Using Formula (17) we can obtain the ζ-value of each pedestrian. It is observed that the pedestrians in the diagram are classified by two groups (i.e., smiling faces and angry faces) with reference to $\nabla T(X,t)$. As illustrated in Figure 11, the "smiling face" indicates the pedestrian whose moving direction is compatible with the expected crowd's moving direction, which is derived from the historical data. The "angry face" indicates the pedestrian whose moving direction conflicts with the expected moving direction.

5.2 Measurement of a Pedestrian's Movement with Reference to the Neighboring Pedestrians in the Scene

Given the pseudo temperature field, pedestrians can be categorized according to their ζ-value. Figure 12 shows the ζ-value histogram of a group of pedestrians. Circles A and C indicate the anomalous pedestrians who move faster than others. Circle B indicates the anomalous pedestrians who stand still or move in the direction orthogonal to $\nabla T(X,t)$.

Figures 13 and 14 show that, with reference to the pseudo temperature field formulated in Figure 10, the pedestrians in the scene at around 9:00am can be categorized into the positive-ζ-value and negative-ζ-value groups. The latter group of pedestrians is regarded as anomalous because only minority of pedestrians in the scene during that period has negative ζ-value.

Furthermore, according to formula (17), the motion of a group of people can be accessed using the following formula:

$$\Psi = \sum_{i=1}^{n} \xi_i = \sum_{i=1}^{n} (\nabla T(X_i,t) \circ \vec{v}_i)$$

where n indicate the number of people of the group.

5.3 Quantitative Measurement of Pedestrians' Movement during Specific Time Period

The movement of a pedestrian during a specific time period can be also measured according to the time-sequence of ζ-value.
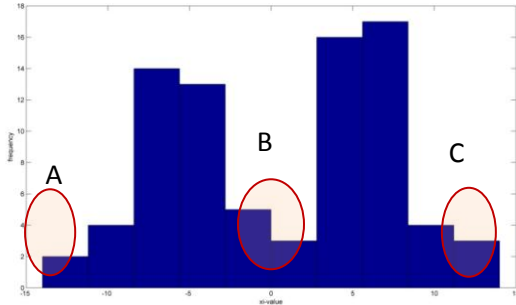
Figure 15 shows the sample time-sequence of ζ-

Fig. 12 Histogram of ζ-value of pedesitrians

Fig. 13 Normal pedestrians (with positive ζ-value, squared in green-color) appeared in the scene at around 9:00am

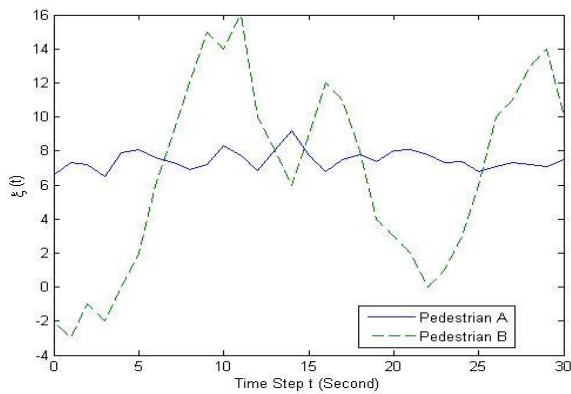Fig. 14 Anomalous pedestrians (with negative ζ-value, squared in red-color) appeared in the scene at around 9:00am

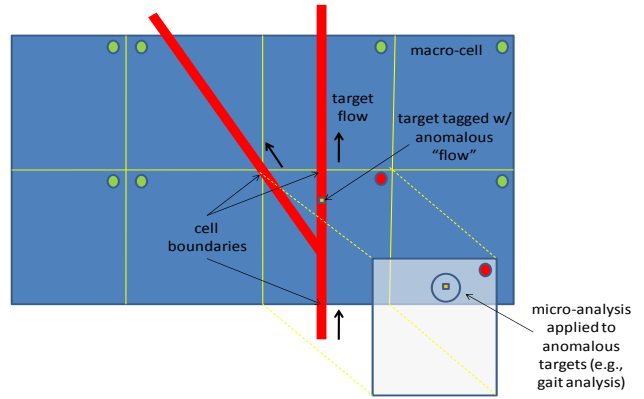Fig. 15 ξ-value about two sampling pedestrians in 30 seconds period

Fig. 16 Enhanced CMA-AHT model: using macro-cell strategy to analyze crowd-behavior within a large-scale scene (the red lines indicate path-ways), which is partitioned into multiple non-overlapping subdomain

value about two pedestrians --Pedestrian A and Pedestrian B within thirty seconds. It is observed that Pedestrian B has a larger ζ-value fluctuation than Pedestrian B. Due to his/her super inconsistent moving behavior, Pedestrian B is regarded as a relatively anomalous pedestrian.

VI.   Using Macro-Cell Strategy to Handle Large-Scale Scene

Instead of only a street block, crowd motion analysis might be needed in a large scene such as a community, a city or even a country. As a result, a CMA-AHT framework should be scalable so as to solve the large-scale problems.

In our future work, a macro-cell strategy, which partitions the global physics domain into multiple subdomains and then manipulates them independently [21,22], will be employed to enhance the capability of the CMA-AHT framework to handle large-scale problems. As illustrated in Figure 16, each cell will work independently. Inter-cell communications only occur between neighboring cells and they are only triggered while somewhat anomalous crowd behavior is observed and detected.

VII.   Summary

This paper develops a sharable self-optimizing and cooperative control sensor network platform for sensor-oriented applications such as persistent surveillance on human behavior and the tracking of environmental phenomenon. As demonstrated in experimental results, the CMA-AHT model can efficiently measure and detect those anomalous individuals within a crowd.

As a multidisciplinary research topic, this work involves image/video processing, partial differential equations, statistics physics, cooperative control, finite-element-method, and optimization, etc., This research integrates the following cutting-edge techniques: processing, analysis, fusion, documenting, compression, storage, and management of heterogeneous sensory data; mathematical modeling of temporal and spatial-dependent (or geographic-dependent) events according to discrete historical observation; evaluation and prediction of observed events according to mathematical modeling of events; full exploitation of the sensor asset according to specific situational awareness; and cooperative control of sensor nodes. Further investigations about these challenging issues need to be made in our future work. This work can be extended to more complicated real-world problems. As a summary, future work includes:

• Introducing stochastic processes to handle uncertainty [26].

• Introducing Lagrange multiplier to handle the influence of geometric constraints (such as train, bus) over the crowd movement.

• Extending the CMA-AHT for three-dimensional scenario.

• Connection with sensor network [8].

• Using macro-cell-based CMA-AHT framework to formulate larger-scale scenes such as community, town or even a city.

• Self-optimization of the exploitation of sensory data and the collective control of sensor asset. Self-optimization will be implemented using principal component analysis (PCA) method. As a cutting-edge research topic in control theory, collective control of sensor asset will also be investigated in our future work.

In addition, the CMA-AHT framework addressed in this paper provides a generic strategy for the processing and analysis of sensory data. Besides crowd-motion analysis, it can be customized and applied in many other sensor-oriented problems such as the simulation of the spread of epidemic disease [20], the management and control of traffic light, prediction of the immigration of locust, etc.

ACKNOWLEDGMENT

REFERENCES

[1] J.K. Aggarwal, J.K., and Cai, Q. (1999). "Human Motion Analysis: A Review". Computer Vision and Image Understanding, 73 (3), 428-440.

[2] AlGadhi, S. and Mahmassani, H. (1991). "Simulation of crowd behavior and movement: Fundamental relations and applications". Transportation Research Record, (1320):260-268.

[3] Al-nasur, S.; Mahmassani, H. (1991). "Simulation of crowd behavior and movement: Fundamental relations and application". Transportation Research Record, 1320, 260-268.

[4] Blue, V. and Adler, J. (1998). "Emergent fundamental pedestrian flows from cellular automata microsimulation". Trasportation Research Record, 1644:29-36.

[5] Boulic, R., Thalmann, M.N., and Thalmann, D. (1990). "A global human walking model with real-time kinematic personification". Visual Computing, 6, 344-358.

[6] Bovik, A. C. (2010), "Handbook of image and video processing (2nd Edition)", ISBN: 978-0-12-119792-6.

[7] Chandler, R.E., Herman, R.; Montroll, E. W. (1958). "Traffic dynamics: studies in car following". Operations Research, 6, 165-184.

[8] Dargie, W. and Poellabauer, C., "Fundamentals of wireless sensor networks: theory and practice", John Wiley and Sons, 2010 ISBN 978-0-470-99765-9, pp. 168-183, 191-192.

[9] Fan, J., El-Kwae E.A., Hacid M-S, and Liang F. (2002). "Novel tracking-based moving object extraction algorithm". J. Electron Imaging, 11, 303.

[10] Fernandes S., Liang, Y., Sritharan S.I., Wei, X., and Kandiah, R., "Real Time Detection of Improvised Explosive Devices using Hyperspectral Image Analysis", 2010 IEEE National Aerospace and Electronics Conference (NAECON 2010).

[11] Gardiner, C., "Handbook of stochastic methods: for physics, chemistry & the natural sciences" (Series in synergetics, Vol. 13), (3rd Ed.)", ISBN-13: 978-3540208822.

[12] Gips, P. G. and Marksjo, B. (1985). "A micro-simulation model for pedestrian flows". Mathmatics and Computer in Simulation, 27:95-105.

[13] Golubitsky, M., "Introduction to Applied Nonlinear Dynamical Systems and Chaos". New York: Springer-Verlag, 1997.

[14] Green, R.D., Guan, L., and Burne, J.A. (2000). "Video analysis of gait for diagnosing movement disorders". J Electron Imaging, 9, 16.

[15] Helbing, D. (1992). "A fluid-dynamic model for the movement of pedestrians". Complex Systems, 6:391-415.

[16] Henderson, L. F. (1974). "On the fluid mechanic of human crowd motion". Transportation research, 8:509-515.

[17] Hughes, R. L. (2002). "A continuum theory for the flow of pedestrians". Transportation Research Part B, 36:507-535.

[18] Incropera F. P., "Fundamentals of Heat and Mass Transfer (6th Edition)". ISBN-13: 978-0470055540.

[19] Liang, Y., Melvin, W., S. I. Sritharan, S. I., Fernandes, S. and Barker, D., "CMA-HT, a Crowd Motion Analysis Framework Based on Heat-transfer-analog Model". Proc. SPIE 8402, Evolutionary and Bio-Inspired Computation: Theory and Applications VI, 84020J (May 1, 2012); doi:10.1117/12.919088.

[20] Liang, Y., Shi, Z.J., Sritharan S. I., and Wan H. (2010). "Simulation of the Spread of Epidemic Disease". COMSOL 2010. Boston.

[21] Liang, Y, Szularz, M., and Yang, L. T., "Finite-element-wise Domain Decomposition Iterative Solvers Based on Polynomial Preconditioning", DOI: 10.1016/j.mcm.2012.11.017, Mathematical and Computer Modeling (in press).

[22] Liang, Y., Weston, J., and Szularz M., "Generalized least-squares polynomial preconditioners for symmetric indefinite linear equations". Parallel computing 28(2): 323-341 (2002).

[23] Lighthill, M.H., Whitham, G.B.,(1955), "On kinematic waves II: A theory of traffic flow on long, crowded roads". Proceedings of the Royal Society of London Ser. A 229, 317-345.

[24] Markos Papageorgiou, "Some remarks on macroscopic traffic flow modeling", Elsevier Science Ltd., Vol. 32, No. 5, pp. 323 to 329, 1998

[25] Nott, M. (2005). "Teaching Brownian motion: demonstrations and role play". School Science Review, 86, 18-28.

[26] Sethna J. P., "Statistical Mechanics Entropy, Order Parameters and Complexity", ISBN13: 9780198566779.

[27] Van Dorp P. and Groen F.C.A. (2008). "Feature-based human motion parameter estimation with radar". IET Proc. on Radar, Sonar, and Navigation, 2 (2), 135-145.

[28] Van Dorp, P. and Groen, F.C.A. (2003). "Human walking estimation with radar'. IET Proc. on Radar, Sonar, and Navigation", 150 (5), 356-365.

[29] Zhang, H. (1998). "A theory of nonequilibrium traffic flow". Transportation Research B, 32:485-498.

Schematic Modelling for Enhancing the Cumulative Network Lifetime of Mobile Adhoc Network using Power Efficient AODV

Rajesh Kumar Yadav and Deepak Panwar
Computer Science & Engineering, Amity University Rajasthan, India

Abstract – the planned scheme highlights one of the schematic modeling for enhancing the cumulative network lifetime of MANET using power efficient AODV. MANET system is always associated with the design constraint from unwanted power drainage during communication. The proposed system considers the intermediate mobile nodes as vital factor which estimates the current mean power of the cumulative network as comparison threshold in order to evaluate the response of route request message along with impact of the proposed system towards the routing performances. Experimented in NS2, the proposed system shows optimal throughputs which can be definitely used for optimizing the energy on overloaded nodes in MANET and enhance the cumulative network lifetime extensively.

Key Words – component; Mobile Adhoc Network, Energy, Routing protocol, AODV

I.   Introduction

The area of mobile adhoc network or commonly termed as MANET [1] is currently under the scanner of extensive research due to the massive advantages it permits on its application. MANET can be termed as a system of wireless mobile nodes which dynamically self-organize in random and impermanent network topologies. Using this technology, various users can therefore be connected in networking areas without any presence of pre-existing networking infrastructure. The mobile nodes can directly communicate with each other within their transmission ranges. In this environment, dual condition can surface where all the mobile nodes which have participated in the transmission involuntarily generate a wireless network, consequently, such types of wireless adhoc network can be visualized as mobile adhoc network. For the purpose of introducing diversified power issues in MANET [2], various power effective routing strategies has already been seen in the review of literature. Any transitional node holds the request packet for an epoch of time previous to forwarding to next mobile node. The time period is set to be inversely proportional to its

existing power so that nodes with lower level of power can be protected. The tradeoff between the cost of power utilization for distribution traffic and the enhanced spatial allocation of power overloads is discussed. Unwanted power utilization due to overhearing is considered, which involved not only the battery reservation of the mobile nodes in the routes, but also the amount of neighbors that these nodes may potentially interfere. The most favorable cardinality of the cooperation sets at each hop on a path to minimize the total power cost per transmitted bit, while a traffic-adaptive routing protocol was proposed to optimally combine the proactive and reactive strategies. The idea that each mobile node evaluates if the route request control message should be forwarded by comparing its residual power with a threshold, where the threshold is attuned incessantly as the network runs. Almost all mobile devices are supported by battery powers, so the power-efficient issue is one of the most important design issues in MANET. Solutions to the energy-efficient issue in MANET can generally be categorized as follows: 1) Low-Power Mode, in which mobile devices can support low-power sleeping mode. The main research challenges in low-power mode are that at what time mobile node can turn to sleeping mode, and at what time it should wake up. Corresponding issues are addressed in [3], [4], [5] and etc; 2) Transmission Power Control: In wireless communication, transmission power has strong impact on transmission range, bit error rate and inter-radio interference, which are typically contradicting factors. By adjusting its transmission power, mobile node can select its immediate neighbors from others, thus the network topology can be controlled in this way. How to determine transmission power of each node so as to determine the best network topology has been addressed in [6], [7], [8] and etc; 3) Power-Aware Routing: Other than the common shortest-hop routing protocols, such as DSDV [9], AODV [10], DSR [9], and etc, power-aware routing protocols take various power metrics or cost functions into account in route selection

The proposed system is designed on the backbone of frequently used AODV (Adhoc on demand distance

vector) routing protocol for enhancing the cumulative lifetime of mobile adhoc network. In section 2, we give an overview of related work which identifies all the major research work being done in this area. Problem description is discussed in Section 3 followed by proposed system illustration in Section 4. Section 5 discusses about result analysis and finally in section 6, we make some concluding remarks.

II.   Related Work

K. Arulanandam and Dr. B. Parthasarathy [11] present Energy is the scarcest resource for the operation of the mobile ad hoc networks. Idle energy consumption is responsible for a large portion of the overall energy consumption in the wireless interfaces of the mobile nodes. Sunsook Jung et.al [12] considers energy constrained routing protocolsand workload balancing techniques for improving MANET routing protocols and energy efficiency. Also, they show new application of energy efficiency metrics to MANET routing protocols for energy efficiency evaluation of the protocols with limited power supply. Rekha Patil [13] proposes a cost based power aware cross layer design to AODV. The discovery mechanism in this algorithm uses Battery Capacity of a node as a routing metric this approach is based on intermediate nodes calculating cost based on Battery capacity. Rutvij H. Jhaveriand Ashish D. Patel [14] has discussed some basic routing protocols in MANET like Destination Sequenced Distance Vector, Dynamic Source Routing, Temporally-Ordered Routing Algorithm and Ad-hoc On Demand Distance Vector.   Main objective of writing this paper is to address some basic security concerns in MANET, operation of wormhole attack and securing the well-known routing protocol Ad-hoc On Demand Distance Vector. Xiangpeng Jing and Myung J. Lee [15] presents a comprehensive energy optimized (locally and globally) routing algorithm and its implementation to AODV this algorithm investigates the combination of device runtime battery capacity and the real propagation power loss information, obtained by sensing the received signal power, without the aid of location information. Abdusy Syarif and Riri Fitri Sari [16] we present some improvement suggestion to AODV routing protocol. Our proposed protocol, called AODV-UI, improved AODV in gateway interconnection, reverse route and in energy consumption. They also measure performance indicators for some metrics, such as energy, routing overhead, end-to-end delay, and packet

delivery ratio. And also they performed simulation scenarios with three mobility model with different maximum speed and sources in hybrid ad hoc network. Preeti Bhati et.al [17] in this paper has tried to remove the existence of misbehaving nodes that may paralyze or slows down the routing operation in MANET. This increases the efficiency of a network. Efficiency can be calculated by the parameters or factors such as transmission capacity, battery power and scalability. They are considering the most crucial factor named as transmission capacity of a node. Annapurna P Patil et.al [18] concentrated on emergency search and rescue operations which rely heavily on the availability of the network.   Dr. Sanjay Sharma and Pushpinder Singh Patheja [19] present some improvement suggestion to AODV routing protocol. They proposed protocol, called AODV-PP, improved AODV in Priority models and in Power consumption. They also measure performance indicators for some metrics, such as energy, routing overhead, end-to-end delay, and packet delivery ratio, in WiMAX adhoc network. Nishant Gupta and Samir R. Das [20] develop a technique to make these protocols energy-aware in order to increase the operational lifetime of an ad hoc network where nodes are operating on battery power alone and batteries cannot be recharged. Jayesh Kataria et.al [21] present Reactive routing protocols like Ad-hoc On-Demand Distance Vector Routing (AODV) and Dynamic Source Routing in Ad-Hoc Wireless Networks (DSR) which are used in Mobile and Ad-hoc Networks (MANETs) work by flooding the network with control packets. P. Latha and R. Ramachandran [22] proposed protocol Energy Reduction Aware Multicast (ERAM) aimed to find a path which utilizes the minimum energy to transmit the packets between the source and the destination. Shivendu Dubey, Prof. Rajesh Shrivastava [23] present A mobile ad hoc network (MANET) is a collection of wireless mobile nodes dynamically forming a network Topology without the use of any existing network infrastructure or centralized administration. Sajjad Ali & Asad Ali [24] present the communication between these mobile nodes is carried out without any centralized control. M. Tamilarasi and T.G. Palanivelu [25] propose a mechanism which integrates the adaptive timeout approach, load balancing approach and transmit power control approach to improve the performance of on-demand routing. they applied this integrated mechanism on Ad hoc On- demand Distance Vector(AODV) routing

protocol to make it as Energy Aware Adaptive AODV (EAA AODV) routing protocol. Lijuan Cao et.al [26] provides a survey and analysis of energy related metrics used for ad hoc routing. First, the most common energy efficient routing protocols are classified into four categories based on the energy cost metrics employed. Chansu Yu et.al [27] purpose of this paper is to facilitate the research efforts in combining the existing solutions to offer a more energy efficient routing mechanism. Tanu Preet Singh et.al [28] present in this paper an ad-hoc network is a local area network (LAN) that is built spontaneously as devices connect. Instead of relying on a base station to coordinate the flow of messages to each node in the network, the individual network nodes forward packets to and from each other. M. Tamilarasi and T.G. Palanivelu [29] present in this paper, a mechanism involving the integration of load balancing approach and transmission power control approach is introduced to maximize the life-span of MANETs. The mechanism is applied on Ad hoc On-demand Vector (AODV) protocol to make it as energy aware AODV (EA_AODV). Mahesh K. Marina and Samir R. Das [30] develop an on-demand, multipath distance vector routing protocol for mobile ad hoc networks. Specifically, they propose multipath extensions to a well-studied single path routing protocol known as ad hoc on-demand distance vector (AODV).

III. Problem Description

In the area of mobile adhoc network, routing is one of the prominent issue which surfaces because of highly dynamic and distributed environment in MANET. The power efficiency in mobile adhoc network has become one of the critical design factors as the mobile nodes will be supported by battery with limited capacity. The failure or degradation of energy in mobile nodes will not only influence the node itself but it will also have impact into its potential to forward the packets on behalf of others and therefore influence the cumulative network lifetime. Hence, majority of the researchers has attempted for designing power aware routing algorithms for specific mobile adhoc network scenario. Unfortunately, it is still in infancy stage as it is still not obvious that which one of the lists of routing protocols is best for majority of scenarios as every routing protocol is designed to work for only specific environment. But, it is also highly feasible to unite and incorporate the current solutions in order to facilitate maximum power efficient routing techniques. As

power efficiency is also vital issue in many other network layers, considerable efforts has already been given for designing power aware MAC as well as transport protocols. Each layer is believed to function in remoteness in layered network architecture but, as some current research suggested, the cross-layer design is indispensable to exploit the highest power performance. In fact, many routing protocols analyzed in survey also deploy the similar concept, i.e. they utilize lesser layer techniques such as transmission energy control and sleep mode methods in their routing layer protocols.

IV. Proposed System

The main aim of the project work is to introduce a novel scheme based on AODV, called Energy Saving Ad-hoc On-demand Distance Vector for routing in MANETs. It also achieves the energy information exchange among neighboring nodes through already-existed signaling packets in AODV and introduces a new network parameter as the comparison threshold, called current average energy of the network, which can fairly accurate estimate the mean power consumption of the cumulative mobile network.

Consider a MANET with n mobile nodes which are uniformly distributed in a field with a size of axb square meters. The wireless transmission range of each node is set to r meters. In such scenario, the distance between two arbitrary nodes, denoted as d, has probability density function (PDF) as

$$f_d(x) \begin{cases} 2x\left[\dfrac{\pi}{ab} - \dfrac{2x}{ab^2} - \dfrac{2x}{ba^2} + \dfrac{x^2}{a^2b^2}\right] & 0 \le x \le b \\[2ex] \dfrac{2x}{ab}\left[\dfrac{\pi}{2} - A\right] - \dfrac{2x}{a^2} - \dfrac{4x}{ab^2}\left[x - \sqrt{x^2 - b^2}\right] & b \le x < a \\[2ex] \dfrac{2x}{ab}B - \dfrac{4x}{ab^2}\left[a - \sqrt{x^2 - b^2}\right] & \\[2ex] \dfrac{2x}{a^2b^2}C + \dfrac{4x}{ba^2}\left[\sqrt{x^2 - a^2} - b\right] & a \le x < \sqrt{a^2 + b^2} \end{cases}$$

where

$A = \left[\arcsin\left(x^2/2 - b^2\right)/\left(x^2/2\right)\right]$,

$B = \left[\arcsin\left((a^2 - x^2/2)/(x^2/2) - A\right)\right]$,

and $C = a^2 + b^2 - x^2$.

Given d, the average number of hops between these two nodes equals l = [d/r]

where [·] is the ceiling operator. Traffic load from each node is generated by a constant-bit-rate (CBR) source with transmission rate λ. Each node moves according to random waypoint model.

The method for determining the energy comparison threshold for each mobile node is also discussed. The main objective is to balance energy consumption among all nodes in the network.

A. Revised Control Packets

For exchanging energy information in the network, we modify the existed control packets in AODV to transmit the required energy information instead of adding new control packets.

HELLO packet: Neighboring nodes in proposed AODV utilize the HELLO packet in AODV to exchange the information of their remaining energy among each other. In order to store such information, a new field called Remaining Energy (RE) is added to the HELLO packet. With the support of the HELLO packet, every node maintains a Local Remaining Energy Table (LRET), which records the remaining energy of all one-hop neighbors.

RREQ packet: Proposed AODV also utilizes the RREQ in AODV which is forwarded in the route discovery process to cumulate the local average remaining energy information along the backward path. A new field denoted as E_{sum} is added to the RREQ packet and each intermediate node will update E_{sum} field by cumulating its own local average energy to it. As the RREQ packet being forwarded, the local energy information of all the intermediate nodes is spread through the backward path. Note that both R_E and E_{sum} fields only occupy 4 bytes in a 28 bytes HELLO packet and a 32 bytes RREQ packet, respectively. Therefore, the introduced traffic load due to energy information exchange is marginal.

B. Route Discovery

Route discovery process is initiated whenever a traffic source needs a route to the destination. Route discovery process includes the operations at source node, intermediate nodes and destination node.

Source Node: When a source node is going to communicate with a destination node and there is no available route, the source node will broadcast a RREQ packet. As the first node along the route, the source node initializes Esum with its local average energy, $\overline{E}^s_l$ which can be calculated as

$$\overline{E}^s_l = \left(E^s_r + \sum_{k=1} \Sigma^k_r \right) \Big/ \left(n_s + 1 \right)$$

where E^s_r denotes the remaining energy of the source node, which can be measured from its physical layer. E^k_r, $k =1\ldots ns$, denotes the remaining energy of its 1-hop neighboring nodes and ns denotes the number of neighbors. Notice that the availability of E^k_r at the source node results from the exchange of modified HELLO packets.

Intermediate Nodes: After receiving the RREQ packet, each intermediate node decides its own response to the received RREQ packet via comparing its remaining energy with an energy threshold which is derived as follows. Consider an intermediate node i. Following the same procedure as that in the source node, node i can calculate its local average energy as

$$\overline{E}^i_l = E^i_t /\left(n_i +1\right)$$

where E^i_t is the total remaining energy of node i and its neighbors, the number of which is defined as ni. We introduce a new parameter, called the current average energy of the network, which can be calculated as

$$\overline{E}^i_c = \left(E_{sum} + \overline{E}^i_l \right)/i$$

where i is the number of nodes along the backward path. Obviously, since the accurate average energy of the network is ordinarily unavailable, such current average energy of the network is an estimation of it. Note that when initiating an RREQ packet, the source node has added its local average energy to the Esum field, and each intermediate node which forwards this RREQ packet has updated Esum by cumulating its own local average energy. Therefore, $\overline{E}^i_c$ is a good approximation of average remaining energy, which takes into account all nodes on the route and their neighbors?

Given current average energy of the network, the energy threshold is set as

$$E^i_{thres} = \alpha \overline{E}^i_c$$

where α is the system parameter, which satisfies $0 < \alpha < 1$. α represents different protection levels. The smaller α is, the more frequently intermediate nodes participate in the route discovery process. If the remaining energy of the intermediate node is larger than the threshold, we treat this node as a node with sufficient energy. Such node should cumulate its local average energy to Esum and forwards it to a next node immediately. Otherwise, the node should wait for a while, and determines whether to forward the RREQ packet based on the number of identical RREQ packets received during the delay period. If the number is larger than a threshold C, which is the effective RREQ packet number in a neighboring area, the node drops the RREQ packet to save energy. Otherwise, if there is not enough RREQ packets received during the delay period, the node should forward this RREQ packet to participate in the route discovery.

Destination Node: Finally, when the RREQ packet has reached the destination, the destination node responds by unicasting a RREP packet back to the source node as AODV does. And the route discovery is finished.

C. Route Maintenance

Route maintenance process is the same as that in AODV protocol. When a link breaks, route error (RERR) packet will be forwarded to erase all routes using the broken link.

V. Research Methodology

In this proposed system, a new architecture based on enhancement in AODV is proposed for conducting energy efficient routing in MANETs. The proposed scheme (Figure 1) achieves the energy information exchange among neighboring nodes through already-existed signaling packets in AODV and introduces a new network parameter as the comparison threshold, called current average energy of the network, which can estimated the mean power utilization of the network. In the proposed scheme, each intermediate node determines whether to forward RREQ packet by comparing its remaining energy with current mean power of network. If the energy of the node is larger than the threshold, it will forward the RREQ packet immediately. Otherwise, the node will wait for a while to decide whether the packet should be forwarded or dropped according to the number of the identical RREQ packets received during the waiting period. After that, effects of the proposed routing protocol on network performance are addressed. Both analytical

Fig. 1 Proposed architecture

and simulation results shows that the proposed routing scheme is comparatively easier for execution and can facilitate a maximized cumulative network lifetime.

The main objectives of the proposed system are as follows:

• To create a network model considering the probability density function with respect to the node and their distance.

• To design an enhanced control packet for exchanging energy information in the network using AODV routing protocol.

• To design an efficient and energy saving route discover scheme using average energy of the network.

• To simulate the proposed energy saving scheme on Network Simulator 2 in Linux OS

• To analyze the throughput with respect to packet delivery ratio, cumulative network lifetime, as well as mean end to end delay.

The routing protocols have multiple operations to be performed apart from instituting correct and resourceful routes among the twosome of mobile nodes. The most prominent aim of the routing protocol is to render the entire networking to operate for as long duration as possible. Such types of parameters are very essential to facilitate the minimum energy path through which the cumulative utilization of power for delivering a packet is reduced. In such experiments, the wireless link is interpreted with the cost of link in terms of transmission power over the link and minimum energy path is another factor which reduces the sum of the cost of link along the same path. But, unfortunately, if such types of routing parameters are selected than it may yield to unbalanced power utilization among the mobile nodes. It was also seen that when certain specific mobile nodes are incorrectly overloaded in order to support majority of packet-relaying operation, such nodes may utilize higher battery power and impede running earlier than other mobile nodes thereby disturbing the cumulative functionality of the mobile adhoc network.

Figure 2 highlights the total of 35 mobile nodes in the simulation environment (green circle). The 1st mobile node is shown by red circle, which is represented with its respective neighbor nodes (yellow circle). Figure 3 highlights the neighboring nodes (yellow circle) of node 2. Figure 4 highlights the progress in transmission as well as draining of energy from the source node 3 via intermediate node 25 to destination node 8.

An estimation algorithm to obtain the network average remaining energy is introduced (Figure 5).
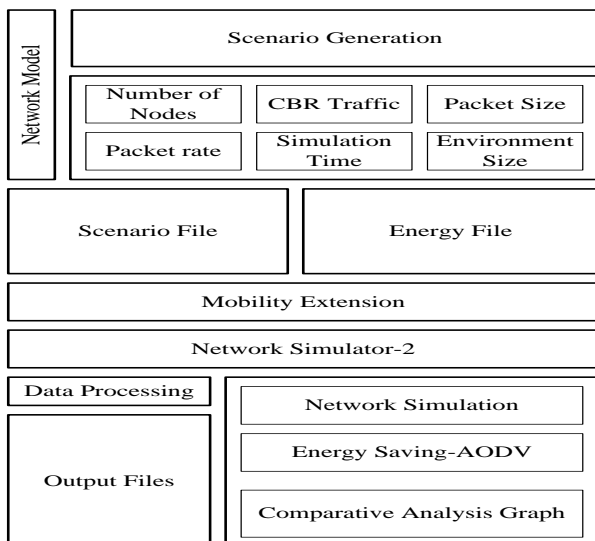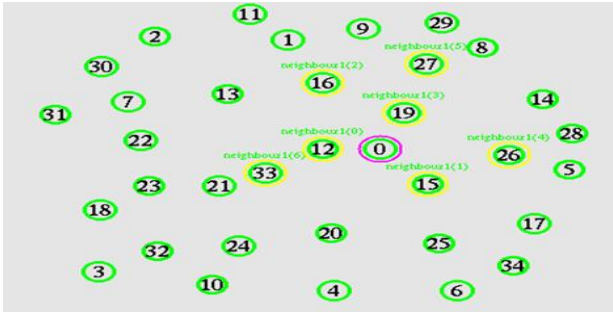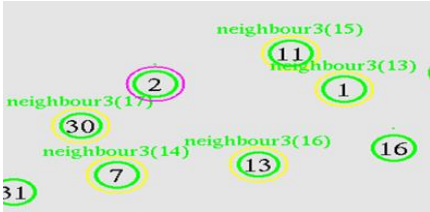
Fig. 2 Simulation Result-I

Fig. 3 Simulation Result-II
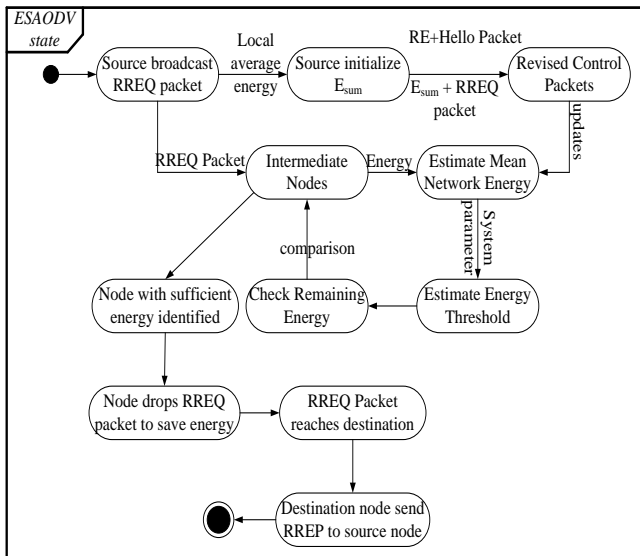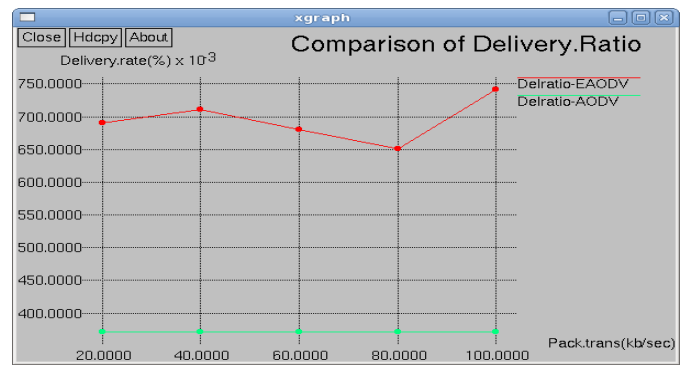

Fig. 4 Simulation Result-III

Fig. 5 Activity Diagram of the proposed system

With such assessment, intermediary mobile nodes can reasonably decide if their residual power is adequate or not. By preventing overused nodes from participating

in route discovery processes, the proposed scheme using AODV effectively balances energy consumption around the network. Simulation results will show that the proposed energy saving schema can evidently increase the lifetime of the network

VI.   Result Analysis

The proposed system is designed on Linux platform using network simulator (NS2), which is an object-oriented, discrete event driven network simulator developed at UC Berkeley. The performance analysis is checked with respect to delivery ratio, network overhead, delay, and cumulative network life time.

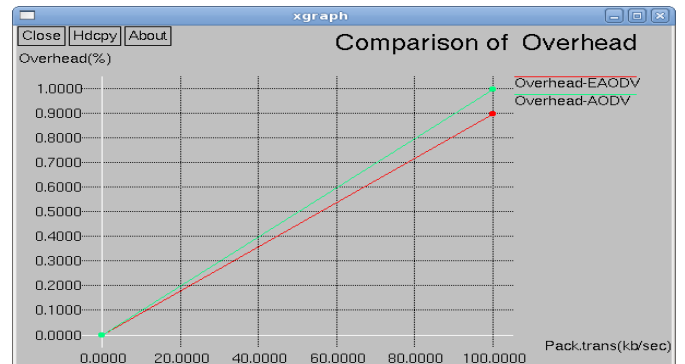Fig. 6 Simulation Result-III

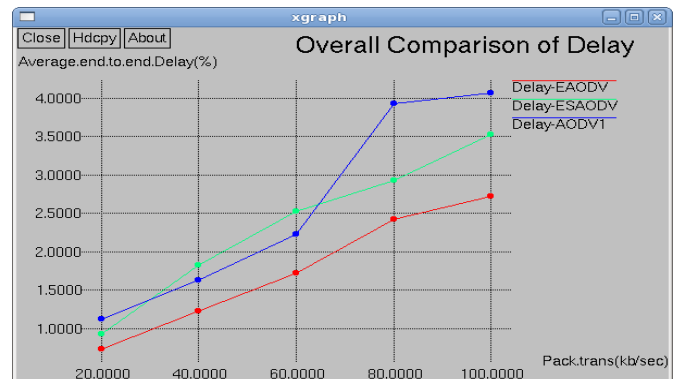Fig. 7 Simulation Result-III

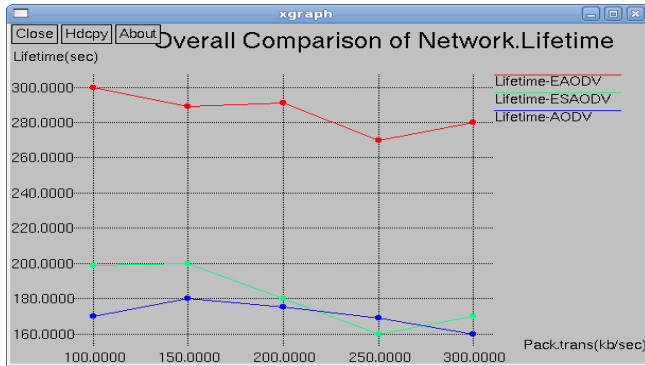Fig. 8 Simulation Result-III

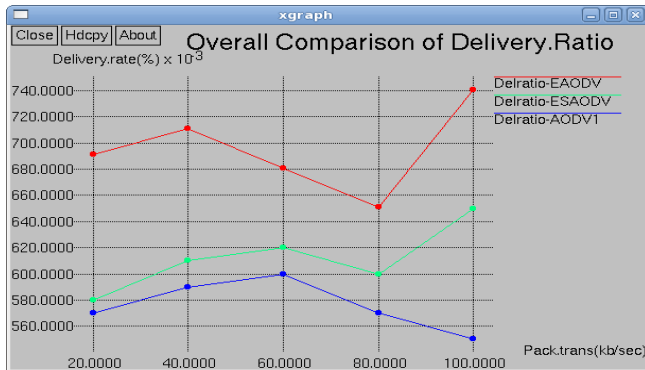Fig. 9 Simulation Result-III

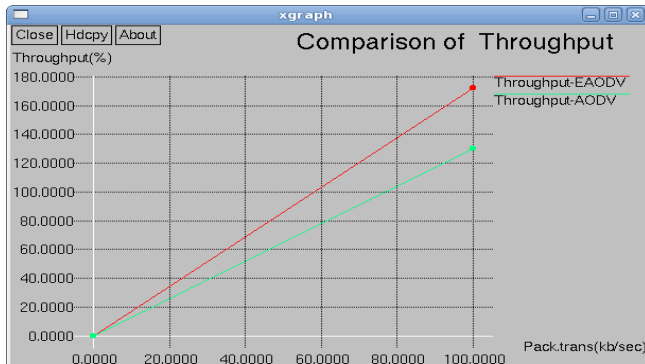
Fig. 10 Simulation Result-III
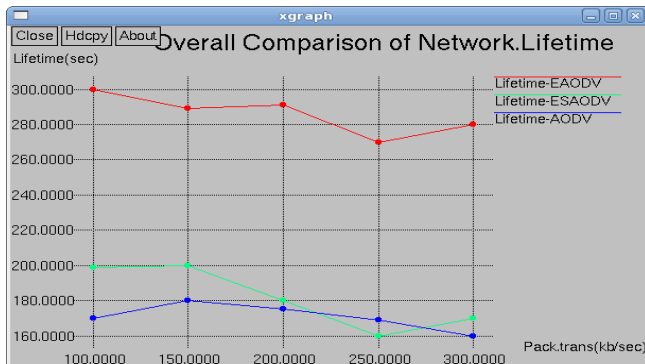

Fig. 11 Simulation Result-III

Fig. 12 Simulation Result-III

The above graph in Figure 6 shows the comparison of packet delivery ratio for proposed energy aware-AODV and AODV with x-axis of packet transmission (kb/sec) and y-axis of Delivery rate (10^3).
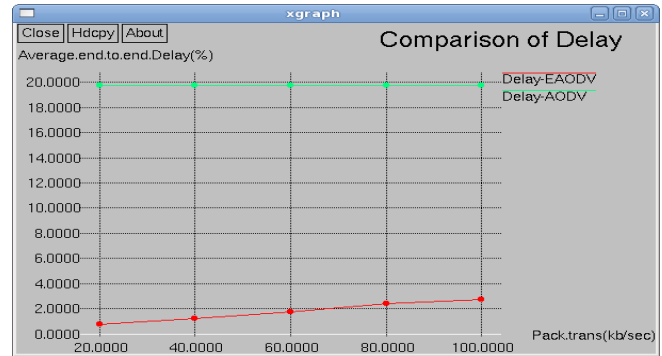

Fig. 13 Simulation Result-III

The above graph in Figure 7 shows the comparison of overhead for proposed energy aware-AODV and AODV with x-axis of packet transmission (kb/sec) and y-axis of Overhead (%).

The above graph in Figure 8 shows the comparison of delay for proposed scheme, energy aware AODV, and AODV with x-axis of packet transmission (kb/sec) and y-axis of average end-to-end Delay(%).

The above graph in Figure 9 shows the comparison of network lifetime for proposed scheme, energy aware AODV, and AODV with x-axis of packet transmission (kb/sec) and y-axis of network lifetime (sec).

The above graph in Figure 10 shows the comparison of overall packet delivery ratio for proposed scheme, energy aware AODV, and AODV with x-axis of packet transmission (kb/sec) and y-axis of delivery rate (10^3).

The above graph in Figure 11 shows the comparison of throughput for energy aware AODV, and AODV with x-axis of packet transmission (kb/sec) and y-axis of throughput (%).

The above graph in Figure 12 shows the comparison of overall network lifetime for proposed scheme, energy aware AODV, and AODV with x-axis of packet transmission (kb/sec) and y-axis of network lifetime (sec).

The above graph in Figure 13 shows the comparison of delay for energy aware AODV, and AODV with x-axis of packet transmission (kb/sec) and y-axis of average end-to-end delay (%).

VII.   Conclusion

In this proposed system, a unique power-aware routing protocol using AODV is presented. In the route discovery process of the proposed scheme, transitional and intermediate nodes estimate the current mean power of the cumulative network as an evaluation threshold to establish how to retort to the received

route request packets. An evaluation algorithm to accomplish the network mean residual power is highlighted. With such estimation, intermediate nodes can reasonably judge whether their residual power is adequate or not. By averting overused nodes from participating in route discovery processes, the proposed routing scheme using AODV efficiently stabilize the power expenditure around the cumulative network. For the analysis of the proposed system, with the assistance of graphical representation, the network lifetime of the proposed scheme and AODV with different levels of mobility and network loads is shown in result analysis. It is also shown that graph of data delivery rates with different levels of mobility and the network performance in term of mean end to end delay.

REFERENCES

[1] http://en.wikipedia.org/wiki/Mobile_ad_hoc_network. Accessed on 28th march, 2012

[2] Yu-Chee Tseng, Chih-Shun Hsu, Ten-Yueng Hsieh, Power-Saving Protocols for IEEE 802.11-Based Multi-Hop Ad Hoc Networks, INFOCOM 2002.

[3] S. Singh and C. S. Raghavendra, "Power Efficient MAC Protocol for Mul-tihop Radio Networks", Proc. of IEEE International Personal, Indoor and Mobile Radio Communications Conference, July 1998, pp. 153–157.

[4] C. F. Chiasserini and R. R. Rao, "A Distributed Power Management Policy for Wireless Ad Hoc Networks", IEEE Wireless Communication and Networking Conference, 2000, pp. 1209–1213.

[5] J. R. Lorch and A. J. Smith, "Software Strategies for Portable Computer Energy Management", IEEE Personal Communications, Jun 1998, pp. 60–73.

[6] L. Hu, "Topology Control for Multihop Packet Radio Networks", IEEE Transactions on Communications, Oct 1993, vol. 41, pp. 1474–1481.

[7] R. Ramanathan and R. Rosales-Hain, "Topology Control of Multihop Wireless Networks using Transmit Power Adjustment", Proc. of IEEE Infocom, Tel Aviv, Israel, March 2000, pp. 404–413.

[8] R. Wattenhofer, L. Li, P. Bahl, and Y. M. Wang, "Distributed Topology Control for Power Efficient Operation in Multihop Wireless Ad Hoc Net-works", Proc. of IEEE Infocom, Anchorage, Alaska, April 2001, pp. 1388–1397.

[9] C.E. Perkins and P. Bhagwat, "Highly Dynamic Destination- Sequenced Distance-Vector Routing (DSDV) for Mobile Computers", Proc. of ACM SIGCOMM, London, U.K., 1994, pp. 234-244.

[10] C. E. Perkins and E. M. Royer, "Ad-hoc On-Demand Distance Vector Routing", Proc. of 2nd IEEE Workshop on Mobile Computing Systems and Applications, New Orleans, LA, 1999, pp. 90-100.

[11] K. ARULANANDAM and Dr. B. PARTHASARATHY "A NEW ENERGY LEVEL EFFICIENCY ISSUES IN MANET" International Journal of Reviews in Computing 2009 IJRIC. All rights reserved

[12] Sunsook Jung, Nisar Hundewale, Alex Zelikovsky "Energy Efficiency of Load Balancing in MANET Routing Protocols"

[13] Rekha Patil , DrA.Damodaram "Cost Based Power Aware Cross Layer Routing Protocol For Manet" IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.12, December 2008

[14] Rutvij H. Jhaveri Ashish D. Patel Ashish D. Patel "MANET Routing Protocols and Wormhole Attack against AODV" IJCSNS International Journal of Computer Science and Network Security, VOL.10 No.4, April 2010.

[15] Xiangpeng Jing and Myung J. Lee "Energy-Aware Algorithms for AODV in Ad Hoc Networks"

[16] Abdusy Syarif and Riri Fitri Sari "Performance Analysis of AODV-UI Routing Protocol With Energy Consumption Improvement Under Mobility Models in Hybrid Ad hoc Network" Vol. 3 No. 7 July 2011

[17] Preeti Bhati "An Efficient Agent-Based AODV Routing Protocol in MANET" Vol. 3 No. 7 July 2011

[18] Annapurna P Patil "Design of an Energy Efficient Routing Protocol for MANETs based on AODV" IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No 1, July 2011 ISSN (Online): 1694-0814

[19] Dr. Sanjay Sharma, Pushpinder Singh Patheja " Improving AODV Routing Protocol with Priority and Power Efficiency in Mobile Ad hoc WiMAX Network" International Journal of Computer Technology and Electronics Engineering (IJCTEE) Volume 2, Issue 1

[20] Nishant Gupta "Energy-Aware On-Demand Routing for Mobile Ad Hoc Networks"

[21] Jayesh Kataria "ACRR: Ad-hoc On-Demand Distance Vector Routing with Controlled Route Requests"

[22] P. Latha and R. Ramachandran "Energy reduction aware multicast routing for mobile ad hoc networks" Vol. 3 No. 3 (Mar 2010)

[23] Shivendu Dubey, Prof. Rajesh Shrivastava "Energy Consumption using Traffic Models for MANET Routing Protocols" International Journal of Smart Sensors and Ad Hoc Networks (IJSSAN) Volume-1, Issue-1, 2011

[24] Sajjad Ali & Asad Ali "Performance Analysis of AODV, DSR and OLSR in MANET" Department of Electrical Engineering with emphasis on Telecommunication Blekinge Institute of Technology, Sweden 2009

[25] M. Tamilarasi1, T.G. Palanivelu, "ADAPTIVE LINK TIMEOUT WITH ENERGY AWARE MECHANISM FOR ON-DEMAND ROUTING IN MANETS" Volume 4 Number 4

[26] Lijuan Cao Teresa Dahlberg Yu Wang "Performance Evaluation of Energy Efficient Ad Hoc Routing Protocols"

[27] Chansu Yu, Ben Lee "Energy Efficient Routing Protocols for Mobile Ad Hoc Networks" This research was supported in part by the Cleveland State University, EFFRD Grant No. 0210-0630-10.

[28] Tanu Preet Singh Energy-Efficient Routing Protocols In Mobile Ad-Hoc Networks" Volume 2, Issue 1, January 2012.

[29] M. Tamilarasi, T.G. Palanivelu "Integrated Energy-Aware Mechanism for MANETs using On-demand Routing" World Academy of Science, Engineering and Technology 38 2008.

[30] Mahesh K. Marina and Samir R. Das "Ad hoc on-demand multipath distance vector routing" Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/wcm.432, 2006