Analysis and Study of Embedded Capability to HSBH by SPA

Yuanzhi Wang
School of Computer and Information,Anqing Normal College,Anhui,China
E-mail: Wangyuanzhi1@sohu.com

Liwei Chen
College of the Computer Science and Technology, Southwest University of S&T,Sichuan,China
E-mail: Chenliwei@sohu.com

Fei Zhang
School of Computer and Information,Anqing Normal College,Anhui,China
E-mail: Zhangfei@aqtc.edu.cn

Shanhe Jiang
School of Physics and Electric Engineering, Anqing Normal College, Anhui,China
E-mail: Jangshanhe@aqtc.edu.cn

Lijuan Sun
College of Computer, Harbin University of Science and Technology, HeiLongJiang,China
E-mail: wyz@aqtc.edu.cn

***Abstract*—Steganography of SPA is one of the important research subjects in the field of information security. And HSBH is an effective image hiding algorithm. It proves that SPA can detect HSBH algorithm in theory and experiment. This method not only can detect the existence of messages embedded by sequential or random replacement, but also can estimate the hiding capability. Experimental results show that the correct rate of detection is above 95% when the embedding rate p>3%..**

***Index Terms*—steganography, LSB, HSBH,SPA**

I. INTROSUCTION

As the popularity of the Internet and the bandwidth increases rapidly, various data in digital form is transmitted over the Internet network. The transmitted data can be a digital representation of text, image, audio and video. In order to ensure the security of the data transmission over the Internet, data encryption and data hiding are two widely used techniques. Data encryption is a technique to protect data from illicit access by transforming important data into meaningless code, which the interceptors know where to obtain something valuable. Nevertheless, data hiding is different from data encryption as it hides the secret data into a meaningful host data to distract the attention of the observers. Recently, hiding data in images has become a hot research topic. The image used to camouflage the secret data is called cover-image while the cover-image with the secret data embedded in is called stego-image. Various techniques about image hiding were proposed. At the same time, some analysis technology is presented, such as χ_2-statistical, RS and SPA. In this paper, through the analysis of using the SPA method to HSBH hiding algorithm, the experiment result shown that the HSBH can not resist the attacking of SPA..

Corresponding author, etc. Corresponding Author: Yuanzhi Wang is with School of Computer and Information, Anqing 246011, Anqing Normal College, CHINA.

II. HSBH ALGORITHM

HSBH algorithm is a high bit information hiding algorithm. It hides information to the four high bits, but not a fixed bit. It hides information to different high bits of carrier image according to Zn calculated by Logistic chaotic map as in (1). As the chaotic map has strong randomness and initial value sensitivity, chaotic system has good cryptographic characteristics.

$$Z_{n+1} = \lambda Z_n (1 - Z_n) \quad (1)$$

Where $Z_n \in (0,1)$, $\lambda \in (0,4]$,if $\lambda >$ 3.57, iterative sequence generated with an initial value Z0 is chaos. If the initial value Z0 has slight difference, generated sequence is entirely different. The sequence has uniform distribution in the interval (0,1) and without period. Based on these characteristics, (0,1) is equally divided

doi:10.4304/jsw.5.11.1256-1261

into four sub-intervals .When embedding secret information bits, we will make sure bit place of carrier pixel in which secret bits will be embedded according to sub-interval.

Pixels which are suitable for high bit hiding are determined by parameters L_k and r. L_k is the number of intervals which pixels are suitable for high bit hiding and is calculated as in (2). And r is the scope of gray value in the interval, namely the length of the interval.

$$L_k = 2^{(9-k)} - 1 \quad (2)$$

Where k=(5,6,7,8), provided secret information is a gray image whose matrix is

$W = \{ w(i,j) 1 \le i \le M_1, 1 \le j \le M_2 \}$, the size of the secret information image is M1 × M2. Carrier image is $F = \{ f(i,j) 1 \le i \le N_1, 1 \le j \le N_2 \}$, the size of the carrier image is N1 × N2. After embedding secret image, carrier image is called camouflage image which is written as $F' = \{ f'(i,j) 1 \le i \le N_1, 1 \le j \le N_2 \}$. Where (i, j) is the coordinate of the secret image pixel; (x, y) represents pixel coordinate of original carrier image and camouflage image respectively, and $f(x, y), w(i, j), f'(x, y)$ represent respectively pixel value of corresponding place. To calculate easily , assuming M1= M2= M , N1= N2=N. Table 1 is the range of fitting hiding; Equation 3 explains the corresponding changes of pixels in the eighth bit.

TABLE.1 THE SCOPE OF EACH HIGH HIDING BIT

Hiding bit	**Suppressible section**	**The range**
The eighth bit	1	[128-r 127] , [128 128+r]
The seventh bit	3	[64-r 63] , [64 64+r]…
The sixth bit	7	[32-r 31] , [32 32+r]…
The fifth bit	15	[16-r 15] , [16 16+r]…

If $w(i,j)_n = 0$:

$$f'(x,y) = \begin{cases} 127 & f(x,y) \in [128, 128+r) \\ f(x,y) & f(x,y) \in [128-r, 127) \end{cases}$$

If $w(i,j)_n = 1$:

$$f'(x,y) = \begin{cases} f(x,y) & f(x,y) \in [128, 128+r) \\ 128 & f(x,y) \in [128-r, 127) \end{cases} \quad (3)$$

III. SPA STEGANALYSIS

A. SPA Steganalysis Method

The principle of SPA method is based on finite-state machine theory. The states of finite-state machine are selected multisets of sample pairs. If sample pairs were drawn from images, there are some inherent relations. But after random embedding, these multisets will change, and it causes changes to these statistics relations. Assuming that the pixel value of an image is represented by the succession of samples $s_1, s_2, \cdots, s_N$, a sample pair means a two-tuple (si, sj), $1 \le i, j \le N$. Let P be a set of sample pairs drawn from an image, then P can be seen as a multiset of two-tuples (u, v), where u and v are the values of two adjacent samples, $0 \le u \le 2^b - 1$, $0 \le v \le 2^b - 1$, and b is the number of bits to represent each sample value. Denote by D_n the submultiset of P that consists of sample pairs of the form (u, u + n) or (u + n, u), i.e., where n is a fixed integer, $0 \le n \le 2^b - 1$.

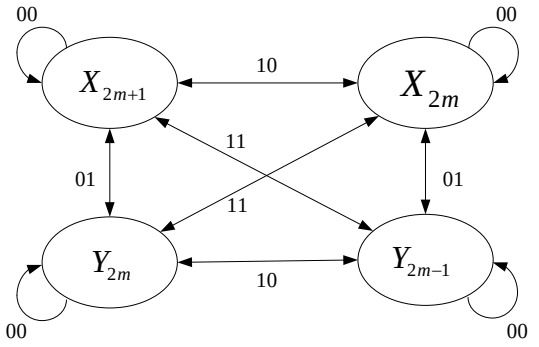
For natural images (normal signals), the probability for a sample pair in D_{2m+1} to have a larger or smaller even component is the same, all the algorithms discussed in literature are based on this important assumption.

$$|X_{2m+1}| = |Y_{2m+1}| \quad (4)$$

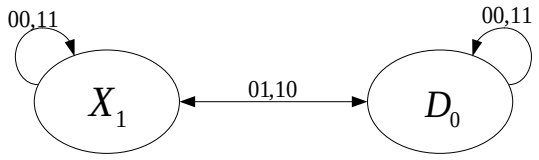
B. SPA Steganalysis to HSBH Algorithm

In HSBH algorithm, the embedded capability is higher when r=4. It will analysis the HSBH by SPA in r=1 and r=4 in the paper.

The finite-state machine in r=1 is as the figure 1:



(a) $m \ge 1$



(b) $m = 0$

Figure 1. Finite-state machine of HSBH $(r = 1)$ embedding

The embedding rate is defined p, so the probility of four modified model are $\rho(00) = (1-\frac{p}{2})^2$, $\rho(01) = (1-\frac{p}{2})\frac{p}{2}$, $\rho(10) = \frac{p}{2}(1-\frac{p}{2})$, $\rho(11) = (\frac{p}{2})^2$. $E_m = X_{2m+1} + X_{2m} + Y_{2m} + Y_{2m-1}$,so

$$\frac{(|E_m|-|E_{m+1}|)p^2}{4}-\frac{\left(\left|D'_{2m}\right|-\left|D'_{2m+2}\right|+2\left|X'_{2m+1}\right|-2\left|Y'_{2m+1}\right|\right)p}{2}+\left|X'_{2m+1}\right|-\left|Y'_{2m+1}\right|=0, m\geq 1$$

$$\frac{(2|E_0|-|E_1|)p^2}{4}-\frac{(2\left|D'_0\right|-\left|D'_2\right|+2\left|X'_1\right|-2\left|Y'_1\right|)p}{2}+\left|X'_1\right|-\left|Y'_1\right|=0, m=0 \quad (6)$$

Through the equation(5) or (6) it can gain the value of embedding rate p.

In order to gain the more accurate value p, the assumption(4) is modified (7):

$$E\left\{\left|\bigcup_{m=i}^{j} X_{2m+1}\right|\right\}=E\left\{\left|\bigcup_{m=i}^{j} Y_{2m+1}\right|\right\} \quad (7)$$

It can gain the formular (8)and (9).

$$\frac{(|E_i|-|E_{j+1}|)p^2}{4}-\frac{(|D'_{2i}|-|D'_{2j+2}|+2\sum_{m=i}^{j}|X'_{2m+1}|-2|Y'_{2m+1}|)p}{2}+\sum_{m=i}^{j}(|X'_{2m+1}|-|Y'_{2m+1}|)=0, i\geq 1 \quad (8)$$

$$\frac{(2|E_0|-|E_{j+1}|)p^2}{4}-\frac{\left[|D'_0|-|D'_{2j+2}|+2\sum_{m=0}^{j}(|X'_{2m+1}|-|Y'_{2m+1}|)\right]p}{2}+\sum_{m=0}^{j}|X'_{2m+1}|-|Y'_{2m+1}|=0, i=0 \quad (9)$$

When r=4, the pixels have five varieties of "0,1,2,3,4". They denote the change of even and odd number. So the sample pairs have 52 . Through the relations of state transfer, formular(10) can be get.

$$\left(1-\frac{p}{2}\right)^2(|X_{2m+1}|-|Y_{2m+1}|)+\frac{p}{8}\left(1-\frac{p}{2}\right)\begin{pmatrix}2|X_{2m-3}|+2|X_{2m-1}|-2|Y_{2m+5}|\\-2|Y_{2m+3}|+|D_{2m-2}|+\\|D_{2m}|-|D_{2m+2}|-|D_{2m+4}|\end{pmatrix}+\frac{p^2}{64}\begin{pmatrix}|X_{2m-7}|+2|X_{2m-5}|+|X_{2m-3}|+|Y_{2m-5}|\\+2|Y_{2m-3}|+|Y_{2m-1}|-|X_{2m+3}|-2|X_{2m+5}|-\\|X_{2m+7}|-|Y_{2m+5}|-2|Y_{2m+7}|-|Y_{2m+9}|+|D_{2m-6}|\\+2|D_{2m-4}|+|D_{2m-2}|-|D_{2m+4}|-2|D_{2m+6}|-|D_{2m+8}|\end{pmatrix}+|Y'_{2m+1}|-|X'_{2m+1}|=0(m\geq 4) \quad (10)$$

$A_m=2|X_{2m-3}|+2|X_{2m-1}|-2|Y_{2m+5}|-2|Y_{2m+3}|+|D_{2m-2}|+|D_{2m}|-|D_{2m+2}|-|D_{2m+4}|$

$B_m=|X_{2m-7}|+2|X_{2m-5}|+|X_{2m-3}|+|Y_{2m-5}|+2|Y_{2m-3}|+|Y_{2m-1}|-|X_{2m+3}|-2|X_{2m+5}|-|X_{2m+7}|-|Y_{2m+5}|-2|Y_{2m+7}|-|Y_{2m+9}|+|D_{2m-6}|+2|D_{2m-4}|+|D_{2m-2}|-|D_{2m+4}|-2|D_{2m+6}|-|D_{2m+8}|$

$C_m=|Y'_{2m+1}|-|X'_{2m+1}|$

Its simplified representation is

$$\left(1-\frac{p}{2}\right)^2(|X_{2m+1}|-|Y_{2m+1}|)+\frac{p}{8}\left(1-\frac{p}{2}\right)A_m+\frac{p^2}{64}B_m+C_m=0 \quad (11)$$

$$\left(1-\frac{p}{2}\right)^2(|X_1|-|Y_1|)+\frac{p}{8}\left(1-\frac{p}{2}\right)A_0+\frac{p^2}{64}B_0+C_0=0, m=0 \quad (12)$$

Where $A_0=2|Y_1|-2|Y_5|+|D_2|-|D_4|$

$B_0=|X_1|+|X_3|+|Y_3|+|Y_5|-|X_5|-|X_7|-|Y_7|-|Y_9|+|D_2|+|D_4|-|D_6|-|D_8|$

$C_0=|Y'_1|-|X'_1|$

$$\left(1-\frac{p}{2}\right)^2(|X_3|-|Y_3|)+\frac{p}{8}\left(1-\frac{p}{2}\right)A_1+\frac{p^2}{64}B_1+C_1=0, m=1 \quad (13)$$

Where $A_1=2|X_1|+2|Y_1|-2|Y_5|-2|Y_7|+2|D_0|+|D_2|-|D_4|-|D_6|$

$B_1=2|X_1|+|X_3|+2|Y_1|+2|Y_3|+|Y_5|-|X_5|-2|X_7|-|X_9|-|Y_7|-2|Y_9|-|Y_{11}|+2|D_0|+2|D_2|-2|D_6|-2|D_8|-|D_{10}|$

$C_1=|Y'_3|-|X'_3|$

$$\left(1-\frac{p}{2}\right)^2\left(|X_5|-|Y_5|\right)+\frac{p}{8}\left(1-\frac{p}{2}\right)A_2 + \frac{p^2}{64}B_2+C_2=0, m=2 \quad (14)$$

Where $A_2=2|X_1|+2|X_3|-2|Y_7| -2|Y_9|+|D_2|+|D_4|-|D_6|-|D_8|$

$B_2=2|X_1|+4|Y_1|+2|Y_3|-|X_7| -2|X_9|-|X_{11}|-|Y_9|-2|Y_{11}|-|Y_{13}| +4|D_0|+2|D_2|-|D_8|-2|D_{10}|-|D_{12}|$

$C_2=|Y'_5|-|X'_5|$

$$\left(1-\frac{p}{2}\right)^2\left(|X_7|-|Y_7|\right)+\frac{p}{8}\left(1-\frac{p}{2}\right)A_3 + \frac{p^2}{64}B_3+C_3=0, m=3 \quad (15)$$

Where $A_3=2|X_3|+2|X_5|-2|Y_9|- 2|Y_{11}|+|D_4|+|D_6|-|D_8|-|D_{10}|$

$B_3=2|X_1|+|X_3|+2|Y_1|+2|Y_3|+|Y_5|- |X_9|-2|X_{11}|-|X_{13}|-|Y_{11}|-2|Y_{13}|-|Y_{15}| +2|D_0|+2|D_2|+|D_4|-|D_{10}|-2|D_{12}|-|D_{14}|$

$C_3=|Y'_7|-|X'_7|$

Using $E\{|X_{2m+1}|\}=E\{|Y_{2m+1}|\}$ can simplify

$$\frac{p}{8}\left(1-\frac{p}{2}\right)A_m+\frac{p^2}{64}B_m+C_m=0\,(m\ge 0) \quad (16)$$

For improve the estimate precision, the hypothesis should be modified to (17),so a more reliable method is (17),(18):

$$E\left\{\left|\bigcup_{m=i}^{j}X_{2m+1}\right|\right\}=E\left\{\left|\bigcup_{m=i}^{j}Y_{2m+1}\right|\right\} \quad (17)$$

$$\frac{p}{8}\left(1-\frac{p}{2}\right)\sum_{m=i}^{j}A_m+\frac{p^2}{64}\sum_{m=i}^{j}B_m+\sum_{m=i}^{j}C_m=0 \quad (18)$$

If it don't know the carrier image, the A_m and B_m are not known too. But (19) (20) can be calculated,

$$|X'_{2m+1}|=|X_{2m+1}|+\Delta|X| \quad (19)$$

$$|Y'_{2m+1}|=|Y_{2m+1}|-\Delta|Y| \quad (20)$$

$\frac{|X'_{2m+1}|+|Y'_{2m+1}|}{2}=\frac{|X_{2m+1}|+|Y_{2m+1}|}{2}+\frac{\Delta|X|-\Delta|Y|}{2}(m\ge 1)$,

$\frac{|X'_{2m+1}|+|Y'_{2m+1}|}{2}\approx|X_{2m+1}|\approx|Y_{2m}|$, $|D'_{2m}|=|D_{2m}|$.

$\frac{|X'_{2m+1}|+|Y'_{2m+1}|}{2}$ replace $|X_{2m+1}|$ 、 $|Y_{2m+1}|$, $|D'_{2m}|$ replace $|D_{2m}|$, so A_m , B_m and the embedded capacity p can be obtained.

IV. THE ANALYSIS OF ALGORITHM CAPABILITY

Fig.2 Experiment images

TABLE.1 EMBEDDING RATES CALCULATED WITH SPA

Embedding rate / Image	0	3.13	6.25	12.5	25	40.28
lena	0.85	3.81	7.26	13.83	26.79	42.24
mandrill	-1.78	2.70	6.86	13.46	29.89	51.32
zelda	0.06	3.76	8.01	15.65	29.94	45.74
boat	-0.72	2.22	5.46	12.42	23.06	40.68
goldhill	-1.07	1.85	5.33	11.66	25.04	43.35
cameraman	0.25	2.58	4.53	9.61	20.75	34.36
peppers	-0.67	2.05	5.17	12.15	26.60	42.73
jet plane	0.9	3.01	5.51	11.15	23.61	38.74
Average value	-0.27	2.75	6.02	12.49	25.71	42.39

For validate the analytic feasibility of SPA to HSBH, the experiment chooses eight images of 512×512, as shown in Fig.2. When r=4, embedded capability are 3.13%,25%, 12.5%, 25%, 40.28% separately. Using the SPA method to analyze HSBH, where i=0, j=25, the final result as shown in Table.1. Experimental results indicate that the algorithm can judge the capability of HSBH exactly.

V. CONCLUSIONS

SPA method is one of the most effective steganalysis for LSB steganography. It proves that SPA can detect HSBH algorithm in theory and experiment. Experimental results show that the correct rate of detection is nearly 100% when the embedding rate p>3%. The emphasis of research in future not only in SPA but also in presenting an improved HSBH algorithm against SPA.

VI. ACKNOWLEDGEMENTS

The work is supported by the Young Scholars plan project of Anhui province Education Department 2006jql208 and the Natural Science project of Anhui province Education Department of China under Grant Nos. KJ2008A18ZC. We thank SU Benyue,QIAN meng, WU Haifeng,LIAN Fei and JIANG Jinjian for their contributions to the work.

REFERENCES

[1] Dumitrescu Sorina, Wu Xiaolin, Wang Zhe. Detection of LSB steganography via sample pair analysis. IEEE Transactions on Signal Processing, 2003, 51(7):1995-2007.

[2] YU X Y, XU G S, ZHANG J.A new method of information high-bit hiding based on chaos[J]. Chinese Journal of Electron Devices, 2007,30(5):1677-1680.

[3] Chandramouli R, Memon N, Analysis of LSB Based Image Stenography Techniques [A]. Proceedings of ICIP 2001. Greece: Thessaloniki. 2001, Voloshynovskiy S, Pun T. Capacity-Security.

[4] Kim K T, Kim J H, Kim E S. Multiple Information Hiding Technique Using Random Sequence and Hadamard Matrix[J]. Optical Engineering, 2004, 40(11): 2489-2494.

[5] LIU Niansheng, GUO Donghui. Transmission Method of Image Information Hiding Based on Chaotic Encryption[J]. Computer Engineering,2006,10(4): 23-25. Scharinger J. “Fast encryption of image data using chaotic kolmogorov flows.” Electron Imageing, vol.7, no.2, pp.318-325, 1998.

[6] Fridrich J. “Symmetric ciphers based on two-dimensional chaotic maps.” Int J Bifurcat Chaos, vol.8, no.6, pp.1259-84, 1998.

[7] Li SJ, Zheng X Mou X, and Cai Y. “Chaotic encryption scheme for real-time digital video.” Proc SPIE on Electronic Imaging, San Jose, CA, USA, vol. 4666, 2002.

[8] Chen G R, Mao Y B and Chui C K. “A symmetric image encryption scheme based on 3D chaotic cat maps”. Chaos, Solitons and Fractals,. vol. 21, pp. 749-761, 2004.

[9] Lu J, Chen G. “A new chaotic attractor coined.” Int. J. of Bifurcation and Chaos, vol.12, no.3, pp.659-661, 2002.

[10] Dohmen, M., de Kraker, K.J. and Bronsvoort, W.F. (1996) Feature validation in a multiple- view modeling system. In: CD-ROM Proceedings of the 1996 ASME Computers in Engineering Conference, 19–22 August, Irvine, CA, USA, McCarthy, J.M. (ed), ASME, New York

[11] De Kraker, K.J., Dohmen, M. and Bronsvoort, W. F. (1997) Maintaining multiple views in feature modeling, In: Solid Modeling '97, Fourth Symposium on Solid Modeling and Applications, 14–16 May, Atlanta, GA, USA, Hoffmann, C.M. and Bronsvoort, W.F.

[12] Van Holland, W. and Bronsvoort, W.F. (2000) Assembly features in modeling and planning.Robotics and Computer Integrated Manufacturing 16(4): 277–294

[13] Noort, A. and Bronsvoort, W.F. (1999) Automatic model adjustment in form feature conversion. In: CD-ROM Proceedings of the 1999 ASME Design Engineering Technical Conferences, 12–16 September, Las Vegas, NV, USA, ASME, New York

[14] Scharinger J. “Fast encryption of image data using chaotic kolmogorov flows.” Electron Imageing, vol.7, no.2, pp.318-325, 1998.

[15] Fridrich J. “Symmetric ciphers based on two-dimensional chaotic maps.” Int J Bifurcat Chaos, vol.8, no.6, pp.1259-84, 1998.

[16] Li SJ, Zheng X Mou X, and Cai Y. “Chaotic encryption scheme for real-time digital video.” Proc SPIE on Electronic Imaging, San Jose, CA, USA, vol. 4666, 2002.

[17] Chen G R, Mao Y B and Chui C K. “A symmetric image encryption scheme based on 3D chaotic cat maps”. Chaos, Solitons and Fractals,. vol. 21, pp. 749-761, 2004.

[18] Lu J, Chen G. “A new chaotic attractor coined.” Int. J. of Bifurcation and Chaos, vol.12, no.3, pp.659-661, 2002.

[19] Yoon S, Salomon B, Gayle R. Quick-VDR: Out-of-core view-dependent rendering of gigantic models. IEEE Transactions on Visualization and Computer Graphics, 2005, 11(4):369-382.

[20] Ulrich T. Rendering massive terrains using chunked level of detail control // Proceedings of SIGGRAPH2002. San Antonio, Texas USA: ACM Press, 2002.

[21] Levenberg J. Fast view-dependent level-of-detail rendering using cached geometry // Gross M, Joy KI, Moorhead RJ, eds. Proceedings of the IEEE Visualization, 2002. Los Alamitos: IEEE Computer Society Press, 2002: 259−266.

[22] Cignoni P, Ganovelli F, Gobbetti E. BDAM-batched dynamic adaptive meshes for high performance terrain visualization. Computer Graphics Forum, 2003, 22(3):505-514.

[23] Losasso F, Hoppe H. Geometry clipmaps: terrain rendering using nested regular grids // Proceedings of SIGGRAPH 2004. Los Angeles : ACM Press, 2004: 769-786.

[24] Schneider J, Westermann R. GPU-friendly high-quality terrain rendering. Journal of WSCG, 2006, 14(1-3): 49-56.

[25] Yotam L, Zvi K, Jihad E S. Seamless patches for GPU-based terrain rendering // Proceedings of WSCG 2007. Plzen Czech Republic, 2007: 201-208.

[26] Jordan G. Adaptive smoothing of valleys in DEMs using TIN interpolation from ridgeline elevations: An application to morphotectonic aspect analysis. Computers & Geosciences, 2007, 33(4): 573-585.

[27] Yotam L, Zvi K, Jihad E S. Seamless patches for GPU-based terrain rendering // Proceedings of WSCG 2007. Plzen Czech Republic, 2007: 201-208.

Yuanzhi Wang

was born in 1977. He is an Associate Professor of School of Computer and Information at Anqing Normal College,Anqing,Anhui,P.R.China.He received his M.S.'s degree from Computer & Control College at Harbin University of Science and Technology in 2005. His main research interests: CAGD,computer graphics,image processing,video processing, , knowledge management, intelligent agent,pattern recognition and machine learning.