An Adaptive Threshold based FPGA Implementation for Object and Face Detection

Sateesh Kumar H.C.[1], Sayantam Sarkar[2], Satish S Bhairannawar[3], Raja K.B.[4] and Venugopal K.R.[5]
[1]Department of Electronics and Communication Engineering, Sai Vidya Institute of Technology, Bangalore, India.
[2]Department of Electronics and Communication Engineering, Vijaya Vittala Institute of Technology, Bangalore, India.
[3]Department of Electronics and Communication Engineering, Dayananda Sagar College of Engineering, Bangalore, India.
[4]Department of Electronics and Communication Engineering, University Visvesvaraya College of Engineering, Bangalore, India.
[5]Principal, University Visvesvaraya College of Engineering, Bangalore, India.
[1]hcsateesh@yahoo.com, [2]sayantam.61@gmail.com, [3]satishbhairannawar@gmail.com

***Abstract*— The moving object and face detection are vital requirement for real time security applications. In this paper, we propose an Adaptive Threshold based FPGA Implementation for Object and Face detection. The input Images and reference Images are preprocessed using Gaussian Filter to smoothen the high frequency components. The 2D-DWT is applied on Gaussian filter outputs and only LL bands are considered for further processing. The modified background with adaptive threshold are used to detect the object with LL band of reference image. The detected object is passed through Gaussian filter to enhance the quality of object. The matching unit is designed to recognize face from standard face database images. It is observed that the performance parameters such as percentage TSR and hardware utilizations are better compared to existing techniques.**

Keywords*— *Discrete Wavelet Transform; Gaussian Filter; Adaptive Threshold; Object Detection and Face Recognition.

I. INTRODUCTION

The biometrics are classified based on their physical and behavioural characteristic parameters. The physical characteristic traits are fingerprint, iris, palm print, DNA etc., of a person and are constant throughout life span. The recognition using physiological traits are easy and require less number of samples to build high speed real time biometric system efficiently with less complexity. The recognition using behavioural traits are not very accurate and require more number of samples to build real time biometric system. The behavioural biometric traits are signature, voice, keystroke, gait etc., and are time variant parameters.

The general biometric system has three sections to recognize a person viz., pre-processing, feature extraction and matching section. In pre-processing section, the images are resized to convert different sizes into uniform size, colour images are converted into gray scale images to reduce computation complexity and noises in the images are removed to enhance the quality of the images. The features like mean, variance, standard deviation and principal component analysis are extracted in spatial domain by directly manipulating enhanced image. The transform domain features are extracted from Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), Dual Tree Complex Wavelet Transform (DTCWT) etc. The Euclidian Distance (ED), Hamming Distance, Neural Network, Support Vector Machine (SVM) etc., are used in matching section to compute similarities and distance and differences among images.

In this paper we propose an Adaptive Threshold based FPGA Implementation of Object and Face Detection. The preprocessing block, modified background subtraction and adaptive threshold techniques are used to detect moving object and recognize face effectively. One of the major advantages of proposed technique is the adaptive threshold approach is used to compute variable reference values for different object and face image of different persons.

The performance parameters are improved since adaptive threshold, modified background subtraction and filters used in the architecture.

Contribution: *The contribution and novel aspects of the proposed techniques are listed as follows*

(i) The threshold values are computed adaptively based on characteristics of the images.

(ii) The modified background subtraction technique is used on filtered LL coefficients of background image/face database and actual image/test face image to compute absolute difference between LL coefficients of two sets of images. The absolute difference is compared with adaptive threshold values to detect object.

(iii) The object detection architecture is extended to face detection by using global threshold and matching unit blocks.

Organization: *Section 2 briefly reviews literature of recent research papers. Section 3 presents proposed architecture. Hardware implementation is given in Section 4. Section 5 explains Performance analysis of proposed object and face detection architecture. The conclusion is given in section 6.*

II. LITERATURE SURVEYS

Andra et al., [1] proposed four-processor architecture for 2D-DWT which is used for block based implementation for 2D-DWT, which requires large memory. Liao et al., [2]

978-1-5090-0148-4/15/$31.00© 2015 IEEE

proposed 2D-DWT dual scan architecture, which requires two lines of data samples simultaneously for forward 2D-DWT and also proposed another 2D-DWT architecture, which accomplished decomposition of all stages resulting in inefficient hardware utilization and more sophisticated control circuitry. Barua et al., [3] proposed folded based architecture for lifting based bi-orthogonal CDF-9/7 2D-DWT. The architecture uses hybrid level at each stage which reduces the controller complexity, but increases the hardware utilization.

A number of defense, security and commercial applications demand real time face recognition system [4]. The algorithm proposed by Fei Wang et al., [5] uses spectral features for face recognition which is a nonlinear method for face feature extraction. The algorithm can detect nonlinear structure present in the face image and this structure is then used for recognition purpose. Ben Niu et al., [6] proposed two dimensional Laplacian face method for face detection. The algorithm based on locality preserved embedding and image based projection techniques which preserve geometric structure locality of sample space to extract face feature. The performance of the algorithm is checked by using FERET and AR database. Sajid et al., [7] proposed FPGA implementation of face recognition system. The system uses Eigen values for recognition purpose. To eliminate floating point Eigen value software-hardware co-design is used with partial re-configurability. Guo and Lu [8] proposed face detection based on Haar classifier which is implemented on FPGA. The uses of pipelined architecture decrease the recognition time. Chen and Lin [9] proposed minimal facial based face detection algorithm. The algorithm first check for skin and hair colored region and then it decide the face area. Veeramanikundan et al., [10] proposed FPGA implementation of face detection and recognition system under light intensity variation. The proposed algorithm based on the Adapting Boosting Machine learning algorithm with Haar transform which increase the accuracy of face recognition.

III. Proposed Architecture

The architectures for Moving Object Detection and Face Detection are shown in Fig. 1 and 2 respectively. The Gaussian filters [11] are used to remove high frequency edges and some amounts of light variations present in the images. The filtered images are then fed to 2D-DWT block [12], where only LL band is considered because most of the valuable information of an image is available. The LL1 is the approximation band coefficients of image_in/test image and LL2 is the approximation band coefficients of image_ref/database. The adaptive threshold block is used to compute the threshold by Modified Background Subtraction simultaneously along with 2D-DWT block. The Modified Background Subtraction block is used to remove the background from LL band of actual image and then image is fed to Gaussian filter to remove any small amounts of light intensity variations present in the image.

The Moving Object Detection block diagram is also used to detect face images by using Matching Unit as shown in Fig. 2. The ORL (Olivetti Research Laboratory) face database [13] is used to test the performance of the proposed face detection architecture.

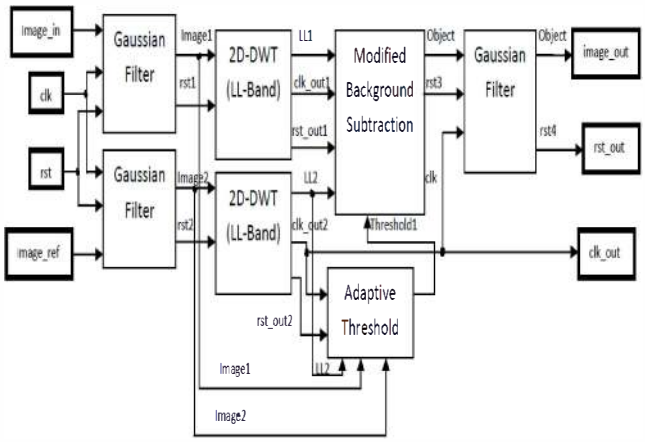

Fig. 1. Proposed Architecture for Moving Object Detection

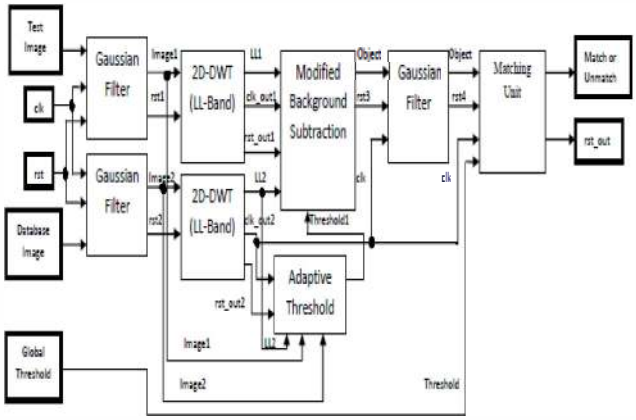

Fig. 2. Proposed Architecture for Face Detection

IV. Hardware Implementation

A. Adaptive Threshold

The adaptive threshold block is used to calculate the threshold which is used to detect object/image efficiently. If the input pixel value is greater than the threshold value then the respective values are passed to the output else the output will be zero. The pixel value difference (S) between background and actual image is given in equation (1) and (2)

$$S = \frac{[A_1-B_1]^2+[A_2-B_2]^2+\cdots+[A_N-B_N]^2}{8N} \quad (1)$$

$$S = \sum_{i=1}^{N} \frac{[A_i-B_i]^2}{8N} \quad (2)$$

Where, N is the total dimension of input image (N=256x256).

A_1, $A_2 \ldots A_N$ are the actual image pixel intensity values after filtering.

B_1, $B_2 \ldots B_N$ are the background image pixel intensity values after filtering.

The Adaptive Threshold (AT_i) is calculated using Pixel Value Difference (S) and LL coefficient values of background image (LL2) are given in equation (3) and (4).

$$AT_i = S + (LL\ Coefficients\ of\ Background\ Image)_i \quad (3)$$

$$AT_i = S + (LL2)_i \quad (4)$$

Where, i=1 to (128x128).

The hardware architecture is used to build adaptive threshold block is given in Fig. 3 where the image size is N=256x256. The presence of feedback in Addition Block in the architecture, the output value changes at every clock pulse, hence D_ff and counter is used to obtain S value for an image at 65536 clock pulse. The right shift by 19 (i.e. *>>19)* is used to implement 8x256x256. The LL coefficients of background image/database face images are added with the value of S to compute final threshold values of each LL coefficients. The threshold value is adaptive since each LL coefficient has different threshold values.

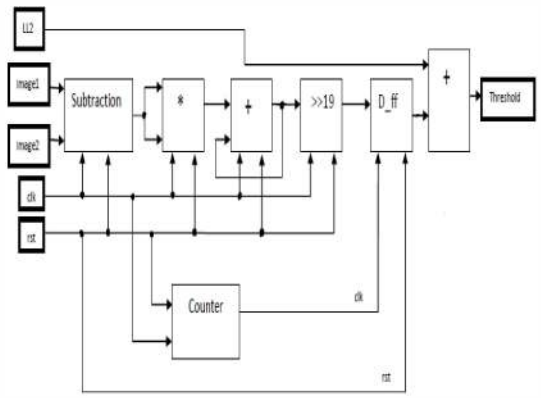

Fig. 3. Hardware Structure for Adaptive Threshold

B. Modified Background Subtraction

In Object Detection, the background information is removed to obtain foreground information. In Background subtraction [14], the LL band coefficients of two images obtained from 2D-DWT block are subtracted and compared with proposed Adaptive Threshold (AT_i) values to obtain segmented foreground image. The background image LL2 coefficients are subtracted and considered absolute values as given in equation (5) and (6).

$$LL_j = mod\left[(LL2)_j - (LL1)_j\right] \quad (5)$$

Where, j=1 to (128x128).

LL2= LL band coefficients of Background Image/Database.

LL1= LL band coefficients of Actual Image/Test.

$$Background\ Subtraction = \begin{cases} LL_j & ; \quad if\ LL_j \geq AT_i \\ 0 & ; \quad Otherwise \end{cases} \quad (6)$$

The hardware architecture of Modified Background Subtraction block is shown in Fig. 4. The two images (i.e. background and foreground) are fed to Max. and Min. Calculations block to find minimum and maximum values of two LL band coefficients. Then use subtraction block to get subtracted values of both LL coefficient values. Now this value is compared with adaptive threshold value. If the adaptive threshold value is less than subtracted LL coefficient value then send the subtracted value (object) to the output else send zero to the output.

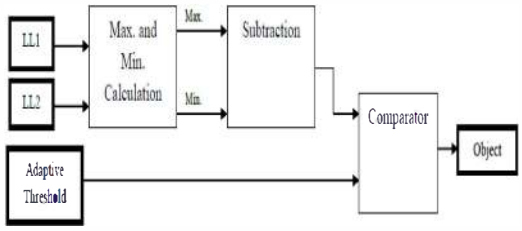

Fig. 4. Hardware Architecture of Background Subtraction

C. Matching Unit

The similarity score between test and database face images are computed using Matching Unit and the block diagram is shown in Fig. 4. The Background Subtraction block subtract two images and as a result similar portions of two images are cancelled at the output. But sometime due to large amount of light intensity variations similar portions of both images produce some small pixel intensity value after background subtraction. To eliminate this problem, a small value of 10 is fixed as tolerance threshold. If input to this block is in between 0 to 10 then the counter will be incremented by amount of one else the counter value stays on the previous value. Now this counter value is compared with the database global threshold value. If the counter value is greater than global threshold value, then person is matched else not matched.

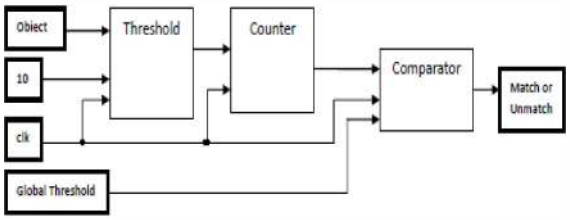

Fig. 4. Hardware Architecture of Matching Unit

Pseudo code for Matching Unit is given below.

```
initilize counter value = 0;
if(rising edge clock)
    if(input ≤ 10)
        increment counter value by 1.
    else
        previous counter value.
    end if;
    if(counter value ≥ global threshold value)
        the person is matched.
    else
        the person is not matched.
    end if;
end if;
```

V. Performance Analysis

In this section, the performance of proposed method is evaluated using various hardware parameters such as slice registers and slice flip-flops and software parameter as percentage TSR.

A. Proposed Object Detection Architecture

The proposed object detection architecture to detect an object from an image effectively as adaptive threshold concept is used. The object detected using proposed architecture is compared with existing architectures presented by Lee and Park [15] and Chu et al., [16] are shown in the Fig. 5. It is observed that, the minute part in the object is also detected effectively in the proposed architecture compare to existing architectures.

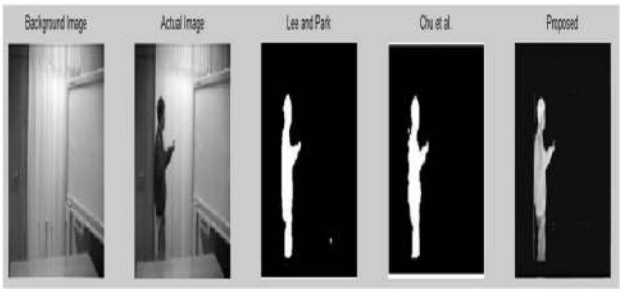

Fig. 5. Image Comparison for Existing and Proposed Moving Object Detection

The hardware requirements of proposed and existing object detection architecture are given in Table I. The architecture presented by Chodwary et al., [17] requires 1180 slice registers and 2118 fully used flip-flop pairs. This architecture is implemented by using Microblaze processor and the scripting is done by using C language. The architecture presented by Mahamuni and Patil [18] requires 961 slice registers and 339 fully used flip-flop pairs. This architecture is implemented by using inbuilt system generator HDL blocksets. Similarly the architecture presented by Susrutha Babu et al., [19] uses 409 slice registers and 269 fully used flip-flop pairs. The proposed object detection architecture requires only 365 slice registers and 98 fully used flip-flop pairs. Hence the proposed architecture is better compared to existing architecture.

TABLE I. HARDWARE COMPARISON OF PROPOSED AND EXISTING OBJECT DETECTION ARCHITECTURES

Logic Utilization	Chodwary et al., [17]	Mahamuni and Patil [18]	Susrutha Babu et al., [19]	Proposed Method
Number of Slice Registers	1180	961	409	365
Number of Fully Used FF-Pairs	2118	339	269	98

B. Proposed Face Detection Architecture

1) Performance Coparisons of Proposed Technique with Existing Techniques

The percentage values of TSR of proposed method is compared with existing methods presented by Ben Niu et al., [6], Sardar and Babu [20], Junjie Yan et al., [21], Thiago H.H. Zavaschi et al., [22], Manchula and Arumugam [23] and Wang and Yin [24]. It is observed that the value of %TSR is high in the case of proposed method compared to existing methods as shown in Table II. The performance parameters are improved since; the proposed method uses adaptive threshold and background subtraction techniques.

TABLE II. COMPARISON OF PERCENTAGE TSR VALUES OF PROPOSED METHOD WITH EXISTING METHODS

Authors	Techniques	% TSR
Ben Niu et al., [6]	Two Dimensional Laplacian Face	97.8
Sardar and Babu [20]	DWT, PCA and Neural Networks	93
Junjie Yan et al., [21]	Hierarchical Part	96.5
Thiago H.H. Zavaschi et al., [22]	Gabor Filter and Local Binary Pattern	96.2
Manchula and Arumugam [23]	PCA and Eigen Vector	96
Wang and Yin [24]	Topographic Context	82.68
Proposed Method	DWT, Background Subtraction and Adaptive Threshold	99

2) Performance Analysis using Hardware Comparison Results

The logic utilizations comparison of proposed method with existing method is given in Table III. The existing technique presented by Sardar and Babu [20] uses 39432 slice registers and the maximum operating frequency is 80MHz. i.e., this architecture uses a large amount of hardware and low operating frequency because the architecture is implemented on Microblaze embedded processor. The proposed architecture uses only 861 slice register and maximum operating frequency is 118.60 MHz i.e., low amount of hardware and high operating frequency because the architecture uses only adders,

shifters and D-flip flop's for implementation and also the adaptive threshold concept.

TABLE III. HARDWARE COMPARISON OF PROPOSED AND EXISTING FACE DETECTION TECHNIQUES

Logic Utilizations	Sardar and Babu [20]	Proposed method
Number of Slice Registers	39432	861
Maximum Operating Frequency (MHz)	80	118.60

VI. CONCLUSION

The object and face detection is essential to detect a person in real time applications. In this paper, we propose an adaptive threshold based object and face detection. The input and reference images are passed through preprocessing filters to remove noise and to generate LL band coefficients. The modified background subtraction is used with adaptive threshold on two LL bands to detect an object. The output of modified background with adaptive threshold is passed through Gaussian filter to obtain smooth image. The face detection can also be performed with the same architecture by additional matching and global threshold unit. The performance parameters are better in the proposed architecture compared to existing architectures in-terms of software and hardware results. In future the proposed architecture can be extended for real time video processing.

References

[1] Kishore Andra, Chaitali Chakrabarti and Tinku Acharya, "A VLSI Architecture for Lifting-Based Forward and Inverse Wavelet Transform", *IEEE Transaction on Signal Processing,* Vol. 50, No. 4, pp. 966–977, April 2002.

[2] H. Liao, M. K. Mandal and B. F. Cockburn, "Efficient Architectures for 1-D and 2-D Lifting-Based Wavelet Transforms", *IEEE Transaction on Signal Processing*, Vol. 52, No. 5, pp. 1315–1326, May 2004.

[3] S. Barua, J. E. Carletta, K. A. Kotteri and A. E. Bell, "An Efficient Architecture for Lifting-Based Two-Dimensional Discrete Wavelet Transform", *Integration VLSI Journal*, Elsevier, Vol. 38, No. 3, pp. 341–352, 2005.

[4] Zhao W., R. Chellappa, P.J. Phillips and A. Rosenfield, "Face Recognition: A Literature Survey", *ACM Computing Surveys*, pp. 399-485, Vol. 35, No. 4, December 2003.

[5] Fei Wang, Jingdong Wang, Changshui Zhang and James Kwok, "Face Recognition using Spectral Features", *International Journal of Pattern Recognition*, Elsevier, Vol. 40, pp. 2786-2797, 2007.

[6] Ben Niu, Qiang Yang, Simon Chi Keung Shiu and Sankar Kumar Pal, "Two Dimensional Laplacian Face Method for Face Recognition", *International Journal of Pattern Recognition*, Elsevier, Vol. 41, pp. 3237-3243, 2008.

[7] I. Sajid, M.M. Ahmed, I. Taj, M. Humayun and F. Hameed, "Design of High Performance FPGA Based Face Recognition System", *Progress in Electromagnetic Research Symposium Proceeding*, pp. 504-510, July 2008, USA.

[8] Changjan Gao and Shih-Lien Lu, "Novel FPGA Based Haar Classifier Face Detection Algorithm Acceleration", *IEEE International Conference on Field Programmable Logic and Applications*, pp. 373-378, 2008, Germany.

[9] Yao-Jiunn Chen and Yen-Chun Lin, "Simple Face Detection Algorithm Based on Minimum Facial Features", *IEEE Annual Conference of the Industrial Electronics Society*, pp.455-460, November 2007, Japan.

[10] K. Veeramanikandan, R. Ezhilarasi and R. Brindha, "An FPGA Based Real Time Face Detection and Recognition System Across Illumination", *International Journal of Engineering Science and Engineering*, Vol. 1, Issue. 5, pp. 66-68, March 2013.

[11] R. A. Haddad and A. N. Akansu, "A Class of Fast Gaussian Binomial Filters for Speech and Image Processing", *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 39, pp. 723-727, 1991.

[12] Satish S Bhairannawar, Sayantam Sarkar, Raja K B and Venugopal K R, "An Efficient VLSI Architecture for Fingerprint Recognition using O2D-DWT Architecture and Modified CORDIC-FFT", *IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems*, pp. 193-197, February 2015, India.

[13] [Online]http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.

[14] Y. Benezeth, P.M. Jodoin, E. Emile, H. Laurent and C. Rosenberger, "Review and Evaluation of Commonly-Implemented Background Subtraction Algorithms", *19th IEEE International Conference on Pattern Recognition,* pp. 1-4, December 2008, Florida.

[15] Jeisung Lee and Mignon Park, "An Adaptive Background Subtraction Method Based on Kernel Density Estimation", *International Journal of Sensors*, Vol. 12, pp. 12279-12300, 2012.

[16] Chiu C.C., Ku M.Y. and Lian, L.W., "A Robust Object Segmentation System using a Probability-Based Background Extraction Algorithm", *IEEE Transaction on Circuit and Systems*, Vol. 20, pp. 518–528, 2010.

[17] M. Kalpana Chowdary, S. Suprashya Babu, S. Susrutha Babu and Habibulla Khan, "FPGA Implementation of Moving Object Detection in Frames by using Background Subtraction Algorithm", *IEEE International Conference on Communication and Signal Processing*, pp. 1032-1036, April, 2013, India.

[18] P. D. Mahamuni and R. P. Patil, "FPGA Implementation of Background Subtraction Algorithm for Image Processing", *IOSR Journal of Electrical and Electronics Engineering*, Vol. 9, Vo. 5, pp. 69-78, 2014.

[19] S. Susrytha Babu, S. Suparshya Babu, Habibulla Khan and M. Kalpana Chowdary, "Implementation of Running Average Background Subtraction Algorithm in FPGA for Image Processing Applications", *International Journal of Computer Applications*, Vol. 73, No. 21, pp. 41-46, 2013.

[20] Santu Sardar and K. Ananda Babu, "Hardware Implementation of Real-Time, High-Performance, RCE-NN Based Face Recognition System", *IEEE International Conference on VLSI Design and Embedded Systems*, pp. 174-179, 2014, India.

[21] Junjie Yan, Xuzong Zhang, Zhen Lei and Stan Z. Li, "Face Detection by Structural Models", *International Journal of Image and Vision Computing*, Elsevier, Vol. 32, Issue. 10, pp. 790-799, December 2013.

[22] Thiago H.H. Zavaschi, Alceu S. Britto Jr., Luiz E.S. Oliveria and Alessandro L. Koerich, "Fusion of Feature Sets and Classifiers for Facial Expression Recognition", *International Journal of Expert Systems with Applications*, Elsevier, Vol. 40, pp. 646-655, 2013.

[23] A. Manchula and S. Arumugam, "Multi-modal Facial Recognition Based on Improved Principle Component Analysis with Eigen Vector Feature Selection", *Middle-East Journal of Scientific Research*, Vol. 22, Issue. 8, pp. 1203-1211, 2014.

[24] Jun Wang and Lijun Yin, "Static Topographic Modeling for Facial Expression Recognition and Analysis", *International Journal of Computer Vision and Image Understanding*, Elsevier, Vol. 108, pp. 19-34, 2007.