2015 IEEE 2015 International Conference on Computer, Communication, and Control Technology (I4CT 2015), April 21 - 23 in Imperial Kuching Hotel, Kuching, Sarawak, Malaysia

An Evaluation of Technical Study and Performance for Real-Time Face Detection Using Web Real-Time Communication

Manop Phankokkruad*, Phichaya Jaturawat
Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand 10520
Email: manop@it.kmitl.ac.th

***Abstract*—This paper proposed the technical study of real-time face detection system and also cover a key technology. In order to find the most appropriate factors, techniques, and algorithm by evaluating the performance that included connection speed, and effectiveness of tracking and detecting a human face in various conditions. WebRTC worked perform in any condition of device and platform independence. Furthermore, WebRTC could operate securely on the HTTPS protocol. In the case of the transferring ability of images, it relied on the connection speed; that LAN and WiFi are the most prefer for this best image quality. The results shown that Haar-like feature and CLM have significantly detected the face area over the web browser in almost light conditions. However, Haar-like has the better precision when operated on the low-speed processor. The real-time face detection system could be identified the person who walked through the capture device. In addition, this proposed system could be applied to any scope of the personal identification system.**

***Index Terms*—Face Detection, WebRTC, Real-Time Face Detection, Algorithms, Internet Protocol**

I. Introduction

Face detection is the most successful method of image analysis, which has received significant attention. It is an extremely reliable method of biometric personal identification based on the analysis of frontal face images that was in the personal identification system [1]. Face detection is a significant step in automatic face recognition system. A powerful and applicable system for human face detection will have great effect on the human identification system. Furthermore, face detection becomes an research interest area which cover the numerous disciplines such as digital image processing, pattern recognition, computer vision, neural networks, and cognitive science [2], etc. Generally, a human face detection system includes the face detection, feature extraction, face database and face recognition. In addition, the face detection can use on the video [3], gray scale image, color image [4]. The task of face detection is to detect the presence of any human face appearing in the image. For past decade, literature reveals that a lot of research work has been carried out on face detection. The problem of using face detection application is that it has many disadvantages related to the appropriate platform, installation method and how to configuration. Since face detection application assumes a fixed architecture for the platform and device dependence, an inappropriate platform can make it have a lot of limitations in such an interoperability, rapid application development, and adaptive to network conditions, etc. For this reason, Web Real-Time Communication (WebRTC) is a solution for this problem. WebRTC is a collection of standards, protocols, and JavaScript APIs, the combination of which enables peer-to-peer audio, video, and data sharing between browsers [5]. WebRTC allows web browsers to communicate with each other through a variety of real-time without any installation, configuration, platform and device independence. Now the problem becomes much harder since any face detection algorithms cannot perform well enough in carry on the web browser and HTML5 technology. Hence, the effective and faster algorithms have to study to overcome this restrict. The applications of face detection has recently received significant attention. There have been many attempts of varying success of the face detection problem, one of the most promising being the Viola-Jones [6], AdaBoost, constrained local model (CLM) [7] and Haar-like feature. Recent face detection algorithm worked by Yang et al [4], they developed verification module to identify the human faces by analysing of shapes and objects . They precisely localize human faces by placing of eyes and mouths position based on mathematics model transform. A valuable reference on real-time face detection algorithm is the recently proposed by Lang and Gu [8], they studied the face detection system based on AdaBoost algorithm. the result shown that, the detection rate of the algorithm is higher, and the detection velocity is also better. In the same way, CLMs were proposed for non-rigid face tracking by Cristinacce and Cootes [7].

Consequently, the purpose of this paper is to find out the most appropriate key technology that enables the reliable face detection from a video over the web browser by studying the essential performance factors. This key factor is the reason for selecting the best algorithm for the real-time face detection system. In order to prove our inference, we will conduct the experiment for system evaluations to demonstrate the most appropriate technology which support desire conditions for the future research. In addition, the significant difference between

978-1-4799-7952-3/15/$31.00 ©2015 IEEE

the face detection algorithms will be discussed in this paper.

The remainder of this paper is organized as follows. First, this paper gives an overview of the face detection problems and related works. In Section 2 provides the problem definition, key technology, face detection algorithms and real-time face detection system architecture. Section 3 describes about the implementation and also its different environments. Moreover, the results of the experimentations that were conducted on an evaluation system and the discussion were described in Section 4. Finally, the summary, conclusions and the future work are presented in Section 5.

II. Algorithms and System Architecture

In this section, the problem definition, key technology, real-time face detection system architecture, and related theoretical of face detection algorithms are described.

A. Problem definition

The problem definition phase is the first phase for determining the existing problem. During this phase, the objectives of using real-time face detection on the desired condition are identified. Likewise, the overall process is designed to define the function and relationship of the system. Usually, the face detection systems are developed and installed as the application on a kind of platform in such a Microsoft Windows. Thus, the system was designed for the specific application platform and device dependence. In additional, the installation processes are supervised by the technical specialist. In order to eliminate this limitation, a new technique and methods to develop the real-time face detection system need to be studied in the way to make the face detection system can operate on any platform and device dependence. Moreover, the new techniques should have more security, reliability and performance. Since, it is well recognized that the performance of a face detection and recognition system depends on the quality of images participating in the face comparison processes [9], [10], the capability of controlling the reference image quality takes advantage to improve face detection and the storage capacity is not excess. The main challenge when using the real-time face detection for tracking is that the person is unreliable moving. Since, the detection output compose of false positives and missed detections, the result of association problem between detections and targets is difficult.

B. WebRTC

WebRTC is an open framework which offers web application able to write the real-time multimedia application on the web without requiring any extensions since this API is inherently supported by a lot of new versions of common browsers [11]. It enables peer-to-peer audio, video, and data sharing between browsers [5]. WebRTC allows web browsers to communicate with each other through a regular standard protocols. Also, WebRTC efforts to put real-time communications capabilities into all web browsers and also make these capabilities accessible through the standard HTML5 and

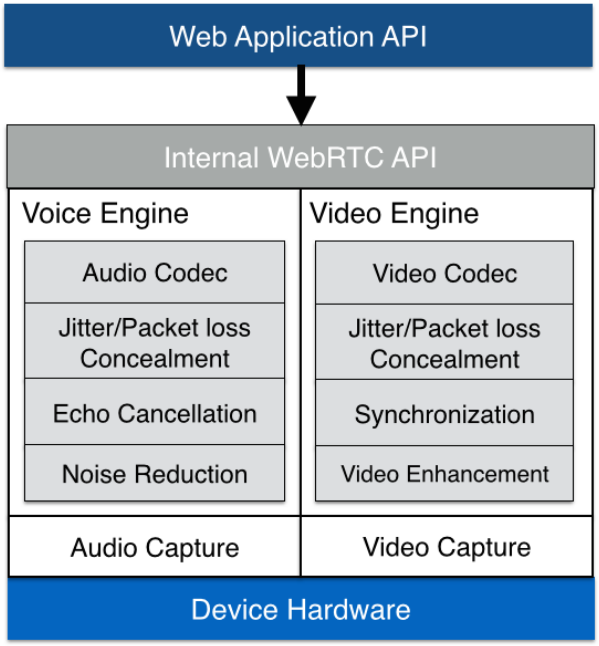


Fig. 1. WebRTC Architecture

JavaScript APIs [5]. For this reason, this is a turning point of an evolution in the web applications technology [12]. An architecture of WebRTC is shown in Figure 1. The transportation of information between the web browser and the web server is provided by the Hyper-Text Transport Protocol (HTTP), which operates over the generally WebSocket protocol.

A WebRTC web application interacts with web browsers through the standardized WebRTC API, also allowing it to properly aptitude and control the real-time browser function as depicted in Figure 1. The WebRTC APIs provide a wide set of functions, such as connection management, encoding and decoding capabilities, media control, and firewall, etc. The design of the WebRTC APIs represent a continuous, and real-time flow of data stream across the network in order to allow direct communication between two browsers. This distinctly represents a revolutionary to web-based communication. In this paper, a real-time video data was captured into a still image, then it was sent across the Internet to a remote server.

C. Face Detection Algorithm

Nowadays, many face detection tasks perform by the computer vision algorithms. A modern computer vision algorithms is a core concept of core computer vision (CCV) development. CCV implement the very fast detection algorithm, and an accurate object detection algorithm for any difficult objects. This paper considers the two powerful of CCV algorithm include Haar-like features and constrained local model(CLM).

1) Constrained Local Model: CLM is a class of locating methods that places the sets of points on a target image, which is constrained by a statistical shape model. Typically, the approach is the sampling area from the image space around the current estimate, then it is projected into a reference frame.

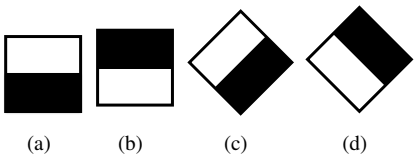


Fig. 2. HAAR-like features of edge

Subsequently, each point is generated a response image, and give a cost at each pixel. The standard CLM framework can be illustrated by the model $M = \{S, D\}$ where S is the shape model and D is a set of linear classifiers for detection of n parts of the face. It can be described as a group of linear classifiers $D = \{w_i, b_i\}_{i=1}^{n}$, where w_i and b_i is the linear detector for the i^{th} part of the face. The probabilistic response map $p(l_i = 1|x, \mathcal{I}))$ for the i^{th} landmark being correctly aligned $(l_i = 1)$ at location x of image $\mathcal{I}$ is modeled by a logistic function.

$$p(l_i = 1|x, \mathcal{I})) = \frac{1}{1 + e^{\{d(w_i^T h(x;\mathcal{I}) + b_i) + c\}}} \quad (1)$$

where d and c are calculated over a cross-validation procedure. h is a feature extraction function, and $h(x; \mathcal{I})$ denote the feature extracted from the patch in image $\mathcal{I}$ centered at x_i [13]. CLM is a specific model, that the response images were created by applying the normalised correlation to fit the current face [7]. It comes to meaning of any method that a set of local models are used to evoke the response images, afterward the shape model is used to find out the best combination of the response image [14]. This approach has been proposed by many research such as Saragih et al [15].

2) Haar-like features: Haar-like features [16] is represented to be the most successful technique for using in object detection algorithm. Due to, it provide an attractive trade-off between accuracy and evaluation speed. In an intuitive way, Viola and Jones [6] proposed the first object detection framework based on Haar-like and it becomes a popular method to provides competitive object detection rates in real-time. Afterward, there are many successful systems based on Haar-like features have been proposed [17]. Haar-like features are the set of two-dimensional Haar functions, that can be used to encode the local appearance of objects [16]. These features use the intensity values, and the change in contrast values between adjacent rectangular groups of pixels. Further, the contrasts values are used to evaluate the relative of light and dark areas. The Haar-like features templates that shows in Figure 2 and 3 are used to detect an image. Haar features can easily be scaled the size of the pixel group being inquired. This allows features to be used to detect objects of various sizes [18].

D. Real-Time Face Detection System Architecture

The proposed system comprises of two major parts include the server-side, and the client-side as shown in Figure 4. In the server-side, it consists of the web server, face detection routine, and image database. Likewise, the client-side is an

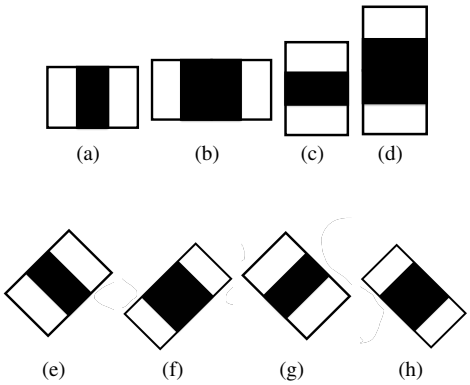


Fig. 3. HAAR-like features of line

web application that consists of image capture device, RTC API, and the fast face detection routine. Focus in the client-side, it was developed as web application, which can operate on any web browsers. An important device in the client-side is the image capture device; it is general WebCam that was built in. Usually, RTC API is built in a modern web browser such as Firefox, Google Chrome, Safari, etc. In that way, the face detection routine is a web application module that is developed based on the face detection algorithms by JavaScript and HTLM5 technology. Face detection algorithms are the key of face segmentation and recognition the human face in the test image. This face detection routine can find an image for faces, and give rectangles of where the faces are found. In the case of server-side, it was set up as the web server with Hypertext Transfer Protocol Secure (HTTPS) connection. On this server, it contains the face detection routine like a routine on the client-side but there is something different such as image quality adjusting and processing. Finally, the processed image will be stored by making the index in the image database.

III. Implementation

In order to implement and evaluate the proposed system, many varieties of conditions were set up. The system was conducted in the real-time face detection environment. The system was separated into two major parts; the client-side and the server-side. In the client-side comprises of the different platform on each client that is Ubuntu Linux desktop, Windows 8 desktop, Windows 8.1 Tablet PC, and Apple MacBook. Likewise, the server-side, it operates the Window Server 2012, Internet Information Server (IIS) with SSL/TLS, and PHP web server. It is well to recognize that face recognition performance is impacted by the image quality. This study uses a range of images size that was obtained by controlled changes of the WebRTC APIs video sequences from the WebCam during the image acquisition step. The quality of captured image was varied in three sizes as follows; QVGA (320x240), VGA(640x480), and 1024x768 pixels. Focus on the connection speed, this work uses various networks to connects the client to the image server such as gigabit LAN, 150 mbps WiFi and 3G cellular network. The variety of

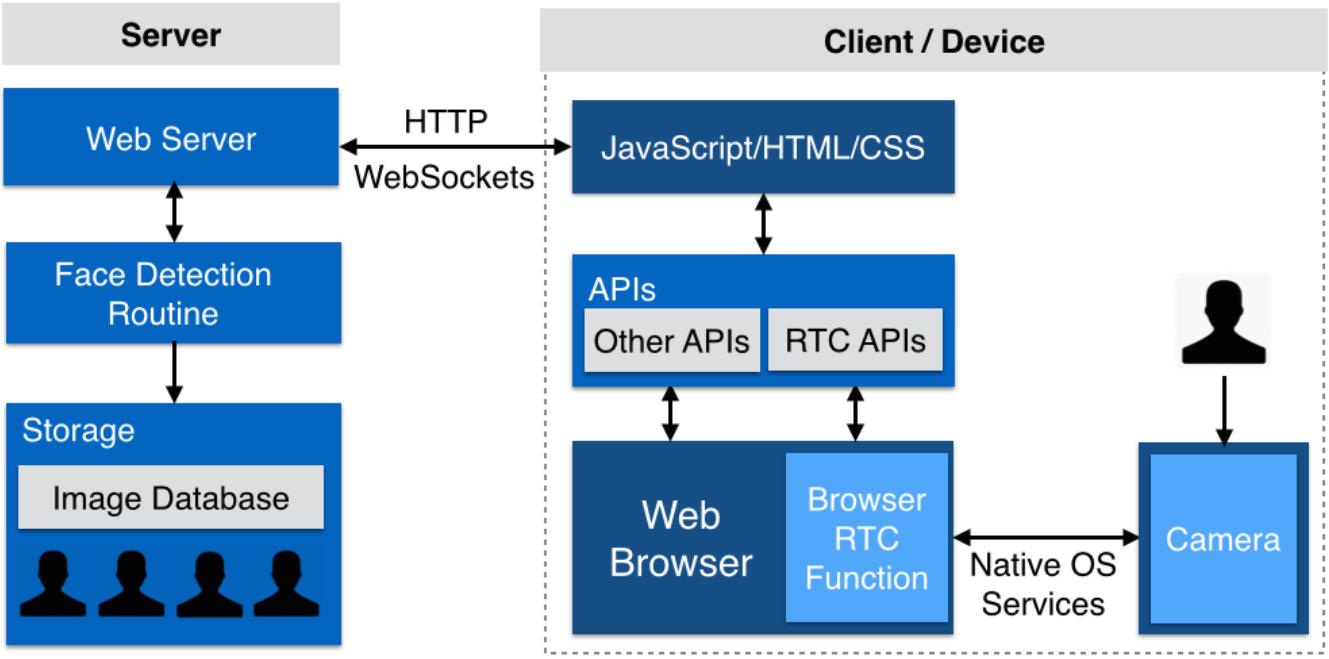


Fig. 4. Real-Time Face Detection System Architecture

TABLE I
A SYSTEM EVALUATION CONDITIONS

	Client	Server
Platform	Windows 8.1, Mac OS, Ubuntu Linux	Windows Server 2012
Web Server	-	IIS
Language	HTML, JavaScript	PHP
Capture device	WebCam	-

evaluation conditions are shown in Table I. Since the face detection is related with the quality of image such as level of light, size of the face in the image, and angle of visibility, thus the image that using in this paper was captured such as normal fluorescent light, light exposure, low light, and side view respectively as depicted in Figure 5. This implementation operates on the Google Chrome browser and HTTPS, that is a communications protocol for secure communication over a computer network, with especially wide deployment on the Internet.

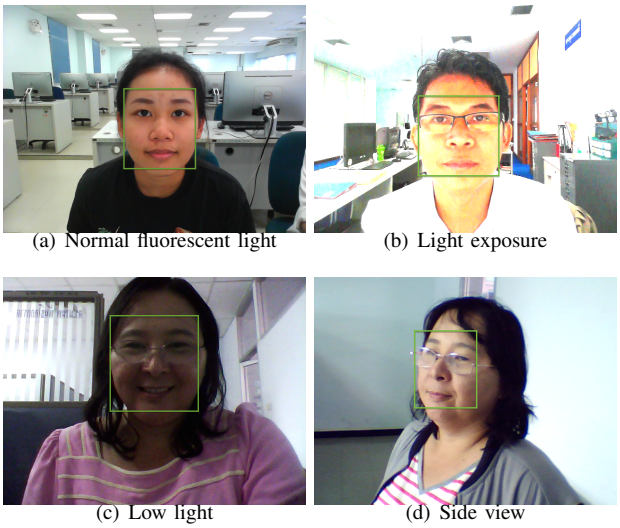
IV. Results and Discussion

In this section, the evaluation factors of our proposed system will be verified. According to the implementation components as shown in Figure 4, the results comprise of six evaluation processes that are the image quality, connection speed, security, performance, platform and device independencies. As results depicted in Table II, where *Excellent* is greater than five images per second, *Good* is four to five image per second, *Fair* is two to three images per second, and *Poor* is one image per second or less than. The image quality is depend on the connection speed; this system works performed well on the high-speed connection such as LAN, and worked quite well on the WiFi. In the case of using the 3G network connection, it cannot send the high resolution image through the network in time.

TABLE II
THE RESULTS OF SPEED CONNECTION IN THE VARIOUS CONDITIONS

Platform	Image Size	Network		
		3G	WiFi	LAN
Windows	320x280	Fair	Excellent	Excellent
	640x480	Poor	Good	Excellent
	1024x768	Poor	Fair	Excellent
Mac OS	320x280	Fair	Excellent	Excellent
	640x480	Poor	Good	Excellent
	1024x768	Poor	Fair	Excellent
Linux	320x280	Fair	Excellent	Excellent
	640x480	Poor	Good	Excellent
	1024x768	Poor	Fair	Excellent

In summary, the implementation results showed that the effectiveness of real-time face detection system depends on the connection speed and the detection algorithm. The proposed system spends a little time, fast processing and high accuracy in detecting and tracking the face area. The main challenge of the real-time face detection is that the person is unreliable moving as a cause of the detection output sometimes consists



(a) Normal fluorescent light (b) Light exposure

(c) Low light (d) Side view

Fig. 5. An example images of real-time face detection in various light conditions.

of some missed detections.

V. Conclusion

This paper proposed the technical study of real-time face detection system and also cover a key technology. In order to find the most appropriate factors, techniques, and algorithm by evaluating the performance that included connection speed, and effectiveness of tracking and detecting a human face in various conditions. WebRTC worked performly in any condition of device and platform independence and also operate securely on the HTTPS protocol. In the case of the images transferring ability that relied on the connection speed; that LAN and WiFi are the most prefer for this best image quality. The results showed that Haar-like feature and CLM have significantly detected the face area over the web browser in almost light conditions. However, Haar-like had the better precision when operated on the low-speed processor and memory such as Windows 8.1 Tablet.

In the future work, the real-time face detection system will have recognition ability to identify the person who walking through the capture device. This technique also can be applied to any scope of the personal identification system. In addition, we will develop the application that allows the users can access the system from the remote devices by detecting the face instead of generally password. Furthermore, this technology can operate on the mobile device such as mobile phone and Tablet PCs, by using HTML, JavaScript programming language, and standard protocols which widely compatible with other mobile platforms.

Acknowledgment

This work was funded by the Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang.

References

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–458, Dec. 2003.

[2] Y. Lu, J. Zhou, and S. Yu, "A survey of face detection, extraction and recognition." *Computers and Artificial Intelligence*, vol. 22, no. 2, pp. 163–195, 2003.

[3] B. Menser and M. Brunig, "Face detection and tracking for video coding applications," in *Signals, Systems and Computers, 2000. Conference Record of the Thirty-Fourth Asilomar Conference on*, vol. 1, Oct 2000, pp. 49–53.

[4] Y. Jie, L. Xufeng, Z. Yitan, and Z. Zhonglong, "A face detection and recognition system in color image series," *Math. Comput. Simul.*, vol. 77, no. 5-6, pp. 531–539, May 2008.

[5] I. Grigorik, *High Performance Browser Networking: What Every Web Developer Should Know about Networking and Web Performance*. O'Reilly Media, 2013.

[6] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1, 2001, pp. I–511–I–518.

[7] D. Cristinacce and T. F. Cootes, "Feature detection and tracking with constrained local models," in *Proceedings of the British Machine Vision Conference*. BMVA Press, 2006, pp. 95.1–95.10.

[8] L. Lang and W. Gu, "Study of face detection algorithm for real-time face detection system," in *Electronic Commerce and Security, 2009. ISECS '09. Second International Symposium on*, vol. 2, May 2009, pp. 129–132.

[9] A. Dutta, R. Veldhuis, and L. Spreeuwers, "The impact of image quality on the performance of face recognition," in *33rd WIC Symposium on Information Theory in the Benelux*. Enschede, the Netherlands: University of Twente, Centre for Telematics and Information Technology (CTIT), May 2012, pp. 141–148.

[10] F. Hua, P. Johnson, N. Sazonova, P. Lopez-Meyer, and S. Schuckers, "Impact of out-of-focus blur on face recognition performance based on modular transfer function," in *Biometrics (ICB), 2012 5th IAPR International Conference on*, March 2012, pp. 85–90.

[11] W. Elleuch, "Models for multimedia conference between browsers based on webrtc," in *Wireless and Mobile Computing, Networking and Communications (WiMob), 2013 IEEE 9th International Conference on*, Oct 2013, pp. 279–284.

[12] S. Loreto and S. P. Romano, *Real-Time Communication with WebRTC: Peer-to-Peer in the Browser*, 1st ed., S. St.Laurent and A. MacDonald, Eds. O'Reilly Media, Inc., May 2014.

[13] S. Cheng, A. Asthana, S. Zafeiriou, J. Shen, and M. Pantic, "Real-time generic face tracking in the wild with cuda," in *Proceedings of the 5th ACM Multimedia Systems Conference*. New York, NY, USA: ACM, 2014, pp. 148–151.

[14] D. Cristinacce and T. Cootes, "Facial feature detection and tracking with automatic template selection," in *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on*, April 2006, pp. 429–434.

[15] J. Saragih, S. Lucey, and J. Cohn, "Deformable model fitting by regularized landmark mean-shift," *International Journal of Computer Vision*, vol. 91, no. 2, pp. 200–215, 2011.

[16] C. Papageorgiou, M. Oren, and T. Poggio, "A general framework for object detection," in *Computer Vision, 1998. Sixth International Conference on*, Jan 1998, pp. 555–562.

[17] D. Zhang, S. Z. Li, and D. Gatica-Perez, "Real-time face detection using boosting in hierarchical feature spaces," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 2, Aug 2004, pp. 411–414 Vol.2.

[18] P. I. Wilson and J. Fernandez, "Facial feature detection using haar classifiers," *J. Comput. Sci. Coll.*, vol. 21, no. 4, pp. 127–133, Apr. 2006.