

Real-Time Person Identification System for Intelligent Digital TV

Min-Cheol Hwang, Le Thanh Ha, Seung-Kyun Kim, and Sung-Jea Ko, *Senior member, IEEE*
Department of Electronic Engineering, Korea University, Seoul, Korea

Abstract—Intelligent digital TV (iDTV) is an enhanced digital TV that can automatically provide user-personalized services for each audience. For the user-personalized services, the iDTV should recognize audiences in real-time. Thus, in this paper, we define a novel structure of the iDTV and propose a real-time person identification system embedded in the iDTV. The proposed system consists of three phases: preprocessing for reducing computational costs of the proposed system, face detection using a statistical approach with Haar-like features, and face recognition using Support Vector Machines (SVMs). Experimental results show that the proposed system achieves efficient performance with high recognition accuracy of above 90% at the speed of 15–20 fps, which is suitable for applying to the iDTV.

I. INTRODUCTION

Recently with the development of compression technology and the digitization of TV programs, the digital TV (DTV) provides not only high quality audio-visual effects as in high definition TV (HDTV) but also datacasting and multimedia interactive services. According to the increasing amount of services provided by the DTV, conventional methods of channel selection such as browsing become impractical [1]. The electronic program guide (EPG) can help viewers check future programs in advance. However, the multi-channel DTV service delivers more programs than viewers can handle, consequently resulting in information overload for users [2]. Thus, the ability of the DTV to provide user-personalized services for the individual person automatically is necessarily required. In order to provide the personalized service, an intelligent DTV (iDTV) needs to identify audiences.

Considering the requirements and problems mentioned previously, we define a novel structure of the iDTV as illustrated in Fig. 1. Our iDTV consists of two main components: intelligent environment sensor manager (iESM) and intelligent EPG manager (iEPGM). Equipped with various sensors the iESM investigates the indoor environment. The iESM also carries out the person identification by using the gathered information through the sensors. By using the information received from the iESM, the iEPGM manages the list of programs which are appropriate for audiences and provides the user-personalized service to each audience.

In this paper, we focus on the person identification module in the iESM, and the others are left for future research. Compared with other existing identification technologies such as voice, fingerprint and iris recognition, face recognition (FR) has several appropriate characteristics for consumer applications [3]. Therefore, we choose the FR method for identifying audiences. The proposed system consists of three phases: preprocessing for reducing computational costs, face

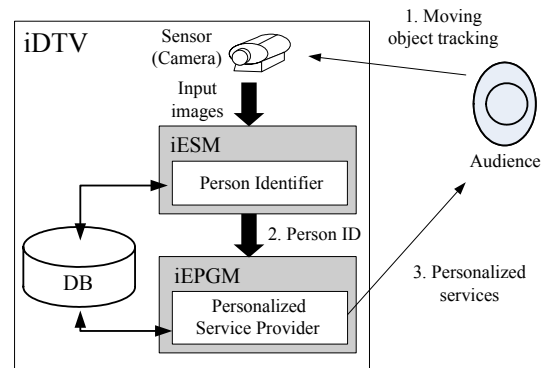


Fig. 1. Proposed structure of the iDTV.

detection (FD) using a statistical approach, and FR using Support Vector Machines (SVMs).

II. THE PROPOSED SYSTEM

Fig. 2 portrays the processing flow of the proposed person identification system for the iESM.

A. Preprocessing

To detect faces anywhere, the detector is applied to every location in the input image. To detect faces with various sizes, the input image is scaled and the detector is applied to all the scaled images. As a result, the computational complexity increases drastically. Generally people are considered as moving objects. The search range of FD can be reduced by considering the area related to the moving objects. To extract moving objects, background subtraction [4] can be applied. Using the stereo image recognition system [5], the distance between each object and the cameras is obtained. A face size can be estimated by using the relationship between object's distance and its size. Therefore, the face detector deals with only one scaled input image.

B. Face Detection

We use a statistical approach for FD, the approach originally developed by Viola and Jones [6]. Cascade of boosted tree classifiers and Haar-like features computed in a similar way to the Haar wavelet coefficients are used in this method. Each feature relates to the face features. The several boosted classifiers with constantly increasing complexity are chained into the cascade. Each classifier is subsequently applied to the face candidate until it gets rejected by one of boosted classifiers or until it passes all.

C. Face Recognition

SVMs [7] shown to be effective learning mechanisms for object recognition are used for FR. Compared with other learning methods such as eigenfaces, hidden markov models,

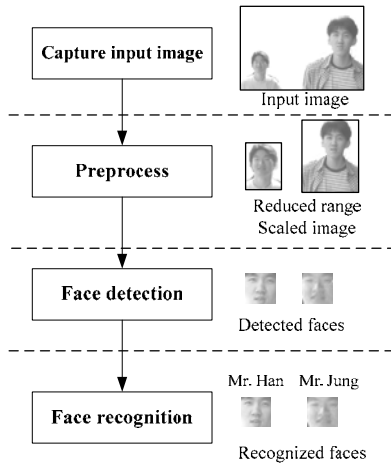


Fig. 2. Processing flow of the proposed iESM.

and neural network, the training time of SVMs is short enough to be applied to the real-time iDTV.

III. EXPERIMENTAL RESULTS

In order to evaluate the accuracy of the proposed system, we have separately examined the FD phase including preprocessing and the FR phase. Each phase has its own accuracy and the accuracy of the whole system is obtained by multiplying above two kinds of accuracy.

Three testing data sets are used to examine the FD phase. SET1¹ and SET2² contain only one frontal face. SET3 created by us consists of 135 images containing a total of 283 faces as well as nature scenery. Table I summarizes results. The overall performance of the FD phase is 96.2% correct rate with 329 false detections. Because the proposed system is attached to the iDTV, the background is fixed and predictable. As a result, false detections can be removed easily.

Considering the application of FR, we take an experiment with 4 people under 3 different lighting conditions: full, half and low light. The accuracy of the FR phase is summarized in Tables II and III. We use two kinds of the training sample sets. One set is taken in full light environment and the other is taken in low light. The FR accuracy depends on the light conditions where the training samples are taken. Table III shows better performance than Table II. It suggests that the training samples should be taken in the condition where light effects are minimized. Two or more faces can be recognized simultaneously and independently as shown Fig. 3 which also shows GUI of our proposed system.

IV. CONCLUSION

In this paper, we have defined a novel structure of the iDTV and proposed the real-time person identification system for the iDTV. To identify audiences, the proposed system has used a FR algorithm which consists of preprocessing, FD, and

¹ The 188 images (are from Website: <http://vis-www.cs.umass.edu/~vidit/AI/dbase.html>) have been created by Vidit Jain and Dr. Amitabha Mukherjee in cooperation with Neeraj Kayal, Pooja Nath and Utkarsh Hriday Shrivastav.

² The 1521 images are from Website: <http://www.BioID.com>.

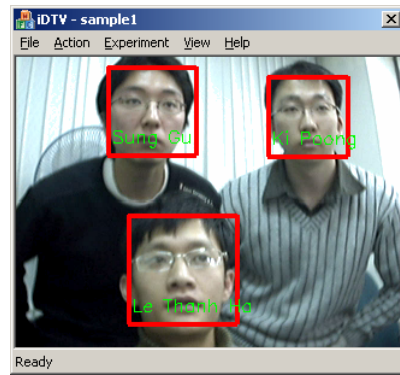


Fig. 3. Recognition results and GUI of the proposed system.

TABLE I
FD PERFORMANCE OF THE PROPOSED SYSTEM

	# of faces	# of detected faces	# of false detections	Detection rate
Set1	188	188	10	100%
Set2	1,521	1,467	263	96.4%
Set3	283	261	56	92.2%
Total	1,992	1,916	329	96.2%

TABLE II
FR ACCURACY USING TRAINING SAMPLES OBTAINED IN FULL LIGHT

Light condition	P1	P2	P3	P4	AVG
Full	96%	97%	87%	85%	91%
Half	85%	85%	73%	86%	82%
Low	74%	27%	37%	54%	48%

TABLE III
FR ACCURACY USING TRAINING SAMPLES OBTAINED IN LOW LIGHT

Light condition	P1	P2	P3	P4	AVG
Full	85%	86%	95%	95%	90%
Half	77%	50%	96%	99%	80%
Low	90%	96%	95%	98%	94%

FR. Experimental results have shown that the proposed system achieves the performance with the accuracy rate of above 90% and 15~20 fps processing speed. We have focused on the person identification module in the iESM, and the others are left for future research.

REFERENCES

- [1] T. Isobe, M. Fujiwara, and H. Kaneta, "Development and features of a TV navigation system," *IEEE Trans. on Consumer electronics*, vol. 50, no. 1, pp. 393-399, Nov. 2003.
- [2] H. Zhang, S. Zheng, and J. Yuan, "A personalized TV guide system compliant with MHP," *IEEE Trans. on Consumer electronics*, vol. 51, no. 2, pp. 731-737, May 2005.
- [3] F. Zuo and P. H. N. de With, "Real-time embedded face recognition for smart home," *IEEE Trans. on Consumer electronics*, vol. 51, no. 1, pp. 183-190, Feb. 2005.
- [4] J. W. Kim, K. S. Choi, B. D. Choi, J. Y. Lee, and S. J. Ko, "Real-time system for counting the number of passing people using a single camera," *Lecture Notes in Computer Science*, vol. 2781, pp. 466-473, Sep. 2003.
- [5] S. Nedeveschi, R. Danescu, D. Frentiu, T. Marita, F. Oniga, C. Pocol, R. Schmidt, and T. Graf, "High accuracy stereo vision system for far distance obstacle detection," *IEEE Proc. Intelligent Vehicles Symposium*, pp. 292-297, Jun. 2004.
- [6] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE Proc. Computer Vision and Pattern Recognition*, vol. 1, pp. 511-518, 2001.
- [7] Cristiana and S. Taylor, *An introduction to Support Vector Machines*, Cambridge University Press, UK, 2000.