

# Image Processing Techniques for Measuring Advertising Effectiveness of Digital Signage

*Tetsuya Kinebuchi<sup>†</sup>, Hiroyuki Arai, Isao Miyagawa, Shingo Ando, Kaori Kataoka, and Hideki Koike*

## Abstract

NTT Cyber Space Laboratories has developed a technique for estimating the number of people in a crowd from images and a technique for detecting faces and estimating their orientation. These techniques will meet the urgent need to establish a way to measure the advertising effectiveness of digital signage and will stimulate the distribution of advertisements using that medium.

## 1. Introduction

Digital signage is attracting attention as an advertising medium that opens up new possibilities such as freely presenting advertisements according to the time band, location, and local conditions. However, there is currently no objective index of advertising effectiveness for digital signage whereas there is for TV broadcasting. Such an index would provide important information for the advertisement distribution business. An objective index of advertising effectiveness and a means of measuring it are important requirements for digital signage to be used as an advertising medium in the future.

There have been various studies on indicators of the advertising effectiveness of digital signage. Daily Effective Circulation, which indicates the number of persons that may come into contact with an advertisement, was proposed in 2001 by the Outdoor Advertising Research Forum as an effectiveness index for outdoor billboards, including large-format video displays on the street. It is calculated on the basis of road traffic sensor data published by Japan's Ministry of Land, Infrastructure and Transport every few years or traffic survey data acquired by the prescribed method.

This index is probably appropriate for billboards, which display content that does not change frequently. For digital signage, however, it may have some inadequacies because the displayed advertisements might be changed daily or even hourly.

There has also been research that took fully into account the characteristics of digital signage. The Digital Signage Consortium in Japan established the Digital Signage Standard Guidelines (version 1.0) in November 2008 and has proposed the addition of a method for evaluating advertising effectiveness on the basis of exposure (number of persons reached and frequency), taking into account the characteristics of digital signage. The Out-of-home Video Advertising Bureau in the USA published "Audience Metrics Guidelines" in 2008, proposing the Average Unit Audience display as an index based on how long people remain in front of the display and pay attention to it for at least a certain time. Although these documents are only at the stage of proposed guidelines and cannot be said to have broad acceptance, they suggest that the acquisition of dynamic time series data about how many people are present and the extent to which they are paying attention to the display may serve as an effectiveness index for digital signage advertising.

After establishing an advertising effectiveness index, one must also implement a means of measur-

<sup>†</sup> NTT Cyber Space Laboratories  
Yokosuka-shi, 239-0847 Japan

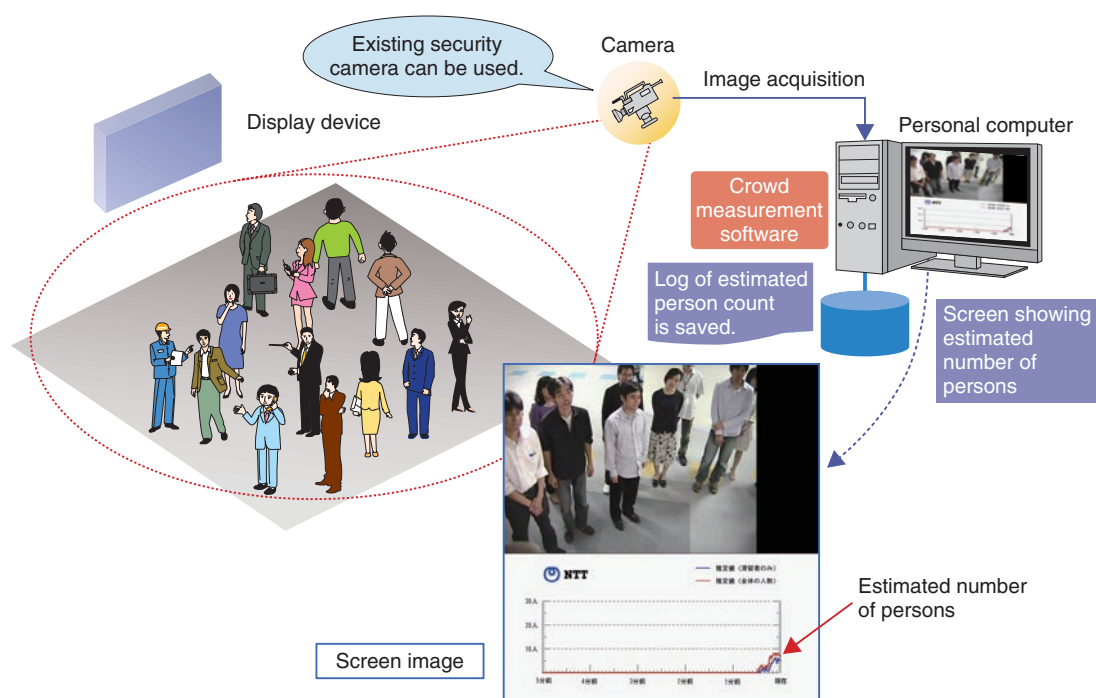


Fig. 1. Overview of crowd measurement technique.

ing it. In the case of TV viewing rate, it is possible to obtain a numerical value that has definite validity by conducting a survey of even an extremely small sample because the main sampled population is home viewers and there is no great variation in viewing conditions for single viewers or families. The viewing conditions for digital signage vary widely for different locations, so ultimately measurement for all signage is necessary. Since it is probably impractical to measure all signage, we believe it is necessary to acquire measurement data for as many displays as necessary to produce a valid index. Consequently, it is important to have a method that can easily measure advertising effectiveness at low expense.

## 2. Image processing techniques for measuring advertising effectiveness

NTT Cyber Space Laboratories has developed a crowd measurement technique and a face detection and orientation estimation technique suitable for measuring the advertising effectiveness of digital signage from images acquired by cameras. This system can use a relatively inexpensive camera and an ordinary personal computer, making possible a low-cost system. It can compute advertising effectiveness

as time series data.

The crowd measurement technique uses image processing to estimate the number of persons at the location photographed by the camera (**Fig. 1**). This technique can achieve successful measurements with an existing camera aimed obliquely downward and can achieve stable processing even for crowded conditions.

The face detection and orientation estimation technique detects human face areas in a camera image and estimates the direction of orientation for each face. It can thus estimate the number of persons who are facing the camera, and hence looking at the advertisement, if the camera is positioned close to the display (**Fig. 2**). By tracking the face areas, the system can compute the cumulative duration for which individual faces looked toward the camera (gaze time). This technique is robust against changes in lighting and face orientation. It can produce time series data for face orientation changes and can estimate face areas even in low-resolution images.

These two techniques are described in more detail below.

### 2.1 Crowd measurement technique

The crowd measurement technique uses image pro-

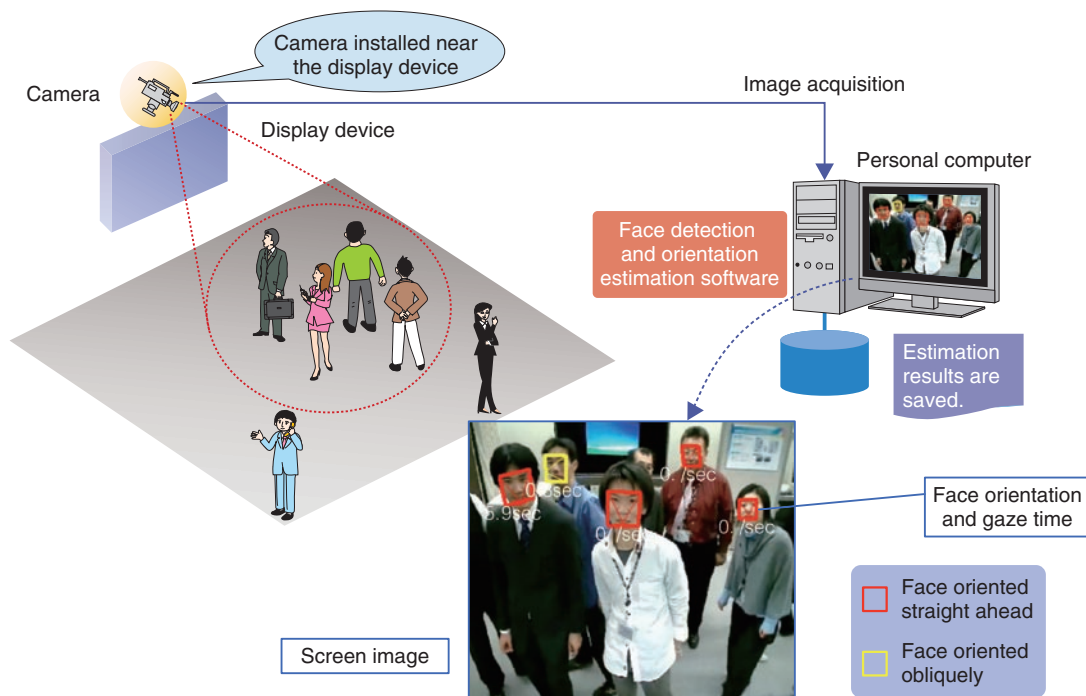


Fig. 2. Overview of the face detection and orientation estimation technique.

cessing to obtain the approximate number of persons in the area photographed by a camera. Among previous methods of counting people in a camera image, the main ones are shape detection and object tracking methods. All of those methods have problems with stable operation when there are many people in the image or when there is any overlapping of people in the image. The shape detection methods fail when the detection target is partially obscured and detection errors occur when there are many falsely similar shapes in areas where human figures overlap in the image. The object tracking methods have difficulty tracking people when they come close to each other or intermingle. Our new technique, however, applies a newly developed algorithm that takes into account the positional relationships of the camera, the plane of the ground, and the people in the image to estimate the number of people from the area covered by the image. The algorithm's concept is illustrated in **Fig. 3**. The algorithm presumes that the geometric relationship between the camera and the ground plane is known (i.e., calibration has been completed).

First, the presence of any object at each pixel location in the image is detected; that is, if the pixel luminance value (brightness) is different from the constant

value of the wall or floor, etc. (referred to as the background), then that is detected. Pixels judged as not part of the background are referred to as foreground (corresponding to the regions of people in the image shown in Fig. 3). Then, for all of the foreground pixels, the proportion of the surface area that the pixel represents (actual spatial surface area) is calculated and the values for all detected foreground pixels are summed. The number of persons can then be estimated by computing the sum as a factor of the standard person surface area ( $S_0$ ). The size of the actual spatial surface area sum can be used to estimate the extent of overlap (occlusion) of persons, so this factor can be included in the final estimation of number of people.

Basically, this algorithm can estimate the number of persons in an image from the cumulative results of signal processing in single-pixel units. That allows stable operation even under crowded conditions. Moreover, because occlusion can be estimated, its effect can be reflected in the estimated person count, so it can be used with ordinary security cameras that point obliquely downward.

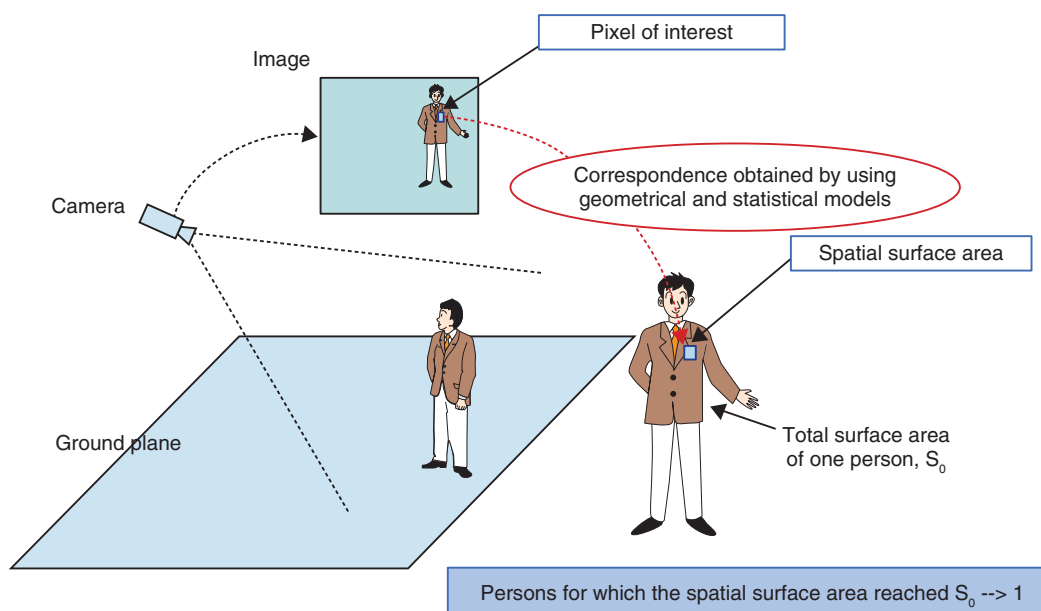


Fig. 3. Crowd estimation algorithm.

## 2.2 Face detection and orientation estimation technique

The face detection and orientation technique can estimate the number of persons who are facing a camera by detecting faces in images acquired by the camera and estimating the orientation of the detected faces. This technique comprises the two techniques of face detection and face orientation estimation. The skillful combination of those two techniques makes it possible to estimate the number of people in real time (Fig. 4).

Face detection is accomplished by a face detector that has been trained with a collection of images of various persons' faces orientated in the up, down, left, and right directions in several tens of angular steps for each direction. In addition, pre-processing with a filtering process called morphological computation is performed. That processing controls the influence of facial shadows to achieve face detection that is robust against lighting conditions.

The face orientation estimation technique estimates the three-dimensional orientation of the detected faces. Face orientation is represented by three parameters (yaw, pitch, and roll). Accordingly, the estimation of those three parameters from a face area image pattern can be framed as a recursion problem. The technique described here applies a pose estimation method that uses principle component analysis and support vector regression (SVR), which solves the

regression problem by using a nonlinear regression method on a linear partial space. Furthermore, robustness against partial occlusion and background changes can be achieved by integrating the results for local areas.

The main approach in conventional face orientation estimation is to use feature spaces with the end points of the eyes and mouth as measurements. That approach involves two problems: 1) high-resolution images are required in order to extract the feature points and 2) all of the feature points must be in the image. Our technique is based on the appearance of the entire face. It can thus be used with relatively low-resolution images and can even measure oblique faces that are nearly horizontal.

## 3. Future plans

We plan to proceed with field trials of the crowd estimation technique and the face detection and orientation estimation technique on the assumption that they will be applied to the measurement of the advertising effectiveness of digital signage in order to clarify problems in practical use and to make technical improvements. For the crowd estimation technique, we also intend to investigate application areas other than digital signage, such as facilities safety management and services that provide congestion information. For the face detection and orientation

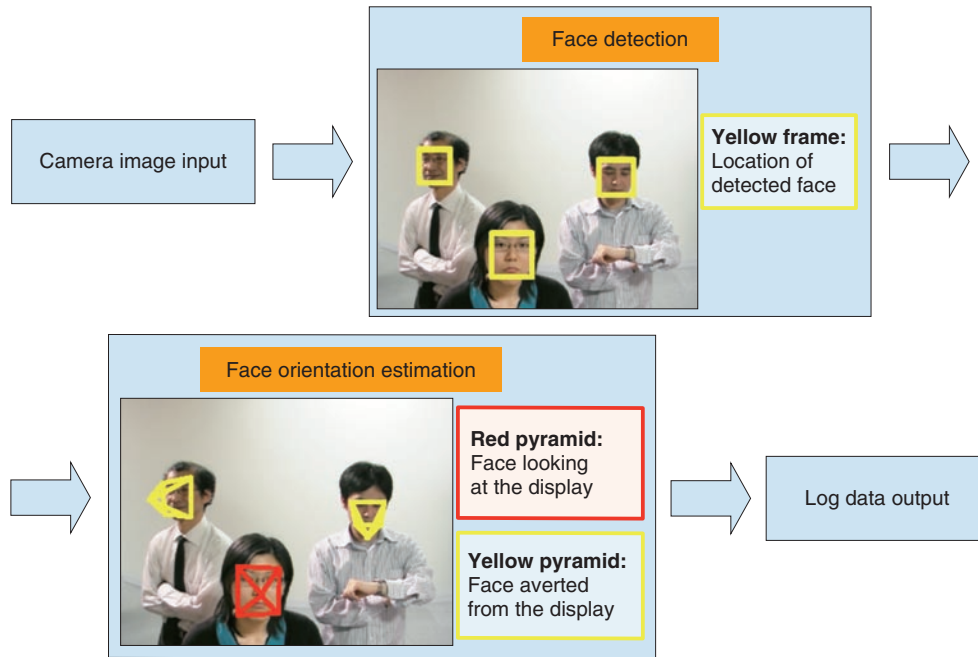


Fig. 4. Flow of face detection and orientation estimation processing.

estimation technique, we will also conduct research and development with an eye to applications in fields other than digital signage, such as remote monitoring and estimation services. Moreover, for application to the measurement of advertising effectiveness, we will investigate whether or not more detailed information can be extracted in addition to the number of people

facing the camera. For example, if it is possible to extract attribute information such as sex or age from face images, a more finely tuned measurement of the effectiveness will be possible and marketing can be based on more detailed results. We believe that will increase the added value of the digital signage medium.


**Tetsuya Kinebuchi**

Senior Research Engineer, Visual Media Communications Project, NTT Cyber Space Laboratories.

He received the M.S. degree in physics from Tohoku University, Miyagi, in 1997. He joined NTT in 1997 and has been engaged in research on image processing and pattern recognition.


**Shingo Ando**

Research Engineer, Visual Media Communications Project, NTT Cyber Space Laboratories.

He received the B.E. degree in electrical engineering and the Ph.D. degree in engineering from Keio University, Kanagawa, in 1998 and 2003, respectively. He joined NTT in 2003. He has been engaged in research and practical application development in the fields of image processing and pattern recognition. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan.


**Hiroyuki Arai**

Senior Research Engineer, Visual Media Communications Project, NTT Cyber Space Laboratories.

He received the M.S. degree in physics from Hokkaido University in 1991. He joined NTT in 1991 and has been engaged in research on map recognition systems, image processing, and pattern recognition. He was a fellowship researcher of the "Natural Vision Project" of the National Institute of Information and Communications Technology of Japan (NICT) from 2000 to 2005. He is a member of the Institute of Image Information and Television Engineers of Japan.


**Kaori Kataoka**

Research Engineer, Visual Media Communications Project, NTT Cyber Space Laboratories.

She received the M.S. degree in physics from Waseda University, Tokyo, in 2000. She joined NTT in 2000 and has been engaged in research on image processing and pattern matching. She is a member of IEICE.


**Isao Miyagawa**

Research Engineer, Visual Media Communications Project, NTT Cyber Space Laboratories.

He received the B.E. degree in electronics engineering and the Ph.D. degree from Fukui University in 1991 and 2005, respectively. He joined NTT in 1991 and from 1991 to 1996, he researched communication technology of color documents and developed some facsimile devices. Since 1997, he has researched computer vision.


**Hideki Koike**

Senior Research Engineer, Supervisor, Group Leader, Visual Media Communications Project, NTT Cyber Space Laboratories.

He received the M.S. degree in mathematics from Tohoku University, Miyagi, in 1985. He joined NTT in 1985 and has been engaged in research on computer vision.