J Real-Time Image Proc (2013) 8:379–388
DOI 10.1007/s11554-012-0255-8

ORIGINAL RESEARCH PAPER

500-fps face tracking system

Idaku Ishii · Tomoki Ichida · Qingyi Gu · Takeshi Takaki

Received: 26 December 2011 / Accepted: 4 May 2012 / Published online: 30 May 2012

Abstract In this paper, we propose a high-speed vision system that can be applied to real-time face tracking at 500 fps using GPU acceleration of a boosting-based face tracking algorithm. By assuming a small image displacement between frames, which is a property of high-frame rate vision, we develop an improved boosting-based face tracking algorithm for fast face tracking by enhancing the Viola–Jones face detector. In the improved algorithm, face detection can be efficiently accelerated by reducing the number of window searches for Haar-like features, and the tracked face pattern can be localized pixel-wise even when the window is sparsely scanned for a larger face pattern by introducing skin color extraction in the boosting-based face detector. The improved boosting-based face tracking algorithm is implemented on a GPU-based high-speed vision platform, and face tracking can be executed in real time at 500 fps for an 8-bit color image of 512 × 512 pixels. In order to verify the effectiveness of the developed face tracking system, we install it on a two-axis mechanical active vision system and perform several experiments for tracking face patterns.

I. Ishii (✉) · T. Ichida · Q. Gu · T. Takaki
1-4-1 Kagamiyama, Higashi-Hiroshima,
Hiroshima 739-8527, Japan
e-mail: iishii@robotics.hiroshima-u.ac.jp

T. Ichida
e-mail: ichida@robotics.hiroshima-u.ac.jp

Q. Gu
e-mail: gu@robotics.hiroshima-u.ac.jp

T. Takaki
e-mail: takaki@robotics.hiroshima-u.ac.jp

1 Introduction

Face detection is a basic image processing technology for detecting and localizing human faces in images [1], and many types of face detection algorithms have been proposed in the 1990s, including knowledge-based methods [2], feature-invariant methods [3, 4], template methods [5, 6], and appearance-based methods [7, 8]. The Viola–Jones face detector [9] is a well-known face detection algorithm for practical uses, based on Haar-like features with cascade AdaBoost classifiers; it can robustly detect human faces even under partial occlusion without heavy computation. Such boosting-based face detection schemes have evolved as the de facto standard of real-time face detection in real-world applications, such as autofocusing in digital cameras, driver monitoring, and biometric authentication.

Field-programmable gate array (FPGA) implementations [10–12] of boosting-based face detection algorithms have been reported for accelerating face detection. A graphical processing unit (GPU) implementation [13] of the Viola–Jones face detector has also been reported as a more effective solution to avoid the complexity of logic synthesis in FPGA design, such as timing constraints. Most of these implementations have been developed for QVGA (320 × 240) or VGA (640 × 480) images at frame rates of dozens of frames per second, and the cameras are constrained by the video signal formats (e.g., NTSC at 30 fps or PAL at 25 fps), which are designed according to the characteristics of the human eye. However in the case of rapid human motion, video images at frame rates of several dozen frames per second are often insufficient for tracking a human face responsively. Therefore, there is a strong demand for high-speed face tracking at frequencies of the order of several hundred hertz.