

Intelligent Three-Dimensional Vision Sensor with Ears

Shigeru Ando

Faculty of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

(Received May 16, 1994; accepted June 8, 1994)

Key words: intelligent sensor, vision sensor, auditory sensor, gradient correlation, image processing, sound localization

This paper describes our newly developed intelligent robot sensor system which comprises two eyes and four ears on a movable head. It can acquire a world image, i.e., a dynamical 3-D map of the environment, while showing humanlike behavior naturally and autonomously. The most important feature of the sensor system is the original and unified sensory architecture of the low-level and intermediate-level visual and auditory functions. We describe in this paper the key items for realizing this sensor: 1) features of the unified sensory architecture "gradient correlation sensing principle," 2) the design of sensory probes and circuits for the vision section, 3) the design of sensory probes and circuits for the auditory section, 4) various algorithms to extract sensory information from each sensor, and 5) autonomous and active sensing procedures by means of visual-auditory fusion and sensory-motor control.

1. Introduction

We are developing a robot head sensor system, "SmartHead," as shown in Fig. 1, in which a binocular vision sensor and a quadaural auditory sensor are mounted on an autonomously movable head. Both the vision and auditory sensors are ready to detect unusual events, i.e., visual motion/accretion and auditory localization of sound sources, in its surrounding environment as an early warning system. If such an event is found, then head motors and a high-speed saccadic eye movement system are activated and they quickly move the field of view (FOV) of the vision sensor in the direction of the object. A tracking and fixation system catches the most salient feature in that direction and fixes the