Neurocomputing 120 (2013) 130–140

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Ego-motion-compensated object recognition using type-2 fuzzy set for a moving robot

Tae-Koo Kang [a], Huazhen Zhang [a], Gwi-Tae Park [a,*], Dong W. Kim [b,*]

[a] Department of Electrical Engineering, Korea University, Seoul, South Korea
[b] Department of Digital Electronics, Inha Technical College, 253 Yong Hyun-dong, Nam-Gu, Incheon 402-752, South Korea

ARTICLE INFO

Article history:
Received 22 December 2011
Received in revised form
8 August 2012
Accepted 10 September 2012
Available online 26 March 2013

Keywords:
Type-2 fuzzy set
Object recognition
Ego-motion compensation
Moving robot

ABSTRACT

This paper presents an efficient stereovision-based motion compensation method for moving robots. The vision system of a moving robot enables it to detect and localize known objects in the images obtained from the camera mounted in its head. However, ego-motion causes some errors that need to be eliminated. We therefore propose an ego-motion compensation method that eliminates the errors in environment recognition caused by the ego-motion of a moving robot, in addition to efficiently improving recognition accuracy. The proposed method uses the disparity map obtained from three-dimensional (3D) vision and can be divided into three modules: segmentation, feature extraction, and estimation. In the segmentation module, we propose the use of extended type-2 fuzzy information theory (ET2FIT) to extract the objects. The results of using ET2FIT are then compared and analyzed to those obtained using type-2 fuzzy set and normal type-1 fuzzy set. The conventional fuzzy information theory [11] can only be applied to the binary image case. Therefore, we need to modify the existing method.

In the feature extraction module, features are extracted using wavelet level-set transform, and least-square ellipse approximation is used in the estimation module to calculate the displacement for the rotation and translation between image sequences. The results of experiments indicate that our proposed method is highly effective when applied to moving robots, especially humanoid robots walking and operating in the real world.

1. Introduction

Humanoid robots are similar in appearance to human beings — with a head, two arms, and two legs — and have some human-like intelligent abilities such as object recognition, tracking, voice identification, and obstacle avoidance. Since they try to imitate human body structure and behavior and they are autonomous systems, humanoid robots are typically more complex than other kinds of robots. In the case of moving over an obstacle or detecting and localizing an object, it is critically important that as much precise information regarding obstacles/object as possible be obtained since the robot establishes contact with an obstacle/object by calculating the appropriate motion trajectories to the obstacle/object. A vision system supplies most of the information to the robot; however, the image sequence from the vision system of a humanoid robot is not static when a humanoid robot is walking; thus, some problems occur due to ego-motion. Therefore, humanoid robots need algorithms that can autonomously determine their actions and paths in unknown environments and compensate for the ego-motion arising in the vision system. The vision system is one of the most important sensors in a humanoid robot system and can supply a large amount of the information required by a humanoid robot. However, for more precise recognition, a stabilization module that can compensate for ego-motion is essential.

Over the years, much research in the motion compensation field has been conducted on the vision systems mounted in robots. Some researchers use a single camera, but stereovision, which can extract information regarding the depth of the environment, is commonly used. From 3D vision, robot motion can be estimated by the 3D rigid transform using a two-dimensional (2D) multi-scale tracker, which projects 3D depth information on the 2D feature space. The scale invariant feature transform (SIFT) [1], a local feature based algorithm used to extract features from images and estimate transformation using their location, and iterative closest point (ICP) [2], which is used for registration of digitized data from a rigid object with an idealized geometric model, have primarily been used for motion estimation using a single camera or stereo camera for video stabilization or autonomous navigation purposes; they have been widely used in wheeled robots [3–5]. Moreover, the optical flow-based method, which can estimate motion by a 3D normal flow

* Corresponding authors. Tel.: +82 1044003689.
E-mail addresses: tkkang@korea.ac.kr (T.-K. Kang), zhanghz@korea.ac.kr (H. Zhang), gtpark@korea.ac.kr (G.-T. Park), dwnkim@inhatc.ac.kr, dongwon.kim25@gmail.com (D.W. Kim).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2012.09.041

constraint using a gradient-based error function, is widely used because of the simplicity of computation [6]. Further, fully affine invariant speed up robust feature (SURF), which resolves the fully affine invariant and computation efficiency problem of conventional SURF [15]; robust co-occurrence histograms of oriented gradients (CoHOGs), which overcomes the shortcomings of conventional CoHOGs [16], and L1-norm-based tensor analysis [17], which overcomes the sensitivity of the conventional tensor analysis method, have been proposed for the motion estimation and analysis in the computer vision field.

However, these methods are not suitable for a biped humanoid robot, as the walking motion of such a robot requires simultaneous vertical and horizontal movements (unlike the motion of a mobile robot) as well as computation cost yielded by its point-to-point operation. Therefore, we propose a more efficient stereovision-based ego-motion estimation method for ego-motion compensation in humanoid robots.

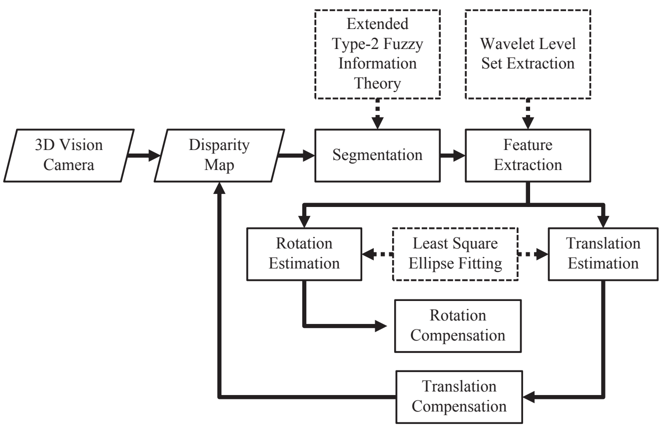

Fig. 1. Architecture of the ego-motion compensation system.

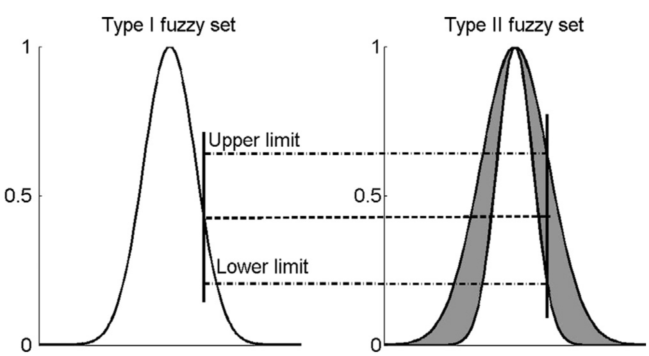

Fig. 2. Example of type-1 and type-2 membership functions.

The proposed ego-motion compensation method using a stereo camera consists of three modules: segmentation, feature extraction, and motion estimation. Stereovision can be used to obtain disparity images in which objects are shown in different gray levels according to the distance between the different objects and the humanoid robot itself. In the segmentation module, objects are extracted by means of image analysis using our proposed fuzzy information-theoretic approach based on type-2 fuzzy sets. The feature extraction module extracts the feature images using wavelet level set, which can obtain horizontal, vertical, and diagonal information for each object. The results of the feature extraction module are used as the input data to the estimation module. The position of each object can be calculated using least-square ellipse approximation. The difference in position between two images is calculated as the compensation parameters. In addition, our proposed type-2 fuzzy method is used to deal with noise data in order to obtain precise rotation and translation date sets.

The remainder of this paper is organized as follows. In Section 2, our proposed stereovision-based motion stabilization system for humanoid robots by means of fuzzy sets is presented. In Section 3, the results of experiments focused on verifying the efficacy of the proposed system is given. Section 4 concludes this paper by presenting the contributions of this study.

2. Proposed ego-motion compensation system

2.1. Proposed ego-motion compensation system

In order to eliminate the error in object recognition caused by the ego-motion of a walking humanoid robot, we propose a novel ego-motion compensation system based on fuzzy set theory using stereovision information. Later, we will also compare its performance using type-2 fuzzy set and normal type-1 fuzzy set, respectively. A vision system that uses an SR4000 camera can supply stereovision information. Stereovision is generated on the basis of the perspectives of our two eyes, leading to slight relative displacements of objects (disparities) in the two monocular views of a scene. These disparities are used to calculate the distance between the object and the camera in a 3D scene to generate a depth image.

The overall architecture of our proposed ego-motion compensation system is illustrated in Fig. 1. The system largely consists of three modules: segmentation, feature extraction, and estimation. The final estimation parameters obtained from the depth image are used to compensate for the ego-motion in the gray image for object recognition.

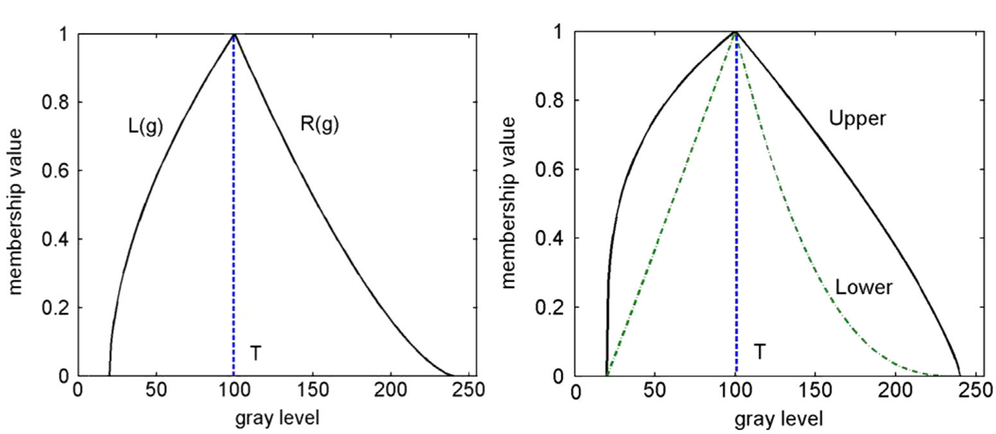

Fig. 3. *LR*-type membership function for image segmentation. Left: type-1 *LR*-type membership function; right: type-2 *LR*-type membership function.