Available online at www.sciencedirect.com
SCIENCE DIRECT®
Engineering Applications of Artificial Intelligence 19 (2006) 269–276
Engineering Applications of
ARTIFICIAL INTELLIGENCE
www.elsevier.com/locate/engappai

Using a genetic algorithm to register an uncalibrated image pair to a 3D surface model

Zsolt Jankó*[,1], Dmitry Chetverikov[1], Anikó Ekárt

Computer and Automation Research Institute, Budapest, Kende u.13-17, H-1111 Hungary and Eötvös Loránd University, Budapest, Hungary

Received 22 December 2004; received in revised form 9 June 2005; accepted 30 October 2005
Available online 4 January 2006

Abstract

In this paper we present a successful application of genetic algorithms to the registration of uncalibrated optical images to a 3D surface model. The problem is to find the projection matrices corresponding to the images in order to project the texture on the surface as precisely as possible. Recently, we have proposed a novel method that generalises the photo-consistency approach by Clarkson et al. to the case of uncalibrated cameras by using a genetic algorithm. In previous studies we focus on the computer vision aspects of the method, while here we analyse the genetic part. In particular, we use semi-synthetic data to study the performance of different GAs and various types of selector, mutation and crossover. New experimental results on real data are also presented to demonstrate the efficiency of the method.

Keywords: Registration; Photo-consistency; Genetic algorithm

1. Introduction

The problem of building and visualising photorealistic 3D models of real-world objects has become an important topic of computer vision during the last years. There are thousands of cultural heritage objects around the world in the danger of being hurt or destroyed. Ambitious projects have been started to preserve these objects by digitalising them. Such projects are: the Michelangelo Project (Levoy et al., 2000), the Pieta Project (Bernardini et al., 2002) and the Great Buddha Project (Ikeuchi et al., 2003).

Photorealistic 3D models must have precise geometry as well as detailed texture on the surface. Active and passive methods for creating such models are discussed in Yemez and Schmitt (2004). The methods are based on different principles. They use different techniques to reconstruct the object surface, acquire its texture and map the texture onto the surface. The geometry can be measured by various methods of computer vision. When precise measurements are needed, laser scanners are often used. However, most of laser scanners do not provide texture and colour information. Even when they do, the data provided are not accurate enough. (See Yemez and Schmitt, 2004 for a detailed discussion.)

Whatever the sources of the geometric and the textural information are, the problem of data fusion, or registration, is to be addressed. In this paper we consider the case when the two sources are independent. We approach the problem of combining precise geometry with high quality images by using genetic algorithms.

A number of approaches to the above registration problem have been proposed. In Jankó and Chetverikov (2004a,b) we introduced a novel method based on photo-consistency. The novelty of our method consists in using uncalibrated cameras—in contrast to Clarkson et al. (2001) who need a calibrated setup—and applying a genetic algorithm. Below we describe the problem of photo-consistency based registration and give a summary of our approach (Jankó and Chetverikov, 2004a,b).

*Corresponding author.
E-mail address: janko@sztaki.hu (Z. Jankó).
[1]Supported by the Hungarian Scientific Research Fund (OTKA) under Grants M28078 and T038355 and the EU Network of Excellence MUSCLE (FP6-507752).

0952-1976/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2005.10.002

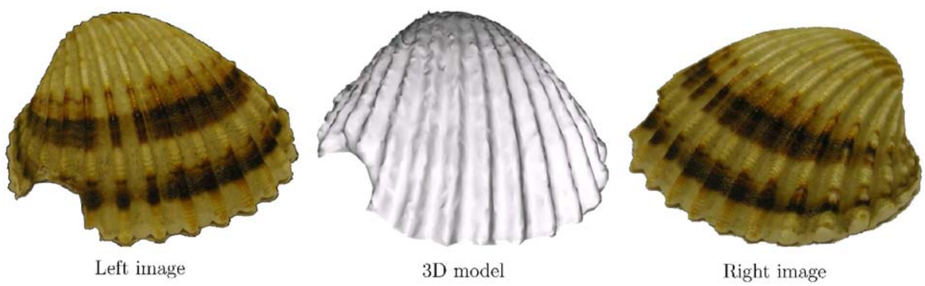

Fig. 1. The Shell dataset.

The mathematical formulation of the registration problem is the following. Two input images, I_1 and I_2, and a 3D model are given. They represent the same object. (See an example in Fig. 1.) The only assumptions about the environment are that the lighting conditions are fixed and the cameras have identical sensitivity.[2] All other camera parameters may differ and are unknown. The 3D model consists of a 3D point set $\mathscr{P}$ and a set of normal vectors assigned to the points. $\mathscr{P}$ is obtained by a hand-held 3D scanner and then triangulated by the robust algorithm of Kós (2001). This algorithm provides the normal vectors as well.

To project the object surface to the image plane, the finite projective camera model (Hartley and Zisserman, 2000) is used: $\mathbf{u} \simeq P\mathbf{X}$, where $\mathbf{u}$ is an image point, P the 3×4 *projection matrix* and $\mathbf{X}$ a surface point ($\simeq$ means that the projection is defined up to an unknown scale).

The task of registration is to determine the precise projection matrices, P_1 and P_2, for both images. The projection matrix P has 12 elements but only 11 degrees of freedom, since it is up to a scale factor. We denote the collection of the 11 unknown parameters by p, which represents the projection matrix P as an 11-dimensional parameter vector.

Values of p_1 and p_2 are sought such that the images are *consistent* in the sense that the corresponding points—different projections of the same 3D point—have the same colour value. Note that the precise mathematical definition is valid only when the surface is Lambertian, that is, the incoming light is reflected equally to every direction on the surface. This is usually true for diffuse surfaces. The formal definition is the following: We say that images I_1 and I_2 are consistent by P_1 and P_2 (or p_1 and p_2) if for each $X \in \mathscr{P}$: $\mathbf{u}_1 = P_1\mathbf{X}$, $\mathbf{u}_2 = P_2\mathbf{X}$ and $I_1(\mathbf{u}_1) = I_2(\mathbf{u}_2)$. (Here $I_i(\mathbf{u}_i)$ is the colour value in point $\mathbf{u}_i$ of image I_i.) This type of consistency is called *photo-consistency* (Clarkson et al., 2001; Kutulakos and Seitz, 1993).

The photo-consistency holds for accurate estimates for p_1 and p_2. Inversely, misregistered projection matrices mean much less photo-consistent images. The cost function introduced in Jankó and Chetverikov (2004a) is the following:

$$C_\phi(p_1, p_2) = \frac{1}{|\mathscr{P}|} \sum_{\mathbf{X} \in \mathscr{P}} \|I_1(P_1\mathbf{X}) - I_2(P_2\mathbf{X})\|^2. \quad (1)$$

Here ϕ stands for *photo-inconsistency* while $|\mathscr{P}|$ is the number of points in $\mathscr{P}$. Difference of the colour values $\|I_1 - I_2\|$ can be defined by a number of different colour models. (For details see Jankó and Chetverikov, 2004b.) Finding the minimum of the cost function (1) over p_1 and p_2 yields estimates for the projection matrices.

The cost function (1) is robustified against occlusion and wrong measurements. Occluded points are eliminated by using the surface normals, and the outliers by rejecting a certain amount of the smallest and largest squares (α-trimmed mean technique).

In spite of the simplicity of the cost function $C_\phi(p_1, p_2)$, finding the minimum is a difficult task. Due to the 22-dimensional parameter space and the unpredictable shape of $C_\phi(p_1, p_2)$, the standard local nonlinear minimisation techniques failed to provide reliable results. We have tested a number of widely used optimisation methods: Newton-like methods, the Broyden–Fletcher–Goldfarb–Shanno (BFGS) variable metric method and the Levenberg–Marquardt algorithm. Experiments have shown that local search techniques terminate every time in local minima quite far from the expected global optimum.

A global nonlinear optimisation technique has also been tested. However, the stochastic optimisation method by Csendes (1988) did not yield acceptable results either. The randomness of a stochastic method is excessive, and it does not save nearly good solutions. In contrast, genetic algorithms preserve the most promising results and try to improve them. (Running a GA without elitism yields unstable and imprecise results, similarly to the stochastic optimisation.)

The methods mentioned above and other popular techniques, such as simulated annealing and tabu search process one single solution. In addition to performing a local search, simulated annealing and tabu search have specific built-in mechanisms to escape local optima. In contrast, genetic algorithms work on a population of potential solutions, which compete for survival. The competition is what makes GAs essentially

[2] The latter can be easily achieved if the images are taken by the same camera.