A Refined Algorithm (GA-Powell) for Camera Parameters Estimation and 3D Modeling Based on Silhouettes

[1]Amine Moua, [1]Rachid Benslimane and [2]Aziza El ouaazizi

[1]Laboratoire de Transmission et Traitement d'Information, École Supérieure de Technologie,
Université Sidi Mohamed Ben Abdellah, Route d'Imouzzer B.P. 2427, Fez, Morocco
[2]LIMAO, Faculté Poly-disciplinaire de Taza, Université Sidi Mohamed Ben Abdellah,
Route d'Oujda, B.P. 1223, Taza, Morocco

Abstract: In this paper, a hybrid method is applied to recover parameters and motion of camera from a set of silhouettes of an object taken under circular motion. Camera parameters can be obtained by maximizing the total coherence between all silhouettes. Two optimization methods, the Powell optimizer (PO) and the Genetic algorithms (GA), are applied to maximize the silhouette coherence and their performances are compared for several experiments. To take advantage of the strengths of the two methods, we developed a hybrid method that combines the genetic algorithm and the Powell optimizer to improve the performances in term of convergence speed and accuracy. The recovered parameters are used for 3D image-based modeling to obtain high fidelity 3D reconstruction.

Key words: Hybrid Genetic Algorithms · Powell optimizer · Silhouette coherence · Parameters estimation · Circular motion

INTRODUCTION

Acquiring 3D information from images has always been a hot research topic in 3D computer vision and recently, it has attracted more and more interest because of its potential applications such as computer games, augmented reality and cultural heritage preservation. In 3D computer vision, it is necessary to know the relationship between the 3D object coordinates and the image coordinates. This transformation is determined in the camera calibration step by recovering the camera intrinsic parameters and the relative pose of the camera.

Recovering camera parameters and motion from image sequences without using any calibration patterns can be classified into two approaches: the feature-based and silhouette-based approaches. In the feature-based approach, structure from motion algorithm [1] determines the camera parameters and the 3D structure of the object simultaneously from the feature correspondences [2, 3].

These methods would therefore be not applicable to smoothed objects with low texture. In addition to feature correspondences, silhouettes also offer important clues for determining both motion and shape. It is especially the case when the object being viewed is composed of non textured smooth surfaces like pottery and sculptures. For this kind of object, silhouettes are the most predominant and stable image feature.

Silhouette-based approaches generally exploit epipolar tangents [4, 5], to locate the images of the frontier points for deriving point correspondences between images. Hernandez *et al.* [6] considered the problem of recovering both the focal length and the camera motion under circular motion from silhouettes. They extended the idea of exploiting the epipolar tangents [5] to the concept of silhouette coherence, which measures how well a set of silhouettes corresponds to the projections of the visual hull. The author performed camera calibration by maximizing the silhouette coherence in optimization procedure.

Corresponding Aurthor: Amine Moua, Laboratoire de Transmission et Traitement d'Information,
École Supérieure de Technologie, Université Sidi Mohamed Ben Abdellah,
Route d'Imouzzer B.P. 2427, Fez, Morocco.

The Powell optimizer [7] was able to quickly reach the optimal solution for silhouette coherence maximization. However, we encountered difficulties in robustness when the initial guess for parameters are far away from the optimal solution and when the desired global maximum was hidden by many local maxima.

Genetic Algorithms (GA) [8, 9], are pseudo-stochastic search methods that derive their fundamental ideas and terminology from the Darwinian ''Natural selection'' theory, according to which individuals that are better fit to a given environment are more likely to survive. GA are problem-independent and can process information generated at previous stages of a search process. They comprise concepts such as natural selection, quick exploration and information collection in a design space. In contrast to most of classical optimization methods, GA require no initial guess for parameters and can avoid being trapped in local optimal solutions as shown in our previous work [10]. These characteristics make the GA powerful tools for solving optimization problems.

In this paper, two optimization methods, the Powell optimizer (PO) and a Genetic algorithm (GA), are applied to maximize the silhouette coherence and their performances are compared for several experiments. To take advantage of the strengths of the two methods, we developed a hybrid method that combines the genetic algorithms (GA) and Powell optimizer (PO) to improve the performance of the optimization procedure.

The remainder of this paper is organized as following: in Section 2 we present the circular motion parameterization. In section 3 we present the silhouette coherence measure and its practical implementation. In section 4 and 5, three optimization methods including a Powell optimizer, Genetic Algorithms and our hybrid GA-PO, are described, applied and compared for several tests in term of convergence and accuracy. In section 6 we build 3D models with the recovered parameters.

Circular Motion Parameterization

Camera Model: We consider a pinhole camera model. The geometry of a pinhole camera model is illustrated in Fig. 1. Let $M = (x, y, z)$ be a 3D point in an object frame and $m = (u, v)$ the corresponding image point in the image frame. The central projection of a 3D scene point M onto its 2D image point m can be written with the following linear equation using homogeneous coordinates:

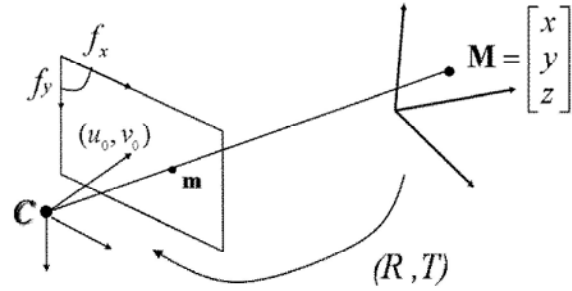$$m \approx PM \approx K\left[R|T\right]M \qquad (1)$$

Fig. 1: The geometry of a pinhole camera model

Where

$$K = \begin{pmatrix} fx & 0 & u_0 \\ 0 & fy & v_0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (2)$$

The projection matrix P is a 3×4 matrix defined up to a scalar factor that captures both the extrinsic and intrinsic camera parameters. R and T representing the rotation and translation between the world coordinate system and the camera coordinate system respectively. K is the camera calibration matrix. The parameters fy and fx represent the focal lengths measured in pixel units, with the aspect ratio defined as r = fy/fx, (u_0, v_0) represents the coordinates of the principal point.

In this paper, the aspect ratio is assumed to be one (r=1), the principal point (u_0, v_0) is considered to be the center of the image. The only intrinsic parameter that we consider is the focal length f.

Circular Motion Parameterization: Circular motion is a practical setup for image-based modeling. A circular motion image sequence can be obtained equivalently in two ways. The most common and the one used in our real image experiments, is the case of a static camera viewing an object rotating on a turntable. A second method is that of a camera rotating around a fixed axis and pointing at a static object. Figure 2 shows the 3D geometry of circular motion. The camera matrix P_1 of the first view can be written as:

$$P_1 = K\left[R_1|t_1\right] \qquad (3)$$

where K is the camera calibration matrix, R_1 and t_1 are the rotation and translation that transform the world coordinate system to the camera coordinate of the first view.
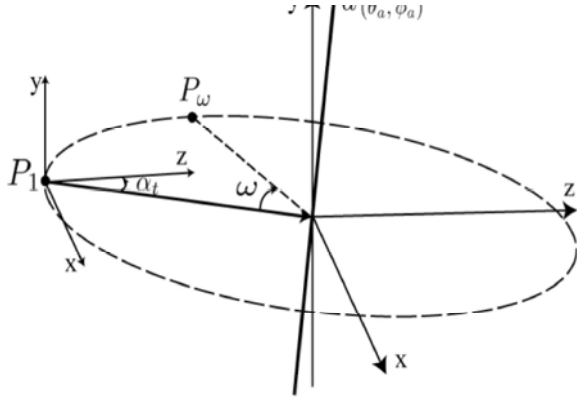
Fig. 2: circular motion parameterization

After rotating by ω about the axis $a(\theta_a,\ \varphi_a)$, the1 camera matrix P_ω of the second view can be achieved by post-multiplying $[R_1|t_1]$ with $R_a(\omega)$:

$$P_\omega = K\,[\,R_1|t_1\,]\,R_a(\omega) \tag{4}$$

Suppose that the circular motion image sequence consists of *n* views and the camera matrices for each view is denoted by $P_i\ i=1,…,n$, from (2) and (3) we have

$$P_i = P_1\,R_a(\omega_i) \tag{5}$$

where ω_i denotes the rotation angle between the ith and the first view, the rotation matrix $R_a(\omega_i)$ is written as a function of ω_i and the axis $a(\theta_a,\ \varphi_a)$ as follow:

$$
\begin{aligned}
R(\omega_i) = {} &(1-\cos\omega_i)\begin{bmatrix} a_x^2 & a_x a_y & a_x a_z \\ a_x a_y & a_y^2 & a_y a_z \\ a_x a_z & a_y a_z & a_z^2 \end{bmatrix} \\
&+ \begin{bmatrix} \cos\omega_i & -a_z\sin\omega_i & a_y\sin\omega_i \\ a_z\sin\omega_i & \cos\omega_i & -a_x\sin\omega_i \\ -a_y\sin\omega_i & a_x\sin\omega_i & \cos\omega_i \end{bmatrix}
\end{aligned} \tag{6}
$$

the rotation axis *a* is written in function of spherical coordinates $(\theta_a,\ \varphi_a)$:

$$a = \left(\sin\theta_a\cos\phi_a, \sin\theta_a\sin\phi_a, \cos\theta_a\right)$$

the translation is written in function of an angle α_t (the angle formed between the camera viewing direction and the z-axis see Fig. 2) as follow:

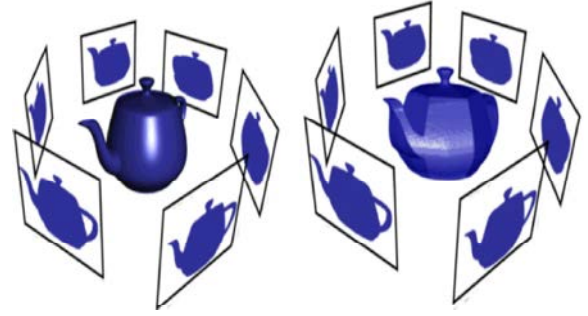$$t_1 = \left(\sin\alpha_t, 0, \cos\alpha_a\right)$$

Fig. 3: Visual Hull reconstruction from a set of silhouettes. Left: silhouettes obtained by projecting the original object back into cameras. Right: the reconstructed visual hull using these silhouettes.

For *n* views, we parameterize the circular motion with *n+3* parameters: the focal length *f,* the translation direction angle α_t, the rotation axis coordinates $(\theta_a,\ \varphi_a)$, the *n-1* camera angle steps ω_i. In this paper, our goal is to recover the projection matrices P_i of a set of silhouettes S_i of an object taken under circular motion as the set of n+3 parameters:

$$v = (\theta_a, \varphi_a, \alpha_t, \omega_i, f),\ i=1,…,n-1$$

Silhouette Coherence: Given a set of silhouettes S_i, $i = 1,…,n$ of a same 3D object taken from different points of view and the corresponding set of camera projection matrices P_i. Let *Vh* denote the reconstructed visual hull[1] using the set of silhouettes S_i and S^v_i denote the reconstructed visual hull silhouettes. We would like to evaluate the coherence between the silhouette S_i and all the other silhouettes $S_{j≠i}$ that contributed to the reconstructed visual hull *Vh*.

We assume that the silhouettes segmentation and the projection matrices are exact. We say that the silhouette S_i is coherent with all the other silhouettes $S_{j≠i}$ if the reconstructed visual hull silhouettes S^v_i and the original silhouette S_i are exactly the same ($S_i = S^v_i$). Two examples of coherent and non-coherent silhouettes are shown in Fig.4 and Fig. 5.

Measure of Silhouette Coherence: To evaluate the coherence between silhouettes, some kind of similarity measurement between the original silhouette S_i and the reconstructed visual hull silhouettes S^v_i is needed. Hernandez [6] defines this measure of coherence as the ratio of the silhouette contour lengths:
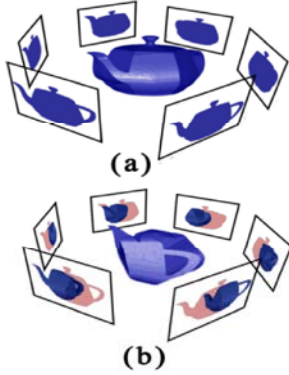
Fig. 4: Two examples of different silhouette coherence. (a) Perfect coherent silhouette. (b) Low silhouette coherence

Fig. 5: The original silhouettes Si superposed with the visual hull Silhouettes Svi. The red region indicates non coherent silhouette pixels.

$$C(S_i, S_i^v) = \frac{\int \left(\partial_i^v \cap \partial_i \right)}{\int \partial_i} \in [0,1] \qquad (7)$$

where ∂_i denote the contour of the original silhouette Si and ∂_i^v the contour of the reconstructed visual hull silhouette S_i^v.

To compute the total coherence between all the silhouettes, we compute the mean coherence of each silhouette with the (n-1) other silhouette [6].

$$C_{SC}(S_i,...,S_j) = \frac{1}{n} \sum_{i=1}^{n} C(S_i, S_i^v) \qquad (8)$$

If the silhouettes segmentation and the projection matrices are exact then:

$$C_{SC}(S_i,...,S_j) = 1 \qquad (9)$$

Optimization Routines: In order to exploit silhouette coherence for recovering camera parameters and motion under circular motion, the idea is to use the silhouette coherence measure C_{SC} as the cost in an optimization procedure. In this section, three optimization methods, including a Powell optimizer, Genetic Algorithms and a hybrid GA-PO, are applied to maximize C_{SC} (Eq. 7) and compared for several tests.

Powell Optimizer (PO): The first optimization method used here is the Powell direction set method. The Powell optimizer applied in this paper is the version described in [12, 13] in which starting points and a set of independent search directions are provided to the program. In each iteration the method serially performs a sequence of line minimizations along the various directions in the space of parameters. At the end of each iteration the method replaces one of the original directions with the line joining the starting and ending points. A special care is taken to ensure that the directions remain linearly independent. This version of the Powell optimizer is applied to the Silhouette coherence maximization problem. Although there are many other implementations of the PO such as described in [14, 15], the current work does not intend to include a comparative study of the merits of each of these implementations.

Genetic Algorithms (GA): When solving an optimization problem using GA, each solution is usually coded as an alphabet string of finite l ength called chromosome. Each string or chromosome is considered as an individual. A collection of N individuals is called population. GA start with a randomly generated population of size N, in each iteration of the algorithm, a new population of the same size is generated from the current population by applying operators, termed selection, crossover and mutation [16], that mimic the corresponding processes of natural selection. Following nature's example the probability p_m of applying the mutation operator is very low compared to the probability of applying the crossover operator p_c.

To improve the search process of the global optimum, an additional operator, elitism, was implemented. The aim of the elitist strategy is to carry the best chromosome from the previous generation into the next. We have implemented this strategy in the following way:

Step 1: Copy the best individual *ind₀* of the initial population *pop₀* in a separate location.

Step 2: Perform selection, crossover and mutation operations to obtain a new population *pop₁*.

$$\boxed{\theta_a} \; \boxed{\phi_a} \; \boxed{\alpha_t} \; \boxed{\Delta\omega_1} \; \boxed{\ldots} \; \boxed{\Delta\omega_{n-1}} \; \boxed{f}$$
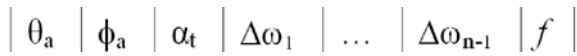
Fig. 6: Encoding of camera parameters as chromosome string

Step 3: Compare the worst individual ind_1 in p_1 with ind_0 in terms of their fitness values. If ind_1 is found to be worse than ind_0, then replace ind_1 by ind_0.

Step 4: Find the best individual ind_2 in pop_1 and replace ind_0 by ind_2.

Note that an individual ind_1 is said to be better than another individual ind_2 if the fitness value of ind_2 is less that of ind_1, since the problem under consideration is a maximization problem.

To adapt the GA to the camera parameters estimation problem, the real-valued Coding GA is utilized. The C_{SC} of the chosen Silhouettes is taken to be the objective function. The camera parameters are encoded as a chromosome string of n+3 genes as shown in Fig. 6. The alleles of each gene are constrained to a bound of values to ensure only feasible solution are adopted for evolution.

There exists no criterion in the literature [17], which ensures the convergence of GA to an optimal solution. But usually, two stopping criteria are used in Genetic Algorithms: In the first, the process is executed for a fixed number of iterations and the best individual obtained is taken to be the optimal one. In the second, the algorithm is terminated if no improvement in the fitness value of the best individual for a fixed number of iterations and the best chromosome is taken to be the optimal one. We have adopted the second stopping criteria with elitist strategy [18].

Hybrid Approach GA-PO: Although GA can quickly locate the region in which the optimal solution exists, it takes a relatively long time to converge to the optimal solution [10]. On the other hand, the Powell optimizer is known for its fast convergence speed but the correctness of solution is very dependent on the quality of the initial guess. Therefore, we exploited in the C_{SC} maximization problem the benefit of combining the Powell Optimizer (PO) and the GA. The proposed hybrid method consists of two steps. We start the search for good initial parameter values using GA followed by the refining process using PO in order to get more accurate solution.
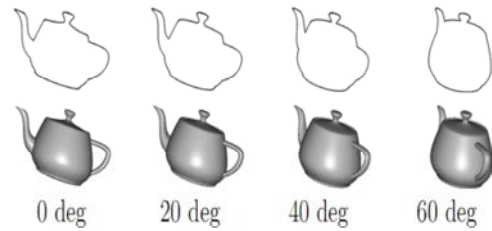

Fig. 7: Some views of synthetic Teapot sequence with their corresponding exact silhouettes and their absolute camera angles

In our Hybrid method, the Genetic Algorithm is responsible to provide good initial values of camera parameters, while the Powell optimizer is responsible for the quality of parameters estimation. The advantage of this hybridization is not only reduces the search space greatly, but also avoids premature convergence of Genetic Algorithm to some extent.

Experimental Results: In order to test the performance of our hybrid method, several parameters of the GA operators need to be determined. As recommended in [9] the crossover probability p_c was set to 0.6 and the mutation probability p_m was set to the inverse of the population size N. The determination of the population size N depends on the number of parameters to optimize. For a simple case (3 parameters), such as the estimation of focal length f and rotation axis coordinates (θ_a, φ_a), a population with N=50 is sufficient. However, for full motion estimation (n+3 parameters) a small population size can drive the GA to converge to a local maximum. To facilitate the implementation of our algorithm, we have used the GAlib version 2.4 developed by MIT.

Comparison between PO and GA: In this experiment, we have used a synthetic Teapot sequence of 18 exact silhouettes shown in Fig. 7. We conducted systematic comparisons between the PO described in [8] and of the GA to estimate the focal length f and the rotation axis coordinates (θ_a, φ_a). We found that in this case the PO performs better than the GA, as shown by comparison of the convergence histories in Fig. 8. Both the PO and the basic GA converge correctly to the optimal solution. However, the PO converges to the optimal solution more rapidly than the GA.

In the second test we have increased the complexity of the optimization problem, we have taken 9 views spaced of 20 degrees and we have computed the full
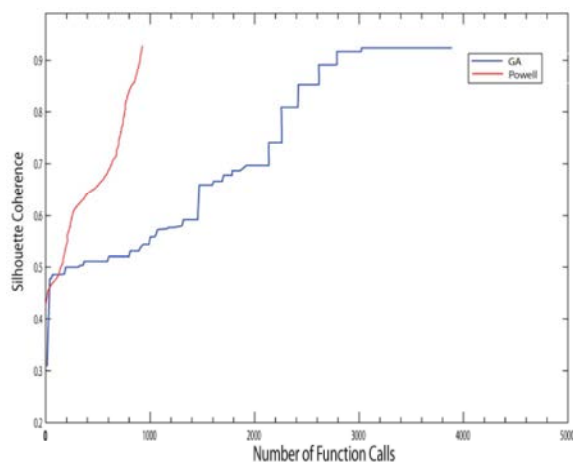
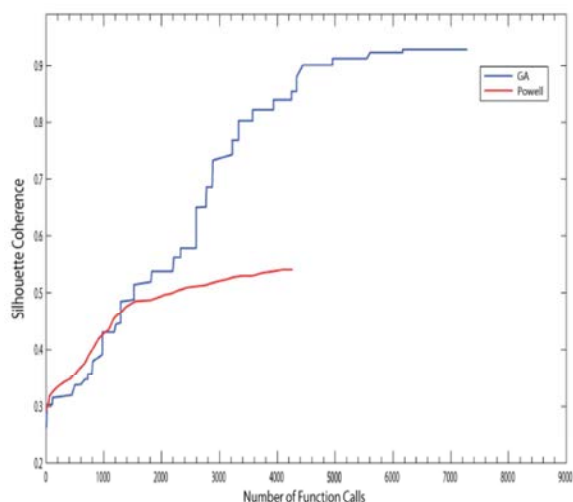Fig. 8: Comparison of the convergence histories of PO and GA for focal length and rotation axis estimation

Fig. 10: Same images of Hannover dinosaur sequence. From top to bottom: color images, binarized silhouettes and contours extracted from silhouettes

Fig. 9: Comparison of the convergence histories of Powell and GA for full motion estimation

Table 1: camera parameter estimated by ga and the proposed hybrid method

	Rotation			
	---------------------		Translation	Focal leght
Parameters	θa	φa	at	f
Ground Truth	92.663	2.261	2.735	3217
Lower range	-180	-90	-5	3000
Upper range	180	90	5	3400
Recovered by GA	92.405	2.272	2.774	3232
Recovered by Hybrid	92.658	2.257	2.743	3246

Comparison between the GA and Hybrid GA-PO: To demonstrate the advantage of the hybrid method over the previous one (basic GA), we applied it to recover camera parameters and motion from real silhouettes. In this experiment, we have used the Hannover dinosaur sequence shown in Fig 10. The dinosaur sequence (36 images) is binarized by a segmentation algorithm and then the contours are extracted from silhouettes using a GVF snake [19].

Table 1 gives the ranges of values for each parameter that we set for this experiment and the estimation of the rotation axis coordinates (θ_a, φ_a), the translation direction α_t and the focal length f by the GA and the Hybrid method. The results are good for both methods. The Hybrid method outperforms the basic GA when computing the rotation axis and the translation direction. The comparison of convergence histories between the both methods is shown in Fig. 11. It is seen that strong improvement is obtained when the PO is launched after 50 generations (1000 function calls). The hybrid method maximizes the C_{SC} faster than the basic GA.

circular motion (translation direction α_t and rotation axis coordinates (θ_a, φ_a), camera angles $\Delta\omega_i$) by keeping the focal length to its calibrated value. Fig 9 shows the convergence histories for the basic GA and the PO for this case. As shown in Fig 9 the GA converges correctly to the optimal solution while PO has been trapped in local optimal solution due to bad initialization of starting points. It is important to note that there are different implementations of the PO. Although the version applied in this paper fails to find the global maximum, there may be other versions of the PO that can improve the result. Though, seeking the best version of the Powell optimizer is not the intent of this paper.
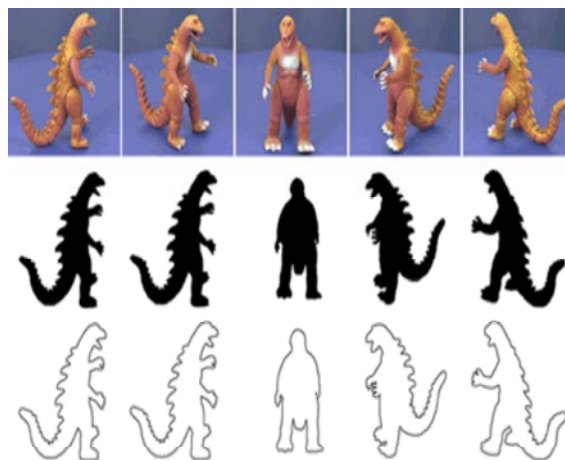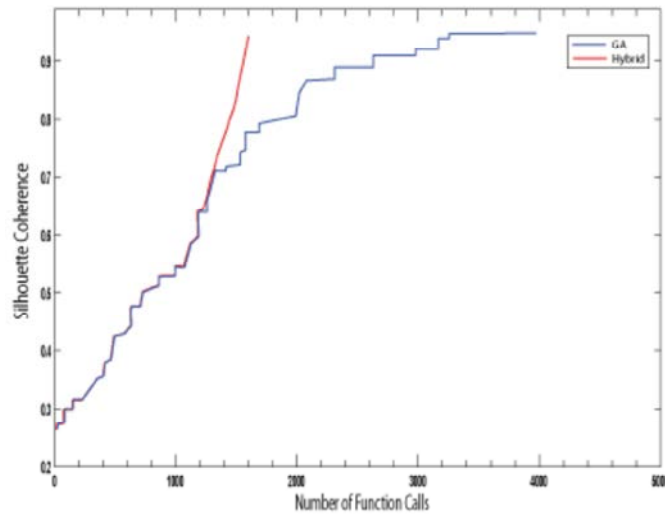
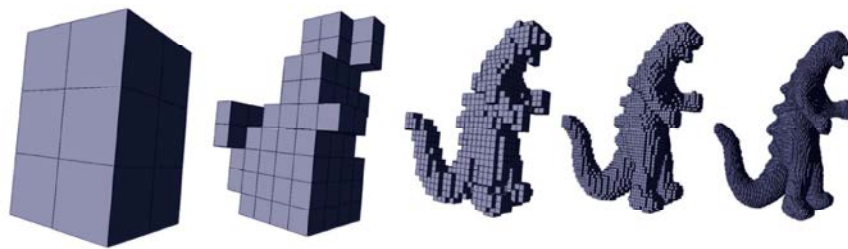Fig. 11: Comparison of the convergence histories of GA and our hybrid method

Fig. 12: Octree generation: the dinosaur octree is carved from a single bounding box given 36 images. From left to right: bounding box, level 3, 5, 6 and 7 of subdivision

Fig. 13: Dinosaur model: marching cubes meshing and decimation steps.

D Model Recostruction: The 3D object surface is determined by an octree based algorithm: We dispose of a set of 36 silhouettes (dinosaur sequence in Fig. 10) and their corresponding projection matrices Pi recovered by our hybrid GA-PO method. The algorithm needs two additional input data: the level of detail (the size of the voxel) and an initial bounding box. Starting from the bounding box, the octree approach subdivides a cube into 8 children whenever it is on the isosurface and iterates the process recursively until the maximum level of depth is attained. To evaluate a given cube, we project it into all the silhouettes to assign it one of the 3 labels [20] depending on whether it lies entirely inside (in), entirely outside (out), or partially intersects the silhouette (on). If the cube is on and the maximum depth is not still reached, we subdivide it and recursively test its children. At the end, only the cubes that are on surface have been subdivided. We can see the result of this step for different levels of resolution in Fig. 12.

Once the octree is constructed, the next step is to mesh it. Marching cubes algorithm [21] provides an initial consistent surface which is then smoothed using a decimation algorithm. Examples are shown in Fig. 13.

CONCLUSION

In this paper we developed a hybrid method GA-PO that combines the genetic algorithms and the Powell optimizer to maximize the silhouette coherence between a

set of silhouettes in circular motion. The Comparison between the GA and the PO shows that both methods converged correctly to the optimal solution. However, the GA was slower than the PO. Important improvements ware obtained with the hybrid method in term of convergence speed and parameters accuracy. The hybrid method can correctly find the optimal parameters without the need of initial values and successfully avoid to be trapped in local maxima. These characteristics will make the silhouette coherence concept more efficient and powerful to work in general motion instead of circular motion.

ACKNOWLEDGMENT

REFRENCES

1. Hartley, R. and A. Zisserman, 2000. Multiple View Geometry, Cambridge University Press.

2. Fitzgibbon, A.W., G. Cross and A. Zisserman, 1998. "Automatic 3D model construction for turn-table sequences," in 3D SMILE, pp: 155-170.

3. Jiang, G., H. Tsui, L. Quan and A. Zisserman, 2002. "Single axis geometry by ?tting conics," in ECCV, 1: 537-550.

4. Wong, K.Y.K. and R. Cipolla, 2001. Structure and motion from silhouettes. In 8th IEEE International Conference on Computer Vision, II: 217-222, Vancouver, Canada.

5. Paulo, R., S. Mendonca, Kwan-Yee K. Wong and Roberto Cipolla, 2011. Epipolar geometry from profiles under circular motion. IEEE Trans. Pattern Anal. Mach. Intell., 23(6): 604616.

6. Carlos Hernndez, Francis Schmitt and Roberto Cipolla, 2007. Silhouette coherence for camera calibration under circular motion. PAMI, 29(2): 343349.

7. Powell, M., 1964. "An efficient method for finding the minimum of a function of several variables without calculating derivatives", Computer Journal, 17: 155162.

8. Holland, J.H., 1975. Genetic algorithms, Scientif. Am., pp: 44.

9. Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization & Machine Learning (Addison-Wesley, Reading, MA, 1989).

10. Mouafi, A., R. Benslimane and A. El-Ouaazizi, 2013. A Genetic Algorithm for Recovering Camera Parameters and Motion from Silhouettes. Telecommunications (ICT), 20th International Conference on, 6-8 May 2013.

11. Matusik, W., C. Buehler, R. Raskar, S. Gortler and L. McMillan, 2000. Image-based visual hulls. SIGGRAPH, pp: 369-374.

12. Press, W.H., S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, 1992. Numerical Recipes, 2nd ed. (Cambrige Univ. Press, Cambridge, UK, 1992).

13. Acton, F.S., 1970. Numerical Methods That Work (Mathematical Association of America, Washington, DC, 1970) pp: 464. [1990 corrected edition].

14. Brent, R.P., 1973. in Algorithms for Minimization without Derivatives (Prentice Hall International, Englewood Cliffs, NJ, 1973), Chap. 7.

15. Dennis, Jr. J.E. and R.B. Schnable, 1983. Numerical Methods for Unconstrained Optimization and Nonlinear Equations (Prentice Hall International, Englewood Cliffs, NJ, 1983).

16. Michalewiez, Z., 1992. "Genetic Algorithms + Data Structure = Evolution Programs", Springer, Berlin.

17. El-Ouaazizi, A., M. Zaim and R. Benslimane, 2011. A Genetic Algorithm for Motion Estimation, International Journal of Computer Science and Network Security, 11: 4.

18. El-Ouaazizi, A., R. Ouremchi and R. Benslimane, 1998. Reconstruction of gray-level image by genetic algorithm , Proceeding of 4th International Conference on Quality Control by Arti?cial Vision, Japan 1998.

19. Xu, C. and J.L. Prince, 1998. Snakes, shapes and gradient vector flow. IEEE Transactions on Image Processing, pp: 359-369.

20. Szeliski, R., 1993. Rapid octree construction from image sequence. CVGIP, 58(1): 23-32.

21. William E. Lorensen and Harvey E. Cline, 1987. Marching cubes: A high resolution 3D surface construction algorithm, ACM Computer Graphics, 21(24): 163-169.