Monitoring of the visual arts degrade by means of the active vision system "3EYES"

D. Cojoc(1), P. Grattoni(2), R. Nerino(3), G. Pettiti(2)

(1) University "Politehnica" of Bucharest, Romania,
(2) Centro di Studio per la Televisione - CNR ,
(3) Istituto Elettrotecnico Nazionale Galileo Ferraris,
Strada delle Cacce, 91
10135 Torino - Italy
E-mail: grattoni@cstv.to.cnr.it

Abstract

Monitoring of some physical characteristics (colour, geometry, size of fissures, ...) of object surfaces is of utmost importance in the field of conservation and restoration of work of arts. In order to get effective results a very high repeatability in the localization of the test points is required. At present, repositioning is a tricky task and is performed by means of manual empirical procedures which can be time consuming and expensive. In this short paper an application of the 3EYES active vision system to the automation of the monitoring is presented. An application of the system to object monitoring in laboratory environment is described.

Keywords: active vision systems, fixation, degrade monitoring, key points, Gabor wavelets, pose estimation

1. Introduction

Monitoring of some characteristics (colour, geometry, size of fissures, ...) of work of arts is of utmost importance in the field of the conservation and restoration.

This task usually involves the execution of measures only over small and well defined areas of the subject (e.g. some squared centimeters) and, sometimes, in almost uniform areas. In addition, results can be heavily influenced by the spatial measuring position, so that very different values will be obtained from different measures on slightly shifted areas even if they are performed closed in time. At present, repositioning is a tricky task and is performed by means of manual empirical procedures (e.g. multiresolution photos) which can be time consuming and expensive [1].

Recently, in the framework of the project P.F. Beni Culturali of the italian C.N.R., an active vision system (called 3EYES)[2] has been designed for high fidelity data acquisition of 3D painted surfaces (frescoes, mosaics, ...). It allows the automatization of the monitoring process. The 3EYES system is composed by three TV cameras (see Fig. 1). Two of them (TL1 and TL2) are equipped with long focal-length lenses to frame parts of the scene at the desired resolution (in the following we will call them teleobjectives). The third one (WA) is equipped with a wide angle lens allowing to frame the whole region of interest of the scene at a time.

The general working principle is as follows. Once the whole scene has been framed by the WA camera, its image is processed to detect and localize regions of interest according to the aim of the current analysis. Then, the results are used to drive the teleobjectives to converge on these regions and to measure their physical and photometric

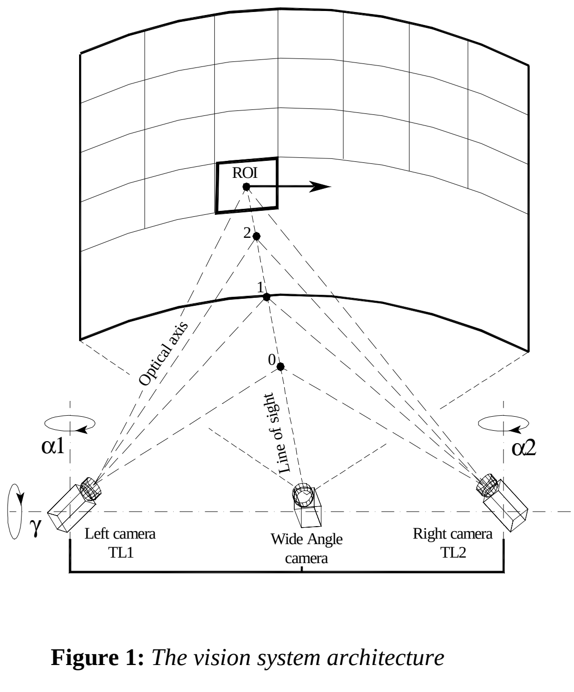

Figure 1: *The vision system architecture*

Published by the Eurographics Association ISSN 1017-4656

properties. The TLs are also used to acquire image details.
To carry out these operations automatically, the system is designed to position the cameras by means of three computer-controlled angular movements. Two independent "pan" movements (α1 and α2) allow TL1 and TL2 to perform the fixation task, while the same tilt movement γ is provided for all the three TV cameras at once.
Depending on the purpose of the analysis, the TLs functions can be carried out by using B/W only, RGB or B/W TV cameras equipped with spectral filters.
Thanks to these features, the system can be used for different applications [3]; in this paper we will describe its use in visual art monitoring.

2. The monitoring procedure

3EYES is placed in front or under the subject, at a suitable distance, and the system is rotated to frame the whole surface of interest by the WA camera; then it starts the measuring procedures.
The first time the 3EYES system analyses a subject, it automatically builds up a description of the subject in terms of characteristic structures with attributes. This allows to establish the relative position between the painted surface and the system (3EYES and illuminating lamps) in a subject centered co-ordinate system, accurately.
The description is carried out in two steps. At first, the WA image is processed in order to get a raw geometric description of the subject; contours, uniform regions and key points can be detected to this purpose. Their spatial location is then measured with high accuracy by means of the two teleobjectives in order to establish a subject-linked reference system.
Then, the measure on the selected test points is performed by a human operator who drives the system to fixate them interactively.
For every test point, its relative angular position with respect to one of the reference (key) points (typically the nearest one) is stored together with the results of the measurement. This is used as an address for future comparisons.
When a comparison with formerly stored measurements is necessary to monitor possible time varying parameters, the system is set again in front of the subject and a new description of it is carried out in the same way as done at the first time.
The relative position of the system in this new setting, compared with the first one taken as reference, allows to compute the roto-translation parameters relating the new position to the initial one in such a way to recover the former measuring conditions. At this point, the system automatically repeats the recorded test. First it converges to fixate the reference point nearest to the test point, then it moves of the formerly stored angular displacement to the test point and performs the desired measures.
Consequently, the reposition accuracy, which is one of the most important requirements for the comparison reliability, relies both on the accuracy of the system in structure localisation, and on the repositioning procedure.

2.a Key point detection and characterization

Measurement of the 3D position of a feature in the space is carried out by the 3EYES system by triangulation (i.e. by intersecting the optical axes of TL1 and TL2 when fixating the feature under examination). Then, the accuracy of the dimensional measurement relies on the accuracy of the fixation process. 3EYES performs this task at first by correlating the images of TLs with the part of the WA image corresponding to the feature of interest for a first raw convergence, and then by cross-correlating the two TLs images directly for the fine registration.
Obviously, not all the points of an image are useful for this purpose. Only those points that have significant locally maximal variation of luminance in more than a single orientation (e.g. various types of junctions, grey level corners, line termination, spots), usually denominated as '*key points*', should be selected.
In order to detect the key points, a method based on the computation of the 2D local energy map (LEM) of an image [4] has been used in this application. The 2D LEM is calculated in two steps:
1 - calculation of the 2D oriented energy maps (OEM) corresponding to N angular orientations. This is achieved by filtering the image with a pair of even- and odd- Gabor functions (G_e, G_o) at first, and then by filtering again the output with (Ge, Go) rotated by $\pi/2$ to get the 2D OEM
2 - calculation of the 2D LEM by summing the 2D OEMs over all the N angular orientations.
The key points are then extracted as local maxima from the 2D LEM.
The pair of 2D Gabor functions used to filter the image is given by the following equations:

$$\mathbf{G_{e,o}(x,y,f,\theta_i)} = \mathbf{exp}\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right]\mathbf{cos}\left[2\pi f\left(x\cos\theta_i + y\sin\theta_i\right)+\phi\right]$$

where f is the frequency, θ_i denotes the angle for orientation i, and $\phi=0$ for G_e and $\phi=\pi/2$ for G_o.
The even (e) and odd (o) Gabor functions defined in the former equation represents a quadrature pair of functions having zero mean, identical L2 norm, and being orthogonal to each other. Gabor have been preferred among others because they allow the signal representation in terms of amplitude

functions that are localized both in space and frequency.
To solve the problem of scale invariance a multiple scale analysis by means of Gabor wavelets has been used (i.e. a family of functions that includes various scaled versions of the basic Gabor functions). The modulus of wavelet transform has been measured for various scales of the Gabor filter. This approach also allows an estimation of the reliability of 2D features in regard to correlation and, as it has been proved elsewhere [5], it can be used to remove the effect of noise from images.

2.b Geometric description of a scene

In this context, the goal of the geometric description of a scene is to establish the relative attitude of the 3EYES system with respect to a subject-centered co-ordinate system, to allow an accurate recovery of the initial measuring setup. At present, this description is based on the spatial (3D) distribution of the key points detected by the method described in the previous subsection. The observation that the key points belong to a rigid body and hence their relative positions do not vary from one view to the other, suggests a global approach to the estimation of the relative attitude changes which does not require to solve the matching problem between couple of corresponding points in the two poses.
Indeed, the center of mass and principal axis of inertia of a rigid body, globally evaluated over all the K points, are invariant to translation and rotation relative to a body-referenced co-ordinate system. The relative translation of the centre of mass and the relative rotation of the axis of inertia can be estimated from one pose to the other one by the relative positions of k key points $\overline{x}_{i\,1,2} \equiv (x_i, y_i, z_i)_{1,2} \quad i = 1,..,k$ referred to the co-ordinate of the geometric center of mass m_{O1}, m_{O2} associated to the same points in the two poses as expressed in the reference frame O1 and O2. These relative positions are used to built the two moment matrices S_1 and S_2 from measurement matrices M_1 and M_2 :

$$M = \begin{bmatrix} x_1 & y_1 & z_1 \\ \cdot & \cdot & \cdot \\ x_K & y_K & z_K \end{bmatrix} \quad ;$$

$$S_1 = M_1 \cdot M_1^T \quad ; \qquad S_2 = M_2 \cdot M_2^T$$

The singular value decomposition (SVD) of the symmetric matrices S_1 and S_2 gives :

$$S_1 = M_1 \cdot M_1^T = Q_1 \cdot W_1 \cdot Q_1^T \quad ;$$

$$S_2 = M_2 \cdot M_2^T = Q_2 \cdot W_2 \cdot Q_2^T$$

The principal axes (PA) orientation associated to the two matrix S_1 and S_2 correspond to the orthogonal matrices Q_1 and Q_2 . The first group of PAs is rotated respect to the other of the same amount of the rotation R between O1 and O2. Then the relative rotation matrix R between the two poses is given by $R = Q_1 \cdot Q_2^T = Q_1 \cdot Q_2^{-1}$, while the relative translation vector between the two poses is given by the relative translation of the two centres of mass, that is : $t = m_{O2} - m_{O2}$

3. A laboratory application

In order to test the proposed monitoring procedure, a set of measures has been carried out in laboratory environment on a 67×100 cm poster. Since the last version of 3EYES is not fully operative yet, a lower accuracy prototype has been used. The poster has been attached to a curved support to simulate the possible shape of a chapel in a suitable scale measure. The test has been carried out as follows. First the poster was placed in front of the 3EYES system at a distance of about 2750 mm, and the WA image was acquired. Using the technique described in Subsection 2a, the key points were detected in the WA image (see Fig. 2). Then, their spatial positions were measured using the triangulation method.

Figure 2: *The WA image with key points (+) and test points (o) superimposed*

The geometric description of the scene was carried out by using the method described in subsection 2b and a poster-centered co-ordinate system was established.
At this point, an operator interactively drove the system to point and acquire the luminance of three test points. The spatial position of these points both in the poster-centered coordinate system and with respect to the nearest reference point were measured and saved. Then the poster-to-3EYES relative position was changed significantly and a new description was carried out using the same procedure.
Comparing the first description, used as the reference one, with the new description we were

able to recover the initial measuring position with a relative error of about 1%.
At this point, the measures were automatically repeated: the description at first and then the pointing of the selected areas. The accuracy of the repositioning procedure was estimated on the base of the angular offset between the crosses of reference and the image center. In Figure 3 the offset positions is shown for one of the three test regions. The angular repositioning error for this three test points were 270 μrad, 120 μrad and 220 μrad respectively.

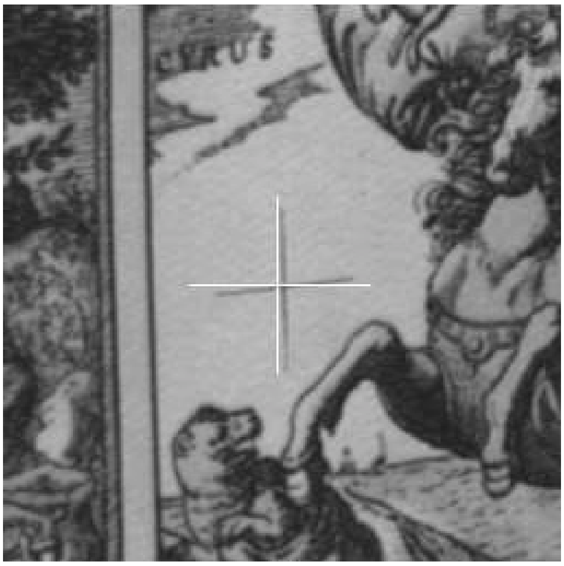

Figure 3: *The repositioning accuracy over one of the monitored regions is shown by the superposition of the white crosses (image centers) to the dark crosses (test points)*

4. Concluding remarks

Obviously this test is too limited to draw conclusions, it can just be used as an indication of validity of the proposed monitoring method. Nevertheless we can make the following remarks. To allow a quantitative evaluation of the system performances it must be outlined that the size of a measuring area for spectral analysis could be of the order of some 10 mm^2 and the reposition accuracy should be of the order of about 1 mm. This means that using the worst result of the previous test (270 *μrad)* this condition is obtained at 3.7 m away from the system. Using the new 3EYES system under development we expect to get the same result at a distance greather than 10 m which is good in many applications.
From the point of view of image processing, the most fragile point of the chain is the geometric description of the scene. Its task is to define a very precise and repeatable reference system on the subject, independent from the viewing point of the scene, in such a way to allow to establish an accurate geometric correspondences between different views, taken at different times.
This aspect involves the well known problem of the search for non-ambiguous feature correspondences in stereo and motion applications. But this should not be a problem since the matching reliability is mainly a matter of software complexity and the processing time is not a severe constraint in this kind of application. Anyway, this will be one of the topics of future developments.

References

1. M. L. Rastello, *Works of Art and Colour Rendering Properties of Light*, 1st Int. Congress on: “Science and Technology for the Safeguard of Cultural Heritage in the Mediterranean Basin”, Nov. 27 - Dec. 2, 1995, Catania, Italy.
2. P. Grattoni, R. Nerino, G. Pettiti, M. L. Rastello, “*Virtual Museum: an Active Vision System for the recording of frescoes and mosaics*”, Proceedings of EVA’96 Conference, Florence, Italy, 9 - 10 Feb., 1996
3. D. Cojoc, P. Grattoni (*), R. Nerino, G. Pettiti, *The 3EYES Active Vision System in Metrological Applications*, 5th Symposium of Optoelectronics SIOEl’98, September 23-25, 1998, Bucharest, Romania.
4. B. Robbins, R. Owens, *2D feature detection via local energy*, Image and Vision Computing 15, pp. 353 – 368, 1997.
5. S. Mallat, W. L. Hwang, *Singularity detection and processing with wavelets*, IEEE Transactions on Information Theory, vol. 38, No. 2, pp. 617 – 643, 1992