SYNTHESIS OF VIRTUAL VIEWS USING NON-LAMBERTIAN REFLECTIVITY MODELS AND STEREO MATCHING

Federico Pedersini, Augusto Sarti and Stefano Tubaro

Dipartimento di Elettronica e Informazione, Politecnico di Milano
Piazza L. da Vinci, 32, 20133 Milano, Italy
`pedersin, sarti, tubaro@elet.polimi.it`

ABSTRACT

A technique for the synthesis of virtual views of a 3D scene, starting from images taken by a calibrated multicamera system, is proposed and tested.

Surface interpolation is performed over a set of 3D edges, computed with stereometric algorithms, and additional curvature-tuning points, scattered where the reflectivity model is sufficiently reliable. The 3D coordinates of the tuning points are computed by minimizing the MSE between the available real views and the corresponding synthetic views. Synthesis is finally carried out by reprojecting on the new image plane the estimated object surface over which texture-mapping of the reflection-corrected luminance function has been performed. Texture correction, which uses an estimate of a non-Lambertian reflectivity model, is done in such a way to simulate the migration of reflections due to the change of viewpoint.

The technique has been tested on real images, producing realistic synthesized views.

1. OVERVIEW

The synthesis of arbitrary views from a set of images taken from different angles is a problem of particular interest in a number of applications such as virtual object manipulation (3D catalogs, tele-surgery, etc.) and 3D television.

Several strategies can be used for the synthesis of a missing view. In general, they can be divided into two large categories: those that operate an "interpolation" between the available views and those that first reconstruct the 3D object surface, and then reproject it onto the image plan of the "virtual" camera. The work we present here uses the second approach. In fact, we extract geometric (stereo correspondence) and radiometric (shading) information from the available views for 3D surface reconstruction, and then use the information on reflectivity and illumination for realistic rendering on the virtual camera.

The complementary nature of stereo and shading has been discussed in several previous articles [3, 6]. It is well known, in fact, that intensity correlation performs better on highly textured regions of the input images, whereas the accuracy of shading is higher on regions corresponding to more regular surfaces.

The information we use for reconstruction comes from a calibrated multicamera (trinocular) system. Stereometric matching is performed between edges, and triangulation allows us to locate such edges in a 3D space with rather good precision. This procedure provides us with an irregular "mesh" of 3D details, but does not provide any clue about the local curvature of smooth surfaces, which is exactly the shape information that can be extracted from shading.

Once the set of 3D edges and points is available, we can proceed with the construction of a first approximation of the 3D object surface. Such a surface can thus be used for a preliminary estimate of the parameters of surface reflectivity and illumination and a measure of their reliability. We then use the reliability mask to rule out regions of the surface over which the reflectivity model is not dependable and, by keeping only the reliable portions of the object surface, proceed with a refinement of the radiometric parameters.

Shading is taken into account by introducing a set of curvature-tuning points, scattered inside the reliability mask of the radiometric model. The 3D coordinates of all tuning points are computed by minimizing the MSE between the available real views and the corresponding synthetic views. A better approximation of the object surface is now possible by using both matched edges and tuning points.

Last step of our technique for the synthesis of a virtual view is the rendering process, which takes into account the fact that reflections may depend on the position of the viewer. In fact, we reproject on the new image plane the estimated object surface over which texture-mapping of the reflection-corrected luminance function has been performed. Texture correction, which

0-8186-7310-9/95 $4.00 © 1995 IEEE

uses an estimate of a non-Lambertian reflectivity model, is done in such a way as to simulate the migration of reflections due to the change of viewpoint. In other words, we modify the image texture through a compensation of the non-lambertian component followed by a simulation of the reflections that would be visible from the virtual viewpoint.

2. RECOVERY OF 3D INFORMATION FROM STEREO

Stereo matching techniques compute, through geometric triangulation, the 3D coordinates of details that originate corresponding edges on two views. Since binocular vision does not guarantee a unique determination of correspondences in a complex scene, we use a set of three cameras, as shown in Figure 1. Trinocular vision, in fact, allows us to select the best pair of cameras for a specific correspondence between elements of two images and to validate this correspondence through a check on the third view [1].

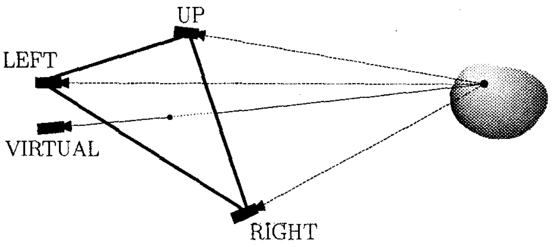


Figure 1: Trinocular camera system and virtual camera.

Preliminary calibration is performed in order to estimate intrinsic (focal length, optic distortion) and extrinsic (relative position of the cameras) parameters of the trinocular system [4, 7]. The 3D edge positioning requires an accurate edge extraction from the three images. This task is performed by a modification of the Canny edge detector [5]. In order to match corresponding edges, an improved version of a technique proposed by Ayache [1, 2] has been developed [7]. Once the correspondences have been found, the construction of the depth map can be easily performed through geometric triangulation.

The final result of this system is an irregular set of 3D edges whose accuracy critically depends on the quality of the calibration procedure and whose density depends on the degree of smoothness of the surfaces of the scene. Highly textured objects, in fact, will produce denser 3D meshes of edges.

3. SHADING-DRIVEN INTERPOLATION

The set of 3D edges determined through stereo matching, is often too sparse for a reliable surface reconstruction. This happens especially when dealing with surfaces that are not very textured. For this reason, wherever a dependable reflectivity model is available and the 3D edges are too sparse, we exploit the radiometric properties of the surface and extract additional information from shading.

The surface reflectivity of real images, besides depending on the direction of illumination, changes according to the viewing direction. Lambertian surfaces, in fact, are very rare, and specular reflections are practically always present. For this reason we use a non-Lambertian radiometric surface description based on a model developed by Torrance and Sparrow [8]. The reflectivity function we use has the following form

$$R = A\cos\vartheta_i + K\frac{e^{-\frac{\alpha^2}{2\sigma^2}}}{\cos\vartheta_r} + D \tag{1}$$

where ϑ_i is the angle between surface normal and incident light, ϑ_r is the angle between surface normal and viewing direction and α is the angle between the surface normal and the plane corresponding to incident light and viewing direction. The first term represents the Lambertian component, the second term is the specular reflection and the third one is an extra constant that we included to take diffused light into account.

The above radiometric model allows us to describe realistic surfaces and illumination conditions. In fact, it is suitable for modeling diffuse and specular reflections produced by one dominant source of light combined with diffused light. Such a model is fully described by A, K, D, σ and the direction of dominant light.

Surface interpolation is performed over the previously determined 3D edges and an additional set of curvature-tuning points. The tuning points are automatically allocated in two steps: first a reliability mask is built for the reflectance model, and then the points are scattered inside the reliable areas in such a way not to lie close to matched edges.

The 3D coordinates of the added points are computed through the minimization of a cost function. Since the ultimate goal is to synthesize virtual views, the cost function chosen to be the MSE between the available real views and the synthetic images in corresponding position. The synthesis is carried out by using an estimate of a viewer-dependent reflectivity model. Notice that, since the reflectivity model is non-Lambertian, the minimization can be performed on all available views, as each of them provides some extra information.

4. RENDERING

The synthesis of the virtual view is performed by using both shape information and radiometric properties of the surface. Since the surface reflectivity is, in general, non-Lambertian, specular reflections can occur; this means that the surface luminance on the virtual view depends on the viewing direction. In fact, observing a non-opaque curved surface from a moving viewpoint, we see that all reflections move accordingly.

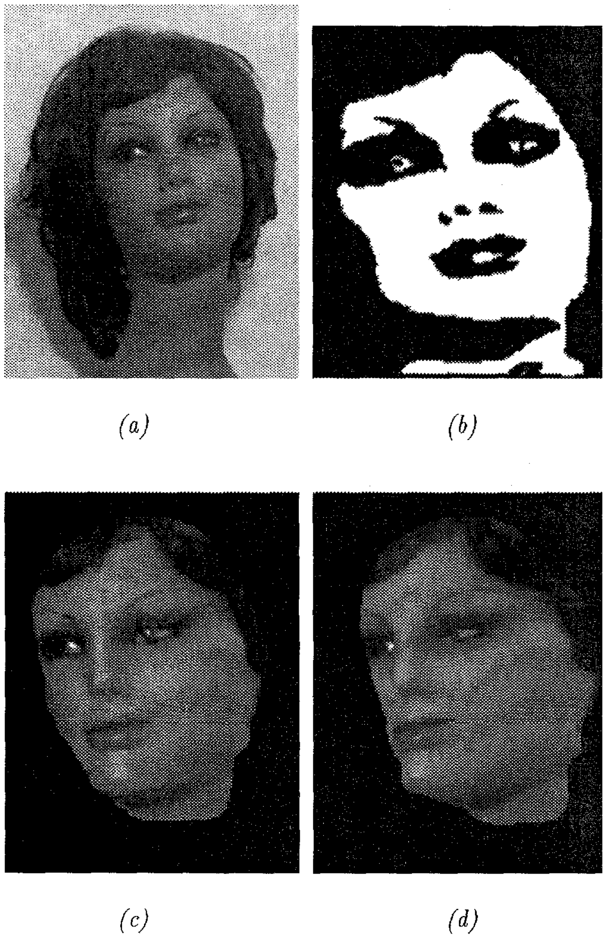
In order to take this effect into account we first eliminate the specular component of eq. 1 from the surface luminance map, then we add the specular component corresponding to the virtual viewing direction. Such an operation is possible as all parameters of the radiometric model have already been estimated before the shading-driven interpolation.

Texture map of the reflection-corrected luminance function is finally performed over the reconstructed 3D surface, and its reprojection onto the image plane of the virtual camera produces the desired view.

5. TESTING AND RESULTS

Tests have been performed on several real images. In Fig. 2a, one of the three views of the face of a dummy is shown. The correspondent reliability mask of the adopted non-Lambertian reflectance model results as in Fig. 2b. The virtual viewpoints in Figs. 2c and d are intentionally chosen to be far away from the triangle of the three cameras (see Fig. 1).

In order to measure the accuracy of the 3D reconstruction algorithm we also used a triplet of images of a ring-shaped portion of circularly symmetric Styrofoam smooth surface of known shape. The shading-driven interpolation of all matched edges produces a surface whose horizontal and vertical sections are shown in Figs. 3a and 3b, respectively. Fig. 3 shows how the curvature of the reconstructed surface (solid line) approximates acceptably well the actual curvature (dotted line). In general, the accuracy of the 3D reconstruction depends on how well the reflectivity model describes both surface and illumination conditions. However, since the minimization is performed on the MSE between real and synthetic views, it is reasonable to expect that the synthesis of the virtual view will give rise to good results even when the accuracy of the reconstruction is not very high. In Fig. 4 the actual view of the test shape and two virtual views, respectively, are shown.


(a) (b)

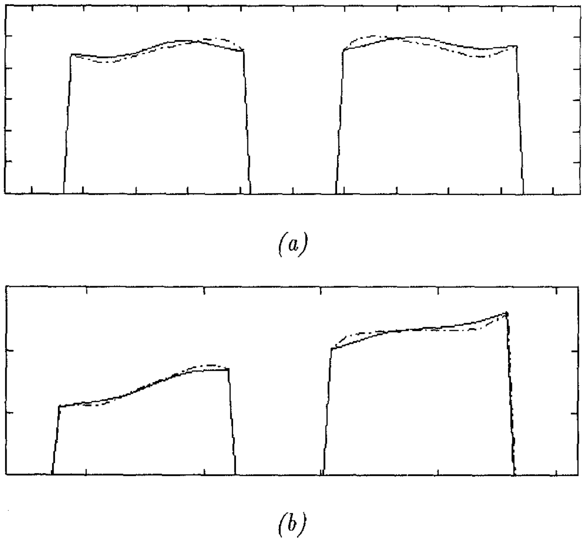
(c) (d)

Figure 2: Reconstruction of the face of a dummy. Original image (a), reliability mask of the reflectivity model (b), virtual views (c,d).

6. CONCLUSIONS

We have presented a technique for the synthesis of virtual views from stereo correspondences and shading, using the output of a calibrated trinocular vision system. The key point of the method is the way texture mapping is performed before reprojection. In fact, a non-Lambertian reflectivity model of the surface is used for computing the luminance that would be seen from the virtual viewpoint. The obtained luminance map is the texture we map onto the surface before reprojection. The method has been successfully tested on real images.

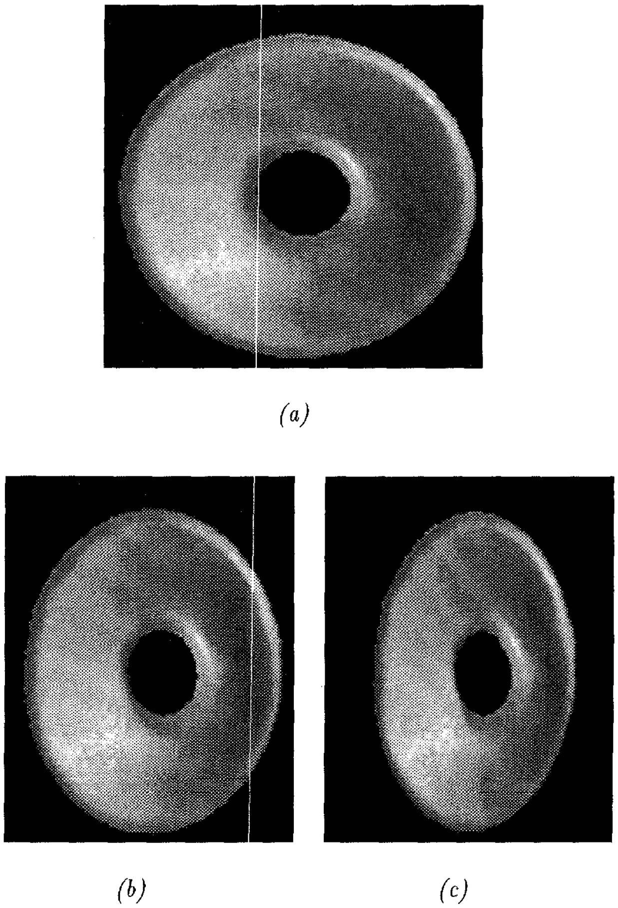
Further improvements are currently being made on the above technique. In particular, preliminary piece-


(a)

(b)

Figure 3: Section profile of a Styrofoam shape (dotted line) and its reconstruction (solid line). Horizontal section (a), vertical section (b)

wise smooth segmentation of surfaces and a "smart" placement of curvature-tuning points are under study.


(a)

(b) (c)

Figure 4: Original view (a) and two virtual views (b,c) of a Styrofoam shape

7. REFERENCES

[1] N. Ayache: "Artificial Vision for Mobile Robots". MIT Press 1990.

[2] N. Ayache, F. Lustman: "Trinocular Stereo Vision for Robotics". In IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 13, No. 1, pp. 73-85, Jan. 1991.

[3] A. Blake, A. Zisserman, and G. Knowles. "Surface descriptions from stereo and shading". Image Vision Computation, Vol. 3, No. 4, pp. 183–191, 1985.

[4] S. Brofferio, F. Pedersini, S. Tubaro, "Calibration of Trinocular System for 3D Measuremts". Proc. 4-th European Workshop on Three-Dimensional Televison (COST 230), Rome, Italy, October 20-21, 1993, pp. 89-96.

[5] J. Canny: "A computational approach to edge detection". IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 8, No. 6, pp. 679-698, Nov. 1986.

[6] Y.G. Leclerc and A.F. Bobick. "The direct computation of height from shading". Conference on Computer Vision and Pattern Recognition, Lahaina, Maui, Hawaii, June 1991.

[7] F. Pedersini, S. Tubaro, F. Rocca: "Camera Calibration and Error Analysis, An Application to Binocular and Trinocular Stereoscopic Systems". Proc. Int. Workshop on Time Varying Image Processing and Moving Object Recognition, Florence, Italy, Jun. 10-11, 1993.

[8] K.E. Torrance, E.M. Sparrow: "Theory for Off-Specular Reflection from Roughened Surface for Ray Reflection". In Journal of Optical Society of America, Vol. 65, pp. 531-536, 1975.