



HAL
open science

Virtual camera synthesis for Micromanipulation and Microassembly.

Julien Bert, Soukalo Dembélé, Nadine Lefort-Piat

► **To cite this version:**

Julien Bert, Soukalo Dembélé, Nadine Lefort-Piat. Virtual camera synthesis for Micromanipulation and Microassembly.. IROS'06, Oct 2006, Beijing, China. pp.1390-1395, 10.1109/IROS.2006.281928 . hal-00333560

HAL Id: hal-00333560

<https://hal.science/hal-00333560v1>

Submitted on 23 Oct 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Virtual Camera Synthesis for Micromanipulation and Microassembly

Julien Bert, Soukalo Dembélé, and Nadine Lefort-Piat

Laboratoire d'Automatique de Besançon

UMR CNRS 6596 - ENSMM - UFC

25000 Besançon, France

Email: {jbert, sdembele, npiat}@ens2m.fr

Abstract—The vision system for micromanipulation and microassembly usually includes at least two cameras allowing top and lateral views of the work field. The top view is used to control the xy position of the microgripper. As for the lateral view, it allows the control of the z position. The paper describes how the lateral camera can be replaced by a virtual one using the trifocal transfer. Then, the work field is set free. The novel view synthesis method used requires only a weak calibration to recover the geometry of the scene, then, it is well suited for the vision at the microscale.

I. INTRODUCTION

Micromanipulation is the manipulation of parts at the microscale, i.e. in the range from $1\ \mu m$ to $1\ mm$, for assembly, sorting or testing. In addition to biomicroparts like cells and pollen seeds, artificial microparts are chemically or mechanically synthesized, or micromachined. Classical examples of the first and second types are respectively grains of powder like drugs or cosmetics, and optomechanical components like balls, pegs, pins, threads, membranes, lenses, shutters and fibres. In some cases these microparts define final products (MEMS), otherwise they must be assembled to lead to the final products. For that purpose some automated microassembly systems have been developed by [1], [2], [3] and [4]. From those results it can be noticed that a microimaging system is always required, and the most used is the photon microscope connected to a camera. The images and their processing and analysis allow the task surveillance, system control or microparts recognition. The field-of-view of the microscope is very narrow that leads to the use of multiple views imaging : global view (usually at the top), left and right

lateral views. The second reason of multiple views use is the fact that top view only allows the access to the xy position of the microgripper. Lateral view is required to get the z position. The third reason is the occurrence of components occlusions during assembly, the microgripper can hide the microparts to pick. However multiple views imaging has a drawback, the microimaging component cannot be positioned anywhere, so some views are not accessible. Sometimes, it is also useful to set free the work field. A view from a virtual imaging system using a novel view synthesis method can allow to overcome that problem.

Novel view synthesis (NVS) is a discipline of computer vision introduced by [5] which deals with the obtaining of a maximum of views of an environment with a minimum of real data of it. For example, a lateral view of an object can be synthesized with only two real top views. There are two classes of methods in NVS: the model-based rendering and the image-based rendering (Fig. 1).

In model-based rendering (MBR), virtual environments are created from mathematical models. A typical example is 3D characters synthesis in movies and video games by modeler softwares. In image-based rendering (IBR), a set of real images of the scene are used to build a novel view. According to the knowledge about scene geometry, [6] proposes the following classification: rendering with no geometry, rendering with explicit geometry and rendering with implicit geometry. Rendering with no geometry i.e. no calibration is used to create a mosaic from a set of local views that leads to a novel global view [7], [8]. Rendering with explicit geometry i.e. with precise calibration is close to MBR. Its purpose is the reconstruction of a 3D view from real views of the scene [9]. This technique needs a precise calibration and is computationally expensive.

Rendering with implicit geometry only needs a weak calibration. Reference [10] presents three techniques of NVS of that type: the line of sight, the epipolar transfer and the trifocal transfer. The line of sight approach is based on ray-tracing [11]. Its drawback is the fact that it needs at least ten images to obtain a synthetic view. The epipolar transfer approach was introduced by [12], it is based on epipolar geometry where the epipolar constraint defines the point-line duality in the pair of images: one

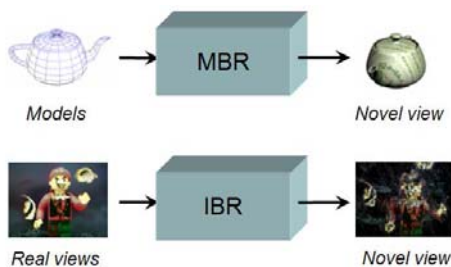


Fig. 1. Model-based Rendering and Image-based Rendering

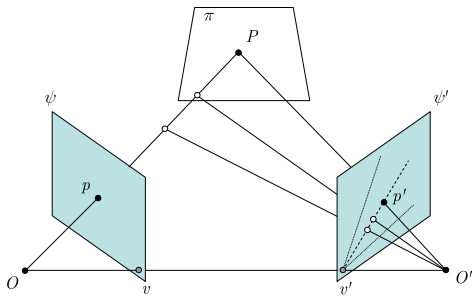


Fig. 2. Epipolar geometry.

point in the left view corresponds to a line in the right view. This concept is used to create a virtual view from two real views. Each point of the virtual view is the intersection of the lines of the points from the real views. The drawbacks of this technique are the definition of the camera (virtual) pose and the estimation of epipolar constraint. The trifocal transfer approach first proposed by [13] is based on the trifocal constraint between three views. The later is defined by a tensor. With two real images and a tensor all the points of the images are transferred to a novel view. The control of the corresponding virtual camera pose is performed through a displacement matrix.

In this paper we use the trifocal approach to synthesized a virtual camera for micromanipulation and microassembly applications. Section 2 and 3 describe respectively the geometry of two (epipolar geometry) and three views (trifocal geometry). Section 4 explains how to perform a novel view synthesis by trifocal geometry. Section 5 describes its application to a micromanipulation example: the picking up of a microgear using a microgripper.

II. EPIPOLAR GEOMETRY

Fig. 2 shows the projective model of two views imaging system. The points O and O' are respectively the optic center of the right and left cameras, then the line (OO') is the baseline of the stereo system. The projection of O in the view ψ' , defines the epipole v' , the projection of O' in the view ψ defines the other epipole v . Both views of this stereovision system are intrinsically linked by the epipolar geometry. If a point P of space belongs to a plane π , it is projected along a line (PO) , in the image plane ψ at the point p . The point P is also projected along a line (PO') , in the image plane ψ' at the point p' . Each point that belongs to the line (Pp) is projected on a point that belongs to the line $(v'p')$ in ψ' : the point p of ψ corresponds to the epipolar line $(v'p')$ of ψ' .

For every point P of an object in \mathcal{P}^3 that belongs to the same plane π , if the point is projected in ψ its correspondent in ψ' can be determined. A projective transformation matrix A in \mathcal{P}^2 (collineation matrix) links the position of the point p of ψ with the position of the point p' of ψ' :

$$p' \cong Ap \quad (1)$$

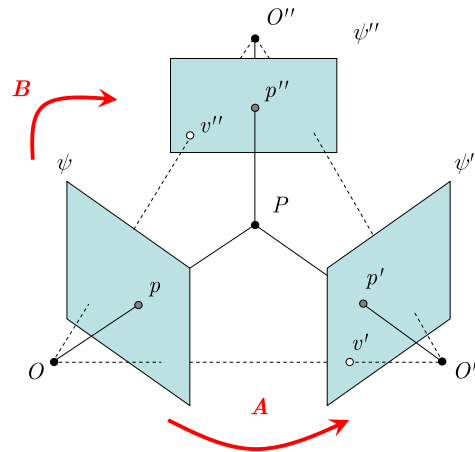


Fig. 3. Trifocal geometry.

where \cong stands for the equality up to a scale factor. If the point P does not belong to the plane π , (1) is not true. To satisfy the projective projection of the point $P \notin \pi$, the point projected p' on ψ' is rectified by the parallax which includes the epipole v' and the relative structure δ . The later does not depend on ψ' : it is constant. The new relation between p and p' , (for the point $P \notin \pi$) is:

$$p' \cong Ap + \delta v' \quad (2)$$

The plane (p, p', O, O') defines the epipolar constraint. It allows the determination of the epipole v' from the point p and p' by the fundamental matrix. That matrix F defines algebraically the coplanar constraint by:

$$p'^T F p = 0 \quad (3)$$

This equation consists of bilinear forms of p and p' .

III. TRIFOCAL GEOMETRY

Trifocal geometry is the extension of epipolar geometry to three views. Let us consider three views of \mathcal{P}^2 ψ , ψ' and ψ'' (Fig. 3). A point $P \in \mathcal{P}^3$ is projected onto the point $p = (x, y, 1)^T$ in ψ , $p' = (x', y', 1)^T$ in ψ' and $p'' = (x'', y'', 1)^T$ in ψ'' . Let us note:

- A and B the collineation matrixes corresponding respectively to the projective transformation $\psi \rightarrow \psi'$ and $\psi \rightarrow \psi''$,
- v' and v'' the projection of the optic center O on respectively ψ' and ψ'' .

A. Definition of Trilinearity

The trilinearity defines the constraint between three views: p , p' and p'' are linked by the same point projected P . Each coordinate of p , p' and p'' satisfies a pair of trilinear equations of the following form [14]:

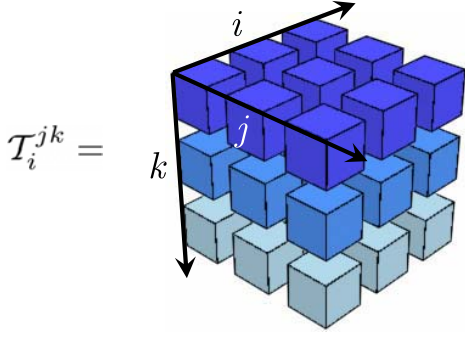


Fig. 4. Trilinear tensor index notation.

$$\begin{aligned}
& x'' (\alpha_1 x + \alpha_2 y + \alpha_3) + x'' x' (\alpha_4 x + \alpha_5 y + \alpha_6) + \\
& x' (\alpha_7 x + \alpha_8 y + \alpha_9) + \alpha_{10} x + \alpha_{11} y + \alpha_{12} = 0 \\
& y'' (\beta_1 x + \beta_2 y + \beta_3) + y'' x' (\beta_4 x + \beta_5 y + \beta_6) + \\
& x' (\beta_7 x + \beta_8 y + \beta_9) + \beta_{10} x + \beta_{11} y + \beta_{12} = 0
\end{aligned} \quad (4)$$

where the coefficients α_i and β_i for $i \in [1, 12]$ are constants i.e. they do not depend on the point P . The equations are trilinear because the constraint is between three points p, p' and p'' . The trilinear coefficients from the equations of the epipolar geometry of (ψ, ψ') and (ψ, ψ'') can be calculated using the following equations:

$$\begin{aligned}
p' &\cong Ap + \delta v' \\
p'' &\cong Bp + \delta v''
\end{aligned} \quad (5)$$

where δ is the relative affine structure of P . The coefficient δ is independent of ψ' , i.e., is invariant according to the choice of the second view [15]. Then δ can be isolated from both (5):

$$\delta = \frac{v'^1 - x'v'^3}{(x'a^3 - a^1)p} = \frac{v'^2 - y'v'^3}{(y'a^3 - a^2)p} = \frac{y'v'^1 - x'v'^2}{(x'a^2 - y'a^1)p'}, \quad (6)$$

$$\delta = \frac{v''^1 - x''v''^3}{(x''b^3 - b^1)p} = \frac{v''^2 - y''v''^3}{(y''b^3 - b^2)p} = \frac{y''v''^1 - x''v''^2}{(x''b^2 - y''b^1)p'}, \quad (7)$$

where b^1, b^2, b^3 and a^1, a^2, a^3 are the row vectors of A and B , and $v' = (v'^1, v'^2, v'^3)^T$, $v'' = (v''^1, v''^2, v''^3)^T$. From these equations all the trilinear equations can be recovered. For example the equality of the first terms of both equations leads to the trilinear form:

$$\begin{aligned}
& x'' (v'^1 b^3 - v''^3 a^1) p - x'' x' (v'^3 b^3 - v''^3 a^3) p + \\
& x' (v'^3 b^1 - v''^1 a^3) p - (v'^1 b^1 - v''^1 a^1) p = 0.
\end{aligned} \quad (8)$$

By developing every equality of (6) and (7), four equations linearly independent with 27 distinct coefficients will be obtained.

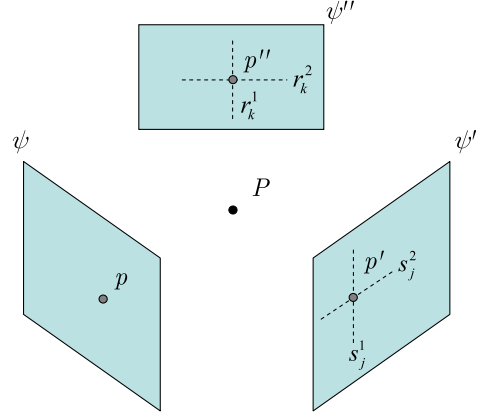


Fig. 5. The description of points as intersection of lines.

B. Trilinear Tensor

The 27 distinct coefficients can be written in a matrix form with a tensor formulation. Each coefficient of trilinear equation can be written in a canonical form:

$$\mathcal{T}_i^{jk} = v'^j b_i^k - v''^k a_i^j \quad (9)$$

where a_i^j and b_i^k are the elements of the collineation matrix A and B with $i, j, k \in [1, 3]$ (j and k are the indexes of the rows and i is the index of the column). The trilinear tensor \mathcal{T}_i^{jk} is a $3 \times 3 \times 3$ array of the 27 trilinear coefficients, the indexes are described at Fig. 4. The four trilinear equations, linearly independent, can be written with the tensor:

$$\begin{aligned}
& x'' \mathcal{T}_i^{13} p^i - x'' x' \mathcal{T}_i^{33} p^i + x' \mathcal{T}_i^{31} p^i - \mathcal{T}_i^{11} p^i = 0 \\
& y'' \mathcal{T}_i^{13} p^i - y'' x' \mathcal{T}_i^{33} p^i + x' \mathcal{T}_i^{32} p^i - \mathcal{T}_i^{12} p^i = 0 \\
& x'' \mathcal{T}_i^{23} p^i - x'' y' \mathcal{T}_i^{33} p^i + y' \mathcal{T}_i^{31} p^i - \mathcal{T}_i^{21} p^i = 0 \\
& y'' \mathcal{T}_i^{23} p^i - y'' y' \mathcal{T}_i^{33} p^i + y' \mathcal{T}_i^{32} p^i - \mathcal{T}_i^{22} p^i = 0
\end{aligned} \quad (10)$$

These equations can be written in a compact form using the following notation:

$$S = (s_j^\mu) = \begin{bmatrix} -1 & 0 & x' \\ 0 & -1 & y' \end{bmatrix} \quad (11)$$

$$R = (r_k^\rho) = \begin{bmatrix} -1 & 0 & x'' \\ 0 & -1 & y'' \end{bmatrix} \quad (12)$$

where $j, k \in [1, 3]$ and $\mu, \rho \in [1, 2]$. p' is defined as the intersection of the horizontal line $s_j^2 = (0, -1, y')$ with the vertical line $s_j^1 = (-1, 0, x')$. p'' is defined as the intersection of the horizontal line $r_k^2 = (0, -1, y'')$ with the vertical line $r_k^1 = (-1, 0, x'')$ (Fig. 5). The four equations of trilinearity can be written in line-line-point correspondence with the tensor and the two matrixes:

$$s_j^\mu r_k^\rho \mathcal{T}_i^{jk} p^i = 0 \quad (13)$$

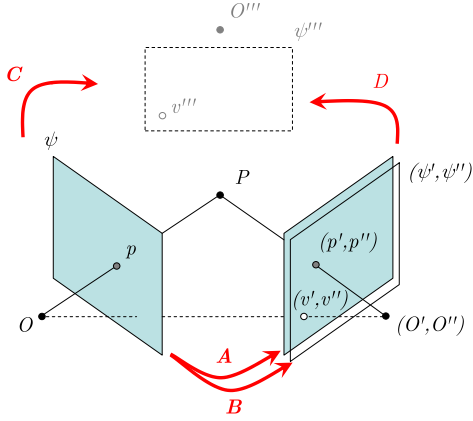


Fig. 6. Three views with two views merged to compute seed tensor, and virtual view after transformation.

IV. NOVEL VIEW SYNTHESIS BY TENSOR TRILINEAR

The first application of NVS by trilinear tensor is reported in [13] where three real views are used to compute the trifocal tensor and the virtual view: three real views lead to a virtual view. Later the authors proposed in [16] a more subtle approach that consists in confusing two of three input views: as a result a novel view (virtual) is obtained from two real views.

A. Seed Tensor from two views

Firstly the trifocal tensor is computed with three views where the second and third views are taken identical: the result defines the seed tensor. Secondly the novel view is defined through a spacial transformation of the third view and then a collineation matrix. Finally the pixels of the first image are matched with the pixels of the second image and projected in the third image that defines the novel view.

Let us consider three real views ψ , ψ' and ψ'' . As explained above the trilinear constraint $\mathcal{T}(\psi, \psi', \psi'')$ can be calculated by (9). Now suppose ψ'' merged with ψ' (Fig. 6). That means the collineation matrixes A ($\psi \rightarrow \psi'$) and B ($\psi \rightarrow \psi''$) and the epipoles v' and v'' are identical. Thus (9) becomes:

$$\mathcal{T}_i^{jk} = v'^j a_i^k - v'^k a_i^j \quad (14)$$

The later defines the seed tensor. Let us suppose the view ψ'' becomes the view ψ''' by a projective transformation D , then a collineation matrix C links ψ and ψ''' . As the same the seed tensor $\mathcal{T}(\psi, \psi', \psi'')$ changes to $\mathcal{G}(\psi, \psi', \psi''')$:

$$\mathcal{G}_i^{jk} = v'^j c_i^k - v'''^k a_i^j \quad (15)$$

with c_i^k and v'''^k the elements of collineation matrix $C(\psi \rightarrow \psi''')$ and the epipole v''' . The new tensor \mathcal{G} can be computed using the seed tensor \mathcal{T} , the collineation matrix

D and the collineation matrixes C and B . C corresponds to the product of D by B , then:

$$c_i^k = d_l^k b_i^l \quad (16)$$

where d_l^k is the element of the collineation matrix D . The seed tensor is defined for $A = B$, then (15) becomes:

$$\begin{aligned} \mathcal{G}_i^{jk} &= v'^j (d_l^k b_i^l) - v'''^k a_i^j \\ &= d_i^k \mathcal{T}_i^{jk} + (d_l^k v'''^k - v'''^k) a_i^j \end{aligned} \quad (17)$$

As $Dv''' - v'''$ represents a translation vector from v''' to v''' , t , the new tensor \mathcal{G} is of the form:

$$\mathcal{G}_i^{jk} = d_l^k \mathcal{T}_i^{jl} + t^k a_i^j \quad (18)$$

where t^k is the element of the vector t . D and t are of the form:

$$D = \begin{bmatrix} d_1^1 & d_2^1 & d_3^1 \\ d_1^2 & d_2^2 & d_3^2 \\ d_1^3 & d_2^3 & d_3^3 \end{bmatrix}, \quad t = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (19)$$

The view ψ''' defines a novel view, a virtual view that can be driven by a function of two sets of parameters d_l^k and t^k . To achieve the synthesis of ψ''' the pixels of ψ and ψ' should be transferred to ψ''' using the trilinear tensor: retroprojection.

B. Reprojection on Virtual View

The trilinear equation is used to compute the position of every pixel of ψ and ψ' in the virtual view ψ''' . From the first equation of (10) the coordinates x''' and y''' of p''' can be calculated according to p , p' and \mathcal{G} as followed:

$$\begin{aligned} x''' &= \frac{x' \mathcal{G}_i^{31} p^i - \mathcal{G}_i^{11} p^i}{x' \mathcal{G}_i^{33} p^i - \mathcal{G}_i^{13} p^i} \\ y''' &= \frac{x' \mathcal{G}_i^{32} p^i - \mathcal{G}_i^{12} p^i}{x' \mathcal{G}_i^{33} p^i - \mathcal{G}_i^{13} p^i} \end{aligned} \quad (20)$$

The point p''' is defined up to a scale factor, the true form is $p''' = (\lambda x''', \lambda y''', \lambda)$. It can be written using (11), (12) and (20):

$$\begin{aligned} \lambda x''' &= s_j^1 \mathcal{G}_i^{j1} p^i \\ \lambda y''' &= s_j^1 \mathcal{G}_i^{j2} p^i \\ \lambda &= s_j^1 \mathcal{G}_i^{j3} p^i \end{aligned} \quad (21)$$

There is a redundancy of p''' with the four equations of trilinearity, i.e. x''' and y''' can be estimated with s_j^1 or s_j^2 . The equation of reprojection can be compacted to:

$$p'''^k \cong s_j^\mu \mathcal{G}_i^{jk} p^i \quad (22)$$

The reprojection requires the knowledge of the correspondence for every point of the real images: p in ψ with p' in ψ' . Consequently the reprojection requires a dense correspondence between the two real images.

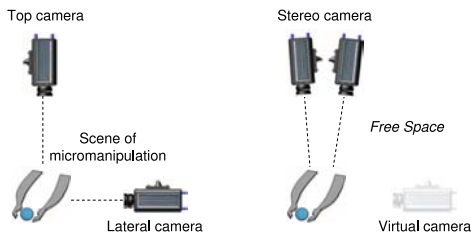


Fig. 7. Real camera versus virtual camera in micromanipulation.

V. EXPERIMENTAL RESULTS

We use the technique of trifocal transfer to perform dynamically the synthesis of novel views in a vision system for micromanipulation and microassembly. Multiple views vision systems are usually used: top view allows the xy position estimation of the microgripper and the lateral view allows its z position estimation (Fig. 7). Unfortunately sometimes it is not possible to position lateral camera or it is useful to set free the workflow. Fortunately with two cameras at the top it is possible to synthesize lateral views. The edges transfer is enough for our experiment since our application is the monitoring and control of micromanipulation activity.

A. The Real Views

We consider a typical task in micromanipulation: the picking up of a micropart (a microgear) by a microgripper. Only one top view is not sufficient to detect whenever or not the microgear is between the gripper tips and can be picked up. The lateral view allows to overcome this drawback. Two cases are considered: the part is not between the tips and the part is between the tips (Fig. 8).

B. Estimation of Trilinear Tensor

The estimation of the trilinear tensor requires the computation of the collineation and fundamental matrixes. A weak calibration is made by using a calibration pattern of two arrays of 6×6 discs. Each disc is localized precisely (sub-pixel accuracy) by its centroid and matched in the two views. The collineation matrix A is estimated using the four points algorithm. The algorithm from [17] is used to compute the fundamental matrix F from which the



Fig. 8. Top, stereo top images and lateral image of the scene with the microgear outside the gripper tips. Bottom, the same scene with the microgear inside the gripper tips.



Fig. 9. The first image represents the original image, the second the gradient and the third the vectoring image.

epipole v and v' are calculated using a Singular Value Decomposition (SVD):

$$F = U\Sigma V^T \quad (23)$$

with Σ the singular value, (u_1, u_2, u_3) the column of U and (v_1, v_2, v_3) the column of V . The epipolar v' corresponds to u_3 and v to v_3 . Knowing v' and a_i^j , the seed tensor can be computed from (14).

C. Matching the Points of Real Views

To reproject a point of the right image in the virtual view by trilinear tensor we must know its correspondence in the left image. For our application we just need to know the relative position between the gripper and the part. Only some features of the gripper and the part are interesting, then the first stage consists in determination an invariant feature in both images for the matching. We choose the feature that defines the edge of the objects, in order to decrease the computation time.

We use the Sobel edge detector [18] to compute the edges of the views (Fig. 9). The pixels of these edges are matched in the views using the algorithm from [19]. It is based on the search of the correspondence along the epipolar line in the other view using a Sum of Squared Difference (SSD) correlation.

D. Reprojection of the microgripper and the microgear

Now we can compute a new tensor to obtain the novel view by modifying the translation matrix according to (18) and compute the position of every pixel of the edges defined above using (20) (reprojection stage). Fig. 10 shows some virtual views for the two configurations considered: the virtual camera moves from top to side. The synthetic views matched with real views, it can be seen with no confusion if the microgear is inside or outside the gripper as soon as the camera leaves the top position. Some noise appears Fig. 10 (surrounded elements). It comes from the wrong matching pixels between the real views.

VI. CONCLUSION

We have summarized the geometry of two views (epipolar geometry) and three views (trifocal geometry), and explained how the later leads to the trifocal constraint. Then we have shown how the trifocal constraint expressed as a tensor can be used to create a virtual view from only two real views. The process includes three stages:

- the computing of a seed tensor from the two real views,

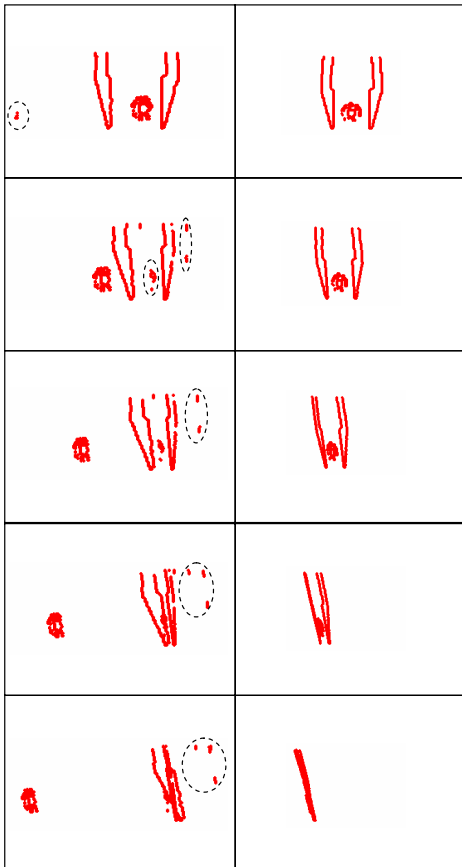


Fig. 10. Left, images from the virtual camera, from top view to lateral view, with the microgear outside the microgripper tips. Right, the same with microgear inside the microgripper tips. The surrounded elements correspond to the error of reprojection.

- the computing of the new tensor according to the seed tensor and the pose of the new view,
- the reprojection onto the new view of the matched pixels of both real views.

Our ideas have been validated by synthesizing a virtual camera in the case of a micromanipulation task: the picking up of a microgear by a microgripper. With two top views of the scene we have generated a virtual camera that allows the obtaining of any lateral view. The position of the camera can be easily changed until the best point of view is not reached. Thus it is possible to see if the microgear is inside or not the microgripper tips. Some noise appears in the new view because of the wrong matching of points, so some attention must be kept to precisely matching the point of the two images. The virtual view includes only the edges of the objects, however a textured image can be created if a dense correspondence is performed between both views.

Finally, the trifocal tensor approach can be used to create virtual cameras that can profitably replace real cameras and set free the work field. That is useful in micromanipulation and microassembly where the work field is

very constrained. By requiring only a weak calibration the method fits the manipulation at the microscale where precise calibration is very difficult to achieve. Instead of the synthesis of textured views using dense correspondence future work will deal with the use of auto-calibration technique to recover the fundamental matrix.

REFERENCES

- [1] G. Yang, J. A. Gaines, and B. J. Nelson, "A supervisory wafer-level 3d microassembly system for hybrid mems fabrication," *Journal of Intelligent and Robotic Systems*, vol. 37, pp. 43–68, 2003.
- [2] A. Matsumoto, T. Akimoto, K. Yoshida, H. Inoue, and K. Kamijo, "Development of mems component assembly machine - application of robotics technology to micromechatronics," in *The International Symposium on Micro-Mechanical Engineering*, December 2003, pp. 83–88.
- [3] D. O. Popa and H. E. Stephanou, "Micro and mesoscale robotic assembly," *Journal of Manufacturing Process*, vol. 6, no. 1, pp. 52–71, 2004.
- [4] L. Sun, H. Xie, W. Rong, and L. Chen, "Task-reconfigurable system for mems assembly," in *IEEE International Conference on Robotics and Automation*, Barcelona, Spain, April 2005, pp. 844–849.
- [5] S. E. Chen and L. Williams, "View interpolation for image synthesis," in *Computer Graphics (SIGGRAPH'93 Proceedings)*, 1993, pp. 279–288.
- [6] H.-Y. Shum and S. B. Kang, "A review of image-based rendering techniques," in *SPIE Proceedings of the Visual Communications and Image Processing*, June 2000, pp. 2–13.
- [7] S. Chen, "Quicktime vr - an image-based approach to virtual environment navigation," in *Computer Graphics (SIGGRAPH'95 Proceedings)*, August 1995, pp. 29–38.
- [8] R. Szeliski and H.-Y. Shum, "Creating full view panoramic image mosaics and environment maps," in *Computer Graphics (SIGGRAPH'97 Proceedings)*, vol. 31, 1997, pp. 251–258.
- [9] H. Saito, S. Baba, and T. Kanade, "Appearance-based virtual view generation from multicamera videos in the 3-d room," in *IEEE Transactions on Multimedia*, vol. 5, September 2003, pp. 303–316.
- [10] K. Connor and I. Reid, "Novel view specification and synthesis," in *Proceeding of the British Machine Vision Conference*, Cardiff, England, 2002, pp. 243–252.
- [11] M. Irani, T. Hassner, and P. Anandan, "What does the scene look like from a scene point?" in *Proceedings of European Conference on Computer Vision (ECCV)*, Copenhagen, 2002, pp. 883–897.
- [12] O. Faugeras and L. Robert, "What can two images tell us about a third one?" INRIA, Technical report 2018, 1993.
- [13] S. Avidan and A. Shashua, "Novel view synthesis in tensor space," in *IEEE Computer Vision and Pattern Recognition*, 1997, pp. 1034–1040.
- [14] A. Shashua, "Algebraic functions for recognition," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, 1994, pp. 779–789.
- [15] A. Shashua and N. Navad, "Relative affine structure: Theory and application to 3d reconstruction from perspective views," in *IEEE Computer Vision and Pattern Recognition (CVPR)*, 1994, pp. 483–489.
- [16] S. Avidan and A. Shashua, "Novel view synthesis by cascading trilinear tensors," in *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, vol. 4, 1998, pp. 293–306.
- [17] R. I. Hartley, "In defense of the eight-point algorithm," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 6, June 1997, pp. 580–593.
- [18] K. Pingle, *Visual Perception by a Computer*. New York: Academic Press, 1969, ch. Automatic Interpretation and Classification of Images, pp. 277–284.
- [19] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong, "A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry," INRIA, Technical report 2273, May 1995.