3D Reconstruction of DNA Molecules

Francesc Piñol i Margalef[1], Patrick Furrer[2]
and Jacques Dubochet[2]

[1]Signal Processing Laboratory, Swiss Federal Institute of Technology, CH-1015 Lausanne, Switzerland
[2]Laboratoire d'Analyse Ultrascturale, Bâtiment de Biologie, CH-1015 Lausanne, Switzerland

RÉSUMÉ

Pour étudier la forme tridimensionnelle des molécules d'ADN, on a développé une méthode qui consiste à vitrifier une fine pellicule de solution observée ensuite par cryo-microscopie électronique. Le spécimen étant très sensible aux dommages causés par le faisceau, seules deux images à faible dose de la même molécule sont enregistrées. Cependant, ces micrographies souffrent à la fois d'un contraste et d'un rapport signal sur bruit extrêmement faibles. L'expérience montre que le suivi manuel effectué par un être humain s'avère assez peu reproductible. Pour améliorer la précision est la fiabilité des mesures on a développé une approche automatique où toute la recherche est faite en trois dimensions (3D). Elle imite la façon dont le cerveau semble procéder en requérant la continuité du filament et en incluant la flexibilité limitée de l'ADN. En plus, le fait de travailler en 3D devrait aider à résoudre quelques incertitudes, comme les croisements et les régions perpendiculaires à l'axe de tilt.

ABSTRACT

To study the three-dimensional shape of DNA molecules, we have developed a method that consists in vitrifying a thin film of solution and observing it by cryo-electron microscopy. As the specimen is rather sensitive to beam-damage, only 2 low-dose pictures (stereo-pair) of the same molecule are recorded. Nevertheless, micrographs are of very low contrast and poor signal-to-noise ratio. Experience shows however that manual tracing by a human observer is not very reproducible. In order to improve the accuracy and the reliability of the measurements we have developed an automatic approach where all the search is made in three dimensions (3D). It mimics the way the brain seems to proceed by requiring the continuity of the filament and by including the known limited flexibility of the DNA. Furthermore, the fact of working in 3D should help to solve remaining uncertainties, like crossings and regions perpendicular to tilt axis.

1 Introduction

Direct visualization of unstained DNA immobilized by vitrification of its aqueous environment has recently become possible by means of cryoelectron microscopy [1]. Stereograms and digital three-dimensional (3D) reconstructions have been obtained from pairs of tilted (30 degrees) images. The major difficulty of the method lies in the fact that DNA is very thin (2nm) and it is bearly visible on the image. Under the best imaging conditions, which are not yet routinely obtainable, the DNA filament has a very low contrast and a weak signal to noise ratio. The human brain seems to have a very good capability to follow such molecules; for this purpose, it uses various preconceived knowledge such as the continuity of the molecule and its restricted bendability. A trained observer is also capable of extending these rules in 3D on the basis of stereoscopic images.

Tracing by manual clicking is unfortunately not very reproducible. It may induce significant errors in the determination of important geometrical parameters such as length, persistence length, crossings in super-coiled DNA and related characteristics. In order to improve the accuracy and the reliability of the measurements we have developed an automatic tracing system. In the approach presented here all the search is made in 3D. It mimics the way the brain proceeds by requiring the continuity of the filament and by including the known limited flexibility of the DNA axis itself.

2 System Description

2.1 Computer System

The computer image processing was performed on Silicon Graphics Workstations as described in [2]. An Eikonix 1412 CCD camera was used to digitize images from illuminated EM (electronic microscopy) negatives. The final magnification was chosen so that the filaments were typically around 5 pixels wide.

2.2 Microscopy System

The digitized images of two micrographs constituting together a stereo pair were oriented with the tilt axis ($0y$-axis) placed vertically on the display.

These pairs of images were aligned along the $0y$ direction on the basis of a reference point which was a precisely defi-

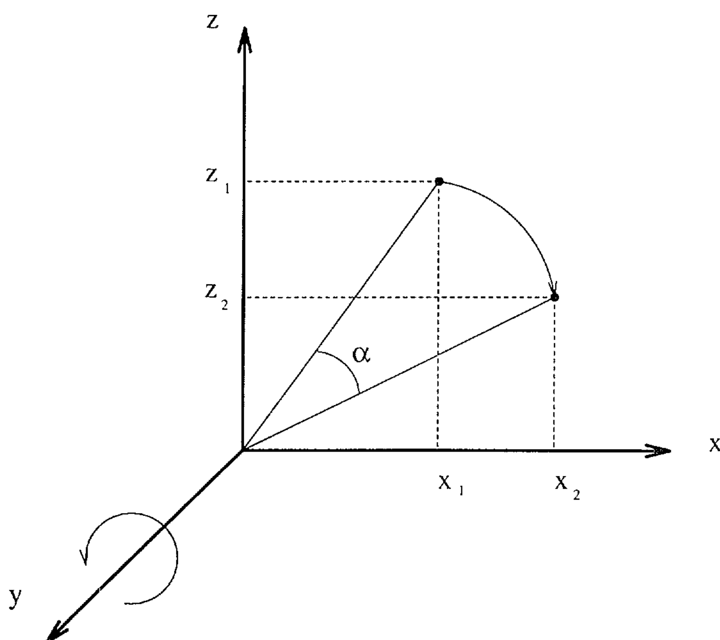


Figure 1: Z coordinate from the projections.

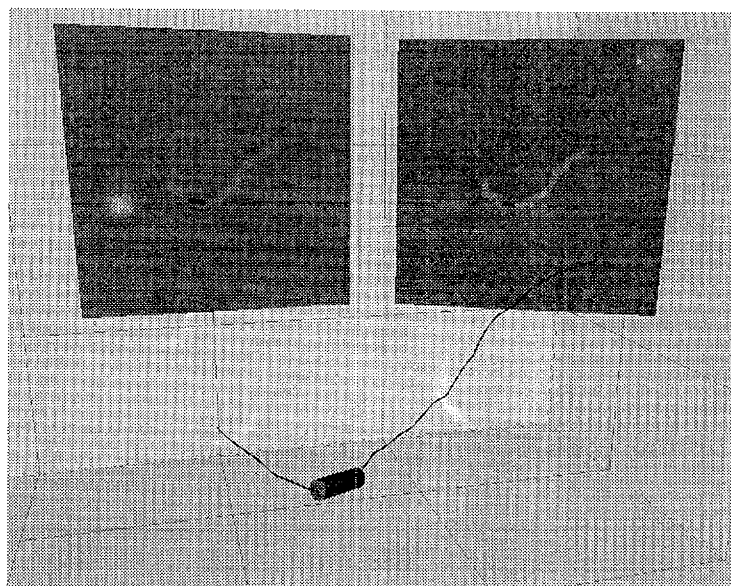


Figure 2: For each position of the cylinder in the 3D space, both projections are compared, in terms of correlation, with the considered region of the real images.

ned feature, clearly visible on both micrographs. Coordinate axes were chosen such as to bring the photographic plane of the second picture perpendicular to the $0z$-axis.

The reference point mentioned above was taken as the origin for the x, y and z coordinates. The coordinate of a point i in space, (x_i, y_i, z_i), projected on the first image plane, is (x_1, y_1), whereas on the second projected plane it becomes (x_2, y_2) (Fig. 1).

Provided that 0y is parallel to the tilt axis, $y_1 = y_2$. A simple calculation gives the z coordinate for this point:

$$z_2 = \frac{x_2 \cos(\theta) - x_1}{\sin(\theta)} \tag{1}$$

where θ is the tilt angle between the two micrographs.

2.3 The Flying Cylinder

Until recently, molecules were traced manually by clicking with a mouse, pairs of corresponding points on the two projections views of the reconstructed molecules [2]. However such procedure was dependent on the operator skill to correctly locate the axis of the molecules along DNA contours on the poorly contrasted images. Therefore sequential reconstructions even performed by the same operator slightly varied in their outcome. To eliminate this subjectivity effect we decided to develop a program of automatic filament recognition. This program called *flying cylinder* places a virtual short cylinder in space in such a way that its virtual projection on the plane of the first and of the second picture of stereo-pair overlaps best with given regions of the two projection views of the photographed molecules (Fig. 2).

This program requires that an operator selects the first two matching points on the two projections (determining thus z-coordinate of the corresponding point according to Eq.(1). This point in space is then used as a center of the short virtual cylinder with a length corresponding to *ca* 7-10 nm. The diameter of this cylinder corresponds to the diameter of the photographed DNA projection views, namely 2-3 nm. Then the program looks for a spatial orientation of this cylinder such that its projection overlaps best with the corresponding regions of the photographed filaments on both pictures of a stereo-pair. The best overlap is estimated by computing the maximum 2D correlation between the projections of the cylinder and the images [3]. Subsequently the program shifts the virtual cylinder by a small distance (1-2 nm) along its axial direction and at this new location the virtual cylinder is again reoriented according to the same maximization rule.

At every step of the flight, the correlations obtained for both projections are multiplied, and this global correlation factor is then compared to a user-defined threshold either for detection of the end of the molecule, or for tracking the exit of an horizontal region. Moreover, the profile of this correlation factor is displayed during the cylinder flight, allowing the user to monitor the tracking in detail.

The shift along the filament is sufficiently short so that the DNA is unlikely to be deflected along this distance by more than 30 degrees. To shorten the time required for calculation, after each shift the program tries only some equispaced angular orientations of the cylinder, typically within a 30 degrees cone (Fig. 3).

Locations of the central points of the flying cylinder after every step provide 3D coordinates of reconstructed trajectories of the analyzed DNA molecules.

There is a particularity inherent to stereoscopy concerning the regions perpendicular to the tilt axis. The presence of such an horizontal segment in the two projections leads to an unfortunate indetermination: an infinity of different paths in 3D may have these horizontal projections, and the real path cannot be determined from these horizontal regions. The adopted solution is to try to detect these regions and to skip them (the end of an horizontal region is reached

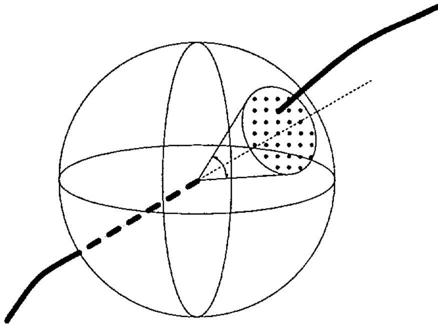


Figure 3: The tracing of the molecule is done with steps that are a part of the cylinder length. For each step, a region with a given solid angle around the previous direction is explored.

when the 2D correlation goes below a given threshold on each picture). The missing segments will be 3D interpolated at the end.

In order to adapt the method to several images with different characteristics, several parameters of the tracking can be user-defined. For the cylinder, we can choose its length – a direct consequence of the DNA persistence length –, and its width. The relative weight between the center (inner width) and the cover (outer width), representing the better quality of signal along central axis, is helping to a better detection of the filament. Concerning its dynamics, the forward step, as well as the searching solid angle, can also be chosen by the user.

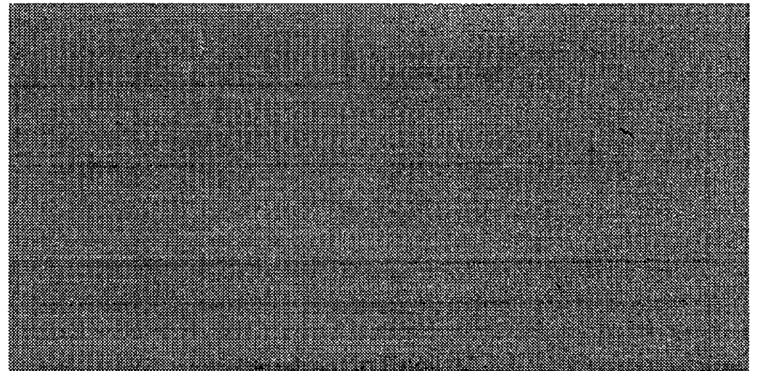
3 Results

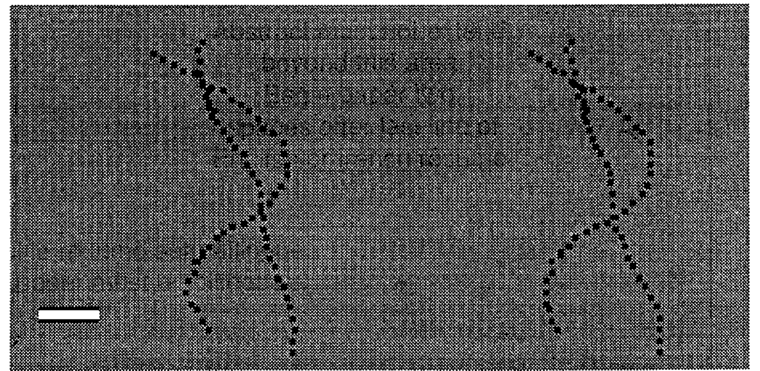
The flying cylinder has been applied for 3D reconstruction of a closed fragment of a DNA molecule. The results for a supercoiled plasmid is shown in Fig 4. Some fragments have been needed to complete the tracking, but the main goal has been achieved. We can determine its shape in space, as well as some physical parameters, such as the persistence length and the writhing of the filament.

From this example we can see the ability of the method to follow correctly the crossing regions, where sometimes the information in only one image would not be enough to determine the correct path.

4 Conclusions

The flying cylinder algorithm is rather outstanding in detecting filaments immersed in noise, provided two pictures of the same filament with different tilt angles are available. The major advantage of the method lies in the fact that it works in three dimensions. Therefore, it takes simultaneously full advantage of the information contained in both images of a stereo-pair. It compares also favorably with visual stereoscopy (stereopsis), which is prone to visual illusions. We have observed that stereograms obtained from images which are hardly discernible have a strong subjective component. For example, it is difficult to override a preconceived idea when deciding which fragment, from two crossing filament, is on the top (for instance, this is crucial when dealing with the handedness of a super-coiled DNA molecule).





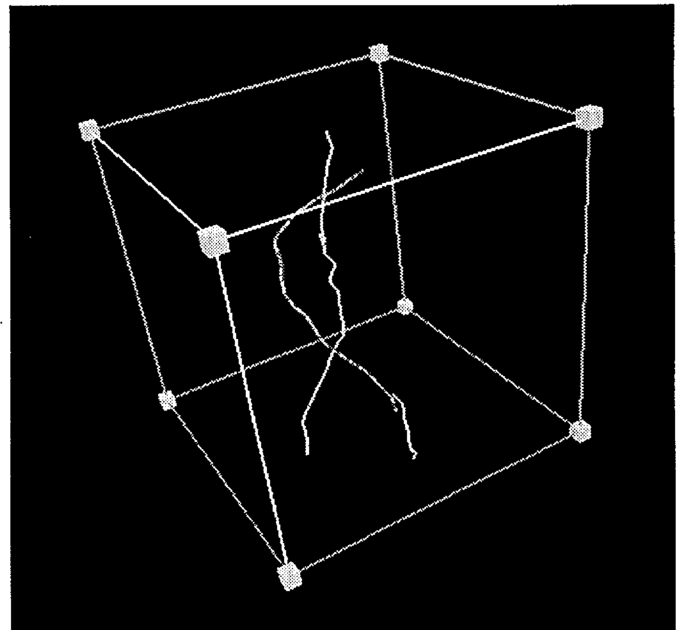


Figure 4: Reconstruction of a super-coiled plasmid with crossings.

The computer program do not suffer from this subjectivity. A further advantage of a reconstruction in 3D comes from knowledge-based tracking by explicitly using the known persistence length of the filament under tracking. This is not possible in 2D projection since on the image, any sharp turn can result from the projection of a smooth bend.

The experimental error of 3D reconstructions using this method is about 1nm. This means that the cylinder does cover DNA efficiently, although it may have some difficulty to remain correctly centered on the DNA axis itself. We are presently developing the flying cylinder algorithm by allowing adaptability for some of the parameters (which could therefore change during tracking), like the length of the cylinder. New degrees of freedom might be needed for helping the cylinder to stay on the axis, as for instance lateral moves at every step. Further improvements of the cylinder's flight should also help to improve the precision of the 3D reconstruction.

References

[1] J. Dubochet, M. Adrian, I. Dustin, P. Furrer, and A. Stasiak. Cryoelectron microscopy of dna molecules in solution. *Methods Enzymol*, (211):507–518, 1992.

[2] I. Dustin, P. Furrer, A. Stasiak, J. Dubochet, J. Langowski, and E. Egelman. Spatial visualization of dna in solution. *Journal of Structural Biology*, 107:15–21, 1991.

[3] A.K. Jain. *Fundamentals of Digital Image Procesing.* Prentice-Hall International Editions, 1989.