Figure 13.12: Computation of the velocity field (minimum of 5 matchings).

The representation is the same as for Figure 13.11, but this time the velocities were computed with a less restrictive matching constraint of 5 images (instead of 12 for Figure 13.11).

13.5.3   Sensitivity to noise

We now want to test how robust to noise our method is : are the DCMA analysis and the induced velocity estimation still reliable when applied to a noisy data ? In order to check this, we took the previous "TREES" movies and corrupted it strongly by replacing 50% of its grey values $u(i, j, k)$ by totally random, uniformly distributed and uncorrelated values. This kind of noise is called *impulse noise* : it is very destructive and impossible to remove efficiently with linear filters. On this corrupted movie, we applied exactly the same processing as in the original one. The figures to follow (to be compared with the corresponding figures for the original movie) show that both the visual aspect and the velocity field are well recovered by the DCMA although half of the original information was lost and replaced with random values.

Figure 13.13: Filtering of the noisy "TREES" movie.

Row 1 :   images 18 and 22 of the noisy "TREES" movie
Row 2 :   images 18 and 22 of the noisy "TREES" movie processed with 31 iterations.

The images on row 1 are very noisy : 50% of their pixel values were chosen by a non-correlated, uniformly distributed random generator (and this 50% amount of pixel was chosen itself by a random generator). When playing the movie, one has the impression of looking at a TV-image received in very poor conditions. In particular, it is almost impossible to see any detail of the ground texture. Filtering this movie with the DCMA gives impressive enhancement results : not only the noise impression is removed, but in addition some details appear that were not visible on the first movie (in particular on the ground and on the left tree). This means that the DCMA takes more advantage of the time coherence and redundancy of information contained in a movie than the human visual system does.
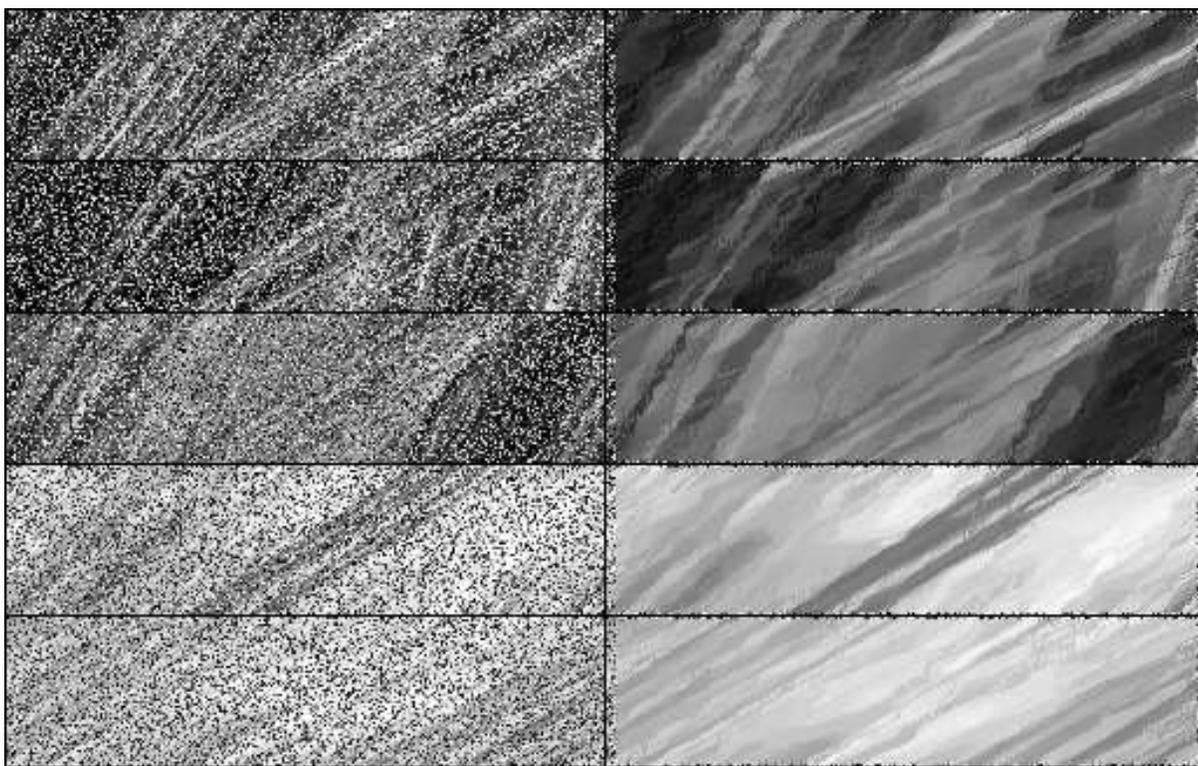
Figure 13.14: Analysis of the epipolar images.

As on Figure 13.5, epipolar images are shown both for the original noisy movie (column 1) and for its processed version after 31 iterations (column 2).
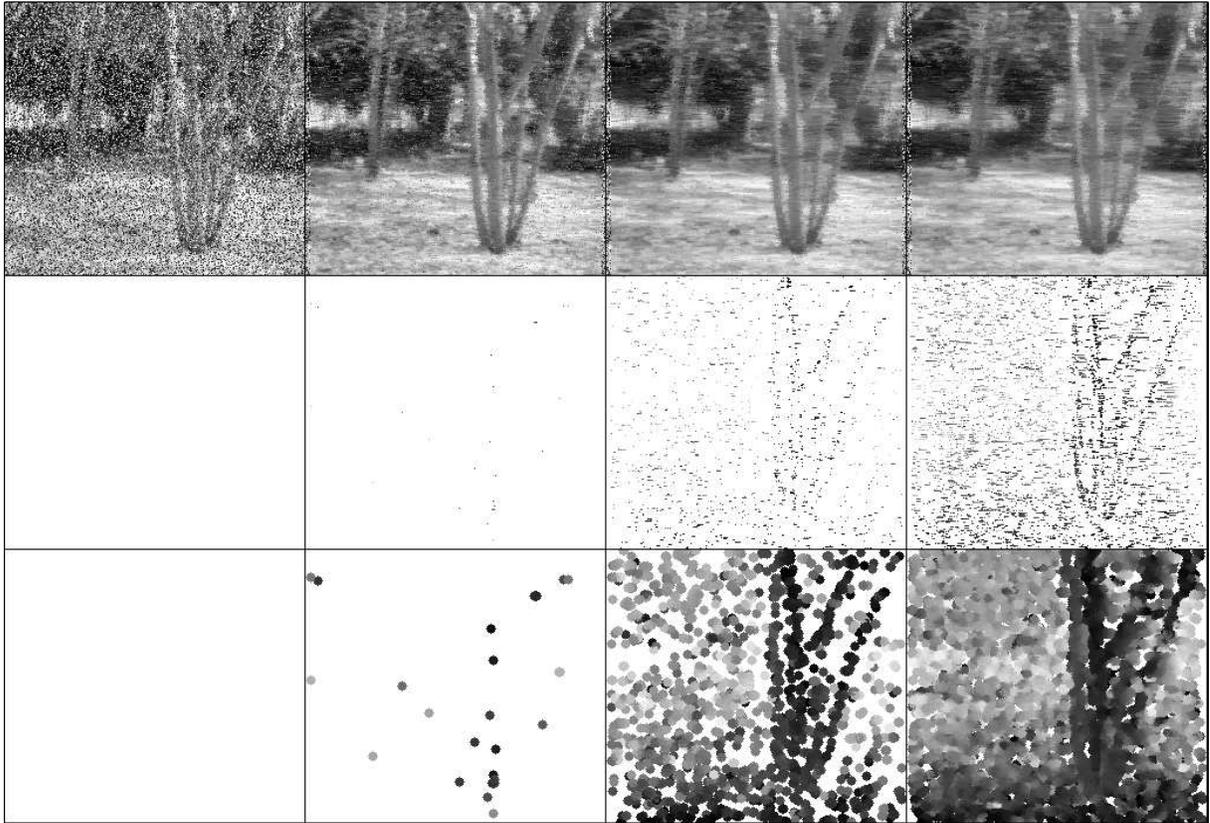
Figure 13.15: Computation of the velocity field (minimum of 15 matchings).

Like on Figure 13.6, the four images on the first row are the 20th image of four different movies :

column 1:	*original noisy "TREES" movie*
column 2:	*processed movie (5 iterations)*
column 3:	*processed movie (15 iterations)*
column 4:	*processed movie (31 iterations)*

Row 2 and 3 represent the extracted velocity field (for a minimum of 15 matching images), in the original (row 2) and dilated (row 3) representation. As expected, not only the movie is filtered, but the velocity of objects is recovered despite a lot of destructed clues due to the large amount of noise put on the movie. Of course, the velocity recovery is not as good as if the movie had not been initially corrupted, but the depth structure of the scene still appear on the bottom-right image.

Chapter 14

Extensions and conclusion

14.1 Extension to more general geometric configurations

In this section, we show that the geometric configuration we adopted throughout our study is not really required. In fact, the DCMA can easily be extended to a more general motion, provided that it is known or that it can be recovered (but we shall not properly investigate the problem of motion recovery here).

14.1.1 The camera motion is not horizontal

Practically, it is difficult to ensure that the camera moves exactly along the direction given by the horizontal axis of the image plane. The consequence is that the y-sections $(x, \theta) \mapsto u(x, y, \theta)$ of the movie should not be processed independently, for the epipolar lines are not contained in the (x, θ) plane. However, if the direction of the camera displacement is known, given by the angle ν with the x axis, then it is theoretically possible to bring the problem back to the ideal case ($\nu = 0$) with the simple rotation of the image plane given by

$$P' = \begin{pmatrix} \cos \nu & \sin \nu \\ -\sin \nu & \cos \nu \end{pmatrix} P.$$

The angle ν may be directly measured by an inertial system on the mobile robot. It can also be easily estimated on the resulting movie since it is a very redundant information.

14.1.2 The camera motion does not lie in the image plane

We now suppose that the motion of the camera is not contained in the image plane, that is to say its component along the direction orthogonal to the image plane is non-zero. We define the (OX) axis as the direction given by the motion of the camera, and the (OY) axis as the only direction orthogonal to (OX) and contained in the image plane. Then, the remaining axis (OZ), naturally defined from (OX) and (OY) in order to form an orthogonal system, makes an angle

α with the direction orthogonal to the image plane (see Figure 14.1). The projection from the scene to the image plane is given by

$$x = \frac{X - C - Z \tan \alpha}{Z + (X - C) \tan \alpha}$$

$$y = \frac{Y}{Z + (X - C) \tan \alpha}.$$

Compared to the ideal case $\alpha = 0$, the case $\alpha \neq 0$ induces a deformation of the image plane given by

$$x' = \frac{x \cos \alpha - \sin \alpha}{\cos \alpha + x \sin \alpha} = \frac{x - \tan \alpha}{1 + x \tan \alpha}$$

$$y' = \frac{y}{\cos \alpha + x \sin \alpha}$$

The map $(x, y) \mapsto (x', y')$ is defined on $(\mathbb{R} - \{-\cot \alpha\}) \times \mathbb{R}$, the singularity $x = -\cot \alpha$ giving a characterization of α. Thus, all previous results should still apply, provided that we rewrite the DCMA evolution equation according to this deformation map.
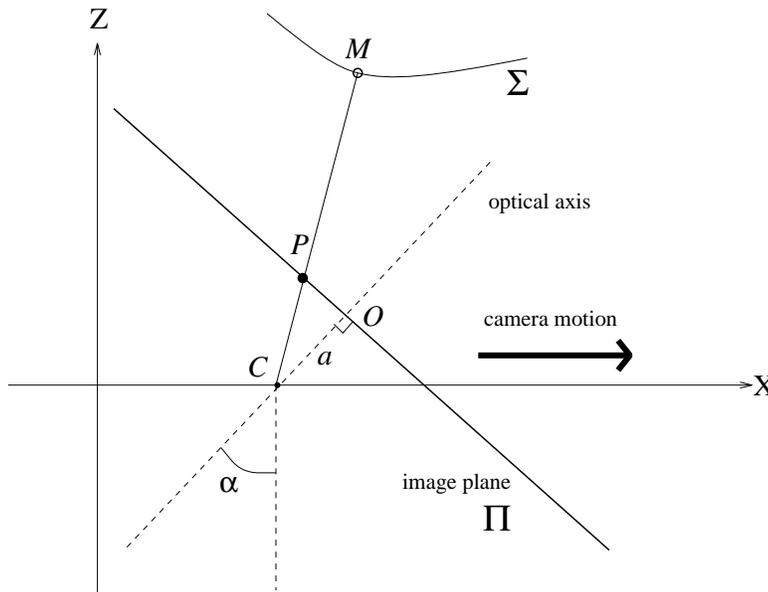
Figure 14.1: Camera motion does not lie in the image plane.

14.1.3   Case of pure zooming

If the camera moves in the direction of the optical axis, which corresponds to a pure "zooming", the movement of a physical point projected in P on the image plane is given by

$$\frac{dP}{d\theta} = \frac{V}{Z} P.$$

Thus, going to polar coordinates, the apparent velocity is

$$v = \frac{r}{Z} V = -\frac{u_\theta}{u_r}$$

everywhere $u_r \neq 0$. This case is formally equivalent to the ideal translation along the X axis as soon as the polar coordinate r is substituted to the x coordinate everywhere. In particular, the apparent acceleration is

$$? = \frac{\mathrm{D}v}{\mathrm{D}\theta} = -\frac{u_{\xi\xi}}{u_r} \quad \text{with} \quad \xi = \left(-\frac{u_\theta}{u_r}, 0, 1\right).$$

Rewriting the axiomatic formulation in that special case, we can expect to obtain the evolution equation

$$\frac{\partial u}{\partial t} = u_{\theta\theta} - 2\frac{u_\theta}{u_r}u_{\theta r} + \left(\frac{u_\theta}{u_r}\right)^2 u_{rr},$$

formally equivalent to the DCMA up to a change of coordinates.

14.2 Case of any rigid motion

The two previous cases can be combined to cover all situations of pure translation motion of the camera. The case of pure rotation with a fixed axis ("radar motion") is not very different from the case of pure translation : the filtering is the same and only the depth interpretation deduced from the velocity field changes.

In case of a general camera motion (translation T + rotation R), there are 6 instantaneous motion parameters : 3 for the translation and 3 for the rotation. More precisely, the movement of a physical point $M(X, Y, Z)$ is given in the camera referential by

$$\frac{dM}{d\theta} = -T - R \wedge M,$$

where we wrote $\wedge$ for the usual vector product in $\mathbb{R}^3$. Then, the perspective projection $(x, y) = \frac{1}{Z}(X, Y)$ induces in the image referential the movement

$$\frac{dP}{d\theta} = \frac{1}{Z}\begin{pmatrix} -1 & 0 & x \\ 0 & -1 & y \end{pmatrix}T + \begin{pmatrix} xy & -(1+x^2) & y \\ 1+y^2 & -xy & -x \end{pmatrix}R = \frac{1}{Z}AT + BR. \tag{14.1}$$

while the well-known Motion Constraint Equation is

$$\nabla u.\frac{dP}{d\theta} + u_\theta = 0, \tag{14.2}$$

∇u standing for the spatial gradient of u. Combining Equations 14.1 and 14.2 yields a scalar equation satisfied by the partial derivatives of u, with one unknow (the depth Z) and six motion parameters (T and R). It permits to compute the disparity $d = 1/Z$ by

$$d = -\frac{u_\theta + (BR).\nabla u}{(AT).\nabla u}.$$

Therefore, depth recovery is still theoretically possible as soon as the camera motion is known. We guess that it is possible to rewrite the DCMA in case of such a general camera motion, by introducing the motion parameters in the evolution equation.

14.3   Occlusions

In this study, we made several allusions to the problem of occlusions, which is not solved by the algorithm we presented. We know precise this point, and try to explain why this is the most important improvement to be brought to our method.

Two kinds of occlusions appear on a movie : the **natural occlusions**, occuring when a part of the scene masks another part (see Figure 14.2), and the **boundary occlusions**, which happen on the border of the image. The natural occlusions are consequences of both the scene geometry and the camera parameters, and they can be theoretically avoided by choosing an optical system with a small field width (or equivalently, with a large focal length). Of course, boundary occlusions cannot be avoided. In addition, avoiding natural occlusions forces the relative depth variations to be small, which prevents the depth estimation from being very accurate. Therefore, being able to deal with occlusions is a key point of the movie analysis, and it is not surprising that the human visual system makes a strong use of occlusions phenomena.

small field width : no occlusion appears   large field width : occlusions appear
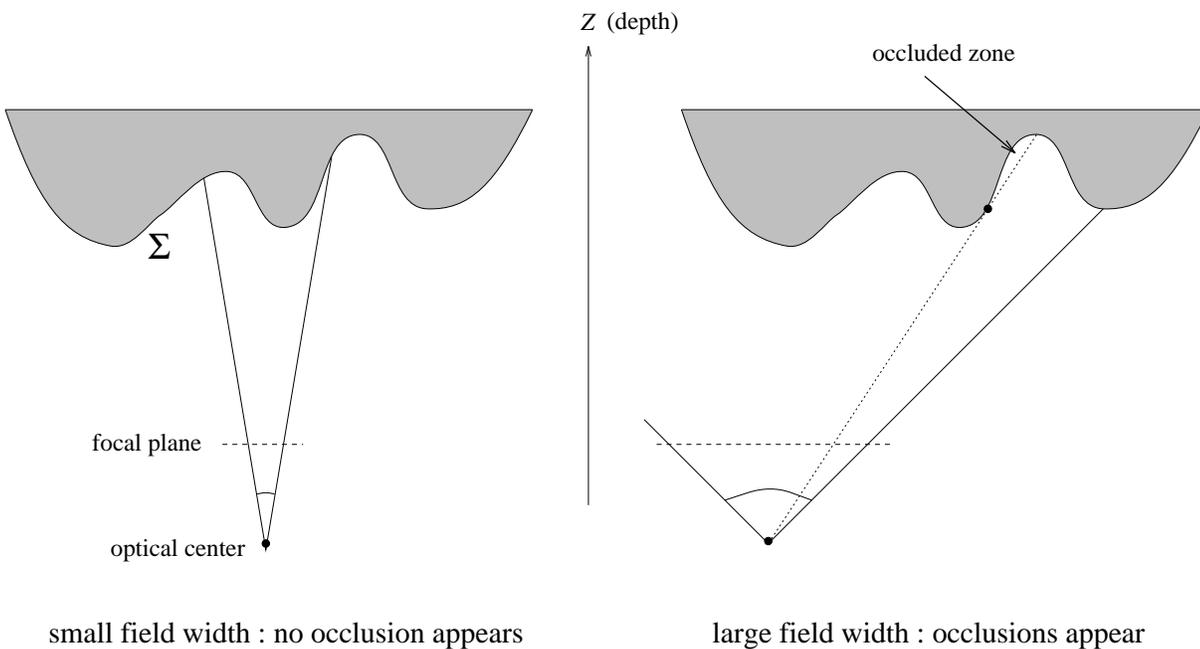
Figure 14.2: Field width and natural occlusions

Figure 14.3 shows what appears in the epipolar plane when occlusions happen : the level lines with the largest velocity (i.e. with the smallest slope on Figure 14.3) occlude the other ones. The reason is simply that when an occlusion arises between two objects, only the nearest one (that is, the one with the largest velocity) remains visible. As in the spatial case (see [21]), the occluding line is characterized by the presence of T-junctions.
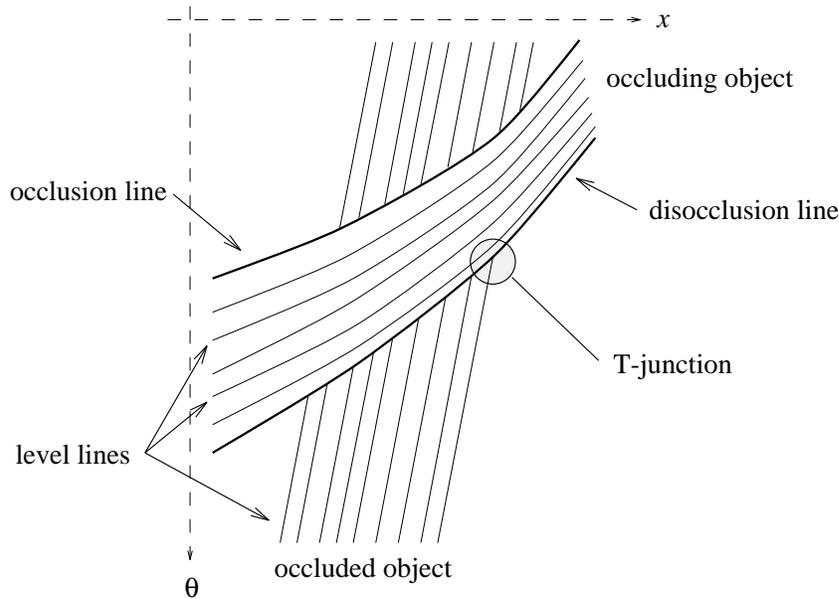
Figure 14.3: Typical occlusions in an epipolar plane

14.4 Conclusion

In this study, we presented a multiscale analysis of movies which is well adapted to the depth recovery. We devised it thanks to an axiomatic formulation in agreement with the depth recovery problem. This multiscale analysis can be viewed as a diffusion process along the movement, with the consequence that it brings time-coherence to movies without performing an undesirable spatial smoothing. In particular, it permits to gather the redundant but incoherent depth information spread among the images of a raw movie into a perfect movie on which the depth can be easily and robustly estimated.

From a theoretical point of view, this multiscale analysis is described by a second order partial differential evolution equation, which presents strong singularities and is not treated by the classical theory of viscosity solutions. We proved uniqueness and existence theorems, although existence is not ensured (at least in a classical sense) in the completely general case. This PDE has interesting properties that can be physically interpreted : in particular, we proved that an ideal movie (that is to say a movie which can be interpreted in terms of a camera movement and a depth map) remains ideal when analyzed by this scale space. We also showed that the corresponding evolution equation is somewhat related to a simple minimization problem.

We provided a very simple numerical scheme which can easily be implemented on parallel machines. By performing numerical experiments on two real movies, we checked the good behaviour of this method, as a movie processing device, and as a depth-recovery preprocessing device.

We think that this study is a good starting point to find robust solutions to the depth recovery

problem. The next step would be to adapt the theory for general movies where occlusions are allowed. Of course, such a generalization should require a non-continuous formulation due to the nature of occlusions. It may also bring new elements to circumvent the strong singularity that appears in the DCMA equation.