Departament d'Enginyeria de Sistemes, Automàtica i Informàtica Industrial

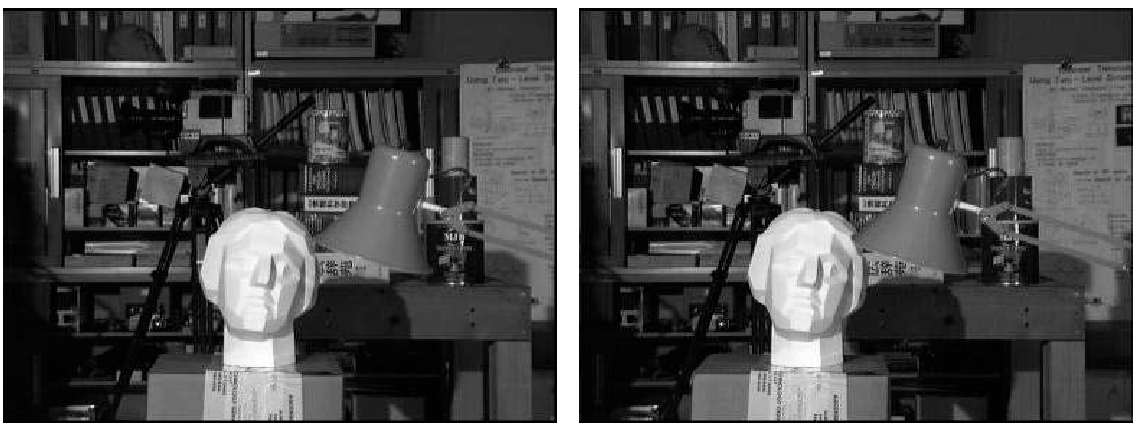
Advanced Topics on Computer Vision

Joan.aranda@upc.edu

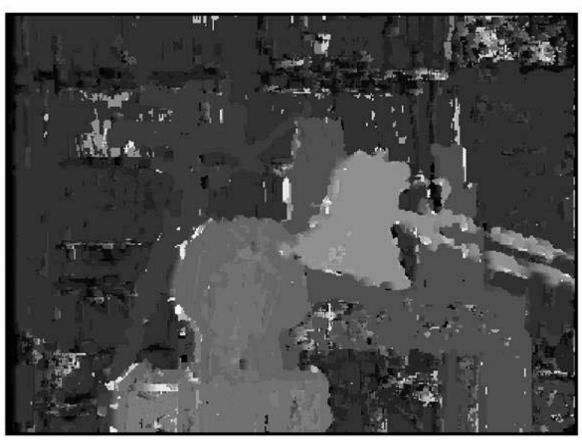
Depth estimation by stereovision

We want to implement a program to estimate the depth of the objects that appear on the scene from pairs of images acquired with some lateral shift, as shown in the next figure.

Left Image Right Image



The first step is to compute the disparity map (showed below), where the gray level indicates the proximity (depth) of this image region to the camera (higher the disparity => higher the proximity). This map can also be viewed as a 3D surface.



Obtained Disparity Map

To do so, implement an algorithm to find correspondences (matching) between homologue points of the left-right images based on correlation, SAD or other similarity function. Assume that the images are already rectified (epipolar lines correspond to the lines of the image), so that a point on the left image will be found on the same line (same Y coordinate) on the right image. Consider a limited range of disparities (depending on input images).

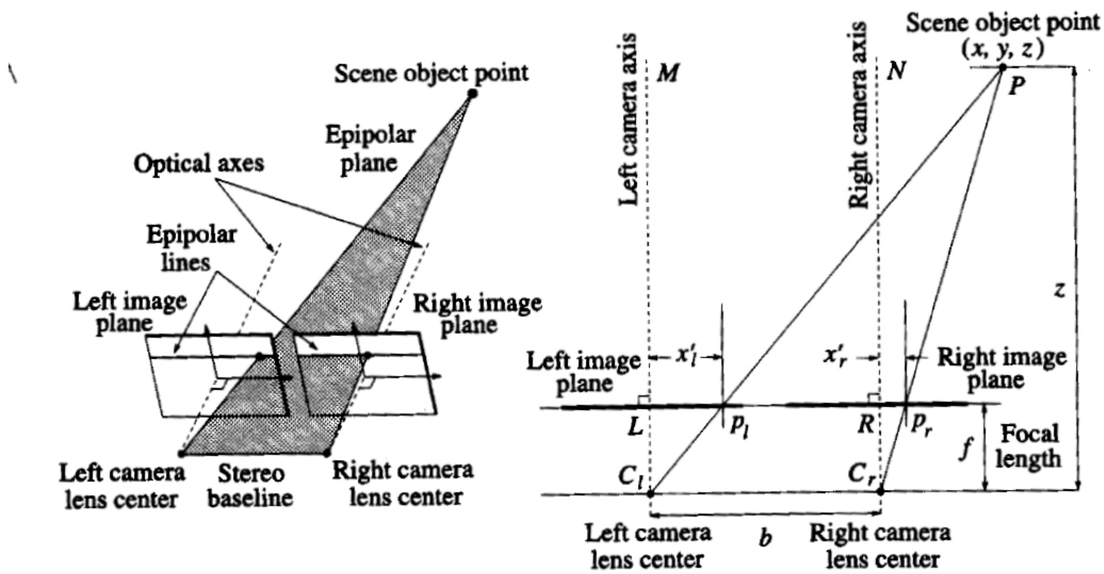


Figure 11.1: Any point in the scene that is visible in both cameras will be projected to a pair of image points in the two images, called a *conjugate pair*. The displacement between the positions of the two points is called the *disparity*.

Try different operator sizes and different similarity functions, indicating those finally chosen (or leave them as function parameters). Try your method with different stereo-pairs and show the input images and the resulting disparity image.

The use of Disparity Space Image (DSI) is recommended to accelerate your code (Szeliski, page 540, fig. 11.5).

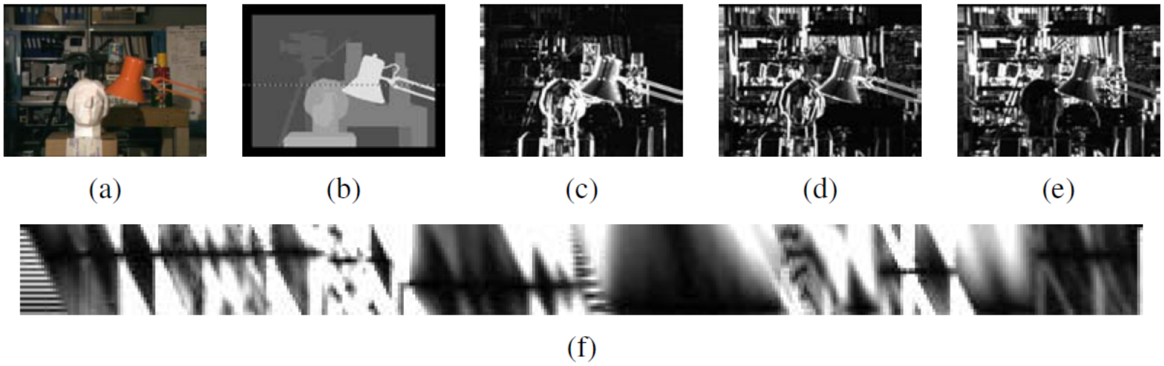
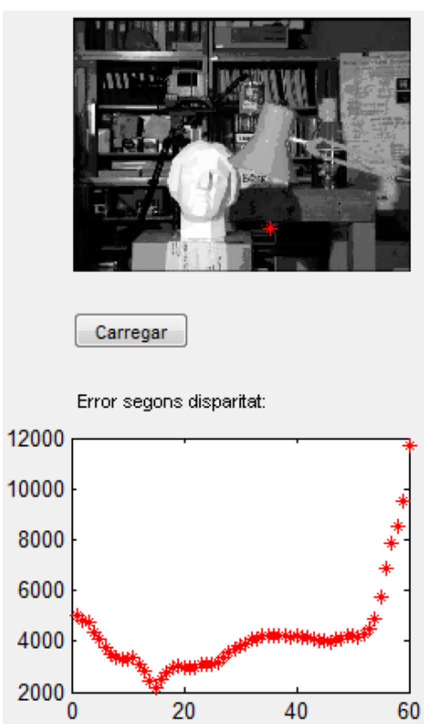
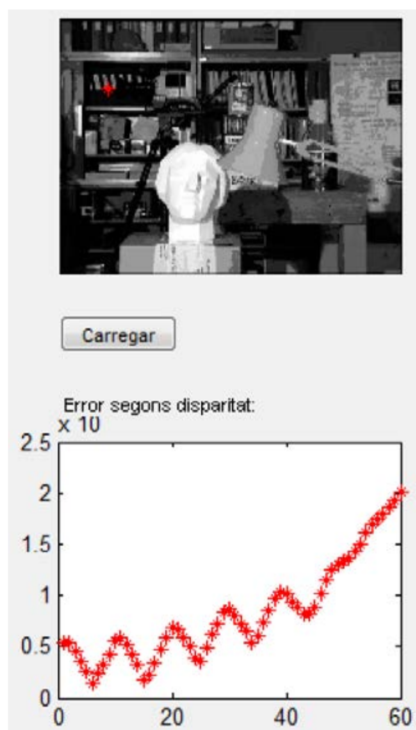


Figure 11.5 Slices through a typical disparity space image (DSI) (Scharstein and Szeliski 2002) © 2002 Springer: (a) original color image; (b) ground truth disparities; (c–e) three (x, y) slices for $d = 10, 16, 21$; (f) an (x, d) slice for $y = 151$ (the dashed line in (b)). Various dark (matching) regions are visible in (c–e), e.g., the bookshelves, table and cans, and head statue, and three disparity levels can be seen as horizontal lines in (f). The dark bands in the DSIs indicate regions that match at this disparity. (Smaller dark regions are often the result of textureless regions.) Additional examples of DSIs are discussed by Bobick and Intille (1999).

During the development of the practice it can be very useful to plot the evolution of the similarity function along the epipolar line. This will give you some clues.





As a first solution, the use of the Shirai algorithm is proposed. Students are encouraged to use other methods of image matching, so that we can also establish a comparison. If you are interested in other methods, refer to the professor.

Algorithm short description:

Shirai algorithm determines a point q in the right image as a corresponding image point to point p in the left image where the similarity measure takes a global minimum within the search interval, and this distinct global minimum must be smaller or equal to an a-priori specified **threshold d1**. If such a global minimum<=d1, does not exist within the search interval then it has to be decided whether to continue the search: if all of the values SIMILARITY (p,q) are greater or equal to an a-priori specified **threshold d2>d1**, then the search is at this position is stopped and the next edge pixel p in the left edge image is selected for the initialization of a further search process. In the other case (there are SIMILARITY values between d1 and d2), the search process is modified: The interval is examined for a possible reduction: All points q at the boundary of the search space are excluded for which it holds SIMILARITY (p,q)>=**d3**, (**d3** is a prior specified threshold>=**d2**). Then the search goes on as previously described. At the beginning you can set **d3=d2**.

SHIRAI ALGORITHM

for (every pixel $(\mathbf{p}, E_{left}(\mathbf{p}))$ of the left image E_{left}) **do**
    if ($(\mathbf{p}, E_{left}(\mathbf{p}))$ is an edge pixel) **then begin**
        initialize the window parameter k and the search interval I;
        loop
        set the window size $n = 2k + 1$ and the new search interval I;
        calculate the similarity measure $SIMILARITY(\mathbf{p}, \mathbf{q})$
            for fixed pixel $(\mathbf{p}, E_{left}(\mathbf{p}))$ in the left image and
            for every pixel $(\mathbf{q}, E_{right}(\mathbf{q}))$ of the search interval;
                                                            { profile analysis }
        if (there is a unique minimum smaller than d_1) **then begin**
            set the disparity value for point $\mathbf{p}$ in the disparity map;
            exit the loop
            end {then}
        else if (all similarity values are larger than d_2) **then begin**
            a disparity assignment is not possible;
            exit the loop
            end {then}
        else if (the window already has maximum size) **then begin**
            a disparity assignment is not possible;
            exit the loop
            end {then}
                                            { preparation of a next search run }
        else begin
            reduce the interval size using d_3;
            $k := k + 1$ { i.e. increase the window size }
            end {else}
        end {loop}
    end {if}

Figure 1.19: Shirai algorithm for a stereo pair E_{left} and E_{right} of scalar images (e.g. of a stereo gray value image pair, or of a stereo pair of related color channels of a stereo color image pair).