Two-Dimensional Regularized Disparity Estimation Based on the Gabor Transform

Xiaodong Huang and Eric Dubois

School of Information Technology and Engineering (SITE)
University of Ottawa, Ottawa, ON, K1N 6N5 Canada

ABSTRACT

This paper presents a disparity estimation algorithm that combines three different kinds of techniques: Gabor transform, variational refinement and region-based affine parameter estimation for disparity calculation. The Gabor transform is implemented using a set of quadrature-pair filters for estimating the two-dimensional correspondences between the two images without the calibration information, and the estimated coarse disparity maps are applied to a variational refinement process which involves solving a set of partial differential equations (PDEs). Then the refined disparity values are used with the image segmentation information so that the parameters of affine transforms for the correspondence of each region can be calculated by singular value decomposition (SVD), and these affine parameters can be applied in turn to bring more refined disparity maps.

Keywords: disparity estimation, Gabor transform, variational regularization, affine parameters

1. INTRODUCTION

In image registration and multiview image analysis, matching is a challenging problem. To analyse object displacements as well as the depth information contained in multiview images for the purpose of 3D scene reconstruction or image-based interpolations, dense matching or feature matching must be performed. Since matching is an ill-posed inverse problem, there are many ambiguities. Even for a stereo image pair with small baseline, the disparity estimations for untextured and slanted surfaces, objects with complex geometries like trees with sky as background, etc., remain as typical problems still unsolved. Our final goal is to solve the matching problem under the framework of multiview images. Such solutions can alleviate or solve most of the ambiguities contained in the matching of a single stereo pair by exploiting the constraint relations among those multiview images. Matching in the general multiview case usually must be performed in two dimensions with possibly large displacements in both horizontal and vertical directions. In addition, for the convenience of image acquisition, the matching should assume no information about the precise relative locations of those multiview images, so that no precise calibrations need to be done in the acquisition stage. Thus, the problem to be solved first in multiview image analysis is how to estimate the horizontal and vertical disparities between two images without calibration information.

The existing approaches for multiview image matching can be classified into two categories: matching with small displacements and with large displacements. For small displacements, the problem is usually handled similarly to motion estimation in video sequences. What we are dealing with here are cases with large displacements involving multiview images rather than video sequences. We mean the difference with at least more than five pixels (usually several dozens of pixels) for the term *large displacement* here.

For multiview image matching with large displacement, most of the algorithms so far assume that the relative camera positions are carefully calibrated. For example, Alvarez et al.[1] and Strecha et al.[2] used some approaches similar to our approach in the refinement stage, and related information from the calibration, such as the epipolar constraint, is used. In the algorithm of Alvarez et al.,[1] the calibration information is assumed to be known for each image in a multiview image sequence, and the dense disparity between each consecutive image pair is estimated first using block-based correlation followed with two-view variational refinement.[3] Then an unstructured set of 3D points is obtained by taking into account the information from these disparity maps.

Further author information: {xhuang, edubois}@site.uottawa.ca

Visual Communications and Image Processing 2006, edited by John G. Apostolopoulos, Amir Said,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6077, 60770Z, © 2005 SPIE-IS&T · 0277-786X/05/$15

In this process, another multiview-based variational refinement exploiting the epipolar relations among those images is applied to regularize those 3D points. One disadvantage in this process is that doing calibration for each image is very inconvenient in the image acquisition process, and sometimes this will limit its applications. For the algorithm of Strecha et al.,[2] the calibration was carried out in an automatic process in which some affine invariant regions among a set of wide-baseline multiview images were detected, and the related parameters for calibration as well as the associated epipolar informations between these images were estimated[4,5]; this information was used in the variational refinement stage for a dense matching. The self-calibration in the algorithm of Strecha et al.[2] is a substantial improvement over doing calibration separately, although it comes at the cost of more complexity in the implementation.

In this paper, we will take a novel approach to use the Gabor transform for the estimation of the initial coarse disparity maps in horizontal and vertical directions, and then use the energy-based variational approach to refine these two coarse disparity maps. The displacements between the images we are dealing in this paper are in the range of several dozens of pixels. In the whole process of our algorithm, we assume no previous knowledge about the relative camera positions, i.e., no calibration information and epipolar constraints are needed. The refinement process in our approach is similar to variational energy-based methods used for disparity estimation as developed by Robert et al.[6] and Alvarez et al.,[3] and used by, e.g., Strecha et al.[2] and Kim et al.[7] However, the novelty of our approach lies in the fact that the energy functional for the refinement is based on the Gabor coefficients combined with the image data. Our algorithm can be applied to the dense disparity estimation for images taken with mostly – but not strictly limited to – translational movements. There could be an amount of rotation in the acquisition movements, although not as large as for Strecha et al.[2] Therefore, our algorithm is an extension of the traditional one-dimensional disparity estimation to the two-dimensional cases, with large displacements in both horizontal and vertical directions, and without the need of a calibration process. This is very helpful for doing dense matching among multiview images in which the displacements are usually in two directions, and the calibration information is not convenient to obtain. Finally, as a procedure to further regularize the refinement for better quality, the variational refined disparity values are used together with image segmentation results[8] so that a region-based disparity refinement scheme is achieved, in which the disparity values within each textureless region become even smoother and vary more linearly than the variational refinement by expressing the disparity values in the region using an affine transform.

The rest of this paper is organized as follows: in section 2, the theory of how the Gabor transform can be used for disparity estimation will be shown together with our new development on how to make use of those Gabor coefficients in a simple and effective way, in order to obtain the coarse disparity maps; in section 3, the refinement process using variational regularization approach performed using Gabor coefficients will be given; and in section 4, the segmentation of the image as well as the affine parameter estimation for each region are presented together with the more refined final disparity maps, followed by a conclusion in section 5.

2. DISPARITY ESTIMATION BASED ON THE GABOR TRANSFORM

The idea of applying the Gabor transform on disparity estimation comes from the fact that the disparity value for one pixel location is related to the difference of the local phase components between the two images. In order to show how the disparity can be obtained from the Fourier phase information of the stereo images, first assume that the right image is a pure horizontal translation of the left image:

$$I_r(x, y) = I_l(x - d, y) \tag{1}$$

where d is constant over the whole image. From the properties of the Fourier transform:

$$\hat{I}_r(\omega_1, \omega_2) = \hat{I}_l(\omega_1, \omega_2) e^{-i\omega_1 d}. \tag{2}$$

Therefore, we could have:

$$\frac{\hat{I}_l(\omega_1, \omega_2)\hat{I}_r^*(\omega_1, \omega_2)}{|\hat{I}_l(\omega_1, \omega_2)||\hat{I}_r(\omega_1, \omega_2)|} = e^{i\omega_1 d} \tag{3}$$

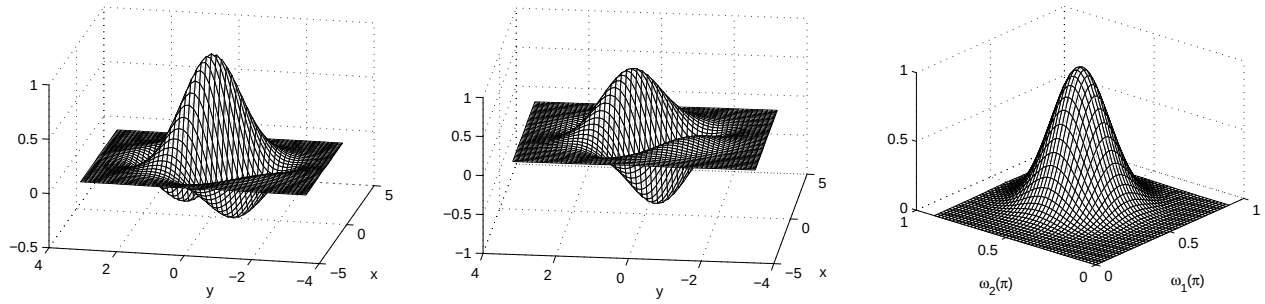


Figure 1. Gabor function with $\omega_1 = \omega_2 = 0.5\pi$: (a) real part; (b) imaginary part; (c) Fourier transform

Hence from the above normalized phase-correlation we can obtain the phase difference between $\hat{I}_l(\omega_1, \omega_2)$ and $\hat{I}_r(\omega_1, \omega_2)$, and the disparity can be obtained by taking the inverse Fourier transform of the correlation product, resulting in an impulse at the location d.

However, in practice, the disparity values vary over the whole image. Thus it is desirable to measure the phase difference locally rather than globally. In order to do this we need to use the windowed Fourier transform. The best choice is to use the Gabor function because the Gabor window performs the localization in both the spatial and the frequency domains simultaneously. The localization in the spatial domain limits the regions taken into account by the Gabor transform to a small neighborhood around that paticular pixel location, while in the frequency domain it could bring band-pass filtered information at different frequencies for detailed analysis. We show these from Fig. 1 in which the real and imaginary parts of a Gabor function with $\omega_1 = \omega_2 = 0.5\pi$ are shown as well as its Fourier transform. We can observe the localizations in both spatial and frequency domains, and the function is tuned to 45° with respect to the x-axis since $\omega_1 = \omega_2$.

The Gabor functions are Gaussian functions modulated by complex sinusoids.[9] For the 2-D case, they are defined as

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)\right] \times \exp\left[i\left(\omega_{10}x + \omega_{20}y\right)\right] \tag{4}$$

where ω_{10} and ω_{20} define the spatial frequencies in the x and y directions respectively. Its Fourier transform is given by

$$G(\omega_1, \omega_2) = e^{-\pi[\sigma_x^2(\omega_1-\omega_{10})^2 + \sigma_y^2(\omega_2-\omega_{20})^2]}. \tag{5}$$

The idea of using band-pass Gabor filters for the problem of disparity estimation originated from Jenkin and Jepson,[10] Sanger[11] and Jepson et al.,[12] and was further developed by Fleet et al.[13] and Fleet.[14] The main advantages of using a phase-based approach include the robustness to different kinds of scenes in which no block size needs to be chosen according to the texture level of images as in the cases using traditional block-based methods (e.g., the sum-of-squared-difference – SSD), as well as the insensitivity to luminance changes. Except for the method used by Fleet,[14] all other references mentioned above[10–13] use a phase-wrapping technique for the processing of Gabor filter results in order to determine the phase difference between the two pixels in left and right images respectively.

2.1. A New Processing Method on Gabor Coefficients for Disparity Estimation

The method that we used for a coarse estimation of the disparity from the local phase information using the Gabor transform is mainly based on the approach of Fleet,[14] in which a set of quadrature-pair Gabor filters is used. Each quadrature-pair Gabor filter is a set of discretized samples of a Gabor function with different tuning frequencies (different ω_{10} and ω_{20}) to different directions, and is used for the filtering of the stereo images to obtain the approximate Gabor transform coefficients at those frequencies. Assume that the outputs of the k^{th} filter

pair are $G_l^k(x,y)$ and $G_r^k(x,y)$ for the left and right images respectively. Instead of doing the phase-wrapping, a local weighted phase-correlation between the two images is calculated as

$$C_k(x,y,\tau) = \frac{W(x,y) \otimes [G_l^k(x,y) G_r^{k*}(x+\tau,y)]}{\sqrt{W(x,y) \otimes |G_l^k(x,y)|^2} \sqrt{W(x,y) \otimes |G_r^k(x+\tau,y)|^2}} \tag{6}$$

where $\otimes$ represents correlation, $W(x,y)$ is a small and localized window around (x,y), and τ is a preshift of the right filter output. Then a summation is obtained over all the filters:

$$S(x,y,\tau) = \sum_k C_k(x,y,\tau) \tag{7}$$

and the disparity for location (x,y) can be estimated by finding a peak in the real part of $S(x,y,\tau)$ and a zero in its imaginary part. This approach avoids the inconvenience of phase-wrapping, but there still remains a problem as to how to choose a selection criterion if the locations of the peak in the real part of $S(x,y,\tau)$ and the zero in its imaginary part do not coincide.

To avoid such uncertainty and to make full use of the information from the Gabor transform, we propose a new method to process these Gabor coefficients.[15] In our approach, the real part and imaginary part of both $G_l^k(x,y)$ and $G_r^k(x,y)$ are compared at each possible location by summing the absolute values of their difference values respectively. Then the disparity $\hat{d_1} \in [0, d_{1,\max}]$ in the horizontal direction and $\hat{d_2} \in [d_{2,\min}, d_{2,\max}]$ in the vertical direction for a position (x,y) in the left image are determined as:

$$(\hat{d_1}, \hat{d_2}) = \arg\min_{d_1,d_2} \sum_k |G_l^k(x,y) - G_r^k(x-d_1, y-d_2)|^2. \tag{8}$$

Here, as a coarse estimation, $\hat{d_1}$ and $\hat{d_2}$ are restricted to integer values.

We use three values $\{\pi/16, \pi/8, \pi/4\}$ for the central frequency $\omega_0 = \sqrt{\omega_{10}^2 + \omega_{20}^2}$ of the Gabor filters. For each frequency, there are four filter pairs tuned to orientations 0°, 45°, 90° and 135° respectively.[16] A sample image pair is shown in Fig. 3(a)(b), where the camera for I_r is shifted to the right and in a backward position with respect to I_l. The maximum horizontal and vertical displacements are 50 pixels and 20 pixels respectively. The coarse dense matching based on (8) for horizontal and vertical directions are shown in Fig. 3(c)(d), where brighter intensities signify larger disparities (and less depth), and the darker intensities signify smaller disparities (and more depth). The horizontal and vertical disparity images are individually scaled to use the entire intensity range for display.

3. ENERGY-BASED VARIATIONAL REFINEMENT

The disparity maps in Fig. 3(c)(d) are very coarse and some untextured areas are very noisy. Therefore, refinement on these disparity maps is needed. To alleviate such errors we need to get the disparity values for continuous surfaces to be changing smoothly, while maintaining the disparity discontinuities at the object boundaries. To achieve such properties, the energy-based variational regularization can be applied as a refinement process for the disparity estimation. The energy functional used to estimate the disparities $d_1(x,y)$ and $d_2(x,y)$ usually contains two terms:

$$E(d_1,d_2) = E_D(d_1,d_2) + \lambda E_S(d_1,d_2) \tag{9}$$

where $E_D(d_1,d_2)$ is a data fidelity term, and $E_S(d_1,d_2)$ is a regularization term which controls the smoothing of the disparity field, together with the regularization coefficient λ.

3.1. Data Fidelity Term of the Energy Functional

Usually, the data fidelity term used in variational refinement can be obtained from image data directly:

$$E_D(d_1,d_2) = \iint [I_l(x,y) - I_r(x-d_1, y-d_2)]^2 \, \mathrm{d}x \, \mathrm{d}y. \tag{10}$$

Since the data of a general image are randomly distributed, with many abrupt changes, therefore they make the variational refinement, which is actually a descent-based optimization process, very easy to be trapped in local minima. To deal with this issue, we propose a new Gabor coefficient-based data fidelity term:

$$E_D(d_1, d_2) = \sum_k \iint |G_l^k(x, y) - G_r^k(x - d_1, y - d_2)|^2 \, \mathrm{d}x \, \mathrm{d}y \tag{11}$$

which makes use of the smooth characteristic of Gabor coefficients as well as of any other frequency transform coefficients. We will show in next subsection the obvious qualitative difference between the 1-D refinement results on a standard stereo image pair using (10) and (11) respectively.

3.2. Regularization Term of the Energy Functional

The variational regularization has had extensive application in optical flow estimation, as well as some for disparity estimations by Robert et al.[6] and Alvarez et al.[3] The idea in Alvarez et al.[3] is to control the smoothing of disparity variations using the values of image gradients. When the value of image gradient is low the disparity would get smoothed, and the smoothing process is stopped when the value of image gradient is high, which represents a possible object boundary. Extending to two-dimensional disparities, the regularization functional we used for the regularization of $d_1(x, y)$ and $d_2(x, y)$ is[15]

$$E_R(d_1, d_2) = \iint \Big\{ \frac{1}{(1 + I_{l,x}^2)^2}(d_{1,x}^2 + d_{2,x}^2) + \frac{1}{(1 + I_{l,y}^2)^2}(d_{1,y}^2 + d_{2,y}^2) \Big\} \mathrm{d}x \, \mathrm{d}y \tag{12}$$

where $I_{l,x}$ and $I_{l,y}$ are derivatives of I_l in x and y directions, $d_{1,x}$ and $d_{1,y}$ are derivatives of $d_1(x, y)$ in x and y directions, and similarly for $d_{2,x}$ and $d_{2,y}$.

Thus the overall energy functional to be minimized for $d_1(x, y)$ and $d_2(x, y)$ is

$$\begin{aligned} E(d_1, d_2) =& E_D(d_1, d_2) + \lambda E_R(d_1, d_2) \\ =& \sum_k \iint |G_l^k(x, y) - G_r^k(x - d_1, y - d_2)|^2 \, \mathrm{d}x \, \mathrm{d}y \\ &+ \lambda \iint \Big\{ \frac{1}{(1 + I_{l,x}^2)^2}(d_{1,x}^2 + d_{2,x}^2) + \frac{1}{(1 + I_{l,y}^2)^2}(d_{1,y}^2 + d_{2,y}^2) \Big\} \mathrm{d}x \, \mathrm{d}y \end{aligned} \tag{13}$$

and the minimization over $d_1(x, y)$ and $d_2(x, y)$ is achieved by applying a gradient descent method to solve the associated Euler-Lagrange equations of (13) with respect to $d_1(x, y)$ and $d_2(x, y)$:

$$\frac{\partial d_1}{\partial t} = \sum_k \{[G_l^k(x, y) - G_r^k(x - d_1, y - d_2)] \times G_{r,x}^k(x - d_1, y - d_2)\} - \lambda \Big\{ \frac{\partial}{\partial x} \Big[\frac{d_{1,x}}{(1 + I_{l,x}^2)^2} \Big] + \frac{\partial}{\partial y} \Big[\frac{d_{1,y}}{(1 + I_{l,y}^2)^2} \Big] \Big\}, \tag{14}$$

$$\frac{\partial d_2}{\partial t} = \sum_k \{[G_l^k(x, y) - G_r^k(x - d_1, y - d_2)] \times G_{r,y}^k(x - d_1, y - d_2)\} - \lambda \Big\{ \frac{\partial}{\partial x} \Big[\frac{d_{2,x}}{(1 + I_{l,x}^2)^2} \Big] + \frac{\partial}{\partial y} \Big[\frac{d_{2,y}}{(1 + I_{l,y}^2)^2} \Big] \Big\}, \tag{15}$$

where t is a pseudo-time variable used in the iteration process.

We first show the improvement of using (11) for the data fidelity term by one-dimensional disparity estimation results on a pair of images from the sequence *Flower Garden* in Fig. 2, in which (a) is the original left image, (b) is the refinement result using the image data directly in the data fidelity term, and (c) is the refinement result using Gabor coefficients in data fidelity term. Comparing (b) and (c) of Fig. 2, we see that using Gabor coefficients in the data fidelity term can bring obvious improvements in some areas around object contours.

(a) (b) (c)

Figure 2. (a) Original left image of *Flower Garden*; (b) refinement by using image data; (c) refinement by using Gabor coefficients

3.3. Results of Variational Refinement

The numerical implementations of solving (14) and (15) iteratively are given by the forward Euler method, and the spatial derivatives are calculated by the central difference scheme. The parameters we used are $\lambda = 2.5$, time step 0.05, 200 iterations. This number of iterations is less than the usual iteration numbers if using (10) for the data fidelity term as a result of the quick convergence when using Gabor coefficients. The initial disparity values are the previous results of the coarse method.

The refinement results are shown in Fig. 3(e)(f). We can see that the noisy areas in coarse disparity maps are smoothed to some extent, objects with depth discontinuities along their boundaries can be approximately recognized, and depth variations from the front objects to the background are clearly reflected. Thus, we have developed a new Gabor-based two-dimensional disparity estimation algorithm, in which Gabor coefficients are used both in the coarse disparity estimation and in the variational regularization refinement stage for better performance, and for which no camera calibration information is required.

4. AFFINE ESTIMATION AND REGULARIZATION

Until now, like most of the existing algorithms for disparity estimation, including the methods we used in this paper – from Gabor-based coarse estimation to variational refinement – are all pixel-based. This is a common feature for various otherwise quite different disparity estimation algorithms. While this approach is convenient, it does not make use of the information associated with the shape of areas in the images. One reason might lie in the fact that there is not a perfect segmentation method so far that can segment the surfaces of the true objects, rather than segment different colored regions on the same surface. Therefore, one can easily imagine that an effective region-based disparity estimation should dramatically increase the quality of disparity maps.

Compared to the large number of papers on pixel-based disparity estimation algorithms, there are only a few dealing with region-based disparity estimation.[17,18] Wei et al.[17] used the mean shift segmentation algorithm developed by Comaniciu et al.[8] to segment the images into different regions. However, in the next steps, like the methods of Birchfield et al.,[18] they applied oversegmentation to each region in order to handle the linear variation problems for untextured and slanted surfaces. Such oversegmentation is a step towards a pixel-based approach reducing the advantage of a region-based approach. We believe that each region in one image of the stereo pair can be considered as an affine transform from the same region in another image, and the region-based disparity estimation is thus converted to the estimation of the affine parameters for each region.

We also applied the mean shift segmentation algorithm[8] to the image I_l, and the segmentation result is shown in Fig. 4. Each region is indicated by one color value.

We assume that the coordinates $(x, y)^T$ of each pixel in a region in I_l is related to its corresponding pixel $(x_r, y_r)^T$ in I_r by an affine transform:

$$\begin{pmatrix} x_r \\ y_r \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}. \tag{16}$$

Therefore, the two-dimensional disparities $d_1(x, y)$ and $d_2(x, y)$ are related to these affine parameters by

$$\begin{aligned} d_1(x, y) &= x - a_{11}x - a_{12}y - a_{13}, \\ d_2(x, y) &= y - a_{21}x - a_{22}y - a_{23}. \end{aligned} \tag{17}$$

Thus, the estimated $d_1(x, y)$ and $d_2(x, y)$ for each pixel in one region from the previous variational refinement can be grouped and used as known variables so that the affine parameters can be estimated from (17). Since each pixel in the region gives a set of equations as in (17), and for most of the cases, the number of pixels in a region is larger than the number of affine parameters (six for 2-D affine transform), therefore the estimation of the six parameters ($a_{11} \sim a_{23}$) can be done by least squares, implemented using singular value decomposition (SVD). Then, once the affine parameters are estimated, a new set of disparities $d_1(x, y)$ and $d_2(x, y)$ for each pixel in the region can be in turn calculated by (17).

The new results for $d_1(x, y)$ and $d_2(x, y)$ from the above procedure are shown in Fig. 5. We can see that this kind of parameterized estimation process can give more reasonable results in which the noise is largely removed, and the linear variation of each surface is clearly and more precisely reflected. Therefore, this kind of disparity calculation from the estimated affine parameters can be seen as another kind of region-based regularization method for disparity estimation.

5. CONCLUSION AND FUTURE WORK

Disparity estimation is a long standing problem. Due to its ill-posed nature, there is much ambiguity in the estimation process. We believe that further improvement will be possible under the multiview framework in which the constraint relations among different images can be exploited to resolve those ambiguities and to further increase the performance. For multiview image correspondences, the disparity estimations should usually be performed in two dimensions and sometimes without calibration information. Thus, the traditional one-dimensional disparity estimation needs to be extended to the two-dimensional case. In this paper, we further developed the Gabor transform method for two-dimensional disparity estimation, and applied the results from this method to two different kinds of regularization schemes. The first is the variational regularization scheme with our improvement on the data fidelity term; the second is the region-based affine parameter regularization scheme which results in more refined disparity maps. Thus, our main contribution is a combined two-dimensional disparity estimation algorithm that takes advantage of different disparity estimation methods and can give better results than the separate disparity estimation scheme, along with our novel developments in each method.

In the future, we will futher develop the affine parameter estimation methods for the region-based disparity calculations, so that the affine parameters can be obtained in an update process along with the matching quality of each region. There are two methods we plan to try: a multi-dimensional searching scheme, as well as the Gauss-Newton method for the optimization of affine parameters.[19]

ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada under the research network *LORNET*.

REFERENCES

1. L. Alvarez, C. Cuenca, and J. Sanchez, "Regularizing a set of unstructured 3D points from a sequence of stereo images," *Proc. Scale Space 2003* **2695**, pp. 449–463, 2003.
2. C. Strecha, T. Tuytelaars, and L. V. Gool, "Dense matching of multiple wide-baseline views," *Proc. IEEE Int. Conf. Computer Vision* **2**, pp. 1194–1201, 2003.
3. L. Alvarez, R. Deriche, J. Sanchez, and J. Weickert, "Dense disparity map estimation respecting image discontinuities: A PDE and scale-space based approach," *Journal of Visual Communication and Image Representation* **13**, pp. 3–21, 2002.
4. T. Tuytelaars and L. V. Gool, "Wide baseline stereo matching based on local, affinely invariant regions," *Proc. British Machine Vision Conference*, pp. 412–422, 2000.

5. V. Ferrari, T. Tuytelaars, and L. V. Gool, "Wide-baseline multiple-view correspondences," *Proc. IEEE Conf. Computer Vision Pattern Recognition*, pp. 718–725, 2003.
6. L. Robert and R. Deriche, "Dense depth map reconstruction: A minimization and regularization approach which preserves discontinuities," *Proc. European Conference on Computer Vision*, 1996.
7. H. Kim and K.Sohn, "3D reconstruction from stereo images for interactions between real and virtual objects," *Signal Processing: Image Communications* **20**, pp. 61–75, 2005.
8. D. Comaniciu and P. Meer, "Mean-shift: a robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.* **24**, pp. 603–619, 2002.
9. D. Gabor, "Theory of communication," *J. IEE* **93**, pp. 429–457, 1946.
10. M. Jenkin and A. Jepson, "The measurement of binocular disparity," in *Computational Processes in Human Vision*, Z. Pylyshyn, ed., Ablex Publ, New Jersey, 1988.
11. T. Sanger, "Stereo disparity computation using Gabor filters," *Biological Cybernetics* **59**, pp. 405–418, 1988.
12. A. Jepson and M. Jenkin, "Fast computation of disparity from phase differences," *Proc. IEEE Conf. Computer Vision Pattern Recognition*, pp. 398–403, 1989.
13. D. Fleet, "Phase-based disparity measurement," *CVGIP: Image Understanding* **53**, pp. 198–210, 1991.
14. D. Fleet, "Disparity from local weighted phase-correlation," *Proc. IEEE Int. Conf. on Systems, Man, and Cybernetics*, pp. 48–56, 1994.
15. X. Huang and E. Dubois, "Three-view dense disparity estimation with occlusion detection," *Proc. IEEE Int. Conf. Image Processing* **3**, pp. 393–396, 2005.
16. X. Huang and E. Dubois, "Disparity estimation for the intermediate view interpolation of stereoscopic images," *Proc. IEEE Int. Conf. Acoustics Speech and Signal Processing* **2**, pp. 881–884, 2005.
17. Y. Wei and L. Quan, "Region-based progressive stereo matching," *Proc. IEEE Conf. Computer Vision Pattern Recognition* **1**, pp. 106–113, 2004.
18. S. Birchfield and C. Tomasi, "Depth discontinuities by pixel-to-pixel stereo," *Proc. IEEE Int. Conf. Computer Vision*, pp. 1073–1080, 1998.
19. C. Bergeron and E. Dubois, "Gradient-based algorithms for block-oriented MAP estimation of motion and application to motion-compensated temporal interpolation," *IEEE Trans. Circ. Sys. Video Tech.* **1**, pp. 72–85, 1991.

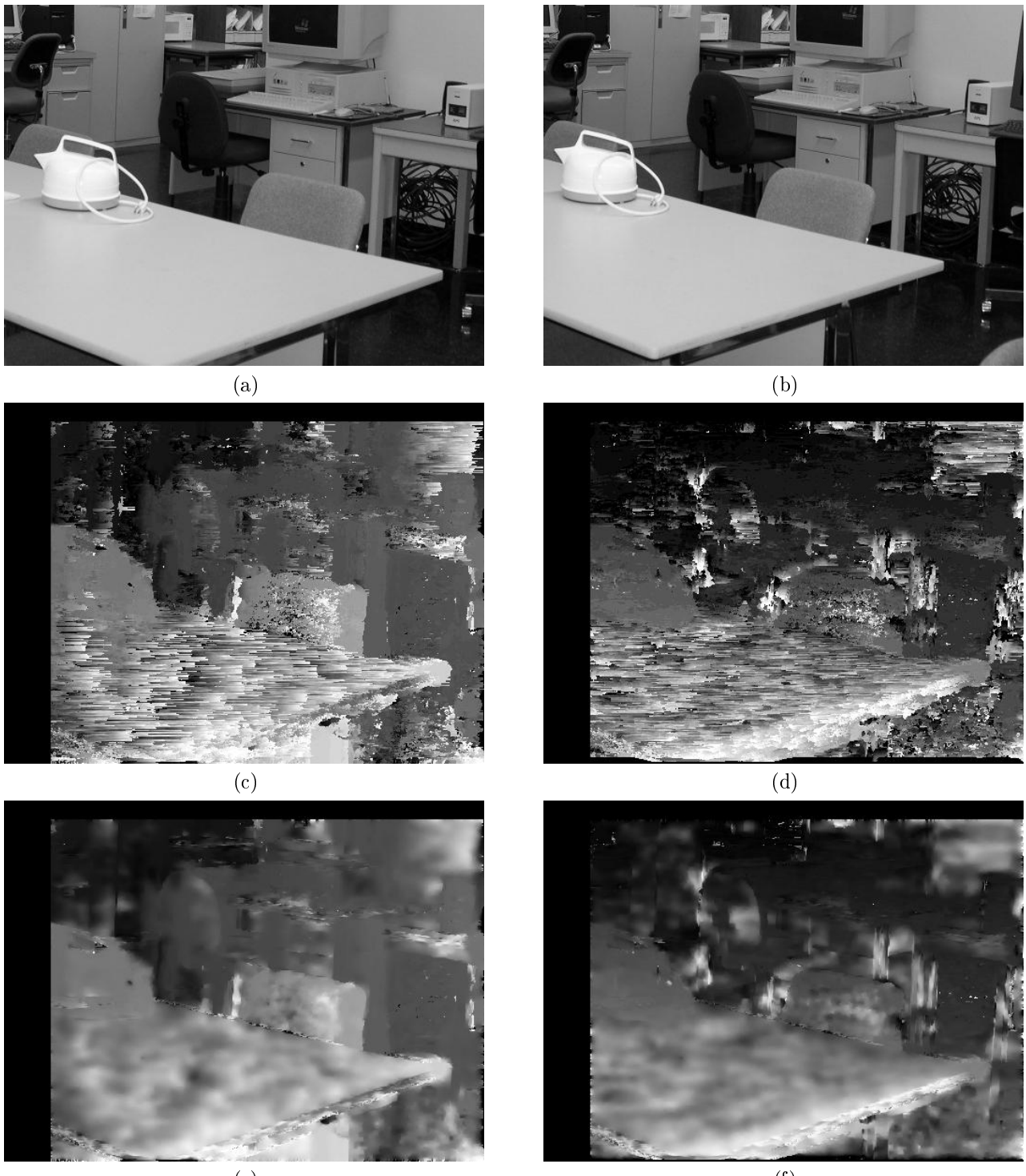
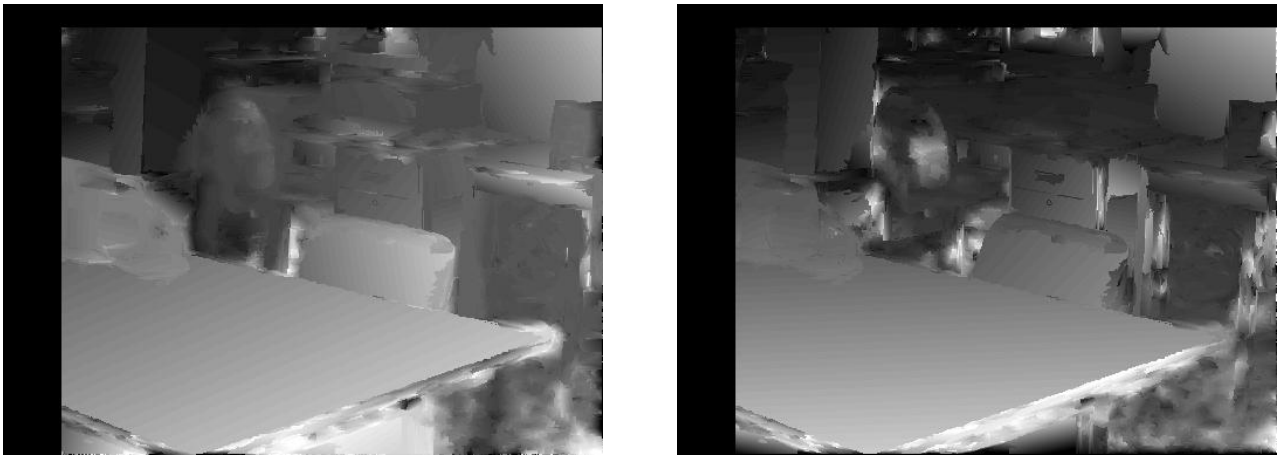


Figure 3. (a) Left view of an image pair I_l; (b) right view of an image pair I_r; (c) horizontal coarse disparities; (d) vertical coarse disparities; (e) refinement result for horizontal disparities; (f) refinement result for vertical disparities.

Figure 4. Segmentation of I_l by mean shift



(a) (b)

Figure 5. New regularized results by applying the affine parameters to the calculation of the disparities for each region: (a) $d_1(x, y)$; (b) $d_2(x, y)$.